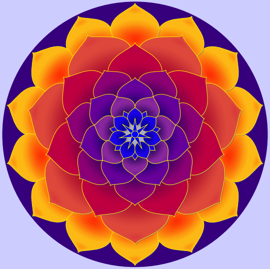
Quantitative Portfolio Management



Textbook for the course

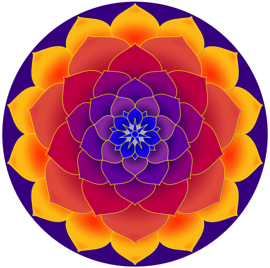
Raman Uppal

2025

Blank page

Table of Contents

Quantitative Portfolio Management



Chapter 1: Managing Financial Data

Raman Uppal

2025

The big picture: Plan for the entire book

Part A: Preliminaries

Chapter 1: Managing financial data

Chapter 2: Performance measurement (especially out of sample)

Part B: Modern portfolio management

Chapter 3: Mean-variance portfolios that *ignore* estimation error

Chapter 4: Mean-variance portfolios that *adjust for* estimation error

Part C: Post-Modern Portfolio Management

Chapter 5: CAPM-based portfolios: Black-Litterman model

Chapter 6: Factor-based portfolios: Parametric portfolio policies

Chapter 7: Volatility-timed factor portfolios

Chapter 8: Portfolios exploiting systematic risk factors *and* unsystematic risk

Table of contents

Road map

What do we want to do in Chapter 1

In this chapter, we first take a big-picture view of the material to be covered in the book.

Then, we study the main kinds of data that one can use to construct optimal portfolios.

You will learn how to use Python to obtain this data and store it in an efficient way so that it can be accessed easily.

Road map

Overview of the entire course

- The objective of this course is to study cutting-edge methods to construct optimal equity portfolios that perform well out of sample.
- Two key questions, therefore, are
  - Q1. How should we construct optimal portfolios?
  - Q2. How should we measure portfolio performance?
- The course provides a solid foundation of the theory of portfolio choice and the knowledge required to implement this theory.
- A key part of the course is learning how to use Python to work with data to implement state-of-the-art portfolio-choice models.

Road map

Mean-variance efficient portfolios

- We will focus on "mean-variance efficient portfolio weights"
- These are portfolio weights that trade-off optimally
  - the return mean of the portfolio,
  - the return variance of the portfolio.
- So, for most of the course, we will not focus on
  - higher returns moments, such as skewness and kurtosis;
  - other risk measures, such as VaR, expected shortfall, etc.
  - other objective functions, such as maximizing expected utility.

Mean-variance portfolio weights when $N = 1$

- When there is only one risky asset (say the market portfolio), and the investor needs to choose the proportion of wealth
  - to allocate to the market portfolio
  - with the rest allocated to the risk-free asset,
- then, the optimal portfolio weight in the risky asset is

$$w_1 = \frac{1}{\gamma} \frac{\mathbb{E}[R_1] - R_f}{\mathbb{V}[R_1]}, \quad \text{where}$$

  - w_1 is the proportion of wealth invested in the single risky asset
  - γ is the risk-aversion of the investor
  - $\mathbb{E}[R_1]$ is the expected (gross) return on the risky asset
  - R_f is the (gross) return on the risk-free asset
  - $\mathbb{V}[R_1]$ is the variance of return on the risky asset, $\sigma^2_{R_1} = \sigma_{R_{11}}$.

Understanding the optimal portfolio weight

$$w_1 = \frac{1}{\gamma} \frac{\mathbb{E}[R_1] - R_f}{\mathbb{V}[R_1]}$$

- The three terms in the expression above make intuitive sense:
  1. If risk aversion γ increases, weight in risky asset decreases;
  2. If risk premium $\mathbb{E}[R_1] - R_f$ increases, weight in risky asset increases;
  3. If the risk $\mathbb{V}[R_1]$ increases, weight in risky asset decreases.
- Whenever we see mathematical expressions in the course, we will want to make intuitive sense of them.

Mean-variance portfolio weights when $N > 1$

- When there are many risky assets ($N > 1$), and the investor needs to choose the proportion of wealth
  - to allocate to the N risky assets
  - with the rest allocated to the risk-free asset,
- then, the optimal vector of portfolio weights in the N risky assets is

$$w = \frac{1}{\gamma}\left(\mathbb{V}[R]\right)^{-1}(\mathbb{E}[R] - R_f 1_N), \quad \text{where}$$

  - w is the N-vector of weights invested in each risky asset
  - γ is the risk-aversion of the investor
  - $\mathbb{V}[R]$ is the $N \times N$ covariance matrix for returns on the N risky assets
  - $\mathbb{E}[R]$ is the N-vector of expected returns on each of the risky assets
  - R_f is the (gross) return on the risk-free asset
  - 1_N is the N-vector of ones

Explicit expressions for components of optimal portfolio

- N-dimensional vectors:

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}; \qquad \mathbb{E}[R] = \begin{bmatrix} \mathbb{E}[R_1] \\ \mathbb{E}[R_2] \\ \vdots \\ \mathbb{E}[R_N] \end{bmatrix}; \qquad 1_N = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

- $(N \times N)$ matrix:

$$\mathbb{V}[R] = \begin{bmatrix} \sigma_{R_{11}} & \sigma_{R_{12}} & \cdots & \sigma_{R_{1N}} \\ \sigma_{R_{21}} & \sigma_{R_{22}} & \cdots & \sigma_{R_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{R_{N1}} & \sigma_{R_{N2}} & \cdots & \sigma_{R_{NN}} \end{bmatrix}.$$

- Note that
  - $\sigma_{R_{nn}} = \sigma^2_{R_n}$ denotes the return variance of asset n, and
  - $\sigma_{R_{nm}}$ denotes the return covariance between assets n and m.

Start of focus

Road map

Estimation error in mean-variance portfolio weights

- Our focus will be to
  - identify mean-variance optimal weights
  - that perform well out-of-sample.

$$w = \frac{1}{\gamma}\big(\mathbb{V}[R]\big)^{-1}(\mathbb{E}[R] - R_f 1_N) \quad \ldots \text{key expression in the course}$$

- Thus, the challenge is to estimate $\mathbb{E}[R]$ and $\mathbb{V}[R]$ precisely.

The challenge in identifying an optimal portfolio

- Estimating the N-vector $\mathbb{E}[R]$ is difficult because its precision does not improve with additional data.

$$\begin{aligned}
\text{Return from year 0 to 100} = R_{0,100} &= \frac{P_{100}}{P_0} \\
&= \frac{P_{100}}{P_{99}} \times \frac{P_{99}}{P_{98}} \times \cdots \times \frac{P_2}{P_1} \times \frac{P_1}{P_0} \\
&= \frac{P_{100}}{P_0} \ldots \text{intermediate prices cancel out}
\end{aligned}$$

- Estimating the $(N \times N)$-matrix $\mathbb{V}[R]$ is difficult because of the large number of parameters to be estimated:
  - N variance parameters on the diagonal of $\mathbb{V}[R]$, and
  - $(N \times N - N)/2$ off-diagonal (unique) covariance parameters, for
  - a total of $N + (N \times N - N)/2 = N(N+1)/2$ parameters.
  - So, if $N = 100$, then $N(N+1)/2 = 5050$.

The problem with poor estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$

- When estimating the portfolio weights

$$w = \frac{1}{\gamma}\big(\mathbb{V}[R]\big)^{-1}(\mathbb{E}[R] - R_f 1_N),$$

- if the elements of $\mathbb{E}[R]$ and $\mathbb{V}[R]$ are estimated poorly, then
  - the vector $\mathbb{E}[R]$ has large error
  - the matrix $(\mathbb{V}[R])^{-1}$ has large error (high condition number)
  - multiplying $\mathbb{E}[R]$ by $(\mathbb{V}[R])^{-1}$ magnifies the error in $\mathbb{E}[R]$
  - leading to badly estimated portfolio weights, w,
  - which perform poorly out of sample.

The solution: Better estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$

$$w = \frac{1}{\gamma}\big(\mathbb{V}[R]\big)^{-1}(\mathbb{E}[R] - R_f 1_N),$$

- Throughout the course, we will focus on studying different ways of estimating $\mathbb{E}[R]$ and $\mathbb{V}[R]$, so as to
  - reduce the error (and dimension) in estimating $\mathbb{E}[R]$, and
  - reduce the error (and dimension) of $\mathbb{V}[R]$.

$\mathbb{E}[R]$ and $\mathbb{V}[R]$ important also for other decisions

- Note that $\mathbb{E}[R]$ and $\mathbb{V}[R]$ required also for other financial decisions
  - for example, cost of capital requires estimating $\mathbb{E}[R]$;
  - for example, risk management requires estimating $\mathbb{V}[R]$.
- A large part of the current research in financial economics is devoted to improving the estimation of $\mathbb{E}[R]$ and $\mathbb{V}[R]$.

End of focus

Road map

Portfolio weights over time

- We wish to identify the portfolio weights not just for one date, but for many dates, $t = \{1, 2, \ldots, T\}$.
- Denote by $w_{t,n}$, with $n = \{1, 2, \ldots, N\}$, the weight at date t on each asset, which is the proportion of wealth invested in that asset.
- Denote by w_t the N-dimensional vector of portfolio weights

$$w_t = \begin{bmatrix} w_{t,1} \\ w_{t,2} \\ \cdots \\ w_{t,N} \end{bmatrix}.$$

- We wish to study how to choose w_t; that is, how to choose
  - the N portfolio weights
  - at each of the t decision dates.

Using the transpose operator

- We have defined the $(N \times 1)$-vector

$$w_t = \begin{bmatrix} w_{t,1} \\ w_{t,2} \\ \cdots \\ w_{t,N} \end{bmatrix}.$$

- The transpose of w_t is given by the $(1 \times N)$-vector

$$w_t^\top = [w_{t,1}, w_{t,2}, \cdots, w_{t,N}], \quad \ldots \text{where } {}^\top \text{ denotes the transpose operator}$$

- which means that the sum of the portfolio weights can be written as

$$\sum_{n=1}^{N} w_{t,n} = w_t^\top 1_N, \quad \text{and}$$

- the expected return on a portfolio p can be written as

$$\mathbb{E}[R_{tp}] = \sum_{n=1}^{N} w_{t,n}\, \mathbb{E}[R_{t,n}] = w_t^\top \mathbb{E}[R_t].$$

Road map

Out-of-sample returns

- When choosing portfolio weights w_t, we want the portfolio to perform well in terms of its risk and return at date $t+1$. That is,
  - the weights w_t are chosen based on information available until t;
  - the portfolio performance depends on returns next period, $t+1$.
  - Therefore, this performance is called out of sample.
- The return at $t+1$ of the portfolio w_t is

  $$[\text{Portfolio return}]_{t+1} = [\text{Portfolio weights}]_t \times [\text{Asset returns}]_{t+1}$$

  $$R_{t+1,p} = w_t^\top R_{t+1} = \sum_{n=1}^{N} w_{t,n}\, R_{t+1,n}, \quad \ldots \text{“}\top\text{” is transpose operator}$$

- So, our choice of w_t depends on our views about $R_{t+1,n}$.

What could the portfolio weights depend on?

- The weight assigned to each asset $w_{t,n}$ will depend on the features of that asset. These features could include:
  - The expected return of this asset;
  - The risk of this asset, which could be measured as its variance, skewness, kurtosis, downside risk, value-at-risk (VaR), expected shortfall, tail risk, etc.;
- The asset's expected return and risk could depend on
  - Past returns of the asset itself and of other assets;
  - Other characteristics of this asset, e.g., its size, profitability, etc.;
  - The sensitivity of the returns of this asset to the returns of other assets, measured by covariance or correlation;
  - The sensitivity of the return of this asset to macroeconomic factors.

Cross-section, time-series, and panel data

- So, the first thing we need to do is to understand how to use data to compute these quantities, which can, potentially, be measured
  - across N assets (cross-section),
  - over T dates (time-series), and
  - across N assets and over T dates (panel).

Road map

Notation: N and T

- Throughout the course, our analysis will depend on two quantities:
  - N with $n = \{1, 2, \ldots, N\}$, representing the number of assets in which we can invest.
  - T, with $t = \{1, 2, \ldots, T\}$ representing the number of observations.

Notation: K and T^{est}

- Throughout the course, we will discuss factors that help us make investment decisions:
  - K with $k = \{1, 2, \ldots, K\}$, representing the number of factors used to make investment decisions.
  - $T^{\text{est}} \leq T$, with $t^{\text{est}} = \{1, 2, \ldots, T^{\text{est}}\}$ representing the number of observations we use for estimating quantities of interest.

Notation: N, K, T, and T^{est}

- In the papers on portfolio management,
  - the notation for N, K, and T is standard; that is, almost all papers use the same notation;
  - the notation for T^{est} is not standard, so you will need to be alert about the notation that is being used.
- Some of the symbols used to represent the length of the estimation window are:
  - $M \leq T$, with $m = \{1, 2, \ldots, M\}$, where M stands for months of data used for estimation (usually 60 or 120 months);
  - $D \leq T$, with $d = \{1, 2, \ldots, D\}$, where D stands for days of data used for estimation (usually 30, 60 or 90 days);
  - $\tau \leq T$, with $t = \{1, 2, \ldots, \tau\}$.

Our notation compared to that used by other people

Our notation	Other choices in other books and papers
w	x or θ
γ	$\gamma = 1/\tau$, where τ denotes risk-tolerance
$\mathbb{V}[R]$ or $\mathbb{V}_R$	V or Σ_R or Σ or $\mathbb{C}$ (for covariance matrix)
$\mathbb{E}[R]$ or $\mathbb{E}_R$	$E[R]$ or μ_R or μ (for mean returns)
R_f	r_f (net risk-free rate)
1_N	$\mathbf{1}$ or e or ι (iota)
T^{est}	M, D, or τ

Road map

Convention regarding data . . . I

- We will adopt the same three principles as in Tidy Finance regarding data:
  1. Every column is a variable.
  2. Every row is an observation.
  3. Every cell is a single value.

Convention regarding data . . . II

- Consider the following table with prices for
  - T dates (in rows) and
  - N assets (in columns),
  - with the typical entry being $P_{t,n}$.

Dates \ Assets	$n=1$	$n=2$	$\cdots$	$n=N$
$t=1$	$P_{1,1}$	$P_{1,2}$	$\cdots$	$P_{1,N}$
$t=2$	$P_{2,1}$	$P_{2,2}$	$\cdots$	$P_{2,N}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$t=T$	$P_{T,1}$	$P_{T,2}$	$\cdots$	$P_{T,N}$

Road map

Prices and returns

- The price of asset n at date t, denoted $P_{t,n}$ is the most common type of data we will work with.

- We will use data on prices to compute returns:

$$\text{Gross returns:} \quad R_{t,n} = \frac{P_{t,n}}{P_{t-1,n}} \tag{1}$$

$$\text{Net returns:} \quad r_{t,n} = R_{t,n} - 1 = \frac{P_{t,n}}{P_{t-1,n}} - 1$$

$$\text{Excess returns:} \quad R_{t,n} - R_{t,0} = \frac{P_{t,n}}{P_{t-1,n}} - \frac{P_{t,0}}{P_{t-1,0}} \tag{2}$$

  where $R_{t,0}$ is the return on some reference asset (usually the risk-free asset, whose return is denoted $R_{t,f}$).

- Note that when looking at excess returns, it does not matter whether we look at gross or net returns: excess gross returns are equal to excess net returns.

Multiperiod compound returns

- Multiperiod returns are called compound returns.
- The return over h periods, between date $t-h$ and date t, is

Multiperiod returns:
$$\begin{aligned} R_{t,n}(h) &= \frac{P_{t,n}}{P_{t-h,n}} \\ &= \frac{P_{t,n}}{P_{t-1,n}} \times \frac{P_{t-1,n}}{P_{t-2,n}} \times \ldots \times \frac{P_{t-h+1,n}}{P_{t-h,n}} \\ &= R_{t,n} \times R_{t-1,n} \times \ldots \times R_{t-h+1,n}. \end{aligned}$$

Annualized returns . . . I

- Returns are always quoted with respect to a time period.
- The convention is to quote returns per year (per annum).
- Monthly net returns are annualized by multiplying by 12.
- Monthly gross returns are obtained by adding 1 to the annualized net return.

Annualized returns ... II

- If each return $R_{t,n}$ is a one-year return, then the multi h-year return is computed as

$$R_{t,n}(h) = R_{t,n}\, R_{t-1,n} \dots R_{t-h+1,n} = \prod_{i=0}^{h-1} R_{t-i,n}$$

in which case, the annualized gross return is

$$R_{t,n}(h)^{1/h} = [R_{t,n}\, R_{t-1,n} \dots R_{t-h+1,n}]^{1/h} = \Big[\prod_{i=0}^{h-1} R_{t-i,n}\Big]^{1/h}$$

and the approximate annualized gross return is

$$\approx \frac{1}{h}\Big[\sum_{i=0}^{h-1} R_{t-i,n}\Big]. \tag{3}$$

Continuous compounding . . . I

- Denote the natural logarithm of prices by $p_{t,n} = \ln P_{t,n}$ so that

$$\ln R_{t,n} = \ln \frac{P_{t,n}}{P_{t-1,n}} = \ln P_{t,n} - \ln P_{t-1,n} = p_{t,n} - p_{t-1,n},$$

and for the multiperiod continuously compounded return

$$\begin{aligned}
\ln R_{t,n}(h) &= \ln \left(R_{t,n}\, R_{t-1,n} \ldots R_{t-h+1,n} \right) \\
&= \ln R_{t,n} + \ln R_{t-1,n} + \ldots + \ln R_{t-h+1,n} && (4) \\
&= (p_{t,n} - p_{t-1,n}) + (p_{t-1,n} - p_{t-2,n}) + \ldots + (p_{t-h+1,n} - p_{t-h,n}) \\
&= p_{t,n} - p_{t-h,n} \ldots \text{all other (intermediate) terms cancel out} && (5)
\end{aligned}$$

- There are two important insights from Equations (4) and (5).

Continuous compounding ... II

- Equation (5) shows that for (continuously compounded) returns, only the first and last price observations matters; the rest drop out!
- That is, if you are computing returns over 100 years, the return that you get is the same whether you use
  - just the data for the first price and the last price,
  - annual data on prices for all 100 years,
  - monthly data on prices for all the months in the last 100 years, or
  - daily data on prices for all the days in the last 100 years.
- That is, additional data is not useful for estimating the return. (This is not true for the variance.)
- Consequently, estimates of expected return are very imprecise. (In contrast, estimates of the variance are much more precise.)

Continuous compounding . . . III

- Equation (4) shows that the multiperiod log return is the sum of one-period returns;
- An advantage of this is that it is easier to derive statistical properties of sums of random variables than products.
- However, there is one disadvantage of log returns:
  - while the gross return of the portfolio, denoted $R_{t,p}$, is a weighted sum of the gross returns of the assets in the portfolio

$$R_{t,p} = \sum_{n=1}^{N} w_{t,n} R_{t,n},$$

  - the log return of the portfolio is not a weighted sum

$$\ln R_{t,p} \neq \sum_{n=1}^{N} w_{t,n} (\ln R_{t,n}).$$

Continuous compounding ... IV

- But, if returns are measured over short intervals, then

$$\ln R_{t,n} \approx \sum_{n=1}^{N} w_{t,n}(\ln R_{t,n}).$$

- When studying returns of the cross-section of assets, we usually use gross returns, $R_{t,n}$ or gross excess returns, $R_{t,n} - R_{t,0}$.

- When studying the time-series properties of returns, we usually use log returns, $\ln R_{t,n}$, or log excess returns, $\ln R_{t,n} - \ln R_{t,0}$.

- In this course, because we wish to ask how best to allocate wealth across assets, we will focus on the cross-section of returns.

Returns with dividend payments

- If an asset n pays dividends, $D_{t,n}$, then its gross return is

$$
\begin{aligned}
R_{t,n} &= \frac{P_{t,n} + D_{t,n}}{P_{t-1,n}} \\
\ln R_{t,n} &= \ln \frac{P_{t,n} + D_{t,n}}{P_{t-1,n}} = \ln(P_{t,n} + D_{t,n}) - \ln P_{t-1,n}.
\end{aligned}
$$

Distribution of returns

- The return of asset n at date t is a random variable.
- Random variables are characterized by their distribution.
- There are at least three types of distributions of interest:
  1. Joint distribution,
  2. Conditional distribution, or
  3. Unconditional distribution.

Unconditional distribution . . . I

- It is convenient to assume that gross simple returns have a distribution that is
  - IID (independent and identical distribution)
  - Lognormal
- The assumption of Lognormal returns implies that the log return has a Normal distribution.
  - log returns, $\ln R_{t,n} \sim \mathcal{N}(\mu_n, \sigma_n)$,
    where μ_n is the mean and σ_n is the volatility (standard deviation).
- The Lognormal distribution has a lower bound of 0;
  thus, gross returns do not violate the requirement of limited liability.
  - Note if $R_{t,n}$ has a lower bound of zero, then $\ln R_{t,n}$ takes values from $-\infty$ to $+\infty$.

Unconditional distribution . . . II

- The general result is that the m-th moment of a Lognormally distributed variable X with mean μ and variance σ^2 is given by

$$\mathbb{E}[X^m] = \exp(m\mu + \frac{1}{2}m^2\sigma^2).$$

- If log returns have a Normal distribution, then returns themselves have the following mean and variance:

$$\mathbb{E}[R_n] = \exp\left(\mu_n + \frac{1}{2}\sigma_n^2\right) \qquad \text{. . . the first moment}$$

$$\mathbb{E}[R_n^2] = \exp\left(2\mu_n + 2\sigma_n^2\right) \qquad \text{. . . the second moment}$$

$$\begin{aligned}
\mathbb{V}[R_n] &= \mathbb{E}[R_n^2] - (\mathbb{E}[R_n])^2 && \text{. . . variance, second central moment} \\
&= \exp\left(2\mu_n + 2\sigma_n^2\right) - \exp\left(2\mu_n + \sigma_n^2\right) && \text{. . . [second moment] - [first moment]}^2 \\
&= \exp\left(2\mu_n + \sigma_n^2\right)\left(\exp(\sigma_n^2) - 1\right) && \text{. . . collecting/rearranging terms}
\end{aligned}$$

- You can read more about the Lognormal distribution at:
  - Wikipedia.

Are asset returns Normally distributed?

- At short horizons, asset returns are not Normal; they have
  - weak skewness, and
  - strong excess kurtosis (fat tails).
- If a random variable X has mean μ and variance σ^2, then:

$$\text{Skewness:} \qquad S[X] = \mathbb{E}\left[\frac{(X-\mu)^3}{\sigma^3}\right]$$

$$\text{Kurtosis:} \qquad K[X] = \mathbb{E}\left[\frac{(X-\mu)^4}{\sigma^4}\right].$$

- For a Normally distributed random variable, $S[X] = 0$ and $K[X] = 3$.
- Excess kurtosis is then $K[X] - 3$.

Distribution of skewness and kurtosis

- If you have T with $t = \{1, \ldots, T\}$ time-series observations, then

$$\text{Estimate of skewness:} \qquad \hat{S}[X] = \frac{1}{\hat{\sigma}^3}\frac{1}{T}\sum_{t=1}^{T}\left[(x_t - \hat{\mu})^3\right]$$

$$\text{Estimate of kurtosis:} \qquad \hat{K}[X] = \frac{1}{\hat{\sigma}^4}\frac{1}{T}\sum_{t=1}^{T}\left[(x_t - \hat{\mu})^4\right],$$

$$\text{where:} \qquad \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} x_t$$

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left[x_t - \hat{\mu}\right]^2.$$

- When T is large, then we have the following distribution result: $\hat{S}[X] \sim \mathcal{N}(0, 6/T)$ and $\hat{K}[X] \sim \mathcal{N}(3, 24/T)$.

Road map

Financial and macroeconomic data

- As mentioned before, we will be working with several types of data:
  1. Stock-price data
  2. Stock-characteristics data
  3. Macroeconomic data
- We look at obtaining these three kinds of data as explained on the Tidy Finance website.

Road map

Stock-price data

Packages we will use to access data

```
# *** IMPORTANT ***
# Note that before you can import these packages,
#     you will need to install them on your computer
# How to install packages depends on your Python environment

import pandas as pd   #For more information, see pandas
import numpy as np    #For more information, see NumPy
import yfinance as yf #For more information, see yfinance
```

Example of downloading price data

Price data for Apple from Yahoo Finance

```
prices = (yf.download(tickers="AAPL",
                      start="2000-01-01",
                      end="2024-12-31")
  .reset_index()
  .assign(symbol = "AAPL")
  .rename(columns = {"Date": "date",
                     "Open": "open",
                     "High": "high",
                     "Low": "low",
                     "Close": "close",
                     "Adj Close": "adjusted",
                     "Volume": "volume"
                    })
)
prices.head()
```

Output: Price data for Apple

	date	open	high	low	close	adjusted	volume	symbol
0	2000-01-03	0.936384	1.004464	0.907924	0.999442	0.848323	535796800	AAPL
1	2000-01-04	0.966518	0.987723	0.903460	0.915179	0.776801	512377600	AAPL
2	2000-01-05	0.926339	0.987165	0.919643	0.928571	0.788168	778321600	AAPL
3	2000-01-06	0.947545	0.955357	0.848214	0.848214	0.719961	767972800	AAPL
4	2000-01-07	0.861607	0.901786	0.852679	0.888393	0.754065	460734400	AAPL

- The adjusted prices are corrected for anything that might affect the stock price after the market closes, e.g., stock splits and dividends.
- Because these actions affect quoted prices but have no direct impact on investors holding the stock, we typically use adjusted prices.
- What was the opening price of Apple stock on
  - 2000-01-04? $0.9665;
  - 2023-01-04? $126.98.

History of stock splits for Apple

Date	Stock split
1987 JUN 16	2-for-1
2000 JUN 21	2-for-1
2005 FEB 28	2-for-1
2014 JUN 09	7-for-1
2020 AUG 28	4-for-1

- Aside: To read more about Apple's stock splits, see Apple stock split 2020: what you need to know.

From prices to returns

- We will be working with returns.
- So we need to convert prices to returns.
- From the daily price data, we can get daily net returns:

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

From prices to returns for AAPL

Converting prices to returns

```
returns = (prices
  .sort_values("date")
  .assign(ret = lambda x: x["adjusted"].pct_change())
  .get(["symbol", "date", "ret"])
)
returns.head()
```

	symbol	date	ret
0	AAPL	2000-01-03	NaN
1	AAPL	2000-01-04	-0.084310
2	AAPL	2000-01-05	0.014633
3	AAPL	2000-01-06	-0.086539
4	AAPL	2000-01-07	0.047369

From a single stock to multiple stocks . . . I

- The return computation above is for a single stock.
- It is straightforward to extend the analysis to multiple stocks.
- For example, the Tidy Finance website shows how to compute the returns for all current constituents of the Dow Jones Industrial Average (DJIA) index. To do this,
  1. first download the ticker symbols for all the current constituents of the DJIA, and
  2. then replace in the old code "symbol = "AAPL""
     with the list of all tickers: "symbol = ["AAPL", . . . , . . . ,]"

From a single stock to multiple stocks . . . II

- You should learn how to compute returns for stock constituents of:
  - S&P 500; see
    - S&P 500 data loader by Hamed-Ahangari, or
    - TowardsDataScience, or
    - InsiderFinanceWire.
  - Nasdaq; see
    - Nasdaq data link, or
    - TowardsDataScience.
  - DAX; use
    - either the same approach as for constituents of DJIA or Nasdaq;
    - or see Section 9.5 of the book by Brugière (2020).
      (Also, Sections 1.5, 3.4, and 5.4 of the book may be useful.)

Sample code

- Sample Python code for analyzing World Stock Indices Performances

Sample code from Intan Dea Yutami's website

```
# Import packages that we will use
import numpy as np
import pandas as pd
import yfinance as yf

# Retrieve list of world major stock indices from Yahoo! Finance
df_list = pd.read_html('https://finance.yahoo.com/world-indices/')
majorStockIdx = df_list[0]
majorStockIdx.head()
```

Sample code (continued)

Sample code from Intan Dea Yutami's website (continued)

```
# Get historical price data for all the stock indices
stock_list = []
for s in majorStockIdx.Symbol: # iterate for every stock indices
    # Retrieve data from Yahoo! Finance
    tickerData = yf.Ticker(s)
    tickerDf1 = tickerData.history(period='1d', start='2010-1-1', end='
     2020-9-30')
    # Save historical data
    tickerDf1['ticker'] = s # don't forget to specify the index
    stock_list.append(tickerDf1)
# Concatenate all data
msi = pd.concat(stock_list, axis = 0)
```

- From this point on, you can repeat the earlier steps to calculate returns from prices, etc.
- Or you can look at Intan Dea Yutami's website, in which case, notice also the nice plots (important to learn *data-visualization* skills).

Road map

Stock characteristics . . . I

- Our objective is to obtain good estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$.
- One simple way to do this is to use the sample counterparts of these moments, i.e.,

$$\hat{\mu}_{R_n} = \frac{1}{T}\sum_{t=1}^{T} R_{t,n} \qquad \text{. . . sample mean}$$

$$\hat{\sigma}_{R_{nn}} = \frac{1}{T}\sum_{t=1}^{T} [R_{t,n} - \hat{\mu}_n]^2 \qquad \text{. . . sample variance}$$

$$\hat{\sigma}_{R_{nm}} = \frac{1}{T}\sum_{t=1}^{T} [(R_{t,n} - \hat{\mu}_n)(R_{t,m} - \hat{\mu}_m)] \qquad \text{. . . sample covariance}$$

Stock characteristics . . . II

- But, one could also use other characteristics of each stock to estimate $\mathbb{E}[R_n]$ and $\mathbb{V}[R_n]$. These characteristics could be:
  - Size
  - Value (ratio of book value to market value)
  - Profitability
  - Investment
  - etc.
- A website used by many academic researchers with data on the above firm characteristics is Ken French's data library.
- On the next few pages, we see how one can use Python to download data from Ken French's website.

Using Python to work with stock-characteristics data

- This material is from the section on "Financial Data" on the Tidy Finance website.
- The website makes the eminently sensible suggestion to organize and store data in a single database.
- We have already loaded the packages "pandas" and "numpy"
- In addition to that, we define the date range for which we will be downloading the data.

Load packages and define data range

```
import pandas as pd    #For more information, see pandas
import numpy as np     #For more information, see NumPy

start_date = "1960-01-01"
end_date = "2022-12-31" #You can change this to the current date
```

Using Python for Fama-French data

- For stock characteristics, we will use data from Ken French's data library.
  - We will study how the Fama-French factors are constructed later on in the course.
- To read this data, we will use "pandas-datareader"

Load package for reading data

```
import pandas_datareader as pdr
#For more information, see pandas-datareader
```

Fama-French data on 3 factors: mkt, size, and value

- The dataset "Fama/French 3 Factors" contains the return time series of the market (mkt_excess), size (smb), and value (hml) alongside the risk-free rates, rf.
- The code below reads the data and does some steps to correctly parse all the columns and scale them appropriately.

Code to download Fama-French 3 factors

```
factors_ff3_monthly_raw = pdr.DataReader(
  name="F-F_Research_Data_Factors",
  data_source="famafrench",
  start=start_date,
  end=end_date)[0]

factors_ff3_monthly = (factors_ff3_monthly_raw
  .divide(100)
  .reset_index(names="month")
  .assign(
    month = lambda x: pd.to_datetime(x["month"].astype(str))
  )
  .rename(str.lower, axis="columns")
  .rename(columns = {"mkt-rf" : "mkt_excess"})
)
```

Fama-French data on 5 factors

- The Tidy Finance website also provides code to download the five Fama-French factors (2x3), which includes the return time series of profitability (rmw) and investment (cma) factors.

Code to download Fama-French 5 factors

```
factors_ff5_monthly_raw = pdr.DataReader(
  name="F-F_Research_Data_5_Factors_2x3",
  data_source="famafrench",
  start=start_date,
  end=end_date)[0]

factors_ff5_monthly = (factors_ff5_monthly_raw
  .divide(100)
  .reset_index(names="month")
  .assign(
    month = lambda x: pd.to_datetime(x["month"].astype(str))
  )
  .rename(str.lower, axis="columns")
  .rename(columns = {"mkt-rf" : "mkt_excess"})
)
```

Daily data and data on other factors

- The Tidy Finance website provides Python code to download also:
  - Daily data for the Fama-French factors;
  - Returns on 10 industry factors from Ken French's data library;
  - Factors for the q-factor model in Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705.

Other code and data on stock characteristics

- Python code to access Fama-French data and also other data is available from a number of other websites:
  - PyAnomaly (efficient data download from WRDS using asynco).
  - pyassetpricing.
  - getFamaFrenchFactors.
- An excellent website with data on 331 predictors is Open Source Asset Pricing; for details, see the paper Chen and Zimmermann (2022).

Sources for data on non-US stocks

- investpy is a free Python package to retrieve data from Investing.com, which provides data retrieval for up to
  - 39952 stocks,
  - 82221 funds,
  - 11403 ETFs,
  - 2029 currency crosses,
  - 7797 indices,
  - 688 bonds,
  - 66 commodities,
  - 250 certificates, and
  - 4697 cryptocurrencies.
- investpy allows you to download both recent and historical data from all the financial products indexed at Investing.com.
  - It includes data from countries such as the US, France, Germany, India, Russia, and Spain, among many others.

Sources for data on characteristics of other asset classes

- Code for other characteristics and asset classes other than stocks
  - Martin Waibel's code in Python to reproduce the results in Gu, Kelly, and Xiu (2020) on machine learning. The code also includes options' features used in the paper Bali, Beckmeyer, Moerke, and Weigert (2023).
    - The paper by Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33 (5): 2223–2273 is an excellent starting point to learn machine learning.
  - Code for bond pricing is available from Open Source Bond Asset Pricing with details in Dickerson, Mueller, and Robotti (2023).
    - See also the paper Dick-Nielsen, Feldhütter, Pedersen, and Stolborg (2023) that constructs an error-free dataset for corporate bonds.
  - Python code to clean academic TRACE data following the procedure outlined in the paper by Dick-Nielsen and Poulsen (2019) is available from Martin Waibel's package "PyCleanTrace".

EDGAR: Accessing financia data on companies . . . I

- In the US, companies are required by law to file forms with the Securities and Exchange Commission ("SEC").
- EDGAR, the Electronic Data Gathering, Analysis, and Retrieval is a database system that automates the collection, validation, and indexing of the information submitted by companies.
  - The database is freely available to the public.
- Details of EDGAR can be read on this US SEC website.
- The types of data available on EDGAR are described here.
  - The different forms that companies have to file are described here.

EDGAR: Accessing financia data on companies . . . II

- Python APIs (Application Programming Interface) to access EDGAR are available from:
  - SEC Edgar Downloader.
  - You can build a master index of SEC filings with python-edgar.
  - OpenEDGAR: Open Source Software for SEC EDGAR Analysis.
  - EDGAR Tools to access and analyze SEC filings.
  - Analyzing stock index data with Python and EDGAR.
  - SEC EDGAR API to stream new filings in real time.
  - There is also commercial software code to access EDGAR. sec-api.io.
- A paper that explains how to scrape EDGAR With Python.
- Background information for accessing EDGAR.

EDGAR: Accessing financia data on companies . . . III

- You can get basic information for French companies from Registre du commerce et des sociétés.
- Information about UK companies is stored on Companies house and a second website, but these are less detailed than EDGAR.
- For EU companies, you can get basic information from:
  - European Business Register (EBR) and a second website;
  - first website and a second website for Germany;
  - Ireland;
  - Italy;
  - Switzerland.
- The Canadian version of EDGAR, is called SEDAR.

Road map

Macroeconomic data

- The last kind of data we discuss is macroeconomic data.
- Macroeconomic variables are often used as predictors for the equity premium on the market portfolio.
- Welch and Goyal (2008) reexamine the performance of a large number of predictor variables considered in the academic literature:
  - dividend-price ratio, dividend yield, earnings-price ratio, dividend-payout ratio,
  - three-month US treasury-bill rate, long-term yield, long-term government bond returns, the term spread (the difference between the long-term yield on government bonds and the Treasury bill), default yield spread (the difference between BAA and AAA-rated corporate bond yields),
  - stock variance, inflation, and net-equity expansion (net issues divided by total end-of-year market capitalization of NYSE stocks).

Macroeconomic predictors on Amit Goyal's website

- Amit Goyal maintains the data for these macroeconomic predictors in an XLSX-file stored on a public Google drive.

Details of file with macroeconomic data on Amit Goyal's website

```
sheet_id = "1g4LOaRj4TvwJr9RIaA_nwrXXWTOy46bP"
sheet_name = "macro_predictors.xlsx"
macro_predictors_link = (
  "https://docs.google.com/spreadsheets/d/" + sheet_id +
  "/gviz/tq?tqx=out:csv&sheet=" + sheet_name
)
```

- The Tidy Finance website explains how to download this data.

Macro predictors

Macro predictors from Amit Goyal's website

```
macro_predictors = (
  pd.read_csv(macro_predictors_link, thousands=",")
  .assign(
    month = lambda x: pd.to_datetime(x["yyyymm"], format="%Y%m"),
    IndexDiv = lambda x: x["Index"] + x["D12"],
    logret = lambda x: (np.log(x["IndexDiv"]) -
                        np.log(x["IndexDiv"].shift(1))),
    Rfree = lambda x: np.log(x["Rfree"] + 1),
    rp_div = lambda x: x["logret"] - x["Rfree"].shift(-1),
    dp = lambda x: np.log(x["D12"]) - np.log(x["Index"]),
    dy = lambda x: (np.log(x["D12"]) -
                    np.log(x["D12"].shift(1))),
    ep = lambda x: np.log(x["E12"]) - np.log(x["Index"]),
    de = lambda x: np.log(x["D12"]) - np.log(x["E12"]),
    tms = lambda x: x["lty"] - x["tbl"],
    dfy = lambda x: x["BAA"] - x["AAA"]
  )
  .get(["month", "rp_div", "dp", "dy", "ep", "de", "svar",
        "b/m", "ntis", "tbl", "lty", "ltr", "tms", "dfy",
        "infl"])
  .query("month >= @start_date and month <= @end_date")
  .dropna()
)
```

Other macroeconomic data

- There is a large amount of macroeconomic data available on the internet, most of which can be accessed efficiently using Python.
- For example, the Federal Reserve Bank of St. Louis provides the Federal Reserve Economic Data (FRED), an extensive database for macroeconomic data, with 817,000 US and international time series from 108 different sources.
- The Tidy Finance website explains how to download this data using Python.

Using Python to download data from FRED

- As an illustration, we reproduce the instructions from the Tidy Finance website to get consumer price index (CPI) data that can be found under the CPIAUCNS key.

Downloading CPI from FRED

```
import pandas_datareader as pdr

cpi_monthly = (pdr.DataReader(
  name="CPIAUCNS",
  data_source="fred",
  start=start_date,           # This was defined earlier
  end=end_date                # This was defined earlier
  )
  .reset_index(names="month")
  .rename(columns = {"CPIAUCNS" : "cpi"})
  .assign(cpi=lambda x: x["cpi"] / x["cpi"].iloc[-1])
)
```

- To download different data, we just need to find its key on FRED; e.g., the key for the producer price index for gold ores is PCU2122212122210.

Setting up a database (optional)

- It is extremely useful to store downloaded data in a database.
- The Tidy Finance website explains how to set up an SQLite database (see the bottom of the web page).
- For our course, it will not be required to store data in an SQLite database (but you may still wish to learn how to do this).

Python code for setting up an SQL database (optional)

Three steps for setting up an SQL database

```
# Step 1: Import sqlite3
import sqlite3

#Step 2: Create an SQLite database ‘‘tidy_finance.db’’
tidy_finance = sqlite3.connect("data/tidy_finance.sqlite")

#Step 3: Convert dates, create remote table, copy table to database
(factors_ff3_monthly
  .assign(                        # convert dates to UNIX integers
    month = lambda x:
      ((x["month"]- pd.Timestamp("1970-01-01"))
          // pd.Timedelta("1d"))
  )
  .to_sql(                        # ‘‘to_sql()’’ creates the remote table
    name="factors_ff3_monthly", # import monthly Fama-French data
    con=tidy_finance,
    if_exists="replace",
    index = False)
)
```

Python code for accessing the SQL database

- To access data stored in an SQL database, follow two steps:
  1. Establish the connection to the SQLite database, and
  2. Execute the query to fetch the data.

- The Tidy Finance website provides the code below for how to access an SQLite database (see bottom of the web page).

Accessing an SQL database

```
import pandas   # package to query the database
import sqlite3  # package to connect to the database

tidy_finance = sqlite3.connect("data/tidy_finance.sqlite") # connection
factors_q_monthly = (pd.read_sql_query(                    # query
    sql="SELECT * FROM factors_q_monthly",
    con=tidy_finance,
    parse_dates={"month": {"unit": "D", "origin": "unix"}})
)
```

WRDS, CRSP, and Compustat (optional)

- WRDS (Wharton Research Data Services) is a convenient interface to access asset- and firm-specific data in CRSP (Center for Research in Security Prices) and Compustat.
- The data at WRDS is organized in an SQL database, although they use the PostgreSQL engine.
- The Tidy Finance website explains how to use WRDS to access CRSP and Compustat data and store the data in an SQL database.
- We will not have time to study how to use WRDS in this course.

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will study how to measure the performance of a portfolio.

Our focus will be on understanding that performance should be measured *out-of-sample*.

We will then study various performance metrics used in finance.

To do for next class

- Readings
  - Please read the section on Financial Data in the book by Scheuch, Voigt, Weiss, and Frey (2024), available online at Tidy Finance.
- Assignment
  - You can also start reading (and working) on the first assignment, even though the assignment also depends on the material to be covered in the next chapter.

Road map

Bibliography . . . I

Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert. 2023. Option return predictability with machine learning and big data. *Review of Financial Studies* 36 (9): 3548–3602. (Cited on page 84).

Brugière, P. 2020. Quantitative portfolio management. *Springer Texts in Business and Economics.* (Cited on page 71).

Chen, A. Y., and T. Zimmermann. 2022. Open source cross-sectional asset pricing. *Critical Finance Review* 11 (2): 207–264. (Cited on pages 82, 776).

Dick-Nielsen, J., P. Feldhütter, L. H. Pedersen, and C. Stolborg. 2023. Corporate bond factors: replication failures and a new framework. Available at SSRN 4586652. (Cited on page 84).

Dick-Nielsen, J., and T. K. Poulsen. 2019. How to clean academic TRACE data. *Available at SSRN 3456082.* (Cited on page 84).

Dickerson, A., P. Mueller, and C. Robotti. 2023. Priced risk in corporate bonds. Forthcoming in Journal of Financial Economics. (Cited on page 84).

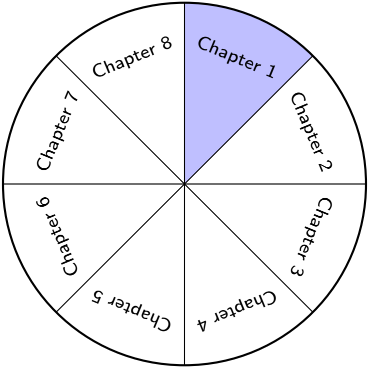
Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33 (5): 2223–2273. (Cited on page 84).

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705. (Cited on pages 81, 132, 775, 945, 966).

Bibliography . . . II

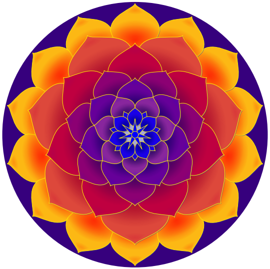
Scheuch, C., S. Voigt, P. Weiss, and C. Frey. 2024. *Tidy finance with Python.* Chapman-Hall/CRC. Available online from this link. (Cited on pages 100, 399, 559).

Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies* 21 (4): 1455–1508. (Cited on page 89).



End of Chapter 1

Quantitative Portfolio Management



Chapter 2:
Performance Measurement

Raman Uppal

2025

The big picture: Plan for the entire book

Part A: Preliminaries

Chapter 1: Properties of asset returns

Chapter 2: Performance measurement (especially out of sample)

Part B: Modern portfolio management

Chapter 3: Mean-variance portfolios that *ignore* estimation error

Chapter 4: Mean-variance portfolios that *adjust for* estimation error

Part C: Post-Modern Portfolio Management

Chapter 5: CAPM-based portfolios: Black-Litterman model

Chapter 6: Factor-based portfolios: Parametric portfolio policies

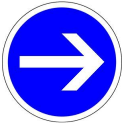
Chapter 7: Volatility-timed factor portfolios

Chapter 8: Portfolios exploiting systematic risk factors *and* unsystematic risk

Table of contents

Road map

What do we want to do in Chapter 2



We first study how to measure the performance of a portfolio strategy; in particular, the importance of evaluating portfolio returns that are out-of-sample, instead of in-sample, which suffer from a *look-ahead* bias.

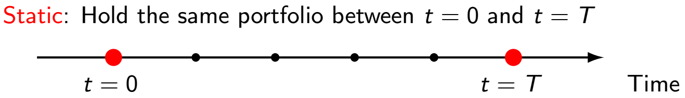
Then, we examine different metrics for evaluating portfolio performance. The key measure we will use is the Sharpe ratio.

We conclude by studying how to test if the difference between the Sharpe ratios of two portfolios is statistically significant.

Road map

Static, myopic, and dynamic portfolios . . . I

- **Static** portfolio: Hold the same portfolio between $t = 0$ and $t = T$:
  - That is, form portfolio at $t = 0$ and hold it until T.

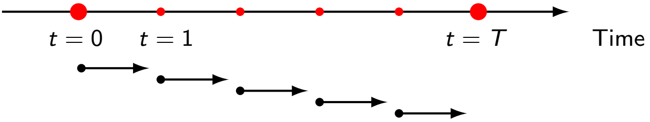


- An example of a static portfolio is:
  - A 25-year old person designs a portfolio for retirement at $t = 0$
  - This person holds the same portfolio for the next 40 years, until T.

Static, myopic, and dynamic portfolios . . . II

- **Myopic** portfolio: Revise portfolio at each date t, but ignore the possibility of rebalancing at future dates.
  - That is, when forming the portfolio at t, ignore the possibility that you will be rebalancing the portfolio again at $t+1$.

Myopic: Rebalance portfolio at each date, but looking only one period ahead

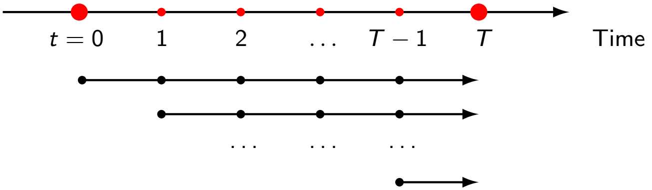


- An example of a myopic portfolio is:
  - A 25-year old person designs a portfolio for retirement at $t = 0$;
  - The person rebalances the portfolio each year for the next 40 years;
  - When rebalancing, the person optimizes for only the next year.

Static, myopic, and dynamic portfolios . . . III

- **Dynamic** portfolio: Choose portfolio at t, knowing we will be rebalancing again at $t+1$.

Dynamic: Rebalance portfolio at each date, anticipating future rebalancing



- An example of a dynamic portfolio is:
  - A 25-year old person designs a portfolio for retirement at $t = 0$
  - The person rebalances the portfolio each year for the next 40 years;
  - When rebalancing, the person optimizes for next year anticipating that she will be rebalancing again the following year.

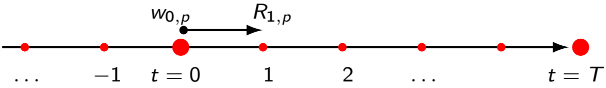
Static, myopic, and dynamic portfolios . . . IV

- We do not study static portfolios because they are sub-optimal relative to myopic (and dynamic) portfolios.
- We do not study dynamic portfolios because to implement them, we need to estimate a larger number of parameters than for myopic portfolios, and so these portfolios perform poorly out of sample.
    - For an excellent description of dynamic portfolios, see: Merton (1971, 1990) and Campbell and Viceira (2002).
- In our course, we will focus on myopic portfolios (even though, if one ignores estimation problems, myopic portfolios are suboptimal relative to dynamic portfolios).

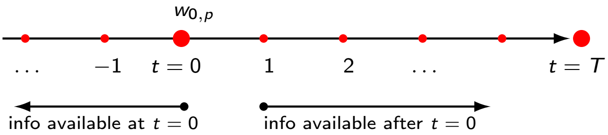
Road map

Portfolio returns

- When evaluating portfolio performance, we will focus on the return of the portfolio.
  - In principle, one could look at other properties of the portfolio, e.g., ESG properties.
- Denote the gross return at date t of portfolio p by $R_{t,p}$.
  - We will adopt the convention of $R_{\text{date,asset-name}}$
- For a myopic portfolio formed at date $t = 0$, $w_{0,p}$, we will want to evaluate its return at date $t = 1$, $R_{1,p}$.

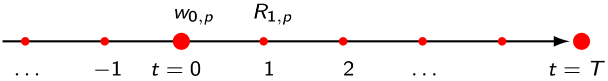


Portfolio weights should depend only on past information



- Consider choosing at date $t = 0$ a myopic portfolio $w_{0,p}$.
- When choosing $w_{0,p}$, it is important that only information available at date $t = 0$ be used to choose the portfolio.
  - $w_{0,p}$ can depend on information available at $t = 0, -1, -2, \ldots$
  - $w_{0,p}$ should not depend on information from the future; that is, no information from $t = 1, 2, \ldots$ should be used.
- It is not realistic to choose weights that depend on future information.

In-sample and out-of-sample portfolio performance



- If $w_{0,p}$ depends only on past information, then $R_{1,p}$ is called the out-of-sample return—because the weights depend on information until $t = 0$ and the return is for $t = 1$, which is out of sample.
- But, if $w_{0,p}$ is chosen using also future information that is available at $t = 1, 2, \dots$, then $R_{1,p}$ is called the in-sample return.
- We will always want to choose portfolio weights that depend on currently available information and not future information.
- Therefore, we will always study out-of-sample performance of the myopic portfolio strategies we consider.

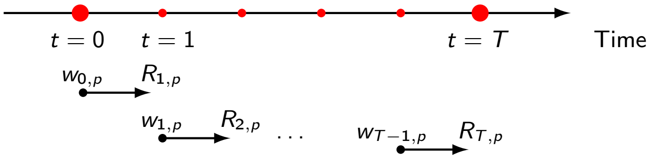
Out-of-sample returns of myopic portfolios

- A myopic portfolio chosen at t will realize its return at $t+1$, where $t = \{0, 1, 2, \ldots, T-1\}$.
- So, the sequence of portfolios

$$w_p = \{w_{0,p}, w_{1,p}, \ldots, w_{T-1,p}\},$$

will generate the sequence of returns

$$R_p = \{R_{1,p}, R_{2,p}, \ldots, R_{T,p}\}.$$



Rolling vs. expanding window . . . I

- Suppose when choosing her portfolio, an investor estimates future expected returns and risk using historical data for a constant interval.
- For example, suppose the investor looks at the past 5 years of monthly data when making her decision at the start of 2020.
  - Then, when making her decision in 2020, the investor will consider data for the past 5 years 2019, 2018, 2017, 2016, and 2015.
  - Similarly, when making her decision in 2021, the investor will consider data for the past 5 years 2020, 2019, 2018, 2017, and 2016.
  - Finally, when making her decision in 2022, the investor will consider data for the past 5 years 2021, 2020, 2019, 2018, and 2017.
- This fixed-length lookback window is called a rolling window.

Rolling vs. expanding window . . . II

- An alternative approach is to estimate future expected returns and risk using historical data for an expanding interval.
  - For example, suppose when making her decision in 2020, the investor considers data until 2015; that is, data for the past 5 years: 2019, 2018, 2017, 2016, and 2015.
  - When making her decision in 2021, the investor will consider data until 2015; that is, data for the past 6 years: 2020, 2019, 2018, 2017, 2016, and 2015.
  - Finally, when making her decision in 2022, the investor will consider data until 2015; that is, data for the past 7 years: 2021, 2020, 2019, 2018, 2017, 2016, and 2015.
- This increasing lookback window is called an expanding window.

Rolling vs. expanding window . . . III

1. We should use a rolling window if we feel that
   - the distant past is not relevant to current decisions,
   - or, if the distant past is different from the more recent past.
2. We should use an expanding window if we feel that
   - the distant past is similar to the recent past.
3. Another alternative is to have an exponentially declining weight on historical data, with the weight decreasing as one goes back in time.

Python functions for rolling and expanding windows

- The pandas library in Python has
  - special functions for doing rolling- and expanding-window analysis;
  - the library also allows for exponentially declining weights on the past.
- Using the links below, you can read about the functions for
  - rolling windows, and
  - expanding windows.
- These functions can be used for some of the course assignments.

Main objective of this course

- The main objective of this course is to understand
  - how to choose the portfolio weights

$$w_p = \{w_{0,p}, w_{1,p}, \ldots, w_{T-1,p}\},$$

  - in order to generate portfolio returns

$$R_p = \{R_{1,p}, R_{2,p}, \ldots, R_{T,p}\}.$$

  - that have good performance.
- In the rest of this class, we will study various metrics of portfolio performance, which we will then use for the rest of the course.

Road map

Measures of portfolio performance

For the return on each portfolio strategy, we compute various performance metrics that can be divided into three groups.

1. Measures of mean (average) returns.
2. Measures of risk of returns.
3. Measures combining mean returns and risk of returns.

Road map

Measures of mean returns

- For the portfolio measures of returns, we use
  1. Total portfolio mean (average) return.
  2. Systematic component of the portfolio mean return based on some factor model for expected returns (see below).
     - The systematic component of portfolio mean returns is the reward for bearing systematic risk.
  3. Alpha component of the portfolio mean return, which is the component of expected returns unexplained by the factor model.
     - The unsystematic component of portfolio mean returns is the reward for bearing unsystematic risk.
  4. Outperformance frequency, which is the average fraction of times that the portfolio being evaluated has a higher cumulative return than the benchmark portfolio within twelve months from the beginning of each such period.

Return factors

- One may believe that the returns on an asset (stocks or portfolios) depend on the exposure of that asset to systematic risk factors.
- One kind of systematic risk factor that is considered in finance is the return on a long-short portfolio.
- Long-short portfolios that have been considered in finance include:
  - $R_{\text{mkt}} - R_f$ = return on market minus return on risk-free asset;
  - R_{smb} = small market capitalization minus big (size);
  - R_{hml} = high book-to-market ratio minus low (value);
  - R_{umd} = up positive momentum minus down momentum;
  - R_{rmw} = robust minus weak profitability;
  - R_{cma} = conservative (low) minus aggressive (high) investment firms.

Factor models of expected returns ... I

- The factor models of expected returns that one can use to compute the systematic and unsystematic component of returns are:

1. One-factor market model of Sharpe (1964),

$$\mathbb{E}[R_p] - R_f = \beta_{p,\mathsf{mkt}}\,\mathbb{E}[R_{\mathsf{mkt}} - R_f].$$

2. Three-factor model of Fama and French (1993),

$$\mathbb{E}[R_p] - R_f = \beta_{p,\mathsf{mkt}}\,\mathbb{E}[R_{\mathsf{mkt}} - R_f] + \beta_{p,\mathsf{smb}}\,\mathbb{E}[R_{\mathsf{smb}}] + \beta_{p,\mathsf{hml}}\,\mathbb{E}[R_{\mathsf{hml}}]$$

3. Four-factor model (three-factors + momentum Carhart (1997),

$$\begin{aligned}\mathbb{E}[R_p] - R_f = {} & \beta_{p,\mathsf{mkt}}\,\mathbb{E}[R_{\mathsf{mkt}} - R_f] + \beta_{p,\mathsf{smb}}\,\mathbb{E}[R_{\mathsf{smb}}] + \beta_{p,\mathsf{hml}}\,\mathbb{E}[R_{\mathsf{hml}}] \\ & + \beta_{p,\mathsf{umd}}\mathbb{E}[R_{\mathsf{umd}}]\end{aligned}$$

Factor models of expected returns ... II

4. Five-factor model of Fama and French (2015),

$$\begin{aligned}\mathbb{E}[R_p] - R_f = {} & \beta_{p,\mathsf{mkt}}\,\mathbb{E}[R_{\mathsf{mkt}} - R_f] + \beta_{p,\mathsf{smb}}\,\mathbb{E}[R_{\mathsf{smb}}] + \beta_{p,\mathsf{hml}}\,\mathbb{E}[R_{\mathsf{hml}}] \\ & + \beta_{p,\mathsf{rmw}}\,\mathbb{E}[R_{\mathsf{rmw}}] + \beta_{p,\mathsf{cma}}\,\mathbb{E}[R_{\mathsf{cma}}]\end{aligned}$$

5. Six-factor model (five-factors + momentum (Fama and French 2018)),

$$\begin{aligned}\mathbb{E}[R_p] - R_f = {} & \beta_{p,\mathsf{mkt}}\,\mathbb{E}[R_{\mathsf{mkt}} - R_f] + \beta_{p,\mathsf{smb}}\,\mathbb{E}[R_{\mathsf{smb}}] + \beta_{p,\mathsf{hml}}\,\mathbb{E}[R_{\mathsf{hml}}] \\ & + \beta_{p,\mathsf{rmw}}\,\mathbb{E}[R_{\mathsf{rmw}}] + \beta_{p,\mathsf{cma}}\,\mathbb{E}[R_{\mathsf{cma}}] \\ & + \beta_{p,\mathsf{umd}}\,\mathbb{E}[R_{\mathsf{umd}}]\end{aligned}$$

6. See also the series of papers on q-factor models that offer an alternative perspective to Fama and French models: Hou, Xue, and Zhang (2015, 2017a, 2017b) and Hou, Mo, Xue, and Zhang (2019).

Factor models of expected returns . . . III

7. K-"factor" model, where factors are principal components (PCs) of a principal component analysis (PCA) of returns (Connor, Goldberg, and Korajczyk 2010; Kozak, Nagel, and Santosh 2018, 2020).
   - These are statistical factors (obtained from an eigenvalue-eigenvector decomposition of the covariance matrix of returns).
   - These factors are often called "latent" (unobservable), in contrast to the observable Fama-French factors.
   - We will study models with latent factors later in the course.

Factor models of expected returns . . . IV

- A recent paper by Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2025) shows that
  - perhaps it is important to adjust mean returns also for unsystematic risk, and
  - once we adjust mean returns for unsystematic risk, the only systematic risk factor needed is the market (or the market and size).
- We will study this paper later in the course.

Road map

Measures of the riskiness of returns . . . I

- We measure the riskiness of portfolio returns using
  1. Volatility (or standard deviation or square root of variance) of portfolio returns.
  2. Semi-volatility (square root of semi-variance) of portfolio returns is calculated in the same manner as the volatility, but only those observations that fall below the mean are included in the calculation.

$$\text{Semi-volatility} = \sqrt{\frac{1}{T}\sum_{\substack{t=1 \\ x_t < \hat{\mu}}}^{T} \left[x_t - \hat{\mu}\right]^2}, \quad \text{where } \hat{\mu} \text{ is the estimated mean.}$$

  3. Skewness of portfolio returns (explained in the previous chapter).
  4. Kurtosis of the portfolio returns (explained in the previous chapter).

Measures of the riskiness of returns . . . II

5. Average maximum drawdown (MDD), defined as the time-series average of the maximum percentage loss of the portfolio value $V(\tau)$ over any period from τ_1 to τ_2 during the last twelve months:

$$\text{MDD} = \frac{1}{T-13} \sum_{t=12}^{T-1} \max_{t-11 \leq \tau_1 < \tau_2 \leq t} \left\{ 0, \frac{V(\tau_1)}{V(\tau_2)} - 1 \right\} \times 100.$$

- This page by Vamshi Jandhyala shows how to use Python to compute drawdown.

Measures of the riskiness of returns . . . III

6. Value at Risk (VaR) measures how much a portfolio may lose (with a given probability, usually 1% or 5%), given normal market conditions, over a set time period such as a day or a month.
   - VaR is typically used by firms and regulators in the financial industry to gauge the amount of assets needed to cover possible losses.
   - For example, if a portfolio has a one-week 99% VaR of $10 million, that means that there is a 1% probability that the portfolio will fall in value by more than $10 million over a one-week period.
   - Thus, a loss of $10 million or more is expected 1 week out of 100 (because of the 1% probability), so about once every two years.
   - See Wikipedia for further details and, in particular, the weaknesses of the VaR measure.

Measures of the riskiness of returns . . . IV

7. Conditional value at risk (CVaR) at $q\%$ level is the expected return on the portfolio in the worst $q\%$ of cases.
   - CVaR is also called: average value at risk (AVaR), expected shortfall, expected tail loss, and superquantile.
   - CVaR is an alternative to value at risk that is more sensitive to the shape of the tail of the loss distribution.
   - See Wikipedia for further details.

Python code

- This web page by Vamshi Jandhyala shows how to use Python to compute:
  - Skewness
  - Kurtosis
  - Semi-volatility
  - Value at Risk, and
  - Conditional Value at Risk (expected shortfall).

Road map

Measures combining mean and risk of returns . . . I

Measures of the risk-return tradeoff include the following.

1. Sharpe ratio (SR), which is the mean return of the portfolio in excess of the risk-free rate divided by the volatility of portfolio returns:

$$\text{SR}_p = \frac{\mathbb{E}[R_p - R_f]}{\sqrt{\mathbb{V}[R_p - R_f]}} = \frac{\text{excess mean return}}{\text{total portfolio risk}}.$$

- Note that (for IID returns) the Sharpe ratio scales with $\sqrt{T}$:

$$\text{SR}_p = \frac{\mathbb{E}[R_p - R_f]}{\sqrt{\mathbb{V}[R_p - R_f]}} = \frac{\text{numerator scales with } T}{\text{denominator scales with } \sqrt{T}} = \text{SR scales with } \sqrt{T}.$$

- Therefore, to annualize the Sharpe ratio of monthly returns, one needs to multiply it by $\sqrt{12} \approx 3.46$.

Measures combining mean and risk of returns . . . II

- Correct scaling to annualize the Sharpe ratio
  - Note that multiplying by $\sqrt{12} \approx 3.46$ is correct only if returns are IID with no serial correlation;
  - Lo (2002, eq. (20) & Table 2) shows that the correct multiplier depends on the serial correlation of the portfolio returns.
  - The correct multiplier can be much smaller than $\sqrt{12}$ (if returns have positive serial correlation);
  - The correct multiplier can be much larger (if returns have negative serial correlation).

Measures combining mean and risk of returns . . . III

- The Sharpe ratio has several limitations:
  - Sharpe ratio is appropriate for efficient portfolios but not for individual assets because it ignores correlation between these assets.
  - Sharpe ratio, because it depends on the mean and variance of returns, is not appropriate if returns have high skewness and kurtosis.
  - Sharpe ratio looks at first two moments, but ignores the distribution of returns. In contrast, the Omega measure uses the entire distribution; see Wikipedia.
  - Despite these limitations, the Sharpe ratio is the main performance measure we will use throughout the course.
- This web page by Vamshi Jandhyala shows how to use Python to compute the Sharpe ratio.
- Links to other websites with code for computing the Sharpe ratio are provided below.

Measures combining mean and risk of returns . . . IV

2. Sortino ratio (SoR) is the mean return in excess of a target rate, divided by the downside variance (DV) of portfolio returns relative to the target rate, R_{target}, which could be the risk-free rate:

$$\text{SoR}_p = \frac{\mathbb{E}[R_p] - R_{\text{target}}}{\sqrt{\text{DV}}}, \quad \text{where}$$

$$\text{DV} = \frac{1}{T} \sum_{i=1}^{T} \Big(\min \left\{ R_p - R_{target}, 0 \right\} \Big)^2,$$

- Compared to Sharpe ratio, Sortino ratio penalizes only downward deviations of the portfolio return from the target rate of return.

Measures combining mean and risk of returns . . . V

- This article by by John Doherty presents a step-wise explanation of computing the Sortino ratio.
- This article by Rollinger and Hoffman presents another example of computing the Sortino ratio.
- This web page from Quantitative Finance explains how to code the Sortino ratio using Python.

Measures combining mean and risk of returns . . . VI

3. Treynor ratio (TR) is the mean return of a portfolio (or individual asset) in excess of the risk-free rate divided by the market beta of returns:

$$\mathrm{TR}_p = \frac{\mathbb{E}[R_p - R_f]}{\beta_{p,\mathrm{mkt}}} = \frac{\text{excess mean return}}{\text{systematic risk}},$$

where $\beta_{p,\mathrm{mkt}} = \frac{\mathrm{Cov}[R_p, R_{\mathbf{mkt}}]}{\mathrm{Var}[R_{\mathbf{mkt}}]} = \frac{\mathbb{C}[R_p, R_{\mathbf{mkt}}]}{\mathbb{V}[R_{\mathbf{mkt}}]}$.

- Because the Treynor ratio ignores unsystematic risk, two portfolios (or assets) with the same Treynor ratio may have very different levels of unsystematic risk.
- Both the Sharpe ratio and Treynor ratio are reasonable for ranking investments, but neither tells us by how much one investment is better than another.

Measures combining mean and risk of returns . . . VII

4. Risk-adjusted performance (RAP or M^2 or M2), developed by Franco Modigliani and his granddaughter, Leah Modigliani, is the excess return of a portfolio adjusted for its risk relative to that of a benchmark, B.

$$
\begin{aligned}
\mathrm{RAP}_p &= \mathbb{E}[R_p - R_f] \times \frac{\mathbb{V}[R_B]}{\mathbb{V}[R_p - R_f]} \\
&= [\text{excess mean return of } p] \times \frac{\text{return variance of benchmark}}{\text{return variance of } p}.
\end{aligned}
$$

- So, if the return variance of the benchmark portfolio is half of that of portfolio p, then the risk-adjusted performance of p will be half of the unadjusted performance, $\mathbb{E}[R_p - R_f]$.
- RAP is useful because, when plotting cumulative returns of different strategies, important to adjust returns so they all have same variance.

Measures combining mean and risk of returns . . . VIII

5. Information Ratio (IR) is the ratio of the active return of a portfolio (or asset) in excess of the benchmark's return to the volatility of the active return (i.e., active risk or benchmark tracking risk)

$$\mathrm{IR}_p = \frac{\mathbb{E}[R_p - R_B]}{\sqrt{\mathbb{V}[R_p - R_B]}} = \frac{\text{mean return in excess of benchmark}}{\text{volatility of excess return}}.$$

- In contrast, the Sharpe ratio uses the risk-free return as the benchmark, whereas IR uses a risky index as a benchmark (such as the market index).
- The Sharpe ratio is useful for evaluating absolute portfolio returns, whereas IR is useful for evaluating relative returns of a portfolio.

Measures combining mean and risk of returns . . . IX

6. Jensen's alpha (α_J) is the ex-post abnormal return of a portfolio(or asset) over the theoretical (model-implied) return

$$\alpha_J = \underbrace{(R_p - R_f)}_{\text{realized excess return}} - \underbrace{(R_{\text{model}} - R_f)}_{\text{model-implied excess return}}.$$

- If the benchmark model is the CAPM: $(R_{\text{model}} - R_f) = \beta_{p,\text{mkt}}(R_{\text{mkt}} - R_f)$, then

$$\alpha_J^{\text{CAPM}} = \underbrace{(R_p - R_f)}_{\text{realized excess return}} - \underbrace{\beta_{p,\text{mkt}}(R_{\text{mkt}} - R_f)}_{\text{CAPM return}} \quad \ldots \text{CAPM alpha}$$

- If the benchmark model is the Fama-French-Carhart four-factor model, then we get the four-factor alpha:

$$\alpha_J^{\text{FFC}} = \underbrace{(R_p - R_f)}_{\substack{\text{realized} \\ \text{excess return}}} - \underbrace{\left[\beta_{p,\text{mkt}}(R_{\text{mkt}} - R_f) + \beta_{p,\text{smb}} R_{\text{smb}} + \beta_{p,\text{hml}} R_{\text{hml}} + \beta_{p,\text{umd}} R_{\text{umd}}\right]}_{\text{FFC return}}$$

Measures combining mean and risk of returns . . . X

7. One could also compute the certainty equivalent return assuming an investor has a particular utility function.
    - The benefit of this measure is that it adjusts returns not just for variance but also for higher moments of return, such as skewness and kurtosis.
    - A disadvantage of this measure is that one needs to specify the parameters of the investor's utility function; in particular, the investor's degree of risk aversion.
    - Because in many cases, the conclusions one draws from evaluating Sharpe ratios are similar to those from evaluating the certainty equivalent return, we will focus on the Sharpe ratio.

Python code

- This page from CodeArmo has Python code to compute the
  - Sharpe ratio,
  - Sortino ratio,
  - Max drawdown, and
  - Calmar ratio.
- This page from Turing Finance has Python code for a large number of measures of risk-adjusted return. In particular:
  - Measures of return adjusted for risk:
    Treynor ratio, Sharpe ratio, Information ratio, Modigliani ratio.
  - Measures of return adjusted for Value-at-Risk
  - Measures of return adjusted for risk based on partial moments:
    Omega, Sortino, Kappa, gain-loss, and upside-potential ratios.
  - Measures of return adjusted for drawdown risk:
    Calmar ratio, Sterling ratio, and Burke ratio.

Road map

Turnover for an individual asset

- When a portfolio is rebalanced, some assets need to be sold while others need to be purchased.
- When the portfolio is rebalanced at time $t+1$, it gives rise to a trade in each asset n of the magnitude

$$\left| w_{t+1,n} - w_{t^+,n} \right|, \quad \text{where}$$

  - $w_{t+1,n}$ is the (new) weight in asset n at date $t+1$, and
  - $w_{t^+,n}$ denotes the (old) date-t weight evaluated at date $t+1$ prices.
- This sale and purchase of assets is called turnover, and it incurs a transaction cost.
- Usually, more sophisticated portfolio strategies have larger turnover.

Turnover for an entire portfolio

- The turnover of a portfolio between dates t and $t+1$ is the sum of absolute changes in weights over the N stocks in the portfolio:

$$\text{Turnover} = \sum_{n=1}^{N} \left(\left| w_{n,t+1} - w_{n,t^+} \right| \right), \quad \text{where}$$

  - $w_{t+1,n}$ is the (new) weight in asset n at date $t+1$, and
  - $w_{t^+,n}$ is the (old) date-t weight evaluated at date $t+1$ prices.

Average turnover for a portfolio over time

- The annualized average monthly turnover of a portfolio is
  - the time-series mean (over the T monthly rebalancing dates) of
  - the sum over the N stocks in the portfolio of the absolute changes in weights
  - multiplied by twelve:

$$\text{Turnover} = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \left(\left| w_{n,t+1} - w_{n,t^+} \right| \right) \times 12,$$

where w_{n,t^+} are the weights of the portfolio using the prices at $t+1$.

Portfolio turnover ratio

- The investment industry often reports the portfolio turnover ratio.
- This is computed as:

$$\text{Portfolio turnover ratio} = \frac{\min(\text{securities sold}, \text{securities purchased})}{\text{average net assets}} \times 100$$

- Example:
  - A portfolio purchased \$10m of securities and sold \$9m of securities, over a one-year time period, so $\min(10, 9) = \$9\text{m}$.
  - Over the one-year period, the portfolio's average net assets were \$100m.
  - The portfolio turnover ratio is: $(\$9\text{m}/\$100\text{m}) \times 100 = 9\%$.
- Note: The measure of turnover (on the previous slide) includes both sales and purchases, while the turnover ratio (on this slide) includes only one side, sales or purchases, depending on which one is smaller.

Turnover and transaction cost . . . I

- To measure the economic cost of turnover, one can compute the transaction costs generated by this turnover.
- One simple approach to measuring transaction cost is to assume that there are only proportional costs for trading (i.e., no fixed costs).
- Then, the transaction cost of a portfolio strategy at $t+1$ is the
  - turnover for date $t+1$ multiplied by
  - κ, the average proportional cost of trading a stock in that portfolio:

$$\kappa \times \sum_{n=1}^{N} \left| w_{t+1,n} - w_{t^+,n} \right|.$$

Turnover and transaction cost . . . II

- Therefore, the net-of-costs evolution of wealth for portfolio p is

$$W_{t+1,p} = W_{t,p} \underbrace{(1 + R_{t+1,p})}_{\text{gross return}} \times \underbrace{\left(1 - \kappa \times \sum_{n=1}^{N} |w_{t+1,n} - w_{t^+,n}|\right)}_{\substack{\text{adjustment for} \\ \text{proportional transaction cost}}},$$

so that the return net of transactions costs is

$$\frac{W_{t+1,p}}{W_{t,p}} - 1 = \underbrace{(1 + R_{t+1,p})}_{\text{gross return}} \times \underbrace{\left(1 - \kappa \times \sum_{n=1}^{N} |w_{t+1,n} - w_{t^+,n}|\right)}_{\substack{\text{adjustment for} \\ \text{proportional transaction cost}}} - 1.$$

Turnover and transaction cost ... III

- One can then set the proportional transaction cost κ equal to an estimate of the average cost of trading the assets in that portfolio.
  - 50 basis points per transaction, as assumed in Balduzzi and Lynch (1999), based on the studies of the cost per transaction for individual stocks on the NYSE by Stoll and Whaley (1983), Bhardwaj and Brooks (1992), and Lesmond, Ogden, and Trzcinka (1999).
  - 10 to 20 basis points if trading large stocks in more recent times.
  - 0 to 5 basis points if you are a hedge fund.
- For a discussion of how the cost of investing has evolved over time, see French (2008).

Limitations of measuring transaction costs this way

- In the slides above, we have said that the three steps for adjusting the return of a portfolio for transaction costs are
  1. First identify the optimal portfolio
  2. Then, measure the turnover of this portfolio
  3. Finally, multiply turnover by the average transaction cost, κ.
- Each of these three steps is not the best way of doing things:
  1. First, transaction costs should be accounted for when identifying the optimal portfolio, not after the portfolio weights are optimized;
  2. Then, the turnover used should be of the portfolio that accounts for transaction costs at the optimization stage (which will be smaller) instead of one that ignores transaction costs (which will be larger).
  3. Finally, one should use the transaction cost for each stock at each date, $\kappa_{t,n}$, instead of a time-series and cross-sectional average, κ.

How to incorporate transaction costs more accurately

- Later on in the course, we will study how to incorporate transaction costs more accurately by
  1. Taking transaction costs into account when optimizing portfolio weights, instead of after choosing the portfolio weights;
  2. Using the transaction cost for each stock at each date, $\kappa_{t,n}$, instead of using some time-series and cross-sectional average;
  3. Finally, netting out trades (across factors) before computing transaction costs.
- For details of how this is done, see, for instance
  - DeMiguel, Martín-Utrera, Nogales, and Uppal (2020),
  - DeMiguel, Martín-Utrera, and Uppal (2024), and
  - Other papers cited in these two papers.

Turnover and capital-gains taxes

- Turnover, in addition to incurring transaction costs, also leads to capital-gains taxes, which we will not study in the course.
- For details on how to account for capital-gains taxes (and transaction costs) when constructing an optimal portfolio, see
  - Dybvig and Koo (1996),
  - DeMiguel and Uppal (2005),
  - and other papers cited in these two papers.

Portfolio turnover for different investment styles

- Active vs. passive management
  - Active managers have higher portfolio turnover, while passive managers, who follow an index, have lower turnover.
- Growth vs. value investing
  - Growth investors invest in rapidly expanding companies so have higher turnover, whereas value investors invest in undervalued companies with long-term potential so have lower turnover.
- Quantitative vs. qualitative analysis
  - Quantitative investors, who rely on data-driven models (as in this course), have higher turnover, while qualitative investors, who focus on fundamentals and management quality, have lower turnover.

Some strategies for controlling portfolio turnover

- When rebalancing the portfolio, account for transaction costs.
  - Taking transaction costs into account when choosing a portfolio can help reduce costs while maintaining an optimal risk-return tradeoff.
- Adopt a long-term investment horizon
  - Focusing on long-term growth can reduce turnover.
- Implement a buy-and-hold strategy
  - Holding assets for longer periods can reduce turnover and taxes.
- Utilize tax-efficient investment vehicles
  - Tax-efficient funds can reduce the impact of turnover on taxes.

Road map

Review of hypothesis testing: Steps involved

- Throughout the course, we will want to evaluate portfolio models.
- This evaluation (inference) will require a test of the hypothesis that our model is better than some benchmark model.
- A typical hypothesis test has the following steps:
  1. Stating the hypothesis.
  2. Finding an appropriate test statistic and its probability distribution.
  3. Specifying the significance level.
  4. Stating the decision rule.
  5. Collecting the data
  6. Calculating the test statistic.
  7. Making the statistical decision.
  8. Making the investment decision.

Review: The null hypothesis . . . I

- If testing a parameter, we want to know if it equals a certain value.
  - For example, we can ask if the mean $\mu = \mu_0$.
- The hypothesis $\mu = \mu_0$ is called the null hypothesis, H_0.
- We can choose the sample mean, $\bar{X}$, as a test statistic.
- If $\bar{X}$ is distributed as a normal with mean μ_0 and variance $\frac{\sigma^2}{T}$, i.e. if the null hypothesis is true, then we expect that $\bar{X}$ should be sufficiently close to μ_0.
- So, the test consists of not rejecting H_0 if $\bar{X}$ is "close" to μ_0.

