Financial Econometrics

Slides-02

Linear Regression

Review and Applications in Finance

R. Ouysse
Economics[1]

UNSW SYDNEY Australia's Global University

UNSW AUSTRALIA AGSM Business School

Linear Regression

- A model where one variable Y_t is linearly explained by a group of variables $(X_{1t}, \cdots, X_{kt}), t = 1, \cdots, T$
  - Easy to Implement
  - Versatile for financial data analysis
  - Foundation for more advanced models

- General formulation
  - $Y_t = \beta_1 + \beta_2 X_{t1} + \beta_3 X_{t2} + \cdots + \beta_K X_{tK} + \mu_t, t = 1 \cdots, T$
  - Y_t: dependent variable
  - $X_{t1}, \cdots X_{tK}$: explanatory variables, regressors
  - $\beta_1, \beta_2, \cdots, \beta_K$: parameters (to be estimated)
  - μ_t: error term
  - T: number of observations

Application 1: Capital Asset Pricing Model *aka CAPM*

One of the most important problems of modern financial economics is the **quantification of the tradeoff between risk and expected return.** Common sensesuggests risky investments (stock market) will generally yield higher returns than investments free of risk!

- Markowitz (1995) casts the investor's portfolio selection problem in terms of expected return and variance of the return.

$\rightarrow$ Investors optimally hold a mean-variance efficient portfolio: a portfolio with the highest expected return for a given level of variance.

$\Longrightarrow$ **The Efficient Frontier** & **Capital Market Line**

- Capital Asset Pricing Model is concerned with the pricing of assets in equilibrium. In equilibrium, all assets must be held by someone.

$\longrightarrow$ How investors determine the expected returns—and thereby asset prices—as a function of risk.

$\Longrightarrow$ **The Security Market Line**

CAPM

- Given that: some risk can be diversified, diversification is easy and costless, and rational investors diversify
- There should be no premium associated with diversifiable risk.
- The question becomes: What is the equilibrium relation between systematic risk and expected return in the capital markets?
- The CAPM is the best-known and most-widely used equilibrium model of the risk/return (systematic risk/return) relation.