Review: The null hypothesis . . . II

- To determine closeness, we can use the following probability statement (based on the standard Normal distribution):

$$P\left[\mu - 1.96\frac{\sigma}{\sqrt{T}} \leq \bar{X} \leq \mu + 1.96\frac{\sigma}{\sqrt{T}}\right] = 0.95.$$

- Under H_0, we have that $\mu = \mu_0$, so we can replace μ by μ_0:

$$P\left[\mu_0 - 1.96\frac{\sigma}{\sqrt{T}} \leq \bar{X} \leq \mu_0 + 1.96\frac{\sigma}{\sqrt{T}}\right] = 0.95.$$

Review: The alternative hypothesis: Rejection region

- In this case, "sufficiently close" is about two times the standard deviation, $\frac{\sigma}{\sqrt{T}}$.
- The rejection region of H_0 is formed by the values of X in
  - $\bar{X} < \mu_0 - 1.96\frac{\sigma}{\sqrt{T}}$, or
  - $\bar{X} > \mu_0 + 1.96\frac{\sigma}{\sqrt{T}}$.
- The rejection region corresponds to the alternative hypothesis H_1, where $\bar{X}$ is Normal with a mean greater in absolute value than μ_0.

Review: Standardizing the parameter to be tested

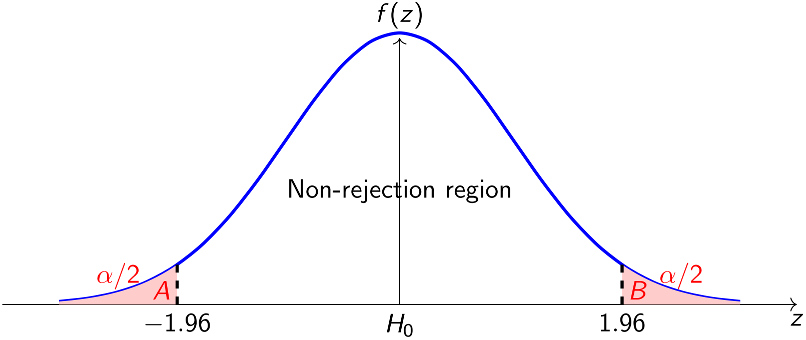
- To evaluate the rejection region, we standardize $\bar{X}$ so that its
  - mean is 0 . . . done by subtracting the mean of $\bar{X}$, μ_0
  - variance is 1 . . . done by dividing by standard deviation of $\bar{X}$, $\sigma/\sqrt{T}$.
- That is, instead of evaluating $\bar{X}$, we evaluate z, where:

$$z = \frac{\bar{X} - \mathbb{E}[\bar{X}]}{\sqrt{\mathbb{V}[\bar{X}]}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{T}}.$$

- The rejection region is then formed by the values of z in
  - $A : \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{T}}} = z < -1.96$, or
  - $B : \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{T}}} = z > +1.96$.

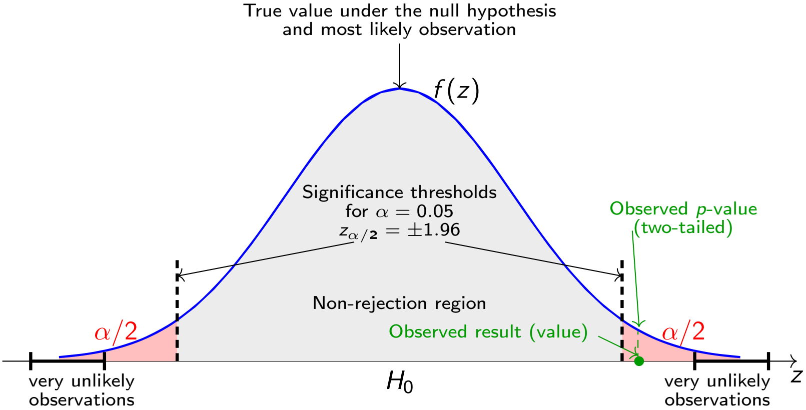
Review: Two-sided critical region

- The two regions A and B form the critical region of the test statistic.
- If the computed value for $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{T}}$ is in this region, we reject H_0.



Review: P-value and rejection region

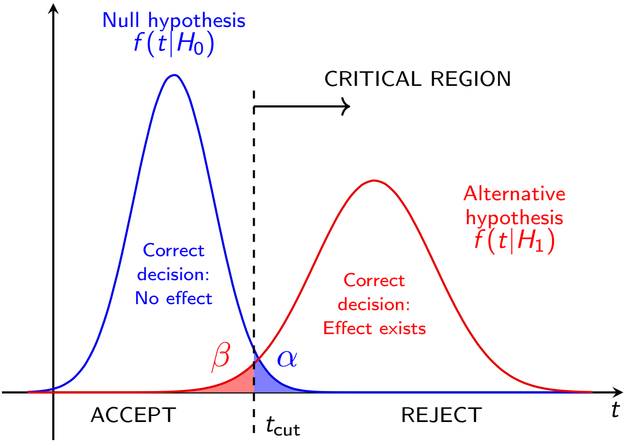
- The *p*-value is the probability of obtaining test results at least as extreme as the result actually observed, assuming the null hypothesis is correct.
- A very small *p*-value means that such an extreme observed outcome is very unlikely under the null hypothesis, so one can reject the null. (Wikipedia)



Review: Type I and Type II errors

- The value 1.96 results from building the interval with a confidence level of 95%.
  - Therefore, there is still a 5% probability that the true value is outside the interval.
- Similarly, we can make errors with hypothesis tests.
  - First, we can reject H_0 while it is true (since we left a 5% probability that it happens). We call this error a Type I error.
  - We can also not reject H_0 while it is false, that is, when one of the alternatives is true. This will be a Type II error.
- The probability of a Type II error depends on the particular alternative that is true in the set described by H_1.

Review: Type I and Type II errors



Type I error: reject H_0 while it is true.

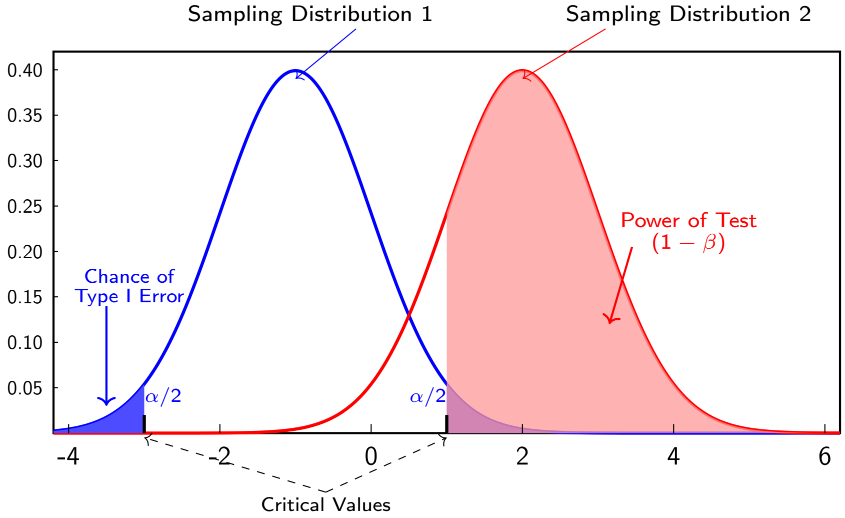
Type II error: not reject H_0 while it is false.

Review: Power of the test . . . I

- Ideally, we wish to minimize both error types, but this is impossible (if we reduce Type I error, we increase Type II error).
- The usual approach is to set Type I error at a certain level (usually at 5%), called the level of the test, and minimize Type II error.
- In fact, we maximize the power of the test, defined as.

$$\text{Power} = 1 - \text{Prob}(\text{Type II error})$$

Review: Power of the test . . . II



Review: Power of the test . . . III

- When drawing inferences using a test statistic, it is important to evaluate the power of the test statistic.
  - The power is the probability that the null hypothesis will be rejected given that an alternative hypothesis is true.
  - Low power indicates that the test is not useful to discriminate between the alternative and the null.
  - For a fixed value of N, considerable increases in power are possible with larger values of T.
  - The power gain is substantial when N is reduced for a fixed alternative.
  - But, when N is large and T is small, power will be low.

Example: Tests with standard Normal distribution . . . I

- Suppose we want to test if the mean of a population is equal to μ_0.
  - The null hypothesis is $H_0 : \mu = 15$.
  - The alternative hypothesis $H_1 : \mu \neq 15$.
- We formulate the hypothesis directly on the standardized statistic:

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{T}}$$

- Suppose we find: $\bar{X} = 13, \sigma = 3.0456$, $T = 30$.
- For these values of $\bar{X}$ and σ and for the given null, the value for z is:

$$z = \frac{13 - 15}{3.0456/\sqrt{30}} = -3.597$$

Example: Tests with standard Normal distribution . . . II

- Should we accept or reject H_0 ?
- For α equal 0.05 :

$$P\left[-1.96 \leq \frac{\bar{X}-\mu_0}{\sigma/\sqrt{T}} \leq 1.96\right] = 0.95.$$

- We reject H_0 because the critical region is $z < -1.96$ and $z > 1.96$, and on the previous slide we calculated -3.597.
- The critical values depend on α and the degrees of freedom.

Example: Single-sided and two-sided test

- In the example above, the alternative hypothesis was of the form $H_1 : \mu \neq \mu_0$, which is
  - two-sided test because it involves the two tails of the distribution.
- If the alternative hypothesis was $H_1 : \mu < \mu_0$, we would have a
  - one-sided test because only the left tail would be relevant.
- In the previous example, for a one-sided test at 5%, the critical region will change from ± 1.96 to

$$z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} < -z_{5\%} = -1.645.$$

Example: Probability value (P-values)

- By convention, hypothesis tests have been carried out at levels of 1%,5% or 10%. However, a test can be carried out at any level.
- The probability value (p-value) is the lowest level at which a null hypothesis can be rejected.
- To find the p-value we evaluate the probability of the value found for z (-3.597 in the example) in the standard Normal distribution. $P[z < -3.597] = 0.000160954$.
- If we set α at 0.000160954, we could reject at this level the null hypothesis that $\mu_x = 15$.
- The advantage of reporting the p-value is that we avoid the arbitrariness of conventional levels, and we let the reader decide whether or not to reject a hypothesis.

Start of focus

Road map

Test of the difference in Sharpe ratios for Normal returns

- For evaluating the performance of a portfolio n relative to a benchmark m, we will compare the Sharpe ratios of n and m; i.e.,
  - instead of testing whether $\bar{X} = \mu_0$,
  - we will now test whether $\mathrm{SR}_n - \mathrm{SR}_m = 0$.
- So, compared to the earlier example,
  - $\bar{X}$ has been replaced by the difference in Sharpe ratios, $\mathrm{SR}_n - \mathrm{SR}_m$;
  - μ_0 has been replaced by 0.
- To test for the difference in Sharpe ratios, we will use two methods.
  1. "Delta method"
  2. "Bootstrap method"

Road map

The delta method

- The test for the difference in Sharpe ratios, using the delta method, was
  - developed by Jobson and Korkie (1981),
  - with a correction provided by Memmel (2003).
- Then, this test was extended to non-IID and non-Normal distributions by Lo (2002), Opdyke (2007), and Bailey and Lopez de Prado (2012).

Underlying logic of the Delta method

From: Wikipedia

- In statistics, the delta method is a method of deriving the asymptotic distribution of a random variable.
- It is applicable when the random variable being considered can be defined as a differentiable function of a random variable that is asymptotically Gaussian.
- The delta method allows us to approximate the distribution of a function of an estimator, assuming the estimator is asymptotically normal.

Central Limit Theorem (CLT) setup

- Let $\theta_0 \in \mathbb{R}^p$ and an estimator $\hat{\theta}$ satisfy a CLT:

$$\sqrt{T}\,(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0,\ \Omega)\,,$$

with Ω positive semidefinite.

- Typically, $\hat{\theta}$ is a sample mean and Ω is a (long-run) variance.
- But, $\hat{\theta}$ could also be a set of moments, not just a single moment.

Univariate Delta Method (1st order)

- If $f : \mathbb{R}^p \to \mathbb{R}$ is a differentiable function at θ_0 with gradient $\nabla f(\theta_0)$, then

$$\sqrt{T}\,\{f(\hat{\theta}) - f(\theta_0)\} \xrightarrow{d} \mathcal{N}(0,\ V), \quad \text{where:} \quad V \equiv \nabla f(\theta_0)^\top \,\Omega\, \nabla f(\theta_0).$$

- Hence, an approximate standard error is

$$\widehat{\mathrm{se}}\{f(\hat{\theta})\} = \sqrt{\widehat{V}/T},$$

where $\widehat{V}$ plugs in consistent estimates.

Obtaining p-values via the delta method

- Test $H_0 : f(\theta_0) = c \quad$ vs. $\quad H_1 : f(\theta_0) \neq c$.
- Statistic:
$$z \equiv \frac{f(\hat{\theta}) - c}{\widehat{\mathrm{se}}\{f(\hat{\theta})\}} \approx \mathcal{N}(0, 1).$$
- Two-sided p-value: $p = 2 \cdot \Phi(-|z|)$.
- One-sided variants are analogous.

Intuition underlying the delta method

- The intuition of the delta method is that any such f function, in a "small enough" range of the function, can be approximated via a first-order Taylor series (which is basically a linear function).
- If the random variable is roughly normal, then a linear transformation of it is also normal.
- A small range can be achieved when approximating the function around the mean, when the variance is "small enough".
- When f is applied to a random variable such as the mean, the delta method tends to work better as the sample size increases, because it helps reduce the variance of the mean.
- Thus, the Taylor approximation is applied over a smaller range of the function f at the point of interest.

Standardizing the difference in Sharpe ratios

- Recall that when testing whether $\bar{X} = \mu_0$, we constructed a standardized variable

$$z = \frac{\bar{X} - \mathbb{E}[\bar{X}]}{\sqrt{\mathbb{V}[\bar{X}]}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{T}}.$$

- Now, to test whether $\mathrm{SR}_n - \mathrm{SR}_m = 0$, we will again need to construct a standardized variable

$$\begin{aligned} z &= \frac{(\mathrm{SR}_n - \mathrm{SR}_m) - 0}{\sqrt{\mathbb{V}[\mathrm{SR}_n - \mathrm{SR}_m]}} \\ &= \frac{(\mathrm{SR}_n - \mathrm{SR}_m)}{\sqrt{\mathbb{V}[\mathrm{SR}_n - \mathrm{SR}_m]}}. \end{aligned}$$

- So, the only thing that we need for the above test is the variance of the difference in the Sharpe ratios, $\mathbb{V}[\mathrm{SR}_n - \mathrm{SR}_m]$.

Difference in Sharpe ratios (Jobson-Korkie-Memmel)

- We start by writing down the final result so that we all know where we are going, after which we will see how this result is derived.
- The test statistic is

$$z = \frac{\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m}{\sqrt{\mathbb{V}\left[\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m\right]}},$$

where the variance of the difference in Sharpe ratios is

$$\boxed{\mathbb{V}\left[\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m\right] = \frac{1}{T}\left[2 - 2\hat{\rho}_{nm} + \frac{1}{2}\left(\widehat{\mathrm{SR}}_n^2 + \widehat{\mathrm{SR}}_m^2 - 2\widehat{\mathrm{SR}}_n\widehat{\mathrm{SR}}_m\hat{\rho}_{nm}^2\right)\right]}$$

in which $\hat{\rho}_{nm}$ denotes the estimated correlation between the excess returns of portfolios n and m.

Understanding the test for difference in Sharpe ratios

- To understand the test for the difference in Sharpe ratios of two portfolios, we will proceed in **three small steps** by studying, for IID normal returns, the:
  1. Distribution of μ_n and σ_n^2;
  2. Distribution of $SR_n = f(\mu_n, \sigma_n) = \mu_n/\sigma_n$;
  3. Distribution of $SR_n - SR_m = f(\mu_n, \mu_m, \sigma_n, \sigma_m)$;
- Then, after this discussion, we will study what to do if returns are not IID normal.

Step 1: Distribution of μ_n and σ_n^2 ... I

- Denote by
  - $u = \{\mu_n, \sigma_n^2\}$ the mean and variance of the excess returns;
  - $\hat{u}$ the empirical counterpart of u, based on T observations;
- Then, the large-sample result for $u = \{\mu_n, \sigma_n^2\}$ is:

$$\sqrt{T}\left(\begin{bmatrix} \hat{\mu}_n \\ \hat{\sigma}_n^2 \end{bmatrix} - \begin{bmatrix} \mu_n \\ \sigma_n^2 \end{bmatrix}\right) \rightarrow \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & 2\sigma_n^4 \end{bmatrix}\right).$$

- This can be written more compactly as:

$$\sqrt{T}(\hat{u} - u) \rightarrow \mathcal{N}(0_2, \Omega), \quad \text{with} \quad \Omega = \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & 2\sigma_n^4 \end{bmatrix}.$$

Step 1: Distribution of μ_n and σ_n^2 ...II

- In plain language, the above says that as T increases,
  - $\sqrt{T}(\hat{u} - u)$ converges to a Normal distribution
  - with a 2-dimensional mean vector of zeros, and
  - a (2×2)-dimensional covariance matrix given by Ω.
- Thus, the asymptotic estimation errors (variances) of $\hat{\mu}_n$ and $\hat{\sigma}_n^2$ are

$$\mathbb{V}[\hat{\mu}_n] \stackrel{a}{=} \frac{\sigma_n^2}{T}, \quad \text{and} \quad \mathbb{V}[\hat{\sigma}_n^2] \stackrel{a}{=} \frac{2\sigma_n^4}{T}.$$

- Thus, as T increases, estimation error decreases.
- For the derivation of the result that $\mathbb{V}[\hat{\sigma}_n^2] = 2\sigma_n^4/T$, look for "Distribution of the sample variance" on this Wikipedia page (toward the bottom of the page).

Step 2: Distribution of SR_n . . . I

- On the previous slide, we have derived the asymptotic distribution of $u = \{\mu_n, \sigma_n^2\}$.
- The Sharpe Ratio is a function of u:

$$\mathrm{SR}_n(u) = \mathrm{SR}_n(\mu_n, \sigma_n^2) = \frac{\mu_n}{\sqrt{\sigma_n^2}} = \frac{\mu_n}{\sigma_n}.$$

- Therefore, we need to find the asymptotic distribution of $\mathrm{SR}_n(u)$.

Step 2: Distribution of SR_n . . . II

- Denote by $f(\cdot)$ a differentiable function of u.
- Then, as we have seen above, a result from statistics tells us

$$\sqrt{T}\big(f(\hat{u}) - f(u)\big) \to \mathcal{N}\left(0, f_u^\top \Omega\, f_u\right), \tag{6}$$

  where f_u is the first derivative (gradient) of $f(u)$ with respect to u.
- In plain language, the above says that as T increases,
  - $\sqrt{T}\big(f(\hat{u}) - f(u)\big)$ converges to a Normal distribution
  - with mean 0, and
  - variance $f_u^\top \Omega\, f_u$ . . . we compute this on the next page.

Step 2: Distribution of SR_n ...III

- Recall that

$$u = \{\mu_n, \sigma_n^2\}. \tag{7}$$

- Set the function $f(\cdot)$ to be

$$f(u) = f(\mu_n, \sigma_n^2) = \mathrm{SR}_n(u) = \frac{\mu_n}{\sqrt{\sigma_n^2}}. \tag{8}$$

- Calculating the first partial derivatives of $f(u)$ (i.e., differentiating $f(u)$ in (8) with respect to the elements of u in (7)), we get

$$\frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f}{\partial \mu_n} \\ \frac{\partial f}{\partial \sigma_n^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{\mu_n}{2\sigma_n^3} \end{bmatrix}. \tag{9}$$

Step 2: Distribution of SR_n . . . IV

- We wish to compute $f_u^\top \Omega\, f_u$, where

$$\frac{\partial f}{\partial u} = f_u = \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{\mu_n}{2\sigma_n^3} \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & 2\sigma_n^4 \end{bmatrix}.$$

- So, we need to multiply: (i) the transpose of f_u; (ii) Ω; (iii) f_u:

$$\begin{aligned} f_u^\top \Omega\, f_u &= \begin{bmatrix} \frac{1}{\sigma_n} & -\frac{\mu_n}{2\sigma_n^3} \end{bmatrix} \cdot \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & 2\sigma_n^4 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{\mu_n}{2\sigma_n^3} \end{bmatrix} \\ &= 1 + \frac{1}{2}\frac{\mu_n^2}{\sigma_n^2} \\ &= 1 + \frac{1}{2}\mathrm{SR}_n^2. \end{aligned}$$

Step 2: Distribution of SR_n ... V

- Then, using the result in (6) that

$$\sqrt{T}\left(f(\hat{u}) - f(u)\right) \to \mathcal{N}\left(0, f_u^\top \Omega \, f_u\right),$$

we get the expression we need:

$$\sqrt{T}\left(\widehat{\mathrm{SR}}_n - \mathrm{SR}_n\right) \to \mathcal{N}\left(0, 1 + \frac{1}{2}\mathrm{SR}_n^2\right).$$

- So the variance of the estimated Sharpe ratio is

$$\mathbb{V}[\widehat{\mathrm{SR}}_n] = \frac{1}{T}\left[1 + \frac{1}{2}\mathrm{SR}_n^2\right].$$

Step 2: Distribution of SR_n ... VI

- Thus, the standard error (SE) for the Sharpe-ratio estimator $\widehat{\mathrm{SR}}_n$ is

$$\mathrm{SE}\left[\widehat{\mathrm{SR}}_n\right] = \sqrt{\frac{1}{T}\left[1 + \frac{1}{2}\mathrm{SR}_n^2\right]}.$$

- Thus, the standardized z-statistic is

$$z = \frac{\mathrm{SR}_n - \mathrm{SR}_0}{\mathrm{SE}\left[\widehat{\mathrm{SR}}_n\right]} = \frac{\mathrm{SR}_n - \mathrm{SR}_0}{\sqrt{\frac{1}{T}\left[1 + \frac{1}{2}\mathrm{SR}_n^2\right]}} \approx \mathcal{N}(0,1).$$

- The 95% Confidence Interval (CI) for the Sharpe ratio is

$$\mathrm{CI} = \widehat{\mathrm{SR}}_n \pm 1.96 \times \sqrt{\frac{1}{T}\left[1 + \frac{1}{2}\mathrm{SR}_n^2\right]}.$$

- Robust variants replace σ^2/T by a HAC long-run variance for the mean and adjust the variance-of-variance accordingly.

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ... I

- We now derive the distribution of the difference in the Sharpe ratios of two portfolio returns.
- Consider two portfolios with excess returns over the risk-free rate given by $R_{t,n}$ and $R_{t,m}$.
- Assume that these excess returns are IID (independently and identically distributed) Normal; that is, there is
  - no change in the distribution over time; i.e., the probability distribution of $R_{t,n}$ is the same as $R_{t',n}$ for any two dates t and t';
  - no serial dependence; i.e., $R_{t,n}$ and $R_{t',n}$ are independent.

Step 3: Distribution of $(\text{SR}_n - \text{SR}_m)$ ... II

- Denote by
  - $u = \{\mu_n, \mu_m, \sigma_n^2, \sigma_m^2\}$ the means and variances of the excess returns;
  - $\hat{u}$ the empirical counterpart of u, based on T observations;
  - σ_{nm} the covariance between $R_{t,n}$ and $R_{t,m}$, and
  - $\rho_{nm} = \sigma_{nm}/(\sigma_n \sigma_m)$ the correlation between the excess returns.
- Note that now $u = \{\mu_n, \mu_m, \sigma_n^2, \sigma_m^2\}$ has four elements in it.

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ...III

- Then, the large-sample result for $u = \{\mu_n, \mu_m, \sigma_n^2, \sigma_m^2\}$ is:

$$\sqrt{T}(\hat{u} - u) \to N\left(0_4, \Omega\right), \quad \text{with}$$

$$\Omega = \begin{bmatrix} \sigma_n^2 & \sigma_{nm} & 0 & 0 \\ \sigma_{nm} & \sigma_m^2 & 0 & 0 \\ 0 & 0 & 2\sigma_n^4 & 2\sigma_{nm}^2 \\ 0 & 0 & 2\sigma_{nm}^2 & 2\sigma_m^4 \end{bmatrix}$$

- In plain language, the above says that as T increases,
  - $\sqrt{T}(\hat{u} - u)$ converges to a Normal distribution
  - with a 4-dimensional mean vector of zeros, and
  - a (4×4)-dimensional covariance matrix given by Ω.

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ...IV

- On the previous slide, we have derived the asymptotic distribution of $u = \{\mu_n, \mu_m, \sigma_n^2, \sigma_m^2\}$.
- We want to find the asymptotic distribution of the function:

$$f(u) = f(\mu_n, \mu_m, \sigma_n^2, \sigma_m^2) = \mathrm{SR}_n - \mathrm{SR}_m = \frac{\mu_n}{\sigma_n} - \frac{\mu_m}{\sigma_m}.$$

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ . . . V

- Let $f(\cdot)$ be a differentiable function of u, then a result from statistics tells us that

$$\sqrt{T}\big(f(\hat{u}) - f(u)\big) \to \mathcal{N}\left(0, f_u^\top \Omega\, f_u\right), \tag{10}$$

  where f_u is the first derivative of $f(u)$ with respect to u.

- In plain language, the above says that as T increases,
  - $\sqrt{T}\big(f(\hat{u}) - f(u)\big)$ converges to a Normal distribution
  - with mean 0, and
  - variance $f_u^\top \Omega\, f_u$ . . . we compute this on the next page.

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ... VI

- Note that now we define

$$u = \{\mu_n, \mu_m, \sigma_n^2, \sigma_m^2\}. \tag{11}$$

- Set the function $f(\cdot)$ to be

$$f(u) = \mathrm{SR}_n - \mathrm{SR}_m = \frac{\mu_n}{\sigma_n} - \frac{\mu_m}{\sigma_m}. \tag{12}$$

- Calculating the first partial derivative of $f(u)$ (i.e., differentiating $f(u)$ in (12) with respect to the elements of u in (11)), we get

$$\frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f}{\partial \mu_n} \\ \frac{\partial f}{\partial \mu_m} \\ \frac{\partial f}{\partial \sigma_n^2} \\ \frac{\partial f}{\partial \sigma_m^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{1}{\sigma_m} \\ -\frac{\mu_n}{2\sigma_n^3} \\ \frac{\mu_m}{2\sigma_m^3} \end{bmatrix}. \tag{13}$$

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m) \ldots$ VII

- We wish to compute $f_u^\top \Omega f_u$, where

$$\frac{\partial f}{\partial u} = f_u = \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{1}{\sigma_m} \\ -\frac{\mu_n}{2\sigma_n^3} \\ \frac{\mu_m}{2\sigma_m^3} \end{bmatrix} \quad \text{and} \quad \Omega = \begin{bmatrix} \sigma_n^2 & \sigma_{nm} & 0 & 0 \\ \sigma_{nm} & \sigma_m^2 & 0 & 0 \\ 0 & 0 & 2\sigma_n^4 & 2\sigma_{nm}^2 \\ 0 & 0 & 2\sigma_{nm}^2 & 2\sigma_m^4 \end{bmatrix}.$$

- So, we need to multiply: (i) the transpose of f_u; (ii) Ω; (iii) f_u:

$$f_u^\top \Omega f_u = \begin{bmatrix} \frac{1}{\sigma_n} & -\frac{1}{\sigma_m} & -\frac{\mu_n}{2\sigma_n^3} & \frac{\mu_m}{2\sigma_m^3} \end{bmatrix} \cdot \begin{bmatrix} \sigma_n^2 & \sigma_{nm} & 0 & 0 \\ \sigma_{nm} & \sigma_m^2 & 0 & 0 \\ 0 & 0 & 2\sigma_n^4 & 2\sigma_{nm}^2 \\ 0 & 0 & 2\sigma_{nm}^2 & 2\sigma_m^4 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sigma_n} \\ -\frac{1}{\sigma_m} \\ -\frac{\mu_n}{2\sigma_n^3} \\ \frac{\mu_m}{2\sigma_m^3} \end{bmatrix}$$

$$= 2 - 2\rho_{nm} + \frac{1}{2}\left(\mathrm{SR}_n^2 + \mathrm{SR}_m^2 - 2\mathrm{SR}_n\mathrm{SR}_m\rho_{nm}^2\right), \quad \text{where } \rho_{nm} = \sigma_{nm}/(\sigma_n\sigma_m).$$

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ... VIII

- Then, using the result in (10) that

$$\sqrt{T}\left(f(\hat{u}) - f(u)\right) \to N\left(0, f_u^\top \Omega\, f_u\right), \tag{14}$$

we get the expression we need:

$$\sqrt{T}\left((\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m) - (\mathrm{SR}_n - \mathrm{SR}_m)\right) \to N\left(0, f_u^\top \Omega\, f_u\right),$$

- so the variance of the difference in Sharpe ratios is

$$\begin{aligned}
\mathbb{V}\left[\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m\right] &= \frac{1}{T}\left(\frac{\partial f}{\partial u}\right)^\top \Omega \left(\frac{\partial f}{\partial u}\right) \\
&= \frac{1}{T}\left[2 - 2\rho_{nm} + \frac{1}{2}\left(\mathrm{SR}_n^2 + \mathrm{SR}_m^2 - 2\mathrm{SR}_n\mathrm{SR}_m\rho_{nm}^2\right)\right] \\
&= \frac{1}{T}\left[\left(1 + \frac{1}{2}\mathrm{SR}_n^2\right) + \left(1 + \frac{1}{2}\mathrm{SR}_m^2\right)\right] \quad \ldots \text{if } \rho_{nm} = 0.
\end{aligned}$$

Step 3: Distribution of $(\mathrm{SR}_n - \mathrm{SR}_m)$ ...IX

- Just as we did for the Sharpe ratio, we can use the expression for $\mathbb{V}\left[\widehat{\mathrm{SR}}_n - \widehat{\mathrm{SR}}_m\right]$ on the previous page to:
  - Compute the Standard Error for the difference in Sharpe ratios;
  - Construct the Confidence Interval for the difference in Sharpe ratios.
  - Construct the p-value for the difference in Sharpe ratios.

Road map

Limitations of the delta method

- The test of Jobson and Korkie (1981) and Memmel (2003) is not valid if
  - **Non-normality:** Returns are heavy-tailed and skewed.
  - **Time dependence:** Autocorrelation and conditional heteroskedasticity (volatility clustering).
  - **Correlation across strategies:** Returns of the two strategies are not independent (e.g., because of common risk factors).
- The problems that arise with non-Normal returns are illustrated in this excellent article by QuantPy, which includes Python code.
  - I strongly recommend that you read this article.
- An alternative approach to the delta method is bootstrapping.

Bootstrapping

From Wikipedia

- Bootstrapping is a procedure for estimating the distribution of an estimator by resampling one's data (usually with replacement).
- Bootstrapping allows estimation of the sampling distribution of almost any statistic using random sampling methods.
- The bootstrap is often used as an alternative to statistical inference based on a parametric model when
  - the assumptions of the parametric model are in doubt, or
  - when parametric inference is impossible, or
  - the parametric model requires complex formulas to calculate standard errors.

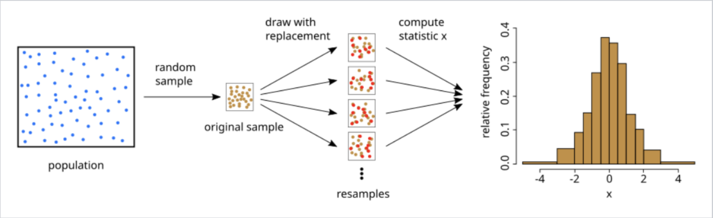
Bootstrap: Intuitive explanation . . . I

- Suppose we are interested in the average height of people in France.
- We sample only a small part—of size N—of the French population.
- From that single sample, we obtain only a single estimate of the mean.
- In order to reason about the population, we need some sense of the variability of the mean that we have computed.
- The simplest bootstrap method involves taking the original data set of heights, and sampling from it to form a new sample (called a 'resample' or 'bootstrap sample') that is also of size N.
- The bootstrap sample is taken from the original by using sampling with replacement (e.g. we might 'resample' 5 times from [1,2,3,4,5] and get [2,5,4,4,1]).

Bootstrap: Intuitive explanation . . . II

- Assuming N is sufficiently large, there is virtually zero probability that it will be identical to the original “real" sample.
- This resampling process is repeated a large number of times (typically 1,000 to 10,000), and for each bootstrap sample, we compute its mean (each of which is called a "bootstrap estimate").
- We can now create a histogram of bootstrap means.
- This histogram provides an estimate of the distribution of the sample mean, from which we can answer questions about how much the mean varies across samples.
- The method described here for the mean can be applied to almost any other statistic or estimator, not just the mean.

Bootstrap: Illustrated



A sample is drawn from a population. From this sample, resamples are generated by drawing with replacement (orange). Data points that were drawn more than once (which happens for approx. 26.4% of data points) are shown in red and slightly offsetted. From the resamples, the statistic x is calculated and, therefore, a histogram can be calculated to estimate the distribution of x.

From: Wikipedia

Bootstrapping: Additional resources

- To learn more about bootstrap sampling, see Introduction to bootstrap sampling, which includes Python code.
- Other articles explaining bootstrapping along with Python code:
  - A few examples of bootstrapping
  - Ditch p-values. Use Bootstrap confidence intervals instead.
- The definitive guide to bootstrapping is the book by Efron and Tibshirani (1994).

Block bootstrap

- The block bootstrap is used when the data (or the errors in a model) are correlated, as is often the case for financial data.
  - In this case, simple resampling will fail, because it is not able to replicate the correlation in the data.
  - The block bootstrap tries to replicate the correlation by resampling inside blocks of data. For additional details, see this article.
  - The block bootstrap has been used mainly with data correlated in time (i.e., time series) but can also be used with data correlated in space, or among groups (so-called cluster data).

Block bootstrap: Motivation

- **The problem with time-series data**
  - The classical bootstrap doesn't apply to time series because time series data has an inherent order to it that i.i.d. methods don't account for.
  - Each data point in a time series depends on previous data points.
    1. So we need a way to incorporate dependency and preserve the order.
    2. And we also want our samples to be random.
  - To achieve both objectives, we divide the data into a series of consecutive blocks, each containing a few observations, in order, and use a bootstrap on the blocks.

Block bootstrap: Details

- Suppose we have the series $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$.
- We first need to determine how many observations to put in each group. Let's say we will use two data points in each block.
- Then our blocked data would look like

$$\overbrace{X_1, X_2}^{block1}, \quad \overbrace{X_3, X_4}^{block2}, \quad \overbrace{X_5, X_6}^{block3}, \quad \overbrace{X_7, X_8}^{block4}, \quad \overbrace{X_9, X_{10}}^{block5}.$$

- Then, we take a random sample of the blocks with replacement.
- A bootstrapped series of blocks might be

block 3, block 1, block 5, block 2, block 5

- This has the effect of giving us a new series with the same (short term) dependence structure.

Moving block bootstrap

- An extension of the block bootstrap is the moving block bootstrap (MBB) where the series is split into overlapping blocks.
  The blocks would therefore look like this:

$$\begin{aligned}
\text{block 1:} &\quad X_1, X_2 \\
\text{block 2:} &\quad X_2, X_3 \\
\text{block 3:} &\quad X_3, X_4 \\
\text{block 4:} &\quad X_4, X_5 \\
\vdots \quad &\quad \vdots \quad\; \vdots \\
\text{block 9:} &\quad X_9, X_{10}.
\end{aligned}$$

- Once the blocks are defined, we can then take a bootstrap sample of the blocks as before.

The stationary bootstrap

- The stationary bootstrap is another variant of the block bootstrap with a different sort of randomness.
- The stationary bootstrap uses a random block length.
- An additional source of randomness comes from selecting the starting points for the blocks at random.

Circular block bootstrap

- Circular block bootstrapping treats data as circular, enabling blocks to wrap around from the end of the series back to the beginning.
- Thus, points on the edges have the same probability of being sampled as points in the middle.
  - For the paper that developed the circular block bootstrap, see Politis and Romano (1994).
  - For a nice explanation with pictures, see this article.
  - For Python code see this link.
- Ledoit and Wolf (2008) apply the circular block bootstrap to study the difference in Sharpe ratios, which is what we wish to study.

Bootstrap p-values: Ledoit and Wolf (2008)

- Ledoit and Wolf (2008) construct a bootstrap confidence interval for the difference in Sharpe ratios, which does not require particular assumptions regarding the return distributions.
  - Link to the paper.
  - Link to code from Michael Wolf's page (in Matlab and R).
  - Python code for the bootstrap method in Ledoit and Wolf (2008) is available at RobustSharpeRatioHAC by Michael Mark.

Bootstrap p-values for ΔSR

- **Hypotheses:**

$$H_0 : SR_n = SR_m \quad \text{vs.} \quad H_1 : SR_n \neq SR_m.$$

- **Define**

$$\Delta SR = SR_n - SR_m, \qquad \text{with estimated values:} \quad \widehat{\Delta SR} = \widehat{SR}_n - \widehat{SR}_m.$$

- **Test statistic:** A robust test statistic is a studentized statistic:

$$\mathcal{S} = \frac{\widehat{\Delta SR}}{\widehat{\mathrm{SE}}(\widehat{\Delta SR})},$$

where $\widehat{\mathrm{SE}}(\widehat{\Delta SR})$ is a heteroskedasticity/autocorrelation consistent (HAC) estimate or a bootstrap standard error.

Resampling design and p-value

- **Block bootstrap (time-series robust):** Resample pairs of the two returns in blocks of data to preserve serial dependence. Common choices:
  - fixed-length blocks,
  - stationary bootstrap with random block lengths, or
  - circular blocks (this is what I recommend).
- Compute bootstrap replicates $\mathcal{S}_b^*$ of the statistic under a resampling scheme that respects the null or, more practically, for SR differences, under the empirical distribution.
- Then, the two-sided p-value is

$$\hat{p} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(|\mathcal{S}_b^*| \geq |\mathcal{S}_{\text{obs}}|\right).$$

Bootstrapping: Practical recipe

1. Compute $\widehat{SR}_A, \widehat{SR}_B$ and $\mathcal{S}_{\text{obs}} = \widehat{\Delta SR}/\widehat{\mathrm{SE}}$.
2. Choose block length ℓ (rule-of-thumb: $\ell \propto T^{1/3}$) or use stationary bootstrap with mean length ℓ.
3. For $b = 1, \dots, B$:
   - 3.1 Draw a bootstrap time series of pairs $\{(R_{t,b}^{n*}, R_{t,b}^{m*})\}_{t=1}^{T}$ by concatenating random blocks from the original series.
   - 3.2 Compute $\widehat{SR}_n^{*(b)}$, $\widehat{SR}_m^{*(b)}$, $\widehat{\Delta SR}^{*(b)}$.
   - 3.3 Studentize: $\mathcal{S}_b^* = \widehat{\Delta SR}^{*(b)}/\widehat{\mathrm{SE}}^{*(b)}$, where $\widehat{\mathrm{SE}}^{*(b)}$ is computed on the bootstrap sample in the same way as in the original sample.
4. Estimated p-value: $\hat{p} = \frac{1}{B}\sum_b \mathbb{I}(|\mathcal{S}_b^*| \geq |\mathcal{S}_{\text{obs}}|)$.

End of focus

Road map

Sample Python code

- The next few slides contain sample Python code.
- Studying this code is optional.
- This code is provided without any guarantees.
- Feel free to play with the code.
- If you like fun and adventure,
  - please reproduce Tables 1 and 2 of Ledoit and Wolf (2008)
  - using your code or code written by ChatGPT or someone else.

Python code for bootstrapping the mean and percentile CI

Code for bootstrapping the mean and a percentile CI

```python
import numpy as np
from numpy.random import default_rng   # rng = random-number-generator

rng = default_rng(42)
n = 200
data = rng.normal(loc=0.0, scale=1.0, size=n)

B = 5000
boot_means = np.empty(B)
for b in range(B):
    sample = rng.choice(data, size=n, replace=True)
    boot_means[b] = sample.mean()

se_mean = boot_means.std(ddof=1)
ci_mean = np.quantile(boot_means, [0.025, 0.975])

print(data.mean(), se_mean, ci_mean)
```

Python code for bootstrapping a quantile

Code for bootstrapping a quantile

```
alpha = 0.05
boot_q = np.empty(B)
for b in range(B):
    sample = rng.choice(data, size=n, replace=True)
    boot_q[b] = np.quantile(sample, alpha)

se_q = boot_q.std(ddof=1)
ci_q = np.quantile(boot_q, [0.025, 0.975])

print(np.quantile(data, alpha), se_q, ci_q)
```

Python code for BCa interval . . . I

BCa interval

```
from math import isfinite
from scipy.stats import norm

def bca_interval(x, stat_fn, conf=0.95, B=5000, rng=None):
    if rng is None:
        rng = default_rng()
    x = np.asarray(x)
    n = x.size
    t0 = stat_fn(x)

    # 1) Bootstrap distribution
    boots = np.empty(B)
    for b in range(B):
        boots[b] = stat_fn(rng.choice(x, size=n, replace=True))

    # 2) Bias-correction z0
    prop = np.mean(boots < t0)  # strictly less than t0
    prop = min(max(prop, 1e-12), 1-1e-12)
    z0 = norm.ppf(prop)

    # 3) Jackknife for acceleration a       ... contd. on next page
```

Python code for BCa interval . . . II

```
    jack = np.empty(n)
    for i in range(n):
        jack[i] = stat_fn(np.delete(x, i))
    jack_mean = jack.mean()
    num = np.sum((jack_mean - jack)**3)
    den = np.sum((jack_mean - jack)**2)**1.5 + 1e-12
    a = (1/6) * (num / den)

    # 4) Adjusted alpha levels
    alpha_lo, alpha_hi = (1 - conf)/2, (1 + conf)/2
    z_lo, z_hi = norm.ppf(alpha_lo), norm.ppf(alpha_hi)
    adj_lo = norm.cdf(z0 + (z_lo - z0)/(1 - a*(z_lo - z0)))
    adj_hi = norm.cdf(z0 + (z_hi - z0)/(1 - a*(z_hi - z0)))

    # 5) Quantiles of bootstrap distribution
    lo, hi = np.quantile(boots, [adj_lo, adj_hi])
    return lo, hi

# Example: BCa for the mean
bca_interval_mean = bca_interval(data, np.mean, conf=0.95, B=5000, rng=
     rng)
print(bca_interval_mean)
```

Code for Sharpe ratio and IID and block bootstrap . . . I

Code for Sharpe ratio, and for IID and block bootstrap

```python
def sharpe(x):
    x = np.asarray(x)
    s = x.std(ddof=1)
    return np.nan if s == 0 else x.mean()/s

def iid_pairs_bootstrap(a, b, rng=None):
    if rng is None:
        rng = default_rng()
    n = len(a)
    idx = rng.integers(0, n, size=n)
    return a[idx], b[idx]

def block_bootstrap_pairs(a, b, ell, rng=None):
    """
    Moving-block bootstrap for paired series a,b (same length).
    Concatenate random starting points of length ell, then truncate to n
     .
    """
    if rng is None:
        rng = default_rng()
```

Code for Sharpe ratio and IID and block bootstrap . . . II

```
    n = len(a)
    num_blocks = int(np.ceil(n/ell))
    starts = rng.integers(0, n-ell+1, size=num_blocks)
    idx = np.concatenate([np.arange(s, s+ell) for s in starts])[:n]
    return a[idx], b[idx]
```

Python code for HAC-style SE for ΔSR . . . I

Code for HAC-style SE for ΔSR

```
def _newey_west_variance_of_mean(series, max_lag=None):
    """
    NW variance for the sample mean of a (possibly) dependent series.
    Returns var(mean(series)).
    """
    x = np.asarray(series)
    n = x.size
    x = x - x.mean()
    if max_lag is None:
        # Same default spirit as Mathematica code: ~ 4*(n/100)^(2/9)
        max_lag = int(np.floor(4*(n/100)**(2/9)))
    # gamma_0
    gamma0 = np.dot(x, x)/n
    # auto-covariances (population-scale)
    def cov_k(k):
        return np.dot(x[:-k], x[k:])/(n - k)
    w = lambda k: 1 - k/(max_lag + 1.0)
    s = gamma0 + 2.0*sum(w(k)*cov_k(k) for k in range(1, max_lag+1))
    return s / n   #                ... continued on the next page
```

Python code for HAC-style SE for ΔSR . . . II

```
def delta_sr_se_hac(a, b, max_lag=None):
    """
    HAC (NW) SE for Delta SR using an influence-function-style proxy.
    """
    a = np.asarray(a); b = np.asarray(b)
    muA, muB = a.mean(), b.mean()
    sA, sB = a.std(ddof=1), b.std(ddof=1)
    if sA == 0 or sB == 0:
        return np.nan
    # Influence-proxy (same form as Mathematica comments)
    zA = (a - muA)/sA
    zB = (b - muB)/sB
    d = (a - muA)/sA - (b - muB)/sB \
        - (muA/(2*sA**3))*((a - muA)**2 - sA**2) \
        + (muB/(2*sB**3))*((b - muB)**2 - sB**2)

    var_mean = _newey_west_variance_of_mean(d, max_lag=max_lag)
    se = np.sqrt(var_mean)
    return se
#                                ... continued on the next page
```

Python code for HAC-style SE for ΔSR ...III

```
def delta_sr_test_statistic(a, b, se_mode="HAC"):
    dSR = sharpe(a) - sharpe(b)
    if se_mode == "HAC":
        se = delta_sr_se_hac(a, b)
    else:
        se = delta_sr_se_hac(a, b)  # placeholder, same as Mathematica
     note
    T = np.nan if (se is None or not np.isfinite(se) or se == 0) else
     dSR/se
    return dSR, se, T
```

Python code for block bootstrap p-value . . . I

Code for block bootstrap p-value

```
def delta_sr_pvalue_block(a, b, ell, B=5000, se_mode="HAC", rng=None):
    if rng is None:
        rng = default_rng()
    dObs, seObs, TObs = delta_sr_test_statistic(a, b, se_mode)
    TBoot = []
    for _ in range(B):
        rA, rB = block_bootstrap_pairs(a, b, ell, rng=rng)
        seB = delta_sr_se_hac(rA, rB)
        dB = sharpe(rA) - sharpe(rB)
        if np.isfinite(seB) and seB != 0:
            TBoot.append(dB/seB)
    TBoot = np.asarray(TBoot)
    if TBoot.size == 0 or not np.isfinite(TObs):
        return np.nan
    return np.mean(np.abs(TBoot) >= np.abs(TObs))
#                               ... continued on the next page
```

Python code for block bootstrap p-value ...II

```
# Example with simple AR(1)-like dependence
rng = default_rng(23)
n = 2000
phiA = 0.2
phiB = 0.1
epsA = rng.normal(0, 0.01, size=n)
epsB = rng.normal(0, 0.01, size=n)

retA = np.zeros(n)
retB = np.zeros(n)
for t in range(1, n):
    retA[t] = phiA*retA[t-1] + epsA[t]
    retB[t] = phiB*retB[t-1] + epsB[t]

ell = max(5, round(n**(1/3)))
pval_block = delta_sr_pvalue_block(retA, retB, ell, B=2000, se_mode="HAC
    ", rng=rng)
print("Block p-value:", pval_block)
```

Python code for stationary bootstrap ... I

Code for stationary bootstrap

```
def stationary_bootstrap_pairs(a, b, mean_block_len, rng=None):
    """
    Stationary bootstrap for paired series (Politis & Romano).
    mean_block_len -> geometric restart prob p = 1/mean_block_len.
    """
    if rng is None:
        rng = default_rng()
    a = np.asarray(a); b = np.asarray(b)
    n = len(a)
    p = 1.0/float(mean_block_len)
    idx = np.empty(n, dtype=int)
    idx[0] = rng.integers(0, n)
    for t in range(1, n):
        if rng.random() < p:
            idx[t] = rng.integers(0, n)    # start new block
        else:
            idx[t] = (idx[t-1] + 1) % n    # continue (circular)
    return a[idx], b[idx]
#                               ... continued on the next page
```

Python code for stationary bootstrap . . . II

```
# Example usage (drop-in replacement inside the bootstrap loop)
def delta_sr_pvalue_stationary(a, b, mean_block_len, B=5000, se_mode="
    HAC", rng=None):
    if rng is None:
        rng = default_rng()
    dObs, seObs, TObs = delta_sr_test_statistic(a, b, se_mode)
    TBoot = []
    for _ in range(B):
        rA, rB = stationary_bootstrap_pairs(a, b, mean_block_len, rng=
    rng)
        seB = delta_sr_se_hac(rA, rB)
        dB = sharpe(rA) - sharpe(rB)
        if np.isfinite(seB) and seB != 0:
            TBoot.append(dB/seB)
    TBoot = np.asarray(TBoot)
    if TBoot.size == 0 or not np.isfinite(TObs):
        return np.nan
    return np.mean(np.abs(TBoot) >= np.abs(TObs))
```

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will study classical mean-variance portfolios that *ignore* estimation error.

To do for next class

- Readings
  - Please read the section on "Working with Data" by Moreira (2021, Chapter 5) available online from this link.
- Assignment
  - You should start working on the first assignment.

Road map

Bibliography . . . I

Bailey, D. H., and M. Lopez de Prado. 2012. The Sharpe ratio efficient frontier. *Journal of Risk* 15 (2): 13. (Cited on page 187).

Balduzzi, P., and A. W. Lynch. 1999. Transaction costs and predictability: Some utility cost calculations. *Journal of Financial Economics* 52 (1): 47–78. (Cited on page 160).

Bhardwaj, R. K., and L. D. Brooks. 1992. The january anomaly: effects of low share price, transaction costs, and bid-ask bias. *Journal of Finance* 47 (2): 553–575. (Cited on page 160).

Campbell, J. Y., and L. M. Viceira. 2002. *Strategic asset allocation: Portfolio choice for long-term investors.* New York: Oxford University Press. (Cited on page 115).

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, no. 1 (March): 57–82. (Cited on pages 131, 469, 966).

Connor, G., L. Goldberg, and R. A. Korajczyk. 2010. *Portfolio risk analysis.* Princeton, New Jersey: Princeton University Press. (Cited on page 133).

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2025. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on page 134).

Bibliography ... II

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. 2020. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Review of Financial Studies* 33, no. 5 (April): 2180–2222. (Cited on pages 162, 733, 852, 885, 957, 1029).

DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. (Cited on pages 162, 734, 886, 906, 953, 955, 1004, 1030, 1154).

DeMiguel, V., and R. Uppal. 2005. Portfolio investment with the exact tax basis via nonlinear programming. *Management Science* 51, no. 2 (February): 277–290. (Cited on pages 163, 393).

Dybvig, P., and H. K. Koo. 1996. Investment with taxes. Working Paper, Washington University. (Cited on page 163).

Efron, B., and R. J. Tibshirani. 1994. *An introduction to the bootstrap.* CRC press. (Cited on page 219).

Fama, E. F., and K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

Bibliography . . . III

Fama, E. F., and K. R. French. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

French, K. R. 2008. Presidential address: The cost of active investing. *Journal of Finance* 63, no. 4 (August): 1537–1573. (Cited on page 160).

Hou, K., H. Mo, C. Xue, and L. Zhang. 2019. Which factors? *Review of Finance* 23 (1): 1–35. (Cited on page 132).

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705. (Cited on pages 81, 132, 775, 945, 966).

———. 2017a. A comparison of new factor models. *Fisher College of Business Working Paper,* nos. 2015-03: 05. (Cited on page 132).

———. 2017b. Replicating anomalies. *The Review of Financial Studies.* (Cited on pages 132, 775).

Jobson, J. D., and B. M. Korkie. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance:* 889–908. (Cited on pages 187, 214, 447).

Kozak, S., S. Nagel, and S. Santosh. 2018. Interpreting factor models. *Journal of Finance* 73 (3): 1183–1223. (Cited on pages 133, 733).

Bibliography . . . IV

Kozak, S., S. Nagel, and S. Santosh. 2020. Shrinking the cross-section. *Journal of Financial Economics* 135 (2): 271–292. (Cited on pages 133, 733).

Ledoit, O., and M. Wolf. 2008. Robust performance hypothesis resting with the Sharpe ratio. *Journal of Empirical Finance* 15:850–859. (Cited on pages 225, 226, 232, 1121).

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka. 1999. A new estimate of transaction costs. *Review of Financial Studies* 12 (5): 1113–1141. (Cited on page 160).

Lo, A. W. 2002. The statistics of Sharpe ratios. *Financial Analysts Journal* 58 (4): 36–52. (Cited on pages 143, 187).

Memmel, C. 2003. Performance hypothesis testing with the Sharpe ratio. Available at SSRN 412588. (Cited on pages 187, 214).

Merton, R. C. 1971. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory* 3:373–413. (Cited on page 115).

———. 1990. *Continuous time finance.* Cambridge MA: Basil Blackwell. (Cited on page 115).

Moreira, A. 2021. *Quantitative investing.* Available online from this link. (Cited on pages 248, 399, 559).

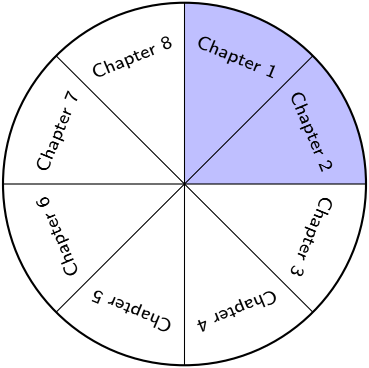
Opdyke, J. D. 2007. Comparing Sharpe ratios: So where are the p-values? *Journal of Asset Management* 8 (5): 308–336. (Cited on page 187).

Bibliography . . . V

Politis, D. N., and J. P. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89 (428): 1303–1313. (Cited on page 225).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).

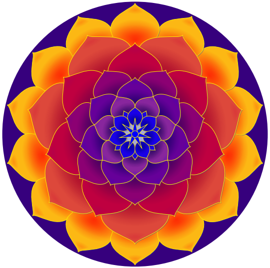
Stoll, H. R., and R. E. Whaley. 1983. Transaction costs and the small firm effect. *Journal of Financial Economics* 12 (1): 57–79. (Cited on page 160).



End of Chapter 2

Blank page 1

Quantitative Portfolio Management



Chapter 3:
Mean-Variance Portfolios that *Ignore* Estimation Error

Raman Uppal

2025

The big picture: Plan for the entire course

Table of contents

Road map

What do we want to do in Chapter 3

In this chapter, we study classical mean-variance portfolios that *ignore* estimation error.

We first understand how to derive the efficient frontier in the absence of a risk-free asset, and then in the presence of a risk-free asset.

Road map

Universe selection . . . I

- Before we study the selection of assets, we need to identify the universe from which we will be selecting these assets.
- Different investors may wish to choose assets from different universes.
- Investors may choose their asset universe based on regions:
  - For a French investor, the universe of investors may be all the assets traded in France, or the EU, or in developed markets.
  - For regulatory reasons, an Indian investor's universe may be the set of assets available in India.
  - Other investors may wish to restrict the asset universe to emerging markets.

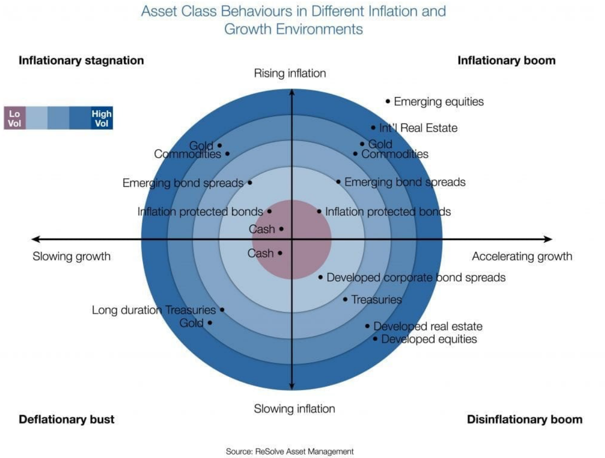
Universe selection ... II

- Investors may have a broad or narrow asset universe, depending on their investment goals and preferences.
  - Some investors may wish to invest only in large-cap companies.
  - Other investors may wish to invest only in a set of riskier companies.
  - Because of the large amount of capital they have, the Norwegian sovereign wealth fund, Norges Bank Investment Management,
    - has holdings in over 11,000 companies,
    - with a holding of 1.5% of all shares in the world's companies,
    - so their universe consists of traded assets all over the world, and
    - includes equities, fixed-income, real-estate, and infrastructure.

Universe selection . . . III

- Investors may restrict their universe to a particular asset class.
  - Some investors may wish to invest only in equities.
  - Other investors may wish to invest only in fixed-income assets.
- Investors' choice of universe may depend on risk-return preferences.
  - Investors with a small appetite for risk may specialize in treasuries.
  - Investors seeking high returns may choose high credit-risk assets.
- The universe could also depend on liquidity.
  - Some investors may want liquid assets.
  - Other investors with a longer horizon (e.g., insurance companies) may be willing to hold illiquid assets.
- The asset class an investor chooses could depend on the investor's views about inflation and growth (see next page).

Asset-class behavior depending on inflation and growth



From: ReSolve Asset Management

List of assets in which one could invest

- Equities
  - US stocks
  - European stocks
  - Asia-Pacific stocks
  - Emerging-market stocks
- Fixed income
  - Short-term treasuries
  - Intermediate-term treasuries
  - Intermediate-term international government bonds
  - Long-term treasuries
  - Long-term TIPs (inflation-hedged bonds)
  - USD-denominated emerging market bonds
- Derivatives
  - Futures
  - Options
- Commodities
  - Gold and other precious metals
  - Other commodities (oil, corn, etc.)
- Real estate
  - Global REITs
- Cryptocurrencies

Asset universe for our course

- For this course, we will choose the asset universe to be equities.
  - Much of the academic literature is about US equities, so many of the results I present will be based on evidence for US equities.
  - But almost everything in the academic literature applies also to equities in France, Germany, the UK, and other developed markets.
  - If you are interested, you may wish to take the evidence for US equities that I present in class and test it on data for French, German, European, Chinese, or Indian equities.
  - You may also wish to take what we learn about equities, and apply it to other asset classes, or apply it across asset classes.

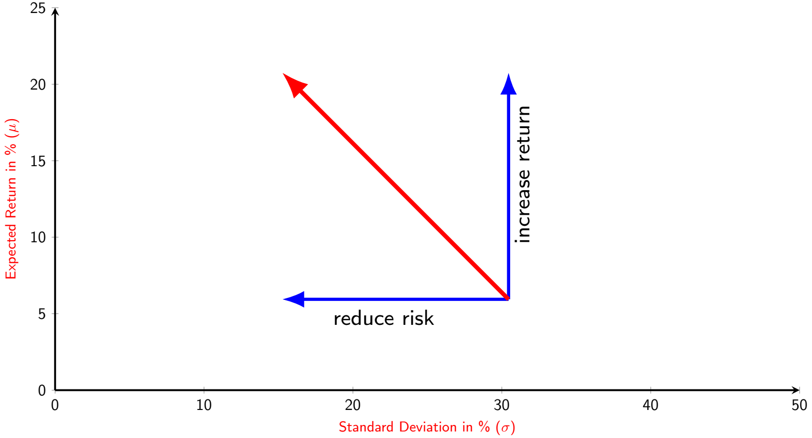
Selecting stocks from the universe of equities

- Having agreed that our asset universe consists of N equities, we now study the modern theory of portfolio choice.

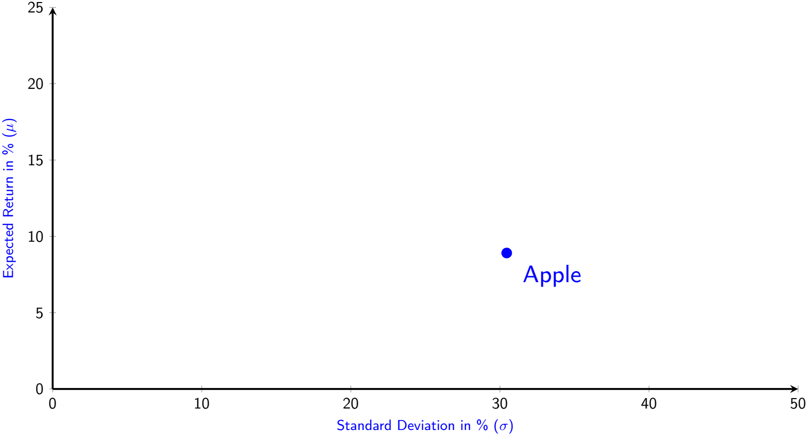
Road map

Understanding portfolio choice through pictures

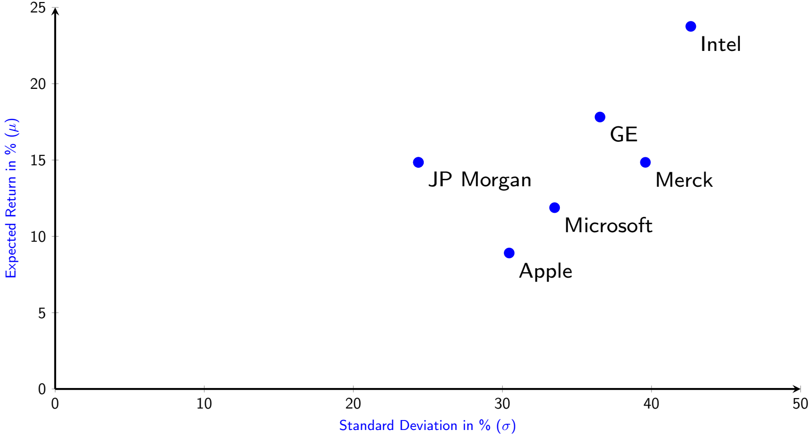
- Modern portfolio theory says that:
  "investors wish to balance their risk and return"



Do not evaluate assets in isolation



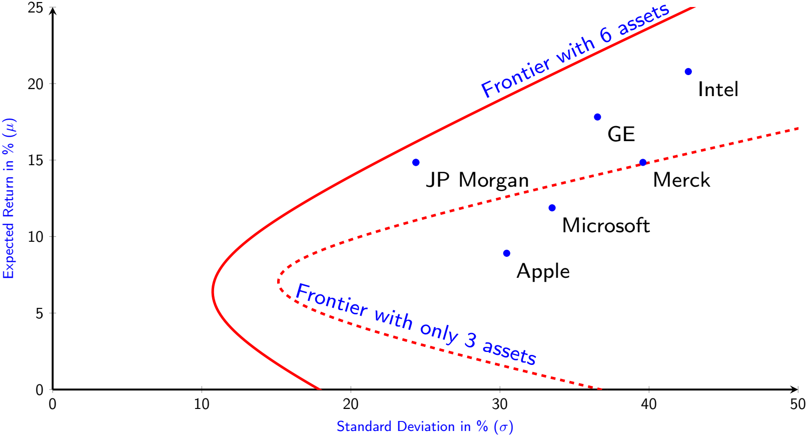
Should consider risk and return for several stocks together



Markowitz (Nobel Prize, 1990)

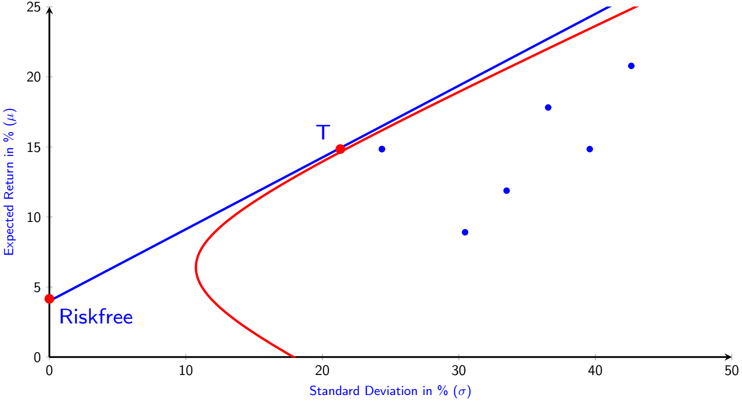
- Look at not just expected return and risk of individual stocks
- But, consider also the "interaction" between stocks: correlation
- Reap the benefits of optimal diversification

Markowitz mean-variance efficient frontier

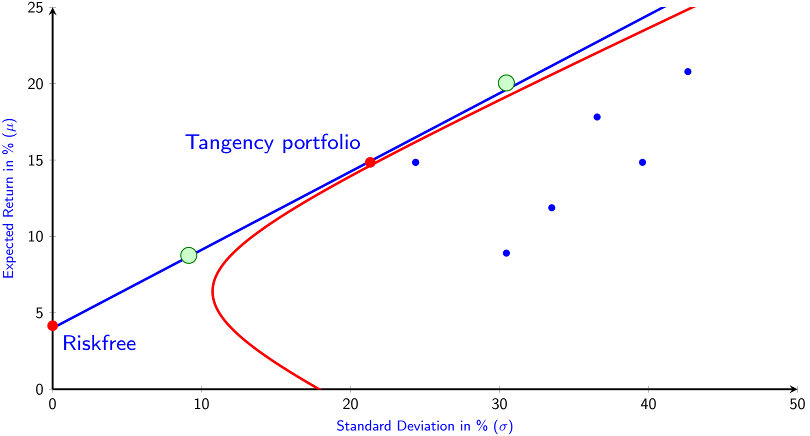


James Tobin (Nobel Prize, 1981)

- Simplify portfolio problem by introducing a risk-free asset.
- If a risk-free asset is available, then frontier portfolios are combinations of
  1. the risk-free asset, and
  2. a portfolio of only-risky assets.



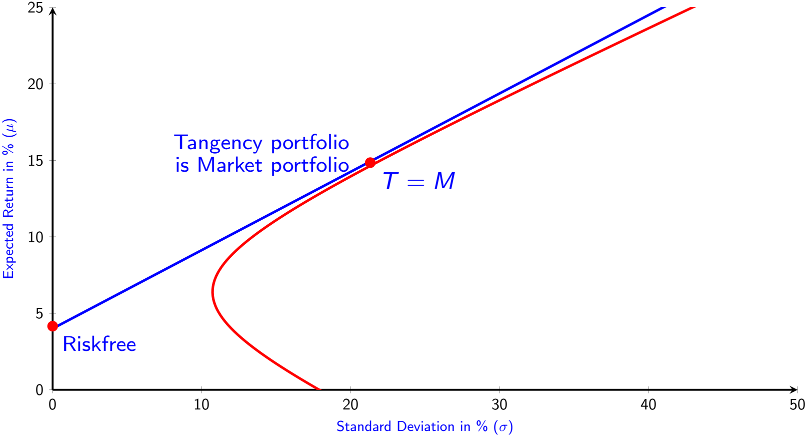
Two-fund separation



- Thus, *every* investor's portfolio consists of only two funds:
  - the risk-free asset
  - the tangency portfolio T, consisting of only-risky assets.

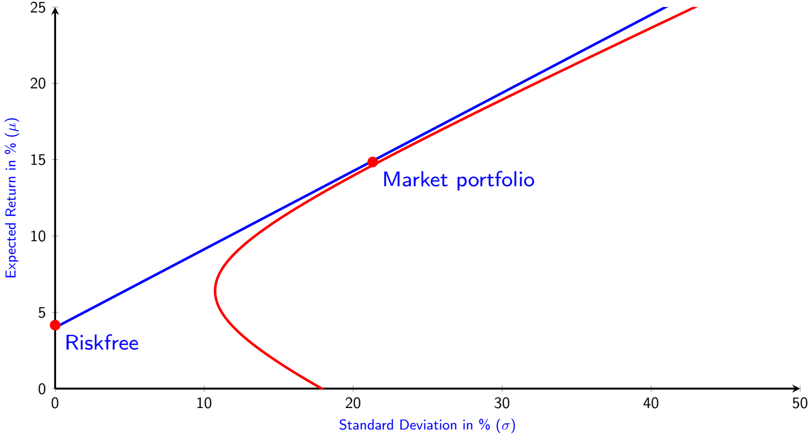
William Sharpe (Nobel Prize, 1990)

- Identified the tangency portfolio as the market portfolio: $T = M$
- This led to the Capital Asset Pricing Model (CAPM), which says: only systematic risk matters



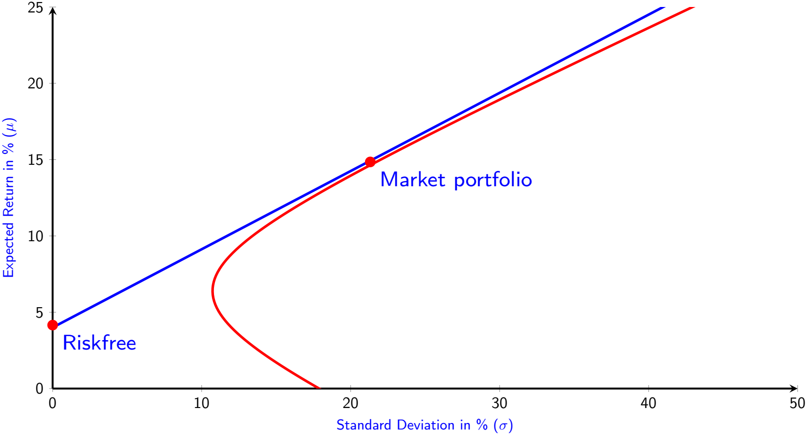
Paul Samuelson (Nobel Prize, 1970)

- What is the optimal portfolio in a multiperiod setting?
  - If the risk-free rate does not change; and,
  - If the slope of tangency line (Sharpe ratio) does not change;
  - Then, the optimal portfolio does not change.



Robert Merton (Nobel Prize, 1990)

- What is the optimal portfolio in a multiperiod (dynamic) setting?
  - If the risk-free rate and slope of the tangency line can change
  - Then, the optimal portfolio is equal to:
    - Markowitz portfolio + intertemporal hedges against these changes



End of discussion of models of portfolio choice
using only pictures.

Road map

Portfolio Diversification

The current finance literature believes that:

- Risks in individual asset returns have two components:
  - Systematic risks—common to many assets; so, non-diversifiable
  - Unsystematic risks—specific to individual assets; so, diversifiable.
- Assume there is no reward for bearing unsystematic risk.
  We will revisit this assumption in the last class.
- Therefore, it is optimal to hold a diversified portfolio that eliminates unsystematic risks.
- If holding a diversified portfolio, then only systematic risks matter.
- Thus, return on an asset compensates only for systematic risks.
  We will revisit this result in the last class.

Portfolio of two assets

We start with two assets, Asset 1 and Asset 2, whose returns, $\{R_1, R_2\}$, are characterized by their mean, variance and covariances.

- Mean returns:

Asset	1	2
Mean Return	$\mathbb{E}[R_1]$	$\mathbb{E}[R_2]$

- Variances and covariances (given by the covariance matrix):

	R_1	R_2
R_1	σ_1^2	σ_{12}
R_2	σ_{21}	σ_2^2

- Covariance of an asset with itself is its variance: $\sigma_{11} = \sigma_1^2$ and $\sigma_{22} = \sigma_2^2$.

Portfolio with two assets

- A portfolio of these two assets is characterized by the value invested in each asset.
- Let V_1 and V_2 be the dollar amount invested in asset 1 and 2, respectively. The total value of the portfolio is

$$V = V_1 + V_2.$$

- Consider a portfolio in which
  - $w_1 = V_1/V$ is the weight on asset 1
  - $w_2 = V_2/V$ is the weight on asset 2.
- Then, $w_1 + w_2 = 1$.
- In this chapter, we consider only portfolios with positive total-value: $V > 0$. When $V = 0$, the portfolio is called an "arbitrage portfolio."

Expected return and variance of a portfolio with two assets

- Expected portfolio return:

$$\mu_p = \mathbb{E}[R_p] = w_1\mathbb{E}[R_1] + w_2\mathbb{E}[R_2].$$

- The variance of the portfolio return:

$$\sigma_p^2 = \mathbb{V}\left[R_p\right] = \mathbb{E}\left[\left(R_p - \mathbb{E}[R_p]\right)^2\right] = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12}.$$

- Note: portfolio variance is the sum of all the entries of the table:

	$w_1 R_1$	$w_2 R_2$
$w_1 R_1$	$w_1^2\sigma_1^2$	$w_1 w_2 \sigma_{12}$
$w_2 R_2$	$w_1 w_2 \sigma_{12}$	$w_2^2\sigma_2^2$

Portfolio of Three Assets

Now consider the case with three assets, 1, 2 and 3. Returns on the three assets are $\{R_1, R_2, R_3\}$ and

- Mean returns:

Asset	1	2	3
Mean Return	$\mathbb{E}[R_1]$	$\mathbb{E}[R_2]$	$\mathbb{E}[R_3]$

- Variances and covariances:

	R_1	R_2	R_3
R_1	σ_1^2	σ_{12}	σ_{13}
R_2	σ_{21}	σ_2^2	σ_{23}
R_3	σ_{31}	σ_{32}	σ_3^2

Expected return of portfolio of three assets

- Consider the following portfolio:
  - w_1 is the weight on asset 1
  - w_2 is the weight on asset 2
  - w_3 is the weight on asset 3 and

$$w_1 + w_2 + w_3 = 1.$$

- The return of the portfolio:

$$R_p = w_1 R_1 + w_2 R_2 + w_3 R_3.$$

- Expected portfolio return:

$$\mu_p = \mathbb{E}[R_p] = w_1 \mathbb{E}[R_1] + w_2 \mathbb{E}[R_2] + w_3 \mathbb{E}[R_3].$$

Variance of return of portfolio of three assets

- The variance of the portfolio return:

$$\begin{aligned}\sigma_p^2 &= \mathbb{V}[R_p] = \mathbb{E}\left[(R_p - \mathbb{E}[R_p])^2\right] \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + w_3^2\sigma_3^2 && \ldots \text{diagonal terms in table below} \\ &\quad + 2w_1w_2\sigma_{12} + 2w_1w_3\sigma_{13} + 2w_2w_3\sigma_{23} && \ldots \text{off-diagonal terms in table.}\end{aligned}$$

- Again, the portfolio variance is the sum of all the table entries:

	w_1R_1	w_2R_2	w_3R_3
w_1R_1	$w_1^2\sigma_1^2$	$w_1w_2\sigma_{12}$	$w_1w_3\sigma_{13}$
w_2R_2	$w_1w_2\sigma_{12}$	$w_2^2\sigma_2^2$	$w_2w_3\sigma_{23}$
w_3R_3	$w_1w_3\sigma_{13}$	$w_2w_3\sigma_{23}$	$w_3^2\sigma_3^2$

Portfolio of N assets

- We now consider the general case of N assets.
- The returns on the N assets, $\{R_1, R_2, \ldots, R_N\}$, are characterized by their mean, variance and covariances.
- Mean returns:

Asset	1	2	$\cdots$	N
Mean Return	$\mathbb{E}[R_1]$	$\mathbb{E}[R_2]$	$\cdots$	$\mathbb{E}[R_N]$

- Variances and covariances:

	R_1	R_2	$\cdots$	R_N
R_1	σ_1^2	σ_{12}	$\cdots$	σ_{1N}
R_2	σ_{21}	σ_2^2	$\cdots$	σ_{2N}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
R_N	σ_{N1}	σ_{N2}	$\cdots$	σ_N^2

Portfolio with N assets . . . I

- Consider a portfolio of N assets in which the proportion of total value invested in asset n is denoted by w_n.

1. The weights sum to one:

$$\sum_{n=1}^{N} w_n = w^\top 1_N = 1.$$

2. The return on the portfolio is:

$$R_p = w_1 R_1 + w_2 R_2 + \cdots + w_N R_n = \sum_{n=1}^{N} w_n R_n = w^\top R.$$

Portfolio with N assets . . . II

3. The expected return on the portfolio is:

$$\mu_p = \mathbb{E}[R_p] = w_1\mathbb{E}[R_1] + w_2\mathbb{E}[R_2] + \cdots + w_N\mathbb{E}[R_N] = \sum_{n=1}^{N} w_n\mathbb{E}[R_n] = w^\top \mu.$$

4. The variance of portfolio return is:

$$\sigma_p^2 = \mathbb{V}[R_p] = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w, \quad \text{where} \quad \sigma_{nn} = \sigma_n^2.$$

5. The volatility (StD) of portfolio return is:

$$\sigma_p = \sqrt{\mathbb{V}[R_p]} = \sqrt{\sigma_p^2}.$$

Variance of portfolio return

- The variance of portfolio return can be computed by summing up all the entries to the following table:

	w_1R_1	w_2R_2	$\cdots$	w_NR_N
w_1R_1	$w_1^2\sigma_1^2$	$w_1w_2\sigma_{12}$	$\cdots$	$w_1w_N\sigma_{1N}$
w_2R_2	$w_2w_1\sigma_{21}$	$w_2^2\sigma_2^2$	$\cdots$	$w_2w_N\sigma_{2N}$
$\cdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
w_NR_N	$w_Nw_1\sigma_{N1}$	$w_Nw_2\sigma_{N2}$	$\cdots$	$w_N^2\sigma_N^2$

- The variance of a sum is not just the sum of variances!
  We also need to account for the covariances.

- In order to calculate the return variance of a portfolio, we need
  (a) portfolio weights
  (b) individual variances
  (c) all covariances.

Diversification with N risky assets . . . I

- From the previous slide:

	$w_1 R_1$	$\cdots$	$w_N R_N$
$w_1 R_1$	$w_1^2 \sigma_1^2$	$\cdots$	$w_1 w_N \sigma_{1N}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$w_N R_N$	$w_N w_1 \sigma_{N1}$	$\cdots$	$w_N^2 \sigma_N^2$

- Now consider an equally weighted portfolio of N assets.
  - Thus, the weight on each asset is $w_n = 1/N$.
- A typical variance term: $\left(\frac{1}{N}\right)^2 \sigma_{nn}$.
  - Total number of variance terms is: N.
- A typical covariance term is: $\left(\frac{1}{N}\right)^2 \sigma_{nm}$ $(n \neq m)$.
  - Total number of covariance terms: $N^2 - N$.
    (Total number of unique covariance terms: $(N^2 - N)/2$.)

Diversification with N risky assets ... II

- Add all the variance and covariance terms:

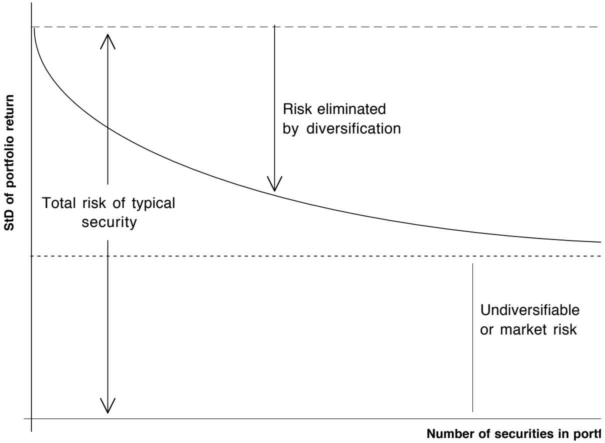
$$
\begin{aligned}
\sigma_p^2 &= \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} \\
&= \sum_{n=1}^{N}\left(\frac{1}{N}\right)^2 \sigma_{nn} + \sum_{n=1}^{N}\sum_{m\neq n}^{N}\left(\frac{1}{N}\right)^2 \sigma_{nm} \\
&= \left(\frac{1}{N}\right)\left(\frac{1}{N}\sum_{n=1}^{N}\sigma_n^2\right) + \left(\frac{N^2-N}{N^2}\right)\left(\frac{1}{N^2-N}\sum_{n=1}^{N}\sum_{m\neq n}^{N}\sigma_{nm}\right) \\
&= \left(\frac{1}{N}\right)(\text{average variance}) + \left(1-\frac{1}{N}\right)(\text{average covariance}) \\
\lim_{N\to\infty}\sigma_p^2 &= (0)\,(\text{average variance}) + (1-0)\,(\text{average covariance}) \\
&= \text{average covariance} \qquad \ldots\text{variance does not matter}
\end{aligned}
$$

Diversification with N risky assets . . . III

- Therefore, as N becomes very large:
  - Contribution of variance terms goes to zero.
  - Contribution of covariance terms goes to "average covariance".
- Conclusion: For a "well-diversified" portfolio:
  - Covariances among assets determine portfolio risk.
  - Variance of each asset contributes little to portfolio risk.
  - Therefore, when considering which asset to add to a portfolio, it is important to look at covariances (or correlations) and not variances.

Impact of Diversification On Portfolio Risk

_ _ _ $\bar{\rho}_{nm} = 1.0$ and __ $\bar{\rho}_{nm} = 0.5$



End of discussion about diversification.

Start of focus

Road map

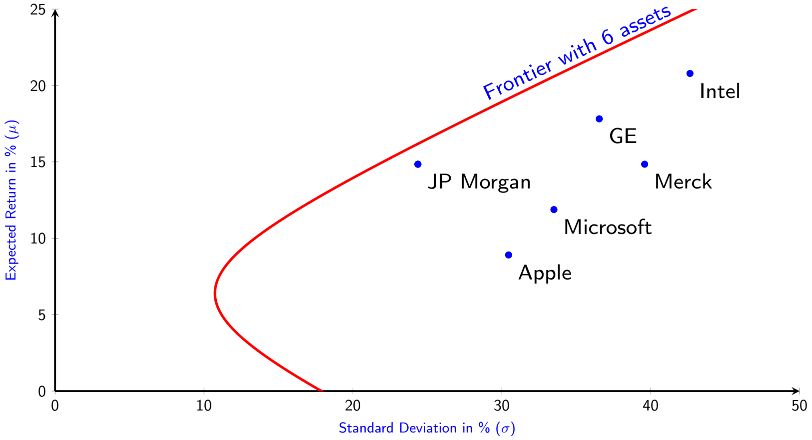
Theory of Markowitz mean-variance portfolios

- We start by studying the Markowitz (1952, 1959) model of portfolio selection.
- Optional reading (not required):
  - The lecture given by Markowitz when receiving the Nobel prize is available here.
  - We will follow the analysis in Merton (1972), which can be downloaded from this link.
  - The best textbook reference is Huang and Litzenberger (1988, ch. 2).

Mean-variance portfolio frontier without a risk-free asset

- How to choose a mean-variance efficient portfolio?
  That is, how to:
  - Minimize risk for a given expected return, or
  - Maximize expected return for a given risk?
- Definition: Given an expected return, the portfolio that minimizes risk (measured by StD) is a *mean-StD frontier portfolio*.
- Definition: The locus of all frontier portfolios in the mean-StD plane is called the portfolio frontier.
  - The upper part of the frontier gives the efficient frontier portfolios.

Portfolio frontier with six risky assets (no risk-free asset)



Notation reminder

- In order to facilitate our discussion, we introduce matrix notation.
- The unit vector with N elements is:

$$1_N = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

- The $(N \times N)$ identity matrix is:

$$1_{N\times N} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Notation for weights, return means, and covariances

- For a portfolio problem we can express portfolio weights, mean returns and covariances of asset returns in matrices:

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}, \quad \mu = \begin{bmatrix} \mathbb{E}[R_1] \\ \mathbb{E}[R_2] \\ \vdots \\ \mathbb{E}[R_N] \end{bmatrix}, \quad V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix},$$

  where μ and w are N-dimensional vectors and V is an $(N \times N)$ symmetric matrix.

- Note that the covariance matrix V is symmetric and is often denoted by Σ in academic papers, and so is often called the "Sigma" matrix.

Mean-variance quadratic optimization problem

- In order to obtain the frontier portfolios, we need to solve the following problem:

$$\begin{aligned}
\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad & \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w \\
\text{subject to:} \quad (1) \quad & \sum_{n=1}^{N} w_n = w^\top 1_N = 1; \\
(2) \quad & \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^\top \mu = \mu_{\text{targ}}.
\end{aligned}$$

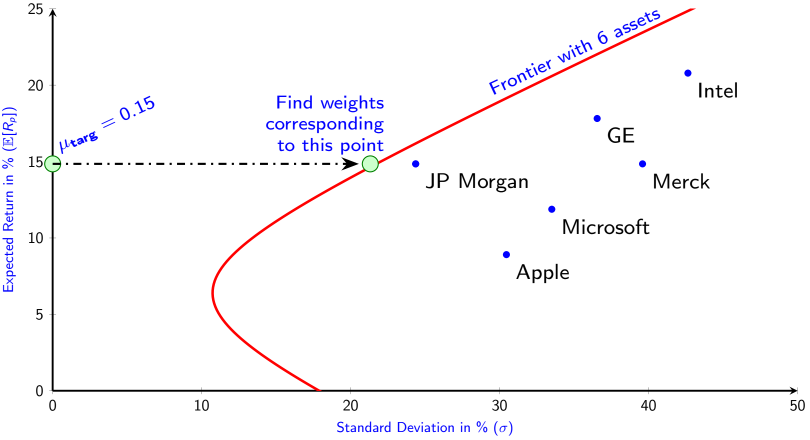
- In plain language, we need to:

  find the weights $\{w_1, \ldots, w_N\}$ that minimize the portfolio variance σ_p^2 while satisfying two constraints:

  1. the weights must add up to 1 and
  2. the portfolio p must have an expected target return μ_{targ} (e.g., $\mu_{\text{targ}} = 15\%$).

- This is a non-linear optimization problem that we will solve below.

Find weights corresponding to $\mu_{\text{targ}} = 0.15$



Lagrangian for mean-variance portfolio problem

- To solve this constrained optimization problem, we construct a Lagrangian function, $\mathcal{L}$:

$$\mathcal{L} = w^\top V w + \lambda_w \left(1 - w^\top 1_N\right) + \lambda_R \left(\mu_{\text{targ}} - w^\top \mu\right)$$

where λ_w and λ_R are two variables called Lagrange multipliers that correspond to the two constraints.

First-order condition for optimal portfolio weights

- We start by re-writing the Lagrangian on the previous page:

$$\mathcal{L} = w^\top V w + \lambda_w \left(1 - w^\top 1_N\right) + \lambda_R \left(\mu_{\text{targ}} - w^\top \mu\right)$$

- The optimality condition is given by the first-order condition

$$\frac{\partial \mathcal{L}}{\partial w} = 2Vw - \lambda_w 1_N - \lambda_R \mu = 0.$$

- Assume that V is not singular; thus, its inverse exists. Then

$$w = V^{-1}\left(\frac{\lambda_w}{2} 1_N + \frac{\lambda_R}{2}\mu\right) \quad \text{. . . multiply both sides by } \frac{V^{-1}}{2} \tag{15}$$

- To get an explicit solution for the optimal weights, we need to:
  1. Find λ_w and λ_R, and
  2. Substitute these values into the expression for w in (15).

First-order condition for the Lagrange multipliers

- We start by re-writing the Lagrangian:

$$\mathcal{L} = w^\top V w + \lambda_w \left(1 - w^\top 1_N\right) + \lambda_R \left(\mu_{\text{targ}} - w^\top \mu\right)$$

- Differentiating the Lagrangian with respect to λ_w gives us:

$$\frac{\partial \mathcal{L}}{\partial \lambda_w} = 0 \implies 1 = w^\top 1_N. \tag{16}$$

- Differentiating the Lagrangian with respect to λ_R gives us:

$$\frac{\partial \mathcal{L}}{\partial \lambda_R} = 0 \implies \mu_{\text{targ}} = w^\top \mu. \tag{17}$$

Going from three equations to two equations

- Substituting the expression for w from Equation (15), which is:

$$w = V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right), \tag{15}$$

into equations (16) and (17), which we rewrite below:

$$1 = w^\top 1_N \tag{16}$$

$$\mu_{\text{targ}} = w^\top \mu \tag{17}$$

- gives:

$$1 = w^\top 1_N = \left[V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right)\right]^\top 1_N = a_3\frac{\lambda_w}{2} + a_2\frac{\lambda_R}{2}, \tag{18}$$

$$\mu_{\text{targ}} = w^\top \mu = \left[V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right)\right]^\top \mu = a_2\frac{\lambda_w}{2} + a_1\frac{\lambda_R}{2}, \tag{19}$$

- where

$$a_1 = \mu^\top V^{-1}\mu, \qquad a_2 = \mu^\top V^{-1}1_N, \qquad a_3 = 1_N^\top V^{-1}1_N.$$

The system of two equations for λ_w and λ_R

- We start by re-writing from the previous page:

$$1 = a_3 \frac{\lambda_w}{2} + a_2 \frac{\lambda_R}{2}, \tag{18}$$

$$\mu_{\text{targ}} = a_2 \frac{\lambda_w}{2} + a_1 \frac{\lambda_R}{2}. \tag{19}$$

- The above system of two equations can be written in matrix notation as follows, where the two unknowns are λ_w and λ_R:

$$\begin{bmatrix} 1 \\ \mu_{\text{targ}} \end{bmatrix} = \begin{bmatrix} a_3 & a_2 \\ a_2 & a_1 \end{bmatrix} \cdot \begin{bmatrix} \frac{\lambda_w}{2} \\ \frac{\lambda_R}{2} \end{bmatrix} \tag{20}$$

Solving the two equations for λ_w and λ_R

- Solving the two equations for $\lambda_w/2$ and $\lambda_R/2$, we have

$$\frac{\lambda_w}{2} = \frac{1}{D}(a_1 - a_2\mu_{\text{targ}})$$

$$\frac{\lambda_R}{2} = \frac{1}{D}(-a_2 + a_3\mu_{\text{targ}}),$$

where we denote by D the determinant (det) of the 2×2 matrix:

$$D = \det\left(\begin{bmatrix} a_3 & a_2 \\ a_2 & a_1 \end{bmatrix}\right) = a_1 a_3 - a_2^2.$$

Solving equation for the optimal portfolio weights

- Substituting this solution into Equation (15) for w, reproduced below:

$$w = V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right) \tag{15}$$

- Gives the frontier portfolio weights for expected return μ_{targ}:

$$w = \frac{1}{D}\left(a_1 V^{-1}1_N - a_2 V^{-1}\mu\right) + \frac{1}{D}\left(a_3 V^{-1}\mu - a_2 V^{-1}1_N\right)\mu_{\text{targ}}.$$

Explicit solution for frontier portfolio weights

- We highlight this important result:
  - Frontier portfolio weights for expected return μ_{targ} are given by:

$$w = \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1} \mu\right) + \frac{1}{D}\left(a_3 V^{-1} \mu - a_2 V^{-1} 1_N\right) \mu_{\text{targ}}.$$

- Note that this explicit solution is possibly only when there are no other constraints on the portfolio weights. For example,
  - if we want all weights to be positive

$$w_n \geq 0, \quad \text{for all } n,$$

  - Or, if we want all weights to be less than some maximum weight

$$w_n \leq w_{\max}, \quad \text{for all } n,$$

- then we will not be able to get an explicit solution; instead, we will have to solve for the weights numerically (which we will do later).

Two-fund separation

- Re-writing the last equation on the previous slide:

$$w = \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1}\mu\right) + \frac{1}{D}\left(a_3 V^{-1}\mu - a_2 V^{-1} 1_N\right)\mu_{\text{targ}}.$$

- Define by w_0 and w_1 the frontier portfolios with expected return $\mu_{\text{targ}} = 0$ and $\mu_{\text{targ}} = 1$, respectively:

$$\begin{aligned} w_0 &= \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1}\mu\right) \\ w_1 &= \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1}\mu\right) + \frac{1}{D}\left(a_3 V^{-1}\mu - a_2 V^{-1} 1_N\right) \end{aligned}$$

- Then any frontier portfolio p with expected return μ_{targ} is equal to

$$\begin{aligned} w_p &= w_0 + (w_1 - w_0)\mu_{\text{targ}} \\ &= w_0(1 - \mu_{\text{targ}}) + w_1\mu_{\text{targ}} \end{aligned}$$

... i.e., $\boldsymbol{w_0}$ and $\boldsymbol{w_1}$ generate the entire frontier.

Global-minimum-variance portfolio

- Given that any frontier portfolio with expected target return μ_{targ} is:

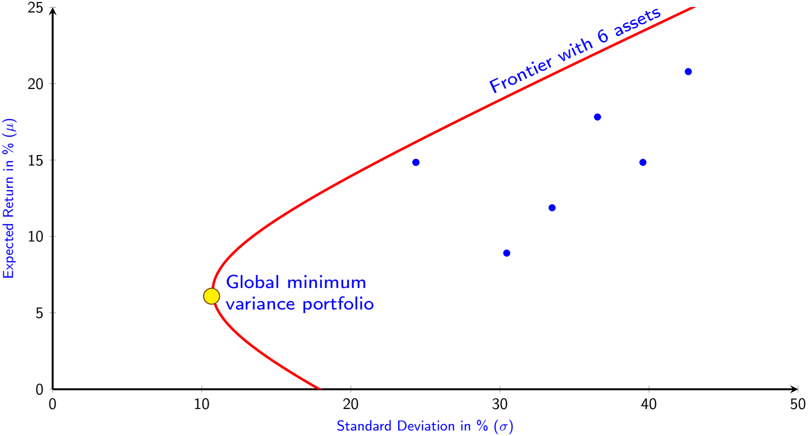
$$w_p = w_0(1 - \mu_{\text{targ}}) + w_1 \mu_{\text{targ}}, \quad \text{then}$$

- The variance of frontier portfolio p is

$$\begin{aligned}
\sigma_p^2 &= w_p^\top V w_p \\
&= \left(w_0(1 - \mu_{\text{targ}}) + w_1 \mu_{\text{targ}}\right)^\top V \left(w_0(1 - \mu_{\text{targ}}) + w_1 \mu_{\text{targ}}\right) \\
&= (1 - \mu_{\text{targ}})^2 w_0^\top V w_0 + 2\mu_{\text{targ}}(1 - \mu_{\text{targ}}) w_0^\top V w_1 + \mu_{\text{targ}}^2 w_1^\top V w_1 .
\end{aligned}$$

- Minimizing σ_p^2 gives the global-minimum-variance (GMV) portfolio.

Portfolio frontier with six risky assets (no risk-free asset)



An example with $N = 3$ assets . . . I

- Python code for this example is provided at the end of the example.
- The unit vector and mean returns on the $N = 3$ assets be:

$$1_N = \begin{bmatrix} 1. \\ 1. \\ 1. \end{bmatrix}; \qquad \mu = \begin{bmatrix} 0.20 \\ 0.30 \\ 0.40 \end{bmatrix}.$$

- And, let the covariance matrix of returns (which is symmetric) be:

$$V = \begin{bmatrix} 0.0625 & 0.0700 & 0.1050 \\ 0.0700 & 0.1225 & 0.0840 \\ 0.1050 & 0.0840 & 0.3600 \end{bmatrix}.$$

An example with $N = 3$ assets ... II

- Then, based on the data given to us, compute:

$$V^{-1} = \begin{bmatrix} 85.06 & -37.61 & -16.03 \\ -37.61 & 26.35 & 4.82 \\ -16.03 & 4.82 & 6.33 \end{bmatrix}.$$

- Then, using the formulas

$$a_1 = \mu^\top V^{-1}\mu, \quad a_2 = \mu^\top V^{-1}1_N, \quad a_3 = 1_N^\top V^{-1}1_N, \quad D = a_1 a_3 - a_2^2,$$

we get that

$$a_1 = 0.865, \qquad a_2 = 2.40, \qquad a_3 = 20.09, \qquad D = 11.63.$$

An example with $N = 3$ assets ... III

- Then, using the formulas

$$w_0 = \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1} \mu\right);$$

$$w_1 = \frac{1}{D}\left(a_1 V^{-1} 1_N - a_2 V^{-1} \mu\right) + \frac{1}{D}\left(a_3 V^{-1} \mu - a_2 V^{-1} 1_N\right),$$

we get that

$$w_0 = \begin{bmatrix} 2.48 \\ -0.96 \\ -0.52 \end{bmatrix}, \quad w_1 = \begin{bmatrix} -5.18 \\ 4.37 \\ 1.82 \end{bmatrix}, \quad w_1 - w_0 = \begin{bmatrix} -7.66 \\ 5.33 \\ 2.34 \end{bmatrix}.$$

An example with $N = 3$ assets ... IV

- Using the expression for optimal weights:

$$w = w_0 + \mu_{\text{targ}}(w_1 - w_0)$$

the weights for a portfolio with a target mean return of μ_{targ} is:

$$w = \begin{bmatrix} 2.48 \\ -0.96 \\ -0.52 \end{bmatrix} + \mu_{\text{targ}} \begin{bmatrix} -7.66 \\ 5.32 \\ 2.34 \end{bmatrix}, \tag{21}$$

- with a portfolio volatility of:

$$\sigma_p = \sqrt{w^\top V w}.$$

The optimal portfolio for $\mu_{\text{targ}} = 35\%$

- Suppose that you want $\mu_{\text{targ}} = 35\%$. Then, what should be your portfolio?
- Using our result in equation (21) on the previous slide,

$$w = \begin{bmatrix} 2.48 \\ -0.96 \\ -0.52 \end{bmatrix} + \begin{bmatrix} -7.66 \\ 5.32 \\ 2.34 \end{bmatrix} 0.35 = \begin{bmatrix} -0.20 \\ 0.90 \\ 0.30 \end{bmatrix}.$$

- The above solution says that if you have \$1, you want to
  - short-sell \$0.20 worth of asset 1
  - buy \$0.90 worth of asset 2
  - buy \$0.30 worth of asset 3,
- which will give a portfolio with $\mu_{\text{targ}} = 0.35$ and $\sigma_p = 0.38$.

Example: The optimal portfolio for $\mu_{\text{targ}} = 30\%$ . . . I

- Suppose that you want $\mu_{\text{targ}} = 30\%$.
- Then, what portfolio should you hold?

The optimal portfolio for $\mu_{\text{targ}} = 30\%$

- Again, using our result in equation (21),

$$w = \begin{bmatrix} 2.48 \\ -0.96 \\ -0.52 \end{bmatrix} + \begin{bmatrix} -7.66 \\ 5.32 \\ 2.34 \end{bmatrix} 0.30 = \begin{bmatrix} 0.18 \\ 0.64 \\ 0.18 \end{bmatrix}.$$

- which then leads to a portfolio with:

$$\mu_{\text{targ}} = 0.30 \quad \text{and} \quad \sigma_p = 0.3256.$$

Conclusion of example

- Note that you can get an expected rate of return of 30% also by investing only in Asset 2;
- But the StD of investing only in Asset 2 is $\sqrt{0.1225} = 0.35$.
- By diversifying, you get same expected return but with lower risk: $0.3256 < 0.35$.

How do you generate the entire portfolio frontier?

- By choosing different levels of μ_{targ}, and
- then finding the corresponding w and σ_p,
- one can identify the entire mean-volatility portfolio frontier.

Python code for the entire example . . . I

Python code for $N = 3$ example

```
# Import numpy
import numpy as np

# Define the vectors and matrices we will need

# Define a column vector of ones
ones = np.ones((3, 1))

# Column vector of mean returns
mu = np.array( [ [0.20], [0.30], [0.40] ] )

# Variance-covariance matrix
V = np.array([ [0.0625, 0.0700, 0.1050],
    [0.0700, 0.1225, 0.0840],
    [0.1050, 0.0840, 0.3600]  ])

# Define inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)
```

Python code for the entire example . . . II

Define the intermediate quantities we need

```
# Compute the intermediate quantities
a1 = mu.T @ Vinverse @ mu    # The ".T" is the transpose operator
a2 = mu.T @ Vinverse @ ones  # The "@" stands for matrix multiplication
a3 = ones.T @ Vinverse @ ones
D  = a1*a3 - a2**2

# Print output of the above to see that there are no errors
print("a1 = ", a1[0,0])  # The "[0,0]" selects number from 1 x 1 matrix
print("a2 = ", a2[0,0])
print("a3 = ", a3[0,0])
print("D  = ", D[0,0])
```

```
a1 =  0.8650305691897012
a2 =  2.3980022388702302
a3 =  20.09299922500646
D  =  11.630643818708927
```

Python code for the entire example . . . III

Compute w0 and w1

```
#Define w0 and w1
w0 = (a1*Vinverse@ones - a2*Vinverse@mu)/D
w1 = (a1*Vinverse@ones - a2*Vinverse@mu)/D + (a3*Vinverse@mu - a2*
     Vinverse@ones)/D
```

Define solution when no risk-free asset

```
# Define function of mu_targ for weights and portfolio volatility
def solution(mu_targ):
    weights = w0 + mu_targ*(w1 - w0)
    portfolioVolatility = np.sqrt(weights.T @ V @ weights)

    print("Optimal weights are:")
    print(weights)
    print(" ")
    print("Volatilitiy of optimal portfolio's return is:",
     portfolioVolatility[0,0])
    return
```

Python code for the entire example . . . IV

Solution for target mean of 0.35

```
solution(0.35)
```

```
Optimal weights are:
[[-0.20335637]
 [ 0.90671273]
 [ 0.29664363]]

Volatilitiy of optimal portfolio’s return is: 0.3764031006457244
```

Solution for target mean of 0.30

```
solution(0.30)
```

```
Optimal weights are:
[[0.17966436]
 [0.64067127]
 [0.17966436]]

Volatilitiy of optimal portfolio’s return is: 0.32580752437866894
```

Road map

Portfolio frontier with a risk-free asset

- When investors can also invest in a risk-free asset, then the optimal portfolio consists of the risk-free asset and risky assets.
- Observation: A portfolio of risk-free and risky assets can be viewed as a portfolio of two sub-portfolios:
  1. a sub-portfolio of *only* the risk-free asset, and
  2. a sub-portfolio of *only* risky assets, called the tangency portfolio.

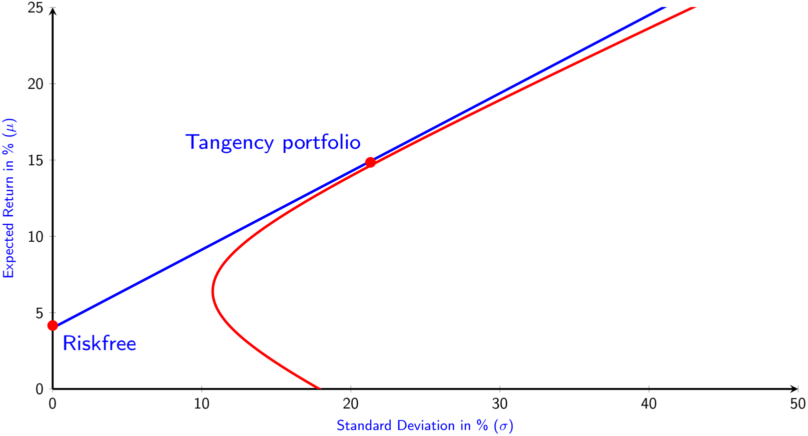
Sub-portfolio of only risky assets

- Consider a portfolio with
  - \$40 invested in the risk-free asset, and
  - \$30 each in two risky assets, IBM and Merck.
- Thus, the portfolio of *only* risky assets is one where 50% is invested in IBM and 50% in Merck.
- We can therefore view the original portfolio in two ways:

  View 1. The original portfolio has 40% invested in the risk-free asset, 30% invested in IBM, and 30% invested in Merck.

  View 2. The original portfolio has 40% invested in a subportfolio of only the risk-free asset, and 60% invested in the subportfolio of only risky assets (where this subportfolio has 50% in IBM and 50% in Merck).

Portfolio frontier with six risky assets and a risk-free asset



Frontier portfolios with a risk-free asset

- Consider a portfolio p with
  - a invested in a risky portfolio q, and
  - $1 - a$ invested in the risk-free asset.
- Then,

  The *mean* of the portfolio return is: $\mu_{\textsf{targ}} = (1 - a)R_f + a\mathbb{E}[R_q]$

  The variance of the portfolio return is: $\sigma_p^2 = a^2\sigma_q^2.$

  The StD of the portfolio return is: $\sigma_p = a\sigma_q.$

- Observation: In the presence of a risk-free asset, frontier portfolios are combinations of
  - the risk-free asset, and
  - the tangent portfolio (of risky assets).

Portfolio Optimization with A Risk-Free Asset . . . I

- Let
  - the return on the risk-free asset be R_f
  - the investment in the risk-free asset be w_f
  - the investment in the n-th risky asset be w_n, $n = 1, \ldots, N$.
- Let the total weight invested in risky assets be a, then

$$a = \sum_{n=1}^{N} w_n$$

and the weight invested in the risk-free asset is

$$w_f = 1 - a = 1 - \sum_{n=1}^{N} w_n.$$

Portfolio Optimization with A Risk-Free Asset . . . II

- With a risk-free asset, the optimal portfolio problem now becomes

$$\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad (1) \quad \sum_{n=1}^{N} w_n + w_f = w^\top 1_N + w_f = 1;$$

$$(2) \quad \sum_{n=1}^{N} w_n \mathbb{E}[R_n] + w_f R_f = w^\top \mu + w_f R_f = \mu_{\mathbf{targ}}.$$

- But, constraint (1) implies that:

$$w_f = 1 - w^\top 1_N.$$

Portfolio Optimization with A Risk-Free Asset . . . III

- Substituting this expression for w_f into constraint (2), we get that the problem to be solved is:

$$\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad w^\top \mu + (1 - w^\top 1_N) R_f = \mu_{\text{targ}}.$$

Portfolio Optimization with A Risk-Free Asset . . . IV

- Solving the above problem gives the optimal portfolio weights:

$$w = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)} \right] V^{-1} (\mu - R_f 1_N). \qquad (22)$$

- We will derive this result below, after looking at an example.

- Note that the term in blue in the square brackets, is a scaling term:
  - the higher is μ_{targ}, the larger is the position in the optimal portfolio of risky assets, given by the term in red, $V^{-1}(\mu - R_f 1_N)$ (and the smaller is the investment in the risk-free asset).
  - Below, we will see that the term in the blue square brackets plays the same role as an investor's risk tolerance, given by the term $1/\gamma$, where γ is the investor's risk aversion.

Example with three risky assets and a risk-free asset . . . I

- Consider the example with three risky assets discussed previously.
- Assume that now there is also a risk-free asset available with $R_f = 6.5\%$.
- Find the optimal portfolio for an investor who requires an expected rate of return of $\mu_{\text{targ}} = 30\%$.

Example with three risky assets and a risk-free asset . . . II

- We know that

$$w = \left[\frac{\mu_{\text{targ}} - R_f}{(\mu - R_f 1_N)^\top V^{-1}(\mu - R_f 1_N)}\right] V^{-1}(\mu - R_f 1_N)$$

$$= \left[\frac{0.30 - 0.065}{0.6382}\right] \begin{bmatrix} 85.06 & -37.61 & -16.03 \\ -37.61 & 26.35 & 4.82 \\ -16.03 & 4.82 & 6.33 \end{bmatrix} \begin{bmatrix} 0.135 \\ 0.235 \\ 0.335 \end{bmatrix}$$

$$= \frac{0.30 - 0.065}{0.6382} \begin{bmatrix} -2.7268 \\ 2.7299 \\ 1.0889 \end{bmatrix}$$

$$= \begin{bmatrix} -1.0041 \\ 1.0052 \\ 0.4010 \end{bmatrix}.$$

Example with three risky assets and a risk-free asset . . . III

- We can use these weights to identify the return volatility of this portfolio:

$$\begin{aligned}\sigma_p &= \sqrt{w^\top V w} \\ &= 0.2942 \quad \text{. . . volatility of this portfolio's return}\end{aligned}$$

Example with three risky assets and a risk-free asset . . . IV

- Comparing the three portfolios we have studied, all of which give an expected return of $\mu_{\text{targ}} = 0.30$:
- we see that as the number of available assets (N) increases,
  - the investor can achieve the same expected return (0.30)
  - at a lower volatility.

Case	$N = 1$ Asset 2 alone	$N = 3$ Three risky assets	$N = 3$ plus risk-free
σ_p	0.3500	0.3256	0.2942

Python code for example with risk-free asset . . . I

Define solution with risk-free asset

```
# This uses the quantities defined in the previous example.
# So, make sure you have defined mu, V, a1, a2, a3, D, w0, and w1

# Define solution function when risk-free asset is available
def solutionWithRf(mu_targ, Rf):
    term1Numerator = (mu_targ - Rf)
    term1Denominator = (mu - Rf*ones).T @ Vinverse @ (mu - Rf*ones)
    term1 = term1Numerator/term1Denominator
    term2 = Vinverse @ (mu - Rf*ones)
    weightsWithRf = term1 * term2
    portfolioVolatilityWithRf = np.sqrt(weightsWithRf.T @ V @
     weightsWithRf)

    print("Optimal weights are:")
    print(weightsWithRf)
    print(" ")
    print("Volatilitiy of optimal portfolio's return is:",
     portfolioVolatilityWithRf[0,0])
    return
```

Python code for example with risk-free asset ... II

Solution for target mean of 0.30 and risk-free rate of 0.065

```
solutionWithRf(0.30,0.065)
```

```
Optimal weights are:
[[-1.00410587]
 [ 1.00522836]
 [ 0.40097203]]

Volatilitiy of optimal portfolio's return is: 0.29416783123749785
```

Example: Portfolio weights in the risk-free asset

- Now that we know the proportion of wealth invested in the three risky assets,
- we can find the proportion invested in the risk-free asset:

$$
\begin{aligned}
w_f &= 1 - \big(w_1 + w_2 + w_3\big) \\
&= 1 - \big(-1.0041 + 1.0052 + 0.4010\big) \\
&= 0.5979 \quad \ldots \text{in risk-free asset}
\end{aligned}
$$

Portfolio weights in the tangency portfolio

- We can also find out the weights in the tangency portfolio, w_{tang}.
- The tangency portfolio is the portfolio with investment in only risky assets (and nothing in the risk-free asset).
- We calculate the tangency portfolio weights by dividing the weight of each risky asset by the sum of the weights in the N risky assets:

$$w_{\text{tang}} = \frac{\text{optimal weight vector}}{\text{sum of weights in all risky assets}} = \frac{w}{w^\top 1_N}. \tag{23}$$

- Note that w and w_{tang} are N-dimensional vectors.

Example: Portfolio weights in the tangency portfolio

- For our numerical example,

$$
\begin{aligned}
w_{\text{tang}} &= \frac{w}{w^\top 1_N} = \frac{1}{w_1 + w_2 + w_3} \times w \\
&= \frac{1}{(-1.0041 + 1.0052 + .4010)} \times \begin{pmatrix} -1.0041 \\ 1.0052 \\ 0.4010 \end{pmatrix} \\
&= \frac{1}{(0.4021)} \times \begin{pmatrix} -1.0041 \\ 1.0052 \\ 0.4010 \end{pmatrix} = \begin{pmatrix} -2.4971 \\ 2.4998 \\ 0.9972 \end{pmatrix}
\end{aligned}
$$

- Note that the weights in the tangency portfolio add up to 1, implying that the tangency portfolio's weight in the risk-free asset is 0.

Python code for weights of tangency portfolio

Code for weights of tangency portfolio

```
# Compute weights in tangency portfolio
def weightsTangencyPortfolio(mu_targ, Rf):
    term1Numerator = (mu_targ-Rf)
    term1Denominator = (mu - Rf*ones).T @ Vinverse @ (mu - Rf*ones)
    term1 = term1Numerator/term1Denominator
    term2 = Vinverse @ (mu - Rf*ones)
    weightsWithRf = term1 * term2
    weightsTangency = weightsWithRf/(weightsWithRf.T @ ones)
    print("Tangency portfolio weights are:")
    print(weightsTangency)
    return
```

Weights in tangency portfolio: target return = 0.30, risk-free rate = 0.065

```
weightsTangencyPortfolio(0.30, 0.065)
```

```
Tangency portfolio weights are:
[[-2.49718869]
 [ 2.49998029]
 [ 0.9972084 ]]
```

Portfolio optimization by maximizing mean-variance utility

- Recall that to find the mean-variance efficient portfolio in the presence of a risk-free asset, we
  - minimize the portfolio variance $\sigma_p^2 = w^\top V w$, subject to the
  - constraint that the expected return of the portfolio is equal to the target portfolio return, μ_{targ}; that is,

$$\min_w \quad \sigma_p^2 = w^\top V w$$
$$\text{subject to:} \quad (1 - w^\top 1_N) R_f + w^\top \mu = \mu_{\text{targ}}.$$

- Markowitz shows that the above problem is equivalent to maximizing mean-variance utility, which is written on the next page.

Maximizing mean-variance utility (MVU)

- Maximizing mean-variance utility (MVU) implies maximizing the difference between the
  - expected return (μ_p) on the portfolio, and
  - variance (σ_p^2) of the portfolio,
  - adjusted for the risk-aversion γ of the investor.

$$\max_w \quad \mu_p - \frac{\gamma}{2}\sigma_p^2$$

$$\Rightarrow \max_w \quad \underbrace{\left[(1 - w^\top 1_N)R_f + w^\top \mu\right]}_{\mu_p} - \frac{\gamma}{2}\underbrace{w^\top V w}_{\sigma_p^2}$$

$$\Rightarrow \max_w \quad \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w \quad \text{...collecting terms in } w$$

Optimal weights from maximizing mean-variance utility

- We start by rewriting the last expression on the previous page:

$$\max_{w} \quad \text{MVU} = \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w.$$

- The first-order condition with respect to w is:

$$\frac{\partial\, \text{MVU}}{\partial w} = (\mu - R_f 1_N) - \gamma V w = 0_N.$$

- Solving this equation (by multiplying both sides by $\frac{1}{\gamma} V^{-1}$) gives:

$$w = \frac{1}{\gamma} V^{-1}(\mu - R_f 1_N) \quad \ldots \text{optimal portfolio}. \tag{24}$$

- In this formulation, one can identify the entire mean-volatility portfolio frontier by varying the investor's level of risk aversion, γ.

Equivalence between weights from the two approaches

- We start by writing the last expression on the previous page again:

$$w = \frac{1}{\gamma} V^{-1}(\mu - R_f 1_N) \quad \ldots \text{optimal portfolio.} \tag{24}$$

- But, in equation (22) on page 341, we had derived the optimal weight:

$$w = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)}\right] V^{-1} (\mu - R_f 1_N) \quad \ldots \text{optimal portfolio.}$$

- We see that the two expressions are identical if we set

$$\frac{1}{\gamma} = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)}\right].$$

Q. Why do the LHS and RHS represent the same economic quantity?

Portfolio weights in the tangency portfolio . . . I

- The vector of relative portfolio weights, w_{tang}, of the N risky assets in the tangency portfolio is:

$$w_{\text{tang}} = \frac{\text{optimal weight}}{\text{sum of weights in all risky assets}} = \frac{V^{-1}(\mu - R_f 1_N)}{1_N^\top V^{-1}(\mu - R_f 1_N)}.$$

- Note that an investor's risk aversion, γ, does not appear in w_{tang} (because it cancels out in the numerator and denominator).

- Thus, all investors will hold risky assets in the same proportion even if they have different risk aversion.

- Thus, all investors demand the same portfolio of risky assets: w_{tang}.

- In equilibrium, the supply of risky assets (market portfolio) must match the demand for risky assets; thus, the tangency portfolio is the market portfolio.

Example with Python code

- Consider the following simple two-asset example.
- Suppose that:
  - the investor's risk aversion is $\gamma = 5$;
  - p.a. gross risk-free rate is 1.01;
  - p.a. gross mean returns for both assets are the same, 1.09;
  - p.a. volatility of returns for both assets are the same, 20%; and
  - the return correlation of the two risky assets is 0.99.
- Find the optimal weights for
  - the mean-variance portfolio, and
  - the tangency portfolio.

Python code for weights in tangency portfolio . . . I

Code for weights in tangency portfolio

```
# Import numpy
import numpy as np

# Define a column vector of ones
ones = np.ones((2, 1))

# Define the gross risk-free rate
Rf = 1.01

# Column vector of gross mean returns
mu = np.array( [     [1.09],    [1.09]    ]  )

# Variance-covariance matrix
V = np.array(
    [     [0.20**2, 0.20*0.20*0.99],
    [0.20*0.20*0.99, 0.20**2]    ]   )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)                        # continued on next slide
```

Python code for weights in tangency portfolio . . . II

Code for weights in tangency portfolio (cont.)

```
# Mean-variance weights for gamma = 5
gamma = 5
w_mv = (1/gamma) * Vinverse @ (mu-Rf*ones)

# Weight in risk-free asset
w_f = 1 - ones.T @ w_mv

# Weights of tangency portfolio
w_Tangency = w_mv/(ones.T @ w_mv)

# Print portfolio weights
# Print portfolio weights
print("Mean-variance weights = ", "\n", w_mv)
print(" ")
print("Weight in risk-free asset = ","\n", w_f)
print(" ")
print("Tangency portfolio weights = ","\n", w_Tangency)
```

Solution from code on the previous pages

```
Mean-variance weights =
 [[0.20100503]
 [0.20100503]]

Weight in risk-free asset =
 [[0.59798995]]

Tangency portfolio weights =
 [[0.5]
 [0.5]]
```

The market portfolio

- The market portfolio is a portfolio in which the weight of asset n equals the market-capitalization value of firm n, relative to the market-capitalization value of all the assets in the market.

- The market-capitalization value of a firm is

  Market capitalization of firm $n =$ [number of shares outstanding of n] $\times$ [price of

- Thus, the market portfolio is defined by the following weights:

$$
\begin{aligned}
w_n^{\text{mkt}} &= \frac{\text{Market capitalization of asset } n}{\text{Total market capitalization of all assets in the market}} \\
&= \frac{\big[\text{number of shares outstanding of } n\big] \times \big[\text{price of share } n\big]}{\sum_{m=1}^{N} \big[\text{number of shares outstanding of } m\big] \times \big[\text{price of share } m\big]} \\
&= \frac{\text{num}_n \times P_n}{\sum_{m=1}^{N} \text{num}_m \times P_m}.
\end{aligned}
$$

Reasons to hold just the market portfolio

- If one believes that the CAPM is a good description of the world, then the only portfolio of risky assets that investors should hold is the market portfolio.
- If one believes in the CAPM, it also suggests that there is little benefit from actively managed funds.
- Thus, the CAPM suggests that investors should pursue a passive strategy of holding only the market portfolio (besides investing in the risk-free asset).

Problems in holding the market portfolio

- From a practical standpoint, it is difficult to hold all the assets in the market.
- Thus, instead of holding the market portfolio,
  - investors usually hold a capitalization-weighted (that is, value-weighted) index,
  - where the value-weighted index invests in a smaller number of securities but is designed to track the market portfolio.
- This was the original motivation for market-capitalization weighted index funds, which then led to the growth of exchange-traded ETFs that are designed to track the market.

Benefits of investing in a capitalization-weighted portfolio

- A capitalization-weighted index provides exposure to a broad set of stocks.
- Capitalization weighting is a passive strategy that requires little trading, so transaction costs are low.
- Market capitalization is highly correlated with trading liquidity (because more is invested in larger firms), which further reduces transaction costs.
- Market capitalization is highly correlated with investment capacity, which makes it easier to scale up this investment strategy.

Limitations of capitalization-weighted portfolios

- As we will see in later chapters,
  - the CAPM may not be the best description of the world, and
  - the returns on the market portfolio may be less than the returns on other portfolios that account for factors such as size, and momentum;
- In light of this, it may be better to invest in other indexes:
  - Indexes based on factors other than capitalization value;
  - Indexes with exposure to factors other than the market;
  - Other assets, in addition to investing in the index.

Relation between returns of individual assets and frontier portfolios . . . I

So far, we have learned two main ideas:

1. Investors hold portfolios to reduce risk.
    - "Unsystematic risks" of individual assets doesn't matter.
    - Only "systematic risks" matter.
2. Investors hold only frontier portfolios.

Relation between returns of individual assets and frontier portfolios . . . II

- The natural questions to ask next are:
  - How does an individual asset contribute to the risk of portfolios, especially the frontier portfolios?
  - Can we be more specific about what is "systematic risk"?
  - What is the relation between an asset's systematic risk and its expected return?

Relation between returns of individual assets and frontier portfolios . . . III

- To simplify the analysis, we first answer the following four questions: What is the . . .
  1. Marginal contribution of an asset to a portfolio's expected return;
  2. Total contribution of an asset to a portfolio's variance;
  3. Marginal contribution of an asset to a portfolio's variance;
  4. Marginal contribution of an asset to a portfolio's volatility.
- Then, we will use the answers from these four steps to derive an extremely useful and powerful result.

Step 1: Marginal contribution of an individual asset to a portfolio's expected return . . . I

- We assume the existence of a risk-free asset.
- The return on a portfolio p is

$$\begin{aligned} R_p &= \left(1 - \sum_{n=1}^{N} w_n\right) R_f + \sum_{n=1}^{N} w_n R_n \\ &= R_f + \sum_{n=1}^{N} w_n \left(R_n - R_f\right) \quad \text{. . . rearranging terms.} \end{aligned}$$

Step 1: Marginal contribution of an individual asset to a portfolio's expected return ... II

- A portfolio's expected return is:

$$\mathbb{E}[R_p] = R_f + \sum_{n=1}^{N} w_n \Big(\mathbb{E}[R_n] - R_f \Big).$$

- Then, the marginal contribution of risky asset n to the portfolio's expected return is

$$\frac{\partial \mathbb{E}[R_p]}{\partial w_n} = \mathbb{E}[R_n] - R_f,$$

  that is, the marginal contribution is the asset's risk premium.

Step 2: Total contribution of an asset to portfolio variance

- Recall that the variance of portfolio return is the sum of all entries of the following table:

	$w_1 R_1$	$w_2 R_2$	$\cdots$	$w_N R_N$
$w_1 R_1$	$w_1^2 \sigma_1^2$	$w_1 w_2 \sigma_{12}$	$\cdots$	$w_1 w_N \sigma_{1n}$
$w_2 R_2$	$w_2 w_1 \sigma_{21}$	$w_2^2 \sigma_2^2$	$\cdots$	$w_2 w_N \sigma_{2n}$
$\cdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$w_N R_N$	$w_N w_1 \sigma_{N1}$	$w_N w_2 \sigma_{N2}$	$\cdots$	$w_N^2 \sigma_N^2$

- The sum of the entries of the n-th-row and the n-th column are:

$$\text{Total contribution of asset } n \text{ to portfolio variance} = w_n^2 \sigma_n^2 + 2 \sum_{m \neq n}^{N} w_n w_m \sigma_{nm}.$$

Step 3: Marginal contribution of asset to portfolio variance

- We now compute the marginal contribution of asset n to the portfolio return variance:

$$\begin{aligned}
\frac{\partial \sigma_p^2}{\partial w_n} &= \frac{\partial}{\partial w_n}\left(w_n^2 \sigma_n^2 + 2 \sum_{m \neq n}^{N} w_n w_m \sigma_{nm} \right) \\
&= 2 w_n \sigma_n^2 + 2 \sum_{m \neq n}^{N} w_m \sigma_{nm} \\
&= 2 \sum_{m=1}^{N} w_m \sigma_{nm} \\
&= 2 \operatorname{Cov}[R_n, R_p].
\end{aligned}$$

- To prove the last step above, note that:

$$\operatorname{Cov}[R_n, R_p] = \operatorname{Cov}[R_n, \sum_{m=1}^{N} w_m R_m] = \sum_{m=1}^{N} w_m \operatorname{Cov}[R_n, R_j] = \sum_{m=1}^{N} w_m \sigma_{nm}.$$

Step 4: Marginal contribution of asset to portfolio volatility

- On the previous slide, we have shown that:

$$\frac{\partial \sigma_p^2}{\partial w_n} = 2\,\text{Cov}[R_n, R_p] = 2\sigma_{np}.$$

- But we wish to find $\frac{\partial \sigma_p}{\partial w_n}$, the marginal contribution of asset n to portfolio volatility (instead of portfolio variance).

- We know that:

$$\frac{\partial \sigma_p^2}{\partial w_n} = 2\sigma_p \frac{\partial \sigma_p}{\partial w_n} = 2\sigma_{np},$$

which implies that

$$\frac{\partial \sigma_p}{\partial w_n} = \frac{2\sigma_{np}}{2\sigma_p} = \frac{\sigma_{np}}{\sigma_p}.$$

- Observation: An individual asset's marginal contribution to portfolio volatility depends only on σ_{np}/σ_p; that is, the ratio of its covariance with the portfolio return (σ_{np}) to the portfolio volatility (σ_p).

Risk-return of individual assets & frontier portfolios . . . I

- We now put together the results we have derived for
  1. the marginal contribution of an asset to a portfolio's expected return;
  2. the marginal contribution of an asset to a portfolio's volatility;
- to identify the risk-return relation between
  1. individual assets and
  2. frontier portfolios.

Risk-return of individual assets & frontier portfolios . . . II

- Definition: The (marginal) return-to-risk ratio (RRR) of risky asset n in a portfolio p is:

$$\text{RRR}_n = \frac{\text{marginal return}}{\text{marginal risk}} = \frac{\partial \mathbb{E}[R_p]/\partial w_n}{\partial \sigma_p / \partial w_n} = \frac{\mathbb{E}[R_n] - R_f}{(\sigma_{np}/\sigma_p)}.$$

- Claim: For any frontier portfolio p, the return-to-risk ratio of all risky assets must be the same:

$$\text{RRR}_n = \text{RRR}_p$$

$$\frac{\mathbb{E}[R_n] - R_f}{(\sigma_{np}/\sigma_p)} = \frac{\mathbb{E}[R_p] - R_f}{(\sigma_{pp}/\sigma_p)} = \frac{\mathbb{E}[R_p] - R_f}{\sigma_p} \quad \ldots \text{recall that } \sigma_{pp} = \sigma_p^2.$$

- Intuition: The RRR of a frontier portfolio cannot be improved.

Risk-return of individual assets & frontier portfolios . . . III

- Re-writing the last result on the previous slide, we have:

$$\frac{\mathbb{E}[R_n] - R_f}{(\sigma_{np}/\sigma_p)} = \frac{\mathbb{E}[R_p] - R_f}{\sigma_p}$$

- Re-arranging the above, we get the following important result:

$$\begin{aligned} \mathbb{E}[R_n] - R_f &= \frac{\sigma_{np}}{\sigma_p^2}\left(\mathbb{E}[R_p] - R_f\right) \\ &= \beta_{np}\left(\mathbb{E}[R_p] - R_f\right), \quad \text{where} \end{aligned}$$

$$\beta_{np} = \frac{\sigma_{np}}{\sigma_p^2} = \frac{\text{Cov}[R_n, R_p]}{\text{Var}[R_p]}.$$

Risk-return of individual assets & frontier portfolios . . . IV

$$\mathbb{E}[R_n] - R_f = \beta_{np}\big(\mathbb{E}[R_p] - R_f\big),$$

- The interpretation of the above important result, for any frontier portfolio p (except the risk-free asset), is that
  - β_{np} gives a measure of asset n's systematic risk.
  - $\mathbb{E}[R_p] - R_f$ gives the premium (price) per unit of systematic risk.
  - The risk premium on asset n then equals the amount of its systematic risk times the premium per unit of the risk.

End of focus

Road map

Portfolio optimization with respect to a benchmark

- Professional money managers are often evaluated based on performance relative to a prespecified benchmark.
- This is a sensible approach because an investor's alternative to an active manager is to invest in an index fund close to the benchmark.
- Fees and transaction costs on index funds are lower, so an active manager provides value only if performance relative to the benchmark is positive on average.
- Because returns are noisy,
  - a long time can elapse before we have a precise estimate of the average return (as discussed already);
  - therefore, investors often focus on the volatility of tracking error.
- Minimization of tracking error volatility (TEV) is often used to evaluate overall manager performance.

Performance evaluation relative to a benchmark

- Thus, performance is evaluated along two dimensions:
    1. Beating the benchmark on average, which is equivalent to having a positive tracking error;
    2. Reducing volatility of tracking error, which is equivalent to minimizing the variance of the difference between managed portfolio returns and benchmark returns.
- This is similar to Markowitz mean-variance analysis, but instead of
    - minimizing total return volatility for a given expected total return, we minimize tracking error volatility for a given expected tracking error.
- Roll (1992) calls minimizing tracking error volatility, for a given expected tracking error, the Tracking Error Variance (TEV) criterion.

Mean-variance optimization with respect to a benchmark

- Let us denote the weights of the benchmark portfolio by w_B.
- Then, the tracking error is the difference between
  - the return R_A of the active portfolio w_A one chooses, and
  - the return R_B of the benchmark portfolio w_B.

$$\begin{aligned}
\text{TrackingError} &= R_A - R_B \\
&= \sum_{i=n}^{N} w_{A,n} R_n - \sum_{n=1}^{N} w_{B,n} R_n \\
&= w_A^\top R - w_B^\top R \\
&= (w_A - w_B)^\top R.
\end{aligned}$$

Expected value and volatility of tracking error

- We start by re-writing the expression for the tracking error:

$$\text{TrackingError} = (w_A - w_B)^\top R.$$

- The expected return of the active strategy over the benchmark is

$$\text{Expected TrackingError} = (w_A - w_B)^\top \mu.$$

- The variance of the tracking error is

$$\text{Variance of TrackingError} = (w_A - w_B)^\top V (w_A - w_B).$$

- The volatility of the tracking error is

$$\text{Volatility of TrackingError} = \sqrt{(w_A - w_B)^\top V (w_A - w_B)}.$$

Objective of investment manager

- The objective of the portfolio manager then is to
  - choose the active portfolio w_A
  - in order to maximize the expected tracking error,
  - while satisfying the constraints that
    - the volatility of the tracking error be no more than some σ_{max}, and
    - the sum of the weight equals 1.
- This problem can be written as:

$$\max_{w_A} \left(w_A - w_B\right)^\top \mu$$

$$\text{subject to:} \quad \sqrt{\left(w_A - w_B\right)^\top V \left(w_A - w_B\right)} < \sigma_{\text{max}},$$

$$\text{and subject to:} \quad w_A^\top 1_N = 1.$$

The Information Ratio

- Performance of an active portfolio strategy, w_A, with respect to a benchmark portfolio, w_B, is often measured by the Information ratio.
- The Information ratio is similar to the Sharpe ratio, but
  - instead of measuring return and risk relative to that of the risk-free rate,
  - we measure return and risk relative to that of the benchmark portfolio:

$$\text{Information ratio} = \frac{\text{expected tracking error}}{\text{volatility of tracking error}}$$

$$= \frac{(w_A - w_B)^\top \mu}{\sqrt{(w_A - w_B)^\top V (w_A - w_B)}}.$$

The Appraisal Ratio

- Just as the Information ratio is the counterpart to the Sharpe ratio when performance is evaluated relative to a benchmark,
- the Appraisal ratio is the counterpart to the Treynor ratio when performance is evaluated relative to a benchmark.
- The Appraisal ratio is
  - similar to the Treynor ratio in that the numerator has the alpha measured with respect to a one-, four-, five, or six-factor model,
  - but, in contrast to the Treynor ratio, risk is measured relative to that of the benchmark portfolio, i.e., by the volatility of the tracking error:

$$\text{Appraisal ratio} = \frac{\text{risk-adjusted excess return relative to a factor model}}{\text{volatility of tracking error}}$$
$$= \frac{\alpha}{\sqrt{(w_A - w_B)^\top V (w_A - w_B)}}.$$

Road map

Portfolio optimization with other constraints

- If there are other constraints, either for regulatory reasons or for other reasons, they can be added to the optimization problem:

$$\min_{\{w_1,\ldots,w_N\}} \quad \text{Portfolio risk} = \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m\, \sigma_{nm}$$

subject to the constraints:

$$
\begin{array}{lll}
(1) & \displaystyle\sum_{n=1}^{N} w_n = 1 & \ldots\text{weights sum to 1} \\
(2) & \displaystyle\sum_{n=1}^{N} w_n\, \mathbb{E}[R_n] = \mu_{\text{targ}} & \ldots\text{expected return is } \mu_{\text{targ}} \\
(3) & w_n \geq w_{\min} & \ldots\text{long-only constraint if } w_{\min} = 0 \\
(4) & w_n \leq w_{\max} & \ldots\text{constraint on max weight.}
\end{array}
$$

- This kind of problem is easy to solve in Python using an optimization library, which we will study in the next chapter.

Road map

Portfolio optimization with transaction costs

- There are at least three ways of incorporating the effect of transaction costs (and taxes) when evaluating returns of an optimal portfolio.

  1. One could ignore transaction costs (and taxes) in the optimization, but account for them when calculating the return on the portfolio.
  2. One could account for the transaction costs and taxes in the optimization but assume that the optimization is done myopically.
     - Theoretically incorrect, but this problem is easy to solve.
  3. One could account for the transaction costs and taxes in the optimization, and do the optimization dynamically, that is, taking into account that one can rebalance in the future.
     - Theoretically correct, but a much more difficult problem to solve.

Ignore transaction costs in optimization, but include when computing returns

- In the first approach, we
  - ignore transaction costs (TC) in the optimization,
  - but subtract them when calculating the return on the portfolio.

$$[\text{Return net of transaction costs}]_t = R_t - \sum_{n=1}^{N} \underbrace{|w_{n,t} - w_{n,t^+}|}_{\text{turnover of asset } n} \times \text{TC}.$$

Myopic optimization with transactions Costs

- Under the second approach, one incorporates transaction costs explicitly in the mean-variance portfolio optimization problem:

$$\max_{w} \quad \underbrace{w^\top \mu}_{\text{mean}} - \frac{1}{\gamma} \underbrace{w^\top V w}_{\text{variance}} - \text{TC}\|w - w_0\|_1,$$

$$\text{s.t.} \quad \sum_{n=1}^{N} w_n = 1,$$

where

- $\text{TC} = \kappa$ is the rate of proportional transactions cost,
- w_0 is the portfolio before trading,
- $\|w - w_0\|_1$ is the difference of the portfolio weights before and after trading,
- $\text{TC}\|w - w_0\|_1$ is the transactions cost.

Dynamic optimization with transactions costs

- Under the third approach, one
  - accounts for the transaction costs in the optimization, and
  - does the optimization dynamically, that is, taking into account that one can rebalance in the future.
- The details are provided in DeMiguel and Uppal (2005).
  - We will not study this approach in the course.

Road map

Summary of main points of modern portfolio theory

1. Investors hold frontier portfolios; large asset base improves portfolio frontier.
2. When there is a risk-free asset, frontier portfolios are linear combinations of the risk-free asset and the tangent portfolio.
3. An individual asset's risk is its contribution to portfolio risk:
   - Variance of each asset contributes little to portfolio risk; covariances between assets determine portfolio risk.
4. For any asset n:
   - Its risk can be measured by its beta with respect to a frontier portfolio.
   - Its return per unit of risk is the premium on the frontier portfolio.

Python code: Links to code available online

- This page has Python code for computing:
  - portfolio return mean, variance, and volatility
  - efficient frontier
  - portfolio with the maximum Sharpe ratio
  - global minimum-variance portfolio.
- This code is from EDHEC's Coursera offering.
- "PyPortfolioOpt" is a library for mean-variance portfolio optimization (and also the Black-Litterman allocation model, which we will study later in the course).
  - See, in particular, Mean-variance optimization.
- "Riskfolio-Lib" is an extensive Python library with a large number of functions for making portfolio optimization decisions.

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will study various methods to reduce the impact of the error in estimating the expected value, variances, and covariances of returns in order to improve the performance of mean-variance portfolios.

To do for next class

- Readings
  - This chapter's material is based on the papers by Markowitz (1952, 1959), Merton (1972), and Roll (1977, 1980). The best textbook reference for this material is Huang and Litzenberger (1988, ch. 2).
  - Please read Moreira (2021, Chapters 7–8), available online from this link.
  - Please read the subsection on "Constrained Optimization and Backtesting" in Scheuch, Voigt, Weiss, and Frey (2024), available online at Tidy Finance.
- Assignment
  - You can also start reading (and working) on the first question of the second assignment (even though the second question depends on the material to be covered in the next chapter.

Road map

Bibliography ... I

DeMiguel, V., and R. Uppal. 2005. Portfolio investment with the exact tax basis via nonlinear programming. *Management Science* 51, no. 2 (February): 277–290. (Cited on pages 163, 393).

Huang, C.-F., and R. Litzenberger. 1988. *Foundations for financial economics.* New Jersey: Prentice Hall. (Cited on pages 302, 399).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

Merton, R. C. 1972. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7 (4): 1851–72. (Cited on pages 302, 399).

Moreira, A. 2021. *Quantitative investing.* Available online from this link. (Cited on pages 248, 399, 559).

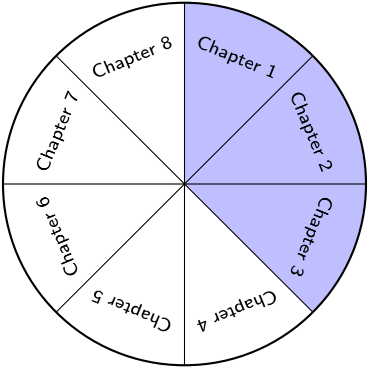
Roll, R. 1977. A critique of the asset pricing theory's tests Part i: On past and potential testability of the theory. *Journal of Financial Economics* 4, no. 2 (March): 129–176. (Cited on page 399).

———. 1980. Orthogonal portfolios. *Journal of Financial and Quantitative Analysis* 15 (5): 1005–1023. (Cited on page 399).

Bibliography . . . II

Roll, R. 1992. A mean/variance analysis of tracking error. *Journal of Portfolio Management* 18 (4): 13–22. (Cited on page 381).

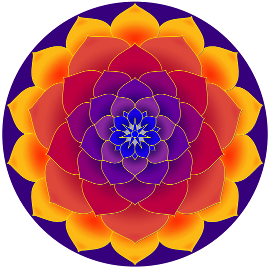
Scheuch, C., S. Voigt, P. Weiss, and C. Frey. 2024. *Tidy finance with Python.* Chapman-Hall/CRC. Available online from this link. (Cited on pages 100, 399, 559).



End of Chapter 3

Blank page 1

Quantitative Portfolio Management



Chapter 4:
Mean-Variance Portfolios that *Adjust for* Estimation Error

Raman Uppal

2025

The big picture: Plan for the entire course

Part A: Preliminaries

Chapter 1: Properties of asset returns

Chapter 2: Performance measurement (especially out of sample)

Part B: Modern portfolio management

Chapter 3: Mean-variance portfolios that *ignore* estimation error

Chapter 4: Mean-variance portfolios that *adjust for* estimation error

Part C: Post-Modern Portfolio Management

Chapter 5: CAPM-based portfolios: Black-Litterman model

Chapter 6: Factor-based portfolios: Parametric portfolio policies

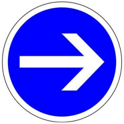
Chapter 7: Volatility-timed factor portfolios

Chapter 8: Portfolios exploiting systematic risk factors *and* unsystematic risk

Table of contents

Road map

What do we want to do in Chapter 4



In this chapter, we first show that the out-of-sample performance of sample-based mean-variance optimal portfolios is poor.

Then, we examine various methods to improve their out-of-sample performance.

Start of focus

Road map

Mean-variance optimal portfolios . . . I

- In the last chapter, we looked at models of optimal portfolio choice.
- The optimal portfolio weight vector is of the form:

$$\text{weights} = [\text{risk tolerance}] \times (\text{covariances})^{-1} \times \text{Expected risk premium}$$

$$w = \frac{1}{\gamma} \left(\mathbb{V}[R]\right)^{-1} \left(\mathbb{E}[R] - R_f 1_N\right), \quad \ldots \text{optimal portfolio}.$$

Estimating inputs for mean-variance portfolios . . . I

- Implementing optimal portfolios requires estimation of parameters:
  - Expected (mean) returns: $\mathbb{E}[R]$
  - Volatilities and correlations (covariances) of returns: $\mathbb{V}[R]$.
- These moments are estimated with error.

Estimating inputs for mean-variance portfolios . . . II

- In this chapter, we first understand the impact of estimation error on the **out-of-sample** performance of portfolios.
- Then, we study models that try to reduce the impact of estimation error on portfolio performance.
  - First, we consider models that focus on the estimation error in expected returns, $\mathbb{E}[R]$;
  - Then, we consider models that focus on the estimation error in the variance-covariance matrix of returns, $\mathbb{V}[R]$;

Road map

Estimation error in $\mathbb{E}[R]$

- The error in estimating expected returns is particularly large.
- Merton (1980) explains why the precision of estimates of expected cannot be improved even if returns are sampled more frequently (e.g., daily instead of monthly).
  - This is because estimates of expected returns depend only on the first and last price observation (i.e., intermediate prices drop out).
  - We saw this in Chapter#1.
- Portfolio weights are very sensitive to estimates of expected returns.

How precise is the estimate of the mean $\mathbb{E}[R_n]$?

- How precise is the estimate of the mean $\mathbb{E}[R_n]$?
- To answer this question, we need to find the standard error (standard deviation) of the estimator of $\mathbb{E}[R_n]$:
- If the annual variance of a return is σ_n^2:
  - Then, the annual variance of the average return is $\frac{\sigma_n^2}{T}$.
  - Thus, the standard deviation of $\mathbb{E}[R_n]$ is $\frac{\sigma_n}{\sqrt{T}}$.

How large is the standard error of $\mathbb{E}[R_n]$?

- The quantity $\sigma_n/\sqrt{T}$ is quite large because
  - For individual stocks, σ_n is about 0.30 p.a.
  - Thus, for nine years of data, $T = 9$; $\sigma_n/\sqrt{T} = 0.30/\sqrt{9} = 10\%$;
  - Thus, a 95% confidence interval (± 2 standard deviations) will be a band of about: $\pm 2 \times 10\% = 4 \times 10\% = 40\%$;
- Given that expected returns are of the order 10%, the 95% confidence interval is

$$0.10 \pm (2 \times 10\%) = 0.10 \pm 0.20 = -10\% \text{ to } 30\%.$$

- Such a large confidence interval implies an estimate that is so imprecise that it is useless for any practical purpose.
- Bottom line:
  Beware of any model that uses expected return as an input.

Estimation error in $\mathbb{V}[R]$ ... I

- The covariance matrix of returns has a large number of elements that need to be estimated:
  - N variances
  - $(N^2 - N)/2$ covariances
  - for a total of $N + (N^2 - N)/2 = (N^2 + N)/2 = N(N+1)/2$.
- The estimated covariance matrix is usually
  - ill-conditioned (it has a large "condition number")
    (for a discussion of "condition number," see Wikipedia);
  - and is often close to being singular.
- These problems get worse as N increases because the number of unique elements in the covariance matrix increases at the rate N^2.

Estimation error in $\mathbb{V}[R]$ ... II

- For a portfolio of $N = 50$ assets, the total number of covariance terms $= 1275$, which, if one were using monthly data, would require $T > 1275/12 = 107$ years of monthly data.
- For a portfolio of $N = 100$ assets, the total number of variance and covariance terms $= 5050$, which,
  - would require $T > 5050/12 = 421$ years of monthly data, or
  - would need $T > 5050/250 = 21$ years of daily data.
- Thus, the estimated covariance matrix is ill-conditioned because,
  - relative to the number of data points,
  - the number of parameters to be estimated is very large.

Effect on weights of error in estimating $\mathbb{E}[R]$ and $\mathbb{V}[R]$

- Above, we have seen that
  - sample-based estimates of $\mathbb{E}[R]$ are not very precise
  - sample-based estimates of $\mathbb{V}[R]$ are ill-conditioned

$$w = \frac{1}{\gamma} \underbrace{(\mathbb{V}[R])^{-1}}_{\text{ill conditioned}} \underbrace{(\mathbb{E}[R] - R_f 1_N)}_{\text{low precision}}$$

- Consequently, expected returns and covariance matrices estimated using classical sample estimators lead to portfolio weights that are
  - Unreasonable—extremely large or small;
  - Unstable—fluctuate a lot over time;
  - Inefficient—perform poorly out of sample.

Road map

Example illustrating effect of estimation error . . . I

- Consider the following simple two-asset example.
  - First, we consider exactly the same setting as in the previous chapter.
  - Then, we change the example slightly to study the change in weights.
- Suppose that:
  - the investor's risk aversion is $\gamma = 5$;
  - p.a. gross risk-free rate is 1.01;
  - p.a. gross mean returns for both risky assets are the same, 1.09;
  - p.a. volatility of returns for both risky assets are the same, 20%; and
  - the return correlation of the two risky assets is 0.99.
- Find the optimal weights of the mean-variance tangency portfolios.

Python code for portfolio weights with *true* parameters

Tangency portfolio weights with *true* parameters

```
# Import numpy
import numpy as np

# Define a column vector of ones
ones = np.ones((2, 1))

# Define the gross risk-free rate of returns
Rf = 1.01

# Column vector of gross mean returns
mu = np.array(  [  [1.09],  [1.09]  ]  )

# Variance-covariance matrix
V = np.array(   [
    [0.20**2, 0.20*0.20*0.99],
    [0.20*0.20*0.99, 0.20**2]
    ]   )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)                 # continued on next page
```

Python code for portfolio weights with *true* parameters (cont.)

Tangency portfolio weights with *true* parameters (cont.)

```python
# Mean-variance weights for gamma = 5
gamma = 5  # gamma represents the investor's risk aversion
w_mv = (1/gamma) * Vinverse @ (mu-Rf*ones)

# Weight in risk-free asset
w_f = 1 - ones.T @ w_mv

# Weights of tangency portfolio
w_Tangency = w_mv/(ones.T @ w_mv)

# Print portfolio weights
print("Mean-variance weights = ", "\n", w_mv)
print(" ")
print("Weight in risk-free asset = ","\n", w_f)
print(" ")
print("Tangency portfolio weights = ","\n", w_Tangency)
```

Solution: weights with *true* parameters

```
Mean-variance weights =
 [[0.20100503]
 [0.20100503]]

Weight in risk-free asset =
 [[0.59798995]]

Tangency portfolio weights =
 [[0.5]
 [0.5]]
```

- The above solution tells us that
  - we should invest about 60% in the risk-free asset;
  - we should invest about 20% in each of the two risky assets;
  - thus, the tangency portfolio has 50% invested in the two risky assets.

Example with estimation error

- Because the two assets are identical, the optimal tangency weights for the two assets are 50%.
- Suppose now that the mean excess return on the first asset is not known and is estimated (with error) to be
  - 1.10%
  - instead of 1.09%.
- There is no other change to the setting of the example.
- Find the optimal weights of the mean-variance tangency portfolios.

Python code for portfolios with *estimated* parameters

Tangency portfolio weights with *estimated* parameters

```
# Import numpy
import numpy as np

# Define a column vector of ones
ones = np.ones((2, 1))

# Define the gross risk-free rate of returns
Rf = 1.01

# Column vector of gross mean returns
# We change the mean return for the first asset from 1.09 to 1.10
mu = np.array([[1.10], [1.09]])

# Variance-covariance matrix
V = np.array(   [
    [0.20**2, 0.20*0.20*0.99],
    [0.20*0.20*0.99, 0.20**2]
    ]   )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)                      # continued on next page
```

Python code for portfolios with *estimated* parameters (cont.)

Tangency portfolio weights with *estimated* parameters (cont.)

```
# Mean-variance weights for gamma = 5
gamma = 5
w_mv = (1/gamma) * Vinverse @ (mu-Rf*ones)

# Weight in risk-free asset
w_f = 1 - ones.T @ w_mv

# Weights of tangency portfolio
w_Tangency = w_mv/(ones.T @ w_mv)

# Print portfolio weights
print("Mean-variance weights = ", "\n", w_mv)
print(" ")
print("Weight in risk-free asset = ","\n", w_f)
print(" ")
print("Tangency portfolio weights = ","\n", w_Tangency)
```

Solution for example with *estimated* parameters

```
Mean-variance weights =
 [[ 2.71356784]
 [-2.28643216]]

Tangency portfolio weights =
 [[ 6.35294118]
 [-5.35294118]]
```

- The above solution tells us that
  - we should invest about $+271\%$ in the first risky asset;
  - we should invest about -228% in the second risky asset;
  - thus, the tangency portfolio has
    - $+635\%$ invested in the first risk asset, and
    - -535% invested in the second risk asset.

Compare solutions with *estimated* and *true* parameters

- Comparing the above solution,
  - which uses estimated parameters,
  - to the solution which uses the true parameters,
- we see that estimation error causes a large change in the weights.

Solution with estimated parameters

```
Mean-variance weights =
 [[ 2.71356784]
 [-2.28643216]]

Tangency portfolio weights =
 [[ 6.35294118]
 [-5.35294118]]
```

Solution with true parameters

```
Mean-variance weights =
 [[0.20100503]
 [0.20100503]]

Tangency portfolio weights =
 [[0.5]
 [0.5]]
```

Magnitude of effect on weights of estimation error . . . I

- Now the mean-variance-optimal weights are:
  - $+271\%$ in the first asset, and
  - -228% in the second asset.
- and the tangency portfolio weights are:
  - $+635\%$ in the first asset, and
  - -535% in the second asset.
- That is, by taking extreme long and short positions in the two assets, the optimized portfolio
  - earns the difference in expected returns, 1.10 vs. 1.09, and also
  - reduces risk. (Because the return correlation is 0.99, going long one asset and short the other leads to a portfolio with low risk.)
- However, the optimization ignores that these differences may be the result of estimation error.

Magnitude of effect on weights of estimation error . . . II

- Compared to the weights in our example, the weights from mean-variance optimization are even more extreme
  - when using actual data, and
  - when there are more than just two assets.
- The estimation error can be so large that
  - the out-of-sample Sharpe ratio is negative;
  - i.e., investing all your money in a risk-free asset would be better!
- We see this evidence on the next few slides.

End of focus

Road map

Performance of mean-variance optimal portfolios

- In this section, we study the evidence on the performance of sample-based mean-variance optimal portfolios.
  - We first describe the methodology for evaluating performance out of sample.
  - Then we describe the datasets used to evaluate performance.
  - Finally, we report the performance metrics; in particular,
    - the portfolio Sharpe ratio
    - the portfolio turnover.

Start of focus

Road map

Methodology to evaluate out-of-sample performance

1. Choose a window $T^{\text{est}} < T$ over which to estimate the parameters. (If using monthly data, usually $T^{\text{est}} = 60$ months or 120 months.)

2. Estimate the parameters for the model.

3. Find optimal portfolio weights using estimated parameters.

4. Measure the return from holding these portfolio weights over the next period, that is, out-of-sample.

5. Repeat this “rolling-window" procedure until the end of the data set.

6. Calculate performance metrics using out-of-sample returns. (The performance metrics will include Sharpe ratio, turnover, etc.)

Measuring out-of-sample portfolio performance . . . I

- On the next few slides, we explain in detail how to measure out-of-sample portfolio performance.
- Assets.
  - Suppose that there are N risky assets with returns $R_t \in \mathbb{R}^N$, and a risk-free asset with monthly rate R_f.
- Rolling estimation of mean and covariance matrix using 60 monthly observations.
  - At month $t = 60, 61, \ldots, T-1$, estimate with the previous $T^{\text{est}} = 60$ months:

$$\hat{\mu}_t = \frac{1}{60} \sum_{s=t-59}^{t} R_s, \qquad \hat{V}_t = \frac{1}{60-1} \sum_{s=t-59}^{t} (R_s - \hat{\mu}_t)(R_s - \hat{\mu}_t)^\top.$$

Measuring out-of-sample portfolio performance . . . II

- Weights of portfolio strategy; e.g., mean-variance optimal portfolios.

$$w_t^{\mathrm{MV}} = \frac{1}{\gamma}\,\hat{V}_t^{\,-1}\big(\hat{\mu}_t - R_f 1_N\big), \qquad w_t^{\mathrm{Rf,MV}} = 1 - 1_N^\top w_t^{\mathrm{MV}}.$$

- Next-month return on portfolio with weights w_t^{MV}.

$$R_{p,t+1}^{\mathrm{MV}} = \big(1 - 1_N^\top w_t^{\mathrm{MV}}\big)\, R_f + \big(w_t^{\mathrm{MV}}\big)^\top R_{t+1}.$$

Measuring out-of-sample portfolio performance . . . III

- Weights in the portfolio of only risk assets (tangency portfolio).

$$w_t^{\text{tang}} = \frac{w_t^{\text{MV}}}{1_N^\top w_t^{\text{MV}}}$$

- Next-month's return on portfolio chosen this month, w_t^{tang}.

$$R_{p,t+1}^{\text{tang}} = (w_t^{\text{tang}})^\top R_{t+1}.$$

Measuring out-of-sample portfolio performance . . . IV

- Suppose we want to compare the performance of the mean-variance portfolio to the equal-weighted (EW) portfolio.

- Weights in the equal-weighted (EW) portfolio (of only risk assets).

$$w_t^{\text{EW}} = \frac{1}{N} 1_N = \frac{1}{N} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

- Next-month's return on the equal-weighted portfolio.

$$R_{p,t+1}^{\text{EW}} = (w_t^{\text{EW}})^\top R_{t+1}.$$

Measuring out-of-sample portfolio performance . . . V

- Rolling-window
  - At each date, drop the oldest observation
  - Add the newest observation
  - Recompute at the new date $\hat{\mu}_{t+1}$ and $\hat{V}_{t+1}$.
  - Recompute at the new date w_{t+1}^{MV} and w_{t+1}^{EW}.
- In this manner, compute all the out-of-sample returns until T.

Measuring out-of-sample portfolio performance . . . VI

- The resulting time-series of out-of-sample returns can be used to compute
  - Average return
  - Volatility
  - Sharpe ratio
  - Drawdowns
  - Turnover
- Then, we can compare different portfolio strategies.
  - This procedure almost all *out-of-sample backtests* in modern empirical asset management.
- On the next slide, we show how to compare Sharpe ratios.