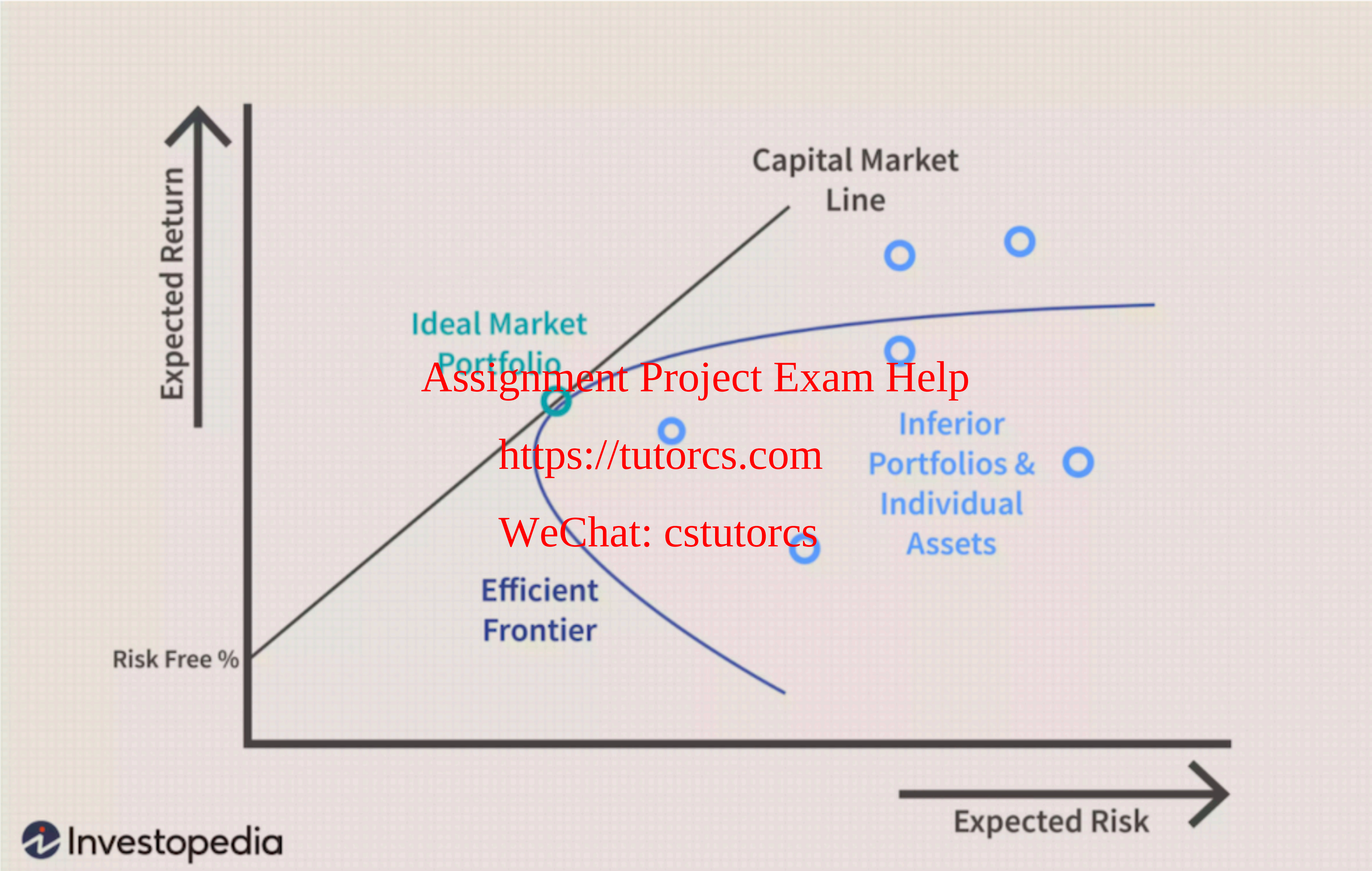


Image by Julie Bang © Investopedia 2019

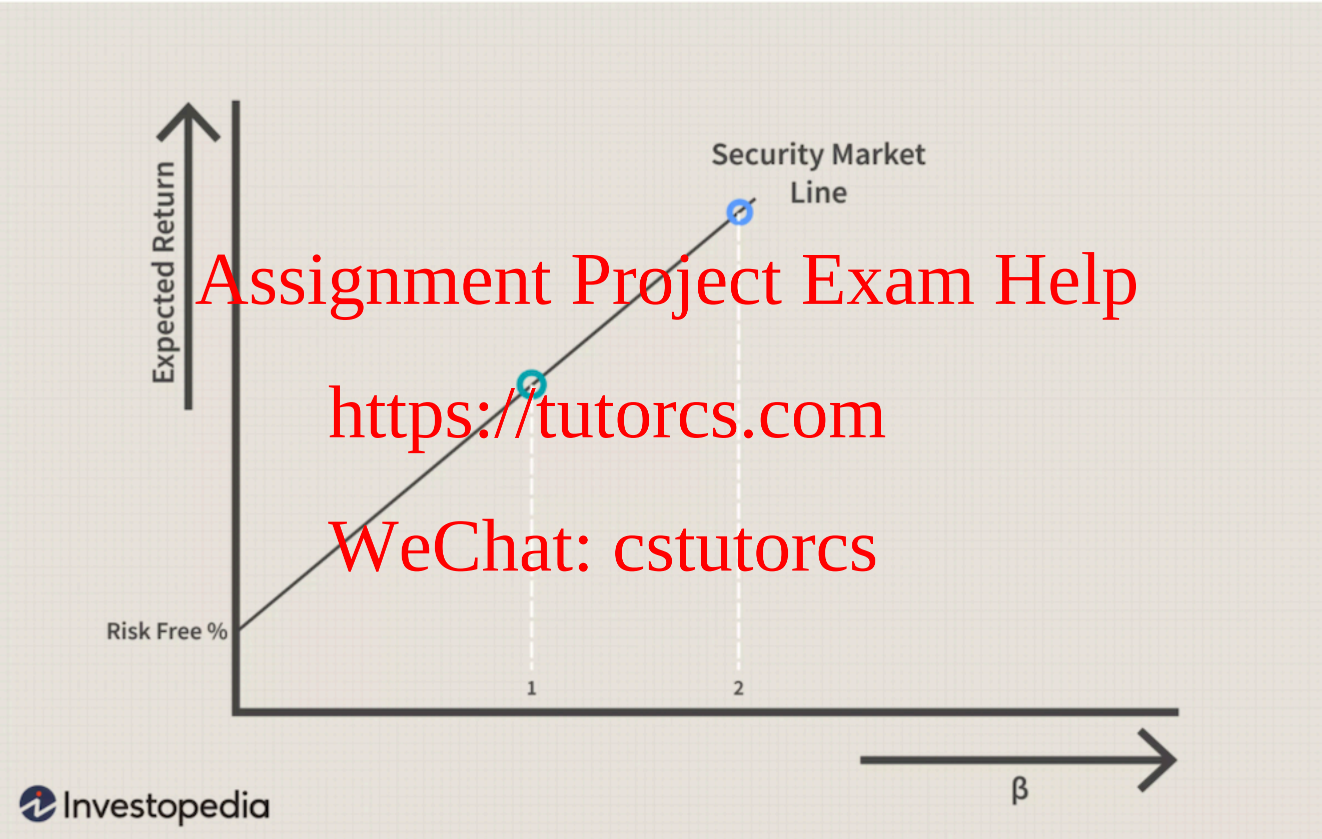


Image by Julie Bang © Investopedia 2019

CAPM Intuition

What would be a "fair" expected return on any stock?

- $E(R_{it}) = R_{ft}$ (risk free) + Risk Premium
- Risk free assets earn the risk-free rate (think of this as a rental rate on capital). The risk free compensate for time.
- If the asset is risky, we need to add a risk premium.
- The size of the risk premium depends on the amount of systematic risk for the asset (stock, bond, or investment project) and the price per unit risk.
- $R_{it} - R_{ft}$ Excess return

CAPM Intuition Formalized

$$E[R_{it}] = R_{ft} + \frac{Cov(R_{it}, R_{mt})}{Var(R_{mt})}[E[R_{mt}] - R_{ft}]$$

$$E[R_{it}] = R_{ft} + \beta_i [E[R_{mt}] - R_{ft}]$$

The expression above is referred to as the "Security Market Line".

- $E[R_{mt}] - R_{ft}$ Market Risk premiun (compensation for risk) or the price per unit of risk
- β_i: number of units of systematic risk
  - $\beta_i > 1$ (or < 1) : the asset is more (less) risky than the market portfolio
  - $\beta_i < 0$: the asset is a hedge against the market portfolio
  - β_i how sensitive the asset to **market movement**