Measuring out-of-sample portfolio performance . . . VII

- Sharpe ratio for the out-of-sample returns R^{MV}

$$\mathrm{SR}^{\mathrm{MV}} = \frac{12}{\sqrt{12}} \frac{\mathbb{E}[R^{\mathrm{MV}} - R_f]}{\sigma[R^{\mathrm{MV}} - R_f]} \quad \ldots (\text{annualized}).$$

- Sharpe ratio for the out-of-sample returns R^{EW}

$$\mathrm{SR}^{\mathrm{EW}} = \frac{12}{\sqrt{12}} \frac{\mathbb{E}[R^{\mathrm{EW}} - R_f]}{\sigma[R^{\mathrm{EW}} - R_f]} \quad \ldots (\text{annualized}).$$

- Then, use either the Jobson and Korkie (1981) formula or a bootstrap test for

$$H_0: \quad \mathrm{SR}^{\mathrm{MV}} = \mathrm{SR}^{\mathrm{EW}}; \qquad H_1: \quad \mathrm{SR}^{\mathrm{MV}} \neq \mathrm{SR}^{\mathrm{EW}}.$$

End of focus

Road map

Datasets used to evaluate performance

- Datasets considered in DeMiguel, Garlappi, and Uppal (2009)

#	Dataset and source	Acronym	N	Time Period
1	Ten sector portfolios of the S&P500 and MKT Source: Roberto Wessels	S&P Sectors	10+1	1981–2002
2	Ten industry portfolios and MKT Source: Ken French's website	Industry Portf.	10+1	1963–2004
3	Eight country indexes and World Mkt Source: MSCI	Intern'l Portf.	8+1	1970–2001
4	SMB and HML portfolios and MKT Source: Ken French's website	Mkt+ SMB/HML	2+1	1963–2004
5	Twenty Size & Book-to-Market portfolios, MKT Source: Ken French's website	FF 1-factor	20+1	1963–2004
6	Twenty Size and Book-to-Market portfolios and the MKT, SMB, HML & UMD portfolios Source: Ken French's website	FF 4-factors	20+4	1963–2004

Road map

Sharpe ratio of mean-variance optimal portfolios

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
MV (in sample)	*0.3848*	*0.2124*	*0.2090*	*0.2851*	*0.5098*	*0.5364*
MV (out of sample)	0.0794	−0.0363	−0.0719	0.2186	−0.0684	−0.0031
$1/N$	**0.1876**	**0.1353**	**0.1277**	**0.2240**	**0.1623**	**0.1753**
p-value	(0.12)	(0.01)	(0.01)	(0.46)	(0.00)	(0.01)

1. Out-of-sample performance of mean-variance portfolio substantially worse than in sample.
    - Out-of-sample Sharpe ratio of mean-variance portfolio is negative for 4 out of the 6 datasets.
2. Out of sample, even the simple $1/N$ strategy performs much better than the mean-variance optimal strategy for all six datasets.

Road map

Turnover of mean-variance optimal portfolios

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
Turnover *relative* to $1/N$ strategy						
MV (out of sample)	38.99	607479.61	4236.10	2.83	9408.16	2672.71

- The above table shows that the turnover of the mean-variance portfolio is
  - 38.99 times that of the $1/N$ portfolio for the S&P sectors dataset;
  - 4236.10 times that of the $1/N$ portfolio for Inter'l portf. dataset;
  - and so on ...
- Out-of-sample turnover of sample-based mean-variance portfolio substantially higher than that of $1/N$ portfolio.

Conclusion about the performance of sample-based mean-variance optimal portfolios

- Out-of-sample performance of mean-variance portfolios is very poor compared to their in-sample performance;
- Mean-variance portfolios fail to outperform simple $1/N$ portfolio:
  - mean-variance portfolios have a lower Sharpe ratio out of sample;
  - mean-variance portfolios have a higher turnover.

Road map

Mean-variance models that adjust for estimation error . . . I

- The poor performance of the Markowitz (1952, 1959) model was recognized as early as the 1970s.
- Considerable research effort has been devoted to dealing with estimation error.
- Some of the approaches proposed to deal with estimation error are listed below.
  - One set of approaches is listed under the title "Bayesian models"
  - Another set is listed under the title "Non-Bayesian models"
- We will study many (but not all) of these models.

Mean-variance models that adjust for estimation error . . . II

1. Bayesian models for dealing with estimation error
   - Purely statistical approach relying on diffuse-priors:
     Barry (1974) and Bawa, Brown, and Klein (1979).
   - Shrinkage estimators:
     Jobson, Korkie, and Ratti (1979), Jobson and Korkie (1980), and Jorion (1985, 1986).
   - Methods to reduce error in estimating the covariance matrix:
     Best and Grauer (1992), Chan, Karceski, and Lakonishok (1999), Ledoit and Wolf (2004a, 2004b).
   - Black-Litterman:
     Black and Litterman (1990, 1992) combine two sets of "priors"—one based on an equilibrium asset-pricing model, the other on subjective views of investor.

Mean-variance models that adjust for estimation error . . . III

2. Non-Bayesian models for dealing with estimation error

- Shortselling constraints:
  Frost and Savarino (1988), Chopra (1993), and Jagannathan and Ma (2003).
- "Robust" portfolio optimization:
  Goldfarb and Iyengar (2003), Uppal and Wang (2003), Garlappi, Uppal, and Wang (2007), Boyle, Garlappi, Uppal, and Wang (2012).
- Optimally diversify across market and estimation risk:
  Kan and Zhou (2005).
- Resampling methods:
  Developed by Michaud (1998); discussed in Scherer (2002) and Harvey, Liechty, Liechty, and Müller (2003).

Shrinkage: Key idea for dealing with estimation error . . . I

- The central idea for dealing with estimation error is shrinkage.
- In the rest of the course, we will be studying several methods for dealing with estimation error, but (almost)
- all of these methods can be interpreted in terms of shrinking the estimated $\mathbb{E}[R]$ and $\mathbb{V}[R]$ towards a reasonable quantity.

Shrinkage: Key idea for dealing with estimation error . . . II

- For example, if you are not sure of your estimate of the expected return on a particular risky asset,
  - then one way to reduce estimation error is to shrink the estimated expected return of this asset toward a sensible value, such as
    - the average of all expected returns, or
    - the expected return on the market portfolio.
- If you are not sure of your estimate of the return covariance matrix, then one way to reduce estimation error is to shrink the estimated covariance matrix toward a sensible matrix, such as
  - the covariance matrix in a one-factor (market) model;
  - the covariance (or correlation) matrix where all covariances (or correlations) are equal.

Central idea that links everything we study in the course

Shrinkage

is a central idea for dealing with estimation error.

Road map

Bayesian approaches to dealing with estimation error

- Under the Bayesian approach, the estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$ are computed using the predictive (posterior) distribution of asset returns.
- This distribution is obtained by
  - integrating the conditional likelihood, $f\left(R \middle| \mathbb{E}[R], \mathbb{V}[R]\right)$, over $\mathbb{E}[R]$ and $\mathbb{V}[R]$,
  - with respect to a certain subjective prior, $p(\mathbb{E}[R], \mathbb{V}[R])$.
- In the literature, the Bayesian approach to estimation error has been implemented in different ways.
- Below, we describe three common implementations.

Approach#1: Bayesian diffuse-prior portfolio

- Barry (1974), Klein and Bawa (1976), and Brown (1979), show that
  - if prior is chosen to be diffuse, that is, $p(\mu, V) \propto |V|^{-(N+1)/2}$, and
  - the conditional likelihood is given by a Normal distribution,
  - then the predictive distribution is a Student-t with
    - mean equal to $\hat{\mu}$, and
    - variance equal to $\hat{V}\,(1 + 1/T^{\text{est}})$, where T^{est} is the number of data points used in the estimation.
- Hence, while still using historical mean to estimate expected returns, this approach increases the covariance matrix by $(1 + 1/T^{\text{est}})$.
- For a long estimation window (e.g., $T^{\text{est}} = 60$ months), the effect of this correction is negligible: $1/60 = 0.0167 < 2\%$.
- Thus, the performance of the Bayesian diffuse-prior portfolio is very similar to that of the sample-based mean-variance portfolio.

Approach#2: Bayes-Stein shrinkage portfolio . . . I

- The Bayes-Stein ("BS") portfolio is an application of shrinkage estimation pioneered by Stein (1956) and James and Stein (1961).

- It is designed to handle the error in estimating expected returns by using estimators of the form:

$$\hat{\mu}_t^{BS} = (1 - \hat{\psi}_t)\,\hat{\mu}_t + \hat{\psi}_t\,\bar{\mu}_t,$$

$$\hat{\psi}_t = \frac{N+2}{(N+2) + T^{\text{est}}(\hat{\mu}_t - \mu_t^{\text{GMV}})^\top \hat{V}_t^{-1}(\hat{\mu}_t - \mu_t^{\text{GMV}})},$$

  in which

  - the shrinkage factor $0 < \hat{\psi}_t < 1$,
  - $\hat{V}_t = \frac{1}{T^{\text{est}} - N - 2} \sum_{s=t-T^{\text{est}}+1}^{t} (R_s - \hat{\mu}_t)(R_s - \hat{\mu}_t)^\top$, and
  - $\bar{\mu}_t$ is the "grand mean" (given by your prior).

Approach#2: Bayes-Stein shrinkage portfolio . . . II

- These estimators "shrink" the sample mean toward a common "grand mean," $\overline{\mu}$; that is,

$$\hat{\mu}_t^{BS} = (1 - \hat{\psi}_t)\,\hat{\mu}_t + \hat{\psi}_t\,\bar{\mu}_t,$$

  where $\hat{\psi}_t$ is the weight on the "grand mean."

- For example, Jorion (1985, 1986), takes the grand mean, $\overline{\mu}$, to be the mean of the minimum-variance portfolio, μ^{GMV}.
  - In addition to shrinking the estimate of the mean, Jorion also accounts for estimation error in the covariance matrix via traditional Bayesian-estimation methods.

- Again, the error in estimating $\hat{\mu}_t$ is so large, that it is optimal to put a lot of weight (around 95%) on the "grand mean," $\bar{\mu}_t$.

Approach#3: Bayesian portfolio based on belief in an asset-pricing model . . . I

- Under the Bayesian “Data-and-Model" (“DM”) approach developed in Pastor (2000) and Pástor and Stambaugh (2000):
  - The shrinkage target depends on the investor’s prior belief in a particular asset-pricing model.
  - The degree of shrinkage is determined by the variability of the prior belief relative to the information contained in the data.
- These portfolios are a further refinement of shrinkage portfolios because they address the arbitrariness of the choice
  - of a shrinkage target, $\overline{\mu}$, and
  - of the shrinkage factor, ψ,

  by using investor’s belief about the validity of an asset-pricing model.

Approach#3: Bayesian portfolio based on belief in an asset-pricing model . . . II

- For example, one could implement the "Data-and-Model" approach using different asset-pricing models:
  - the Capital Asset Pricing Model (CAPM),
  - the Fama and French (1993) three-factor model, and
  - the Fama and French (1993) + Carhart (1997) four-factor model.
- The Bayesian investor's belief in the asset-pricing model is captured by a prior about the extent of mispricing.
  - Let the variable α reflect this mispricing.
  - Assume the prior to be normally distributed around $\alpha = 0$, and with the benchmark value of its tightness being $\sigma_\alpha = 1\%$ per annum.
  - Intuitively, this implies that the investor believes with 95% probability the mispricing is approximately between -2% and $+2\%$ p.a.

Performance of Bayesian models

- On the next few pages, we compare the out-of-sample performance of these Bayesian models to the $1/N$ portfolio.
- We study the performance across six datasets.
- Performance is measured by the Sharpe ratio.
- We also report turnover, which provides a sense of the transaction costs associated with implementing the portfolio strategy.
- For each dataset, the best-performing strategy is highlighted in **bold**.

Road map

Sharpe ratios of Bayesian portfolios . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
MV (in sample)	0.3848	0.2124	0.2090	0.2851	0.5098	0.5364
MV (out of sample)	0.0794	−0.0363	−0.0719	0.2186	−0.0684	−0.0031
$1/N$	**0.1876**	**0.1353**	**0.1277**	0.2240	**0.1623**	0.1753
	(0.12)	(0.01)	(0.01)	(0.46)	(0.00)	(0.01)
Bayesian models						
Bayes-Stein	0.0811	−0.0319	−0.0528	**0.2536**	−0.0636	−0.0042
	(0.09)	(0.01)	(0.01)	(0.25)	(0.00)	(0.01)
Data-Model ($\sigma_\alpha = 1\%$)	0.1410	−0.0475	0.1032	0.0181	0.0745	**0.2355**
	(0.12)	(0.00)	(0.35)	(0.00)	(0.06)	(0.25)

Sharpe ratios of Bayesian portfolios . . . II

- Out-of-sample performance Sharpe ratios of portfolios based on Bayesian shrinkage is not very different from that of mean-variance portfolios.
  - Out-of-sample Sharpe ratio of portfolios based on Bayesian shrinkage is negative for 4 out of the 6 datasets.
  - The Data-and-Model approach performs better than the "BS" approach, but only for 1 out of the 6 datasets does it outperform the $1/N$ strategy.
- Conclusion:
  Out-of-sample, the Sharpe ratio of the two Bayesian strategies fail to outperform the $1/N$ portfolio.

Turnover of mean-variance optimal portfolios . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
Turnover *relative* to $1/N$ strategy						
MV (out of sample)	38.99	607479.61	4236.10	2.83	9408.16	2672.71
Bayesian models						
Bayes-Stein	22.41	10092.20	1760.67	1.85	10510.09	2783.58
Data-Model ($\sigma_\alpha = 1\%$)	1.72	20468.41	45.24	76.85	762.83	32.46

Turnover of mean-variance optimal portfolios ... II

- In terms of turnover,
  - Both Bayesian strategies have better (smaller) turnover than the mean-variance portfolio,
  - Turnover of Data-and-Model approach is typically better (smaller) than that of the Bayes-Stein.
  - However, both Bayesian strategies have substantially higher turnover than the $1/N$ portfolio.

Conclusion about performance of Bayesian portfolios

- Out-of-sample performance of portfolios based on Bayesian shrinkage is not very different from that of mean-variance portfolios.
  - Out-of-sample Sharpe ratio of portfolios based on Bayesian shrinkage is negative for 4 out of the 6 datasets.
  - The Data-and-Model approach Sharpe ratio is better than the Bayes-Stein approach, but only for 1 out of the 6 datasets does it outperform the $1/N$ strategy.
  - The turnover of the Bayesian portfolios is better than that of the mean-variance model but still much higher (worse) than for $1/N$.
- Conclusion:
  Out-of-sample, the $1/N$ strategy usually performs better than the two Bayesian strategies.

Start of focus

Road map

Effect of constraints on short-sellling . . . I

- In this section, we consider short-sale constraints.
- Recall the solution for the example with estimation error:
  - $+271\%$ weight in the first asset, and
  - -228% weight in the second asset.
- Imposing the constraint that the weights on the assets have to be non-negative, limits how extreme the weights can be.

Effect of constraints on short-sellling . . . II

- Thus, constraints are an ad-hoc way to limit the effect of estimation error;
  - this approach is popular with practitioners;
  - short-sales may also be constrained by regulations (e.g., pension funds);
  - mutual funds may find it prohibitively costly to go short.
- But, one can also show formally (mathematically) that short-sale constraints lead to shrinkage.

Shortsale-constrained portfolios

- We consider a number of strategies that constrain short-selling:
- These are:
  - Sample-based mean-variance-constrained (MV-C),
  - Bayes-Stein-constrained (BS-C), and
  - Global-minimum-variance-constrained (GMV-C) (we will study this one in the next section.)
- These policies are obtained by
  - imposing an additional non-negativity constraint
  - on the portfolio weights in the corresponding optimization problems.

Mean-variance portfolio with short-sale constraints

- In the last chapter, we saw that one can obtain the optimal (unconstrained) weights by maximizing mean-variance utility (MVU):

$$\max_{w} \quad \text{MVU} = \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w.$$

- To constrain short selling, i.e., $w_n \geq 0$, $n = \{1, \ldots, N\}$, one can rewrite the above problem by adding a short-sale constraint:

$$\max_{w} \quad \text{MVU} = \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w$$

subject to the short-sale constraint:

$$w_n \geq 0, \text{ for all } n = \{1, \ldots, N\}.$$

Showing that short-sale constraints lead to shrinkage . . . I

- To interpret the effect of short-sale constraints, observe that
- imposing the constraint $w_n \geq 0$, $i = \{1, \ldots, N\}$ in the basic mean-variance optimization yields the following Lagrangian,

$$\begin{aligned}\mathcal{L} &= \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w + (\lambda_1 w_1 + \lambda_2 w_2 + \ldots + \lambda_N w_N) \\ &= \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w + \underbrace{w^\top \lambda_{ss}}_{\text{new term}},\end{aligned}$$

- where $\lambda_{ss} = \{\lambda_1, \ldots, \lambda_N\}$ is the $(N \times 1)$ vector of Lagrange multipliers for the N constraints on short-selling (ss); i.e., $w_n \geq 0$.

Showing that short-sale constraints lead to shrinkage . . . II

- We can now rearrange terms in the Lagrangian

$$\begin{aligned}\mathcal{L} &= \left[R_f + w^\top(\mu - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w + \underbrace{w^\top \lambda_{ss}}_{\text{new term}}, && \text{. . . old way of writing} \\ &= \left[R_f + w^\top(\mu + \underbrace{\lambda_{ss}}_{\substack{\text{new}\\ \text{term}}} - R_f 1_N)\right] - \frac{\gamma}{2} w^\top V w && \text{. . . new way of writing}\end{aligned}$$

- Thus, the constrained mean-variance weights are equivalent to the unconstrained weights but with the adjusted mean vector: $\mu + \lambda$.

Showing that short-sale constraints lead to shrinkage . . . III

- We start by rewriting the last equation on the previous slide:

$$\mathcal{L} = \left[R_f + w^\top \big(\underbrace{\mu + \lambda_{ss}}_{\text{adjusted mean}} - R_f 1_N \big)\right] - \frac{\gamma}{2} w^\top V w$$

- To understand why the above expression implies a shrinkage of expected returns:
  - Note that the short-sale constraint on asset n is likely to be binding when its expected return μ_n is low.
  - When the constraint for a particular asset n binds, $\lambda_n > 0$ and the expected return is increased from μ_n to $\mu_n + \lambda_n$.
  - Hence, imposing a short-sale constraint on the sample-based mean-variance problem is equivalent to "shrinking" the expected return toward the average.

Performance of short-sale-constrained portfolios

- On the next few pages, we compare the out-of-sample performance of these short-sale-constrained portfolios to the $1/N$ portfolio.
- We study the performance across six datasets.
- Performance is measured by the Sharpe ratio.
- We also report turnover, which provides a sense of the transaction costs associated with implementing the portfolio strategy.
- For each dataset, the best-performing strategy is highlighted in **bold**.

Sharpe ratios of short-sale-constrained portfolios . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
MV (in sample)	0.3848	0.2124	0.2090	0.2851	0.5098	0.5364
MV (out of sample)	0.0794	−0.0363	−0.0719	0.2186	−0.0684	−0.0031
$1/N$	**0.1876**	**0.1353**	**0.1277**	0.2240	0.1623	0.1753
	(0.12)	(0.01)	(0.01)	(0.46)	(0.00)	(0.01)
Without constraints						
Bayes-Stein	0.0811	−0.0319	−0.0528	**0.2536**	−0.0636	−0.0042
With short-sale constraints						
MV-Constrained	0.0892	0.0678	0.0848	0.1084	**0.1977**	**0.2024**
	(0.09)	(0.03)	(0.17)	(0.02)	(0.02)	(0.27)
BS-Constrained	0.1075	0.0819	0.0848	0.1514	**0.1955**	**0.2062**
	(0.14)	(0.06)	(0.15)	(0.09)	(0.03)	(0.25)

Sharpe ratios of short-sale-constrained portfolios . . . II

- Performance of mean-variance short-sale-constrained portfolios is better than that of the unconstrained mean-variance portfolio in 5 out of 6 datasets (only exception is for dataset with $N = 3$).
- Similarly, the Bayes-Stein model with constraints perform better than the unconstrained Bayes-Stein model in 5 out of 6 datasets (only exception is for dataset with $N = 3$).
- The Bayesian short-sale-constrained portfolios is better than the mean-variance (constrained) portfolio in 5 out of 6 cases.
- The $1/N$ strategy, however, performs better out-of-sample than the two short-sale constrained strategies in 4 out of the 6 datasets.

Turnover of shortsale-constrained portfolios . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
Turnover *relative* to $1/N$ strategy						
MV (out of sample)	38.99	607479.61	4236.10	2.83	9408.16	2672.71
Bayes-Stein	22.41	10092.20	1760.67	1.85	10510.09	2783.58
Models with short-sale constraints						
MV-Constrained	4.53	7.17	7.23	4.12	17.53	13.82
Bayes-Stein-Constrained	3.64	7.22	6.10	3.65	17.32	13.07

Turnover of shortsale-constrained portfolios . . . II

- In terms of turnover,
  - short-sale constraints are very successful in reducing turnover;
  - However, even with short-sale constraints, the turnover is still higher than that of the $1/N$ portfolio.

Conclusion about short-sale constrained portfolios

- Shortsale constraints are very effective at reducing turnover.
- Shortsale constraints are less successful at improving Sharpe ratios.
- Overall, short-sale-constrained portfolios fail to achieve a:
  - higher Sharpe ratios than the $1/N$ portfolio
  - lower turnover than the $1/N$ portfolio.

End of discussion
of models dealing with error in
estimating expected returns ($\mathbb{E}[R]$)

Next, we look at models focusing
on reducing the error in estimating
the return covariance matrix ($\mathbb{V}(R)$)

Road map

Global-minimum-variance (GMV) portfolios . . . I

- In the last chapter, we saw that one can obtain the mean-variance optimal (unconstrained) weights by minimizing:

$$\min_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad (1) \quad \sum_{n=1}^{N} w_n = w^\top 1_N = 1;$$

$$(2) \quad \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^\top \mu = \mu_{\text{targ}}.$$

Global-minimum-variance (GMV) portfolios . . . II

- But, as we saw in the previous section, portfolios that rely on sample estimates of the expected returns, perform poorly
  - even after adjusting the sample mean for estimation error (using Bayesian and non-Bayesian methods).
- For academic research papers showing this, see, e.g.,
  - DeMiguel, Garlappi, and Uppal (2009) and
  - Jacobs, Müller, and Weber (2014).
- So, one approach is to remove from the optimization problem the term that cannot be estimated precisely, $\mathbb{E}[R]$.

Global-minimum-variance (GMV) portfolios . . . III

- That is, instead of solving the following optimization problem

$$\begin{aligned}
\min_{\{w_1,\dots,w_N\}} \quad & \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w \\
\text{subject to:} \quad & (1) \quad \sum_{n=1}^{N} w_n = w^\top 1_N = 1; \\
& (2) \quad \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^\top \mu = \mu_{\mathbf{targ}}.
\end{aligned}$$

- One could assume that expected returns for all N assets are equal;
- Therefore, in the optimization problem $\mathbb{E}[R]$ can be ignored;
- Thus, constraint (2) can be removed from the optimization problem.

Global-minimum-variance (GMV) portfolios . . . IV

- This reduces the above optimization problem to the following:

$$\text{Minimize}_{\{w_1,\dots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad (1) \quad \sum_{n=1}^{N} w_n = w^\top 1_N = 1;$$

with the solution to this problem being

$$w = \frac{V^{-1} 1_N}{1_N^\top V^{-1} 1_N}$$

- where the numerator of w has the "raw" weights that are inversely related to their risk (so, more risky assets have a smaller weight);
- and the denominator of w has the sum of these weights, which then scales the "raw" weights in the numerator so that they add up to 1.

Global-minimum-variance (GMV) portfolios . . . V

- In this section, we study global-minimum-variance portfolios:
  - These portfolios have optimal weights based only on risk estimates.
  - So, these portfolios can be viewed as shrinking completely the expected returns on all assets to the same value.

Motivation for global-minimum-variance portfolios

- The global-minimum-variance portfolio does not depend on expected returns, which we now know cannot be estimated precisely.
- At the same time, the weights of the global-minimum-variance portfolio are optimized to have the lowest risk, which can be estimated precisely.
- So, possibly, these portfolios will perform well out of sample.

Variants of global-minimum-variance portfolios . . . I

1. Global minimum-variance portfolios
   Jorion (1985, 1986)
   No shrinkage of the variance-covariance matrix.

2. Minimum-variance portfolios with variance estimated by imposing a factor structure:
   Green and Hollifield (1992) and Chan, Karceski, and Lakonishok (1999).
   Shrink variance-covariance matrix from $(N \times N)$ to $(K \times K)$ matrix, where $K < N$ is the number of factors driving the N returns.

3. Minimum-variance portfolios with short-sale constraints
   Short-sale constraints are an indirect form of shrinkage of the variance-covariance matrix.

Variants of global-minimum-variance portfolios . . . II

4. Minimum-variance portfolios estimated using shrinkage
   Ledoit and Wolf (2004b, 2004a, 2017, 2022)
   Explicit shrinkage of the variance-covariance matrix.

5. Minimum-variance portfolios with norm constraints
   DeMiguel, Garlappi, Nogales, and Uppal (2009)
   Generalized shrinkage of the variance-covariance matrix.

- Below,
  - we will discuss Approaches 1, 3, 4, and 5;
  - Approach 2 (based on factor models of returns), performs less well than Approaches 3, 4, and 5 (see Jagannathan and Ma 2003), and we will study factor models later.

Weights of the global-minimum-variance portfolio

- Recall that to obtain the global-minimum-variance portfolio weights, the optimization problem to be solved is:

$$\begin{aligned} \min_{w} \quad & w^\top V w && \ldots \text{portfolio variance to be minimized} \\ \text{s.t.} \quad & w^\top 1_N = 1 && \ldots \text{adding-up constraint} \end{aligned}$$

- The solution, available in closed form, is:

$$w = \frac{V^{-1} 1_N}{1_N^\top V^{-1} 1_N}.$$

Interpreting the weights of the GMV . . . I

- To understand the minimum-variance portfolio weights, we study the special case with zero correlation between the returns of all assets.
- That is, instead of the general covariance matrix

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix},$$

- we consider the special case of a diagonal matrix

$$V = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix}.$$

Interpreting the weights of the GMV ...II

- For the special case where V is a diagonal matrix, its inverse is

$$V^{-1} = \begin{bmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sigma_N^2 \end{bmatrix}.$$

- For this special case, one can show that the optimal portfolio weights of the GMV portfolio, $w = \frac{V^{-1} 1_N}{1_N^\top V^{-1} 1_N}$, simplify to:

$$w_n = \frac{\frac{1}{\sigma_n^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_n^2}},$$

which shows that there is an inverse relation between the variance of an asset and its weight in the portfolio.

When will the GMV weights be reasonable? . . . I

- The GMV portfolio is given, in general, by the expression

$$w = \frac{V^{-1}\,1_N}{1_N^\top\,V^{-1}\,1_N}.$$

- The GMV portfolio weights will be reasonable if the elements of the covariance matrix V can be estimated precisely, i.e., if return variances and covariances can be estimated precisely.
  - Note that we have already agreed that elements of the covariance matrix can be estimated more precisely than expected excess returns.

When will the GMV weights be reasonable? . . . II

- There is a great deal of empirical evidence that realized volatility, i.e., volatility over the recent past (e.g., the last 30 days) is a good estimator of future volatility;
  - see, for instance, Ghysels, Santa-Clara, and Valkanov (2006) and Andersen and Benzoni (2010).
- It is also possible to improve estimates of variances and covariances by using implied volatility and correlations from option prices.
  - See Driessen, Maenhout, and Vilkov (2009).
- DeMiguel, Plyakha, Uppal, and Vilkov (2013) show how to use the information in option prices to form portfolios.
- Python code to compute option-implied volatility is available from this link to an article by Roi Polanitzer on Medium.

When will the GMV weights be reasonable? ...III

- Finally, when we have a large number of assets, then
  - even if some elements of the variance-covariance matrix V are estimated precisely,
  - the matrix V^{-1} may be ill-conditioned.
- In this case, estimation of V^{-1} can be improved, as we show below, by
  1. Imposing short-sale constraints, or
  2. Using shrinkage to estimate V.

When will the GMV earn high Sharpe ratios? ...I

- Recall that for the special case of zero return correlations,

$$w_n = \frac{\frac{1}{\sigma_n^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_n^2}}.$$

- From the above expression, we see that our portfolio will assign high weights to assets with low volatilities.

- Thus, the minimum-variance portfolio will earn a high Sharpe ratio if low-volatility assets earn a high return.

When will the GMV earn high Sharpe ratios? ...II

- For the general case of nonzero return correlation
  - Scherer (2011) for the case of unconstrained portfolios (assuming a CAPM world), and
  - Clarke, Silva, and Thorley (2010) for short-sale constrained portfolios (assuming a single-factor covariance matrix)

  show that the weight of the minimum-variance portfolio are inversely related to the volatility and beta of each asset.

- Thus, the minimum-variance portfolio will earn a high Sharpe ratio if assets with low volatility and low beta earn a high return.

When will the GMV earn high Sharpe ratios? ...III

- Dangl and Kashofer (2013), in a very nice paper, discuss
  - the empirical performance of minimum-variance portfolios,
  - how this performance has evolved over time, and
  - how this performance is related to the results on the returns of low-volatility and low-beta stocks.
- Plyakha, Uppal, and Vilkov (2021) discuss the literature that studies the relation between low-volatility stocks and their returns.
  - This discussion highlights that the results about the relation between volatility and returns is extremely sensitive to
  - how one measures return, and
  - how one measures volatility.

Example of global-minimum-variance portfolio

- Suppose that:
  - the investor's risk aversion is $\gamma = 5$;
  - p.a. gross risk-free rate is 1.01;
  - p.a. gross mean returns for the first asset is, 1.10;
  - p.a. gross mean returns for the second asset is, 1.09;
  - p.a. volatility of returns for both assets are the same, 20%; and
  - the return correlation of the two risky assets is 0.99.
- Find the optimal weights of the global-minimum-variance portfolio.

Python code for global-minimum-variance portfolio

Code for finding global-minimum-variance portfolio

```python
# Import numpy
import numpy as np

# Column vector of ones
ones = np.ones((2, 1))

# Variance-covariance matrix
V = np.array(   [
    [0.20**2, 0.20*0.20*0.99],
    [0.20*0.20*0.99, 0.20**2]
    ]   )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)

# Weights of the global minimum variance (GMV) portfolio
w_gmv = (Vinverse @ ones)/(ones.T @ Vinverse @ ones)

print("\nWeights of the global minimum variance (GMV) portfolio = ")
print(w_gmv)
```

Solution for global-minimum-variance portfolio

```
Weights of the global minimum variance (GMV) portfolio =
[[0.5]
 [0.5]]
```

- We see from the above solution that
  - The optimal weights are exactly the weights we would get if we knew the true parameters;
  - That is, these weights are independent of mean returns, which are estimated with error; in particular, the mean return for the first asset, $\mathbb{E}[R_1] = 1.10$.
- Thus, it is clear from this example that the GMV is not affected by the error in estimating expected returns.

Second example of GMV portfolio for $N = 3$

- Find the global-minimum-variance portfolio if the covariance matrix of returns is the following diagonal matrix (i.e., special case of zero correlations):

$$V = \begin{bmatrix} 0.0625 & 0.0 & 0.0 \\ 0.0 & 0.1225 & 0.0 \\ 0.0 & 0.0 & 0.3600 \end{bmatrix}.$$

Python code for GMV portfolio with $N = 3$

Code for finding GMV portfolio with $N = 3$

```
# Import numpy
import numpy as np

# Column vector of ones
ones = np.ones((3, 1))

# Variance-covariance matrix
V = np.array( [
    [0.0625, 0.0, 0.0],
    [0.0, 0.1225, 0.0],
    [0.0, 0.0, 0.3600]  ] )

# Inverse of the variance-covariance matrix using the linear-algebra
     command "inv"
Vinverse = np.linalg.inv(V)

# Weights of the global minimum variance (GMV) portfolio
w_gmv = (Vinverse @ ones)/(ones.T @ Vinverse @ ones)

print("Weights of the global minimum variance (GMV) portfolio = ")
print(w_gmv)
```

Solution for GMV portfolio with $N = 3 \ldots$ I

```
Weights of the global minimum variance (GMV) portfolio =
[[0.5938894]
 [0.3030048]
 [0.1031058]]
```

- On the next slide, we verify, using Python, that the weights satisfy:

$$w_n = \frac{\frac{1}{\sigma_n^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_n^2}}.$$

Solution for GMV portfolio with $N = 3 \ldots$ II

Verification of weights

```
(
(1/0.0625)/(1/0.0625 + 1/0.1225 + 1/0.3600),
(1/0.1225)/(1/0.0625 + 1/0.1225 + 1/0.3600),
(1/0.3600)/(1/0.0625 + 1/0.1225 + 1/0.3600)
)
```

```
(0.5938894,  0.3030048,  0.1031058)
```

- You can also verify the weights with just a single line of Python code

Verification of weights – single line of code

```
# Numerator   =  diagonal of the inverse of V
# Denominator =  sum of the diagonal of the inverse of V
print(np.diag(Vinverse)/np.sum(np.diag(Vinverse)))
```

```
[0.5938894   0.3030048   0.1031058]
```

Solving problem of ill-conditioned covariance matrix

- Recall that the optimal mean-variance portfolio,

$$w = \frac{1}{\gamma} V^{-1}(\mathbb{E}[R] - R_f 1_N).$$

- is plagued by two problems:
  1. error in estimating expected returns, $\mathbb{E}[R]$;
  2. ill-conditioned variance-covariance matrix, V.

- The GMV portfolio, given by,

$$w = \frac{V^{-1}1_N}{1_N^\top V^{-1} 1_N}$$

  does not depend on expected returns, $\mathbb{E}[R]$, and so is free of the error in estimating expected returns.

- So, now we need to address the second problem of an ill-conditioned variance-covariance matrix, V.

Improving ill-conditioned covariance matrix . . . I

- There are several approaches for improving the behavior of the covariance matrix, V.
- We will discuss three approaches:
  1. Imposing a short-sale constraint on the portfolio weights (Jagannathan and Ma 2003);
  2. Shrinking the estimated covariance matrix toward a target matrix (Ledoit and Wolf 2003, 2004a, 2004b, 2017, 2022).
  3. Imposing a norm constraint, a more general version of the short-sale constraint (DeMiguel, Garlappi, Nogales, and Uppal 2009).

Improving ill-conditioned covariance matrix . . . II

- It turns out that all three approaches can be interpreted as applying shrinkage or regularization to the covariance matrix.
  - DeMiguel, Garlappi, Nogales, and Uppal (2009) show this formally.
- For quantitative asset management, most people use the Ledoit-Wolf approach.
- For investment products, Amundi sells an ETF based on the approach in DeMiguel, Garlappi, Nogales, and Uppal (2009).

Road map

Approach 1: GMV with short-sale constraints

- Jagannathan and Ma (2003) study the effect of imposing short-sale constraints on the global minimum-variance portfolio.
- Shortsales-constrained global minimum-variance portfolio is

$$\begin{aligned} \min_{w} \quad & w^\top \hat{V} w && \ldots \text{ portfolio variance} \\ \text{s.t.} \quad & w^\top 1_N = 1 && \ldots \text{ adding-up constraint} \\ & w \geq 0 && \ldots \text{ long-only constraint that prohibits short selling} \end{aligned}$$

- We see on the next page why the short-sale constraint, again,
    - can be interpreted as inducing shrinkage,
    - but this time of the covariance matrix.

Shrinkage interpretation of short-sales-constrained GMV portfolio . . . I

- Shortsales-constrained global minimum-variance portfolio is

$$\begin{aligned}
\min_{w} \quad & w^\top \hat{V} w && \ldots \quad \text{portfolio variance} \\
\text{s.t.} \quad & w^\top 1_N = 1 && \ldots \quad \text{adding-up constraint} \\
& w \geq 0 && \ldots \quad \text{long-only constraint that prohibits short selling}
\end{aligned}$$

- The solution coincides with unconstrained global-minimum-variance portfolio

$$\begin{aligned}
\min_{w} \quad & w^\top \hat{V}_{JM} w && \ldots \quad \text{portfolio variance} \\
\text{s.t.} \quad & w^\top 1_N = 1 && \ldots \quad \text{adding-up constraint}
\end{aligned}$$

- where the sample covariance matrix $\hat{V}$ has been replaced by $\hat{V}_{JM}$, the cov matrix *after shrinkage*:

$$\hat{V}_{JM} = \hat{V} - \lambda_{ss} 1_N^\top - 1_N \lambda_{ss}^\top .$$

Shrinkage interpretation of short-sales-constrained GMV portfolio . . . II

- We start by rewriting $\hat{V}_{JM}$:

$$\underbrace{\hat{V}_{JM}}_{N\times N} = \underbrace{\hat{V}}_{N\times N} - \underbrace{\lambda_{ss}\, 1_N^\top}_{N\times N} - \underbrace{1_N \lambda_{ss}^\top}_{N\times N}.$$

- Note that elements of the vector of Lagrange multipliers can take only non-negative values, i.e., $\lambda_{ss} \geq 0_N$:
  - $\lambda_{ss,n} = 0$ if $w_n > 0$
  - $\lambda_{ss,n} > 0$ if $w_n \leq 0$.

- Therefore, the matrix $\hat{V}_{JM}$ may be interpreted as the sample covariance matrix after shrinkage, because
  - if the short-sale constraint for the n^{th} asset is binding ($\lambda_{ss,n} > 0$)
  - then, the sample covariance of this asset with any other asset is reduced by $\lambda_{ss,n}$.

Example of shrinkage interpretation with $N = 3$ . . . I

- Suppose that there $N = 3$ risky assets, and
- only the third risky asset in the GMV portfolio has a negative weight.
- This means that $\lambda_1 = 0$, $\lambda_2 = 0$, and only $\lambda_3 > 0$.
- Then, the short-sale constrained GMV can be obtained from

$$
\begin{aligned}
\hat{V}_{JM} &= \hat{V} - \lambda_{ss} 1_N^\top - 1_N \lambda_{ss}^\top \\
&= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \lambda_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \lambda_3 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \lambda_3 & \lambda_3 & \lambda_3 \end{bmatrix} - \begin{bmatrix} 0 & 0 & \lambda_3 \\ 0 & 0 & \lambda_3 \\ 0 & 0 & \lambda_3 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} - \lambda_3 \\ \sigma_{12} & \sigma_{22} & \sigma_{23} - \lambda_3 \\ \sigma_{13} - \lambda_3 & \sigma_{23} - \lambda_3 & \sigma_{33} - 2\lambda_3 \end{bmatrix}
\end{aligned}
$$

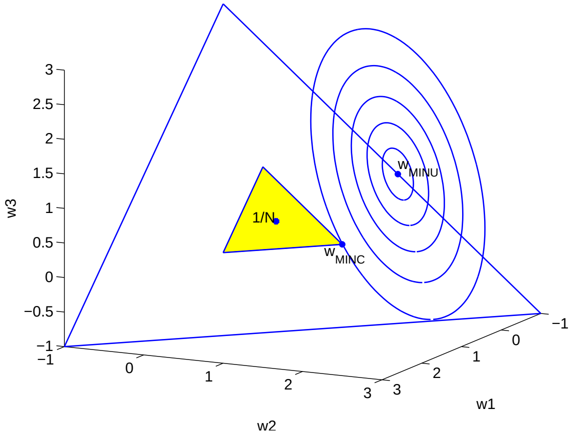
Example of shrinkage interpretation with $N = 3$ ...II

- We start by writing the last expression from the previous page:

$$\hat{V}_{JM} = \hat{V} - \lambda_{ss}\, 1_N^\top - 1_N \lambda_{ss}^\top$$
$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} - \lambda_3 \\ \sigma_{12} & \sigma_{22} & \sigma_{23} - \lambda_3 \\ \sigma_{13} - \lambda_3 & \sigma_{23} - \lambda_3 & \sigma_{33} - 2\lambda_3 \end{bmatrix}$$

- So, the effect of subtracting λ_3 in V is that the riskiness of the third asset is reduced, so that its weight is increased.
- The value of λ_3 is exactly what is required to increase the weight of the third asset from negative to zero.

Illustrating short-sale-constrained portfolio ($N = 3$) . . . I



Illustrating short-sale-constrained portfolio ($N = 3$) . . . II

- The ellipses in the figure represent the minimum-variance isoquants; i.e., each ellipse has the same portfolio variance.
- The triangle represents the constraint space, which is where
  - the sum of the weights equals 1, and
  - no short positions are allowed.
- The optimal portfolio is given by the intersection of
  - the isoquant with the lowest variance
  - that also satisfies the constraints.
- The figure shows that one limitation of imposing the short-sale constraint is that you may get zero holdings in some assets.

Road map

Performance of short-sale-constrained GMV portfolio

- On the next few pages, we compare the out-of-sample performance of the $1/N$ portfolio to the GMV portfolio (without and with short-sale constraints).
- We study the performance across six datasets.
- Performance is measured by the Sharpe ratio.
- We also report turnover, which provides a sense of the transaction costs associated with implementing the portfolio strategy.
- For each dataset, the best-performing strategy is highlighted in **bold**.

Sharpe ratios of GMV portfolios without and with short-sale constraints . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
MV (in sample)	0.3848	0.2124	0.2090	0.2851	0.5098	0.5364
MV (out of sample)	0.0794	−0.0363	−0.0719	0.2186	−0.0684	−0.0031
MV-C	0.0892	0.0678	0.0848	0.1084	0.1977	0.2024
$1/N$	**0.1876**	0.1353	0.1277	0.2240	0.1623	0.1753
	(0.12)	(0.01)	(0.01)	(0.46)	(0.00)	(0.01)
GMV portfolios (without and with constraints)						
GMV	0.0820	**0.1554**	0.1490	**0.2493**	**0.2778**	-0.0183
	(0.05)	(0.30)	(0.21)	(0.23)	(0.01)	(0.01)
GMV-C	0.0834	**0.1425**	**0.1501**	**0.2493**	0.1546	**0.3580**
	(0.01)	(0.41)	(0.16)	(0.23)	(0.35)	(0.00)

Sharpe ratios of GMV portfolios without and with short-sale constraints ... II

- The above table suggests that it is optimal to ignore estimates of expected returns
  - Performance of GMV portfolio is almost always better than that of the MV portfolio.
  - Performance of the short-sale constrained GMV-C portfolio is always better than that of the MV-C portfolio.
- The effect of short-sale constraints on the performance of the GMV portfolio is not substantial when $N < 25$ because
  - the performance of the GMV-C portfolio is not much better than that of GMV, except for the last dataset with $N = 24$.
- But, as we will see in the next table, short-sale constraints are useful for reducing turnover.

Turnover of GMV portfolios . . . I

Strategy \ Dataset	(1) S&P Sectors	(2) Industry Portf.	(3) Inter'l Portf.	(4) Mkt + SMB/HML	(5) FF 1-factor	(6) FF 4-factor
	$N = 11$	$N = 11$	$N = 9$	$N = 3$	$N = 21$	$N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
Turnover *relative* to $1/N$ strategy						
MV (out of sample)	38.99	607479.61	4236.10	2.83	9408.16	2672.71
MV-C	4.53	7.17	7.23	4.12	17.53	13.82
GMV portfolios (without and with constraints)						
GMV	6.54	21.65	7.30	1.11	45.47	6.83
GMV-C	2.47	2.58	2.27	1.11	3.93	1.76

Turnover of GMV portfolios . . . II

- The table on the previous slide shows that the turnover of the GMV portfolio is much smaller than the mean-variance portfolio.
- Turnover of GMV-C portfolio (GMV with short-sale constraints) is
  - even smaller than that of the GMV portfolio;
  - only slightly higher than that of the $1/N$ portfolio (about double).
- Thus, ignoring the mean and constraining the weights, reduces turnover substantially.

Road map

Approach 2: GMV with Ledoit-Wolf shrinkage . . . I

- Instead of "ad-hoc" shrinkage via the short-sale constraint, Ledoit and Wolf propose "optimal" shrinkage.
- Ledoit and Wolf propose replacing the sample covariance matrix $\hat{V}$ by a weighted average of $\hat{V}$ and a low-variance estimator, $\hat{V}_{target}$

$$\hat{V}_{LW} = (1 - \psi)\hat{V} + \psi\hat{V}_{target}, \quad \ldots \psi \text{ is the Greek letter "psi"}$$

- Using the weighted average $\hat{V}_{LW}$ is equivalent to shrinking the estimated sample covariance matrix toward $\hat{V}_{target}$.
- The degree of shrinkage is controlled by the shrinkage parameter ψ.

  (Most papers and Python codes use α instead of ψ for the shrinkage parameter.)

Approach 2: GMV with Ledoit-Wolf shrinkage . . . II

- To compute the shrinkage parameter ψ, Ledoit and Wolf minimize the expected Frobenius norm of the difference between $\hat{V}_{LW}$ and the true covariance matrix.
  - If the difference of the two matrices is the matrix A with elements a_{nm} and dimensions $N \times M$, its Frobenius norm, denoted $\|A\|_F$, is:

$$\|A\|_F = \sqrt{\sum_{n=1}^{N} \sum_{m=1}^{M} (a_{nm}^2)}.$$

- Mathematically, this shrinkage reduces the ratio between the smallest and largest eigenvalues of the estimated covariance matrix.
  - This shrinkage can be done by shifting every eigenvalue according to a given offset, which is equivalent to finding the ℓ^2-penalized Maximum Likelihood Estimator of the covariance matrix.

Different shrinkage targets suggested by Ledoit and Wolf

- Ledoit and Wolf have suggested several possibilities for $\hat{V}_{target}$:

$$\hat{V}_{LW} = (1-\psi)\hat{V} + \psi\hat{V}_{target},$$

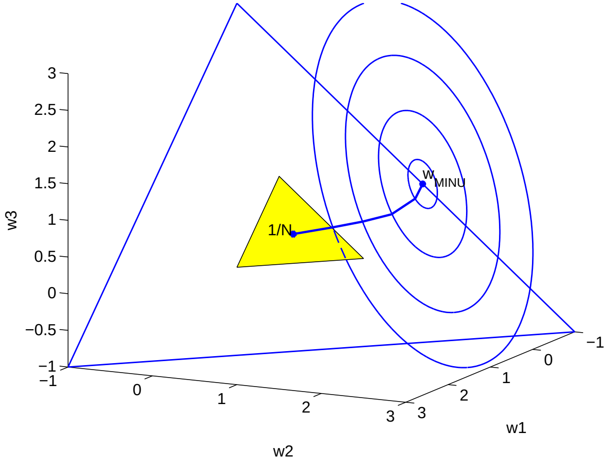
- Ledoit and Wolf (2003)
  - $\hat{V}_{target}$ is the 1-factor covariance matrix under the market model.
- Ledoit and Wolf (2004a)
  - $\hat{V}_{target}$ is the diagonal matrix with average variance on the diagonal.
- Ledoit and Wolf (2004b)
  - $\hat{V}_{target}$ is a matrix where all pairwise correlations are set equal to the average correlation while the sample variances are unchanged.

Example of shrinkage estimator (Ledoit and Wolf 2004a)

- If $\hat{V}_{target}$ is the diagonal matrix with average variance on the diagonal, then for $N = 3$, the Ledoit-Wolf covariance matrix is

$$
\begin{aligned}
\hat{V}_{LW} &= (1-\psi)\hat{V} + \psi \hat{V}_{target} \\
&= (1-\psi)\hat{V} + \psi \frac{\text{trace}(\hat{V})}{N} I_{3\times 3} \\
&= (1-\psi)\hat{V} \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} & \hat{\sigma}_{23} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_{33} \end{bmatrix} + \psi \frac{(\hat{\sigma}_{11} + \hat{\sigma}_{22} + \hat{\sigma}_{33})}{3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= (1-\psi)\hat{V} \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} & \hat{\sigma}_{23} \\ \hat{\sigma}_{13} & \hat{\sigma}_{23} & \hat{\sigma}_{33} \end{bmatrix} \\
&\quad + \psi \begin{bmatrix} \frac{(\hat{\sigma}_{11}+\hat{\sigma}_{22}+\hat{\sigma}_{33})}{3} & 0 & 0 \\ 0 & \frac{(\hat{\sigma}_{11}+\hat{\sigma}_{22}+\hat{\sigma}_{33})}{3} & 0 \\ 0 & 0 & \frac{(\hat{\sigma}_{11}+\hat{\sigma}_{22}+\hat{\sigma}_{33})}{3} \end{bmatrix}.
\end{aligned}
$$

Illustrating the Ledoit-Wolf portfolios for 3 assets . . . I



Illustrating the Ledoit-Wolf portfolios for 3 assets . . . II

- The ellipses in the figure represent the minimum-variance isoquants; i.e., each ellipse has the same portfolio variance.
- The optimal portfolio is given by the
  - weighted average of the GMV portfolio (w_{MINU} in the figure) and the $1/N$ portfolio,
  - where the weights are $(1-\psi)$ and ψ;
  - these weights determine where exactly the portfolio lies between w_{MINU} and $1/N$.
- The value of ψ is obtained by
  - making a distributional assumption about the variances and covariances and
  - then choosing ψ to minimize the distance between $\hat{V}_{LW}$ and the true covariance matrix.

Python code for Ledoit-Wolf shrinkage of covariance matrix

- Python code to implement Ledoit-Wolf shrinkage is available from a variety of sources:
  - For Ledoit and Wolf (2004a).
  - For Ledoit and Wolf (2004b).
  - For Chen, Wiesel, Eldar, and Hero (2010), whose method "Oracle Approximating Shrinkage," improves on the convergence properties of Ledoit and Wolf.
  - Code to compare Ledoit-Wolf vs OAS Estimation in Scikit Learn.

Ledoit and Wolf: Nonlinear shrinkage

- The Ledoit and Wolf (2003, 2004a, 2004b) papers propose linear shrinkage methods.
- In more recent work, Ledoit and Wolf (2017, 2022) have proposed nonlinear shrinkage methods.
- I have not included nonlinear shrinkage in the course because
  - the performance gains from nonlinear shrinkage over linear shrinkage are modest.
- For those of you who are interested, an excellent easy-to-follow exposition is provided in this article by Jeffrey Näf.
  - The article includes Python code (and also code in R and Matlab).
  - Python code for the nonlinear shrinkage proposed in Ledoit and Wolf (2020) is available from this link.
  - If you want, you can write a master's project based on nonlinear shrinkage of covariance matrices.

Road map

Approach 3: GMV with norm constraints

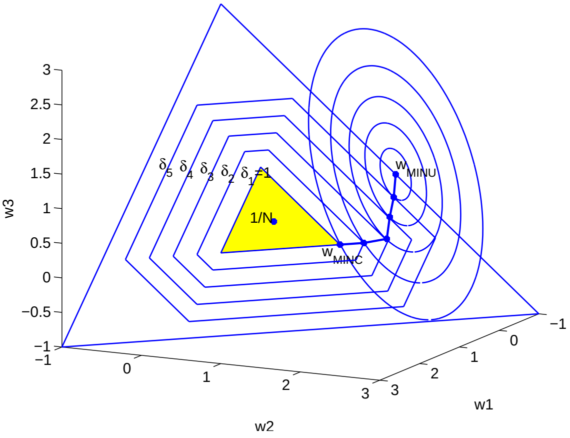
- Solve minimum-variance problem subject to the constraint that norm of portfolio-weight vector is smaller than a threshold δ:
- General norm-constrained portfolio

$$
\begin{aligned}
\min_{w} \quad & w^\top \hat{V} w && \ldots \text{portfolio variance} \\
\text{s.t.} \quad & w^\top 1_N = 1 && \ldots \text{adding-up constraint} \\
& \|w\| \leq \delta && \ldots \text{constraint on portfolio norm}
\end{aligned}
$$

where $\|w\|$ is the norm of the portfolio-weight vector.

- We consider the following norms for the constraint:
  - The 1-norm constraint: $\|w\|_1 = \sum_{n=1}^{N} |w_n| \leq \delta$.
  - The 2-norm constraint: $\|w\|_2 = \left(\sum_{n=1}^{N} w_n^2\right)^{1/2} \leq \delta$.
  - The general A-norm $\|w\|_A = \left(w^\top A w\right)^{1/2}$.

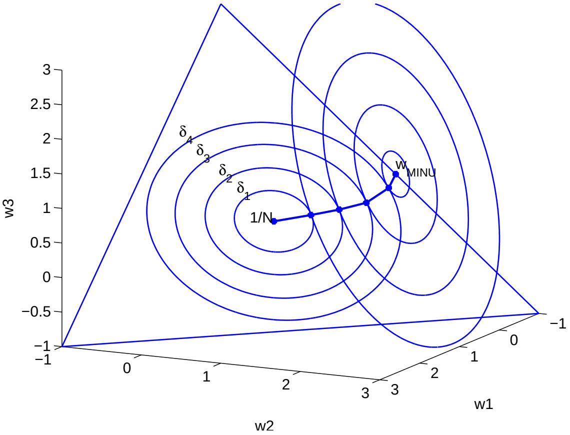
Illustrating 1-norm-constrained portfolio ($N = 3$) . . . I



Illustrating 1-norm-constrained portfolio ($N = 3$) . . . II

- The ellipses in the figure represent the minimum-variance isoquants; i.e., each ellipse has the same portfolio variance.
- The polygons represents the constraint space, which is where
  - the sum of the absolute value of the weights is less than δ.
- The optimal portfolio is given by the intersection of
  - the isoquant with the lowest variance
  - that also satisfies the constraints.
- The figure shows that one limitation of imposing the 1-norm constraint is that you may get zero holdings in some assets.

Illustrating 2-norm-constrained portfolio ($N = 3$) . . . I



Illustrating 2-norm-constrained portfolio ($N = 3$) ... II

- The ellipses in the figure represent the minimum-variance isoquants; i.e., each ellipse has the same portfolio variance.
- The circles represents the constraint space, which is where
  - the sum of the square of the weights is less than δ.
- The optimal portfolio is given by the intersection of
  - the isoquant with the lowest variance
  - that also satisfies the constraints.
- The figure shows that one advantage of imposing the 2-norm constraint is that you are unlikely to get zero holdings of assets.

How to choose the value of δ?

- How should one choose δ in the norm-constraint?

$$\|w\| \leq \delta \qquad \ldots \text{constraint on portfolio norm}$$

- The choice of δ here is similar to Ledoit-Wolf's choice of ψ.
    - Ledoit-Wolf make distributional assumptions about returns to derive the optimal value of ψ;
    - Here, we do not make any distributional assumption and we let the data guide our choice of δ.
- DeMiguel, Garlappi, Nogales, and Uppal (2009) explain how cross-validation can be used to choose the optimal value of δ.
- Note: when choosing δ, it is important not to have a look-ahead bias; i.e., we should use only historical data when choosing δ.

A brief summary of cross validation

- Cross-validation is a resampling method from machine learning.
- Cross-validation uses different portions of the data to estimate and test the choice of parameters.
- For example, you could take the historical sample data and
  - divide it into five folds (parts);
  - define possible values of δ, also known as the grid for δ;
  - estimate your model with each δ using four folds;
  - test which delta delivers the optimal model performance in the fifth (holdout) fold using a selection criterion (e.g., minimizing portfolio volatility or maximizing Sharpe ratio)
  - then, "rotate the folds" (i.e., use a different set of four folds to estimate δ and test it on the fifth fold).
  - choose a value of δ that delivers the best average performance across all the hold-out folds.
- For more details about cross-validation, see Wikipedia.

Python code for cross validation

- Cross-validation techniques are used widely in machine learning.
- Python code for cross-validation is available from many sources:
  - scikit-learn.
  - W3schools.com.
  - Analytics Vidhya (log-in required to read web site).
- For this course, you do not need to learn cross-validation methods. (But, it is an extremely useful technique to know.)

Road map

Sharpe ratios of LW and norm-constrained portfolios

Strategy	10Ind	48Ind	6FF	25FF	500CRSP
Simple benchmarks					
1/N	0.2541	0.2508	0.2563	0.2565	0.3326
Minimum-variance portfolios					
GMV-C	0.2852	**0.2914**	0.2629	0.2720	0.3985
	(0.06)	(0.44)	(0.00)	(0.00)	(0.08)
Ledoit and Wolf shrinkage portfolios					
LW-diagonal	0.2962	0.2620	0.3226	0.3974	0.4086
	(0.11)	(0.05)	(0.00)	(0.10)	(0.11)
LW-1-factor	0.2902	0.2544	0.3296	0.3927	0.4500
	(0.13)	(0.04)	(0.01)	(0.09)	(0.56)
Norm-constrained portfolios					
NC-1-norm	0.2890	0.2831	0.3374	0.3553	0.3706
	(0.05)	(0.19)	(0.00)	(0.00)	(0.04)
NC-2-norm	**0.3193**	0.2891	**0.3922**	**0.4278**	**0.4672**
	(0.40)	(0.05)	(0.93)	(0.36)	(0.55)

- 2-norm constrained portfolios best, but not very different from LW.

Turnover of LW and norm-constrained (NC) portfolios

Strategy	10Ind	48Ind	6FF	25FF	500CRSP
Simple benchmarks					
1/N	0.0232	0.0311	0.0155	0.0174	0.0595
Minimum-variance portfolio policies					
GMV	0.1656	0.8286	0.2223	0.7953	0.7769
GMV-C	**0.0552**	**0.0741**	**0.0461**	**0.0841**	**0.4222**
Ledoit and Wolf shrinkage portfolios					
LW-identify	0.1132	0.4029	0.0905	0.3144	0.5594
LW-1-factor	0.1428	0.4034	0.1455	0.4893	0.5463
Norm-constrained portfolio policies					
NC-1-norm	0.6013	0.8232	1.0064	0.9767	0.9753
NC2-norm	1.0177	2.7556	1.6594	3.6275	1.0443

- Excluding $1/N$, turnover lowest (best) for GMV-C, then LW-diagonal, then LW-1-factor;
- Norm-constrained portfolios have the highest (worst) turnover.

End of focus

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will study how to take advantage of factor models of asset returns to reduce the error in estimating expected returns and to reduce the dimensionality of the covariance matrix.

To do for next class: Readings

- Readings
  - DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. Download link.
    The data and code (in Matlab) to replicate the results in this paper are available from this link to Lorenzo Garlappi's website.
  - Please read Moreira (2021, Chapter 9), which is available online from this link.
  - Please read the subsection on "Constrained Optimization and Backtesting" in Scheuch, Voigt, Weiss, and Frey (2024), available online at Tidy Finance.
- Assignment
  - You should start working on the next assignment.

Road map

Appendix: Numerical optimization in Python

- In this appendix, we see how to do numerical optimization in Python.
- We look at the problems of finding portfolio weights that:
  - Maximize mean-variance utility;
  - Minimize portfolio variance;
  - Other problems discussed in this chapter can then be coded similarly.
- We look at both constrained and unconstrained optimization.

Road map

Maximizing mean-variance utility with risk-free asset . . . I

- With a risk-free asset, the optimal portfolio problem is

$$\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad (1) \quad \sum_{n=1}^{N} w_n + w_f = w^\top 1_N + w_f = 1;$$

$$(2) \quad \sum_{n=1}^{N} w_n \mathbb{E}[R_n] + w_f R_f = w^\top \mu + w_f R_f = \mu_{\mathbf{targ}}.$$

Maximizing mean-variance utility with risk-free asset . . . II

- But, constraint (1) implies that:

$$w_f = 1 - w^\top 1_N.$$

- Substituting this expression for w_f into constraint (2), we get that the problem to be solved is:

$$\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad w^\top \mu + (1 - w^\top 1_N) R_f = \mu_{\text{targ}}.$$

- Solving the above problem gives the optimal portfolio weights:

$$w = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)}\right] V^{-1} (\mu - R_f 1_N). \qquad (25)$$

Mean-variance utility formulation

- Markowitz shows that minimizing the portfolio variance subject to a target-mean constraint is equivalent to maximizing mean-variance utility (MVU):

$$
\begin{aligned}
&\max_w \quad \mu_p - \frac{\gamma}{2}\sigma_p^2 \\
\Rightarrow &\max_w \quad \underbrace{\Big[(1 - w^\top 1_N)R_f + w^\top \mu\Big]}_{\mu_p} - \frac{\gamma}{2}\underbrace{w^\top V w}_{\sigma_p^2} \\
\Rightarrow &\max_w \quad \Big[R_f + w^\top(\mu - R_f 1_N)\Big] - \frac{\gamma}{2} w^\top V w \quad \ldots\text{collecting terms in } w
\end{aligned}
$$

- Solving this equation gives:

$$
w = \frac{1}{\gamma} V^{-1}(\mu - R_f 1_N) \quad \ldots\text{optimal portfolio.} \tag{26}
$$

Equivalence between weights from the two approaches

- The solutions for the two different formulations are

$$w = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)} \right] V^{-1} (\mu - R_f 1_N) \quad \text{and}$$

$$w = \frac{1}{\gamma} V^{-1} (\mu - R_f 1_N),$$

- We see that the two expressions for w are identical if we set

$$\frac{1}{\gamma} = \left[\frac{(\mu_{\text{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1} (\mu - R_f 1_N)} \right].$$

- Similarly, we note that both formulations lead to the same expression for the tangency portfolio

$$w_{\text{tang}} = \frac{\text{optimal weight vector}}{\text{sum of weights in all risky assets}} = \frac{V^{-1}(\mu - R_f 1_N)}{1_N^\top V^{-1}(\mu - R_f 1_N)}.$$

Python code for the mean-variance optimal portfolios

- We now provide Python code to solve an example of the mean-variance optimization problem when a risk-free asset is available,
  - first without any constraints and
  - then with constraints (when the above solution is no longer valid).

Optimal mean-variance portfolios (MVP)

Code common to all the MVP problems we solve in this appendix

```
# Import the packages we will need
import numpy as np
from scipy.optimize import minimize

# Column vector of ones
ones = np.ones((3, 1))

# Investor's risk aversion
# The value chosen matches that for the risk-minimization formulation
# The value we choose will not affect the tangency portfolio weights
gamma = 2.715745

# Gross risk-free rate
Rf = 1.065

# Variance-covariance matrix of returns
V = np.array(    [
    [0.0625, 0.0700, 0.1050],
    [0.0700, 0.1225, 0.0840],
    [0.1050, 0.0840, 0.3600]
    ]  )

# Column vector of gross mean returns
mu = np.array(  [ [1.20], [1.30], [1.40] ] )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)
```

Analytic solution for MVP weights (no constraints) . . . I

Analytic solution for MVP weights (no constraints)

```
# Weights for the mean-variance portfolio (MVP)
w_mvp = (1/gamma) * Vinverse @ (mu - Rf * ones)

#Weights in the risk-free asset
w_riskFree = 1 - ones.T @ w_mvp

#Weights of the tangency portfolio
w_tangency = w_mvp / (ones.T @ w_mvp)

print("\nBased on analytic solution ")

print("\nWeights of the mean-variance portfolio (MVP) are = ")
print(w_mvp)

print("\nWeight in the risk-free asset = ")
print(w_riskFree)

print("\nWeights of the tangency portfolio (based on MVP) are = ")
print(w_tangency)
```

Analytic solution for MVP weights (no constraints) . . . II

```
Based on analytic solution

Weights of the mean-variance portfolio (MVP) are =
[[-1.00407932]
 [ 1.00520178]
 [ 0.40096142]]

Weight in the risk-free asset =
[[0.59791612]]

Weights of the tangency portfolio (based on MVP) are =
[[-2.49718869]
 [ 2.49998029]
 [ 0.9972084 ]]
```

Numerical solution using minimization (unconstrained)

Code for Numerical solution using minimization (unconstrained)

```
# Remember to first execute the code common to all problems

# Define the objective function to be minimized
# Note that the optimizer only minimizes,
#   so reverse sign of objective because it is to be maximized
def objective(weights):
    portfolio_return = weights.T @ (mu - Rf * ones)
    portfolio_risk = weights.T @ V @ weights
    return (gamma/2) * portfolio_risk -  portfolio_return

# Note that we need to give some starting value for each unknown weight
# We set the starting value to be our initial guess 1/N
initial_weights = np.ones(len(mu))/len(mu)

# Use scipy.optimize.minimize to find the optimal weights
result = minimize(objective, initial_weights, method='SLSQP')

# Code continues on the next slide ...
```

Code (contd.)

Code (contd.)

```
# Extract the optimal weights from results
w_mvp = result.x

#Weights in the risk-free asset
w_riskFree = 1 - ones.T @ w_mvp

#Weights of the tangency portfolio
w_tangency = w_mvp / (ones.T @ w_mvp)

print("\n* Using numerical solver")

print("\nWeights of the mean-variance portfolio (MVP) are = ")
print(w_mvp)

print("\nWeight in the risk-free asset = ")
print(w_riskFree)

print("\nWeights of the tangency portfolio (based on MVP) are = ")
print(w_tangency)
```

Numerical solution using minimization (unconstrained)

- Solution from using numerical optimization:

```
* Using numerical solver

Weights of the mean-variance portfolio (MVP) are =
[-1.00418375  1.00507421  0.40096424]

Weight in the risk-free asset =
[0.5981453]

Weights of the tangency portfolio (based on MVP) are =
[-2.49887271  2.5010886   0.99778412]
```

- The above solution is very close to the analytic (closed-form) solution:

```
* Based on analytic solution

Weights of the mean-variance portfolio (MVP) are =
[[-1.00407932]  [ 1.00520178]  [ 0.40096142]]

Weight in the risk-free asset =
[[0.59791612]]

Weights of the tangency portfolio (based on MVP) are =
[[-2.49718869] [ 2.49998029]  [ 0.9972084 ]]
```

Adding position constraints to the portfolio problem

- So far, we solved the problem without requiring the constraints to be nonnegative and without requiring the weights to be less than a particular value.
- We now see how to impose bounds on the choice variables when using scipy optimize.

Mean-variance optimization with short-sale constraint

Code for mean-variance optimization with short-sale constraint

```
# Remember to first execute the code common to all MVP problems

# Define the objective function to be minimized
# Note that the optimizer only minimizes,
#   so reverse sign of objective if it is to be maximized

def objective(weights):
    portfolio_return = weights.T @ (mu - Rf * ones)
    portfolio_risk = weights.T @ V @ weights
    return (gamma/2) * portfolio_risk -  portfolio_return

# Define the initial guess for weights
initial_weights = np.ones(len(mu))/len(mu)

# Add a lower-bound constraint of 0 to the above problem
# There is no upper bound, so I have set it to 999
bounds = tuple((0, 999) for _ in range(len(mu)))

# When calling minimize, we add ‘‘bounds’’ to the minimization function
result = minimize(objective, initial_weights, method='SLSQP', bounds =
     bounds)
```

Code with shortsale constraints (contd.)

Code (contd.)

```
# Extract the optimal weights from results
optimal_weights = result.x

#Weights in the risk-free asset
w_riskFree = 1 - ones.T @ optimal_weights

#Weights of the ‘‘tangency’’ portfolio
w_tangency = optimal_weights / (ones.T @ optimal_weights)

w_mvp_shortsaleConstrained = list(map(lambda num: format(num, ’.6f’),
     optimal_weights))
w_tangency_formatted = list(map(lambda num: format(num, ’.6f’),
     w_tangency))

print("\n* Using numerical solver and with short-sale constraints")
print("\nWeights of the mean-variance portfolio (MVP) are = ")
print(w_mvp_shortsaleConstrained)
print("\nWeight in the risk-free asset = ")
print(w_riskFree)
print("\nWeights of the tangency portfolio (based on MVP) are = ")
print(w_tangency_formatted)
```

Solution for short-sale-constrained mean-variance weights

```
* Using numerical solver and with short-sale constraints

Weights of the mean-variance portfolio (MVP) are =
['0.000000', '0.560923', '0.211491']

Weight in the risk-free asset =
[0.22758546]

Weights of the tangency portfolio (based on MVP) are =
['0.000000', '0.726194', '0.273806']
```

Road map

Minimizing portfolio variance . . . I

- We now look at the problem of finding the global minimum-variance (GMV) portfolio weights.
- This optimization problem, as we saw earlier, is:

$$\text{Minimize}_{\{w_1,\ldots,w_N\}} \quad \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n w_m \sigma_{nm} = w^\top V w$$

$$\text{subject to:} \quad (1) \quad \sum_{n=1}^{N} w_n = w^\top 1_N = 1;$$

- This problem is different from finding mean-variance optimal weights because we now need to have an explicit constraint, which is that the sum of the weights equals 1.
  - So, this problem gives us the opportunity to learn how to add constraints to `scipy.optimize`.

Minimizing portfolio variance . . . II

- The analytic solution to this problem is

$$w = \frac{V^{-1}1_N}{1_N^\top V^{-1}1_N},$$

- where the numerator of w has the "raw" weights that are inversely related to their risk (so, more risky assets have a smaller weight).
- and the denominator of w has the sum of these weights, which then scales the "raw" weights in the numerator so that they add up to 1.

Minimizing portfolio variance . . . III

- Just as before,
  1. We will first solve the unconstrained problem analytically;
  2. Then we will solve the same problem numerically;
  3. Finally, we will solve the problem numerically when there are also position bounds on the weights (and so, no analytic solution exists).

Analytic solution for finding GMV weights (no bounds)

Code common to all the GMV problems we solve in this appendix

```
# Import the packages we will need
import numpy as np
from scipy.optimize import minimize

# Column vector of ones
ones = np.ones((3, 1))

# Variance-covariance matrix of returns
V = np.array(    [
    [0.0625, 0.0700, 0.1050],
    [0.0700, 0.1225, 0.0840],
    [0.1050, 0.0840, 0.3600]
    ]  )

# Inverse of the variance-covariance matrix
Vinverse = np.linalg.inv(V)
```

Analytic solution for GMV weights (no bounds)

Code for analytic solution for GMV weights (no bounds)

```
# Remember to execute the common code for GMV first

# Weights of the global minimum variance (GMV) portfolio
w_gmv = (Vinverse @ ones)/(ones.T @ Vinverse @ ones)

print("\n* Using analytic solution")
print("\nWeights of the global minimum variance (GMV) portfolio = ")
print(w_gmv)
```

```
* Using analytic solution

Weights of the global minimum variance (GMV) portfolio =
[[ 1.56355533]
 [-0.32056227]
 [-0.24299306]]
```

Numerical solution for GMV weights (no bounds)

Code for solving GMV weights numerically (no bounds)

```
# Remember to execute the common code for GMV first

# Define the objective function to be minimized (portfolio variance)
def objective(weights, V):
    portfolio_variance = weights.T @ V @ weights
    return portfolio_variance

# Define the initial guess for weights
initial_weights = np.ones(len(V)) / len(V)

# This is the new step: define constraint
# Define the equality constraint (weights sum to 1)
constraints = ({'type': 'eq', 'fun': lambda weights: np.sum(weights) -
     1})

# Use scipy.optimize.minimize to find the optimal weights
result = minimize(objective, initial_weights, args=(V,), method='SLSQP',
      constraints=constraints)

# Code continues on the next page
```

Numerical solution for GMV weights (no bounds) ... contd.

Code for solving GMV weights numerically (no bound) ... contd.

```
# Extract the optimal weights
optimal_weights = result.x

# Print the results
print("\n* Using numeric solution with no short-sale constraints")
print("\nOptimal Weights:")
for i, weight in enumerate(optimal_weights):
    print(f"Asset {i + 1}: {weight:.6f}")
print("\nMinimum Portfolio Variance is:", result.fun)
```

Solution to GMV weights (no bounds)

- Solution from using numerical optimization (with constraint that the weights sum to 1).

```
* Using numeric solution with no short-sale constraints

Optimal Weights:
Asset 1: 1.563574
Asset 2: -0.320576
Asset 3: -0.242997
```

- We can see below that this matches the analytic solution.

```
* Using analytic solution

Weights of the global minimum variance (GMV) portfolio =
[[ 1.56355533]
 [-0.32056227]
 [-0.24299306]]
```

Numeric solution for GMV with short-sale constraints

Code for GMV with short-sale constraints

```
# Remember to execute the common code for GMV first

# Define the objective function to be minimized (portfolio variance)
def objective(weights, V):
    portfolio_variance = weights.T @ V @ weights
    return portfolio_variance

# Define the equality constraint (weights sum to 1)
constraints = ({'type': 'eq', 'fun': lambda weights: np.sum(weights) -
     1})

# Define the initial guess for weights
initial_weights = np.ones(len(V)) / len(V)

# We now add to the code the short-sale constraint
# Define the bounds for each weight (0 <= weight <= 999)
bounds = tuple((0, 999) for _ in range(len(V)))

# Code continues on the next page
```

Numeric solution for GMV with short-sale constraints

Code for GMV with short-sale constraints (contd.)

```
# Use scipy.optimize.minimize to find the optimal weights
result = minimize(objective, initial_weights, args=(V,), method='SLSQP',
      bounds=bounds, constraints=constraints)

# Extract the optimal weights
optimal_weights = result.x

# Print the results
print("\n* Using numeric solution with short-sale constraints")

print("\nOptimal Weights:")
for i, weight in enumerate(optimal_weights):
    print(f"Asset {i + 1}: {weight:.6f}")

print("\nMinimum Portfolio Variance is:", result.fun)
```

Numeric solution for GMV with short-sale constraints

```
* Using numeric solution with short-sale constraints

Optimal Weights:
Asset 1: 1.000000
Asset 2: 0.000000
Asset 3: 0.000000
```

- This solution illustrates the problem with short-sale constraints
  - A short-sale constraint leads to a zero position in many assets;
  - As a result, the portfolio is not well diversified.

GMV weights bounded between min and max values

- We conclude our exploration of scipy.optimize by looking at GMV weights that are bounded between minimum and maximum values.
- In our example, we set the
  - minimum value to be 0.0, and
  - maximum value to be 0.5.

Numeric solution for GMV with bounds on weights

Code for numeric solution for GMV with bounds on weights

```
# Remember to execute the common code for GMV first

# Define the objective function to be minimized (portfolio variance)
def objective(weights, V):
    portfolio_variance = weights.T @ V @ weights
    return portfolio_variance

# Define the equality constraint (weights sum to 1)
constraints = ({'type': 'eq', 'fun': lambda weights: np.sum(weights) -
     1})

# Define the initial guess for weights
initial_weights = np.ones(len(V)) / len(V)

# We have changed the next line of code to bound the weights
# Define the bounds for each weight (0 <= weight <= 0.50)
bounds = tuple((0, 0.50) for _ in range(len(V)))

# code continued on next page
```

Numeric solution for GMV with bounds on weights (contd.)

Code for numeric solution for GMV with bounds on weights (contd.)

```python
# Use scipy.optimize.minimize to find the optimal weights
result = minimize(objective, initial_weights, args=(V,), method='SLSQP',
      bounds=bounds, constraints=constraints)

# Extract the optimal weights
optimal_weights = result.x

# Print the results
print("\n* Using numeric solution with short-sale constraints")

print("\nOptimal Weights:")
for i, weight in enumerate(optimal_weights):
    print(f"Asset {i + 1}: {weight:.6f}")

print("\nMinimum Portfolio Variance is:", result.fun)
```

Numeric solution for GMV with bounds on weights

```
* Using numeric solution with short-sale constraints and an upper bound

Optimal Weights:
Asset 1: 0.500000
Asset 2: 0.494436
Asset 3: 0.005564
```

- From the above solution, we see that
  - the portfolio has an investment in both the first and second asset.
- So, this portfolio is better diversified compared to the case with only short-sale constraints (and no upper bound on the maximum weight of an asset), where
  - there was an investment in only the first asset.

References for additional details

- There is a rich literature on numerical optimization, because this is a problem that appears in a large number of fields.
- For an easy-to-understand guide to optimization in Python, with several worked-out examples, see Optimization in SciPy, which is a chapter in the online book, "Scientific Computing with Python".
- For an excellent and deeper discussion see Section 2.7 on Mathematical Optimization by Gaël Varoquaux, which is from the book "Scientific Python Lectures."
- You can also see the official documentation at scipy for
  - scipy.optimize,
  - scipy.optimize.minimize, and
  - Optimization.
- Note that Chat-GPT is quite good at providing a "first version" of Python code, which you can then correct/improve.

Aside: Information about "norms" from Wikipedia . . . I

- A norm is a function from a (real or complex) vector space to the non-negative real numbers, which
  - behaves in certain ways like the distance from the origin:
  - it commutes with scaling, obeys a form of the triangle inequality, and is zero only at the origin.
- The Euclidean norm or 2-norm or L2 is the magnitude of a vector.
  - On the N-dimensional Euclidean space $\mathbb{R}^N$, the length of the vector $x = (x_1, x_2, \ldots, x_N)$ is given by the Euclidean distance

$$\|x\|_2 := \sqrt{x_1^2 + \cdots + x_N^2}.$$

- The absolute value or 1-norm or L1 is a norm on the 1-dimensional vector spaces is given by the Manhattan distance or taxi-cab norm.

$$\|x\|_1 = |x_1| + \cdots + |x_N|.$$

Aside: Information about "norms" from Wikipedia . . . II

- The *p*-norm or ℓ^p-norm of the vector $x = (x_1, \dots, x_N)$, where $p \geq 1$ be a real number, is

$$\|x\|_p := \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}.$$

  - For $p = 1$, we get the taxicab norm,
  - for $p = 2$ we get the Euclidean norm, and
  - as p approaches ∞ the p-norm approaches the infinity norm or maximum norm:

$$\|x\|_\infty := \max_n |x_n|.$$

- For additional details, see Wikipedia.

Aside: Matrix norms . . . I

- Just like we did for vectors, we can define norms also for matrices.
- Here is an explanation provided by Chat-GPT using the prompt: "Can you please explain what is a Frobenius norm at a level that is suitable for university students."
- The Frobenius norm, also known as the matrix norm, is a mathematical concept used in linear algebra.
- It is a way to measure the "size" or "magnitude" of a matrix.

Aside: Matrix norms . . . II

- The Frobenius norm of a matrix is a way to calculate the overall "length" or "size" of the matrix, similar to how you might calculate the length of a vector in three-dimensional space.
- If you have a matrix A with elements a_{nm} and dimensions $N \times M$, you can calculate its Frobenius norm (denoted $\|A\|_F$) as follows:

$$\|A\|_F = \sqrt{\sum_{n=1}^{N} \sum_{m=1}^{M} (a_{nm}^2)}.$$

- The Frobenius norm is particularly useful for measuring the difference between two matrices.
- For additional details, see Wikipedia (toward bottom of page).

Road map

Bibliography . . . I

Andersen, T. G., and L. Benzoni. 2010. Stochastic volatility. *CREATES Research Paper,* nos. 2010-10. (Cited on page 506).

Barry, C. B. 1974. Portfolio analysis under uncertain means, variances, and covariances [in English]. *Journal of Finance* 29 (2): 515–522. (Cited on pages 458, 465).

Bawa, V. S., S. Brown, and R. Klein. 1979. *Estimation risk and optimal portfolio choice.* North Holland, Amsterdam. (Cited on pages 458, 732, 884, 1028).

Best, M. J., and R. R. Grauer. 1992. Positively weighted minimum-variance portfolios and the structure of asset expected returns. *Journal of Financial and Quantitative Analysis* 27 (4): 513–537. (Cited on page 458).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

———. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Boyle, P., L. Garlappi, R. Uppal, and T. Wang. 2012. Keynes meets Markowitz: The trade-off between familiarity and diversification. *Management Science* 58 (2): 253–272. (Cited on page 459).

Brown, S. 1979. The effect of estimation risk on capital market equilibrium. *Journal of Financial and Quantitative Analysis* 14 (2): 215–220. (Cited on page 465).

Bibliography . . . II

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, no. 1 (March): 57–82. (Cited on pages 131, 469, 966).

Chan, L. K. C., J. Karceski, and J. Lakonishok. 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies* 12 (5): 937–74. (Cited on pages 458, 500).

Chen, Y., A. Wiesel, Y. C. Eldar, and A. O. Hero. 2010. Shrinkage algorithms for mmse covariance estimation. *IEEE Transactions on Signal Processing* 58 (10): 5016–5029. (Cited on page 542).

Chopra, V. K. 1993. Improving optimization. *Journal of Investing* 8 (Fall): 51–59. (Cited on page 459).

Clarke, R. G., H. de Silva, and S. Thorley. 2010. Minimum variance portfolio composition. *SSRN, www.* (Cited on page 509).

Dangl, T., and M. Kashofer. 2013. Minimum-variance stock picking – a shift in preferences for minimum-variance constituents. Working Paper, Vienna University of Technology. (Cited on page 510).

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal. 2009. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science* 55 (5): 798–812. (Cited on pages 501, 519, 520, 550).

Bibliography . . . III

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. (Cited on pages 450, 495, 559, 619, 724, 1120).

DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov. 2013. Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis* 48 (6): 1813–1845. (Cited on page 506).

Driessen, J., P. Maenhout, and G. Vilkov. 2009. The price of correlation risk: Evidence from equity options. *Journal of Finance* 64 (3): 1375–1404. (Cited on page 506).

Fama, E. F., and K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

Frost, P. A., and J. E. Savarino. 1988. For better performance constrain portfolio weights. *Journal of Portfolio Management* 15 (1): 29–34. (Cited on page 459).

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies* 20:41–81. (Cited on pages 459, 835).

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131 (1-2): 59–95. (Cited on page 506).

Bibliography . . . IV

Goldfarb, D., and G. Iyengar. 2003. Robust portfolio selection problems. *Mathematics of Operations Research* 28 (1): 1–38. (Cited on pages 459, 835).

Green, R., and B. Hollifield. 1992. When will mean-variance efficient portfolios be well diversified? *Journal of Finance* 47:1785–1809. (Cited on page 500).

Harvey, C. R., J. Liechty, M. Liechty, and P. Müller. 2003. Portfolio selection with higher moments. Working Paper, Duke University. (Cited on page 459).

Jacobs, H., S. Müller, and M. Weber. 2014. How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets* 19:62–85. (Cited on pages 495, 620, 621, 623, 724).

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance* 58 (4): 1651–1684. (Cited on pages 459, 501, 519, 522).

James, W., and C. Stein. 1961. Estimation with quadratic loss. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability:* 361–379. (Cited on page 466).

Jobson, J. D., and B. M. Korkie. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance:* 889–908. (Cited on pages 187, 214, 447).

Jobson, J. D., and B. Korkie. 1980. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association* 75:544–554. (Cited on page 458).

Bibliography . . . V

Jobson, J., B. Korkie, and V. Ratti. 1979. Improved estimation for Markowitz portfolios using James-Stein type estimators. *Proceedings of the American Statistical Association:* 279–284. (Cited on page 458).

Jorion, P. 1985. International portfolio diversification with estimation risk [in English]. *Journal of Business* 58 (3): pp. 259–278. (Cited on pages 458, 467, 500, 732, 884, 1028).

———. 1986. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 458, 467, 500).

Kan, R., and G. Zhou. 2005. Optimal estimation for economic gains: Portfolio choice with parameter uncertainty. Working paper, University of Toronto. (Cited on page 459).

Klein, R. W., and V. S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3 (3): 215–31. (Cited on pages 465, 732, 884, 1028).

Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10 (5): 603–621. (Cited on pages 519, 538, 543, 1120).

———. 2004a. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88 (January): 365–411. (Cited on pages 458, 501, 519, 538, 539, 542, 543).

Bibliography . . . VI

Ledoit, O., and M. Wolf. 2004b. Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management* 30, no. 4 (December): 110–119. (Cited on pages 458, 501, 519, 538, 542, 543).

———. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *Review of Financial Studies* 30 (12): 4349–4388. (Cited on pages 501, 519, 543).

———. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Working Paper. (Cited on page 543).

———. 2022. The power of (non-) linear shrinking: a review and guide to covariance matrix estimation. *Journal of Financial Econometrics* 20 (1): 187–218. (Cited on pages 501, 519, 543).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

Merton, R. C. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8 (4): 323–361. (Cited on page 417).

Bibliography . . . VII

Michaud, R. O. 1998. *Efficient asset management: A practical guide to stock portfolio management and asset allocation.* Survey and Synthesis Series, HBS Press, Boston: Financial Management Association. (Cited on page 459).

Moreira, A. 2021. *Quantitative investing.* Available online from this link. (Cited on pages 248, 399, 559).

Pastor, L. 2000. Portfolio selection and asset pricing models. *The Journal of Finance* 55:179–223. (Cited on page 468).

Pástor, Ľ., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56 (3): 335–381. (Cited on pages 468, 732, 884, 1028).

Plyakha, Y., R. Uppal, and G. Vilkov. 2021. Equal or value weighting? Implications for asset-pricing tests. In *Risk management and modeling,* edited by R. Benkraiem, I. Kalaitzoglou, and C. Zopounidis. Springer. (Cited on pages 510, 779).

Scherer, B. 2002. Portfolio resampling: Review and critique. *Financial Analysts Journal* 58 (6): 98–109. (Cited on page 459).

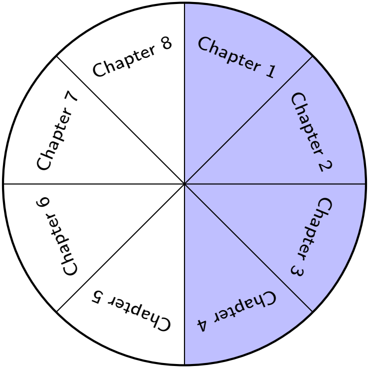
———. 2011. A note on the returns from minimum variance investing. *Journal of Empirical Finance* 18 (4): 652–660. (Cited on page 509).

Scheuch, C., S. Voigt, P. Weiss, and C. Frey. 2024. *Tidy finance with Python.* Chapman-Hall/CRC. Available online from this link. (Cited on pages 100, 399, 559).

Bibliography ... VIII

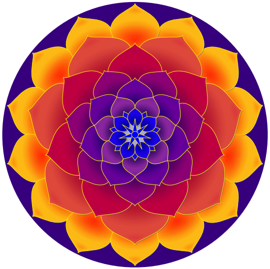
Stein, C. 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proceedings of the Third Berkeley Symposium.* (Cited on page 466).

Uppal, R., and T. Wang. 2003. Model misspecification and underdiversification. *Journal of Finance* 58 (6): 2465–2486. (Cited on page 459).



End of Chapter 4

Quantitative Portfolio Management



Chapter 5:
CAPM-Based Portfolios: Black-Litterman Model

Raman Uppal

2025

The big picture: Plan for the entire course

Table of contents

Road map

What do we want to do in Chapter 5

In this chapter and the next two chapters, we study how *asset-pricing factor models* can be used to build portfolios that perform well out of sample.

Today, we study the Black-Litterman model, which is based on the Capital Asset Pricing Model.

In the next two chapters, we will study parametric portfolio policies, which are based on factor asset-pricing models, such as the ones developed by Fama and French.

Road map

Error in estimating $\mathbb{E}[R]$ and $\mathbb{V}[R]$ from sample moments

- In the last-to-last chapter, we saw that mean-variance weights depend on estimates of return means and covariances:

$$w = \frac{1}{\gamma}(\mathbb{V}[R])^{-1}(\mathbb{E}[R] - R_f 1_N).$$

- In the last chapter, we saw that
  - Sample-based estimates of $\mathbb{E}[R]$ are imprecise;
  - Sample-based estimates of $\mathbb{V}[R]$ are ill-conditioned;
  - Consequently, mean-variance portfolios perform poorly out of sample.
- We then studied shrinkage methods to improve the properties of the sample estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$.

Shrinkage using only sample moments

- The shrinkage methods we studied relied only on sample moments but did not take advantage of asset-pricing theory.
- For example, Bayesian shrinkage of expected returns relies on shrinking sample estimate of $\mathbb{E}[R]$ toward a "grand mean," which is
  - either the average of all mean returns
  - or the expected return on the GMV portfolio.
- Similarly, the Ledoit and Wolf methods relies on shrinking the sample estimate of $\mathbb{V}[R]$ toward
  - either a diagonal matrix with the average variance on its diagonal
  - or a matrix where all the cross-asset correlations are replaced by the average correlation.

Empirical performance of sample-based models . . . I

- When we evaluated the empirical performance of models relying only on sample moments of returns, we saw that
  - models shrinking sample estimates of expected returns fail to outperform the simple $1/N$ benchmark;
  - models that ignore expected returns altogether and choose weights based on minimization of portfolio variance perform better,
    - especially when shrinking the covariance matrix of returns using either a short-sale constraint or the Ledoit and Wolf approach,
  - however, even these models do not always outperform the simple $1/N$ portfolio.

Empirical performance of sample-based models . . . II

- For the empirical performance of models based on shrinkage of sample moments, we studied the evidence reported in
  - DeMiguel, Garlappi, and Uppal (2009).
- But several other papers confirm that optimizing models based on sample moments performs poorly.
- Some of these papers are described on the next few slides.

Empirical performance of sample-based models . . . III

- Jacobs, Müller, and Weber (2014)
  - Take the perspective of a Euro investor
  - Extend the data period: 1973 to 2013
  - Extend the analysis across countries
  - Extend the analysis across asset classes (stocks, bonds, and commodities)
  - Extend the list of models of optimal portfolio selection studied
  - Extend the performance-evaluation metrics.

Empirical performance of sample-based models . . . IV

- Jacobs, Müller, and Weber (2014) find that:

  "Analyzing more than 5,000 heuristics, our results show that in fact almost any form of well-balanced allocation over asset classes offers similar diversification gains as even recently developed portfolio optimization approaches."

- Jacobs, Müller, and Weber (2014) conclude that:
  - Estimation error leads to poor performance of "optimal" models, relative to fixed-weight portfolios.

Empirical performance of sample-based models . . . V

- However, it is possible to find more sophisticated optimizing portfolios that do outperform the $1/N$ portfolio.
- For example, Ao, Li, and Zheng 2019
  - use machine-learning methods (in particular, lasso constraints) to
  - find shrinkage portfolios that outperform $1/N$ out of sample.
- Another example, which we will study in the last chapter, is the paper by Raponi, Uppal, and Zaffaroni (2023), which
  - shows how to build portfolios that outperform the $1/N$ portfolio
  - by exploiting the compensation for bearing unsystematic risk.

Ideas for Master's projects

- As a Master's project, you could extend the data until 2024 in the paper by Jacobs, Müller, and Weber (2014)
- As a Master's project, you could
  - reproduce the results in Ao, Li, and Zheng (2019), and
  - extend the analysis in Ao, Li, and Zheng (2019) along various dimensions, as in Jacobs, Müller, and Weber (2014).
- As a Master's project, you could
  - reproduce the results in Raponi, Uppal, and Zaffaroni (2023)
  - extend the analysis in Raponi, Uppal, and Zaffaroni (2023) to new asset classes (beyond equities) and along other dimensions, as in Jacobs, Müller, and Weber (2014).

End of discussion of
portfolios based on *sample-moments of returns*

. . . .

Start of discussion of
portfolios based on *asset-pricing models*

Shrinkage using asset-pricing factor models . . . I

- So far, in the shrinkage models we studied, we did not use any information from asset-pricing models.
- We will now study two "shrinkage" methods based on asset-pricing factor models.
  1. The first is based on the Capital Asset Pricing Model of Sharpe (1964) . . . which we will study in this chapter.
  2. The second can be interpreted as being based on the models of Fama and French (1992, 1993, 2012, 2015, 2018) . . . which we will study in the next chapter.
- After that, in the last two chapters of our course, we will see how to go beyond Fama-French models.

Shrinkage using asset-pricing factor models . . . II

1. The first asset-pricing-based model for portfolio choice we study is the famous Black-Litterman model developed at Goldman Sachs, described in a series of papers,
   - Black and Litterman (1990, 1991a, 1991b, 1992); the 1992 paper is available from this link.
   - for the intuition underlying the Black-Litterman model, see He and Litterman (1999), which is available from this link;
   - and, for a historical perspective, see the book by Litterman (2003).
2. The second asset-pricing-based model for portfolio choice we will study (in the next class) is the parametric portfolio policy of Brandt, Sant-Clara, and Valkanov (2009).

- Both models are based on very clever insights, are widely used in industry, and are straightforward to implement using Python.

Road map

Road map

First motivation for Black-Litterman model

- Recall that the solution to the Markowitz problem is given by:

$$w = \frac{1}{\gamma} \left(\mathbb{V}[R]\right)^{-1} \left(\mathbb{E}[R] - R_f 1_N\right),$$

  where $\mathbb{E}[R]$ denotes the $N \times 1$ vector of expected returns and $\mathbb{V}[R]$ is the $N \times N$ the variance-covariance matrix for returns.

- In the Markowitz problem, $\mathbb{E}[R]$ and $\mathbb{V}[R]$ are typically estimated using historical sample data.

- As we have seen, the sample-based Markowitz portfolio performs very poorly out-of-sample.

- This is the first motivation for the Black-Litterman model: how to obtain better estimates of expected returns using an asset-pricing model, the CAPM.

Portfolio choice using the CAPM

- The Capital Asset Pricing Model (CAPM) says that
  - every investor should hold the same portfolio of risky assets,
  - this common portfolio then must be the market portfolio.
- From the CAPM-implied market portfolio weights, we will obtain estimates of expected returns (as shown on the next slide).
  - Fischer Black adored the CAPM model.
  - You can read about this in his biography, "Fischer Black and the Revolutionary Idea of Finance" link to book on Amazon.

Expected returns implied by CAPM weights (without proof)

- Start with the Markowitz mean-variance optimal solution:

$$w_{\text{markowitz}} = \frac{1}{\gamma}\mathbb{V}[R_{\text{sample}}]^{-1}(\mathbb{E}[R_{\text{sample}}] - R_f 1_N) \quad \text{... from our earlier class.}$$

- Apply Markowitz to the Market (mkt) portfolio:

$$w_{\text{mkt}} = \frac{1}{\gamma_{\text{mkt}}}\mathbb{V}[R_{\text{sample}}]^{-1}(\mathbb{E}[R_{\text{capm}}] - R_f 1_N)$$

- Get CAPM-implied expected excess returns from market weights:

$$\underbrace{\mathbb{E}[R_{\text{capm}}]}_{N\times 1} - R_f 1_N = w_{\text{mkt}}\gamma_{\text{mkt}}\,\mathbb{V}[R_{\text{sample}}] \quad \text{... we will see how to eliminate } \gamma_{\text{mkt}}$$

Second motivation for the Black-Litterman model

- Imagine that you work for a quant fund.
- A client walks in and, based on asset-pricing theory (CAPM), you recommend that the client should hold the market portfolio.
- The client says, "but I have views about some of the asset returns."
- Question:
  How exactly should you tilt the client's portfolio away from the market portfolio in response to the client's views?
- It is not obvious how to adjust the weights of assets.
  - Because asset returns are correlated, changing the weight on one asset will also change the other weights.
- The Black-Litterman model tells us how to incorporate the client's views, which is the second motivation for this model.

Deviating from the market portfolio based on "views"

- The client could have absolute views:
  - Apple's return will be 2% below its historical mean.
  - Intel's return will be 4% above its historical mean.
- Or, the client could have relative views:
  - Meta will underperform Google by 5%.
  - NVIDIA will outperform both Apple and Intel by 3%.
- The confidence (inverse of variance) about each absolute and relative view could differ, so we need to account for this also.
  - The more confident the investor is about a particular view, the smaller the variance of that view.

Expressing Absolute and Relative Views

- The investor can have K views, denoted by the $K \times 1$ vector q, about the N expected returns, μ.
- These K views can be expressed as a linear combination of returns, through a "pick" matrix P as follows:

$$P\,\mu = q + \epsilon_q, \quad \text{where:} \quad \epsilon_q \sim \mathcal{N}(0_K, V_{\epsilon_q}), \tag{27}$$

where
  - P is $K \times N$ "pick" matrix;
  - μ is an $N \times 1$ vector of investor's expected returns;
  - q is $K \times 1$ vector of views about future absolute or relative returns;
  - ϵ_q is $K \times 1$ vector of "errors" with multivariate normal distribution,
    - with a mean given by the $K \times 1$ vector of 0_K, and
    - covariances given by the $K \times K$ matrix V_{ϵ_q}.

Example of views expressed via P matrix ... I

- Suppose that the market has only $N = 5$ securities.
- Suppose that the investor has the following $K = 3$ views:
  1. An absolute view that the return on Security 1 will be 10%, with a view-variance of 9%.
  2. An absolute view that the return on Security 2 will be 8%, with a view-variance of 4%.
  3. A relative view that the return on Security 3 will exceed that on Security 4 by 2%, with a view-variance of 1%.

  - No absolute view about Security 3.
  - No absolute or relative view about Security 5.

Example of views expressed via P matrix . . . II

- To express these $K = 3$ views using Equation (27), which is:

$$P\,\mu = q + \epsilon_q, \quad \text{where:} \quad \epsilon_q \sim \mathcal{N}(0, V_{\epsilon_q}), \tag{27}$$

- we specify the following:

$$\begin{array}{ccccccc} \textit{pick matrix} & \times & \textit{means} & = & \textit{views} & + & \textit{errors} \\ \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}}_{\substack{P \\ K \times N}} & \times & \underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{bmatrix}}_{\substack{\mu \\ N \times 1}} & = & \underbrace{\begin{bmatrix} 0.10 \\ 0.08 \\ 0.02 \end{bmatrix}}_{\substack{q \\ K \times 1}} & + & \underbrace{\begin{bmatrix} \epsilon_{q,1} \\ \epsilon_{q,2} \\ \epsilon_{q,3} \end{bmatrix}}_{\substack{\epsilon_q \\ K \times 1}} \end{array}$$

- and the $K \times K$ diagonal matrix of "view variances":

$$V_{\epsilon_q} = \begin{bmatrix} 0.09 & 0 & 0 \\ 0 & 0.04 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}.$$

Blend/condition expected returns from CAPM with investor's views (without proof)

- Blend (i.e., condition) the CAPM-implied expected returns with the investor's views to get
  - Expected returns, conditional on views: $\mathbb{E}[R_{\text{capm}}|\text{views}]$, and
  - Return-covariance matrix, conditional on views: $\mathbb{V}[R_{\text{sample}}|\text{views}]$.
- Compute the Black-Litterman (BL) weights using blended views:

$$w_{\text{BL}} = \frac{1}{\gamma}\mathbb{V}[R_{\text{sample}}|\text{views}]^{-1}(\mathbb{E}[R_{\text{capm}}|\text{views}] - R_f 1_N).$$

Road map

Main advantages of Black-Litterman model

- The main advantages of the Black-Litterman model are that
  - You can provide views on only a subset of assets and the model will make the correct adjustments for the covariance with other assets.
  - You can provide confidence (inverse of variance) about your views, which will be reflected in the degree of shrinkage.
  - Using Black-Litterman posterior returns leads to more reasonable portfolios than those from using sample moments of returns.

The importance of the Black-Litterman model

- The He and Litterman (1999, page 13) article says the following:
  - In the Quantitative Strategies group at Goldman Sachs Asset Management, we develop quantitative models to manage portfolios.
  - The Black-Litterman model is the central framework for our modeling process.
  - Our process starts with finding a set of views that are profitable.
    - For example, it is well known that portfolios based on certain value and momentum factors are consistently profitable.

      Note that Mark Carhart, who developed the momentum factor, worked at Goldman Sachs Asset Management; Wikipedia link.
    - We forecast returns for portfolios incorporating these factors and construct a set of views.
  - The Black-Litterman model takes these views and constructs a set of expected returns on each asset.

Relation/link to what we did in the last chapter

- In the last chapter, we dealt with unreasonable sample estimates of asset-return moments (means and covariances)
  - by shrinking them
  - toward a reasonable "value."
- In the Black-Litterman model, we are going to
  - start with a reasonable portfolio, the market portfolio, and
  - tilt away from this portfolio based on an investor's views.

Is the Black-Litterman model a Bayesian model?

- The Black and Litterman (1990, 1992) model is a combination of a model-based (CAPM) and view-based approach.
- Strictly speaking, the Black-Litterman model is not a Bayesian model.
  - A Bayesian model combines a prior with the data;
  - The Black-Litterman model combines the CAPM with a view.
- The similarity with the Bayesian approach is that
  - the equation for combining the CAPM with the investor's views
  - is the same as that for updating the prior with the data in the Bayesian approach.

Road map

The Black-Litterman model in two steps . . . I

- We will study the Black-Litterman model in two steps.
  - Both steps are simple.
  - Only the notation is cumbersome.

1. How to back out CAPM-implied expected returns?
2. From those returns, how to get
   2.1 $\mathbb{E}[R_{\text{capm}}|\text{views}]$, i.e., CAPM-implied expected returns conditional on the investor's views;
   2.2 $\mathbb{V}[R_{\text{sample}}|\text{views}]$, i.e., sample covariance matrix conditional on the investor's views.

The Black-Litterman model in two steps . . . II

- To help bring the theory to life, we will then use an example from a paper by Idzorek (2007), which you can download from this link.
  - Note that there are some minor differences in the theory presented in the Idzorek (2007) paper and what we will do in class.
  - The key difference is how the confidence (variance) about the views is specified.
- Finally, the next assignment will ask you to compute the Black-Litterman portfolio for a given set of views.

Notation used in most Black-Litterman papers/websites

Symbol	What it represents
Returns are random	
$\mu = \mathbb{E}(R)$	$N \times 1$ vector of expected (excess) returns
V_{ϵ_R}	$N \times N$ sample covariance matrix of asset-return residuals
V_R	$N \times N$ sample covariance matrix of asset returns
Means of returns are also random	
$\Pi = \mathbb{E}[R_{\text{capm}}]$	$N \times 1$ vector of prior views of $\mu =$ expected returns from CAPM
$\tau\Sigma = V_\mu$	Variance of mean returns (μ)
τ	(scalar) tuning constant, usually less than 1
Views are also uncertain	
q	$K \times 1$ vector of views, where $K \leq N$
P	$K \times N$ pick matrix which captures views about the N assets
$\Omega = V_{\epsilon_q}$	$K \times K$ uncertainty matrix of views

Explaining our notation . . . I

- When discussing the Black-Litterman model, we will be dealing with several random variables. Thus, we need to be careful with notation.
  - The returns, R, have two components, both of which are random:

$$R = \mu + \epsilon_R, \quad \text{where}$$

    - The residual is random, with distribution: $\epsilon_R \sim \mathcal{N}(0, V_{\epsilon_R})$.
    - The mean returns, μ, are also random, with covariance matrix V_μ.
    - Thus, $R \sim \mathcal{N}(\mu, V_\mu + V_{\epsilon_R})$.
  - The investor has views that have errors, ϵ_q; the covariance matrix of the errors is denoted by V_{ϵ_q}.
  - The views of the investor are uncertain, with covariance matrix V_q.

Explaining our notation . . . II

- We will design notation that is slightly different from what is standard, but one that is easier to understand.

- In particular, all covariance matrices are denoted by $\mathbb{V}[x] = V_x$.

- So
  - if returns are: $R = \mu + \epsilon_R$, with both μ and ϵ_R random, and
  - if views are: $q = P\,\mu - \epsilon_q$, (P is a matrix of constants), then:

Our notation	Equal to (using our notation)	Equal to (using notation used by others)
V_{ϵ_R}	—	Σ
V_μ	$\tau\, V_{\epsilon_R}$	$\tau\Sigma$
V_R	$V_\mu + V_{\epsilon_R}$	$\tau\Sigma + \Sigma = (1+\tau)\Sigma$
V_{ϵ_q}	—	Ω
V_q	$PV_\mu P^\top + V_{\epsilon_q}$	$P(\tau\Sigma)P^\top + \Omega.$

The two steps for deriving the Black-Litterman result

- We now study how to
  1. Back out the CAPM-implied expected returns;
  2. From those returns, get $\mathbb{E}[R_{\text{capm}}|\text{views}]$ and $\mathbb{V}[R_{\text{sample}}|\text{views}]$, and use these to solve the Markowitz mean-variance problem.

The steps that summarize the Black-Litterman model

- The explanation of the Black-Litterman model is divided into the steps below.

Step 1.1 Back out expected returns from the CAPM.

Step 1.2 Obtain aggregate risk aversion of the market, γ_{mkt}.

. . . .

Step 2.1 Start with a prior distribution for expected returns; that is,
- i.e., assume that the vector of mean returns is itself random,
- with the mean of the mean given by expected returns from CAPM.

Step 2.2 Specify subjective views of the investor regarding expected returns.

Step 2.3 Use Bayes' Theorem to condition the prior distribution of expected returns on the investor's views to obtain its posterior distribution.

Step 2.4 From the posterior distribution of *expected returns*, get the posterior mean and variance of returns.

Step 2.5 Use the posterior mean and variance of returns, conditional on the investor's views, to solve for the mean-variance optimal portfolio.

Step 1.1 of Black-Litterman model

Back out expected returns from the CAPM.

Step 1.1: Get expected returns from CAPM

- Recall optimality condition for any mean-variance efficient portfolio:

$$w = \frac{1}{\gamma} V_R^{-1} \left(\mathbb{E}[R] - R_f \, 1_N\right) \qquad \text{. . . condition for optimality;}$$

  Then, if the CAPM is true, the above equation for the market-portfolio weights is

$$w_{\text{mkt}} = \frac{1}{\gamma_{\text{mkt}}} V_R^{-1} \left(\mathbb{E}[R_{\text{capm}}] - R_f \, 1_N\right) \qquad \text{. . . under the CAPM;}$$

$$\gamma_{\text{mkt}} V_R \, w_{\text{mkt}} = \left(\mu_{\text{capm}} - R_f \, 1_N\right) \qquad \text{. . . isolating expected return}$$

$$\mu_{\text{capm}} - R_f \, 1_N = \gamma_{\text{mkt}} V_R \, w_{\text{mkt}} \qquad \text{. . . rearranging LHS and RHS}$$

- where $\mu_{\text{capm}} \equiv \mathbb{E}[R_{\text{capm}}]$ is the $N \times 1$ vector with expected returns for the N risky assets according to the CAPM.
- Note that R_f, V_R, and w_{mkt} are observable from market data.
- But γ_{mkt} is not observable, so we need to identify this.

Step 1.2 of Black-Litterman model

Obtain aggregate risk aversion of the market, γ_{mkt}.

Step 1.2: Obtaining aggregate risk aversion of the market

- We start by re-writing the expression from the previous page:

$$\mu_{\text{capm}} = R_f \, 1_N + \gamma_{\text{mkt}} \, V_R \, w_{\text{mkt}} \tag{28}$$

then, multiplying both sides by $w_{\text{mkt}}^\top$, we get

$$\underbrace{w_{\text{mkt}}^\top \mu_{\text{capm}}}_{=\mu_{\text{mkt}}} = R_f \, \underbrace{w_{\text{mkt}}^\top 1_N}_{=1} \; + \; \gamma_{\text{mkt}} \, \underbrace{w_{\text{mkt}}^\top V_R \, w_{\text{mkt}}}_{=\sigma^2_{\text{mkt}}}$$

$$\mu_{\text{mkt}} = R_f + \gamma_{\text{mkt}} \, \sigma^2_{\text{mkt}} \quad \ldots \quad \text{definition of } \mu_{\text{capm}} \text{ and } \sigma^2_{\text{mkt}}$$

$$\gamma_{\text{mkt}} = \frac{\mu_{\text{mkt}} - R_f}{\sigma^2_{\text{mkt}}} \quad \ldots \quad \text{one expression for } \gamma_{\text{mkt}}$$

$$\gamma_{\text{mkt}} = \underbrace{\left(\frac{\mu_{\text{mkt}} - R_f}{\sigma_{\text{mkt}}} \right)}_{\text{SR}_{\text{mkt}}} \frac{1}{\sigma_{\text{mkt}}} \quad \ldots \quad \text{split denominator into two}$$

$$\gamma_{\text{mkt}} = \frac{\text{SR}_{\text{mkt}}}{\sigma_{\text{mkt}}} \quad \ldots \quad \text{another expression for } \gamma_{\text{mkt}}.$$

CAPM-implied expected returns from Steps 1.1 and 1.2 . . . I

- Putting together our results,

$$\mu_{\text{capm}} - R_f\, 1_N = \gamma_{\text{mkt}}\, V_R\, w_{\text{mkt}} \qquad \ldots \text{Step 1.1}$$

$$\gamma_{\text{mkt}} = \frac{\mu_{\text{mkt}} - R_f}{\sigma^2_{\text{mkt}}} = \frac{\text{SR}_{\text{mkt}}}{\sigma_{\text{mkt}}} \qquad \ldots \text{Step 1.2}$$

leads to the final expression for the CAPM-implied stock returns:

$$\mu_{\text{capm}} - R_f\, 1_N = \frac{\text{SR}_{\text{mkt}}}{\sigma_{\text{mkt}}} V_R\, w_{\text{mkt}} \qquad \ldots \text{in terms of observables.}$$

CAPM-implied expected returns from Steps 1.1 and 1.2 . . . II

- From the last expression on the previous page,

$$\mu_{\text{capm}} - R_f \, 1_N = \left(\frac{\text{SR}_{\text{mkt}}}{\sigma_{\text{mkt}}} \right) V_R \, w_{\text{mkt}},$$

- we see one can get the CAPM-implied expected excess returns, from
  - SR_{mkt}, the market Sharpe ratio;
  - σ_{mkt}, the volatility of the market;
  - V_R, the $N \times N$ sample-covariance matrix of returns;
  - w_{mkt}, the $N \times 1$ vector of market-capitalization weights.

End of Step 1 of the Black-Litterman model

Getting expected returns from CAPM

Step 2 of the Black-Litterman model

Combine expected returns from CAPM with the views of the investor

Step 2.1 of Black-Litterman model

Recognize that the mean returns themselves are random, and identify the distribution of mean returns

Step 2.1: Recognize that mean returns are random . . . I

- Just like the Markowitz model, we start with a vector of random returns;
- We specify the distribution of these returns as follows:

$$R = \mu + \epsilon_R, \quad \text{where}$$
$$\epsilon_R \sim \mathcal{N}\big(0, V_{\epsilon_R} = \Sigma\big);$$
$$R \sim \mathcal{N}\big(\mu, V_R\big)$$

- Black and Litterman assume that expected returns μ themselves are random, where the prior distribution of expected returns is

$$\mu \sim \mathcal{N}\big(\mu_{\mathsf{capm}}, V_\mu\big) \quad \ldots \text{i.e., prior mean is given by CAPM.}$$

Step 2.1: Recognize that mean returns are random . . . II

- Therefore, given returns

$$R = \mu + \epsilon_R,$$

- if μ is random, then the distribution of returns, R, is

$$R \sim \mathcal{N}\big(\mathbb{E}[R], V_R\big), \tag{29}$$

where

$$\mathbb{E}[R] = \mathbb{E}[\mu_{\mathsf{capm}} + \epsilon_R] = \mu_{\mathsf{capm}} \quad \text{. . . because } \mathbb{E}[\epsilon_R] = 0_N,$$

$$V_R = V_\mu + V_{\epsilon_R}, \quad \text{. . . assuming that } \mu \text{ and } \epsilon_R \text{ are independent.}$$

Step 2.1: Recognize that mean returns are random ... III

- Black and Litterman suggest that the mean of expected returns should be less variable than the returns themselves; therefore,

$$V_\mu = \tau V_{\epsilon_R}, \quad \text{where} \quad \tau \leq 1.$$

- In classical statistics, $\tau = 1/T^{\text{est}}$, where T^{est} is the number of observations used to estimate sample moments
  - So, if you are using 5 years of monthly observations, $T^{\text{est}} = 60$;
  - implying that τ is a small number, close to 0.
- Thus,

$$\begin{aligned} V_R &= V_\mu + V_{\epsilon_R}, && \ldots \quad \text{assuming that } \mu \text{ and } \epsilon_R \text{ are independent} \\ &= \tau\Sigma + \Sigma \\ &= (1+\tau)\Sigma \\ &\approx \Sigma && \ldots \quad \text{if } \tau = 1/T^{\text{est}} \text{ is small.} \end{aligned}$$

Step 2.2 of Black-Litterman model

Specify subjective views of investor regarding expected returns.

Step 2.2: Views of the investor and their distribution

- In the absence of views that differ from the implied equilibrium return, the investor should hold the market portfolio.
- The views of the investor may be expressed in
  - absolute terms:
    Asset i will have a return of 10%, with a variance of 50%;
  - relative terms:
    Asset n's return will exceed Asset m's, with a variance of 20%.
- The more confident the investor is about a particular view, the smaller the variance of that view.

Expressing Absolute and Relative Views

- The investor can have K views, $q \in \mathbb{R}^K$, about the N returns.
- These K views can be expressed as a linear combination of returns, through a "pick" matrix P as follows:

$$P\,\mu = q + \epsilon_q, \quad \text{where:}$$
$$q \sim \mathcal{N}\big(\mathbb{E}[q], V_q\big) \;\text{ and }\; \epsilon_q \sim \mathcal{N}(0, V_{\epsilon_q}),$$

  where
  - P is $K \times N$ matrix;
  - μ is $N \times 1$ vector of investor's expected returns;
  - q is $K \times 1$ vector of views about future absolute or relative returns;
  - ϵ_q is $K \times 1$ vector of errors with multivariate normal distribution,
    - with a mean given by the $K \times 1$ vector of 0_K, and
    - covariances given by the $K \times K$ matrix V_{ϵ_q}.

The Distribution of Views

- We wish to determine the distribution of the investor's views, q.
- We can obtain this from Equation (27), which is reproduced below:

$$P\,\mu = q + \epsilon_q, \tag{27}$$

- Re-arranging the terms in the equation above, we get:

$$q = P\,\mu - \epsilon_q, \quad \text{where,}$$
$$\epsilon_q \sim \mathcal{N}(0, V_{\epsilon_q}) \ \text{ and } \ q \sim \mathcal{N}\big(\mathbb{E}[q], V_q\big),$$

- From the expressions above, we get the distribution of views:

$$\mathbb{E}[q] = \mathbb{E}[P\,\mu - \epsilon_q] = P\,\mu \qquad \ldots \quad \mathbb{E}[\epsilon_q] = 0$$
$$V_q = \mathbb{V}[P\,\mu - \epsilon_q] = P\,V_\mu\,P^\top + V_{\epsilon_q} = P(\tau\Sigma)P^\top + \Omega,$$
$$\mathbb{C}[\mu, q] = \mathbb{C}[\mu, P\,\mu - \epsilon_q] = V_\mu\,P^\top \quad \ldots \mathbb{C}[\mu, \epsilon_q] = 0, \text{note: } \mathbb{C} \text{ denotes Cov}$$
$$\mathbb{C}[q, \mu] = \mathbb{C}[P\,\mu - \epsilon_q, \mu,] = P\,V_\mu \quad \ldots \mathbb{C}[\epsilon_q, \mu] = 0.$$

Step 2.3 of Black-Litterman model

Combine expected returns from CAPM with the views of the investor

Use Bayes' Theorem to condition the prior distribution of *expected returns* on the investor's views to obtain its posterior distribution

Step 2.3: Posterior distribution of *expected returns* . . . I

- From standard results about multivariate normal variables, we know
- the expectation of a vector of random variables Y conditional on a particular realization x of the random vector X is:

$$\mathbb{E}[Y|x] = \mathbb{E}[Y] + \underbrace{\mathbb{C}[Y,X]V_X^{-1}}_{\text{beta}}(x - \mathbb{E}[X]),$$

- with the conditional variance of Y given by

$$\mathbb{V}[Y|x] = \mathbb{V}[Y] - \underbrace{\mathbb{C}[Y,X]\,V_X^{-1}}_{\text{beta}}\mathbb{C}[X,Y].$$

Step 2.3: Posterior distribution of *expected returns* . . . II

- In our context,
  - $Y = \mu$, and
  - $X = q$.
- Thus, in our context, the conditional expectation of mean returns is

$$\mathbb{E}[\mu|q] = \mathbb{E}[\mu] + \mathbb{C}[\mu, q] V_q^{-1} \big(q - \mathbb{E}[q]\big), \tag{30}$$

with the conditional variance of μ given by

$$\mathbb{V}[\mu|x] = V_\mu - \mathbb{C}[\mu, q] \, V_q^{-1} \, \mathbb{C}[q, \mu]. \tag{31}$$

Step 2.3: Posterior distribution of *expected returns* . . . III

- In the context of the Black-Litterman model, $Y = \mu$, with:

$$\mathbb{E}[\mu] = \mu_{\text{capm}} \quad \text{. . . assumption that prior of mean is given by CAPM,}$$

$$V_\mu = \tau\, V_{\epsilon_R} = \tau\, \Sigma.$$

- Similarly, X corresponds to $q = P\,\mu - \epsilon_q$, with:

$$\mathbb{E}[q] = \mathbb{E}[P\,\mu - \epsilon_q] = P\,\mathbb{E}[\mu] = P\,\mu_{\text{capm}},$$

$$V_q = \mathbb{V}[P\,\mu - \epsilon_q] = P\,V_\mu\,P^\top + V_{\epsilon_q} = P\,(\tau\,\Sigma)\,P^\top + \Omega,$$

$$\mathbb{C}[q, \mu] = \mathbb{C}[P\,\mu - \epsilon_q, \mu] = P\,V_\mu = P\,(\tau\,\Sigma)$$

$$\mathbb{C}[\mu, q] = \mathbb{C}[\mu, P\,\mu - \epsilon_q] = V_\mu\,P^\top = (\tau\,\Sigma)\,P^\top.$$

Step 2.3: Posterior distribution of *expected returns* . . . IV

- Making these substitutions we get the conditional expectation of μ

$$\begin{aligned}
\mathbb{E}[\mu|q] &= \mathbb{E}[\mu_{\text{capm}}|\text{views}] \\
&= \mathbb{E}[\mu] + \mathbb{C}[\mu, q] V_q^{-1} (q - \mathbb{E}[q]), && (30) \\
&= \mu_{\text{capm}} + \left[(\tau \Sigma) P^\top \right] \left[P(\tau \Sigma) P^\top + \Omega \right]^{-1} (q - P \mu_{\text{capm}}) && (32)
\end{aligned}$$

with the conditional variance of μ

$$\begin{aligned}
\mathbb{V}[\mu|q] &= \mathbb{V}[\mu_{\text{capm}}|\text{views}] \\
&= V_\mu - \mathbb{C}[\mu, q] V_q^{-1} \mathbb{C}[q, \mu]. && (31) \\
&= \left[\tau \Sigma \right] - \left[(\tau \Sigma) P^\top \right] \left[P(\tau \Sigma) P^\top + \Omega \right]^{-1} \left[P (\tau \Sigma) \right]. && (33)
\end{aligned}$$

- **Note** that the results we have in equations (32) and (33) are for
  - the distribution of expected returns, μ,
  - while what we need is the distribution of returns, R.

Step 2.4 of Black-Litterman model

From the posterior distribution of expected returns,
get the posterior distribution of returns.

More precisely, use the posterior distribution of expected returns to obtain the posterior mean and posterior variance of returns (conditional on investor's views).

Step 2.4: Posterior (conditional) distribution of returns . . . I

- In the previous step, we have computed the posterior (conditional) distribution of expected returns.

- We now use that result for expected returns to derive the posterior (conditional) distribution of returns.

Step 2.4: Posterior (conditional) distribution of returns . . . II

- Recall also from Equation (29) on Page 661 that

$$R = \mu + \epsilon_R, \qquad \epsilon_R \sim \mathcal{N}\big(0, V[\epsilon_R]\big), \tag{29}$$

- which implies that conditional on the views q

$$\mathbb{E}[R|q] = \mathbb{E}[\mu|q] + \mathbb{E}[\epsilon_R|q] = \mathbb{E}[\mu|q], \quad \ldots \text{where } \mathbb{E}[\epsilon_R|q] = 0;\ \mathbb{E}[\mu|q] \text{ is in Eqn. (32)}$$

$$= \mu_{\text{capm}} + \left[(\tau\,\Sigma)\,P^\top\right]\left[P(\tau\,\Sigma)P^\top + \Omega\right]^{-1}(q - P\,\mu_{\text{capm}}), \tag{34}$$

$$\mathbb{V}[R|q] = V_{\epsilon_R} + \mathbb{V}[\mu|q], \quad \ldots \text{where } \mathbb{V}[\mu|q] \text{ is in (33) and } V_{\epsilon_R} = \Sigma$$

$$= \Sigma + \left[\tau\,\Sigma - \Big((\tau\,\Sigma)\,P^\top\Big)\Big(P(\tau\,\Sigma)P^\top + \Omega\Big)^{-1}\Big(P\,(\tau\,\Sigma)\Big)\right].$$

$$= (1+\tau)\,\Sigma - \Big((\tau\,\Sigma)\,P^\top\Big)\Big(P(\tau\,\Sigma)P^\top + \Omega\Big)^{-1}\Big(P\,(\tau\,\Sigma)\Big). \tag{35}$$

Interpreting expression for conditional mean of the return

- The expected return conditional on views can be interpreted as a weighted average of the model-return and the view-return.

$$\mathbb{E}[R|q] = \mathbb{E}[R|\text{views}]$$

$$= \underbrace{\mu_{\text{capm}}}_{\text{model mean}} + \underbrace{\left[(\tau\,\Sigma)\,P^\top\right]\left[P(\tau\,\Sigma)P^\top + \Omega\right]^{-1}}_{\text{weight on your view}} \underbrace{(q - P\,\mu_{\text{capm}})}_{\text{your views vs. model}}, \qquad (34)$$

which, after some algebra, can be written in another way:

$$= \left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\underbrace{\mathbb{E}[R_{\text{capm}}]}_{\text{model}} + P^\top\Omega^{-1}\underbrace{q}_{\text{views}}\right] \qquad (36)$$

which is the weighted average of the "model" and the "views":

$$= \underbrace{\left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}(\tau\Sigma)^{-1}}_{\text{weight on CAPM-implied return}}\underbrace{\mathbb{E}[R_{\text{capm}}]}_{\text{model}}$$

$$+ \underbrace{\left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}P^\top\Omega^{-1}}_{\text{weight on views}}\underbrace{q}_{\text{views}}.$$

Another way of writing the conditional variance of returns

- The expression for the conditional variance can also be simplified (for details of the derivation, see Meucci 2010):

$$\begin{aligned} \mathbb{V}[R|q] &= \mathbb{V}[R|\text{views}] \\ &= (1+\tau)\,\Sigma - \left((\tau\,\Sigma)\,P^\top\right)\,\left(P(\tau\,\Sigma)P^\top + \Omega\right)^{-1}\,\left(P\,(\tau\,\Sigma)\right). \qquad (35) \\ &= \Sigma + \left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}. \qquad (37) \end{aligned}$$

Step 2.5 of Black-Litterman model

Use the posterior mean and variance of returns, conditional on the investor's views, to solve for the Black-Litterman portfolio weights.

Black-Litterman portfolio weights . . . I

- We now know how to obtain the expected returns (and covariance matrix of returns)
  1. implied by the CAPM (Part 1),
  2. conditional on the views of the investor (Part 2).
- To find the optimal portfolio, we proceed just as before, but
  - instead of using the sampling distribution of returns,
  - we use the posterior distribution.

Black-Litterman portfolio weights . . . II

- That is, in contrast to Markowitz portfolio weights,

$$w_{\text{markowitz}} = \frac{1}{\gamma} \left(\mathbb{V}[R_{\text{sample}}]\right)^{-1} \left(\mathbb{E}[R_{\text{sample}}] - R_f 1_N\right),$$

- the Black-Litterman portfolio weights, w_{BL}, are given by:

$$w_{\text{BL}} = \frac{1}{\gamma} \left(\mathbb{V}[R|\text{q}]\right)^{-1} \left(\mathbb{E}[R|\text{q}] - R_f 1_N\right), \tag{38}$$

with

- $\mathbb{E}[R|\text{q}]$ defined in Equations (34) or (36), and
- $\mathbb{V}[R|\text{q}]$ defined in Equations (35) or (37).

Black-Litterman combined with constraints or shrinkage

- Note that the Black-Litterman approach can be combined with the methods we studied in the last class. For example:
  - When choosing the portfolio weights using mean-variance optimization with the Black-Litterman μ_{BL} and Σ_{BL}, one can impose portfolio constraints;
  - The covariance matrix used can be Σ_{BL} with Ledoit-Wolf shrinkage applied to it.

End of the Black-Litterman model with the detailed derivations

Road map

Example of Black-Litterman with one asset, one view . . . I

- To confirm our understanding and gain further intuition, we consider the special case of the Black-Litterman model with
  - one risky asset, and
  - one view about this asset.
- We show that expected returns conditional on views
  - are a weighted average of the returns from the CAPM model and the investors' views about returns;
  - with the weights depending on the relative confidence in the model and in the return views.

Example of Black-Litterman with one asset, one view . . . II

- For the special case of one asset and one view:
  - $V_{\epsilon_R} = \Sigma = \sigma^2_{\epsilon_R}$
  - $P = 1$
  - $V_{\epsilon_q} = \Omega = \sigma^2_{\epsilon_q}$.

Posterior mean of returns . . . I

- Recall the general result in Equation (32) that

$$\begin{aligned}\mathbb{E}[R|q] &= \mu_{\text{capm}} + \left[(\tau\, V_{\epsilon_R})\, P^\top\right] \left[P(\tau\, V_{\epsilon_R})P^\top + V_{\epsilon_q}\right]^{-1} (q - P\,\mu_{\text{capm}})\,. \\ &= \mu_{\text{capm}} + \left[(\tau\,\Sigma)\, P^\top\right] \left[P(\tau\,\Sigma)P^\top + \Omega\right]^{-1} (q - P\,\mu_{\text{capm}})\,. \qquad (32)\end{aligned}$$

- Making the substitutions for
  - $V_{\epsilon_R} = \Sigma = \sigma^2_{\epsilon_R}$
  - $P = 1$
  - $V_{\epsilon_q} = \Omega = \sigma^2_{\epsilon_q}$.

  in Equation (32) leads to the result on the next slide:

Posterior mean of returns . . . II

$$\mathbb{E}[R|q] = \mu_{\text{capm}} + \left(\tau\, \sigma^2_{\epsilon_R}\right) \left[\left(\tau\, \sigma^2_{\epsilon_R}\right) + \sigma^2_{\epsilon_q}\right]^{-1} \left(q - \mu_{\text{capm}}\right)$$

$$= \mu_{\text{capm}} \underbrace{\left[1 - \frac{\tau\, \sigma^2_{\epsilon_R}}{\tau\, \sigma^2_{\epsilon_R} + \sigma^2_{\epsilon_q}}\right]}_{\text{weight on model}} + q \underbrace{\left[\frac{\tau\, \sigma^2_{\epsilon_R}}{\tau\, \sigma^2_{\epsilon_R} + \sigma^2_{\epsilon_q}}\right]}_{\text{weight on view}} \quad \text{. . . collect terms}$$

$$= \mu_{\text{capm}} \underbrace{\left[\frac{\sigma^2_{\epsilon_q}}{\tau\, \sigma^2_{\epsilon_R} + \sigma^2_{\epsilon_q}}\right]}_{\text{weight on model}} + q \underbrace{\left[\frac{\tau\, \sigma^2_{\epsilon_R}}{\tau\, \sigma^2_{\epsilon_R} + \sigma^2_{\epsilon_q}}\right]}_{\text{weight on view}} \quad \text{. . . simplify first fraction}$$

$$= \mu_{\text{capm}} \underbrace{\left[\frac{\frac{1}{\tau\, \sigma^2_{\epsilon_R}}}{\frac{1}{\tau\, \sigma^2_{\epsilon_R}} + \frac{1}{\sigma^2_{\epsilon_q}}}\right]}_{\substack{\text{relative confidence}\\ \text{in model}}} + q \underbrace{\left[\frac{\frac{1}{\sigma^2_{\epsilon_q}}}{\frac{1}{\tau\, \sigma^2_{\epsilon_R}} + \frac{1}{\sigma^2_{\epsilon_q}}}\right]}_{\substack{\text{relative confidence}\\ \text{in view}}} \quad \text{. . . divide by } \tau\, \sigma^2_{\epsilon_R} \times \sigma^2_{\epsilon_q}$$

$$= \mu_{\text{capm}}\, \psi + q\,(1 - \psi) \quad \text{. . . weighted average of view and CAPM}$$

Road map

Example of Black-Litterman model: Numerical

- To check our understanding of the Black-Litterman model, we will now consider a numerical application of the model.
  - We will consider an example from the paper by Idzorek (2007), which you can download from this link.
- The next assignment will give you another opportunity to code the Black-Litterman model in Python.
  - Writing the code for the Black-Litterman model is simple, so I recommend you write the code yourself rather than using a package.
  - In my view, writing the code will be simpler than figuring out how to use a package written by someone else.
  - If you decide to use a package and find one that is simple to use, please let me know (so I can learn about the package).

Inputs needed for implementing Black-Litterman model

- Data from the market:
  - Risk-free rate, R_f.
  - Market-capitalization weights, w_{mkt}.
  - Volatility of the return on the market, $\sqrt{\mathbb{V}[R_{\text{mkt}}]}$.
  - Sharpe ratio of the return on the market, $\frac{\mathbb{E}[R_{\textbf{mkt}}]-R_f}{\sqrt{\mathbb{V}[R_{\textbf{mkt}}]}}$.
  - Variance-covariance matrix of sample returns, $V_{\epsilon_R} = \Sigma$.
- Subjective parameters:
  - Parameter to shrink variance-covariance matrix of sample returns, τ.
  - Matrix reflecting absolute and relative views, P.
  - Estimate of returns based on absolute and relative views, q_{views}.
  - Estimates of variance of the error in the views, $V_{\epsilon_q} = \Omega$.

Data provided to us . . . I

- Assume that $\tau = 0.025$ ($\tau \approx= 1/T^{\text{est}}$, where T^{est} is the number of data points used in the estimation of the return moments, so this corresponds to 40 years of annual observations.)
- Number of risky assets = $N = 8$, which are listed on the next page.
- All return data for the assets is in excess of the risk-free rate.
- The market
  - Sharpe ratio is 0.426169, and
  - the market return volatility is 14.0789%.

Data provided to us . . . II

- Historical sample excess mean returns (μ_{sample}) and market-capitalization weights, w_{mkt}.

#	Asset Class	μ_{sample}	w_{mkt}
1	US Bonds	3.15%	0.180409
2	Int'l Bonds	1.75%	0.268921
3	US Large Growth	−6.39%	0.119896
4	US Large Value	−2.86%	0.124435
5	US Small Growth	−6.75%	0.016023
6	US Small Value	−0.54%	0.010849
7	Int'l Dev. Equity	−6.75%	0.243523
8	Int'l Emerg. Equity	−5.26%	0.035942

Data provided to us . . . III

- Covariance matrix of excess returns, Σ is:

Asset Class	US Bonds	Intern Bonds	US Large Growth	US Large Value	US Small Growth	US Small Value	Intern Dev. Equity	Inter Emerg. Equity
US Bonds	0.00100	0.00132	-0.00057	-0.00067	0.00012	0.00012	-0.00044	-0.00043
Intern Bonds	0.00132	0.00727	-0.00130	-0.00061	-0.00223	-0.00098	0.00144	-0.00153
US Large Growth	-0.00057	-0.00130	0.05985	0.02758	0.06349	0.02303	0.03296	0.04803
US Large Value	-0.00067	-0.00061	0.02758	0.02960	0.02657	0.02146	0.02069	0.02985
US Small Growth	0.00012	-0.00223	0.06349	0.02657	0.10248	0.04274	0.03994	0.06599
US Small Value	0.00012	-0.00098	0.02303	0.02146	0.04274	0.03205	0.01988	0.03223
Intern Dev. Equity	-0.00044	0.00144	0.03296	0.02069	0.03994	0.01988	0.02835	0.03506
Inter Emerg. Equity	-0.00043	-0.00153	0.04803	0.02985	0.06599	0.03223	0.03506	0.07995

Data provided to us . . . IV

- The investor has three views, the first absolute, the others relative.
  - View 1: International Developed Equity will have an absolute excess return of 5.25% (with a view variance of = 0.000709).
  - View 2: International Bonds will outperform US Bonds by 25 basis points (view variance = 0.000141).
  - View 3: US Large Growth and US Small Growth will outperform US Large Value and US Small Value by 2% (view variance = 0.000866).
- Thus,

$$q + \epsilon_q = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{q_1} \\ \epsilon_{q_2} \\ \epsilon_{q_3} \end{bmatrix} = \begin{bmatrix} 5.25 \\ 0.25 \\ 2.00 \end{bmatrix} + \begin{bmatrix} \epsilon_{q_1} \\ \epsilon_{q_2} \\ \epsilon_{q_3} \end{bmatrix}$$

- And, the $K \times N$ pick matrix corresponding to the $K = 3$ views for the $N = 8$ assets is

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5 & 0.5 & -0.5 & 0 & 0 \end{bmatrix}$$

Data provided to us . . . V

- We now need to specify the matrix Ω, which captures the uncertainty in the views.

$$\Omega = \begin{bmatrix} \Omega_{11} & 0 & 0 \\ 0 & \Omega_{22} & 0 \\ 0 & 0 & \Omega_{33} \end{bmatrix} = \begin{bmatrix} 0.000709 & 0 & 0 \\ 0 & 0.000141 & 0 \\ 0 & 0 & 0.000866 \end{bmatrix}$$

- These numbers are obtained from $P_k(\tau\Sigma)P_k^\top$, where $k = \{1, 2, 3\}$ is the k-th row of matrix P.

- I have used one method for specifying the Ω matrix, but there are others (and you may wish to use one of the other methods).

We are ready to start analyzing the data given to us.

Numerical example: Markowitz portfolio weights

- We start by computing the Markowitz portfolio weights (w_{MVU}) for $\gamma_{mkt} = 3.0271189$.

$$w_{\text{MVU}} = \frac{1}{\gamma}\Sigma^{-1}\mu_{\text{sample}}$$

Asset Class	μ_{sample} %	w_{mkt} %	w_{MVU} %
US Bonds	3.15	18.04	1161.24
Int'l Bonds	1.75	26.89	−106.56
US Large Growth	−6.39	11.98	56.00
US Large Value	−2.86	12.44	−5.72
US Small Growth	−6.75	1.60	−61.61
US Small Value	−0.54	1.08	82.81
Int'l Dev. Equity	−6.75	24.35	−105.48
Int'l Emerg. Equity	−5.26	3.59	14.78

- It is clear from the last column that the mean-variance weights based on historical estimates of the sample mean are not reasonable.

Numerical example: Implied CAPM excess returns

- Next, we compute the expected (excess) returns implied by the CAPM, using the expression:

$$w_{mkt} = \frac{1}{\gamma_{mkt}} \Sigma^{-1} \mu_{\text{capm}} \quad \text{which implies} \quad \mu_{\text{capm}} = \gamma_{mkt} \, \Sigma \, w_{mkt}$$

Asset Class	μ_{sample} %	w_{mkt} %	w_{MVU} %	μ_{capm} %
US Bonds	3.15	18.04	1161.24	0.08
Int'l Bonds	1.75	26.89	−106.56	0.67
US Large Growth	−6.39	11.98	56.00	6.41
US Large Value	−2.86	12.44	−5.724	4.08
US Small Growth	−6.75	1.60	−61.61	7.43
US Small Value	−0.54	1.08	82.81	3.70
Int'l Dev. Equity	−6.75	24.35	−105.48	4.80
Int'l Emerg. Equity	−5.26	3.59	14.78	6.60

- Clearly, the CAPM-implied expected excess returns seem much more reasonable than the sample-based historical mean returns, μ_{sample}.

Numerical example: Posterior mean returns . . . I

- We now combine the CAPM-implied expected excess returns with the investor's views to obtain the posterior expected excess returns:

$$\begin{aligned}\mu_{\mathrm{BL}} &= \mathbb{E}[R_{\mathrm{capm}}|\text{views}] - R_f 1_N \\ &= \left[(\tau\Sigma)^{-1} + P^\top \Omega^{-1} P\right]^{-1} \left[(\tau\Sigma)^{-1} \underbrace{\mu_{\mathrm{capm}}}_{\text{model}} + P^\top \Omega^{-1} \underbrace{q}_{\text{views}}\right]. \qquad (36)\end{aligned}$$

Asset Class	μ_{sample} %	w_{mkt} %	w_{MVU} %	μ_{capm} %	μ_{BL} %
US Bonds	3.15	18.04	1161.24	0.08	0.0621
Int'l Bonds	1.75	26.89	−106.56	0.67	0.5036
US Large Growth	−6.39	11.98	56.00	6.41	6.2827
US Large Value	−2.86	12.44	−5.72	4.08	4.3383
US Small Growth	−6.75	1.60	−61.61	7.43	7.2545
US Small Value	−0.54	1.08	82.81	3.70	3.9105
Int'l Dev. Equity	−6.75	24.35	−105.48	4.80	4.8576
Int'l Emerg. Equity	−5.26	3.59	14.78	6.60	6.6881

Numerical example: Posterior return covariance matrix

- Similarly, the posterior covariance matrix of returns is:

$$\begin{aligned}\Sigma_{\mathrm{BL}} &= \mathbb{V}[R_{\mathrm{sample}}|\mathrm{views}] \\ &= \Sigma + \left[(\tau\Sigma)^{-1} + P^{\top}\Omega^{-1}P\right]^{-1}. \end{aligned} \quad (37)$$

$$\Sigma_{\mathrm{BL}} = \begin{bmatrix} 0.00102 & 0.00135 & -0.00058 & -0.00068 & 0.00013 & 0.00013 & -0.00045 & -0.00044 \\ 0.00135 & 0.00738 & -0.00133 & -0.00063 & -0.00226 & -0.00100 & 0.00144 & -0.00156 \\ -0.00058 & -0.00133 & 0.06061 & 0.02801 & 0.06415 & 0.02331 & 0.03330 & 0.04857 \\ -0.00068 & -0.00063 & 0.02801 & 0.03015 & 0.02692 & 0.02181 & 0.02096 & 0.03030 \\ 0.00013 & -0.00226 & 0.06415 & 0.02692 & 0.10385 & 0.04344 & 0.04034 & 0.06681 \\ 0.00013 & -0.00100 & 0.02331 & 0.02181 & 0.04344 & 0.03267 & 0.02013 & 0.03272 \\ -0.00045 & 0.00144 & 0.03330 & 0.02096 & 0.04034 & 0.02013 & 0.02868 & 0.03546 \\ -0.00044 & -0.00156 & 0.04857 & 0.03030 & 0.06681 & 0.03272 & 0.03546 & 0.08131 \end{bmatrix}$$

Numerical example: Black-Litterman portfolio weights

- We can use our estimates of μBL and Σ_{BL} to compute the Black-Litterman portfolio weights

$$w_{\text{BL}} = \frac{1}{\gamma_{mkt}} \Sigma_{\text{BL}}^{-1} \mu_{\text{BL}}.$$

Asset Class	μ_{sample} %	w_{mkt} %	w_{MVU} %	μ_{capm} %	μ_{BL} %	w_{BL} %
US Bonds	3.15	18.04	1161.24	0.08	0.0621	28.35
Int'l Bonds	1.75	26.89	−106.56	0.67	0.5036	15.48
US Large Growth	−6.39	11.98	56.00	6.41	6.2827	9.07
US Large Value	−2.86	12.44	−5.72	4.08	4.3383	14.76
US Small Growth	−6.75	1.60	−61.61	7.43	7.2545	−1.05
US Small Value	−0.54	1.08	82.81	3.70	3.9105	3.68
Int'l Dev. Equity	−6.75	24.35	−105.48	4.80	4.8576	28.85
Int'l Emerg. Equity	−5.26	3.59	14.78	6.60	6.6881	3.50

Road map

Python code for the Black-Litterman model

- Python code for the Black-Litterman model is available from:
  - PyPortfolioOpt.
  - Luís Fernando Torres.
  - Python for Finance.
  - Robert Martin
  - Cardiel - A portfolio allocation tool based on Black-Litterman with very nice visualization of the portfolio weights.
- My advice to you is to write your own code – you will learn a lot more from writing the code, which is straightforward to do.

Road map

1. Overview of this chapter
2. Motivation for material in this chapter
3. Black-Litterman (BL) model
4. Python code for the Black-Litterman model
5. To do for next class: Readings and assignment
6. Bibliography

What we plan to do in the next chapter

What's next?

In the next chapter, we will study parametric portfolio policies developed by Brandt, Santa-Clara, and Valkanov (2009).

To do for next class: Readings

- Readings
  - To get just the intuition for the Black-Litterman model, you can read He and Litterman (1999), which is available from this link.
    - The main text of the article offers a non-mathematical discussion;
    - The maths underlying the model is in Appendix B of the article.
  - To read a well-written description of the Black-Litterman model, I recommend Meucci (2010), which provides a careful and detailed analysis. The article can be downloaded from this link.

Road map

Bibliography . . . I

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies* 32 (7): 2890–2919. (Cited on pages 622, 623, 1118–1120).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

———. 1991a. Combining investor views with market equilibrium. *Journal of Fixed Income* 1 (2): 7–18. (Cited on pages 626, 732, 884, 1028).

———. 1991b. Global asset allocation with equities, bonds, and currencies. *Fixed Income Research* 2 (15-28): 1–44. (Cited on pages 626, 732, 884, 1028).

———. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Brandt, M. W., P. Sant-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. *Review of Financial Studies:* Forthcoming. (Cited on pages 626, 726, 739, 854).

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. (Cited on pages 450, 495, 559, 619, 724, 1120).

Bibliography . . . II

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2 (June): 427–465. (Cited on pages 625, 726, 733, 759, 760, 885, 1029).

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105 (3): 457–472. (Cited on pages 625, 726, 733, 885, 1029).

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

———. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

He, G., and R. Litterman. 1999. The intuition behind Black-Litterman model portfolios. *Investment Management Research (Goldman, Sachs & Company).* (Cited on pages 626, 640, 705, 732, 884, 1028).

Idzorek, T. 2007. A step-by-step guide to the Black-Litterman model: incorporating user-specified confidence levels. In *Forecasting expected returns in the financial markets,* 17–38. Elsevier. (Cited on pages 645, 688).

Bibliography . . . III

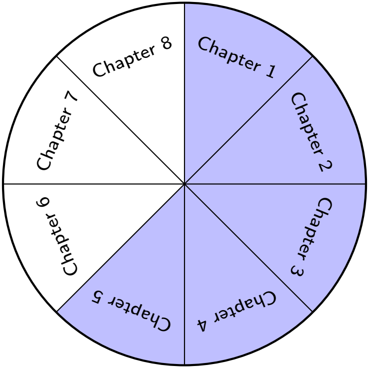
Jacobs, H., S. Müller, and M. Weber. 2014. How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets* 19:62–85. (Cited on pages 495, 620, 621, 623, 724).

Litterman, R. 2003. *Modern investment management: An equilibrium approach.* New York: Wiley. (Cited on pages 626, 732, 884, 1028).

Meucci, A. 2010. The Black-Litterman approach: Original model and extensions. Available at SSRN. (Cited on pages 676, 705).

Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary. (Cited on pages 622, 623, 734, 886, 1030, 1111, 1155).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).



End of Chapter 5

Blank page 1

Blank page 2

Quantitative Portfolio Management

Chapter 6:
Factor-based Portfolios: Parametric Portfolio Policies

Raman Uppal

2025

The big picture: Plan for the entire course

Part A: Preliminaries

Chapter 1: Properties of asset returns

Chapter 2: Performance measurement (especially out of sample)

Part B: Modern portfolio management

Chapter 3: Mean-variance portfolios that *ignore* estimation error

Chapter 4: Mean-variance portfolios that *adjust for* estimation error

Part C: Post-Modern Portfolio Management

Chapter 5: CAPM-based portfolios: Black-Litterman model

Chapter 6: Factor-based portfolios: Parametric portfolio policies

Chapter 7: Volatility-timed factor portfolios

Chapter 8: Portfolios exploiting systematic risk factors *and* unsystematic risk

Table of contents

Road map

What do we want to do in Chapter 6

In this chapter, we study how asset-pricing factor models can be used to construct optimal portfolios.

We study the use of factor models for estimating both expected returns and the covariance matrix.

We conclude by studying the "parametric portfolio policies" of Brandt, Santa-Clara, and Valkanov.

Road map

Motivation for the material in this chapter

- We start by explaining why asset-pricing factor models may be important.
- First we explain the problems in implementing Markowitz mean-variance portfolios
  - by now, you should understand this perfectly.
- Then, we explain how factor models can help resolve the problems
  - for estimating expected returns
  - for obtaining well-conditioned covariance matrices.
- which allows us to construct portfolios that perform well out-of-sample.

Distinction between factor models and portfolio construction

- Throughout this chapter, remember there is a difference between:
  - Factor models (relation between risk and return)
  - Portfolio choice (weights that exploit optimally risk-return tradeoff)

Error in estimating $\mathbb{E}[R]$ and $\mathbb{V}[R]$ from sample moments

- Mean-variance-optimal portfolio weights depend on estimates of return means and covariances:

$$w = \frac{1}{\gamma}(\mathbb{V}[R])^{-1}(\mathbb{E}[R] - R_f 1_N).$$

- In earlier chapters, we saw that
  - Sample-based estimates of $\mathbb{E}[R]$ are imprecise;
  - Sample-based estimates of $\mathbb{V}[R]$ are ill-conditioned;
  - Consequently, mean-variance portfolios perform poorly out of sample.

First shrinkage method: Using only sample moments

- Having identified the problems in sample-based estimates of $\mathbb{E}[R]$ and $\mathbb{V}[R]$, we study shrinkage methods to solve these problems.
- The first kind of shrinkage methods we studied relied only on sample moments but did not take advantage of asset-pricing theory.
- For example, Bayesian shrinkage of expected returns relies on shrinking sample estimate of $\mathbb{E}[R]$ toward a "grand mean," which is
  - either the average of all mean returns
  - or the expected return on the GMV portfolio.
- Similarly, the Ledoit and Wolf methods relies on shrinking the sample estimate of $\mathbb{V}[R]$ toward
  - either a diagonal matrix with the average variance on its diagonal
  - or a matrix where all the cross-asset correlations are replaced by the average correlation.

Empirical performance of sample-based models . . . I

- When we evaluated the empirical performance of models relying only on sample moments of returns, we saw that
  - models shrinking sample estimates of expected returns fail to outperform the simple $1/N$ benchmark;
  - models that ignore expected returns altogether and choose weights based on minimization of portfolio variance perform better,
    - especially when shrinking the covariance matrix of returns using either a short-sale constraint or the Ledoit and Wolf approach,
- However, even these models fail to consistently outperform the $1/N$ portfolio (DeMiguel, Garlappi, and Uppal 2009; Jacobs, Müller, and Weber 2014).

Second shrinkage method: Using asset-pricing models . . . I

- In the last class, we studied how to take advantage of asset-pricing theory to construct portfolios.
- The Black-Litterman model showed how to
  1. use the Capital Asset Pricing Model of Sharpe (1964) to obtain estimates of expected returns from market-portfolio weights;
  2. combine these CAPM-implied expected returns with the views of the investor to obtain the posterior distribution of asset returns, which are then used to construct mean-variance optimal weights.

Second shrinkage method: Using asset-pricing models . . . II

- In this class, we study other asset-pricing models, besides the CAPM, to see how they can be used for portfolio construction.
- In particular, we will look at the asset-pricing factor models proposed by Fama and French (1992, 1993, 2012, 2015, 2018).
- We will then study parametric portfolio policies
  - developed by Brandt, Sant-Clara, and Valkanov (2009),
  - who find a very clever way to build portfolios using the factors identified by Fama and French and other researchers.

Importance of factor models

- Factor models are the core of modern
  - empirical asset pricing and
  - portfolio construction.
- For asset pricing, factor models provide the link between
  - covariances (risk) and
  - expected returns.
- For portfolio construction, factor models reduce the task of
  - searching among thousands of assets
  - to the more tractable problem of finding
  - the optimal risk-return tradeoff among a small number of factors.

Portfolio construction with factor models

- Recall that in the Markowitz model, the expression for the optimal portfolio is given by

$$w = \frac{1}{\gamma}(\mathbb{V}[R])^{-1}(\mathbb{E}[R] - R_f 1_N),$$

  where we estimate $\mathbb{E}[R]$ and $\mathbb{V}[R]$ from their sample moments.

- When using a factor model to construct the optimal portfolio,
  - we still use the expression for w in the equation above,
  - but estimate $\mathbb{E}[R]$ and $\mathbb{V}[R]$ using the factor model.

- As we show below, using a K-factor model substantially reduces the number of parameters we need to estimate, from $N(N+3)/2$ to $2N + 2K + (N \times K)$.

- After that, we will see if it is possible to reduce the number of parameters to be estimated to only K, no matter how large is N.

Road map

Road map

Timeline: Quantitative portfolio management ideas . . . I

- We can see how ideas about investment have progressed over time.
- 4th century: $1/N$
  - "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand [cash].
    Rabbi Issac bar Aha, Babylonian Talmud: Tractate Baba Mezi'a, folio 42a
  - "My ventures are not in one bottom trusted"
    "Merchant of Venice, "Shakespeare ((1564–1616) on the importance of diversification in investing
  - Do not put all your eggs in one basket

Timeline: Quantitative portfolio management ideas ...II

- 1950s: Mean-variance optimization
  (Markowitz 1952, 1959)

- 1964: CAPM
  (Sharpe 1964)

- 1970–2000s: Bayesian shrinkage
  (Klein and Bawa 1976; Bawa, Brown, and Klein 1979; Jorion 1985; Jorion 1988; Jorion 1992; Pástor and Stambaugh 2000)

- 1990s: Black-Litterman model
  (Black and Litterman 1990, 1991a, 1991b, 1992; He and Litterman 1999; Litterman 2003)

.... *This is the point where we are in the course* ....

Timeline: Quantitative portfolio management ideas . . . III

. . . . *Today's class* . . . .

- 1970s: Factor models
  (Ross 1976, 1977)

- 1980s Macro factor models
  (Chen, Roll, and Ross 1986)

- 1990–2020s: Fundamental (firm-characteristic-based) factor models
  (Fama and French 1992, 1993, 2012, 2015, 2018).

- 2015–2020: Principal-component-based factor models
  (Kozak, Nagel, and Santosh 2018, 2020; Lettau and Pelger 2018, 2020).

- 2009-2025: Parametric portfolio policies
  (Brandt, Santa-Clara, and Valkanov 2009; DeMiguel, Martín-Utrera, Nogales, and Uppal 2020).

Timeline: Quantitative portfolio management ideas . . . IV

. . . . *Next two classes* . . . .

- 2017-2023: Volatility-timing of factors
  (Moreira and Muir 2017, 2019; Cederburg, O'Doherty, Wang, and Yan 2020; Barroso and Detzel 2021; DeMiguel, Martín-Utrera, and Uppal 2024).

- 2023-2024: Portfolio construction: Beyond systematic risk
  (Raponi, Uppal, and Zaffaroni 2023; Dello-Preite, Uppal, Zaffaroni, and Zviadadze 2024).

- For a more detailed history of the development of ideas about investment, see the book by Rubinstein (2006).

Road map

An overview of factor models

- The theory of factor models was developed
  - by Ross (1976, 1977),
  - with important advances by Huberman (1982), Chamberlain (1983), Chamberlain and Rothschild (1983), and Ingersoll (1984).
- The theory is completely silent about
  - how many factors to include in the model;
  - which factors to include in the model.
- Thus, in practice, we need to determine the
  - number (K) of factors and
  - identity of these factors.

- The objective is to identify factors that explain well the variation in stock returns.

Road map

Main idea: Shrinking the "asset space" . . . I

- In quantitative portfolio management, factor investing aims to simplify the problem of portfolio construction by reducing the dimension of the space to search in.
- If a model has $K < N$ factors, then one can view these models as shrinking the asset space from N to K.
  - Usually, $N \approx 100$; and
  - Usually, $K \approx 5$ and rarely larger than 7.

Main idea: Shrinking the "asset space" ... II

- Below, we will see how the Fama-French-type factor models reduce the number of parameters that need to be estimated.
- The Brandt, Sant-Clara, and Valkanov (2009) parametric portfolios "shrink" further the number of parameters to be estimated.
  - This approach leads to substantial performance gains.

Road map

A single-factor model

- To understand factor models, we start with a single-factor model.
- Usually, the single factor is assumed to be the return on the market portfolio, $R_{\text{mkt}} = R_m$.
  - Therefore, the single-factor model is sometimes called a single index model, where the index is the market index.
- It is important to understand the difference between
  - CAPM: an equilibrium model of asset returns
  - Market model: a statistical model of asset returns
- The CAPM implies the market model, but the reverse is not true.

Implications of single-factor model for returns . . . I

- In a single-factor model, the (excess) return on asset n is

$$R_n - R_f = \alpha_n + \beta_n(R_m - R_f) + e_n, \quad \mathbb{E}[e_n] = 0, \quad \text{where}$$

  - α_n is the component of security n's return that is independent of the market's return;
  - β_n measures the expected change in R_n given a change in R_m;
  - e_n is uncorrelated with R_m, $\text{Cov}[e_n, R_m] = 0$; i.e., how well the model describes the return of security n is independent of R_m.

Implications of single-factor model for returns . . . II

- If (excess) returns are given by a single-index model,

$$R_n - R_f = \alpha_n + \beta_n(R_m - R_f) + e_n, \quad \mathbb{E}[e_n] = 0, \quad \text{then}$$

- The expected excess return on asset n is:

$$\begin{aligned} \mathbb{E}[R_n] - R_f &= \alpha_n + \beta_n(\mathbb{E}[R_m] - R_f) \\ &= \alpha_n + \beta_n \lambda_m. \qquad \qquad \ldots N \text{ alphas, } N \text{ betas, one } \lambda_m \end{aligned}$$

- Note that $\lambda_m = \mathbb{E}[R_m] - R_f$ in the above equation is the price of risk, i.e., the compensation for bearing beta risk.

Implications of single-factor model for returns . . . III

- Important not to confuse this λ_m with the Lagrange multipliers on the constraints in the Markowitz mean-variance optimization that
  - the weights sum to 1, λ_w;
  - the expected portfolio return has to equal a target mean return, λ_R;
  - the weights cannot take short positions, λ_{ss}.
- We are using the same Greek letter, λ, but the subscripts tell you:
  - in this chapter λ_m is the price of market risk,
  - while in the earlier chapters, λ_w, λ_R, and λ_{ss} were Lagrange multipliers.