CAPM Formalized

Three inputs are required:

1. An estimate of the risk free interest rate. The current yield on short term treasury bills is one proxy. Practitioners tend to favor the current yield on longer-term treasury bonds but this may be a fix for a problem we don't fully understand.

2. An estimate of the market risk premium, $E[R_{mt}] - R_{ft}$. Expectations are not observable. Generally use a historically estimated value.

   The market is defined as a portfolio of all wealth including real estate, human capital, etc. In practice, a broad based stock index, such as the S&P 500 or the portfolio of all NYSE stocks, is generally used.

3. An estimate of beta.

CAPM: Econometric model

Let $X_{mt} = R_{mt} - R_{ft}$ and $X_{it} = R_{it} - R_{ft}$ and consider the econometric model:

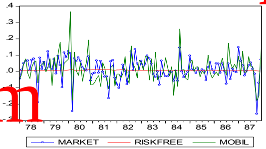
$$X_{it} = \alpha_i + \beta_i X_{mt} + \mu_{it}$$

- The CAPM can be examined by testing $H_0 : \alpha_i = 0$
- If $\alpha_i > 0$, asset i beats the market by earning more than $\beta_i E[X_{mt}]$
- This has been used to test the performance of mutual funds (application in the Brooks textbook)

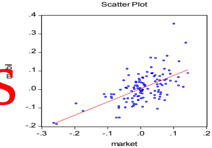
CAPM: Application

What determines the expected return of an asset?