Implications of single-factor model for returns . . . IV

- Rewriting that (excess) returns are given by a single-index model,

$$R_n - R_f = \alpha_n + \beta_n(R_m - R_f) + e_n, \quad \mathbb{E}[e_n] = 0, \quad \text{then}$$

- The variance of a security's return is:

$$\sigma_n^2 = \beta_n^2 \sigma_m^2 + \sigma_{e_n}^2. \qquad \ldots N \text{ betas, } N \text{ terms of } \sigma_{e_n}^2, \text{ one } \sigma_m^2$$

- The covariance of returns between securities i and j is:

$$\sigma_{i,j} = \beta_i \beta_j \sigma_m^2. \qquad \ldots \text{no new quantity to be estimated!}$$

Key restriction (assumption) of the single-index model

- The key assumption of the single-index model is that e_i is independent of e_j; i.e.,

$$\mathbb{E}[e_i e_j] = 0, \quad \text{for all pairs } i \neq j.$$

- This means that the only reason for assets to vary together, i.e., systematically, is because of their common comovement with the market return.
- So, this model rules out any effects beyond the market (e.g., industry effects) that account for comovement between securities.

Diversification & portfolio risk in a single-factor model . . . I

- The last item we discuss in the context of the single-factor model is the variance of a portfolio p.
- Let a portfolio be defined by its weights in the N available assets:

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}$$

- Then, the alpha of the portfolio and the beta of the portfolio are the weighted averages of the asset alphas and betas:

$$\alpha_p = \sum_{n=1}^{N} w_n \alpha_n; \quad \text{and} \quad \beta_p = \sum_{n=1}^{N} w_n \beta_n.$$

Diversification & portfolio risk in a single-factor model ... II

- Recall that for the single-index model

$$R_n - R_f = \alpha_n + \beta_n(R_m - R_f) + e_n, \qquad \mathbb{E}[e_n] = 0.$$

- The portfolio variance can then be written as

$$\begin{aligned}
\sigma_p^2 &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \beta_i \beta_j \sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2 && \text{... from market-model returns} \\
&= \left(\sum_{i=1}^{N} w_i \beta_i\right)\left(\sum_{j=1}^{N} w_j \beta_j\right)\sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2 && \text{... grouping terms} \\
&= \beta_p \beta_p \sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2 && \text{... definition of average beta} \\
&= \beta_p^2 \sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2. && \text{... collecting the } \beta_p \text{ terms}
\end{aligned}$$

Diversification & portfolio risk in a single-factor model ...III

- Now consider a portfolio with equal weights, $w_n = 1/N$.
- The portfolio variance in this case can be written as

$$
\begin{aligned}
\sigma_p^2 &= \beta_p^2 \sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2 && \text{... from previous page} \\
&= \beta_p^2 \sigma_m^2 + \sum_{i=1}^{N} \left(\frac{1}{N}\right)^2 \sigma_{e_i}^2 && \text{... replace } w_i \text{ by } 1/N \\
&= \beta_p^2 \sigma_m^2 + \left(\frac{1}{N}\right) \left(\sum_{i=1}^{N} \frac{1}{N} \sigma_{e_i}^2\right) && \text{... grouping terms} \\
&= \beta_p^2 \sigma_m^2 + \left(\frac{1}{N}\right) [\text{average residual risk}] && \text{... defn. avg. resi. risk} \\
\lim_{N\to\infty} \sigma_p^2 &= \beta_p^2 \sigma_m^2 && \text{... taking the limit}
\end{aligned}
$$

- Thus, a diversified portfolio has only beta (systematic) risk.

Comparing number of parameters to be estimated . . . I

- In the absence of a single-factor model, we needed to estimate:
  - N mean returns
  - N variances
  - $(N^2 - N)/2$ covariances
  - for a total of $\frac{N(N+3)}{2}$, which, for $N = 100$, is 5150.
- In the presence of a single-factor model, we needed to estimate:
  - N alphas
  - N betas
  - N asset-specific volatilities, $\sigma^2_{e_n}$
  - mean and volatility of the market excess return, i.e., λ_m and σ_m
  - for a total of only $(3N + 2)$, which, for $N = 100$, is 302.

Comparing number of parameters to be estimated . . . II

- So, in terms of the number of parameters to be estimated, we have made considerable progress:
  - In the Markowitz sample-based model,

$$[\text{Number of parameters}] = \frac{1}{2}N(N+3) \approx \frac{1}{2}N^2.$$

  - In the single-factor model,

$$[\text{Number of parameters}] = 3N + 2 \approx 3N$$

- Thus, as N increases, the number of parameters to be estimated will increase at much slower for the single-factor model.

Comparing number of parameters to be estimated . . . III

- For an investor holding about 11,000 assets (e.g., Norges Bank)
  - The sample-based approach would need to estimate

$$[\text{Number of parameters}] = \frac{1}{2}N(N+3) = 60,516,500.$$

  - The single-factor model would need to estimate

$$[\text{Number of parameters}] = 3N + 2 = 33,002.$$

How to reduce the number of parameters even further?

- Can you think of a way to reduce the number of parameters to be estimated even further?
- Is it possible to de-link entirely
  - the number of parameters to be estimated
  - from the number of assets, N?
- Is it possible to formulate the portfolio problem in such a way that
  - the number of parameters to be estimated is small and constant
  - even when the number of assets, N, is large and increasing?
- Is it possible that, even if $N = 11,000$, we need to estimate only $K = 5$ parameters?
  - We will answer this question later today.

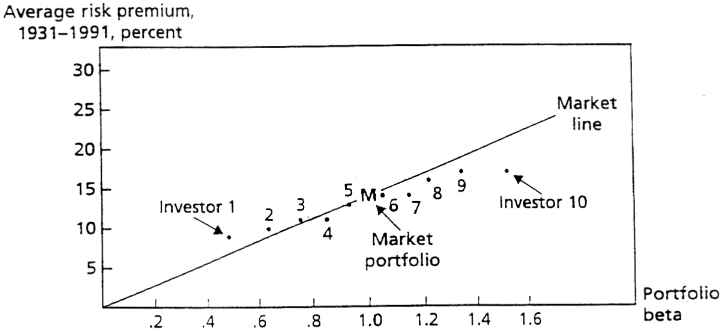
Performance of the market model

- We now examine the empirical evidence on the single-factor model, with the factor being the return on the market portfolio.
- A single-factor model that uses the market return as the factor performs poorly at explaining the cross-section of stock returns.
- As early as the 1990's, Fischer Black had found that the market model did not explain differences in stock returns very well:
  - low-beta stocks had returns higher than those predicted by the market-model;
  - high-beta stocks had returns lower than those predicted by the market-model;

Empirical evidence on performance of market model . . . I

- Actual average risk premiums from portfolios with different betas do not line up with the Security Market Line.

1931–1991
(full sample)

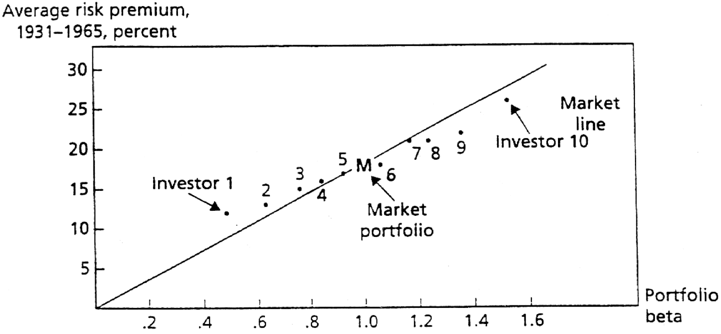


(Source: "Beta and Return" Black (1993))

Empirical evidence on performance of market model . . . II

- CAPM works moderately well over some periods of time

1931–1965
(first half of sample)

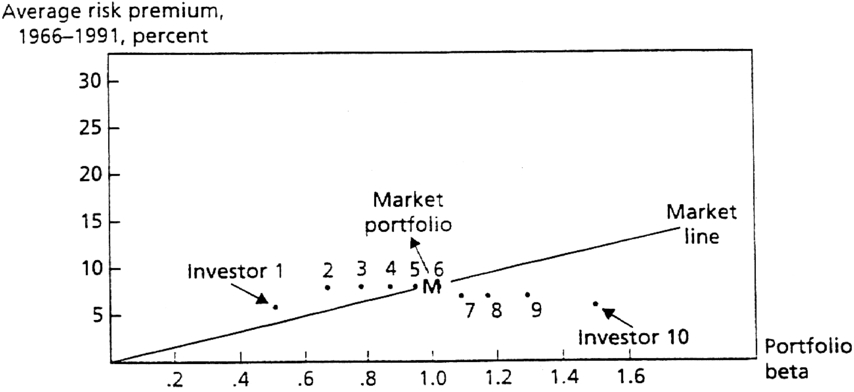


(Source: "Beta and Return" Black (1993))

Empirical evidence on performance of market model . . . III

- CAPM does not work well over significant periods of time

1961–1991
(second half of sample)



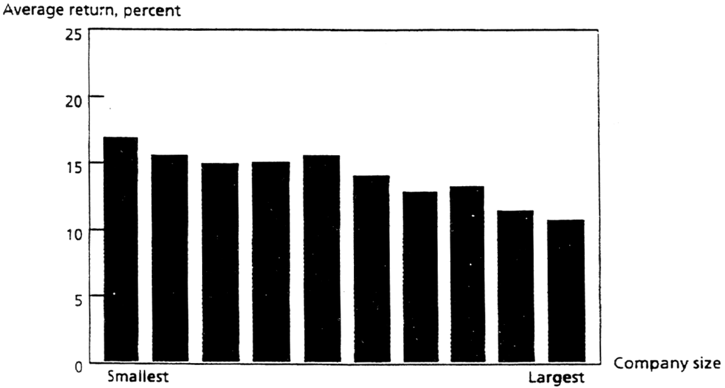
(Source: "Beta and Return" Black (1993))

Empirical evidence on performance of market model . . . IV

- Fama and French showed that factors other than beta seem important in pricing assets; these factors include
  - Size
  - Value (ratio of market value to book value)
- Following their finding, follow-up work has identified hundreds of other factors that seem to be related to returns.

Empirical evidence on performance of market model . . . V

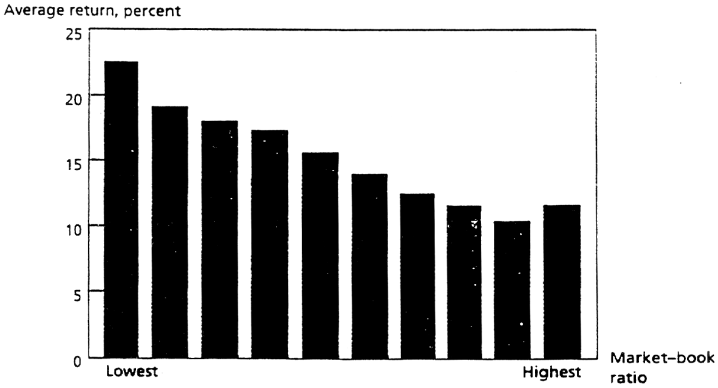
- Since the 1960s, small stocks have outperformed large stocks



Source: "The Cross-Section of Expected Stock Returns" Fama and French (1992)

Empirical evidence on performance of market model . . . VI

- Since mid-1960s, stocks with low ratios of market-to-book value, have outperformed stocks with high ratios.



Source: "The Cross-Section of Expected Stock Returns" Fama and French (1992)

Overall assessment of CAPM (and market model)

- CAPM is a good theoretical model
  - It is simple and sensible.
  - It is built on modern portfolio theory
  - It distinguishes diversifiable risk and non-diversifiable risk
  - It provides a simple pricing model.
  - It is relatively easy to implement in practice.
- But, empirical evidence supporting the market model is weak.

- Recall that our objective was to identify factors that explain well the variation in stock returns.
- The market model is not very successful at explaining the variation in stock returns.

Underlying assumptions of CAPM

- Assumptions of the CAPM:
  - Investors care about returns over the next short horizon
  - Financial market is perfect:
    - all assets are traded
    - no frictions such as trading costs and taxes
    - all investors have perfect information
  - Investors hold fully diversified mean-variance frontier portfolios.
- These assumptions are often violated in practice.
  - E.g., investors do not hold fully diversified portfolios.
- Deviations from these assumptions lead to alternative models.
  - E.g., the model in Merton (1987), in which investors do not hold fully diversified portfolios.

From single-factor to multifactor models

- The poor empirical performance of the single-factor market model motivates the development of multifactor models.
- The hope is that by including multiple factors, we will be able to do a better job of explaining the variation in stock returns.

Start of focus

Road map

A multifactor model in terms of K factors

- We now look at a model with K factors.

$$R_n - R_f = \alpha_n + \beta_{n,1} F_1 + \beta_{n,2} F_2 + \ldots + \beta_{n,K} F_K + e_n, \quad \mathbb{E}[e_n] = 0,$$

  - where $\beta_{n,k}$ is the nth row of a $N \times K$ matrix of betas with respect to the K factors, and
  - $F_1, F_2, \ldots, F_K$ are the K factors,
  - $\text{Cov}[F_k, F_\ell] = 0$; that is, the factors are assumed to be orthogonal to one another.
  - If the original factors are not orthogonal, they can always be transformed so that they become orthogonal, without affecting any implications for asset returns.

A multifactor model in terms of K factor returns

- If the factors are tradable (as opposed to being, e.g., macro factors), then the model can be written in terms of factor returns.
- The K-factor model in terms of factor returns R_{F_k} is:

$$R_n - R_f = \alpha_n + \beta_{n,1} R_{F_1} + \beta_{n,2} R_{F_2} + \ldots + \beta_{n,K} R_{F_K} + e_n, \quad \mathbb{E}[e_n] = 0,$$

  which implies that expected (excess) returns, are

$$\begin{aligned} \mathbb{E}[R_n] - R_f &= \alpha_n + \beta_{n,1}\lambda_1 + \beta_{n,2}\lambda_2 + \ldots + \beta_{n,K}\lambda_K \\ &= \alpha_n + \beta_n \lambda, \quad \text{where} \end{aligned}$$

  - $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ is the $K \times 1$ vector of factor prices of risk.

Return variance and covariance in a multifactor model

- The K-factor model in terms of factor returns is

$$R_n - R_f = \alpha_n + \beta_{n,1} R_{F_1} + \beta_{n,2} R_{F_2} + \ldots + \beta_{n,K} R_{F_K} + e_n, \quad \mathbb{E}[e_n] = 0,$$

  which implies that the return variance is

$$\sigma_n^2 = \beta_{n,1}^2 \sigma_{F_1}^2 + \beta_{n,2}^2 \sigma_{F_2}^2 + \ldots + \beta_{n,K}^2 \sigma_{F_K}^2 + \sigma_{e_n}^2, \quad \text{where}$$

  $\sigma_{F_k}^2$ is the variance of the return on the kth factor.

- The K-factor model implies that the return covariance between assets i and j is a

$$\sigma_{i,j}^2 = \beta_{i,1}\beta_{j,1}\sigma_{F_1}^2 + \beta_{i,2}\beta_{j,2}\sigma_{F_2}^2 \ldots + \beta_{i,K}\beta_{j,K}\sigma_{F_K}^2$$

  where we have assumed that $\text{Cov}(e_i, e_j) = 0$ and also the factor returns are orthogonal.

Number of parameters to be estimated for portfolio construction

- For portfolio construction, we need estimates of expected returns and risk (variances and covariances).
- To estimate the mean return and risk for N assets in a K-factor model, we need to estimate the following parameters:
  - α_n for each of the N stocks ... N;
  - $\sigma^2_{e_n}$ for each of the N stocks ... N;
  - $\beta_{n,k}$ for each of the N stocks for each of the K factors; ... $N \times K$;
  - λ_k and $\sigma^2_{F_k}$ for the K factors. ... $2K$;
- This is a total of: $2N + 2K + (N \times K)$.
  - For $N = 100$ and $K = 5$, we need to estimate 710 parameters;
  - For $N = 11,000$ and $K = 5$, we need to estimate 77,010 parameters.

Road map

Three types of factors in asset-pricing factor models

- To improve the market model's performance for explaining the cross-section of stock returns, these models add to it new factors.
- Depending on the type of factors we add, multifactor models of security market returns can be divided into three types.
  1. Macroeconomic factor models use observable economic time series, such as inflation and interest rates, as measures of pervasive shocks to security returns, i.e., as factors.
  2. Fundamental factor models use the returns on portfolios associated with observed security characteristics such as dividend yield, the book-to-market ratio, size, value, and industry identifiers.
  3. Statistical factor models use factors from the principal-components analysis (PCA) of the dataset of security returns.

Road map

First type: Macroeconomic factor models

- The first empirical application of a macroeconomic factor model was a paper by Chen, Roll, and Ross (1986).
  - Cochrane (2017) provides a review of the factors that emerge from macro-economic theories.
- Some of the macroeconomic variables typically used as factors are
  - inflation,
  - the percentage change in industrial production,
  - the excess return to long-term government bonds, and
  - the realized return premium of low-grade corporate bonds relative to high-grade bonds.
- To estimate the risk-premia (λ_k) for macroeconomic factors, we estimate the risk-premia on factor-mimicking portfolios.
- Macroeconomic factor models are not very good at explaining stock returns, compared to the other types of factors models.

Road map

Second type: Fundamentals-based factor models . . . I

- The literature on fundamentals-based factor models started with the
  - three-factor model of Fama and French (1993),
  - extended to five factors by Fama and French (2015), and then
  - extended to six factors by Fama and French (2018).
- Python code to access Fama-French data is available from:
  - PyAnomaly (efficient data download from WRDS using asynco).
  - pyassetpricing.
  - getFamaFrenchFactors.
- Hou, Xue, and Zhang (2015, 2017b) propose new fundamentals and evaluate how well different factor models explain many “anomalies.”

Second type: Fundamentals-based factor models ...II

- The finance literature has "discovered" hundreds of fundamentals or firm characteristics that drive returns.
  - An excellent website with data on 331 predictors is Open Source Asset Pricing; for details, see Chen and Zimmermann (2022).
- In the last few years, machine-learning techniques have been applied to factor investing; for a review, see Giglio, Kelly, and Xiu (2022).
  - Cazalet and Roncalli (2014) and Clarke, Silva, and Thorley (2016) discuss the practical issues in using machine-learning factor models.

How to construct a fundamental factor . . . I

- Define the company-specific characteristic:
  - This could be a fundamental metric like earnings, dividend yield, or a financial ratio, or it could be based on non-financial attributes like ESG criteria.
- Data collection:
  - Collect historical data for the selected company-specific characteristics for a sample of companies.
- Selection criteria:
  - Establish criteria for selecting companies to include in the portfolio.
  - In academic research, we use all companies listed on NYSE, NASDAQ, and Amex exchanges.
  - But, you could use other criteria to select firms.

How to construct a fundamental factor . . . II

- Characteristic portfolio formation
  - Create "two" sub-portfolios: one for companies with high exposure to the chosen factor (High Factor Portfolio) and another for companies with low exposure (Low Factor Portfolio).
  - But, you could also form 4 quartile portfolios and consider only the top and bottom quartile.
  - Or, you could form 10 decile portfolios and consider only the top and bottom decile.
  - Or you could rank all firms and consider all of them.

How to construct a fundamental factor . . . III

- Equal-weighted or value-weighted characteristic portfolio
  - Decide whether the characteristic will be equally weighted (each stock has the same weight) or value-weighted (weights based on market capitalization).
  - In most cases, it is better to use value weights so that bigger firms receive more weight (see Plyakha, Uppal, and Vilkov 2021).
- Rebalancing:
  - Determine the rebalancing frequency of the portfolio.
  - Rebalancing ensures that the portfolio continues to represent the chosen factor over time.
  - In academics, we usually rebalance the portfolio annually (to limit turnover).
  - But, you could rebalance quarterly or monthly.

How to construct a fundamental factor . . . IV

- Benchmark selection:
  - Identify a benchmark index or portfolio against which you will compare the performance of your characteristic portfolio.
- Performance measurement:
  - Calculate performance metrics for both the "High" and "Low" portfolios. (Or, for all quartile/decile portfolios.)
  - Common performance metrics include annualized returns, cumulative returns, volatility, and the Sharpe ratio.
- Sensitivity analysis:
  - Perform sensitivity analyses by varying parameters such as portfolio construction criteria, rebalancing frequency, and benchmark selection to assess the robustness of your findings.

Portfolio choice with the Fama-French factor model

- Suppose that the N stock returns are driven by the three-factor ($K = 3$) Fama-French model,
- where the three factors are long-short portfolios
  - returns on the market (mkt) minus the risk-free asset,
  - size (smb), and
  - value (hml).

$$R_j - R_f = \alpha_j + \beta_{j,\text{mkt}}(R_{\text{mkt}} - R_f) + \beta_{j,\text{smb}} R_{\text{smb}} + \beta_{j,\text{hml}} R_{\text{hml}} + e_j,$$

  - where $\mathbb{E}[e_j] = 0, \mathbb{V}[e_j] = \sigma^2_{e_j}, \text{Cov}[e_i, e_j] = 0$, and
  - σ^2_{mkt}, σ^2_{smb}, and σ^2_{hml} are the return variances of the three factors, and
  - the factors are assumed to be orthogonal to one another.

Moments of asset returns

- Returns are described by

$$R_j - R_f = \alpha_j + \beta_{j,\text{mkt}}(R_{\text{mkt}} - R_f) + \beta_{j,\text{smb}} R_{\text{smb}} + \beta_{j,\text{hml}} R_{\text{hml}} + e_j,$$

- From the returns above, we can compute:
  - Expected value (mean) of the returns of asset j is given by:

$$\mathbb{E}[R_j] - R_f = \alpha_j + \beta_{j,\text{mkt}}(\mathbb{E}[R_{\text{mkt}}] - R_f) + \beta_{j,\text{smb}}\mathbb{E}[R_{\text{smb}}] + \beta_{j,\text{hml}}\mathbb{E}[R_{\text{hml}}]$$

  - Variance of the returns of asset j is given by:

$$\mathbb{V}[R_j] = \beta^2_{j,\text{mkt}}\sigma^2_{\text{mkt}} + \beta^2_{j,\text{smb}}\sigma^2_{\text{smb}} + \beta^2_{j,\text{hml}}\sigma^2_{\text{hml}} + \sigma^2_{e_j}.$$

  - Covariance between the returns on asset i and j is given by:

$$\mathbb{C}[R_i, R_j] = \beta_{i,\text{mkt}}\beta_{j,\text{mkt}}\sigma^2_{\text{mkt}} + \beta_{i,\text{smb}}\beta_{j,\text{smb}}\sigma^2_{\text{smb}} + \beta_{i,\text{hml}}\beta_{j,\text{hml}}\sigma^2_{\text{hml}}.$$

Optimal portfolio weights based on Fama-French model

- To find the optimal portfolio weights when returns are give by the Fama-French model, we need to use the expression

$$w = \frac{1}{\gamma}(\mathbb{V}[R])^{-1}(\mathbb{E}[R] - R_f 1_N),$$

  but where the vector of expected returns, $\mathbb{E}[R]$, and the return covariance matrix, $\mathbb{V}[R]$, are obtained as shown on the previous slide.

Performance of firm-characteristic based factor models

- Recall that our objective was to identify factors that explain well the variation in stock returns.
- Factor models based on firm characteristics
  - perform better than the market model
  - but still have limited success in explaining variation in stock returns.
- In our final class, we will see how to construct portfolios that deliver much higher Sharpe ratios than those based on firm characteristics.

Road map

Third type: Statistical factor models (based on principal components) . . . I

- Recall that our objective is to identify factors that explain well the variation in stock returns.
- Principal-components analysis (PCA) is a statistical technique designed to do exactly that.
- PCA identifies the principal components (PCs) that explain the largest variation in stock returns.

Third type: Statistical factor models (based on principal components) . . . II

- PCA is a statistical technique for factor construction and dimensionality reduction.
- In the context of factor-based investing,
  - PCA helps identify underlying factors in a dataset by transforming the original variables
  - into a set of linearly uncorrelated variables, called PCs.
- Python is very very convenient for doing PCA.
  - See the "PCA" function in the "sklearn.decomposition" library.
- For an introduction to PCA, see
  - this excellent article, or
  - this article for an even simpler exposition.

Steps for using PCA for factor construction . . . I

- Note: factor construction is different from portfolio construction.
  - In factor construction, we construct the factors using the PCs.
  - In portfolio construction, we construct portfolio weights using the factors selected from the PCs.
- In this section, we discuss factor construction.
- In the next section, we discuss portfolio construction.

Steps for using PCA for factor construction . . . II

1. Data preparation for factor construction
   - Collect relevant financial data, such as stock returns for a set of assets (e.g., stocks).
     - One can also apply PCA to other quantities, such as financial ratios, or other firm characteristics.
   - Decide whether to standardize the data to have a mean of zero and a standard deviation of one.
2. Covariance matrix calculation
   - Calculate the covariance matrix of the (standardized) return data.

Steps for using PCA for factor construction . . . III

3. Eigen-system decomposition

- Perform eigenvalue decomposition on the (standardized) return covariance matrix, i.e., decompose it into a set of eigenvectors and eigenvalues:

$$\mathbb{V}[R] = E \Lambda E^\top,$$

where

$$\begin{aligned} \mathbb{V}[R] &= \text{original } N \times N \text{ return covariance matrix} \\ E &= (E_1, E_2, \ldots E_N) = \text{matrix of } N \text{ (column) eigenvectors} \\ \Lambda &= \text{diag}(\Lambda_1^2, \Lambda_2^2, \ldots \Lambda_N^2) = \text{diagonal matrix with } N \text{ eigenvalues} \end{aligned}$$

and, its inverse is given by

$$(\mathbb{V}[R])^{-1} = E \Lambda^{-1} E^\top.$$

Steps for using PCA for factor construction . . . IV

4. Interpretation of the eigen-system
    - Eigenvectors represent the directions (principal components = PCs) of maximum variance;
    - Eigenvalues represent the magnitude of the variance in those directions.
5. Principal-components extraction
    - PCA arranges the eigenvectors in descending order of their corresponding eigenvalues.
    - The first PC explains the most variance in the data, the second PC explains the second most, and so on.

Steps for using PCA for factor construction . . . V

6. PC loadings
   - The PCs obtained from PCA are linear combinations of the original variables.
   - These linear combinations are represented by PC loadings, which indicate the contribution of each original variable to the PC.
   - The PC loadings provide insights into which variables are driving the variation captured by each PC.
7. "Factor" construction
   - The first few PCs that capture significant variance are considered potential "factors" (use cross-validation to decide how many PCs to retain).
   - These "factors" represent the dominant sources of variation in the dataset.

Steps for using PCA for factor construction . . . VI

8. PC interpretation
   - Interpret the PC based on the characteristics of the original variables with high factor loadings.
   - For example, if the first PC has high loadings on the market return, it might be interpreted as a "market factor."
   - Interpreting factors is crucial for understanding their economic significance.

We have finished our discussion of
using PCA for building factor models.

Now we study
how to use PCA to construction portfolios.

Steps for using PCA for portfolio construction . . . I

1. Download historical stock data: Use the yfinance library to download historical stock data for a set of companies.
   - This data should be organized in a matrix, where rows represent time periods and columns represent different assets.
2. Calculate returns: Calculate monthly returns from the adjusted closing prices.
3. Standardize returns: Standardize the returns by subtracting the mean and dividing by the standard deviation for each asset, so that each variable has a mean of zero and a standard deviation of one.
4. Split data into training and testing sets: Split the data into a training set (used for constructing portfolios) and a testing set (used for evaluating out-of-sample performance).

Steps for using PCA for portfolio construction . . . II

5. Perform PCA: Use the training set to perform PCA to decompose the covariance matrix into its principal components; i.e., find the eigenvalues and eigenvectors of the covariance matrix.

6. Select principal components: Determine the number of principal components to retain based on the explained variance.

   - You may choose a certain percentage of total variance to retain (e.g., 95%).

7. Construct principal components portfolio: Construct portfolios using the retained principal components.

   - Each portfolio will be a linear combination of the original assets,
   - with the weights given by the corresponding eigenvector of the retained PCs.

Steps for using PCA for portfolio construction . . . III

8. Determine portfolio weights for each asset: Determine the weights *for each asset* in the portfolio by combining the weights across the retained principal components.

9. Normalize weights: Normalize the portfolio weights to ensure that they sum to one, making the portfolio fully invested.

10. Apply weights to testing data:
    - Apply the weights to the testing set to evaluate the performance of the PCA portfolio using a metric such as the Sharpe ratio.
    - Choose the set of weights that perform best out-of-sample.

GMV with PCA-based dimensionality reduction

- Let us denote the return-covariance matrix by $V = \mathbb{V}[R]$.
- We know that the global minimum-variance portfolio (GMV) is

$$w_{\text{GMV}} = \frac{V^{-1}1_N}{1_N^\top V^{-1}1_N}.$$

- We also know that the return-covariance matrix can be written as

$$V = E\Lambda E^\top, \quad \text{with its inverse given by:} \quad V^{-1} = E\Lambda^{-1}E^\top,$$

  where E contains the N principal components (eigenvectors).
- Putting the two together,

$$w_{\text{GMV}} = \frac{(E\Lambda^{-1}E^\top)1_N}{1_N^\top(E\Lambda^{-1}E^\top)1_N},$$

  which would just be another way of getting exactly the old result.

Sample code for PCA-based portfolios ... I

- The code below shows that, exactly as one would expect, you get the same results if you use
  - the standard $N \times N$ covariance matrix
  - all N PCs from the PCA decomposition of the covariance matrix
- The objective of this code is to
  - illustrate this (obvious) result, and
  - set the stage for the case where we consider $K < N$ PCs.

Sample code for PCA-based portfolios . . . II

```python
import numpy as np

# Simulate returns for 10 Assets
np.random.seed(0)
T, N, K = 100, 10, 3
F = np.random.normal(0, 1, size=(T, K))      # Latent factors
B = np.random.uniform(-1, 1, size=(N, K))    # Factor loadings
Err = np.random.normal(0, 0.05, size=(T, N))   # Idiosyncratic noise
R = F @ B.T + Err                              # Simulated returns

#Sample mean vector and covariance matrix of returns
mu = R.mean(axis=0)                    # Sample mean returns
Sigma_full = np.cov(R, rowvar=False)   # Sample covariance matrix
inv_Sigma = np.linalg.inv(Sigma_full) # Inverse of cov. matrix

# Minimum variance portfolio (no expected return constraint)
ones = np.ones(N)
w_full = inv_Sigma @ ones
w_full /= ones @ inv_Sigma @ ones
```

Sample code for PCA-based portfolios . . . III

```
# PCA-based portfolio choice
from sklearn.decomposition import PCA

pca = PCA(n_components=N)  # No truncation: use all N components
pca.fit(R)
E_N = pca.components_.T                           # Eigenvectors
Lambda_N = np.diag(pca.explained_variance_)  # Eigenvalues

# Untruncated covariance and its inverse
Sigma_pca = E_N @ Lambda_N @ E_N.T     # Covariance matrix from PCA
Sigma_pca_inv = E_N @ np.linalg.inv(Lambda_N) @ E_N.T

# PCA-based weights
w_pca = Sigma_pca_inv @ ones
w_pca /= ones @ Sigma_pca_inv @ ones
```

Sample code for PCA-based portfolios . . . IV

```
print("\nWeights of the GMV portfolio using the standard (full)
     covariance matrix = ")
print(w_full)

print("\nWeights of the GMV portfolio using the PCA (full) covariance
     matrix = ")
print(w_pca)

Weights of GMV portfolio using standard (full) covariance matrix =
[ 0.08867118  0.23587618  0.06426487  0.06258271  0.08986071
0.25729893  0.15513514  0.08365772  0.01134367 -0.0486911 ]

Weights of GMV portfolio using PCA (full) covariance matrix =
[ 0.08867118  0.23587618  0.06426487  0.06258271  0.08986071
0.25729893  0.15513514  0.08365772  0.01134367 -0.0486911 ]
```

Truncated PCA covariance estimation

- To improve out-of-sample performance, we reduce noise by retaining only the top $K < N$ principal components.
  - You can choose K using cross-validation.
- Let $E_K \in \mathbb{R}^{N \times K}$ and $\Lambda_K \in \mathbb{R}^{K \times K}$, so the reduced-dimension matrix is

$$\hat{V}_K = E_K \Lambda_K E_K^\top.$$

- Then the (pseudo) inverse of $\hat{V}_K$ is given by:

$$\hat{V}_K^{-1} = E_K \Lambda_K^{-1} E_K^\top.$$

- The GMV weights, based on the reduced-dimension matrix, are

$$w_{\text{GMV}} = \frac{\hat{V}_K^{-1} 1_N}{1_N^\top \hat{V}_K^{-1} 1_N}.$$

Sample code for *truncated*-PCA GMV portfolios . . . I

- We now look at the code for the case where we do not use all the PCs.
- Instead, we truncate the number of PCs, keeping only K PCs.
- The code below shows the results for the GMV portfolio if you truncate the PCA to keep only $K = 3$ PCs.

Sample code for *truncated*-PCA GMV portfolios . . . II

```
import numpy as np
# Simulate returns for 10 Assets
np.random.seed(0)
T, N, K = 100, 10, 3
F = np.random.normal(0, 1, size=(T, K))       # factor returns
B = np.random.uniform(-1, 1, size=(N, K))     # betas
Err = np.random.normal(0, 0.1, size=(T, N))  # residual errors
R = F @ B.T + Err  # simulated excess returns using factor model
ones = np.ones(N)  # vector of ones

# Sample moments
mu = R.mean(axis=0)                       # Sample mean returns
Sigma_full = np.cov(R, rowvar=False)   # Sample return covariance matrix
inv_Sigma = np.linalg.inv(Sigma_full)  # Inverse of covariance matrix
print("\nSample mean returns = ")
print(mu)
print("\nSample covariance matrix = ")
print(Sigma_full)

# Minimum variance portfolio (no expected return constraint)
w_full = inv_Sigma @ ones
w_full /= ones @ inv_Sigma @ ones
```

Sample code for *truncated*-PCA GMV portfolios . . . III

```
# PCA-based portfolio choice
from sklearn.decomposition import PCA

# PCA with only K = 3 elements included (i.e., with truncation)
pca = PCA(n_components=K)
pca.fit(R)   # fit the simulated returns keeping only K components
E_K = pca.components_.T   # eigenvalues
Lambda_K = np.diag(pca.explained_variance_)   # eigenvectors
print("\nEigenvectors when K = ", K)
print(E_K)
print("\nEigenvalues when K = ", K)
print(Lambda_K)

# Reconstruct rank-K covariance'' matrix
V_K = E_K @ Lambda_K @ E_K.T
V_K_inv = E_K @ np.linalg.inv(Lambda_K) @ E_K.T

print("\nSample covariance matrix reconstructed from truncated PCA = ")
print(V_K)

# PCA-based weights
w_pca = V_K_inv @ ones
w_pca /= ones @ V_K_inv @ ones
```

Sample code for *truncated*-PCA GMV portfolios . . . IV

```
print("\nWeights of the GMV portfolio using the standard (full)
    covariance matrix = ")
print(w_full)

print("\nWeights of GMV portfolio with K = 3 truncated PCA covariance
    matrix = ")
print(w_pca)

# Output of the print statements

Sample mean returns =
[ 0.12708671 -0.16183823  0.04197566 -0.05191858  0.16836941  0.16672355
 -0.1176492  -0.17905506 -0.08783624 -0.0241036 ]
```

Sample code for *truncated*-PCA GMV portfolios . . . V

```
Sample covariance matrix =
[[ 0.78095393 -0.49041375 -0.12489251  0.6365252   0.42636192
0.09346918  -0.14552655  0.11854582  0.37169097  0.82012935]
 [-0.49041375  0.70871751 -0.56471794  0.0583193  -0.9199419
 -0.46095861   0.4286234   0.41375973 -0.07788141 -0.26635618]
 [-0.12489251 -0.56471794  2.02442023 -0.34594391  1.36384133
 -0.23111022  -0.21837588  0.16318862  0.69051056  0.24033844]
 [ 0.6365252   0.0583193  -0.34594391  1.37999485 -0.22606286
 -0.90015408   0.46308246  1.24660673  1.0640816   1.42801229]
 [ 0.42636192 -0.9199419   1.36384133 -0.22606286  1.46660552
 0.38634751  -0.52256429 -0.37653416  0.32575217  0.3299647 ]
 [ 0.09346918 -0.46095861 -0.23111022 -0.90015408  0.38634751
 1.27531849  -0.66979766 -1.44249204 -1.08167927 -0.92001436]
 [-0.14552655  0.4286234  -0.21837588  0.46308246 -0.52256429
 -0.66979766   0.44018513  0.74632079  0.41538559  0.34165468]
 [ 0.11854582  0.41375973  0.16318862  1.24660673 -0.37653416
 -1.44249204   0.74632079  1.72328278  1.34745599  1.3035032 ]
 [ 0.37169097 -0.07788141  0.69051056  1.0640816   0.32575217
 -1.08167927   0.41538559  1.34745599  1.35972644  1.36714473]
 [ 0.82012935 -0.26635618  0.24033844  1.42801229  0.3299647
 -0.92001436   0.34165468  1.3035032   1.36714473  1.72964436]]
```

Sample code for *truncated*-PCA GMV portfolios . . . VI

```
Eigenvectors when K = 3
[[ 0.12668987  0.16152746 -0.5232372 ]
 [ 0.04005712 -0.36740922  0.26813076]
 [ 0.06822867  0.5931274   0.52954895]
 [ 0.41130711 -0.08183566 -0.31543207]
 [-0.02205572  0.6012429  -0.05156336]
 [-0.38439408  0.12290759 -0.31998217]
 [ 0.18044483 -0.19896489  0.15881839]
 [ 0.48314812 -0.1173505   0.20341851]
 [ 0.42571837  0.16574378  0.09315237]
 [ 0.46156536  0.15163668 -0.30672813]]

Eigenvalues when K = 3
[[6.74909868 0.         0.        ]
 [0.         4.01792971 0.        ]
 [0.         0.         2.05425031]]
```

Sample code for *truncated*-PCA GMV portfolios . . . VII

```
Sample covariance matrix reconstructed from truncated PCA =
[[ 0.77556432 -0.49240332 -0.12590944  0.63761857  0.42677506
0.09503061  -0.14554898  0.11830449  0.37145019  0.82276092]
 [-0.49240332  0.70089637 -0.56546393  0.05826203 -0.9219337
 -0.46160855   0.42997837  0.41589898 -0.07827325 -0.26801405]
 [-0.12590944 -0.56546393  2.02098337 -0.34876189  1.36660032
 -0.23218501  -0.21830363  0.16410182  0.69236057  0.24024777]
 [ 0.63761857  0.05826203 -0.34876189  1.37306986 -0.22550843
 -0.90013196   0.46341728  1.24797173  1.06691515  1.43017637]
 [ 0.42677506 -0.9219337   1.36660032 -0.22550843  1.46119851
 0.38802748  -0.52433279 -0.3769562   0.32715791  0.33009957]
 [ 0.09503061 -0.46160855 -0.23218501 -0.90013196  0.38802748
 1.26826651  -0.67078115 -1.44510115 -1.08382827 -0.92094232]
 [-0.14554898  0.42997837 -0.21830363  0.46341728 -0.52433279
 -0.67078115   0.43062574  0.74857606  0.41634801  0.34081939]
 [ 0.11830449  0.41589898  0.16410182  1.24797173 -0.3769562
 -1.44510115   0.74857606  1.71579077  1.34896518  1.30540805]
 [ 0.37145019 -0.07827325  0.69236057  1.06691515  0.32715791
 -1.08382827   0.41634801  1.34896518  1.35138257  1.36846366]
 [ 0.82276092 -0.26801405  0.24024777  1.43017637  0.33009957
 -0.92094232   0.34081939  1.30540805  1.36846366  1.72350068]]
```

Sample code for *truncated*-PCA GMV portfolios . . . VIII

```
Weights of the GMV portfolio using the standard (full) covariance matrix
      =
[ 0.0903513   0.23522936  0.06315848  0.06223911  0.09010167  0.25788985
  0.15465218  0.08350277  0.01110026 -0.04822498]

Weights of GMV portfolio with K = 3 truncated PCA covariance matrix =
[ 0.18394367 -0.15265068  0.13217576  0.16635709  0.20041182 -0.03796389
 -0.03046346  0.09304727  0.18555295  0.25958947]
```

Mean-variance portfolio using truncated PCA . . . I

- As before, we start by doing a PCA of returns, and retain only the top $K < N$ principal components.
  - You can choose K using cross-validation.
- Then, as before, write the factor model for returns on the N assets,

$$R = B\, R_{\text{PCA}} + \epsilon,$$

  but now we are using matrix notation, so
  - $R \in \mathbb{R}^N$,
  - $B \in \mathbb{R}^{N \times K}$ is the matrix of betas,
  - $R_{\text{PCA}} \in \mathbb{R}^K$ is the vector of K statistical-factor returns, and
  - $\epsilon \in \mathbb{R}^N$:

Mean-variance portfolio using truncated PCA . . . II

- Using the eigenvalues $E_K \in \mathbb{R}^{N\times K}$, the return on the K statistical factors, denoted R_{PCA}, is:

$$\hat{R}_{\text{PCA}} = E_K^\top R.$$

- Use the returns on the statistical factors to estimate expected returns on the N assets:

$$\mathbb{E}[R] = \hat{B}\,\mathbb{E}[\hat{R}_{\text{PCA}}].$$

- Similarly, estimate the covariance matrix of N asset returns:

$$\mathbb{V}[R] = \hat{B}\,\mathbb{V}[\hat{R}_{\text{PCA}}]\,\hat{B}^\top + D \quad \ldots D \text{ is diagonal matrix of idio (residual) variances}$$

- Then, the mean-variance optimal portfolio based on PCA is:

$$w_{\text{MV}} = \frac{1}{\gamma}(\mathbb{V}[R])^{-1}(\mathbb{E}[R] - R_f 1_N).$$

Python Code: Expected Returns and Optimal Weights

```
# Estimate factor returns and loadings
R_PCA = R @ E_K  # R, E_K are from earlier code for PCA decomposition
B_hat = E_K
mu_R_PCA = R_PCA.mean(axis=0)   # mean returns for statistical factors
mu = B_hat @ mu_R_PCA   # mean returns for N risky assets

# Factor covariance and idiosyncratic variance
cov_R_PCA = np.cov(R_PCA, rowvar=False)
D = np.diag(np.var(R - R_PCA @ B_hat.T, axis=0))

# Total covariance of returns for the $N$ assets
Sigma = B_hat @ cov_R_PCA @ B_hat.T + D

# Optimal mean-variance weights (gamma = 1 for simplicity)
inv_Sigma = np.linalg.inv(Sigma)
w_MV = inv_Sigma @ mu
w_MV /= np.sum(w_MV)  # normalize to sum to 1
```

Papers and Python code for portfolio construction using principal components

- Papers explaining portfolio construction using PCA:
  - Partovi and Caputo (2004).
  - Meucci (2009).
  - Pasini (2017).
- Python code to construct portfolios using PCA
  - "Eigen-portfolio construction using PCA".
  - "A principal component analysis of portfolio risk shift-pandemic period".

Advantages of PCA for factor construction . . . I

- Balance:
  - PCA balances structure (via statistical factor models) and robustness (via dimension reduction).
- Dimensionality reduction:
  - PCA allows for the reduction of the dimensionality of the dataset while retaining most of the important information.
  - This is particularly useful when dealing with a large number of variables.
- Uncorrelated factors:
  - The principal components obtained through PCA are uncorrelated, which can simplify subsequent analyses.
  - This is valuable when constructing factors that are intended to be independent of each other.

Advantages of PCA for factor construction . . . II

- Capturing commonalities:
  - PCA effectively captures commonalities across variables and identifies the dominant sources of variation in the data.
  - This is particularly relevant in finance when seeking to identify common factors affecting asset returns.
  - Expected returns can be incorporated through latent factor structures estimated by PCA.
- Quantitative approach:
  - PCA provides a quantitative and data-driven approach to factor construction.
  - It does not require a priori assumptions about the nature of the factors, making it a flexible method.

Limitations of PCA for factor construction . . . I

- Interpretability:
  - While PCA identifies statistical factors, the economic interpretation of these factors may not always be straightforward.
  - Researchers often need to rely on additional information to interpret and name the factors meaningfully.
- Assumption of linearity:
  - PCA assumes a linear relationship between variables.
  - If the relationship is highly nonlinear, PCA may not capture the underlying factors accurately.

Limitations of PCA for factor construction . . . II

- Data quality:
  - The effectiveness of PCA depends on the quality and relevance of the input data.
  - Noisy or irrelevant variables may lead to less meaningful factors.
- Factor stability:
  - PCA assumes that the factors derived are stable over time.
  - In dynamic market conditions, factors may change, and ongoing monitoring and adjustment may be necessary.
  - Recent papers show how to do conditional PCA; see Kelly, Pruitt, and Su (2020). Python code for this paper is available from Pruitt's homepage.

Evaluation of truncated-PCA portfolios

- Full mean-variance uses all eigen-directions, including noisy ones.
- Truncated-PCA approach filters out small-eigenvalue directions—usually dominated by estimation error.
  - Choice of truncation parameter depends on trade-off between *fit* and *stability*.
- PCA portfolios tend to be more stable, diversified, and robust in practice.

Road map

Comparison of the three types of factor models ... I

- Macroeconomic and statistical factor models both estimate a firm's factor-beta by time-series regression.
    - Given the nature of security returns data, this limitation is substantial.
    - Time-series regression requires a long and stable history of returns to estimate the factor betas accurately.

Comparison of the three types of factor models . . . II

- Fundamental factor models do not require time series regression; it uses observed company attributes as factor betas.
  - These models rely on the empirical finding that company attributes such as firm size, dividend yield, book-to-market ratio, and industry classification explain a substantial proportion of common return.
  - The factors in a fundamental factor model are the realized returns to a set of mimicking portfolios designed to capture the marginal returns associated with a unit exposure to each attribute.
    - For example, the dividend yield factor is the realized return per extra unit of dividend yield, holding other attributes constant.
  - The betas are exogenously determined, firm-specific attributes rather than estimated sensitivities to random factors, and

Comparison of the three types of factor models . . . III

- Portfolios relying on PCA perform well empirically.
- With improving computing capabilities, in the last 10 years there have been substantial advances in
  - the use of PCA-based methods along with
  - the use of techniques from machine learning.
- In our last class, we will look at a PCA-based model in detail.

Comparison of the three types of factor models ...IV

- In theory, i.e., in the absence of estimation error and with no limits on data availability, the three types of models are fully consistent with one another;
- That is, the three models are simply restatements or (to use a technical term from factor modeling) rotations of one another.
- In this sense, the three types of factor models can all hold simultaneously.
- In practice, as we have discussed above, their performance will vary.

End of focus

Road map

How are factors identified?

- Factors are identified in various ways – depending on type of factors.

1. Factor models could be based on macroeconomic equilibrium asset pricing model.
   - E.g., long-run consumption growth, past consumption habits, and macroeconomic uncertainty.
2. The Fama-French factors are based on economic intuition for firm characteristics that may be indicators of expected returns.
   - E.g., size, value, profitability, investment.

   Other factors originate as a trading strategy that is discovered to produce abnormal returns:
   - E.g., momentum, price reversal, low beta, and idiosyncratic volatility.
3. A purely statistical approach starts by estimating the principal components from a panel of returns.

Challenges in implementing factor models

- While conceptually elegant, the empirical identification of factors is a difficult problem.
- One problem is that there is a very large number of factors.
  - Cochrane (2011) coined the term "factor zoo" referring to the literally hundreds of factors discovered so far.
- Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2024) give a list of
  - 457 firm-specific characteristics in the finance literature;
  - 103 macro factors.

Questions when using factor models

- How many true factors are there?
  - Given the large number of factors "discovered" so far, how should we "tame the factor zoo" (Feng, Giglio, and Xiu 2020).
- Which factors are priced and so should be included?
  - For portfolio management, there is no point taking factor risk for which there is no reward (Daniel, Mota, Rottke, and Santos 2020).
- How stable is the factor structure?
  - A single factor model with time-varying beta's is often equivalent to a multifactor model, and vice versa.
- What is the economic interpretation of the factors?
  - Why does a particular factor drive returns?

Estimating how many factors are needed

- Suppose that one wishes to test whether K factors are sufficient to explain all the pervasive movements in security returns.
- Then, one could estimate two models on the same data set.
  - One model with K factors;
  - The other model with $K + 1$ factors.
- Comparing the unexplained residual cross-sectional variance by the two models will then tell us whether we need the $K + 1$ factors.
- If $K + 1$ factors are necessary, then the difference in the average asset-specific variance should be strictly positive.
- Thus, the increase in explanatory power from adding a factor is the basis for testing whether the additional factor is needed.

Road map

Portfolio construction using a factor model . . . I

- Having selected a factor model, the next question is how to use it to construct a portfolio.
- To use factors to construct a portfolio, we need to compute the returns associated with the factor.
  - That is, each factor needs to be translated into a trading strategy so that we can compute its returns.
- For some factors, this is straightforward because the factor is itself defined as the trading strategy that generates (abnormal) returns.
- However, for macro factors, such as inflation, a trading strategy that replicates (mimics) the factor needs to be constructed.

Portfolio construction using a factor model . . . II

- Lamont (2001) explains how to use regression models to construct factor-mimicking portfolios.
- But when the number of assets or number of trading strategies is large, standard regression models do not work well.
- Implementing regressions in such a setting requires regularization (shrinkage) techniques, such as Lasso, Boosting, or Bayesian priors.

Portfolio construction using a factor model . . . III

- In the last ten years, machine-learning techniques have been used to regularize large-scale portfolio problems.
  - One example is Fan, Zhang, and Yu (2012), who use the LARS (Least Angle Regression) algorithm (explained in the appendix to these slides).
  - Heaton, Polson, and Witte (2016) apply "deep learning" techniques to portfolio optimization; Goodfellow, Bengio, and Courville (2016) is a textbook on deep learning.
  - DeMiguel, Martin-Utrera, Nogales, and Uppal (2017) use the lasso.
- Other approaches to regularization of portfolio weights have been motivated by model uncertainty
  (see Goldfarb and Iyengar 2003; Garlappi, Uppal, and Wang 2007).

Portfolio construction using a factor model . . . IV

- Once you have selected the K factors, there are two steps:
  1. Estimate the factor model for asset returns and compute the means and covariance matrix for the N asset returns;
  2. Use mean-variance portfolio optimization to select how much weight to put on each of the N assets.

Start of focus

Road map

Parametric portfolio policies: The big picture

- Parametric portfolio policies (PPP) were developed by Brandt, Santa-Clara, and Valkanov (2009) and rely on a very clever insight.
- Their insight is that,
  - instead of having two steps:
    1. first estimate a factor model for returns and
    2. then, in a second step, use the factor model to identify the optimal weights,
  - to specify, directly in one step,
    1. a factor model for the portfolio *weights*.

Parametric portfolio policies: The details . . . I

- That is, instead of using a factor model for *asset returns*, where we use two steps to find the optimal portfolio weights:
  1. Start with a K-factor model for the N asset returns

     $$\mathbb{E}[R - R_f 1_N] = \alpha + \beta\lambda;$$

     and find the asset return means and covariances in terms of the parameters of the factor model
     - N alphas for the assets, α
     - N residual variances for the assets, $\sigma^2_{e_n}$
     - $N \times K$ asset betas for the K factors, $\beta_{n,k}$
     - K means and K variances for the factors, λ_k and $\sigma^2_{F_k}$;
  2. Then, in a second step, use these in a mean-variance optimization to find the optimal portfolio weights.

Parametric portfolio policies: The details . . . II

- Brandt, Santa-Clara, and Valkanov (2009) propose, in a single step,
- specifying a parametric (factor) model for the $N_t \times 1$ vector of portfolio weights, $w_t(\theta)$

$$w_t(\theta) = w_{b,t} + (F_{1,t}\theta_1 + F_{2,t}\theta_2 + \ldots + F_{K,t}\theta_K)/N_t, \quad \text{where}$$

  - $w_{b,t}$ is the $N_t \times 1$ vector of benchmark portfolio weights at time t,
  - $F_{k,t}$ is the $N_t \times 1$ long-short characteristic portfolio obtained by standardizing the kth firm-specific characteristic at time t,
  - θ_k is the (scalar) weight on the kth characteristic portfolio in the parametric portfolio,
  - N_t is the number of firms at time t.
- Note θ does not depend on time, but firm characteristics, $F_{k,t}$, do.

Parametric portfolio policies: The details . . . III

- In the expression for parametric portfolio weight

$$w_t(\theta) = w_{b,t} + (F_{1,t}\theta_1 + F_{2,t}\theta_2 + \ldots + F_{K,t}\theta_K)/N_t, \quad \text{where}$$

  - $w_{b,t}$ is determined by the client (could be market portfolio);
  - $F_{k,t}$ is determined by the data on firm characteristics (e.g., size, value, momentum, etc.);
  - N_t is determined by how many firms we have data for;
  - $\theta = \{\theta_1, \ldots, \theta_K\}$ is a $K \times 1$ vector to be chosen by the investor.
- So, the only unknown is the $K \times 1$ vector $\theta = \{\theta_1, \ldots, \theta_K\}$.
- So, even if $N = 11,000$, we have to estimate only K parameters.
  - Whether $K = 3, 4, 5, \ldots$ will depend on your choice of factor model; but K is usually a small number (less than 10).

Parametric portfolio policies: The details . . . IV

- The weights of the characteristics in the parametric portfolio are scaled by the number of stocks N_t
  - so that they are meaningful for the case where the number of stocks is varying over time;
  - otherwise, increasing the number of stocks while keeping the weights fixed would result in more aggressive portfolio allocations.
- On the next few slides we explain (and illustrate) the design of parametric portfolio policies.

Standardizing characteristic portfolios . . . I

- Brandt, Santa-Clara, and Valkanov (2009) suggest that you standardize each characteristic so that
  - it has a cross-sectional mean of zero, and
  - a cross-sectional standard deviation of one.
- The resultant standardized characteristic is a long-short portfolio that is
  - long stocks whose characteristic is above the cross-sectional average;
  - short stocks whose characteristic is below cross-sectional average.

Standardizing characteristic portfolios . . . II

- Each row represents a particular stock.
- Each number represents the value of the characteristic for that stock.

Firm-specific characteristics

	Value long-short	Momentum long-short	Size long-short
Stock 1:	0.01%	-0.02%	0.03%
Stock 2:	-0.02%	+0.01%	-0.01%
Stock 3:	-0.03%	+0.12%	0.13%
⋮	.	.	.
	.	.	.
	.	.	.
Stock N:	0.03%	-0.02%	-0.01%

Illustration of parametric portfolios

	Parametric portfolio		Benchmark portfolio		Value long-short		Momentum long-short		Size long-short
Stock 1:	0.05%		0.09%		0.01%		-0.02%		0.03%
Stock 2:	0.05%		0.05%		-0.02%		+0.01%		-0.01%
Stock 3:	0.17%		0.21%		-0.03%		+0.12%		0.13%
⋮	.	=	.	$+\ \Theta_1$	.	$+\ \Theta_2$	.	$+\ \Theta_3$	.
	.		.		.		.		.
	.		.		.		.		.
Stock N:	0.17%		0.15%		0.03%		-0.02%		-0.01%

Full-invested in risky assets

- Brandt, Santa-Clara, and Valkanov (2009) consider a portfolio that is fully invested in risky assets.
- Thus, the parametric portfolio weights on the N_t stocks sum to one.
- Because the weights of the stocks in each characteristic long-short portfolio sum to zero,
- it implies that the parametric weight on the benchmark portfolio must be equal to one.

Returns on the parametric portfolio . . . I

- The parametric portfolio can be written in compact matrix notation.
- Define F_t, the $N_t \times K$ matrix whose kth column is $F_{k,t}$.
- Then,
$$w_t(\theta) = w_{b,t} + F_t\theta/N_t, \quad \text{where} \tag{39}$$
- θ is the $K \times 1$ parameter vector, whose kth component is the weight of the kth characteristic θ_k, and
- $F_t\theta/N_t$ is the characteristic portfolio at time t.

Returns on the parametric portfolio ... II

- The return of the parametric portfolio at time $t+1$, $r_{p,t+1}(\theta)$, is

$$\begin{aligned} r_{p,t+1}(\theta) &= w_{b,t}^{\top} r_{t+1} + \theta^{\top} F_t^{\top} r_{t+1}/N_t \\ &= r_{b,t+1} + \theta^{\top} r_{c,t+1}, \quad \text{where} \end{aligned} \tag{40}$$

  - r_{t+1} is the $N_t \times 1$ return vector at time $t+1$,
  - $r_{b,t+1} = w_{b,t}^{\top} r_{t+1}$ is the benchmark portfolio return at time $t+1$,
  - $r_{c,t+1} = F_t^{\top} r_{t+1}/N_t$ is the characteristic return vector at time $t+1$, which contains the returns of the long-short portfolios corresponding to the K characteristics scaled by the number of firms N_t.

- Equation (40) shows that the parametric-portfolio return is
  - the benchmark-portfolio return
  - plus the return of the characteristic portfolio.

Mean and variance of parametric-portfolio returns

- Once you have the expression we derived above for the return of a parametric portfolio

$$r_{p,t+1}(\theta) = r_{b,t+1} + \theta^\top r_{c,t+1}, \tag{40}$$

- we can find the mean and variance of the portfolio returns using standard methods for finding the moments of portfolio returns.
- Once we have the mean and variance of the portfolio return, we can write the optimization problem of a mean-variance investor.

Determining the weights on the characteristic portfolios

- Assume that the investor choses the weights θ by maximizing mean-variance utility:

$$\max_{\theta} \quad \mathbb{E}_t[r_{p,t+1}(\theta)] - \frac{\gamma}{2}\mathbb{V}_t[r_{p,t+1}(\theta)], \quad \text{where} \tag{41}$$

  - γ is the investor's risk-aversion parameter
  - $\mathbb{E}_t[r_{p,t+1}(\theta)]$ is the mean of the parametric-portfolio return.
  - $\mathbb{V}_t[r_{p,t+1}(\theta)]$ is the variance of the parametric-portfolio return.
  - Note, again, that θ does not depend on time.

- If you want, you can
  - add nonnegativity constraints on θ;
  - apply shrinkage to the covariance matrix;
  - adjust for transaction costs.

Regularized parametric portfolios

- To deal with the large number of characteristics in Finance, DeMiguel, Martín-Utrera, Nogales, and Uppal (2020) develop a new class of parametric portfolios.
- These are called regularized parametric portfolios.
- These portfolios are obtained by imposing a lasso (least absolute shrinkage and selection operator).
- The lasso constraint
  - reduces the impact of estimation error and
  - acts as a variable-selection method that helps to reduce problem dimensionality.

Interpretation of parametric portfolio policies

- Parametric portfolios use firm-specific characteristics to tilt the benchmark portfolio toward stocks that improve performance.
  - Parametric portfolios are similar to Black-Litterman portfolios in the sense that both portfolios are a tilt of a benchmark portfolio.
- Parametric portfolios are obtained by
  - adding to the benchmark portfolio
  - a linear combination of long-short characteristic portfolios, obtained by standardizing K firm-specific characteristics cross-sectionally.
- Black-Litterman portfolios are obtained by
  - adding to the benchmark (market) portfolio
  - the views of the investor.

Python code for parametric portfolio policies

- Python code for parametric portfolio policies is available from:
  - The Tidy Finance website (in the section on "Portfolio optimization").
  - But, to run exactly what they run, you may also need access to WRDS.
  - Alternatively, you can use the data file I provided for the assignment.
  - Note that the Tidy Finance website shows optimization for a power utility function, which is what Brandt, Sant-Clara, and Valkanov (2009) use, while we use a mean-variance utility function.

End of focus

Road map

Factor models for other asset classes

- Factor models are relevant not just for stocks, but also other asset classes.
- The challenge is to identify a single set of factors that explain returns for stocks and also other asset classes.
  - Note that stocks, corporate bonds, and equity options have the same underlying: the assets of the firm.
- Some of the current research is focused on the search for a factor model that works across many asset classes:
  - equities
  - bonds
  - options
  - currencies
  - commodities
- See, e.g., Bali, T. G., H. Beckmeyer, and A. Goyal. 2024. A joint factor model for bonds, stocks, and options. Available at SSRN 4589282, from this link.

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will study if it is possible to time the factors we have found to be important for explaining the cross-section of stock returns.

Specifically, we will study if we can use the volatility of a factor to time our investment in that factor.

Then we will examine what this implies for the risk-return trade-off; i.e., is risk related to returns?

To do for next class: Readings

- Readings
  - A good reference for "factor investing" is the book by Ang (2014).
    - Ang was Head of Factors, Sustainable and Solutions for BlackRock.
    - You can read more about him on this website.
  - For parametric portfolio policies, see the Tidy Finance website (in the section on "Portfolio optimization").
    - This website includes Python code for parametric portfolios.

Road map

Least Angle Regression (LARS) algorithm . . . I

- The Least Angle Regression (LARS) algorithm is a method
  - used for linear regression,
  - particularly in the context of high-dimensional data
  - that was introduced by Efron, Hastie, Johnstone, and Tibshirani (2004) and can be downloaded from this link.
- The main goal of LARS is to find a sparse linear regression model
  - when the number of predictors (features) is much larger than the number of observations.
- It is particularly useful in situations where the number of predictors is comparable to or even exceeds the number of observations.

Key concepts underlying LARS . . . I

- Forward Selection: LARS builds the regression model by adding predictors to the model one at a time. At each step, it identifies the predictor most correlated with the response variable.
- Equal Correlation (Least Angle): LARS moves towards the predictor with the highest correlation with the residual, but it does so in a way that maintains all predictors with equal absolute correlations with the residual.
  - This is where the name "Least Angle" comes from; the algorithm takes the least angle necessary to include a new predictor.

Key concepts underlying LARS . . . II

- Regularization: LARS involves a regularization parameter that controls the shrinkage of coefficients.
  - As the algorithm progresses, it shrinks the coefficients of variables not in the model toward zero.
- Multiple Coefficient Paths: LARS can trace the entire path of the solution for different values of the regularization parameter.
  - This allows for a comprehensive understanding of how the coefficients change as the regularization strength varies.

Road map

Bibliography . . . I

Ang, A. 2014. *Asset management: A systematic approach to factor investing.* Oxford University Press. (Cited on page 860).

Bali, T. G., H. Beckmeyer, and A. Goyal. 2024. A joint factor model for bonds, stocks, and options. Available at SSRN 4589282. (Cited on page 857).

Barroso, P., and A. L. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140 (3): 744–767. (Cited on pages 734, 886, 906, 939, 949–953, 963, 966, 1030, 1153).

Bawa, V. S., S. Brown, and R. Klein. 1979. *Estimation risk and optimal portfolio choice.* North Holland, Amsterdam. (Cited on pages 458, 732, 884, 1028).

Black, F. 1993. Beta and return. *Journal of Portfolio Management* 20 (1): 8–18. (Cited on pages 755–757).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

———. 1991a. Combining investor views with market equilibrium. *Journal of Fixed Income* 1 (2): 7–18. (Cited on pages 626, 732, 884, 1028).

———. 1991b. Global asset allocation with equities, bonds, and currencies. *Fixed Income Research* 2 (15-28): 1–44. (Cited on pages 626, 732, 884, 1028).

Bibliography . . . II

Black, F., and R. Litterman. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Brandt, M. W., P. Sant-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. *Review of Financial Studies:* Forthcoming. (Cited on pages 626, 726, 739, 854).

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22 (9): 3411–3447. (Cited on pages 733, 839, 841, 844, 847, 885, 1029, 1151).

Cazalet, Z., and T. Roncalli. 2014. Facts and fantasies about factor investing. Available at SSRN 2524547. (Cited on page 776).

Cederburg, S., M. S. O'Doherty, F. Wang, and X. Yan. 2020. On the performance of volatility-managed portfolios. *Journal of Financial Economics* 138 (1): 95–117. (Cited on pages 734, 886, 906, 939, 940, 942, 943, 945–947, 1030, 1153).

Chamberlain, G. 1983. Funds, factors and diversification in arbitrage pricing models. *Econometrica* 51:1305–1324. (Cited on page 736).

Chamberlain, G., and M. Rothschild. 1983. Arbitrage, factor structure and mean-variance analysis on large asset markets. *Econometrica* 51:1281–1304. (Cited on page 736).

Bibliography ... III

Chen, A. Y., and T. Zimmermann. 2022. Open source cross-sectional asset pricing. *Critical Finance Review* 11 (2): 207–264. (Cited on pages 82, 776).

Chen, N.-F., R. Roll, and S. Ross. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403. (Cited on pages 733, 773, 885, 1029).

Clarke, R., H. de Silva, and S. Thorley. 2016. Fundamentals of efficient factor investing. *Financial Analysts Journal* 72:9–26. (Cited on page 776).

Cochrane, J. 2017. Macro-finance. *Review of Finance* 21:1–42. (Cited on page 773).

Daniel, K., L. Mota, S. Rottke, and T. Santos. 2020. The cross-section of risk and returns. *Review of Financial Studies* 33 (5): 1927–1979. (Cited on page 830).

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2024. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on pages 734, 829, 886, 1030, 1069, 1155).

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. (Cited on pages 450, 495, 559, 619, 724, 1120).

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal. 2017. A portfolio perspective on the multitude of firm characteristics. Working Paper, London Business School. (Cited on page 835).

Bibliography . . . IV

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. 2020. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Review of Financial Studies* 33, no. 5 (April): 2180–2222. (Cited on pages 162, 733, 852, 885, 957, 1029).

DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. (Cited on pages 162, 734, 886, 906, 953, 955, 1004, 1030, 1154).

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani. 2004. Least angle regression. *Annals of Statisitcs* 32 (2): 407–499. (Cited on page 862).

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2 (June): 427–465. (Cited on pages 625, 726, 733, 759, 760, 885, 1029).

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105 (3): 457–472. (Cited on pages 625, 726, 733, 885, 1029).

Bibliography . . . V

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

———. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

Fan, J., J. Zhang, and K. Yu. 2012. Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association* 107:592–606. (Cited on page 835).

Feng, G., S. Giglio, and D. Xiu. 2020. Taming the factor zoo: a test of new factors. *Journal of Finance* 75 (3): 1327–1370. (Cited on page 830).

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies* 20:41–81. (Cited on pages 459, 835).

Giglio, S., B. Kelly, and D. Xiu. 2022. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics* 14:337–368. (Cited on page 776).

Goldfarb, D., and G. Iyengar. 2003. Robust portfolio selection problems. *Mathematics of Operations Research* 28 (1): 1–38. (Cited on pages 459, 835).

Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep learning.* MIT Press. (Cited on page 835).

Bibliography . . . VI

He, G., and R. Litterman. 1999. The intuition behind Black-Litterman model portfolios. *Investment Management Research (Goldman, Sachs & Company).* (Cited on pages 626, 640, 705, 732, 884, 1028).

Heaton, J., N. Polson, and J. Witte. 2016. Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry* 33:3–12. (Cited on page 835).

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705. (Cited on pages 81, 132, 775, 945, 966).

———. 2017b. Replicating anomalies. *The Review of Financial Studies.* (Cited on pages 132, 775).

Huberman, G. 1982. A simple approach to the Arbitrage Pricing Theory. *Journal of Economic Theory* 28:183–191. (Cited on page 736).

Ingersoll, J. E., Jr. 1984. Some results in the theory of arbitrage pricing. *Journal of Finance* 39:1021–1039. (Cited on page 736).

Jacobs, H., S. Müller, and M. Weber. 2014. How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets* 19:62–85. (Cited on pages 495, 620, 621, 623, 724).

Bibliography . . . VII

Jorion, P. 1988. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 732, 884, 1028).

Jorion, P. 1985. International portfolio diversification with estimation risk [in English]. *Journal of Business* 58 (3): pp. 259–278. (Cited on pages 458, 467, 500, 732, 884, 1028).

———. 1992. Portfolio optimization in practice. *Financial Analysts Journal* 48 (1): 68–74. (Cited on pages 732, 884, 1028).

Kelly, B. T., S. Pruitt, and Y. Su. 2020. Instrumented principal component analysis. *Available at SSRN 2983919.* (Cited on page 819).

Klein, R. W., and V. S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3 (3): 215–31. (Cited on pages 465, 732, 884, 1028).

Kozak, S., S. Nagel, and S. Santosh. 2018. Interpreting factor models. *Journal of Finance* 73 (3): 1183–1223. (Cited on pages 133, 733).

———. 2020. Shrinking the cross-section. *Journal of Financial Economics* 135 (2): 271–292. (Cited on pages 133, 733).

Lamont, O. 2001. Economic tracking portfolios. *Journal of Econometrics* 105:161–184. (Cited on page 834).

Bibliography . . . VIII

Lettau, M., and M. Pelger. 2018. *Estimating latent asset-pricing factors.* Technical report. National Bureau of Economic Research. (Cited on page 733).

———. 2020. Factors that fit the time-series and cross-section of stock returns. *Review of Financial Studies* 33:2274–2325. (Cited on page 733).

Litterman, R. 2003. *Modern investment management: An equilibrium approach.* New York: Wiley. (Cited on pages 626, 732, 884, 1028).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information [in English]. *Journal of Finance* 42 (3): 483–510. (Cited on pages 762, 1081, 1082).

Meucci, A. 2009. Managing diversification. *Risk:* 74–79. (Cited on page 815).

Moreira, A., and T. Muir. 2017. Volatility-managed portfolios. *Journal of Finance* 72 (4): 1611–1644. (Cited on pages 734, 886, 905–907, 909, 911, 912, 914, 918, 921, 923, 927, 936–939, 942, 945, 947, 955, 966, 1008, 1030, 1152, 1153).

———. 2019. Should long-term investors time volatility? *Journal of Financial Economics* 131 (3): 507–527. (Cited on pages 734, 886, 1030).

Bibliography . . . IX

Partovi, M. H., and M. Caputo. 2004. Principal portfolios: Recasting the efficient frontier. *Economics Bulletin* 7 (3): 1–10. (Cited on page 815).

Pasini, G. 2017. Principal component analysis for stock portfolio management. *International Journal of Pure and Applied Mathematics* 115 (1): 153–167. (Cited on page 815).

Pástor, Ľ., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56 (3): 335–381. (Cited on pages 468, 732, 884, 1028).

Plyakha, Y., R. Uppal, and G. Vilkov. 2021. Equal or value weighting? Implications for asset-pricing tests. In *Risk management and modeling,* edited by R. Benkraiem, I. Kalaitzoglou, and C. Zopounidis. Springer. (Cited on pages 510, 779).

Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary. (Cited on pages 622, 623, 734, 886, 1030, 1111, 1155).

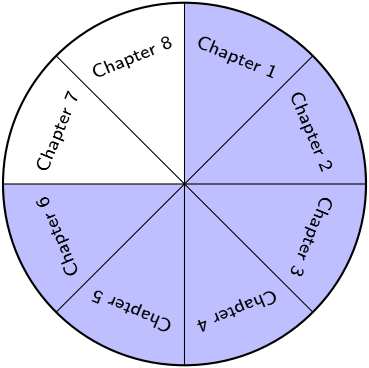
Ross, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13:341–360. (Cited on pages 733, 736, 885, 1029, 1082).

———. 1977. Return, risk, and arbitrage. In *Risk and return in finance,* edited by I. Friend and J. Bicksler. Cambridge, MA: Ballinger. (Cited on pages 733, 736, 885, 1029, 1082).

Bibliography . . . X

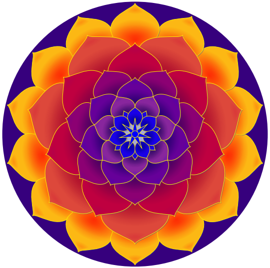
Rubinstein, M. 2006. *A history of the theory of investments: My annotated bibliography.* Hoboken, NJ: John Wiley & Sons. (Cited on pages 734, 887, 1031).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).



End of Chapter 6

Quantitative Portfolio Management



Chapter 7: Volatility-Timed Factor Portfolios

Raman Uppal

2025

The big picture: Plan for the entire book

Part A: Preliminaries

Chapter 1: Properties of asset returns

Chapter 2: Performance measurement (especially out of sample)

Part B: Modern portfolio management

Chapter 3: Mean-variance portfolios that *ignore* estimation error

Chapter 4: Mean-variance portfolios that *adjust for* estimation error

Part C: Post-Modern Portfolio Management

Chapter 5: CAPM-based portfolios: Black-Litterman model

Chapter 6: Factor-based portfolios: Parametric portfolio policies

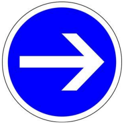
Chapter 7: Volatility-timed factor portfolios

Chapter 8: Portfolios exploiting systematic risk factors *and* unsystematic risk

Table of contents

Road map

What do we want to do in Chapter 7



In this chapter, we study if it is possible to *time* our investment in risk factors.

Specifically, first we study if we can use the volatility of a factor to time the investment in that factor.

Then, we study the performance of a portfolio of *many* factors where the weight in each factor depends on its own volatility or on market volatility.

Timeline: Quantitative portfolio management ideas . . . I

- We can see how ideas about investment have progressed over time.

. . . . *The thinking in ancient times* . . . .

- 4th century: $1/N$
  - "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand [cash].
    Rabbi Issac bar Aha, Babylonian Talmud: Tractate Baba Mezi'a, folio 42a
  - "My ventures are not in one bottom trusted"
    "Merchant of Venice, "Shakespeare ((1564–1616) on the importance of diversification in investing
  - Do not put all your eggs in one basket

Timeline: Quantitative portfolio management ideas . . . II

. . . . *Below are the topics we covered in Chapters 3–5* . . . .

- 1950s: Mean-variance optimization
  (Markowitz 1952, 1959)

- 1964: CAPM
  (Sharpe 1964)

- 1970–2000s: Bayesian shrinkage
  (Klein and Bawa 1976; Bawa, Brown, and Klein 1979; Jorion 1985; Jorion 1988; Jorion 1992; Pástor and Stambaugh 2000)

- 1990s: Black-Litterman model
  (Black and Litterman 1990, 1991a, 1991b, 1992; He and Litterman 1999; Litterman 2003)

Timeline: Quantitative portfolio management ideas . . . III

. . . . *Last time: Chapter 6* . . . .

- 1970s: Factor models
  (Ross 1976, 1977)

- 1980s Macro factor models
  (Chen, Roll, and Ross 1986)

- 1990–2020s: Fundamental (firm-characteristic-based) factor models
  (Fama and French 1992, 1993, 2012, 2015, 2018).

- 2009-2023: Parametric portfolio policies
  (Brandt, Santa-Clara, and Valkanov 2009; DeMiguel, Martín-Utrera, Nogales, and Uppal 2020).

Timeline: Quantitative portfolio management ideas . . . IV

. . . . *Today: Chapter 7* . . . .

- 2017-2024: Volatility-timing of factors
  - Moreira and Muir (2017, 2019)
  - Cederburg, O'Doherty, Wang, and Yan (2020)
  - Barroso and Detzel (2021)
  - DeMiguel, Martín-Utrera, and Uppal (2024).

. . . . *Next: Chapter 8* . . . .

- 2023-2024: Portfolio construction: Beyond systematic risk factors
  - Raponi, Uppal, and Zaffaroni (2023)
  - Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2024).

Timeline: Quantitative portfolio management ideas . . . V

- For a more detailed history of the development of ideas about investment, see the book by Rubinstein (2006).

Road map

Road map

Motivation for the material in this chapter

- Factor investing: Quantitative-investment approach that exploits firm characteristics that predict expected stock returns.
- Smart beta: *low-cost* approach that exploits *commonly known* characteristics with the aim to outperform a capitalization-weighted benchmark in a rules-based, transparent manner.
- Smart beta strategies are designed to provide investors with exposure to specific factors that have historically been associated with
  - higher returns,
  - lower risk, or
  - other desirable investment properties (e.g., lower drawdown or higher ESG score).

Some common firm characteristics

Characteristic	Acronym	**Buy**	**Sell**
Size	SMB	Small firms	Large firms
Value	HML	High book-to-market	Low book-to-market
Momentum	UMD	Winners (Up)	Losers (Down)
Investment	CMA	Conservative	Aggressive
Profitability	RMW	Robust	Weak

- In the assignment based on this chapter, we will look at 9 factors.

Key components of smart beta strategies . . . I

- Factor exposure
  - Smart beta strategies are often based on exposure to specific factors, which are characteristics or risk premiums that have been shown to drive returns over the long term.
  - Common factors include value, momentum, size, and low volatility.
  - E.g., a value-based smart beta strategy might involve selecting undervalued stocks based on price-to-earnings ratios or book value.
- Rules-based methodology
  - Smart beta strategies use a predetermined set of rules for portfolio construction and rebalancing,
  - as opposed to relying on market-capitalization-weighted portfolios.

Key components of smart beta strategies . . . II

- Systematic and transparent
  - Smart beta strategies are systematic, meaning they follow a consistent and predefined methodology.
  - They are also transparent, allowing investors to understand the underlying factors and rules guiding the portfolio.
  - E.g., a low-volatility smart beta strategy might select stocks with historically lower price volatility.
- Diversification
  - Smart beta strategies often maintain a level of diversification to reduce specific stock or sector risk.
  - E.g., a multifactor smart beta strategy might combine exposure to value, momentum, and quality factors to create a diversified portfolio that aims to outperform traditional market-cap-weighted indices.

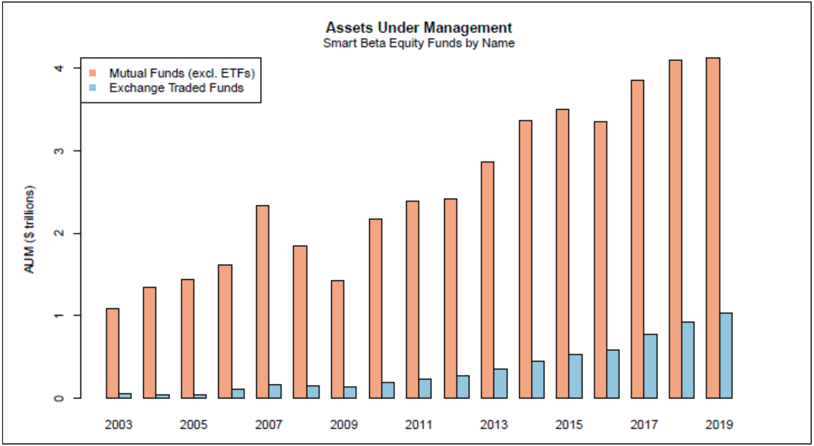
Key components of smart beta strategies . . . III

- Performance evaluation and risk considerations
  - Smart beta strategies are evaluated based on their historical performance and risk characteristics relative to traditional benchmarks.
  - E.g., a smart beta strategy focusing on dividend yield might attract investors seeking income, with the performance evaluated against a dividend-focused benchmark.
- Huang, Song, and Xiang (2020) find that smart beta strategies have underperformed by 1% on average since launch.

Factor investing growth

- Assets under management growing fast:

  Johansson, Sabbatucci, Tamoni (2020) *since 2010, the total assets under management for the U.S. Smart Beta ETF market have grown 30% per year.*"



Factor investing and parametric portfolio policies

- In the last class, we saw how to construct portfolios using factor models.
- In particular, we studied parametric portfolio policies, which showed how to reduce the dimensionality of the portfolio problem when investing in factors.
  - Instead of modeling asset returns in terms of factors, and then constructing the optimal portfolio,
  - parametric portfolio policies directly specify portfolio weights in terms of factors.

$$w_t(\theta) = w_{b,t} + (\theta_1 F_{1,t} + \theta_2 F_{2,t} + \ldots + \theta_K F_{K,t})/N_t.$$

Start of focus

Road map

Parametric portfolios whose weights depend on volatility

- In this class, we build on parametric portfolio policies,
  - by modeling the portfolio weight on each factor, θ_k,
  - to depend on factor or market volatility.

$$\theta_{k,t} = a_k + \frac{b_k}{\sigma_{k,t}} \qquad \text{or} \qquad \theta_{k,t} = a_k + \frac{b_k}{\sigma_{\text{mkt},t}}.$$

Motivation for studying volatility-timing of factors

A fundamental premise in finance:

strong risk-return tradeoff

$$\mathbb{E}\left[\text{Return}\right] \propto \text{Risk}$$

If volatility timing was successful,
it would imply breakdown of risk-return tradeoff

Importance of volatility-timing strategies

- Volatility-timing strategies have important implications for:
  - investors,
  - asset managers, and, therefore,
  - they received considerable attention in the financial press.

Articles in the financial press about volatility timing

- "Reassessing the classic risk-return tradeoff," The Financial Times, March 9, 2016. Link to article.
- "When markets get scary, panicking is smart," CNBC, March 23, 2016. Link to article and video.
- "Authers' Note: Healthy correction?," The Financial Times, February 7, 2018. Link to article.
- Video interview of Alan Moreira by CEPR & VideoVox Economics in 2017. Link to video.

Volatility-timing strategies in practice

- BlackRock offers the following description of the investment strategy for its Managed Volatility V.I. Fund:
  - "In periods of heightened volatility, the portfolio will de-risk into less volatile assets like fixed income and cash and re-risk when market turbulence subsides."
  - Performance plots available from this link to FT.

Can you time factors?
Cliff Asness (AQR) believes you cannot . . .

The Journal of Portfolio Management

INVITED EDITORIAL COMMENT

The Siren Song of Factor Timing
aka "Smart Beta Timing" aka "Style Timing"

CLIFFORD S. ASNESS

Link to Cliff Asness (2016) article

(look for the "View" button on the right side of the page)

Can you time factors?
Moreira and Muir (2017) believe you can ...

- Moreira and Muir (2017) show investors can increase Sharpe ratios by reducing exposure to risk factors when their volatility is high.
- The success of the volatility-timing strategy implies that:
  - changes in factor volatility are not offset
  - by proportional changes in expected returns.
- This is a challenge to key insight about the risk-return tradeoff.

How this class is structured

- In today's class, we will study four papers:
  1. Moreira, A., and T. Muir. 2017. Volatility-managed portfolios. *Journal of Finance* 72 (4): 1611–1644. Available from this link.
  2. Cederburg, S., M. S. O'Doherty, F. Wang, and X. Yan. 2020. On the performance of volatility-managed portfolios. *Journal of Financial Economics* 138 (1): 95–117. Available from this link.
  3. Barroso, P., and A. L. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140 (3): 744–767. Available from this link.
  4. DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. Available from this link.
- The first and last papers in detail, the other two papers briefly.

Start of our discussion of Moreira and Muir (2017).

Road map

Outline for our discussion of Moreira and Muir (2017)

- Moreira and Muir (2017) is described in the following steps:
  1. The empirical evidence on volatility-timing strategies
  2. Data used by Moreira and Muir (2017)
  3. Analysis of gains from timing individual factors
  4. Analysis of gains from timing portfolio of multiple factors
  5. The implications of the performance of volatility-timing strategies
- Python code to implement volatility-timing strategies is available from this link, which is written by Alan Moreira.
  - The code includes an excellent explanation of volatility timing.
  - The code relies on access to WRDS, but you can use it instead with the data I have provided for the next assignment.

Road map

What Moreira and Muir (2017) do? . . . I

- Volatility managed portfolio: scale aggregate priced factor by $1/\sigma_t^2$
- Motivation: we know from Markowitz that the demand for a risky asset is:

$$w_t = \frac{1}{\text{[risk aversion]}} \frac{\text{[excess mean returns]}}{\text{[return variance]}} = \frac{1}{\gamma} \frac{\mu_t}{\sigma_t^2}.$$

- If volatility doesn't forecast returns, then a volatility-timing strategy will outperform an unconditional strategy with constant weight, w_t.
  - That is, if μ_t does not change fully when σ_t^2 changes,
  - then, this strategy will make returns in excess of the returns of an unconditional strategy (i.e., it will earn an alpha).

What Moreira and Muir (2017) find?

- Volatility-timing
  - increases Sharpe ratios and
  - generates large alphas relative to the original factors.
- The strategy takes less risk in recessions when σ is high.
  - Strategy sells after the market crashes (1929, 1987, 2008).
  - That is, when volatility increases, returns do not increase proportionately.
  - So, reducing your position when volatility is high reduces risk more than it reduces returns.

Road map

Data used by Moreira and Muir (2017)

- Daily and monthly return data for several factors
- Factors:
  - Market, SMB, HML, UMD, Profitability, ROE, Investment, Carry (FX), BAB.
- Sample periods
  - 1926–2015 (Mkt, SMB, HML, Momentum),
  - Post-1960 for the remaining factors.
- All numbers are annualized.

Road map

How to construct volatility-managed factors

- Define f_{t+1} to be an excess return
  - If you invest one unit of the factor, your return is f_{t+1}
- Construct a new volatility-managed factor, whose return is

$$f^{\sigma}_{t+1} = \frac{c}{\sigma^2_t(f)} \times f_{t+1}, \quad \text{where}$$

  - $\sigma_t(f)$ is the previous month's realized volatility, estimated using daily data
  - choose c so f^{σ} has the same unconditional volatility as f;
  - If you invest $\frac{c}{\sigma^2_t(f)}$ units, your return is $f^{\sigma}_{t+1} = \frac{c}{\sigma^2_t(f)} \times f_{t+1}$.
  - Note that quantity invested decreases as σ^2_t increases.

Calculation of mean-variance weights

- You can then do mean-variance optimization to combine
  1. The original (without timing) factor, f_{t+1};
  2. The volatility-managed version of this factor, f^{σ}_{t+1}.
- The mean-variance optimization will give you the weight to put on
  - the original factor and
  - the volatility-timed factor.
- Comparing the Sharpe ratios of
  - the portfolio with just the original factor and
  - the portfolio that includes also the volatility-timed factor.
- will tell you whether volatility-timing improves performance.
- *This is what the next assignment asks you to do.*

Calculation of alphas

- An "equivalent" way to assess the performance gains from volatility timing is to use regression analysis (Moreira and Muir 2017).
  - Later on we will see that the two are not "equivalent." Why?
- Run the following regression:

$$f^{\sigma}_{t+1} = \alpha + \beta f_{t+1} + \epsilon_{t+1}.$$

- Compute α, which Moreira and Muir (2017) show is theoretically equal to

$$\alpha = -\operatorname{Cov}\left(\frac{\mu_t}{\sigma_t^2}, \sigma_t^2\right)\frac{c}{\mathbb{E}\left[\sigma_t^2\right]};$$

which we interpret on the next slide.

Market return mean and variance

- Consider the market factor
- When studying Black-Litterman model, we showed for the market

$$\gamma_{\mathsf{mkt}} = \frac{\text{expected excess return of market}}{\text{variance of market return}} = \frac{\mu_{t,\mathsf{mkt}}}{\sigma^2_{t,\mathsf{mkt}}}.$$

- If the above result is true, then

$$\begin{aligned}
\alpha &= -\operatorname{Cov}\left(\frac{\mu_{t,\mathsf{mkt}}}{\sigma^2_{t,\mathsf{mkt}}}, \sigma^2_{t,\mathsf{mkt}}\right) \frac{c}{\mathbb{E}\left[\sigma^2_{t,\mathsf{mkt}}\right]} && \ldots \text{from previous page} \\
&= -\operatorname{Cov}\left(\gamma_{\mathsf{mkt}}, \sigma^2_{t,\mathsf{mkt}}\right) \frac{c}{\mathbb{E}\left[\sigma^2_{t,\mathsf{mkt}}\right]} && \ldots \text{setting } \gamma_{\mathsf{mkt}} = \frac{\mu_{t,\mathsf{mkt}}}{\sigma^2_{t,\mathsf{mkt}}} \\
&= 0 && \ldots \text{if } \gamma_{\mathsf{mkt}} \text{ is constant, cov} = 0.
\end{aligned}$$

- So, if $\mu_{t,\mathsf{mkt}} = \gamma_{\mathsf{mkt}}\sigma^2_{t,\mathsf{mkt}}$, we should find $\alpha = 0$.
  - But, if $\alpha \neq 0$, then it implies risk is not proportional to return.

Road map

Univariate α and β for volatility-managed portfolios . . . I

- Moreira and Muir (2017, Panel A of Table 1 – see next page) run a
  - time-series regressions of each volatility-managed factor
  - on the unmanaged factor:

  $$f_t^{\sigma} = \alpha + \beta f_t + \epsilon_t,$$

  where f_t plays the role of the benchmark portfolio return.
- The table shows that the betas are significant, as one would expect.
- The surprising result is that the estimated alphas are significant: they are more than two standard errors away from zero for MKT, UMD, RMW, ROE, IA, BAB (but not for SMB, HML, CMA).

Univariate α and β for volatility-managed portfolios . . . II

	(1) Mkt^{σ}	(2) SMB^{σ}	(3) HML^{σ}	(4) UMD^{σ}	(5) RMW^{σ}	(6) CMA^{σ}	(7) FX^{σ}	(8) ROE^{σ}	(9) IA^{σ}	(10) BAB^{σ}
	Regression betas and standard errors (in parenthesis)									
MktRF	0.61 (0.05)									
SMB		0.62 (0.08)								
HML			0.57 (0.07)							
UMD				0.47 (0.07)						
RMW					0.62 (0.08)					
CMA						0.68 (0.05)				
Carry							0.71 (0.08)			
ROE								0.63 (0.07)		
IA									0.68 (0.05)	
BAB										0.57 (0.05
	Regression alphas and standard errors (in parenthesis)									
α	4.86	-0.58	1.97	12.51	2.44	0.38	2.78	5.48	1.55	5.67
std. error	(1.56)	(0.91)	(1.02)	(1.71)	(0.83)	(0.67)	(1.49)	(0.97)	(0.67)	(0.98)
N	1,065	1,065	1,065	1,060	621	621	360	575	575	996
R^2	0.37	0.38	0.32	0.22	0.38	0.46	0.33	0.40	0.47	0.33
RMSE	51.39	30.44	34.92	50.37	20.16	17.55	25.34	23.69	16.58	29.73

Multivariate alphas controlling for Fama-French 3-factors

- Moreira and Muir (2017, Panel B of Table 1) show that the alphas remain significant
  - for MKT, HML, UMD, RMW, ROE, BAB
  - even after controlling for the Fama-French three factors.

	(1) Mkt^σ	(2) SMB^σ	(3) HML^σ	(4) UMD^σ	(5) RMW^σ	(6) CMA^σ	(7) FX^σ	(8) ROE^σ	(9) IA^σ	(10) BAB^σ
Panel B of Table 1: Regression alphas and standard errors (in parenthesis)										
α	5.45	-0.33	2.66	10.52	3.18	-0.01	2.54	5.76	1.14	5.63
std. error	(1.56)	(0.89)	(1.02)	(1.60)	(0.83)	(0.68)	(1.65)	(0.97)	(0.69)	(0.97)

Performance gains from volatility-managed factors . . . I

- How much do we increase Sharpe ratio/expand the MVE frontier?

- Note that the improvement in Sharpe ratio from using a volatility-managed factor is given by the appraisal ratio;

$$\begin{aligned}\text{Appraisal ratio} &= \text{Additional Sharpe ratio relative to benchmark} \\ &= \frac{\text{additional mean return relative to the benchmark}}{\text{additional risk relative to the benchmark}} \\ &= \frac{\alpha}{\sigma_\epsilon} \\ &= \frac{\alpha}{\text{RMSE}},\end{aligned}$$

  where RMSE is the root mean squared error; i.e., the volatility of the regression residual.

Performance gains from volatility-managed factors . . . II

Additional Sharpe ratio from volatility timing

Factor	α	RMSE	$\frac{\alpha}{\text{RMSE}} \times \sqrt{12}$
MktRf	4.86	51.39	0.32
SMB	−0.58	30.44	−0.06
HML	1.97	34.92	0.19
UMD	12.51	50.37	0.86
RMW	2.44	20.16	0.41
CMA	0.38	17.55	0.07
Carry (FX)	2.78	25.34	0.38
ROE	5.48	23.69	0.80
IA	1.55	16.58	0.32
BAB	5.67	29.73	0.66

- The additional Sharpe ratio generated from volatility-timing is large:
  - For MktRf factor it is 32%;
  - For UMD factor it is 86%;
  - For ROE factor it is 80%;
  - For BAB factor it is 66%.

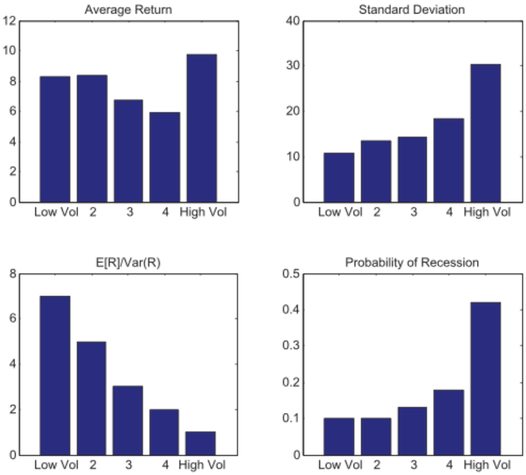
Why does volatility-timing work? . . . I

- To explain why volatility timing performs well, Moreira and Muir provide a nice explanation using pictures for the market factor.
- They use the monthly time series of realized volatility to sort the *following* month's returns into five buckets.
  - The lowest, "low vol," looks at the properties of returns over the month following the lowest 20% of realized volatility months.
- Note that for the market, the average (excess) return per unit of variance represents the optimal weight of a mean-variance investor;
  - it also represents "market's risk-aversion," as we saw when studying the Black-Litterman model:

$$\underbrace{\mathbb{E}_t\left[R_{\mathrm{mkt},t+1}\right]}_{\text{excess return}} = \gamma_{\mathrm{mkt}}\,\sigma^2_{\mathrm{mkt},t}.$$

Why does volatility-timing work? Explained in pictures

- Top, right figure shows:
  Market volatility persists
- Top, left figure shows:
  Lagged market volatility uncorrelated with future returns
- Bottom, left figure shows:
  Market risk-return tradeoff deteriorates with lagged market volatility
- Conclusion: Volatility timing (for market) works due to weak risk-return trade-off



From: Moreira and Muir (2017, Figure 1)

Road map

Empirical results: Gains from timing multiple factors . . . I

- Some investors invest in multiple factors beyond the market
- So, Moreira-Muir extend their analysis to a (static) mean-variance efficient (MVE) portfolio of multiple risk factors.
- Moreira-Muir construct the MVE portfolio to match exactly the procedure for individual factors.
  - Recall that for individual factors, the volatility-managed factor's return is

$$f_{t+1}^{\sigma} = \frac{c}{\sigma_t^2(f)} \times f_{t+1}.$$

Empirical results: Gains from timing multiple factors . . . II

- Let F_{t+1} be a vector of factor returns and b the static weights that produce the maximum in-sample Sharpe ratio.
- Then the MVE portfolio is

$$f_{t+1}^{MVE} = b^{\top} F_{t+1} \qquad \text{. . . optimal combination of } K \text{ factors}$$

- Then, the volatility-timed MVE portfolio is

$$\begin{aligned} f_{t+1}^{MVE,\sigma} &= \frac{c}{[\text{variance of MVE portfolio}]}[\text{return of MVE portfolio}] \\ &= \frac{c}{\hat{\sigma}_t^2\left(f_{t+1}^{MVE}\right)} f_{t+1}^{MVE}, \end{aligned}$$

where again c is a constant that normalizes the variance of the volatility-managed portfolio so it is equal to the MVE portfolio.

Empirical results: Gains from timing multiple factors ... III

- We start by rewriting the volatility-time MVE portfolio:

$$
\begin{aligned}
f_{t+1}^{MVE,\sigma} &= \frac{c}{[\text{variance of MVE portfolio}]}[\text{return of MVE portfolio}] \\
&= \frac{c}{\hat{\sigma}_t^2\left(f_{t+1}^{MVE}\right)} f_{t+1}^{MVE},
\end{aligned}
$$

- From the above, we see that the volatility-managed portfolio
  - shifts the conditional weight on the entire MVE portfolio
  - but does not change the relative weights across the individual factors.

Volatility-timed mean-variance efficient portfolios

- Moreira-Muir consider seven sets of risk factors.
  - FF means Fama and French
  - HXZ means Hou, Xue, and Zhang
- Find substantial performance gains from volatility timing.

Statistic	(1) MKT	(2) FF3	(3) FF3&UMD	(4) FF5	(5) FF5&UMD	(6) HXZ	(7) HXZ&UMD
α	4.86	4.99	4.04	1.34	2.01	2.32	2.51
Standard error	(1.56)	(1.00)	(0.57)	(0.32)	(0.39)	(0.38)	(0.44)
Original Sharpe	0.42	0.52	0.98	1.19	1.34	1.57	1.57
Vol-Managed Sharpe	0.51	0.69	1.09	1.20	1.42	1.69	1.73
Appraisal Ratio	0.33	0.50	0.69	0.56	0.77	0.91	0.91
Observations	1,065	1,065	1,060	621	621	575	575
R^2	0.37	0.22	0.25	0.42	0.40	0.46	0.43
RMSE	51.39	34.50	20.27	8.28	9.11	8.80	9.55

Road map

Implications of profitability of volatility-timing strategies

- Portfolio choice for long term investors
  - Large wealth gains for both short and long-term oriented investors.
- Reduced-form pricing
  - Risk-adjust mutual fund/ hedge fund strategies.
- General-equilibrium asset pricing models
  - Puzzle: the price of risk is low when volatility is high.

Road map

Python code for volatility-timing strategies

- Alan Moreira has written Python code for volatility timing.
  - The code is available from this link.
- Alan Moreira has also written an excellent online book on Quantitative Investing that is available from this link.
  - This book parallels closely our course.
- The Moreira and Muir (2017) paper is an excellent example of very well-executed research.
  - Please read it if you are interested in advanced material.

End of our discussion of Moreira and Muir (2017)

Now we look at two papers criticizing
the result in Moreira and Muir (2017)

The finding of Moreira and Muir (2017) has been criticized

1. Out of sample
   - Cederburg, O'Doherty, Wang, and Yan (2020) show gains from volatility timing cannot be realized out of sample.
2. Transaction costs
   - Barroso and Detzel (2021) show that transaction costs entirely erode the gains.
3. Sentiment
   - Barroso and Detzel (2021) also show that gains from volatility timing the market portfolio are achieved only during periods of "high sentiment."

We first look at the critique in the paper by Cederburg, O'Doherty, Wang, and Yan (2020).

Road map

Cederburg, O'Doherty, Wang, and Yan (2020) . . . I

- Using a comprehensive set of 103 equity strategies, they analyze the value of volatility-managed portfolios for real-time investors.
- Volatility-managed portfolios do not systematically outperform their corresponding unmanaged portfolios in direct comparisons.
- Consistent with Moreira and Muir (2017), volatility-managed portfolios tend to exhibit significantly positive alphas in spanning regressions.

Cederburg, O'Doherty, Wang, and Yan (2020) . . . II

- However, the trading strategies implied by these regressions are not implementable in real time.
    - The regression alpha is based on the entire sample,
    - but an investor would not have access to future data.
- Out-of-sample versions generally earn lower Sharpe ratios than the simple investments in the original, unmanaged portfolios.
- This poor out-of-sample performance for volatility-managed portfolios stems primarily from structural instability in the underlying spanning regressions.

Understanding why the alpha estimated from regression analysis may not be achievable

- Consider the regression to measure the alpha from volatility timing:

$$f_t^\sigma = \alpha + \beta f_t + \epsilon_t.$$

- Gibbons, Ross, and Shanken (1989) explain that to actually earn this alpha, you need to choose the optimal mean-variance portfolio that is a combination of
  1. the factor, f_t, and
  2. the volatility-timing strategy, f_t^σ.
- But the weights of this portfolio have to be chosen in real time
  - using only past (historical) data;
  - without knowing what will happen in the future (no look-ahead bias).
- The regression alpha, because it is estimated from the entire data, suffers from look-ahead bias.

Data used by Cederburg, O'Doherty, Wang, and Yan

- 9 factors used by Moreira and Muir (2017),
- plus 94 anomaly returns from
  - Hou, Xue, and Zhang (2015) and
  - McLean and Pontiff (2016).

Results of Cederburg, O'Doherty, Wang, and Yan

- They directly compare the Sharpe ratios earned by
  - volatility-scaled strategies with those of
  - unscaled strategies.
- Find no systematic evidence that volatility-managed portfolios outperform their unmanaged versions.
  - Volatility scaling generates a higher Sharpe ratio for 5 of the 9 equity factors examined by Moreira and Muir (2017).
  - Volatility-managed versions outperform in 53 cases out of 103.
  - Find that only eight strategies in the broad sample yield statistically significant Sharpe ratio differences in favor of volatility management, which are concentrated among momentum-related strategies.

End of our discussion of
Cederburg, O'Doherty, Wang, and Yan (2020)

Next, we look at the second paper
criticizing Moreira and Muir (2017)

Road map

Barroso and Detzel (2021)

- Barroso and Detzel (2021) study whether "limits to arbitrage" can explain the benefits of volatility-timing strategies.
- In particular, they investigate if
  - transaction costs,
  - arbitrage risk, and
  - short-sale impediments
- explain the abnormal returns of volatility-managed equity portfolios.

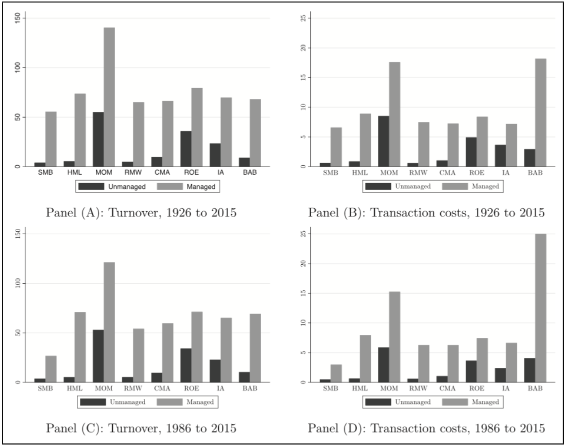
Findings of Barroso and Detzel (2021)

- After transaction costs, volatility management of factors (other than the market) generally significantly reduces Sharpe ratios.
- In contrast, the volatility-managed market portfolio is
  - robust to transaction costs (which are very small for market factor),
  - concentrated in the most easily arbitraged stocks (those with low arbitrage risk and few impediments to short selling).
- But, the volatility-timed market strategy has superior performance
  - only when sentiment is high,
  - consistent with theory that sentiment traders underreact to volatility.

Findings of Barroso and Detzel (2021)

- Factors other than the market have very high transaction costs
  - because they take relatively large positions in small-cap stocks
  - that are expensive to trade.
- Time-varying leverage inherent to volatility-managed portfolios further increases these trading costs by
  - increasing the maximum possible trade size and
  - forcing trades when none would otherwise exist in corresponding unmanaged portfolios.
- Turnover of volatility-managed factors is up to 15 times higher.

Turnover and transaction costs of unmanaged and managed strategies



From: Barroso and Detzel (2021, Figure 2)

End of our discussion of
Barroso and Detzel (2021)

Next, we look at the fourth paper
that resurrects the gains from volatility timing,
DeMiguel, Martín-Utrera, and Uppal (2024)

Road map

DMU address these criticisms of Moreira and Muir (2017)

- DMU = DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. Available from this link.

- DMU show that volatility-timing can improve performance even
  1. Out of sample
  2. Net of transaction costs
  3. Independent of sentiment

Outline of discussion of DMU

1. Introduction . . . motivation and objective
2. Methodology . . . data and modeling approach
3. Results—performance gains . . . what DMU find
4. Source of performance gains . . . what drives DMU's results

Contribution of DMU

- DMU propose a new volatility-timing strategy, with four distinguishing features:

  1. Multifactor, instead of individual-factor portfolios.
  2. Relative factor weights can vary (as a function of market volatility), instead of having a fixed-weight multifactor portfolio.
  3. Account for trading diversification (netting of trades across factors) when computing transaction costs, as in DeMiguel, Martín-Utrera, Nogales, and Uppal (2020).
  4. Optimize factor weights accounting for transaction costs.

DMU: Key finding

- DMU's proposed volatility-managed multifactor strategy outperforms its unconditional counterpart (and Moreira and Muir strategies) even
  1. out-of-sample,
  2. net of transaction costs, and
  3. during both high- and low-sentiment periods.
- Implication: breakdown of risk-return tradeoff even more puzzling.

Road map

Related papers: Timing individual factors

- Factor-momentum strategies, which rely on the positive autocorrelation of factor returns (Ehsani and Linnainmaa 2019; Gupta and Kelly 2019).
- Timing the market using a business-cycle predictor derived from macroeconomic data (Gómez-Cram 2021).
- In contrast, DMU's strategy times multiple factors in combination.

Related papers: Timing combination of factors . . . I

- Dynamic portfolio approach using classification-tree analysis (Miller, Li, Zhou, and Giamouridis 2015).
- Factor portfolios conditional on macroeconomic state variables (Bass, Gladstone, and Ang 2017; Hodges, Hogan, Peterson, and Ang 2017; Amenc, Esakia, Goltz, and Luyten 2019; Bender, Sun, and Thomas 2018).
- Conditioning on macroeconomic regimes identified using Nowcasting (Blin, Ielpo, Lee, and Teiletche 2018).

Related papers: Timing combination of factors . . . II

- Multivariate Markov regime-switching model for the three traditional Fama-French factors (De Franco, Guidolin, and Monnier 2017).
- Timing the market and first five principal components of a large set of equity factors using value spread as timing variable (Haddad, Kozak, and Santosh 2020).
- In contrast, DMU study multifactor portfolios whose
  - weights change with market volatility,
  - which allows us to study breakdown of the risk-return tradeoff.

Related papers: Factor portfolios with transaction costs

- Gupta and Kelly (2019) study performance of factor momentum net of transaction costs.
- Barroso and Detzel (2021) study performance of volatility-managed portfolios net of transaction costs.
- In contrast, DMU's strategy
  - combines factors optimally and
  - accounts for trading diversification (netting out trades across factors before computing the transaction costs).

Road map

Data . . . I

- DMU's data contains all firms traded on
  - NYSE,
  - AMEX, and
  - NASDAQ.
- Use CRSP & Compustat data from Jan. 1967–Dec. 2020.
  - For out-of-sample analysis, DMU use an expanding-window approach, with the first window = 120 months, starting 1967.
  - To ensure fair comparison with out-of-sample results, the in-sample results are reported for the same period, Jan. 1977 to Dec. 2020.

Data . . . II

- Nine factors, as in Moreira and Muir & Barroso and Detzel:
  - Fama and French (2015):
    1. market (MKT),
    2. small-minus-big (SMB),
    3. high-minus-low (HML),
    4. robust-minus-weak (RMW),
    5. conservative-minus-aggressive (CMA)
  - Carhart (1997):
    6. momentum (UMD, up-minus-down)
  - Hou, Xue, and Zhang (2015):
    7. profitability (ROE),
    8. investment (IA).
  - Frazzini and Pedersen (2014)
    9. betting-against-beta (BAB)
- Robustness check: consider larger set of 60 risk factors.

Road map

Methodology in existing literature: (1) Individual factors

- Unmanaged factor return: f_{t+1}
- Volatility-managed factor return: f^{σ}_{t+1}

$$\overbrace{f^{\sigma}_{t+1}}^{\text{managed}} = \text{constant} \times \frac{\overbrace{f_{t+1}}^{\text{unmanaged}}}{\underbrace{(\sigma^f_t)^2}_{\text{scaling}}}$$

  - $(\sigma^f_t)^2$: variance of factor in month t (estimated using realized daily volatility over previous month).
  - "constant" equates volatility of unmanaged & managed factors.
- Combine managed and unmanaged factors to maximize mean-variance utility.

Methodology in existing literature: (2) Fixed-weight multifactor portfolios

- Optimal combination of
  - unconditional mean-variance multifactor portfolio and its
  - volatility-managed counterpart,
    - scaled using unconditional portfolio's past-month return variance.
- This portfolio assigns same relative weight to each factor as the unconditional mean-variance multifactor portfolio
  - thus, DMU call this "conditional fixed-weight multifactor portfolio."

DMU's methodology: General idea

- DMU consider a conditional mean-variance multifactor portfolio that
  - allows relative weights $\theta_{k,t}$ of different factors to vary over time, and
  - uses as the conditioning variable inverse market volatility, σ_t.

$$\theta_{k,t} = a_k + \frac{b_k}{\sigma_t}.$$

Conditional multifactor portfolio return

- Return of conditional multifactor portfolio is

$$r_{p,t+1}(\theta_t) = \sum_{k=1}^{K} r_{k,t+1}\theta_{k,t} = \sum_{k=1}^{K} r_{k,t+1}\left(a_k + \frac{b_k}{\sigma_t}\right).$$

- DMU use this return to do mean-variance optimization.

Conditional mean-variance multifactor portfolio

- Conditional mean-variance multifactor portfolio (CMV) solves

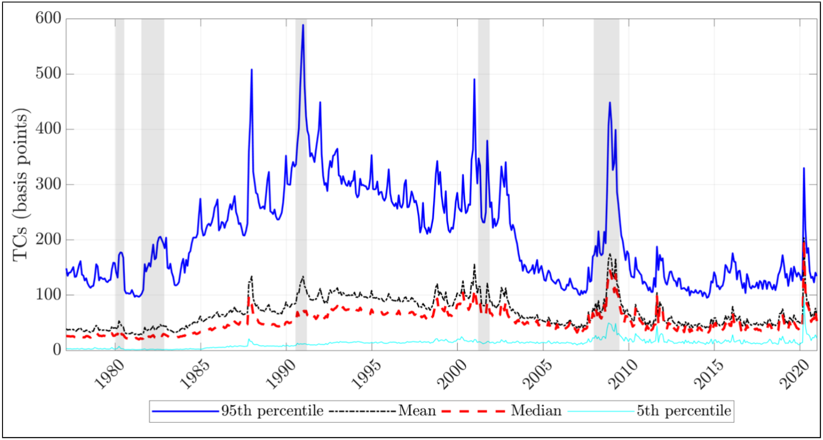
$$\max_{\{a_k \geq 0, b_k \geq 0\}} \quad \mathbb{E}[r_{p,t+1}(\theta_{k,t})] - \mathrm{TC}(\theta_{k,t}) - \frac{\gamma}{2}\mathrm{Var}[r_{p,t+1}(\theta_{k,t})]$$

  - $\mathbb{E}[r_{p,t+1}(\theta_{k,t})]$ is mean of multifactor portfolio return.
  - $\mathrm{TC}(\theta_{k,t})$ is average proportional transaction cost
  - $\mathrm{Var}[r_{p,t+1}(\theta_{k,t})]$ is the variance the multifactor portfolio return
- That is, to find the optimal portfolio, we need to maximize the above objective function over $a_k \geq 0$ and $b_k \geq 0$:
  - a_k: the portfolio component that is unconditional
  - b_k: the portfolio component that is conditional on volatility

Big picture understanding of transaction costs

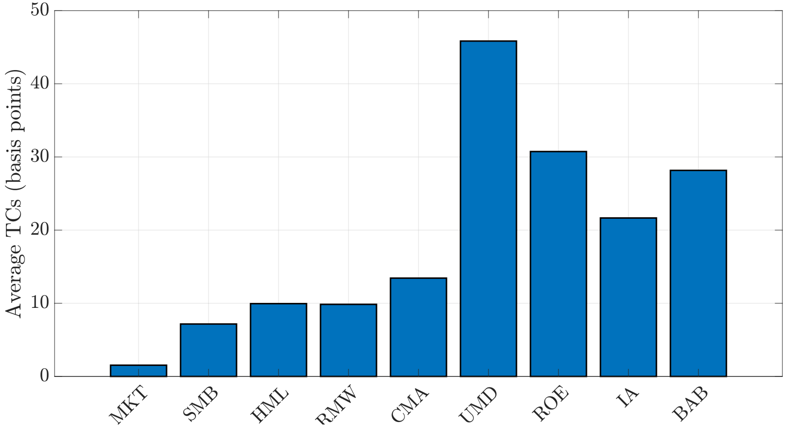
1. Estimate stock-level transaction-cost, as in Abdi and Ranaldo (2017) – figure on next slide shows why it is important to do this.
2. For each factor, determine trading in all the stocks required to rebalance the factor.
3. Net out trades in each stock across the nine factors in the portfolio.
4. Compute aggregate transaction cost for trading all stocks to rebalance all nine factors – figure on next-to-next slide shows that the cost for trading various factors is very different.

Transaction costs at level of individual stocks



- Transaction costs of individual stocks are highly time varying.
- The time variation is particularly strong for less liquid stocks.

Transaction costs for various factors



- As one would expect,
  - the cost for trading the market factor is the lowest;
  - the cost for trading the momentum factor is the highest.

Details: Transaction costs of rebalancing trades

- With trading diversification, transaction cost are estimated as

$$\text{TC} = \frac{1}{T-1}\sum_{t=1}^{T-1} \|\Lambda_t (w_{t+1} - w_t^+)\|_1,$$

  - Λ_t, diagonal transaction-cost matrix whose ith diagonal element contains transaction cost parameter $\kappa_{i,t}$ of stock i,
  - $\|a\|_1 = \sum_{i=1}^{N} |a_i|$ is the 1-norm, and
  - w_t^+ is the conditional mean-variance multifactor portfolio before rebalancing at time $t+1$.

Details: Transaction costs of rebalancing trade

- With trading diversification, transaction cost are estimated as

$$\mathrm{TC} = \frac{1}{T-1} \sum_{t=1}^{T-1} \|\Lambda_t (w_{t+1} - w_t^+)\|_1,$$

  - That is, first net out the trades in each stock across the nine factors.
  - Then, add up the transaction costs to execute the net trades.

- If one ignores trading diversification, transaction costs are given by

$$\mathrm{TC}(\eta) = \frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{k=1}^{K} \|\Lambda_t \left(x_{k,t+1}\theta_{k,t+1} - x_{k,t}^+ \theta_{k,t} \right) \|_1,$$

where $x_{k,t}^+ = x_{k,t} \circ (e_t + r_{t+1})$ is the portfolio before rebalancing.

  - That is, first compute transaction cost for rebalancing each factor;
  - Then, add up the transaction costs across factors;
  - This is suboptimal.

Road map

Start by showing results in the existing literature
but using the DMU dataset.

Performance of individual-factor portfolios

- Performance measured using Sharpe ratio (SR)

	MKT	SMB	HML	RMW	CMA	UMD	ROE	IA	BAB
Panel I: In-sample without transaction costs									
SR(f)	0.530	0.208	0.170	0.506	0.399	0.474	0.722	0.508	0.880
SR(f, f^σ)	**0.585**	**0.246**	**0.215**	**0.739**	**0.419**	**1.088**	**1.153**	**0.621**	**1.397**
p-value($SR(f, f^\sigma) - SR(f)$)	0.244	0.376	0.338	0.038	0.308	0.000	0.001	0.099	0.000
Panel II: In-sample net of transaction costs (but without trading diversification)									
SR(f)	0.519	0.126	0.054	0.357	0.162	0.117	0.313	0.109	0.627
SR(f, f^σ)	**0.521**	0.126	0.054	0.357	0.162	**0.261**	**0.335**	0.109	**0.740**
p-value($SR(f, f^\sigma) - SR(f)$)	0.464	0.500	0.500	0.500	0.500	0.223	0.389	0.500	0.127
Panel III: Out-of-sample and net of transaction costs (but without trading diversification)									
SR(f)	0.519	0.126	0.054	0.357	0.162	0.117	0.313	0.109	0.627
SR(f, f^σ)	0.325	-0.292	-0.038	-0.442	-0.043	**0.204**	0.274	-0.122	**0.727**
p-value($SR(f, f^\sigma) - SR(f)$)	0.979	1.000	0.879	0.999	1.000	0.324	0.672	1.000	0.249

Road map

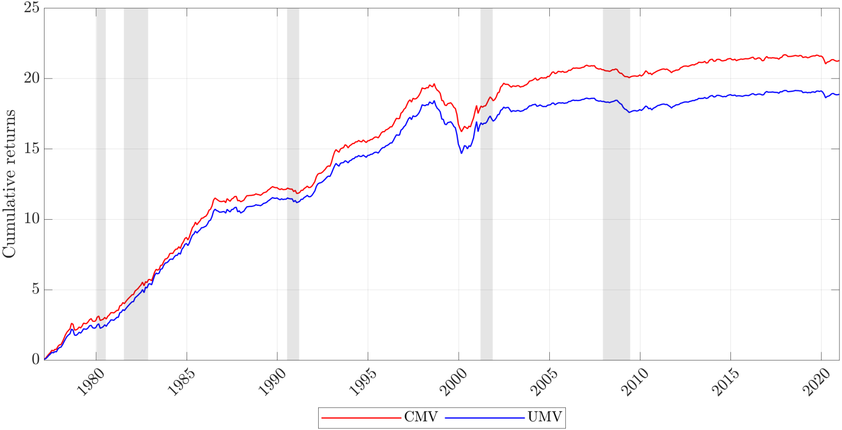
DMU's main results

- DMU's main result: for the new multifactor mean-variance portfolio, volatility timing improves performance even
  - OOS
  - net of TC
  - in both high- and low-sentiment periods

Main result: Performance of conditional multifactor portfolio (out of sample & net of transaction costs)

	unconditional UMV	conditional CMV
Mean	0.446	0.507
Standard deviation	0.459	0.450
Sharpe ratio	0.971	1.126
p-value($SR_{\text{CMV}} - SR_{\text{UMV}}$)	—	0.000
α (%)	—	8.195
$t(\alpha)$	—	4.416
TC	0.149	0.188

Cumulative wealth (OOS & net of TC)



Strategy	Return p.a.	Amt: 1977–2020
Unconditional multifactor portfolio	14.87%	\$445
Conditional multifactor portfolio	17.60%	\$1,253

Road map

Another key finding:

Volatility timing improves portfolio performance
in both high- and low-sentiment periods

Performance (Sharpe ratios) in different sentiment regimes

- Definition of high-sentiment years: sentiment in December of prior year is above its median value for entire sample.

	Entire sample			High sentiment			Low sentiment		
	Uncond.	Cond.	p-val.	Uncond.	Cond.	p-val.	Uncond.	Cond.	p-val.
Panel A: In sample									
Market	0.519	**0.532**	0.295	0.178	**0.217**	0.093	0.954	0.940	0.656
Multifactor	1.130	**1.339**	0.000	1.250	**1.594**	0.000	1.102	**1.150**	0.282
Panel B: Out of sample									
Market	0.519	0.449	0.889	0.178	0.082	0.887	0.954	0.952	0.496
Multifactor	0.971	**1.126**	0.001	1.403	**1.600**	0.003	0.412	**0.569**	0.016

- Out-of-sample, conditional multifactor portfolio significantly outperforms unconditional portfolio during both high and low sentiment.
- Thus, sentiment does not explain the out-of-sample and net-of-costs performance of DMU's conditional multifactor portfolio.

Road map

To identify sources of good performance
decompose the main result
into several small steps

Disentangling the source of the gains

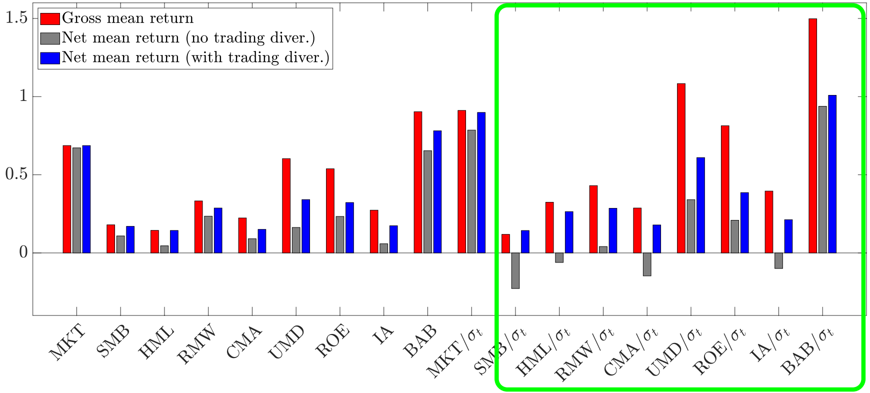
- In the paper,
  - DMU decompose performance into **16 steps**;
  - here, will show only 4 steps.
- Each of the columns reports the performance of multifactor portfolios evaluated in a different way:
  1. in-sample without TC,
  2. out-of-sample (OOS) without TC,
  3. out-of-sample with TC but ignoring trading diversification,
  4. out-of-sample with TC and with trading diversification.

Disentangling source of gains of CMV relative to UMV

	(1) In-sample without TC		(2) Out-of-sample without TC		(3) Out-of-sample with TC no trading diver.		(4) Out-of-sample with TC with trading diver.	
	UMV	CMV	UMV	CMV	UMV	CMV	UMV	CMV
Sharpe ratio	1.379	1.729	1.296	1.543	0.761	0.748	0.971	1.126
p-value($SR_{\text{CMV}} - SR_{\text{UMV}}$)		0.000		0.000		0.682		0.001
α		10.915		12.885		0.357		8.193
$t(\alpha)$		7.625		6.433		0.193		4.416

1. In-sample, CMV significantly outperforms UMV.
2. OOS, UMV & CMV perform less well, but CMV still outperforms UMV.
3. OOS & net of costs, without trading divers., CMV underperforms UMV.
4. OOS & net of costs with trading diversification, CMV outperforms UMV.

Gross and net-of-costs mean factor returns



- TCs reduce returns substantially, especially for managed factors.
- Net mean returns are negative for four factors (ignoring trading diver.)
- But, they have positive net mean returns with trading diversification.

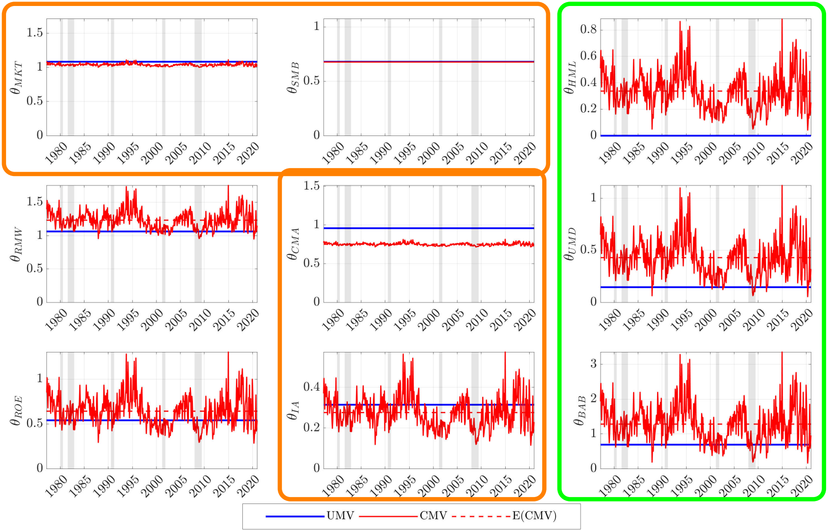
Trading diversification across and within factors

	(1) with full trading div. within & across factors		(2) with trading div. only within factors	(3) without any trading div.
	UMV	CMV	CMV	CMV
Sharpe ratio	0.971	1.126	0.790	0.748
p-value wrt SR_{UMV}		0.000	1.000	1.000
α		8.193	−7.028	−8.879
$t(\alpha)$		4.416	−3.664	−4.589
TC	0.149	0.188	0.339	0.358

- Without trading diversification, $SR_{CMV} < SR_{UMV}$.
- Trading diversification only within factors does not reduce TC much.
- Thus, trading diversification across factors is the main driver of TCs reduction, which can only be achieved with multifactor portfolios.

Time-variation of portfolio weights:

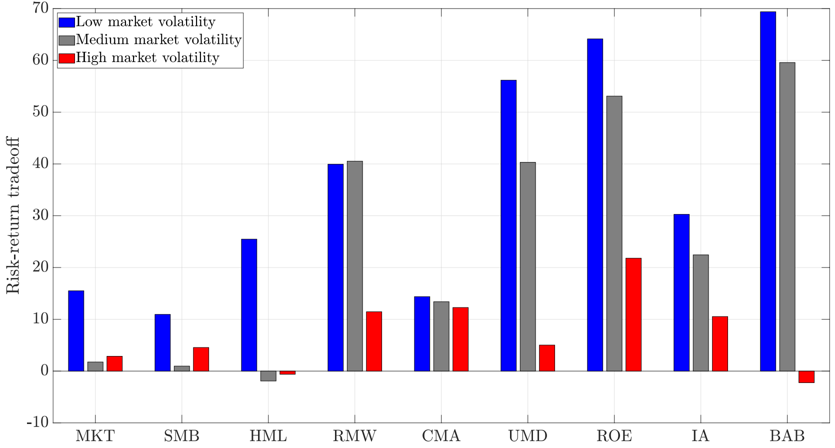
Unconditional (UMV), Conditional mean-variance (CMV)



Market volatility and returns

- DMU wish to study how the risk-return tradeoff for the nine factors in their dataset varies with realized market volatility.
- DMU use the monthly time series of realized market volatility to sort the months in their sample into terciles.
- For each factor, DMU then estimate the risk-return tradeoff for month t as the
  - realized factor return for month $t+1$
  - divided by the monthly realized factor variance estimated as the sample variance of daily returns for month t.
- Finally, they report the risk-return tradeoff averaged across the months in each tercile.

Key takeaway from experiment . . . I



Key takeaway from experiment . . . II

- For all nine individual factors the risk-return tradeoff weakens with market volatility.
- Moreover, the weakening of the risk-return tradeoff is
  - substantial for some of the factors (UMD, ROE, and BAB) but
  - less striking for others (SMB and CMA).
- This motivates using
  - a conditional multifactor portfolio that
  - allows the relative weights of the different factors to vary with market volatility.

Market volatility and factor returns: Regressions . . . I

- Study more formally the tradeoff between realized market volatility and returns of
  - nine individual factors
  - unconditional multifactor portfolio (UMV) using regression analysis.

$$r_{k,t+1} = \alpha + \beta\,\sigma_t + \epsilon_{t+1}$$

- If strong risk-return tradeoff, β should be positive.

Market volatility and factor returns: Regressions . . . II

Factor	MKT	SMB	HML	RMW	CMA	UMD	ROE	IA	BAB	UMV
Intercept	0.709	−0.221	0.567	0.127	0.108	1.248	0.622	0.188	1.471	2.670
t-stat	[1.311]	[−1.046]	[2.200]	[0.735]	[0.741]	[2.472]	[2.547]	[1.434]	[4.965]	[4.578]
Slope (β)	−0.041	0.361	− 0.569	0.118	−0.018	− 1.202	−0.426	−0.141	− 0.904	− 0.742
t-stat	[−0.064]	[1.593]	[−1.953]	[0.554]	[−0.121]	[−1.861]	[−1.389]	[−1.045]	[−2.503]	[−1.145]

- Of the slopes for individual factors, none are significantly positive.
- Slope for UMV portfolio is negative, with a t-stat of −1.145.
- Slopes for HML, UMD, and BAB are negative, with large t-statistics.
- CMV exploits heterogeneity in risk-return tradeoff across factors: times HML, UMD, and BAB more aggressively.

Road map

Robustness checks

1. Consider a large set of 60 factors, instead of just 9.
2. Performance during periods of high volatility.
3. Exclude the market or BAB from the multifactor portfolio.
4. Relax non-negativity constraints.
5. Constrain leverage.
6. Condition on each factor's own volatility, instead of market volatility.
7. Condition on the value spreads in addition to market volatility.
8. Condition on business-cycle variables, besides market volatility.
9. Alternative multifactor portfolio with a general covariance matrix.
10. Evaluate performance using alternative risk measures.
11. Consider investment horizons of up to 18 months.

Road map

Takeaways

- What DMU do: New strategy to time investment in multiple factors
  - Allow relative weight on each factor to vary with market volatility.
  - Optimize for transaction costs with trading divers. across factors.
- What DMU find: Volatility timing leads to large performance gains
  - out-of-sample,
  - net of transaction costs, and
  - for both low- and high-sentiment periods.
- What this means: Breakdown of risk-return tradeoff
  - is even more puzzling than previously thought.

End of discussion of
DeMiguel, Martín-Utrera, and Uppal (2024)

End of focus

Road map

What we plan to do in the next chapter

What's next?

In the next chapter, we will test if it is actually optimal to diversify away unsystematic risk, or, instead, is it better to construct a portfolio to earn the compensation for bearing unsystematic risk.

To do for next class

- Readings
  - The best reference for volatility-timed portfolios is Moreira and Muir (2017); you can download the original article from this link.
  - Python code to implement volatility-timing strategies is available from this link, which is written by Alan Moreira.

Road map

Bibliography . . . I

Abdi, F., and A. Ranaldo. 2017. A simple estimation of bid-ask spreads from daily close, high, and low prices. *Review of Financial Studies* 30 (12): 4437–4480. (Cited on page 973).

Amenc, N., M. Esakia, F. Goltz, and B. Luyten. 2019. Macroeconomic risks in equity factor investing. *Journal of Portfolio Management* 45 (6): 39–60. (Cited on page 961).

Barroso, P., and A. L. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140 (3): 744–767. (Cited on pages 734, 886, 906, 939, 949–953, 963, 966, 1030, 1153).

Bass, R., S. Gladstone, and A. Ang. 2017. Total portfolio factor, not just asset, allocation. *Journal of Portfolio Management* 43 (5): 38–53. (Cited on page 961).

Bawa, V. S., S. Brown, and R. Klein. 1979. *Estimation risk and optimal portfolio choice.* North Holland, Amsterdam. (Cited on pages 458, 732, 884, 1028).

Bender, J., J. Sun, and R. Thomas. 2018. Asset allocation vs. factor allocation—Can we build a unified method? *Journal of Portfolio Management* 45 (2): 9–22. (Cited on page 961).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

Bibliography . . . II

Black, F., and R. Litterman. 1991a. Combining investor views with market equilibrium. *Journal of Fixed Income* 1 (2): 7–18. (Cited on pages 626, 732, 884, 1028).

———. 1991b. Global asset allocation with equities, bonds, and currencies. *Fixed Income Research* 2 (15-28): 1–44. (Cited on pages 626, 732, 884, 1028).

———. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Blin, O., F. Ielpo, J. Lee, and J. Teiletche. 2018. Factor timing revisited: alternative risk premia allocation based on nowcasting and valuation signals. SSRN Working Paper. (Cited on page 961).

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22 (9): 3411–3447. (Cited on pages 733, 839, 841, 844, 847, 885, 1029, 1151).

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, no. 1 (March): 57–82. (Cited on pages 131, 469, 966).

Cederburg, S., M. S. O'Doherty, F. Wang, and X. Yan. 2020. On the performance of volatility-managed portfolios. *Journal of Financial Economics* 138 (1): 95–117. (Cited on pages 734, 886, 906, 939, 940, 942, 943, 945–947, 1030, 1153).

Chen, N.-F., R. Roll, and S. Ross. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403. (Cited on pages 733, 773, 885, 1029).

Bibliography . . . III

De Franco, C., M. Guidolin, and B. Monnier. 2017. The robustness of the volatility factor: linear versus nonlinear factor model. *Journal of Index Investing* 8 (3): 75–88. (Cited on page 962).

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2024. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on pages 734, 829, 886, 1030, 1069, 1155).

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. 2020. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Review of Financial Studies* 33, no. 5 (April): 2180–2222. (Cited on pages 162, 733, 852, 885, 957, 1029).

DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. (Cited on pages 162, 734, 886, 906, 953, 955, 1004, 1030, 1154).

Ehsani, S., and J. T. Linnainmaa. 2019. Factor momentum and the momentum factor. NBER Working Paper. (Cited on page 960).

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2 (June): 427–465. (Cited on pages 625, 726, 733, 759, 760, 885, 1029).

Bibliography . . . IV

Fama, E. F., and K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105 (3): 457–472. (Cited on pages 625, 726, 733, 885, 1029).

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

———. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

Frazzini, A., and L. H. Pedersen. 2014. Betting against beta. *Journal of Financial Economics* 111 (1): 1–25. (Cited on page 966).

Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. A test of the efficiency of a given portfolio. *Econometrica:* 1121–1152. (Cited on page 944).

Gómez-Cram, R. 2021. Late to recessions: stocks and the business cycle. *Journal of Finance* 77 (2): 923–966. (Cited on page 960).

Gupta, T., and B. Kelly. 2019. Factor momentum everywhere. *Journal of Portfolio Management* 45 (3): 13–36. (Cited on pages 960, 963).

Bibliography . . . V

Haddad, V., S. Kozak, and S. Santosh. 2020. Factor timing. *Review of Financial Studies* 33 (5): 1980–2018. (Cited on page 962).

He, G., and R. Litterman. 1999. The intuition behind Black-Litterman model portfolios. *Investment Management Research (Goldman, Sachs & Company).* (Cited on pages 626, 640, 705, 732, 884, 1028).

Hodges, P., K. Hogan, J. R. Peterson, and A. Ang. 2017. Factor timing with cross-sectional and time-series predictors. *Journal of Portfolio Management* 44 (1): 30–43. (Cited on page 961).

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705. (Cited on pages 81, 132, 775, 945, 966).

Huang, S., Y. Song, and H. Xiang. 2020. The smart beta mirage. *Journal of Financial and Quantitative Analysis:* 1–69. (Cited on page 894).

Jorion, P. 1988. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 732, 884, 1028).

Jorion, P. 1985. International portfolio diversification with estimation risk [in English]. *Journal of Business* 58 (3): pp. 259–278. (Cited on pages 458, 467, 500, 732, 884, 1028).

———. 1992. Portfolio optimization in practice. *Financial Analysts Journal* 48 (1): 68–74. (Cited on pages 732, 884, 1028).

Bibliography . . . VI

Klein, R. W., and V. S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3 (3): 215–31. (Cited on pages 465, 732, 884, 1028).

Litterman, R. 2003. *Modern investment management: An equilibrium approach.* New York: Wiley. (Cited on pages 626, 732, 884, 1028).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

McLean, R. D., and J. Pontiff. 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71:5–31. (Cited on page 945).

Miller, K. L., H. Li, T. G. Zhou, and D. Giamouridis. 2015. A risk-oriented model for factor timing decisions. *Journal of Portfolio Management* 41 (3): 46–58. (Cited on page 961).

Moreira, A., and T. Muir. 2017. Volatility-managed portfolios. *Journal of Finance* 72 (4): 1611–1644. (Cited on pages 734, 886, 905–907, 909, 911, 912, 914, 918, 921, 923, 927, 936–939, 942, 945, 947, 955, 966, 1008, 1030, 1152, 1153).

———. 2019. Should long-term investors time volatility? *Journal of Financial Economics* 131 (3): 507–527. (Cited on pages 734, 886, 1030).

Bibliography . . . VII

Pástor, Ľ., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56 (3): 335–381. (Cited on pages 468, 732, 884, 1028).

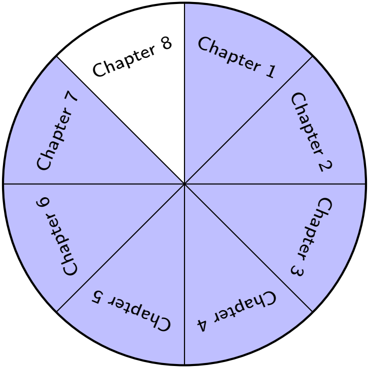
Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary. (Cited on pages 622, 623, 734, 886, 1030, 1111, 1155).

Ross, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13:341–360. (Cited on pages 733, 736, 885, 1029, 1082).

———. 1977. Return, risk, and arbitrage. In *Risk and return in finance,* edited by I. Friend and J. Bicksler. Cambridge, MA: Ballinger. (Cited on pages 733, 736, 885, 1029, 1082).

Rubinstein, M. 2006. *A history of the theory of investments: My annotated bibliography.* Hoboken, NJ: John Wiley & Sons. (Cited on pages 734, 887, 1031).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).



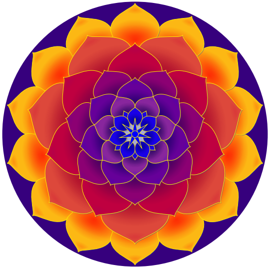
End of Chapter 7

Blank page 1

Blank page 2

Blank page 3

Quantitative Portfolio Management



Chapter 8:
Portfolios Exploiting Systematic and Unsystematic Risk

Raman Uppal

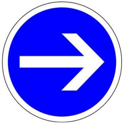
2025

The big picture: Plan for the entire course

Table of contents

Road map

What do we want to do in Chapter 8



In this chapter, first we study the concept of a stochastic discount factor (SDF).

Then we show that an SDF that is good at pricing assets much include compensation for bearing unsystematic risk.

Finally, we study how to build portfolios that exploit both systematic and unsystematic risk.

Timeline: Quantitative portfolio management ideas . . . I

- We can see how ideas about investment have progressed over time.

.... *The thinking in ancient times* ....

- 4th century: $1/N$
  - "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand [cash].
    Rabbi Issac bar Aha, Babylonian Talmud: Tractate Baba Mezi'a, folio 42a
  - "My ventures are not in one bottom trusted"
    "Merchant of Venice, "Shakespeare ((1564–1616) on the importance of diversification in investing
  - Do not put all your eggs in one basket

Timeline: Quantitative portfolio management ideas . . . II

. . . . *Below are the topics we covered in Chapters 3–5* . . . .

- 1950s: Mean-variance optimization
  (Markowitz 1952, 1959)

- 1964: CAPM
  (Sharpe 1964)

- 1970–2000s: Bayesian shrinkage
  (Klein and Bawa 1976; Bawa, Brown, and Klein 1979; Jorion 1985; Jorion 1988; Jorion 1992; Pástor and Stambaugh 2000)

- 1990s: Black-Litterman model
  (Black and Litterman 1990, 1991a, 1991b, 1992; He and Litterman 1999; Litterman 2003)

Timeline: Quantitative portfolio management ideas . . . III

. . . . *Chapter 6* . . . .

- 1970s: Factor models
  (Ross 1976, 1977)

- 1980s Macro factor models
  (Chen, Roll, and Ross 1986)

- 1990–2020s: Fundamental (firm-characteristic-based) factor models
  (Fama and French 1992, 1993, 2012, 2015, 2018).

- 2009-2023: Parametric portfolio policies
  (Brandt, Santa-Clara, and Valkanov 2009; DeMiguel, Martín-Utrera, Nogales, and Uppal 2020).

Timeline: Quantitative portfolio management ideas . . . IV

. . . . *Last time's class: Chapter 7* . . . .

- 2017-2024: Volatility-timing of factors
  - Moreira and Muir (2017, 2019)
  - Cederburg, O'Doherty, Wang, and Yan (2020)
  - Barroso and Detzel (2021)
  - DeMiguel, Martín-Utrera, and Uppal (2024).

. . . . *Today's class: Chapter 8* . . . .

- 2023-2024: Portfolio construction: Beyond systematic risk factors
  - Raponi, Uppal, and Zaffaroni (2023)
  - Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2024).

Timeline: Quantitative portfolio management ideas . . . V

- For a more detailed history of the development of ideas about investment, see the book by Rubinstein (2006).

Road map

Motivation for the material in this chapter

- Existing factor models of asset pricing perform poorly in explaining the cross-section of stock returns;
- This means that portfolios constructed relying on these models are missing an important part of the driver of stock returns.
- In this chapter, we want to find out
  - what is missing in standard models of asset pricing;
  - how to exploit the missing component for portfolio construction.

This chapter is divided into three parts:

1. Understanding the theory of stochastic discount factors (SDFs);
2. Using SDFs to identify what is missing in asset-pricing factor models.
3. Exploiting the missing component for portfolio construction.

Part 1 of this chapter

Understanding the theory of stochastic discount factors (SDFs).

These notes are based on material from the book:

Back, K. E. 2017. *Asset pricing and portfolio choice theory.* Oxford University Press.

Road map

Road map

The stochastic discount factor (SDF) and beta-pricing models

- A key objective of finance is to show how to adjust cash flows
  - for risk and
  - for time.
- We will now study how all the traditional results that you are familiar with can be expressed in terms of the SDF.
- Every asset-pricing model/theory implies a particular SDF:
  - So, all we need to know about any asset-pricing model is the particular SDF implied by it;
  - Conversely, knowing the SDF fully characterizes the asset-pricing model it comes from.

SDF and gross returns

- The price at date $t = 0$ of a non-dividend paying asset, $P_{0,n}$, in terms of the SDF at $t = 1$, m_1, can be expressed as:

$$P_{0,n} = \mathbb{E}\left[m_1 \, P_{1,n}\right].$$

  - The above expression tells us that m_1 has adjusted $P_{1,n}$ for risk and time appropriately, which, therefore, allows us to get $P_{0,n}$.

- If $P_{0,n} > 0$, then we can divide the above equation by $P_{0,n}$ to get:

$$1 = \mathbb{E}\left[m_1 \, \frac{P_{1,n}}{P_{0,n}}\right] = \mathbb{E}\left[m_1 \, R_{1,n}\right], \quad \ldots \text{gross return } R_{1,n} = P_{1,n}/P_{0,n}.$$

  - The above equation says that after we adjust for time and risk via m_1, the expected gross return R_n on every asset is 1.

SDF and expected excess returns . . . I

- In finance, we often study excess returns. So, next we identify the relation between the SDF and excess returns.
- We start by reproducing the last equation on the previous slide:

$$1 = \mathbb{E}\left[m_1 \frac{P_{1,n}}{P_{0,n}}\right] = \mathbb{E}\left[m_1\, R_{1,n}\right], \quad \ldots \text{gross return } R_{1,n} = P_{1,n}/P_{0,n}.$$

- Taking the difference of the two expressions above, we get an expression in terms of excess returns:

$$0 = \mathbb{E}\left[m_1 \left(R_{1,n} - R_{1,m}\right)\right], \quad \ldots \text{where } R_m \text{ is the gross return on asset } m.$$

- The expression above is true for any asset or portfolio.
- It says that once we adjust for risk and time, the excess return on every asset is 0.

Road map

Expected returns in terms of covariance with the SDF . . . I

- Recall from basic statistics that:

$$\mathrm{Cov}[x, y] = \mathbb{E}[x\, y] - \mathbb{E}[x]\,\mathbb{E}[y],$$

which implies that:

$$\mathbb{E}[x\, y] = \mathrm{Cov}[x, y] + \mathbb{E}[x]\,\mathbb{E}[y].$$

- Starting with our basic result and writing m_1 as m and $R_{1,n}$ as R_n:

$$\begin{aligned}
1 &= \mathbb{E}\left[m\, R_n\right], && \text{. . . our basic result} \\
1 &= \mathrm{Cov}[m, R_n] + \mathbb{E}[m]\,\mathbb{E}[R_n] && \text{. . . because } \mathbb{E}[x\, y] = \mathrm{Cov}[x, y] + \mathbb{E}[x]\,\mathbb{E}[y] \\
\mathbb{E}[R_n]\,\mathbb{E}[m] &= 1 - \mathrm{Cov}[m, R_n] && \text{. . . rearranging terms} \\
\mathbb{E}[R_n] &= \frac{1}{\mathbb{E}[m]} - \frac{1}{\mathbb{E}[m]}\,\mathrm{Cov}[m, R_n] && \text{. . . dividing by } \mathbb{E}[m].
\end{aligned}$$

Road map

SDF and the risk-free rate . . . I

- We start with our result on the previous slide:

$$\mathbb{E}[R_n] = \frac{1}{\mathbb{E}[m]} - \frac{1}{\mathbb{E}[m]} \mathrm{Cov}[m, R_n].$$

- Now consider the special case of the risk-free asset, so that
  - $R_n = R_f$, and $\mathrm{Cov}[m, R_f] = 0$,
  - which, using the equation above, then implies that

$$\mathbb{E}[R_f] = R_f = \frac{1}{\mathbb{E}[m]}$$

. . . a useful result that will appear often.

Road map

SDF, risk premia, and covariance . . . I

- Substituting $R_f = \frac{1}{\mathbb{E}[m]}$ in the result

$$\mathbb{E}[R_i] = \frac{1}{\mathbb{E}[m]} - \frac{1}{\mathbb{E}[m]} \operatorname{Cov}[m, R_n]$$

leads to

$$\mathbb{E}[R_n] = R_f - R_f \operatorname{Cov}[m, R_n]$$

which, upon rearranging terms, gives

$$\mathbb{E}[R_n] - R_f = -R_f \operatorname{Cov}[m, R_n] \quad \text{. . . useful result.} \tag{42}$$

SDF, risk premia, and covariance . . . II

- The equation

$$\mathbb{E}[R_n] - R_f = -R_f \,\mathrm{Cov}[m, R_n]$$

  says that the risk-premium on asset n over R_f, is negatively related to the covariance of R_n with the SDF, m.

- The reason for the negative relation between the risk-premium and the covariance of R_n with the SDF m is that
  - when wealth decreases, SDF increases (SDF is like marginal utility);
  - Thus, assets that pay off in states of the world where wealth is low (and SDF is high),
  - need to pay only a small risk-premium (because they have a higher payoff in states where we value this payoff much more).

Road map

Single-factor beta pricing model

- A single-factor beta-pricing model with factor f is said to exist if there are constants R_z and λ such that for each return R_n

$$\mathbb{E}[R_n] = R_z + \lambda \underbrace{\frac{\text{Cov}[f, R_n]}{\mathbb{V}[f]}}_{beta}.$$

- Note that if a risk-free asset exists, then $\text{Cov}[f, R_f] = 0$ and so setting $R_n = R_f$ in the equation above, gives us $R_z = R_f$.

- More generally, R_z (where the "z" stands for "*zero* beta") is the expected value of the return satisfying $\text{Cov}[f, R_z] = 0$.

- In the absence of a risk-free asset, different beta-pricing models can have different expected zero-beta returns (that is, each model may imply a different return for the asset with a beta of zero).

Multifactor beta pricing model

- It is straightforward to extend the single-factor result to a multi-factor setting.

- A multi-factor beta-pricing model with factors $F = (f_1, f_2, \ldots, f_K)^\top$ is said to exist if there are constants R_z and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_K)^\top$ such that for each return R_n

$$\underbrace{\mathbb{E}[R_n]}_{1\times 1} = \underbrace{R_z}_{1\times 1} + \underbrace{\lambda^\top}_{1\times K} \overbrace{\underbrace{V_F^{-1}}_{K\times K} \underbrace{\text{Cov}[F, R_n]}_{K\times 1}}^{\text{betas}},$$

where

  - V_F is the $K \times K$ invertible covariance matrix of the vector F;
  - $\text{Cov}[F, R_n]$ is a $K \times 1$ column vector with elements $\text{Cov}[f, R_n]$;
  - $V_F^{-1}\ \text{Cov}[F, R_n]$ is the vector of multiple regression betas of R_n on F;
  - λ are the $K \times 1$ factor risk premia.

Beta-Pricing Models in Terms of Covariances . . . I

- One can always express beta-pricing models in terms of covariances instead of betas:

- For instance, we can write the single-factor beta model

$$\mathbb{E}[R_n] = R_z + \lambda \underbrace{\frac{\mathrm{Cov}[f, R_n]}{\mathbb{V}[f]}}_{beta}$$

in terms of a covariance

$$\mathbb{E}[R_n] = R_z + \psi \, \mathrm{Cov}[f, R_n], \quad \text{where } \psi = \frac{\lambda}{\mathbb{V}[f]}.$$

Beta-Pricing Models in Terms of Covariances . . . II

- Similarly, we can write the multifactor beta model

$$\mathbb{E}[R_n] = R_z + \lambda^\top V_F^{-1} \operatorname{Cov}[F, R_n],$$

in terms of covariances instead of betas

$$\mathbb{E}[R_n] = R_z + \psi^\top \operatorname{Cov}[F, R_n], \quad \text{where } \psi = V_F^{-1}\lambda.$$

Number of factors is not uniquely determined

- It is important to realize that the number of factors in a beta-pricing model is not uniquely determined.
- For example, given a K-factor model, one can always use $\lambda^\top V_F^{-1} F$ as a single factor.
- One can also use the SDF m as the single systematic factor, provided $\mathbb{E}[m] \neq 0$, as we showed earlier, and show again below:

$$\mathbb{E}[R_n] = \frac{1}{\mathbb{E}[m]} - \frac{1}{\mathbb{E}[m]} \operatorname{Cov}[m, R_n]$$

with $1/\mathbb{E}[m] = R_f$ if there exists a risk-free asset.

Single-factor Models with *returns* as factors . . . I

- In general, a factor need not be a return.
  - For example, macro-factors, such as inflation, are not returns.
- But some factors may be returns; for example, the return on the market factor.
- If a factor is a return, then its factor risk premium is its ordinary risk premium, treating R_z as a proxy for the risk-free return.

Single-factor Models with *returns* as factors . . . II

- To see this, suppose that there is a single-factor beta-pricing model with the factor being the return R_*. Then,

$$\mathbb{E}[R_n] = R_z + \lambda \frac{\text{Cov}[R_*, R_n]}{\mathbb{V}[R_*]}.$$

- But, the above equation must be true for $R_n = R_*$

$$\begin{aligned}
\mathbb{E}[R_*] &= R_z + \lambda \frac{\text{Cov}[R_*, R_*]}{\mathbb{V}[R_*]} && \text{. . . the general result from earlier} \\
&= R_z + \lambda \frac{\mathbb{V}[R_*]}{\mathbb{V}[R_*]} && \text{. . . because } \text{Cov}[R_*, R_*] = \mathbb{V}[R_*] \\
&= R_z + \lambda, && \text{. . . } \mathbb{V}[R_*]/\mathbb{V}[R_*] = 1
\end{aligned}$$

and so, rearranging terms, we get that

$$\lambda = \mathbb{E}[R_*] - R_z.$$

Beta-pricing model with a return as a single factor

- Substituting $\lambda = \mathbb{E}[R_*] - R_z$ into our earlier result

$$\mathbb{E}[R_n] = R_z + \lambda \frac{\text{Cov}[R_*, R_n]}{\mathbb{V}[R_*]},$$

  leads to the following new result:

$$\mathbb{E}[R_n] = R_z + (\mathbb{E}[R_*] - R_z) \frac{\text{Cov}[R_*, R_n]}{\mathbb{V}[R_*]}.$$

- Thus, there is a beta-pricing model with a return as a single factor if and only if
  - the return is on mean-variance frontier and
  - is not equal to the risk-free rate, if one exists, and
  - otherwise not equal to the global minimum-variance portfolio (GMV).

The CAPM

- The Capital Asset Pricing Model (CAPM) states that the result on the previous slide, reproduced below

$$\mathbb{E}[R_n] = R_z + \lambda \frac{\mathrm{Cov}[R_*, R_n]}{\mathbb{V}[R_*]},$$

is true if the factor is the return on the market portfolio, $R_* = R_{mkt}$:

$$\mathbb{E}[R_n] = R_z + (\mathbb{E}[R_{mkt}] - R_z) \frac{\mathrm{Cov}[R_{mkt}, R_n]}{\mathbb{V}[R_{mkt}]},$$

which is equivalent to

  - the market portfolio being on the mean-variance frontier, and
  - not equal to the risk-free rate, if one exists, and
  - otherwise not equal to global minimum-variance portfolio.

Road map

SDF and beta pricing

- There is a beta-pricing model with respect to some factors with the expected zero-beta return being non-zero
- if and only if there is an SDF m such that:
  1. $\mathbb{E}[m] \neq 0$; and,
  2. m is an affine function of the factors: $m = a + b\,f$.
- This is true for both single- and multi-factor models.
- In a single-factor model such as

$$\mathbb{E}[R_n] = R_z + \psi \ \mathrm{Cov}[f, R_n], \quad \text{where: } \psi = \frac{\lambda}{\mathbb{V}[f]} \text{ and } R_z \neq 0.$$

- the SDF is given by:

$$m = \frac{1}{R_z} - \frac{1}{R_z}\psi\,(f - \mathbb{E}[f]).$$

How to prove that m is a valid SDF? ... I

- Suppose that there is a one-factor beta-pricing model

$$\mathbb{E}[R_n] = R_z + \psi \operatorname{Cov}[f, R_n].$$

- We now wish to show that the following is a valid SDF:

$$m = \frac{1}{R_z} - \frac{1}{R_z}\psi\left(f - \mathbb{E}[f]\right).$$

Proof of being a valid SDF . . . I

- To show that $m = \frac{1}{R_z} - \frac{1}{R_z}\psi\left(f - \mathbb{E}[f]\right)$ is a valid SDF,
  - we need to show that $\mathbb{E}[m\, R_n] = 1$.
- To prove this, start by multiplying m by R_n:

$$m\, R_n = \frac{R_n}{R_z} - \frac{1}{R_z}\left(\psi\, f\, R_n - \mathbb{E}[\psi\, f] R_n\right).$$

- Now take expectations;

$$\mathbb{E}[m\, R_n] = \frac{\mathbb{E}[R_n]}{R_z} - \frac{1}{R_z}\Big(\mathbb{E}[\psi\, f\, R_n] - \mathbb{E}[\psi\, f]\mathbb{E}[R_n]\Big).$$

- On the next page, we simplify this expression.

Proof of being a valid SDF . . . II

- We write the expression from the previous page and then simplify it:

$$
\begin{aligned}
\mathbb{E}[m\,R_n] &= \frac{\mathbb{E}[R_n]}{R_z} - \frac{1}{R_z}\Big(\mathbb{E}[\underbrace{\psi\, f}_{x}\ \underbrace{R_n}_{y}] - \mathbb{E}[\underbrace{\psi\, f}_{x}]\mathbb{E}[\underbrace{R_n}_{y}]\Big) \\
&= \frac{\mathbb{E}[R_n]}{R_z} - \frac{1}{R_z}\,\mathrm{Cov}[\psi\, f, R_n] \qquad \ldots \text{because } \mathrm{Cov}[x, y] = \mathbb{E}[x\,y] - \mathbb{E}[x]\,\mathbb{E}[y] \\
&= \frac{1}{R_z}\,(\mathbb{E}[R_n] - \psi\ \mathrm{Cov}[f, R_n]) \quad \ldots \text{because } \psi \text{ is a constant} \\
&= \frac{1}{R_z}\,(R_z) \qquad \ldots \text{from } \mathbb{E}[R_n] = R_z + \psi\ \mathrm{Cov}[f, R_n] \\
&= 1.
\end{aligned}
$$

- So, we have proved that indeed $\mathbb{E}[m\,R_n] = 1$.