Example: Mobil (a US petroleum firm), 1978:01-1987:12 with $T = 120$



Dependent Variable: E_MOBIL				
Method: Least Squares				
Sample: 1978M01 1987M12				
Included observations: 120				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.004241	0.005881	0.721087	0.4723
E_MARKET	0.714695	0.085615	8.347761	0.0000
R-squared	0.371287	Mean dependent var		0.0[illegible]9[illegible]53
Adjusted R-squared	0.365959	S.D. dependent var		0.0[illegible]0[illegible]68
S.E. of regression	0.0640[illegible]4	Akaike info criterion		-2.[illegible]41[illegible]19
Sum squared resid	0.484452	Schwarz criterion		-2.594561
Log likelihood	160.4612	F-statistic		69.68511
Durbin-Watson stat	2.087124	Prob(F-statistic)		0.000000



Application 2: The term structure of interest rates

- Interest rates are the price of money and, in equilibrium, interest rates equate the amount of borrowing to the amount of saving.
- The Term Structure of Interest Rates shows the relation between interest rates for different term-to-maturity loans.
- In the most basic sense, theories to explain the term structure are still based on interest rates equating the supply and demand for loanable funds.
- Different rates may exist over different terms because of expectations of changing inflation and differing preferences regarding longer-term vs. shorter-term saving.

The term structure of interest rate

The relation of long and short bonds?

- Monthly return on the $n-$month bond at time t: $R_{n,t}$, $n = 1, 3$

  Rule of one price (Expectations Hypothesis)

$$\left(1 + R_{3,t}\right)^3 = \left(1 + R_{1,t}\right)\left(1 + E_t R_{1,t+1}\right)\left(1 + E_t R_{1,t+2}\right)$$

$$R_{3,t} \approx \left[R_{1,t} + E_t R_{1,t+1} + E_t R_{1,t+2}\right]/3$$

  where E_t is expectation formed at time t

- If $R_{1,t}$ follows a random walk: $R_{1,t+1} = R_{1,t} + v_{t+1}$ with $E_t v_{t+1} = 0$ then $E_t R_{1,t+1} = E_t R_{1,t+2} = R_{1,t}$ and

$$R_{3,t} \approx R_{1,t}$$

- Test the null $H_0 : \beta_0 = 0, \beta_1 = 1$ in $R_{3,t} = \beta_0 + \beta_1 R_{1,t} + u_t$.

Term structure of interest rate: example

- Consider that given expectations for inflation over the next year, investors require 4% for a one year loan.
- Suppose investors currently expect inflation for the next year (the second year) to be higher so that they expect to require 6% for a one year loan (starting one year from now).

- Then, the Pure-Expectations Hypothesis, is consistent with the current 2-year spot rate defined as follows:

$$(1 + R_{2,t})^2 = (1 + R_{1,t})(1 + E_t[R_{1,t+1}) = (1.04)x(1.06)$$

so $R_{2,t} = 4.995238\%$

- Restated, if we observe $R_{1,t} = 4\%$ and $R_{2,t} = 4.995238\%$, then, under the Pure-Expectations Hypothesis, we would have $E_t[R_{1,t+1}]$ to be 6%.

Term Structure

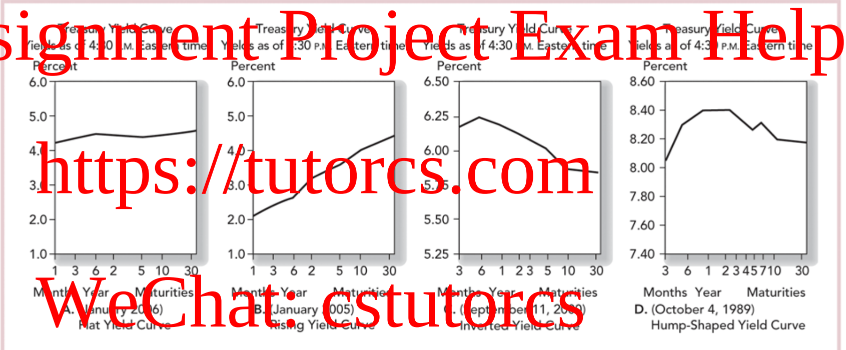


FIGURE 15.1 Treasury yield curves

Source: Various editions of *The Wall Street Journal*. Reprinted by permission of *The Wall Street Journal*, © 1989, 2000, 2006 Dow Jones & Company, Inc. All Rights Reserved Worldwide.

Application 3: Present Value (Gordon) Model

- The price of the asset today is a discounted sum of all possible future cash flows (or dividends D)

$$P_t = \sum_{j}^{\infty} \frac{E_t\left(D_{t+j}\right)}{(1+R)^j}$$

$$E_t\left(D_{t+j}\right) = D_t$$

Present Value (Gordon) Model

$$
\begin{aligned}
P_t &= D_t \left[\frac{1}{1+R} + \frac{1}{(1+R)^2} + K \right] \\
&= \frac{D_t}{1+R} \left[1 + \frac{1}{1+R} + \frac{1}{(1+R)^2} + K \right] \\
&= \frac{D_t}{1+R} \left[\frac{1}{1 - 1/(1+R)} \right] \\
&= \frac{D_t}{R}
\end{aligned}
$$

- we used the property of the infinite converging geometric progression series: $\sum_{k=0}^{\infty} a^k = 1/(1-a)$

Present Value (Gordon) Model

- the model is still nonlinear, but we may take the logs: $\log(P_t) = -\log(R_t) + \log(D_t)$
- And once again use OLS to test whether the model is correct

$$\log(P_t) = \alpha + \beta \log(D_t) + u_t$$

Test the null hypothesis $H_0 : \beta = 1$

Ordinary Least Squares (OLS) Estimation

Consider a K-variable linear regression model

$$Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \ldots + \beta_K X_{iK} + \mu_i, \quad i = 1, \ldots, N$$

The OLS estimator minimises the sum of squared residuals of the sample regression function, i.e.

$$\min_{\hat{\beta}_1, \ldots, \hat{\beta}_K} \sum \hat{\mu}_i^2 = \sum \left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{i2} - \hat{\beta}_3 X_{i3} - \ldots - \hat{\beta}_K X_{iK} \right)^2$$

This yields a system of K equations (first order conditions) in K unknowns which can be solved for the OLS estimator.

Model in matrix Form

The equation above is a shorthand expression for the system

$$
\begin{aligned}
Y_1 &= \beta_1 + \beta_2 X_{12} + \beta_3 X_{13} + \ldots + \beta_K X_{1K} + \mu_1 \\
Y_2 &= \beta_1 + \beta_2 X_{22} + \beta_3 X_{23} + \ldots + \beta_K X_{2K} + \mu_2 \\
&\vdots \\
Y_N &= \beta_1 + \beta_2 X_{N2} + \beta_3 X_{N3} + \ldots + \beta_K X_{NK} + \mu_N
\end{aligned}
$$

Alternatively, this system can be written in matrix notation

$$
\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & X_{12} & X_{13} & \cdots & X_{1K} \\ 1 & X_{22} & X_{23} & \cdots & X_{2K} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{N2} & X_{N3} & \cdots & X_{NK} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix}
$$

Model in matrix Form

This matrix notation can be simplified to

$$y = X\beta + \mu$$

by defining

- y is a $N \times 1$ column vector holding N observations on Y
- X is a $N \times K$ matrix holding N observations on the K explanatory variables
- β is a $K \times 1$ vector holding K unknown parameters
- μ is a $N \times 1$ column vector holding N disturbances μ