Road map

What is the SDF in the CAPM?

- If the CAPM is true, then

$$f = R_{mkt} \qquad \text{and} \qquad \psi = \frac{\mathbb{E}[R_{mkt}] - R_z}{\mathbb{V}[R_{mkt}]}.$$

- Substituting the above expression into our general result, which is reproduced below

$$m = \frac{1}{R_z} - \frac{1}{R_z}\psi\left(f - \mathbb{E}[f]\right),$$

- gives us the following result for the SDF in the CAPM:

$$m = \frac{1}{R_z} - \frac{1}{R_z}\underbrace{\left(\frac{\mathbb{E}[R_{mkt}] - R_z}{\mathbb{V}[R_{mkt}]}\right)}_{\psi}\left(\underbrace{R_{mkt}}_{f} - \mathbb{E}[\underbrace{R_{mkt}}_{f}]\right).$$

Road map

Mean-variance portfolio weights in terms of the SDF

- We now show how the mean-variance portfolio weights are related to the SDF.
- Denote by R^e the vector of excess returns on N assets (i.e., in excess of the risk-free rate)
- Recall our condition for mean-variance optimal portfolio weights:

$$w = \frac{1}{\gamma} V_{R^e}^{-1} \mathbb{E}[R^e].$$

- Recall our result in Equation (42) on page 1046:

$$\mathbb{E}[R^e] = -R_f \operatorname{Cov}[m, R^e].$$

- Substituting $\mathbb{E}[R^e]$ into the expression for the optimal weights gives

$$w = (-)\frac{1}{\gamma} V_{R^e}^{-1} R_f \operatorname{Cov}[m, R^e] \quad \ldots \text{mean-variance weights in terms of } m.$$

SDF in terms of mean-variance portfolio returns

- Conversely, the SDF can be expressed in terms of the return on the mean-variance optimal portfolio, R^e_{MV}; i.e.,

$$m = 1 - \gamma R^e_{MV} = 1 - \gamma(w^\top R^e), \quad \text{where} \quad w = \frac{1}{\gamma} V_{R^e}^{-1} \mathbb{E}[R^e]$$

and $R^e_{MV} = (w^\top R^e)$ is the excess return on the MV optimal portfolio.

- **Proof:**

$$\begin{aligned}
\mathrm{Cov}[m, R^e] &= \mathrm{Cov}[1 - \gamma(w^\top R^e), R^e] && \text{. . . substituting the definition of } m \\
&= -\gamma\, \mathrm{Cov}[(w^\top R^e), R^e] && \text{. . . moving } \gamma \text{ out of cov operator} \\
&= -\gamma w\, \mathrm{Cov}[R^e, R^e] && \text{. . . moving } w \text{ out of cov operator} \\
&= -\gamma w \mathbb{V}[R^e] && \text{. . . } \mathrm{Cov}[x, x] = \mathbb{V}[x],
\end{aligned}$$

which, upon rearranging (i.e., multiplying both sides by $-\gamma^{-1}(\mathbb{V}[R^e])^{-1}$), gives

$$w = (-)\frac{1}{\gamma} V_{R^e}^{-1}\, \mathrm{Cov}[m, R^e].$$

This brings us to an end to Part 1 of this chapter on understanding the SDF.

We now apply this knowledge to characterize the SDF empirically.

Part 2 of this chapter

Use SDF to show that expected returns include compensation for unsystematic risk.

This part of the chapter is based on the paper:

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2024. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146 and can be downloaded using this link.

Start of focus

Road map

The SDF with systematic and unsystematic risk

- In this part of the chapter, we study the drivers of the SDF.
- In particular,
  - we study which systematic risk factors drive the SDF;
  - whether unsystematic risk is an important driver of the SDF.

Road map

Objective and Motivation

- Major challenge in finance is to price the cross-section of stock returns.
  - That is, to explain why stocks differ in their expected returns?
- The answer to this question is essential for
  - Portfolio-investment: optimal portfolio design;
  - Risk-management: how to adjust returns for risk.
  - Capital-budgeting: cost of capital;

The world of Sharpe (1964) . . . perfect diversification

- In the world of **frictionless markets** assumed by the CAPM of Sharpe (1964),
  - investors hold a perfectly diversified portfolio of risky assets,
  - which, in equilibrium, is the market portfolio.
  - Thus, the only source of risk is the systematic risk of the market.

In the real world . . . less than perfect diversification

- In the data, however,
  - investors do not hold diversified portfolios, and
  - risk compensation for asset exposures to the market factor, performs poorly in explaining the cross-section of expected stock returns.

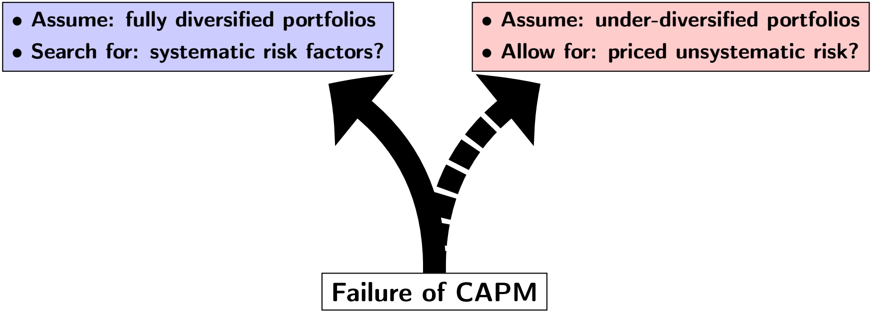
Post-CAPM . . . factor zoo

- When the CAPM performed poorly in the cross-section, researchers
  - continued to assume that investors hold perfectly diversified portfolios,
  - empirically examined a large number of alternative proxies for systematic risk,
  - leading to the **factor zoo** (Cochrane 2011).

Status of models with systematic risk factors

- Of the more than four hundred systematic risk factors from the factor zoo, none explains the cross-section of stock returns.
- Virtually all models—more than 2000 trillion—featuring factors from this zoo have sizable pricing errors (Bryzgalova, Huang, and Julliard 2023).
- Bottom line: Existing factor models cannot explain differences in expected stock returns.

Which road to take?



What is unsystematic risk?

- Our research focuses on a single question:

Q. Is unsystematic risk priced (i.e., rewarded)?

Unsystematic risk = **Total risk** − **Systematic risk**

- What is unsystematic risk?
- Unsystematic risk can represent
  - Asset-specific risk (i.e., risk specific to a particular stock).
  - Weak factors (i.e., factors that explain returns of only a small number of assets).

The world of Merton (1987) . . . imperfect diversification

- Merton (1987) relaxes the assumption of frictionless markets and derives an equilibrium in which investors hold underdiversified portfolios.
  - consistent with a large body of empirical evidence.
- If investors hold underdiversified portfolios, they will demand compensation for bearing unsystematic risk.
- Building on Merton (1987), we study how compensation for unsystematic risk
  - provides an avenue for explaining the cross-section of expected asset returns and
  - resolving the factor zoo.

Our work is founded on the Arbitrage Pricing Theory (APT)

- Instead of specifying a particular equilibrium model like Merton (1987),
- we use instead the Arbitrage Pricing Theory (APT) of Ross (1976, 1977)
- to explore the possibility that unsystematic risk is compensated (priced),
- So we call our model "Priced Unsystematic Risk (PUR) model."

Benefits of using the APT

- There are several benefits to using the APT:
  - It is agnostic about the relevant systematic risk factors, which are latent.
  - It allows risk premium to deviate from reward for bearing only systematic risk.

$$\begin{array}{c}\text{Total expected}\\ \text{excess returns}\end{array} = \begin{array}{c}\text{Expected return for}\\ \text{systematic risk}\end{array} + \begin{array}{c}\text{Expected return for}\\ \text{unsystematic risk}\end{array}$$

Arbitrage Pricing Theory (APT) . . . our starting point

- APT makes only two assumptions:
  1. Gross unexpected returns are given by a K latent linear factor model

$$R_{t+1} - \mathbb{E}(R_{t+1}) = \beta\big(f_{t+1} - \mathbb{E}(f_{t+1})\big) + e_{t+1},$$

     - f_{t+1} is $K \times 1$ vector of common (latent) risk factors with risk premia λ and covariance matrix V_f;
     - $\beta = (\beta_1, \beta_2, \ldots\ , \beta_N)'$ is the $N \times K$ full-rank matrix of loadings;
     - e_{t+1} is vector of unsystematic shocks with zero mean and $N \times N$ covariance matrix $V_e > 0$.
  2. Asymptotic no-arbitrage: as N increases, it is not possible to have a portfolio whose risk goes to zero, but its return is strictly positive.

From the two assumptions of the APT, we get ...

- Expected excess returns are:

$$\begin{array}{ccccc} \text{Expected excess returns} & = & \text{Expected return for systematic risk} & + & \text{Expected return for unsystematic risk} \\ \mathbb{E}(R^e_{t+1}) & = & \beta\lambda & + & \alpha \end{array}$$

with the vector α satisfying the asymptotic no-arbitrage restriction: as $N \to \infty$:

$$\underbrace{\alpha' V_e^{-1} \alpha}_{(\mathsf{SR}^{\alpha})^2} \leq \delta_{\mathsf{apt}}{}^2 < \infty$$

- By construction, unsystematic risk is *orthogonal* (uncorrelated) to systematic risk.

Road map

Our estimation approach

- Our estimation approach is . . . simple and exhaustive.

Model selection and parameter estimation

- To estimate the data-generating process of asset returns with K latent factors

$$R_{t+1} - \mathbb{E}(R_{t+1}) = \beta(f_{t+1} - \mathbb{E}(f_{t+1})) + e_{t+1} \tag{43}$$

$$\text{subject to: } \alpha' V_e^{-1} \alpha \leq \delta_{\text{apt}}^2, \tag{44}$$

- we need to determine only two parameters:
  - K = number of systematic (latent) risk factors;
  - δ_{apt} = magnitude of the Sharpe ratio of the alpha portfolio.

Our estimation approach

- Estimate K using
  - Statistical methods of factor analysis:
    1. Ahn and Horenstein (2013),
    2. Onatski (2012), and
    3. Scree plots
  - Cross-validation (CV), to avoid overfitting
- Estimate δ_{apt} by five-fold cross-validation with four performance metrics
  1. Hansen-Jagannathan distance (HJ) . . . Financial economist
  2. Root-mean-square error (RMSE) . . . Statistician
  3. Sharpe ratio (SR) . . . Asset manager
  4. Generalized-least-square (R2GLS) . . . Econometrician

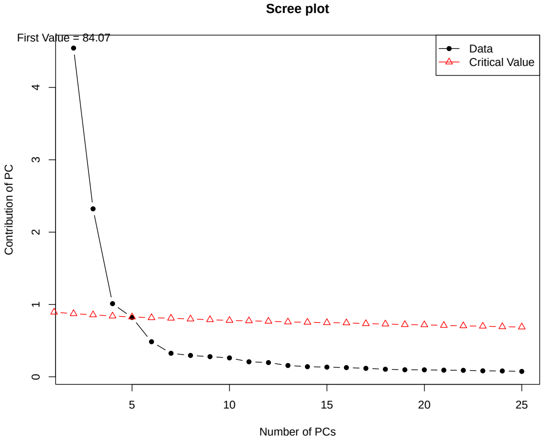
Road map

Data

- Basis assets: 303 characteristic-based double-sorted portfolios with monthly returns from 1963:07 to 2024:12.
  - 25 portfolios each sorted by Size and: Book-to-Market, Operating profitability, Investment, Momentum, Short-term reversal, Long-term reversal, Accruals, Market beta, Variance or Residual variance, (for a total of 25 x 10)
  - 35 portfolios sorted by Size and: Net Share Issues
  - 6 portfolios sorted by Size and: Earnings/Price, Cashflow/Price, Dividend yield
- We also consider several other datasets (including, in a companion paper, data on individual stock returns from the S&P 500).

Road map

Number of systematic (latent) risk factors: $K = \{1, 3, 4\}$



- Ahn and Horenstein: 1 factor
- Onatski: 3 factors
- Scree plot: 4 factors

How to find if unsystematic risk is compensated?

- Q. Do expected excess returns include compensation for unsystematic risk?
- Compare performance of two models across four metrics
  1. PCA model that includes only common (latent) risk factors (i.e., $\delta_{\text{apt}} = 0$)
  2. Our model (DUZZ) that includes
     - the same risk factors as the PCA model
     - and compensation for unsystematic risk: $\delta^2_{\text{apt}} = \underbrace{\alpha' V_e^{-1} \alpha}_{(\text{SR}^\alpha)^2} > 0$

Is $\delta_{\text{apt}} = (\text{SR}^{\alpha}) > 0$ when K fixed by statistical criteria? Yes

- Fix number of factors K as per statistical criteria: $K = \{1, 3, 4\}$
- Look for optimal δ_{apt} in cross-validation using the four metrics

		Metric for estimating δ_{apt} and evaluating performance							
		HJ_{reg}		RMSE$\times\sqrt{N}\times 100$		SR		R^2_{gls}	
Model (1)	K (2)	δ_{apt} (3)	Value (4)	δ_{apt} (5)	Value (6)	δ_{apt} (7)	Value (8)	δ_{apt} (9)	Value (10)
Panel A: When K is estimated by ER									
PUR-ER	1	0.20	**0.67**	0.80	**7.11**	0.63	**0.57**	0.72	**0.21**
PCA-ER	1	0.00	0.78	0.00	7.95	0.00	0.14	0.00	−0.16
Panel B: When K is estimated by GR									
PUR-GR	1	0.20	**0.67**	0.80	**7.11**	0.63	**0.57**	0.72	**0.21**
PCA-GR	1	0.00	0.78	0.00	7.95	0.00	0.14	0.00	−0.16
Panel C: When K is estimated by Onatski									
PUR-Onatski	3	0.16	**0.71**	0.89	**7.15**	0.88	**0.55**	0.73	**0.21**
PCA-Onatski	3	0.00	0.78	0.00	7.57	0.00	0.23	0.00	0.00
Panel D: When K is estimated by scree plots									
PUR-scree plot	4	0.15	**0.70**	0.85	**7.15**	0.54	**0.55**	0.71	**0.19**
PCA-scree plot	4	0.00	0.77	0.00	7.46	0.00	0.27	0.00	0.04

Is $\delta_{\text{apt}} > 0$ if we choose both K and δ_{apt} in cross-validation? "Yes"

	Metric for estimating K, δ_{apt}, and K^{pc} and for evaluating performance											
	HJ_{reg}			RMSE$\times\sqrt{N}\times 100$			SR			R^2_{gls}		
Model (1)	K, K^{pc} (2)	δ_{apt} (3)	Value (4)	K, K^{pc} (5)	δ_{apt} (6)	Value (7)	K, K^{pc} (8)	δ_{apt} (9)	Value (10)	K, K^{pc} (11)	δ_{apt} (12)	Value (13)
PUR	1	0.20	**0.67**	1	0.80	**7.11**	1	0.63	0.57	1	0.72	**0.21**
PCA	4	0.00	0.77	153	0.00	7.21	123	0.00	**0.58**	51	0.00	0.13

- Performance of PCA model improves with larger K;
- Recall, however, that number of latent risk factors is at most $K = 4$;
- So, remaining risk factors represent weak factors (unsystematic risk).

Is $\delta_{\text{apt}} > 0$ if we define "systematic factor" more generally?

1. Lettau and Pelger (2020) define a factor to be systematic if
   - large contribution to the covariance of returns
   - or large contribution to cross-sectional variation of *expected returns*
2. Kozak, Nagel, and Santosh (2020) define a factor to be systematic if
   - large contribution to the covariance of returns
   - and high price of risk (i.e., high expected returns)

Is $\delta_{\text{apt}} > 0$ if we define "systematic factor" differently? Yes

- Each model was estimated as recommended in the original paper.

		Metric for estimating $\delta_{\text{apt}}/\kappa/\gamma$ and evaluating performance							
		HJ_{reg}		RMSE$\times$ $\sqrt{N}\times 100$		SR		R^2_{gls}	
Model (1)	K (2)	$\delta_{\text{apt}}/\kappa/\gamma$ (3)	Value (4)	$\delta_{\text{apt}}/\kappa/\gamma$ (5)	Value (6)	$\delta_{\text{apt}}/\kappa/\gamma$ (7)	Value (8)	$\delta_{\text{apt}}/\kappa/\gamma$ (9)	Value (10)
Panel A: PUR models with K estimated by statistical criteria									
PUR-ER	1	0.20	**0.67**	0.80	**7.11**	0.63	**0.57**	0.72	**0.21**
PUR-GR	1	0.20	**0.67**	0.80	**7.11**	0.63	**0.57**	0.72	**0.21**
PUR-Onatski	3	0.16	0.71	0.89	7.15	0.88	0.55	0.73	0.21
PUR-scree plot	4	0.15	0.70	0.85	7.15	0.54	0.55	0.71	0.19
Panel B: Traditional models with observable risk factors									
CAPM	1	—	0.79	—	8.58	—	0.14	—	−0.12
FF3	3	—	0.79	—	8.01	—	0.19	—	−0.11
FFC	4	—	0.78	—	7.78	—	0.29	—	0.00
FF5	5	—	0.80	—	7.54	—	0.30	—	0.03
FF6	6	—	0.85	—	7.45	—	0.27	—	0.06
Panel C: State-of-the-art models with a more general definition of systematic risk									
KNS	6	1.18	0.76	16.00	7.36	16.00	0.28	5.23	0.09
LP	5	10.00	0.95	10.00	7.19	10.00	0.33	10.00	0.15
LP	5	20.00	0.97	20.00	7.19	20.00	0.35	20.00	0.14

Road map

Out-of-sample analysis using six subperiods

- Divide the entire sample into six periods, approximately ten years each (with 123 monthly observations).
- Then, estimate the models on one ten-year window.
- Evaluate it on the next ten-year window.

Is $\delta_{\text{apt}} > 0$ even out-of-sample? Yes

Model	Metric for estimating and evaluating models			
	$\text{HJ}_{\textbf{reg}}$	RMSE× $\sqrt{N}$ × 100	SR	$R^2_{\textbf{gls}}$
(1)	(2)	(3)	(4)	(5)
Panel A: When K is estimated by ER, and $\delta_{\textbf{apt}}$ by CV with each metric				
PUR	**0.40**	**2.76**	**0.51**	**0.73**
PCA	0.55	5.02	0.14	-0.11
Panel B: When K is estimated by GR, and $\delta_{\textbf{apt}}$ by CV with each metric				
PUR	**0.42**	**2.23**	**0.51**	**0.76**
PCA	0.55	3.95	0.15	0.02
Panel C: When K is estimated by Onatski, and $\delta_{\textbf{apt}}$ by CV with each metric				
PUR	**0.46**	**2.28**	**0.50**	**0.73**
PCA	0.54	3.81	0.18	0.03
Panel D: When K is estimated by scree plot, and $\delta_{\textbf{apt}}$ by CV with each metric				
PUR	**0.45**	**2.29**	**0.52**	**0.74**
PCA	0.57	4.15	−0.04	-0.03
Panel E: When K and $\delta_{\textbf{apt}}$ for PUR model and K^{pc} for PCA model are estimated by CV				
PUR	**0.40**	**2.70**	**0.46**	**0.69**
PCA	0.55	3.11	−0.24	0.45

Results of other out-of-sample are even stronger.

Is unsystematic risk persistent over time? Yes

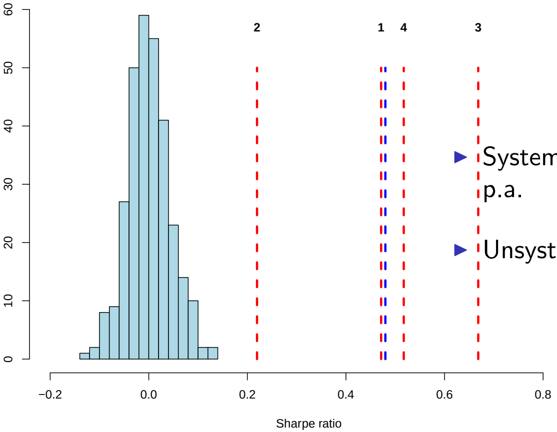
Correlations of unsystematic risk across subperiods

	1973–1984	1984–1994	1994–2004	2004–2014	2014–2024
Panel: K is chosen by Ahn-Horenstein					
1973–1984	1.00	0.62	0.56	0.34	0.31
1984–1994		1.00	0.24	0.26	0.07
1994–2004			1.00	0.30	0.68
2004–2014				1.00	0.09
2014–2024					1.00

Road map

Is priced unsystematic risk important?

- To be conservative, look at the case of $K = 4$ and use the HJ metric.



- Systematic SR ($K = 4$) is 0.99 p.a.
- Unsystematic SR is 0.48 p.a.

Properties of unsystematic component of present-value operator (SDF)

- Small stocks contribute most to the unsystematic SDF component.
    - The most prominent portfolios are those sorted by Momentum, Short-term reversal, Variance, Residual FF3 variance, and Investment.
- Variation of the unsystematic SDF component,
    - Short-Term Reversal Factor explains less than 5%,
    - Investment factor explains less than 6%,
    - Idiosyncratic-volatility factor of Ang et. al (2006) explains less than 8%, and
    - Momentum factor explains less than 12%.

Properties of systematic component of present-value operator (SDF)

- Market factor explains most of the variation (88%) of the first PC, PC_1.
- Size factor of FF explains most of the variation in PC_2;
- Value and Profitability factors of FF explain most of the variation in PC_3; and,
- Momentum factor explains most of the variation in PC_4.

Road map

Conclusion

- Developed a method to examine potential significance of unsystematic risk.
- Key result: establish quantitative importance of priced unsystematic risk
- Our finding has important implications for any question involving the risk-return tradeoff; particularly, portfolio selection, asset allocation, and risk management.

This brings us to the end to Part 2 of this chapter on the importance in the cross-section of stock returns of not just compensation for systematic risk but also unsystematic risk.

Part 3 of this chapter

Exploiting unsystematic risk for portfolio construction.

This part of the chapter is based on the paper:

Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary which can be downloaded using this link.

Road map

Portfolio construction with systematic & unsystematic risk

- The optimal mean-variance portfolio with unsystematic risk is

$$w^{\mathsf{mv}} = \phi^{\alpha}\, w^{\alpha} + \phi^{\beta}\, w^{\beta}, \quad \text{where}$$

$\phi^{\alpha} = \frac{\gamma^{\alpha}}{\gamma}$, $\phi^{\beta} = \frac{\gamma^{\beta}}{\gamma}$, w^{α} and w^{β} are orthogonal to each other.

- The beta portfolio has the usual mean-variance form, but where $\beta\lambda$ is only the systematic component of returns:

$$w^{\beta} = \frac{1}{\gamma^{\beta}} V^{-1}(\beta\lambda)$$

- The alpha portfolio also has the usual mean-variance form, but where α is the unsystematic component of returns:

$$w^{\alpha} = \frac{1}{\gamma^{\alpha}}\Sigma^{+}\alpha + \mathcal{O}(1), \quad \text{where: } \Sigma^{+} = \left[\Sigma^{-1} - \Sigma^{-1}\beta(\beta^{\top}\Sigma^{-1}\beta)^{-1}\beta^{\top}\Sigma^{-1}\right].$$

1. Two inefficient funds span the efficient frontier

- Under the APT, the entire set of mean-variance efficient portfolios is generated by two inefficient portfolios:
  - "beta" portfolio: depends on observed factor risk premia;
  - "alpha" portfolio: depends on unsystematic risk.

$$(\mathrm{SR}^{\beta})^2 + (\mathrm{SR}^{\alpha})^2 \to (\mathrm{SR}^{mv})^2.$$

- If one ignores alpha portfolio,s resulting portfolio will not be efficient, regardless of how many risk factors are included in the model.

2. Alpha portfolio weights dominate beta portfolio weights

- As N increases, elements of the alpha portfolio dominate, in terms of magnitude, the corresponding elements of the beta portfolio.

$$\frac{w_i^\beta}{w_i^\alpha} \to 0 \quad \text{if} \quad a \neq 0_N.$$

- This is a consequence of: $w_i^\beta = O\left(\frac{1}{N}\right)$ but $w_i^\alpha = O\left(\frac{1}{\sqrt{N}}\right)$.

3. No need to estimate factor moments

- For beta portfolio: identify conditions under which, as N increases,
  - beta portfolio can be replaced, without any loss of performance,
  - by a benchmark portfolio (e.g., equal- or value-weighted portfolio),
  - which is functionally independent, and hence immune to misspecification, in
    - mean vector (risk premia) of observed factors and
    - covariance matrix of these factors' returns.

$$\boxed{(\mathrm{SR}^{\text{bench}})^2} + (\mathrm{SR}^{\alpha})^2 \to (\mathrm{SR}^{mv})^2 \tag{45}$$

$$\text{instead of:} \qquad (\mathrm{SR}^{\beta})^2 + (\mathrm{SR}^{\alpha})^2 \to (\mathrm{SR}^{mv})^2. \tag{46}$$

Empirical results
Evaluate out-of-sample performance

Data: Four datasets

- Datasets of monthly stock returns: two emprical, two simulated.
- For empirical datasets, follow closely Ao, Li, and Zheng (2019).
  - Monthly returns from 1977 to 2016.
    1. $N = 30$ stocks comprising the Dow Jones that month.
    2. $N = 100$ randomly selected stocks from S&P 500 that month.
- For simulated returns, match the empirical datasets:
  - Monthly returns
    1. $N = 30$ stocks;
    2. $N = 100$ stocks.
- In all cases, returns augmented by Fama-French $K = 3$ factors.

Design of experiment

- Compare out-of-sample performance of various portfolios that have a target volatility of 5% per month, as in Ao, Li, and Zheng (2019).
  - Estimate APT model at date t
    - Employ pseudo-MLE (maximum-likelihood-estimation) with the no-arbitrage constraint.
    - Use cross-validation to identify when no-arbitrage constraint binding.
    - Use a rolling window of 120 months
  - Compute portfolio weight at date t.
  - Evaluate out-of-sample portfolio return at date $t + 1$.
  - And, so on ...

Use five reference portfolios from existing literature

- Compare performance of our strategies, with five reference portfolios:

1. MV: Mean-variance efficient portfolio, using sample return moments.
2. GMV-LW: Global minimum-variance portfolio, with covariance matrix estimated using shrinkage (Ledoit and Wolf 2003) .
3. PCA$_n$, with $n = \{1, 2, \ldots, 10\}$: Treat PCs as observed factors.
4. EW: Equally weighted portfolio (DeMiguel, Garlappi, and Uppal 2009).
5. MAXSER (Ao, Li, and Zheng 2019)

⋆ DeMiguel, Garlappi, and Uppal (2009) and Ao, Li, and Zheng (2019) show, respectively, that the EW and MAXSER portfolios outperform fourteen other strategies proposed in the literature.

Our portfolio strategies: Robust-Mean-Variance (RMV)

1. RMV using V (covariance matrix of N stock returns)
2. RMV using Ω (covariance matrix of K factor returns), because beta portfolio return from investing in N assets is equal to that from investing in only the $K = 3$ Fama-French factors.
3. RMV using V: OptComb combines optimally alpha and beta portfolios, recognizing that with finite N the two are not necessarily orthogonal.
4. RMV using Ω: OptComb combines optimally alpha and beta-equivalent portfolios using only $K = 3$ Fama-French factors.

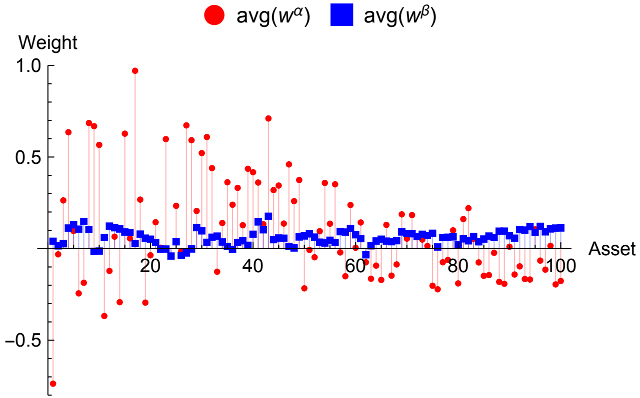
⋆ Because returns may not be normal, compute t-statistic for difference in Sharpe ratios using heteroskedasticity and autocorrelation robust (HAC) kernel estimation approach (Ledoit and Wolf 2008).

- Display results only for S&P stocks ($N = 100$)
- Results for other three datasets are similar

Out-of-sample performance: For S&P stock returns

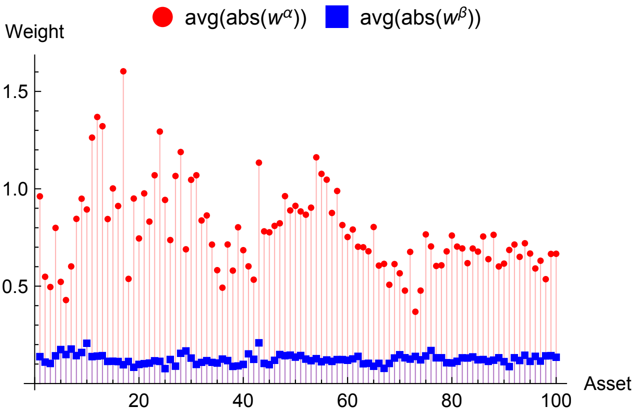
	Mean p.a.	SR p.a.	$\frac{(SR_k - SR_{EW})}{SR_{EW}}$	$\frac{(SR_k - SR_{MAXS})}{SR_{MAXS}}$	t-stat for diff in SR wrt EW	t-stat for diff in SR wrt MAXSER
MV	−0.024	− 0.190	−1.385	−1.283	−2.052	−3.112
GMV-LW	0.019	0.146	−0.704	−0.782	−1.176	−1.816
PCA2	−0.001	−0.001	−1.003	−1.002	−2.206	−2.985
PCA3	0.056	0.334	−0.324	−0.502	−0.712	−1.497
PCA4	0.075	0.460	−0.068	−0.314	−0.150	−0.937
PCA10	0.033	0.203	−0.587	−0.696	−1.292	−2.075
EW	0.070	0.494	0.000	−0.265	—	−0.495
MAXSER	0.094	0.672	0.360	0.000	0.495	—
RMV using V	0.116	0.763	0.546	0.137	0.703	0.406
RMV using Ω	0.137	1.016	1.055	0.511	1.959	1.623
RMV using V: OptComb	0.114	0.642	0.298	−0.046	0.952	0.875
RMV using Ω: OptComb	0.206	1.222	1.472	0.818	2.169	2.203

Properties of alpha and beta portfolio weights



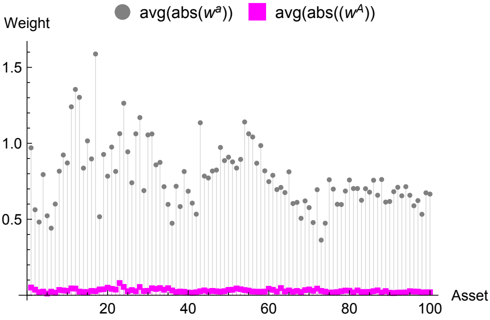
- Weights of beta portfolio are small and mostly positive
- Weights of alpha portfolio are large and have long-short positions

Importance of alpha portfolio relative to beta portfolio



- Weights of alpha portfolio dominate weights of beta portfolio.

Relative size of missing systematic & unsystematic components



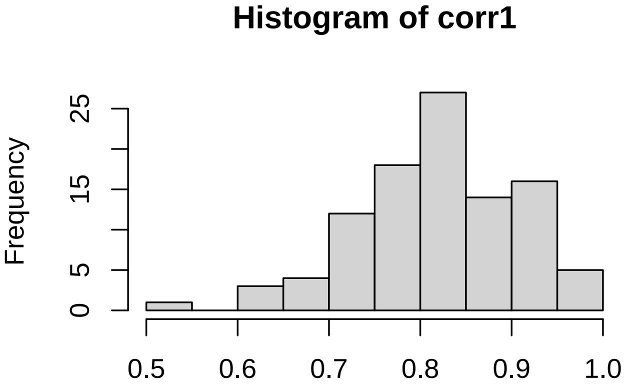
- When beta portfolio is based on Fama-French three-factor model,
- Then, the alpha portfolio weights show that:
  - The effect of omitted systematic risk factors is negligible;
  - The effect of omitted unsystematic risk dominates.

Source of performance of our portfolio (RMV using V)

Portfolio	w^{rmv}	w^{β}	w^{α}	$w^{\beta_{\mathbf{mis}}}$
Mean	0.1160	0.0088	0.1070	0.0001
Standard deviation	0.1520	0.0302	0.1520	0.0136
Sharpe ratio	0.7630	0.2920	0.7020	0.0090
$\frac{(\mathrm{SR}^k)^2}{(\mathrm{SR}^{rmv})^2}$	100%	14.65%	84.65%	0.01%

- When beta portfolio is based on Fama-French three-factor model, then to the squared SR of the mean-variance optimal portfolio:
  - Fama-French factors contribute only: 14.65%
  - Omitted unsystematic risk contributes: 84.65%
  - Omitted systematic risk factors contribute: 0.01%

How persistent is w^α (or α)?



- Average serial correlation is greater than 0.80.

Key takeaway:
Unsystematic risk is priced

Key implication:
Optimal portfolio should exploit unsystematic risk

Conclusion

- Our theoretical and empirical results
  - highlight the importance of priced unsystematic risk.
- What has typically been viewed as a pricing error,
  - should instead be viewed as an integral part of asset-pricing models.
- An optimal portfolio should exploit unsystematic risk.

End of focus

This brings us to an end to Part 3 of this chapter
on how to construct portfolios that exploit
both systematic and unsystematic risk.

- The rest of the chapter contains
  1. first, a summary of all the chapters in the book,
  2. then, some advice about the final information, and,
  3. finally, some concluding remarks.

Road map

Road map

One-page summary of the entire book

Part A: Preliminaries

Chapter 1: Properties of asset returns (when managing financial data)

Chapter 2: Out-of-sample Sharpe ratio (for performance measurement)

Part B: Modern portfolio management

Chapter 3: Markowitz model (which ignores estimation error)

Chapter 4: Shrinkage in the Markowitz model (to adjust for estimation error)

Part C: Post-Modern Portfolio Management

Chapter 5: Black-Litterman model (portfolio choice using the CAPM)

Chapter 6: Parametric portfolio policies (portfolio choice with FF characteristics)

Chapter 7: Volatility-timed factor portfolios (*timing* systematic risk factors)

Chapter 8: Exploiting unsystematic risk (portfolio choice beyond systematic factors)

Road map

Summary of Chapter 1: Managing financial data

- The Tidy Finance website tells us how we can obtain
  1. Stock-price data
  2. Stock-characteristics data
  3. Macroeconomic data.
- From daily stock-price data, we can construct
  - daily returns, monthly, returns, and annual returns;
  - we can then estimate return means, volatilities, and covariances.
- Estimate of expected returns based on sample mean is imprecise.
  - Additional data does not lead to improvement in precision.
- Estimate of return variance based on sample data is precise.
  - Additional data does lead to an improvement in precision.

Summary of Chapter 2: Performance measurement

- Performance of myopic portfolios can be evaluated by
  - estimating their out-of-sample returns
  - based on a rolling or expanding estimation window
  - performance metrics include Sharpe ratio, Sortino ratio, Treynor ratio, alpha, maximum drawdown, VaR, and CVaR.
- When evaluating performance, we
  - test the difference in Sharpe ratios,
  - for which the test-statistic must be computed correctly,
  - given the properties of the distribution of the underlying returns.

Summary of Chapter 3: Markowitz model . . . I

- It is important to look at an asset as part of a portfolio instead of in isolation; i.e., it is important to take into account asset correlations.
- The variance of the return on a portfolio with many assets equals the average covariance of returns in the portfolio.
- The entire efficient frontier can be generated from two efficient portfolios:
  - if a risk-free asset is not available, then we need two portfolios on the efficiency frontier;
  - if a risk-free asset is available, then we need the tangency portfolio and the risk-free asset.

Summary of Chapter 3: Markowitz model . . . II

- In the absence of a risk-free asset, every portfolio on the frontier can be written as

$$w_p = w_0 + (w_1 - w_0)\mu_{\mathrm{targ}} = w_0(1 - \mu_{\mathrm{targ}}) + w_1\mu_{\mathrm{targ}}$$

where w_0 is a portfolio with $\mu_{\mathrm{targ}} = 0$ and w_1 with $\mu_{\mathrm{targ}} = 1$.

- In the presence of a risk-free asset, the optimal portfolio weights are:

$$\begin{aligned} w &= \left[\frac{(\mu_{\mathrm{targ}} - R_f)}{(\mu - R_f 1_N)^\top V^{-1}(\mu - R_f 1_N)}\right] V^{-1}(\mu - R_f 1_N) \\ &= \frac{1}{\gamma} V^{-1}(\mu - R_f 1_N). \end{aligned}$$

Summary of Chapter 3: Markowitz model . . . III

- The marginal contribution of an asset to the portfolio's expected return and to the portfolio's volatility are, respectively

$$\frac{\partial \mathbb{E}[R_p]}{\partial w_n} = \mathbb{E}[R_n] - R_f \quad \text{and} \quad \frac{\partial \sigma_p}{\partial w_n} = \frac{\sigma_{np}}{\sigma_p}.$$

- For any frontier portfolio p, the return-to-risk ratio of all the risky assets in it must be the same:

$$\frac{\mathbb{E}[R_n] - R_f}{(\sigma_{np}/\sigma_p)} = \frac{\mathbb{E}[R_p] - R_f}{(\sigma_{pp}/\sigma_p)} = \frac{\mathbb{E}[R_p] - R_f}{\sigma_p} \qquad \ldots \text{recall that } \sigma_{pp} = \sigma_p^2,$$

which implies that

$$\mathbb{E}[R_n] - R_f = \beta_{np}\big(\mathbb{E}[R_p] - R_f\big),$$

which links optimal portfolio choice to beta-pricing models!

Summary of Chapter 3: Markowitz model . . . IV

- The Markowitz analysis extends in a straightforward manner
  - to mean-variance optimization with respect to a benchmark portfolio,
  - with the Information ratio replacing the Sharpe ratio, and
  - the Appraisal ratio replacing the Treynor ratio.

Summary of Chapter 4: Shrinking Markowitz model . . . I

- The mean-variance portfolio weights

$$w = \frac{1}{\gamma} V^{-1} (\mu - R_f 1_N)$$

perform very poorly out-of-sample because

  - expected returns are estimated imprecisely;
  - the covariance matrix is ill-conditioned.

- Bayesian shrinkage of sample mean returns is not very effective at reducing the effect of estimation error on out-of-sample performance of mean-variance portfolios.

- Shortsale constraints on mean-variance weights reduce turnover but are less effective at improving its out-of-sample Sharpe ratio.

Summary of Chapter 4: Shrinking Markowitz model . . . II

- The global-minimum-variance (GMV) portfolio, which ignores expected returns entirely, achieves a higher Sharpe ratio than mean-variance portfolios;
- Performance of the GMV portfolio can be improved further by
  - imposing short-sale constraints or
  - Ledoit-Wolf shrinkage of the covariance matrix.

Summary of Chapter 5: Black-Litterman model . . . I

- Instead of starting with sample moments and shrinking them,
- Black and Litterman's approach has two steps:
  1. start with the weights of the market portfolio
  2. tilt these weights to reflect the views of the investor.

Summary of Chapter 5: Black-Litterman model . . . II

- The Black-Litterman model has two steps:
  1. Back out expected returns from the CAPM.

$$w_{\text{mkt}} = \frac{1}{\gamma}\mathbb{V}[R_{\text{sample}}]^{-1}(\mathbb{E}[R_{\text{CAPM}}] - R_f 1_N).$$

$$\underbrace{\mathbb{E}[R_{\text{CAPM}}]}_{N\times 1} - R_f 1_N = w_{\text{mkt}}\gamma\,\mathbb{V}[R_{\text{sample}}] \quad \ldots \text{CAPM-implied expected returns}$$

  2. Blend investor's views with CAPM-implied expected returns to get $\mathbb{E}[R_{\text{CAPM}}|\text{views}]$ and $\mathbb{V}[R_{\text{sample}}|\text{views}]$, to compute

$$w = \frac{1}{\gamma}\mathbb{V}[R_{\text{sample}}|\text{views}]^{-1}(\mathbb{E}[R_{\text{CAPM}}|\text{views}] - R_f 1_N).$$

Summary of Chapter 5: Black-Litterman model ...III

1. You can get the CAPM-implied expected excess returns from

$$\mathbb{E}[R_{\text{CAPM}}] - R_f 1_N = \gamma V_{\epsilon_R} w_{\text{mkt}}, \quad \text{where } \gamma = \frac{\mathbb{E}[R_{\text{mkt}} - R_f]}{\sigma^2_{\text{mkt}}}.$$

2. Conditional on an investor's views, the Black-Litterman formula
   2.1 for posterior mean returns is:

$$\begin{aligned}\mu_{\text{BL}} &= \mathbb{E}[R_{\text{CAPM}}|\text{views}] \\ &= \left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\underbrace{\mathbb{E}[R_{\text{CAPM}}]}_{\text{model}} + P^\top\Omega^{-1}\underbrace{q}_{\text{views}}\right].\end{aligned}$$

   2.2 for posterior covariance matrix of returns is:

$$\Sigma_{\text{BL}} = \mathbb{V}[R_{\text{sample}}|\text{views}] = \Sigma + \left[(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P\right]^{-1}.$$

Summary of Chapter 6: Parametric portfolios . . . I

- In the absence of a single-factor model, we needed to estimate:
  - N mean returns
  - N variances
  - $(N^2 - N)/2$ covariances
  - for a total of $\frac{N(N+3)}{2} \approx \frac{N^2}{2}$ parameters.
- In the presence of a single-factor model, we needed to estimate:
  - N alphas
  - N betas
  - N asset-specific volatilities, $\sigma^2_{e_n}$
  - mean and volatility of the market excess return, i.e., λ_m and σ_m
  - for a total of $(3N + 2) \approx N$ parameters.

Summary of Chapter 6: Parametric portfolios . . . II

- But the market model is bad at explaining the cross-section of stock returns.
- Thus, we need a model with $K > 1$ factors, which could be
  1. macroeconomic factors,
  2. fundamental factors (firm characteristics), or
  3. statistical factors (from principal-component analysis).
- For a K-factor model for N assets, we need to estimate:
  - α_n for each of the N stocks . . . N;
  - $\sigma^2_{e_n}$ for each of the N stocks . . . N;
  - $\beta_{n,k}$ for each of the N stocks for each of the K factors; . . . $N \times K$;
  - λ_k and $\sigma^2_{F_k}$ for the K factors. . . . $2K$;
- which is a total of: $2N + 2K + (N \times K)$.

Summary of Chapter 6: Parametric portfolios . . . III

- Instead of a factor model for asset returns,
- Brandt, Santa-Clara, and Valkanov (2009) propose, a K-factor model for portfolio weights:

$$w_t(\theta) = w_{b,t} + (F_{1,t}\theta_1 + F_{2,t}\theta_2 + \ldots + F_{K,t}\theta_K)/N_t.$$

  - Note that θ does not depend on time.
- Then, the investor choses the weights θ by maximizing mean-variance utility:

$$\max_{\theta} \quad \mathbb{E}_t[r_{p,t+1}(\theta)] - \frac{\gamma}{2}\mathbb{V}_t[r_{p,t+1}(\theta)].$$

- If you want, you can impose nonnegativity or other constraints on θ and apply shrinkage to the covariance matrix.

Summary of Chapter 7: Volatility-timed factors . . . I

- Moreira and Muir (2017) show that the mean return of a factor does not change proportionately with its variance.

- Thus, conditioning the weight on return volatility will allow you to earn an "alpha" relative to the return of the unconditional factor.

- The return on the volatility-managed factor is

$$f^{\sigma}_{t+1} = \frac{c}{\sigma^2_t(f)} \times f_{t+1}, \quad \text{where}$$

  - $\sigma_t(f)$ is the previous month's realized volatility, estimated using daily data
  - choose c so f^{σ} has the same unconditional volatility as f.

Summary of Chapter 7: Volatility-timed factors ... II

- The findings of Moreira and Muir (2017) have been criticized:
  - Cederburg, O'Doherty, Wang, and Yan (2020) show gains from volatility timing cannot be realized out of sample.
  - Barroso and Detzel (2021) show that transaction costs erode the gains entirely.
  - Barroso and Detzel (2021) show that gains from volatility timing the market are achieved only during periods of "high sentiment."

Summary of Chapter 7: Volatility-timed factors . . . III

- DeMiguel, Martín-Utrera, and Uppal (2024) address these criticisms by proposing a volatility-timing strategy, with four distinct features:
  1. Multifactor, instead of individual-factor portfolios.
  2. Relative factor weights can vary (as a function of market volatility), instead of having a fixed-weight multifactor portfolio.

  $$\theta_{k,t} = a_k + \frac{b_k}{\sigma_t}.$$

  3. Account for trading diversification (netting of trades across factors) when computing transaction costs;
  4. Optimize factor weights accounting for transaction costs.

Summary of Chapter 8: Harvesting unsystematic risk . . . I

- Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2024) show that:
  - Expected stock returns include compensation for unsystematic risk.
  - A large part of the SDF's variation is explained by unsystematic risk.
  - Contrary to the factor-zoo literature, a single systematic risk factor, the market return, is sufficient to explain most of the variation in the SDF's systematic component.
- Raponi, Uppal, and Zaffaroni (2023) show how the optimal mean-variance efficient portfolio must have two components:
  - beta portfolio that depends only on systematic risk factors;
  - alpha portfolio that depends only on unsystematic risk.
- They also show as N increases, the weight of each asset in the alpha portfolio dominates the weight of that asset in the beta portfolio.

Road map

What we plan to do in the next chapter

What's next?

Today's class is the last class for this course.

I hope you continue to enjoy learning about quantitative portfolio management (and Python) in your working life.

I wish you the very best in your journey of learning.

Road map

Bibliography . . . I

Ahn, S. C., and A. R. Horenstein. 2013. Eigenvalue ratio test for the number of factors. *Econometrica* 81 (3): 1203–1227. (Cited on page 1089).

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies* 32 (7): 2890–2919. (Cited on pages 622, 623, 1118–1120).

Back, K. E. 2017. *Asset pricing and portfolio choice theory.* Oxford University Press. (Cited on page 1035).

Barroso, P., and A. L. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140 (3): 744–767. (Cited on pages 734, 886, 906, 939, 949–953, 963, 966, 1030, 1153).

Bawa, V. S., S. Brown, and R. Klein. 1979. *Estimation risk and optimal portfolio choice.* North Holland, Amsterdam. (Cited on pages 458, 732, 884, 1028).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

———. 1991a. Combining investor views with market equilibrium. *Journal of Fixed Income* 1 (2): 7–18. (Cited on pages 626, 732, 884, 1028).

———. 1991b. Global asset allocation with equities, bonds, and currencies. *Fixed Income Research* 2 (15-28): 1–44. (Cited on pages 626, 732, 884, 1028).

Bibliography . . . II

Black, F., and R. Litterman. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22 (9): 3411–3447. (Cited on pages 733, 839, 841, 844, 847, 885, 1029, 1151).

Bryzgalova, S., J. Huang, and C. Julliard. 2023. Bayesian solutions for the factor zoo: we just ran two quadrillion models. *The Journal of Finance* 78 (1): 487–557. (Cited on page 1078).

Cederburg, S., M. S. O'Doherty, F. Wang, and X. Yan. 2020. On the performance of volatility-managed portfolios. *Journal of Financial Economics* 138 (1): 95–117. (Cited on pages 734, 886, 906, 939, 940, 942, 943, 945–947, 1030, 1153).

Chen, N.-F., R. Roll, and S. Ross. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403. (Cited on pages 733, 773, 885, 1029).

Cochrane, J. 2011. Discount rates. *Journal of Finance* 56:1047–1108. (Cited on page 1077).

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2024. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on pages 734, 829, 886, 1030, 1069, 1155).

Bibliography . . . III

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. (Cited on pages 450, 495, 559, 619, 724, 1120).

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. 2020. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Review of Financial Studies* 33, no. 5 (April): 2180–2222. (Cited on pages 162, 733, 852, 885, 957, 1029).

DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. (Cited on pages 162, 734, 886, 906, 953, 955, 1004, 1030, 1154).

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2 (June): 427–465. (Cited on pages 625, 726, 733, 759, 760, 885, 1029).

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105 (3): 457–472. (Cited on pages 625, 726, 733, 885, 1029).

Bibliography . . . IV

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

———. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

He, G., and R. Litterman. 1999. The intuition behind Black-Litterman model portfolios. *Investment Management Research (Goldman, Sachs & Company).* (Cited on pages 626, 640, 705, 732, 884, 1028).

Jorion, P. 1988. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 732, 884, 1028).

Jorion, P. 1985. International portfolio diversification with estimation risk [in English]. *Journal of Business* 58 (3): pp. 259–278. (Cited on pages 458, 467, 500, 732, 884, 1028).

———. 1992. Portfolio optimization in practice. *Financial Analysts Journal* 48 (1): 68–74. (Cited on pages 732, 884, 1028).

Klein, R. W., and V. S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3 (3): 215–31. (Cited on pages 465, 732, 884, 1028).

Bibliography . . . V

Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10 (5): 603–621. (Cited on pages 519, 538, 543, 1120).

———. 2008. Robust performance hypothesis resting with the Sharpe ratio. *Journal of Empirical Finance* 15:850–859. (Cited on pages 225, 226, 232, 1121).

Litterman, R. 2003. *Modern investment management: An equilibrium approach.* New York: Wiley. (Cited on pages 626, 732, 884, 1028).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information [in English]. *Journal of Finance* 42 (3): 483–510. (Cited on pages 762, 1081, 1082).

Moreira, A., and T. Muir. 2017. Volatility-managed portfolios. *Journal of Finance* 72 (4): 1611–1644. (Cited on pages 734, 886, 905–907, 909, 911, 912, 914, 918, 921, 923, 927, 936–939, 942, 945, 947, 955, 966, 1008, 1030, 1152, 1153).

———. 2019. Should long-term investors time volatility? *Journal of Financial Economics* 131 (3): 507–527. (Cited on pages 734, 886, 1030).

Bibliography . . . VI

Onatski, A. 2012. Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics* 168 (2): 244–258. (Cited on page 1089).

Pástor, Ľ., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56 (3): 335–381. (Cited on pages 468, 732, 884, 1028).

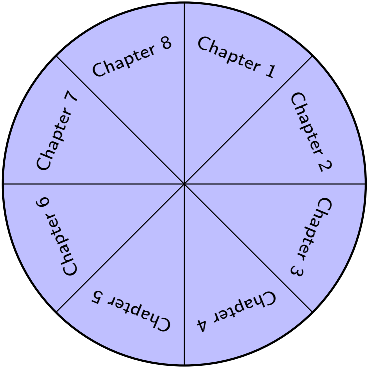
Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary. (Cited on pages 622, 623, 734, 886, 1030, 1111, 1155).

Ross, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13:341–360. (Cited on pages 733, 736, 885, 1029, 1082).

———. 1977. Return, risk, and arbitrage. In *Risk and return in finance,* edited by I. Friend and J. Bicksler. Cambridge, MA: Ballinger. (Cited on pages 733, 736, 885, 1029, 1082).

Rubinstein, M. 2006. *A history of the theory of investments: My annotated bibliography.* Hoboken, NJ: John Wiley & Sons. (Cited on pages 734, 887, 1031).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).



End of Chapter 8

Blank page 1

Blank page 2

Blank page 3

Bibliography

for the entire course

Bibliography . . . I

Abdi, F., and A. Ranaldo. 2017. A simple estimation of bid-ask spreads from daily close, high, and low prices. *Review of Financial Studies* 30 (12): 4437–4480. (Cited on page 973).

Ahn, S. C., and A. R. Horenstein. 2013. Eigenvalue ratio test for the number of factors. *Econometrica* 81 (3): 1203–1227. (Cited on page 1089).

Amenc, N., M. Esakia, F. Goltz, and B. Luyten. 2019. Macroeconomic risks in equity factor investing. *Journal of Portfolio Management* 45 (6): 39–60. (Cited on page 961).

Andersen, T. G., and L. Benzoni. 2010. Stochastic volatility. *CREATES Research Paper,* nos. 2010-10. (Cited on page 506).

Ang, A. 2014. *Asset management: A systematic approach to factor investing.* Oxford University Press. (Cited on page 860).

Ao, M., Y. Li, and X. Zheng. 2019. Approaching mean-variance efficiency for large portfolios. *Review of Financial Studies* 32 (7): 2890–2919. (Cited on pages 622, 623, 1118–1120).

Back, K. E. 2017. *Asset pricing and portfolio choice theory.* Oxford University Press. (Cited on page 1035).

Bailey, D. H., and M. Lopez de Prado. 2012. The Sharpe ratio efficient frontier. *Journal of Risk* 15 (2): 13. (Cited on page 187).

Bibliography . . . II

Balduzzi, P., and A. W. Lynch. 1999. Transaction costs and predictability: Some utility cost calculations. *Journal of Financial Economics* 52 (1): 47–78. (Cited on page 160).

Bali, T. G., H. Beckmeyer, and A. Goyal. 2024. A joint factor model for bonds, stocks, and options. Available at SSRN 4589282. (Cited on page 857).

Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert. 2023. Option return predictability with machine learning and big data. *Review of Financial Studies* 36 (9): 3548–3602. (Cited on page 84).

Barroso, P., and A. L. Detzel. 2021. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *Journal of Financial Economics* 140 (3): 744–767. (Cited on pages 734, 886, 906, 939, 949–953, 963, 966, 1030, 1153).

Barry, C. B. 1974. Portfolio analysis under uncertain means, variances, and covariances [in English]. *Journal of Finance* 29 (2): 515–522. (Cited on pages 458, 465).

Bass, R., S. Gladstone, and A. Ang. 2017. Total portfolio factor, not just asset, allocation. *Journal of Portfolio Management* 43 (5): 38–53. (Cited on page 961).

Bawa, V. S., S. Brown, and R. Klein. 1979. *Estimation risk and optimal portfolio choice.* North Holland, Amsterdam. (Cited on pages 458, 732, 884, 1028).

Bender, J., J. Sun, and R. Thomas. 2018. Asset allocation vs. factor allocation—Can we build a unified method? *Journal of Portfolio Management* 45 (2): 9–22. (Cited on page 961).

Bibliography . . . III

Best, M. J., and R. R. Grauer. 1992. Positively weighted minimum-variance portfolios and the structure of asset expected returns. *Journal of Financial and Quantitative Analysis* 27 (4): 513–537. (Cited on page 458).

Bhardwaj, R. K., and L. D. Brooks. 1992. The january anomaly: effects of low share price, transaction costs, and bid-ask bias. *Journal of Finance* 47 (2): 553–575. (Cited on page 160).

Black, F. 1993. Beta and return. *Journal of Portfolio Management* 20 (1): 8–18. (Cited on pages 755–757).

Black, F., and R. Litterman. 1990. Asset allocation: Combining investor views with market equilibrium. Goldman, Sachs & Co. (Cited on pages 458, 626, 642, 732, 884, 1028).

———. 1991a. Combining investor views with market equilibrium. *Journal of Fixed Income* 1 (2): 7–18. (Cited on pages 626, 732, 884, 1028).

———. 1991b. Global asset allocation with equities, bonds, and currencies. *Fixed Income Research* 2 (15-28): 1–44. (Cited on pages 626, 732, 884, 1028).

———. 1992. Global portfolio optimization. *Financial Analysts Journal* 48:28–43. (Cited on pages 458, 626, 642, 732, 884, 1028).

Blin, O., F. Ielpo, J. Lee, and J. Teiletche. 2018. Factor timing revisited: alternative risk premia allocation based on nowcasting and valuation signals. SSRN Working Paper. (Cited on page 961).

Bibliography . . . IV

Boyle, P., L. Garlappi, R. Uppal, and T. Wang. 2012. Keynes meets Markowitz: The trade-off between familiarity and diversification. *Management Science* 58 (2): 253–272. (Cited on page 459).

Brandt, M. W., P. Sant-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross section of equity returns. *Review of Financial Studies:* Forthcoming. (Cited on pages 626, 726, 739, 854).

Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22 (9): 3411–3447. (Cited on pages 733, 839, 841, 844, 847, 885, 1029, 1151).

Brown, S. 1979. The effect of estimation risk on capital market equilibrium. *Journal of Financial and Quantitative Analysis* 14 (2): 215–220. (Cited on page 465).

Brugière, P. 2020. Quantitative portfolio management. *Springer Texts in Business and Economics.* (Cited on page 71).

Bryzgalova, S., J. Huang, and C. Julliard. 2023. Bayesian solutions for the factor zoo: we just ran two quadrillion models. *The Journal of Finance* 78 (1): 487–557. (Cited on page 1078).

Campbell, J. Y., and L. M. Viceira. 2002. *Strategic asset allocation: Portfolio choice for long-term investors.* New York: Oxford University Press. (Cited on page 115).

Bibliography . . . V

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52, no. 1 (March): 57–82. (Cited on pages 131, 469, 966).

Cazalet, Z., and T. Roncalli. 2014. Facts and fantasies about factor investing. Available at SSRN 2524547. (Cited on page 776).

Cederburg, S., M. S. O'Doherty, F. Wang, and X. Yan. 2020. On the performance of volatility-managed portfolios. *Journal of Financial Economics* 138 (1): 95–117. (Cited on pages 734, 886, 906, 939, 940, 942, 943, 945–947, 1030, 1153).

Chamberlain, G. 1983. Funds, factors and diversification in arbitrage pricing models. *Econometrica* 51:1305–1324. (Cited on page 736).

Chamberlain, G., and M. Rothschild. 1983. Arbitrage, factor structure and mean-variance analysis on large asset markets. *Econometrica* 51:1281–1304. (Cited on page 736).

Chan, L. K. C., J. Karceski, and J. Lakonishok. 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies* 12 (5): 937–74. (Cited on pages 458, 500).

Chen, A. Y., and T. Zimmermann. 2022. Open source cross-sectional asset pricing. *Critical Finance Review* 11 (2): 207–264. (Cited on pages 82, 776).

Chen, N.-F., R. Roll, and S. Ross. 1986. Economic forces and the stock market. *Journal of Business* 59:383–403. (Cited on pages 733, 773, 885, 1029).

Bibliography . . . VI

Chen, Y., A. Wiesel, Y. C. Eldar, and A. O. Hero. 2010. Shrinkage algorithms for mmse covariance estimation. *IEEE Transactions on Signal Processing* 58 (10): 5016–5029. (Cited on page 542).

Chopra, V. K. 1993. Improving optimization. *Journal of Investing* 8 (Fall): 51–59. (Cited on page 459).

Clarke, R. G., H. de Silva, and S. Thorley. 2010. Minimum variance portfolio composition. *SSRN, www.* (Cited on page 509).

Clarke, R., H. de Silva, and S. Thorley. 2016. Fundamentals of efficient factor investing. *Financial Analysts Journal* 72:9–26. (Cited on page 776).

Cochrane, J. 2011. Discount rates. *Journal of Finance* 56:1047–1108. (Cited on page 1077).

———. 2017. Macro-finance. *Review of Finance* 21:1–42. (Cited on page 773).

Connor, G., L. Goldberg, and R. A. Korajczyk. 2010. *Portfolio risk analysis.* Princeton, New Jersey: Princeton University Press. (Cited on page 133).

Dangl, T., and M. Kashofer. 2013. Minimum-variance stock picking – a shift in preferences for minimum-variance constituents. Working Paper, Vienna University of Technology. (Cited on page 510).

Daniel, K., L. Mota, S. Rottke, and T. Santos. 2020. The cross-section of risk and returns. *Review of Financial Studies* 33 (5): 1927–1979. (Cited on page 830).

Bibliography . . . VII

De Franco, C., M. Guidolin, and B. Monnier. 2017. The robustness of the volatility factor: linear versus nonlinear factor model. *Journal of Index Investing* 8 (3): 75–88. (Cited on page 962).

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze. 2024. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on pages 734, 829, 886, 1030, 1069, 1155).

———. 2025. Cross-sectional asset pricing with unsystematic risk. Available at SSRN 4135146. (Cited on page 134).

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal. 2009. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science* 55 (5): 798–812. (Cited on pages 501, 519, 520, 550).

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies* 22 (5): 1915–1953. (Cited on pages 450, 495, 559, 619, 724, 1120).

DeMiguel, V., A. Martin-Utrera, F. J. Nogales, and R. Uppal. 2017. A portfolio perspective on the multitude of firm characteristics. Working Paper, London Business School. (Cited on page 835).

Bibliography . . . VIII

DeMiguel, V., A. Martín-Utrera, F. J. Nogales, and R. Uppal. 2020. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Review of Financial Studies* 33, no. 5 (April): 2180–2222. (Cited on pages 162, 733, 852, 885, 957, 1029).

DeMiguel, V., A. Martín-Utrera, and R. Uppal. 2024. A multifactor perspective on volatility-managed portfolios. *Journal of Finance* 79 (6): 3859–3891. (Cited on pages 162, 734, 886, 906, 953, 955, 1004, 1030, 1154).

DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov. 2013. Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis* 48 (6): 1813–1845. (Cited on page 506).

DeMiguel, V., and R. Uppal. 2005. Portfolio investment with the exact tax basis via nonlinear programming. *Management Science* 51, no. 2 (February): 277–290. (Cited on pages 163, 393).

Dick-Nielsen, J., P. Feldhütter, L. H. Pedersen, and C. Stolborg. 2023. Corporate bond factors: replication failures and a new framework. Available at SSRN 4586652. (Cited on page 84).

Dick-Nielsen, J., and T. K. Poulsen. 2019. How to clean academic TRACE data. *Available at SSRN 3456082.* (Cited on page 84).

Dickerson, A., P. Mueller, and C. Robotti. 2023. Priced risk in corporate bonds. Forthcoming in Journal of Financial Economics. (Cited on page 84).

Bibliography . . . IX

Driessen, J., P. Maenhout, and G. Vilkov. 2009. The price of correlation risk: Evidence from equity options. *Journal of Finance* 64 (3): 1375–1404. (Cited on page 506).

Dybvig, P., and H. K. Koo. 1996. Investment with taxes. Working Paper, Washington University. (Cited on page 163).

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani. 2004. Least angle regression. *Annals of Statisitcs* 32 (2): 407–499. (Cited on page 862).

Efron, B., and R. J. Tibshirani. 1994. *An introduction to the bootstrap.* CRC press. (Cited on page 219).

Ehsani, S., and J. T. Linnainmaa. 2019. Factor momentum and the momentum factor. NBER Working Paper. (Cited on page 960).

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47, no. 2 (June): 427–465. (Cited on pages 625, 726, 733, 759, 760, 885, 1029).

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56. (Cited on pages 131, 469, 625, 726, 733, 775, 885, 1029).

———. 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105 (3): 457–472. (Cited on pages 625, 726, 733, 885, 1029).

Bibliography . . . X

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. (Cited on pages 132, 625, 726, 733, 775, 885, 966, 1029).

———. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. (Cited on pages 132, 625, 726, 733, 775, 885, 1029).

Fan, J., J. Zhang, and K. Yu. 2012. Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association* 107:592–606. (Cited on page 835).

Feng, G., S. Giglio, and D. Xiu. 2020. Taming the factor zoo: a test of new factors. *Journal of Finance* 75 (3): 1327–1370. (Cited on page 830).

Frazzini, A., and L. H. Pedersen. 2014. Betting against beta. *Journal of Financial Economics* 111 (1): 1–25. (Cited on page 966).

French, K. R. 2008. Presidential address: The cost of active investing. *Journal of Finance* 63, no. 4 (August): 1537–1573. (Cited on page 160).

Frost, P. A., and J. E. Savarino. 1988. For better performance constrain portfolio weights. *Journal of Portfolio Management* 15 (1): 29–34. (Cited on page 459).

Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies* 20:41–81. (Cited on pages 459, 835).

Bibliography . . . XI

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131 (1-2): 59–95. (Cited on page 506).

Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. A test of the efficiency of a given portfolio. *Econometrica:* 1121–1152. (Cited on page 944).

Giglio, S., B. Kelly, and D. Xiu. 2022. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics* 14:337–368. (Cited on page 776).

Goldfarb, D., and G. Iyengar. 2003. Robust portfolio selection problems. *Mathematics of Operations Research* 28 (1): 1–38. (Cited on pages 459, 835).

Gómez-Cram, R. 2021. Late to recessions: stocks and the business cycle. *Journal of Finance* 77 (2): 923–966. (Cited on page 960).

Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep learning.* MIT Press. (Cited on page 835).

Green, R., and B. Hollifield. 1992. When will mean-variance efficient portfolios be well diversified? *Journal of Finance* 47:1785–1809. (Cited on page 500).

Gu, S., B. Kelly, and D. Xiu. 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33 (5): 2223–2273. (Cited on page 84).

Gupta, T., and B. Kelly. 2019. Factor momentum everywhere. *Journal of Portfolio Management* 45 (3): 13–36. (Cited on pages 960, 963).

Bibliography . . . XII

Haddad, V., S. Kozak, and S. Santosh. 2020. Factor timing. *Review of Financial Studies* 33 (5): 1980–2018. (Cited on page 962).

Harvey, C. R., J. Liechty, M. Liechty, and P. Müller. 2003. Portfolio selection with higher moments. Working Paper, Duke University. (Cited on page 459).

He, G., and R. Litterman. 1999. The intuition behind Black-Litterman model portfolios. *Investment Management Research (Goldman, Sachs & Company).* (Cited on pages 626, 640, 705, 732, 884, 1028).

Heaton, J., N. Polson, and J. Witte. 2016. Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry* 33:3–12. (Cited on page 835).

Hodges, P., K. Hogan, J. R. Peterson, and A. Ang. 2017. Factor timing with cross-sectional and time-series predictors. *Journal of Portfolio Management* 44 (1): 30–43. (Cited on page 961).

Hou, K., H. Mo, C. Xue, and L. Zhang. 2019. Which factors? *Review of Finance* 23 (1): 1–35. (Cited on page 132).

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28 (3): 650–705. (Cited on pages 81, 132, 775, 945, 966).

———. 2017a. A comparison of new factor models. *Fisher College of Business Working Paper,* nos. 2015-03: 05. (Cited on page 132).

Bibliography . . . XIII

Hou, K., C. Xue, and L. Zhang. 2017b. Replicating anomalies. *The Review of Financial Studies.* (Cited on pages 132, 775).

Huang, C.-F., and R. Litzenberger. 1988. *Foundations for financial economics.* New Jersey: Prentice Hall. (Cited on pages 302, 399).

Huang, S., Y. Song, and H. Xiang. 2020. The smart beta mirage. *Journal of Financial and Quantitative Analysis:* 1–69. (Cited on page 894).

Huberman, G. 1982. A simple approach to the Arbitrage Pricing Theory. *Journal of Economic Theory* 28:183–191. (Cited on page 736).

Idzorek, T. 2007. A step-by-step guide to the Black-Litterman model: incorporating user-specified confidence levels. In *Forecasting expected returns in the financial markets,* 17–38. Elsevier. (Cited on pages 645, 688).

Ingersoll, J. E., Jr. 1984. Some results in the theory of arbitrage pricing. *Journal of Finance* 39:1021–1039. (Cited on page 736).

Jacobs, H., S. Müller, and M. Weber. 2014. How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets* 19:62–85. (Cited on pages 495, 620, 621, 623, 724).

Jagannathan, R., and T. Ma. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance* 58 (4): 1651–1684. (Cited on pages 459, 501, 519, 522).

Bibliography . . . XIV

James, W., and C. Stein. 1961. Estimation with quadratic loss. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability:* 361–379. (Cited on page 466).

Jobson, J. D., and B. M. Korkie. 1981. Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance:* 889–908. (Cited on pages 187, 214, 447).

Jobson, J. D., and B. Korkie. 1980. Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association* 75:544–554. (Cited on page 458).

Jobson, J., B. Korkie, and V. Ratti. 1979. Improved estimation for Markowitz portfolios using James-Stein type estimators. *Proceedings of the American Statistical Association:* 279–284. (Cited on page 458).

Jorion, P. 1988. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 732, 884, 1028).

Jorion, P. 1985. International portfolio diversification with estimation risk [in English]. *Journal of Business* 58 (3): pp. 259–278. (Cited on pages 458, 467, 500, 732, 884, 1028).

———. 1986. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* 21 (3): 279–292. (Cited on pages 458, 467, 500).

———. 1992. Portfolio optimization in practice. *Financial Analysts Journal* 48 (1): 68–74. (Cited on pages 732, 884, 1028).

Bibliography . . . XV

Kan, R., and G. Zhou. 2005. Optimal estimation for economic gains: Portfolio choice with parameter uncertainty. Working paper, University of Toronto. (Cited on page 459).

Kelly, B. T., S. Pruitt, and Y. Su. 2020. Instrumented principal component analysis. *Available at SSRN 2983919.* (Cited on page 819).

Klein, R. W., and V. S. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3 (3): 215–31. (Cited on pages 465, 732, 884, 1028).

Kozak, S., S. Nagel, and S. Santosh. 2018. Interpreting factor models. *Journal of Finance* 73 (3): 1183–1223. (Cited on pages 133, 733).

———. 2020. Shrinking the cross-section. *Journal of Financial Economics* 135 (2): 271–292. (Cited on pages 133, 733).

Lamont, O. 2001. Economic tracking portfolios. *Journal of Econometrics* 105:161–184. (Cited on page 834).

Ledoit, O., and M. Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10 (5): 603–621. (Cited on pages 519, 538, 543, 1120).

———. 2004a. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88 (January): 365–411. (Cited on pages 458, 501, 519, 538, 539, 542, 543).

Bibliography . . . XVI

Ledoit, O., and M. Wolf. 2004b. Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management* 30, no. 4 (December): 110–119. (Cited on pages 458, 501, 519, 538, 542, 543).

———. 2008. Robust performance hypothesis resting with the Sharpe ratio. *Journal of Empirical Finance* 15:850–859. (Cited on pages 225, 226, 232, 1121).

———. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *Review of Financial Studies* 30 (12): 4349–4388. (Cited on pages 501, 519, 543).

———. 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Working Paper. (Cited on page 543).

———. 2022. The power of (non-) linear shrinking: a review and guide to covariance matrix estimation. *Journal of Financial Econometrics* 20 (1): 187–218. (Cited on pages 501, 519, 543).

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka. 1999. A new estimate of transaction costs. *Review of Financial Studies* 12 (5): 1113–1141. (Cited on page 160).

Lettau, M., and M. Pelger. 2018. *Estimating latent asset-pricing factors.* Technical report. National Bureau of Economic Research. (Cited on page 733).

———. 2020. Factors that fit the time-series and cross-section of stock returns. *Review of Financial Studies* 33:2274–2325. (Cited on page 733).

Bibliography . . . XVII

Litterman, R. 2003. *Modern investment management: An equilibrium approach.* New York: Wiley. (Cited on pages 626, 732, 884, 1028).

Lo, A. W. 2002. The statistics of Sharpe ratios. *Financial Analysts Journal* 58 (4): 36–52. (Cited on pages 143, 187).

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91. (Cited on pages 302, 399, 457, 732, 884, 1028).

———. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley. (Cited on pages 302, 399, 457, 732, 884, 1028).

McLean, R. D., and J. Pontiff. 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71:5–31. (Cited on page 945).

Memmel, C. 2003. Performance hypothesis testing with the Sharpe ratio. Available at SSRN 412588. (Cited on pages 187, 214).

Merton, R. C. 1971. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory* 3:373–413. (Cited on page 115).

———. 1972. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7 (4): 1851–72. (Cited on pages 302, 399).

———. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8 (4): 323–361. (Cited on page 417).

Bibliography . . . XVIII

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information [in English]. *Journal of Finance* 42 (3): 483–510. (Cited on pages 762, 1081, 1082).

———. 1990. *Continuous time finance.* Cambridge MA: Basil Blackwell. (Cited on page 115).

Meucci, A. 2009. Managing diversification. *Risk:* 74–79. (Cited on page 815).

———. 2010. The Black-Litterman approach: Original model and extensions. Available at SSRN. (Cited on pages 676, 705).

Michaud, R. O. 1998. *Efficient asset management: A practical guide to stock portfolio management and asset allocation.* Survey and Synthesis Series, HBS Press, Boston: Financial Management Association. (Cited on page 459).

Miller, K. L., H. Li, T. G. Zhou, and D. Giamouridis. 2015. A risk-oriented model for factor timing decisions. *Journal of Portfolio Management* 41 (3): 46–58. (Cited on page 961).

Moreira, A. 2021. *Quantitative investing.* Available online from this link. (Cited on pages 248, 399, 559).

Moreira, A., and T. Muir. 2017. Volatility-managed portfolios. *Journal of Finance* 72 (4): 1611–1644. (Cited on pages 734, 886, 905–907, 909, 911, 912, 914, 918, 921, 923, 927, 936–939, 942, 945, 947, 955, 966, 1008, 1030, 1152, 1153).

Bibliography . . . XIX

Moreira, A., and T. Muir. 2019. Should long-term investors time volatility? *Journal of Financial Economics* 131 (3): 507–527. (Cited on pages 734, 886, 1030).

Onatski, A. 2012. Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics* 168 (2): 244–258. (Cited on page 1089).

Opdyke, J. D. 2007. Comparing Sharpe ratios: So where are the p-values? *Journal of Asset Management* 8 (5): 308–336. (Cited on page 187).

Partovi, M. H., and M. Caputo. 2004. Principal portfolios: Recasting the efficient frontier. *Economics Bulletin* 7 (3): 1–10. (Cited on page 815).

Pasini, G. 2017. Principal component analysis for stock portfolio management. *International Journal of Pure and Applied Mathematics* 115 (1): 153–167. (Cited on page 815).

Pastor, L. 2000. Portfolio selection and asset pricing models. *The Journal of Finance* 55:179–223. (Cited on page 468).

Pástor, Ľ., and R. F. Stambaugh. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* 56 (3): 335–381. (Cited on pages 468, 732, 884, 1028).

Plyakha, Y., R. Uppal, and G. Vilkov. 2021. Equal or value weighting? Implications for asset-pricing tests. In *Risk management and modeling,* edited by R. Benkraiem, I. Kalaitzoglou, and C. Zopounidis. Springer. (Cited on pages 510, 779).

Bibliography . . . XX

Politis, D. N., and J. P. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89 (428): 1303–1313. (Cited on page 225).

Raponi, V., R. Uppal, and P. Zaffaroni. 2023. Robust portfolio choice. Working Paper, SSRN eLibrary. (Cited on pages 622, 623, 734, 886, 1030, 1111, 1155).

Roll, R. 1977. A critique of the asset pricing theory's tests Part i: On past and potential testability of the theory. *Journal of Financial Economics* 4, no. 2 (March): 129–176. (Cited on page 399).

———. 1980. Orthogonal portfolios. *Journal of Financial and Quantitative Analysis* 15 (5): 1005–1023. (Cited on page 399).

———. 1992. A mean/variance analysis of tracking error. *Journal of Portfolio Management* 18 (4): 13–22. (Cited on page 381).

Ross, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13:341–360. (Cited on pages 733, 736, 885, 1029, 1082).

———. 1977. Return, risk, and arbitrage. In *Risk and return in finance,* edited by I. Friend and J. Bicksler. Cambridge, MA: Ballinger. (Cited on pages 733, 736, 885, 1029, 1082).

Rubinstein, M. 2006. *A history of the theory of investments: My annotated bibliography.* Hoboken, NJ: John Wiley & Sons. (Cited on pages 734, 887, 1031).

Bibliography ... XXI

Scherer, B. 2002. Portfolio resampling: Review and critique. *Financial Analysts Journal* 58 (6): 98–109. (Cited on page 459).

———. 2011. A note on the returns from minimum variance investing. *Journal of Empirical Finance* 18 (4): 652–660. (Cited on page 509).

Scheuch, C., S. Voigt, P. Weiss, and C. Frey. 2024. *Tidy finance with Python.* Chapman-Hall/CRC. Available online from this link. (Cited on pages 100, 399, 559).

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 1919 (3): 425–442. (Cited on pages 131, 625, 725, 732, 884, 1028, 1075).

Stein, C. 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proceedings of the Third Berkeley Symposium.* (Cited on page 466).

Stoll, H. R., and R. E. Whaley. 1983. Transaction costs and the small firm effect. *Journal of Financial Economics* 12 (1): 57–79. (Cited on page 160).

Uppal, R., and T. Wang. 2003. Model misspecification and underdiversification. *Journal of Finance* 58 (6): 2465–2486. (Cited on page 459).

Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies* 21 (4): 1455–1508. (Cited on page 89).

End of bibliography

End of textbook for

Quantitative Portfolio Management