Deriving OLS Estimator

The OLS estimator can be obtained by minimising the sum of squared residuals from the sample regression function. In matrix notation, this amounts to minimising $\hat{\mu}'\hat{\mu}$ since:

$$\hat{\mu}'\hat{\mu} = \begin{bmatrix} \hat{\mu}_1 & \hat{\mu}_2 & \cdots & \hat{\mu}_N \end{bmatrix} \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \vdots \\ \hat{\mu}_N \end{bmatrix} = \hat{\mu}_1^2 + \hat{\mu}_2^2 + \cdots + \hat{\mu}_N^2 = \sum \hat{\mu}_i^2$$

with $\hat{\mu} = y - X\hat{\beta}$.
Therefore

$$\begin{aligned} \hat{\mu}'\hat{\mu} &= \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right) \\ &= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= y'y - 2y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

Deriving OLS Estimator

The OLS estimator for β can be derived from the first order conditions

$$\frac{\partial \hat{\mu}'\hat{\mu}}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

as

$$X'X\hat{\beta} = X'y$$
$$\hat{\beta} = \left(X'X\right)^{-1}X'y$$

Features of the $(X'X)$ matrix:

- Symmetrical (see example below)
- Invertible if there is no exact multicollinearity

Linear Regression: Basic Assumptions

1. No collinearity in X: crucial for inverse of $X'X$ to exist

   Intuitively: all regressors are non-redundant

2. Zero conditional mean $E(\mu|X) = 0$ (μ independent of X)

   Violation means endogeneity, a serious problem

$$
\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + \mu) \\
&= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'\mu = \beta + (X'X)^{-1}X'\mu \\
E(\hat{\beta}) &= \beta + E\left[(X'X)^{-1}X'\mu\right] \\
&= \beta + (X'X)^{-1}X'E[\mu|X]
\end{aligned}
$$

   The latter assuming X deterministic. Otherwise, we can use the law of iterated expectations.

3. Homoskedasticity: $Var(\mu_i|X) = \sigma^2$ for all i

Linear regression: variance-covariance

$$
\begin{aligned}
var\left(\hat{\beta}\right) &= E\left[\left(\hat{\beta}-\beta\right)\left(\hat{\beta}-\beta\right)'\right] \\
&= E\left[\left(\left(X'X\right)^{-1}X'\mu\right)\left(\left(X'X\right)^{-1}X'\mu\right)'\right] \\
&= E\left[\left(X'X\right)^{-1}X'\mu\mu'X\left(X'X\right)^{-1}\right] \\
&= \left(X'X\right)^{-1}X'E\left[\mu\mu'\right]X\left(X'X\right)^{-1} \\
&= \left(X'X\right)^{-1}X'\sigma^2 I_N X\left(X'X\right)^{-1} \\
&= \sigma^2\left(X'X\right)^{-1}
\end{aligned}
$$

An unbiased estimator of σ^2 is given by

$$
\hat{\sigma}^2 = \frac{\sum \hat{\mu}_i^2}{N-K} = \frac{\hat{\mu}'\hat{\mu}}{N-K} = \frac{y'y-\hat{\beta}'X'y}{N-K}
$$

Linear regression: Properties (BLUE)

1. **Unbiased**: $E(\hat{\beta}) = \beta$
2. **Consistent**: plim$\hat{\beta} = \beta$ as T goes to infinity (∞).

   $P\left(\hat{\beta} - \beta > \epsilon\right) \to 0$ as T goes to ∞.

   Intuitively $\hat{\beta}$ gets closer and closer to β as $T \to \infty$.

   (does not imply unbiaseness, may sill be that $E(\hat{\beta}) \neq \beta$)
3. **Asymptotically** normal (Why?): $\hat{\beta} = \beta + (X'X)^{-1}X'\mu$

   (or **Exat** normality if μ are normally distributed)

   $$\hat{\beta} \sim N\left(\beta, \sigma^2(X'X)^{-1}\right)$$
4. **Efficient among linear estimators:**

   OLS has smallest variance among linear estimators