Extensions of the Two-Variable Linear Regression Model
Part I

$Y_i \rightarrow GPDI_i = \hat{\beta}_1 + \hat{\beta}_2 GDP_i \rightarrow X_i$

Scaling and Units of Measurement

Consider the data given in Table 6.2, which refers to U.S. gross private domestic investment (GPDI) and gross domestic product (GDP) in billions as well as millions of (chained) 2000 dollars.

Suppose in the regression of GPDI on GDP one researcher uses data in billions of dollars but another expresses data in millions of dollars.

- **Will the regression results be the same in both cases?**
- **Do the units in which the regressand and regressor are measured make any difference in the regression results?**

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

$$where\ Y = GDPI\ \ and\ \ X = GDP$$

$$Y_i^* = w_1 Y_i$$

$$X_i^* = w_2 X_i$$

Where w_1 and w_2 are constants, call the **Scale factors**

- If Y_i and X_i are measured in billions of dollars and one wants to express them in millions of dollars, we will have

$$Y_i^* = 1000\ Y_i$$
$$X_i^* = 1000\ X_i$$

$$w_1 = w_2 = 1000$$

Y_i^* million dollars

X_i^* " ——— "

$$Y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* X_i + \hat{u}_i$$

$$where\ Y_i^* = w_1 Y_i, X_i^* = w_2 X_i, and\ \hat{u}_i^* = w_1 \hat{u}_i$$

Y_i → billion dollars

X_i → billion "

Original unit

Billion dollars

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_2 = \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\text{var}(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2} \sigma^2$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

New Unit

New unit - Million

$$\hat{\beta}_1^* = \bar{Y}^* - \hat{\beta}_2^* \bar{X}^*$$

$$\hat{\beta}_2^* = \frac{\sum x_i^* y_i^*}{\sum x_i^{*2}}$$

$$\text{var}(\hat{\beta}_1^*) = \frac{\sum X_i^{*2}}{n \sum x_i^{*2}} \sigma^{*2}$$

$$\text{var}(\hat{\beta}_2^*) = \frac{\sigma^{*2}}{\sum x_i^{*2}}$$

$$\hat{\sigma}^{*2} = \frac{\sum \hat{u}_i^{*2}}{n-2}$$

$Y \to w_1 = 1000$

$X \to w_2 = 1000$

$$\hat{\beta}_2^* = \left(\frac{w_1}{w_2}\right)\hat{\beta}_2$$

$$\hat{\beta}_1^* = w_1\hat{\beta}_1$$

$$\hat{\sigma}^{*2} = w_1^2\hat{\sigma}^2$$

$$\text{var}(\hat{\beta}_1^*) = w_1^2\ \text{var}(\hat{\beta}_1)$$

$$\text{var}(\hat{\beta}_2^*) = \left(\frac{w_1}{w_2}\right)^2 \text{var}(\hat{\beta}_2)$$

$$r_{xy}^2 = r_{x^*y^*}^2$$

$u_i = $

Billion → Million

$w_1 = 1000$

$w_2 = 1000$

$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 X_i$

$\hat{\beta}_2 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$

$\sum \hat{u}_i^2 \simeq 0$

Example

Gross Private Domestic Investment and GDP, United States, 1990-2005

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

$$where\ Y_i = GPDI\ \ and\ \ X_i = GDP$$

TABLE 6.2
Gross Private Domestic Investment and GDP, United States, 1990–2005 (Billions of chained [2000] dollars, except as noted; quarterly data at seasonally adjusted annual rates)

Source: *Economic Report of the President*, 2007, Table B-2, p. 328.

Year	GPDIBL	GPDIM	GDPB	GDPM
1990	886.6	886,600.0	7,112.5	7,112,500.0
1991	829.1	829,100.0	7,100.5	7,100,500.0
1992	878.3	878,300.0	7,336.6	7,336,600.0
1993	953.5	953,500.0	7,532.7	7,532,700.0
1994	1,042.3	1,042,300.0	7,835.5	7,835,500.0
1995	1,109.6	1,109,600.0	8,031.7	8,031,700.0
1996	1,209.2	1,209,200.0	8,328.9	8,328,900.0
1997	1,320.6	1,320,600.0	8,703.5	8,703,500.0
1998	1,455.0	1,455,000.0	9,066.9	9,066,900.0
1999	1,576.3	1,576,300.0	9,470.3	9,470,300.0
2000	1,679.0	1,679,000.0	9,817.0	9,817,000.0
2001	1,629.4	1,629,400.0	9,890.7	9,890,700.0
2002	1,544.6	1,544,600.0	10,048.8	10,048,800.0
2003	1,596.9	1,596,900.0	10,301.0	10,301,000.0
2004	1,713.9	1,713,900.0	10,703.5	10,703,500.0
2005	1,842.0	1,842,000.0	11,048.6	11,048,600.0

Note: GPDIBL = gross private domestic investment, billions of 2000 dollars.
GPDIM = gross private domestic investments, millions of 2000 dollars.
GDPB = gross domestic product, billions of 2000 dollars.
GDPM = gross domestic product, millions of 2000 dollars.

GPDIBL = Gross private domestic investment, billions of 2000 dollars
GPDIM = Gross private domestic investment, millions of 2000 dollars
GDPB = Gross domestic product, billions of 2000 dollars
GDPM= Gross domestic product, millions of 2000 dollars

Y			① $w_1 = 1$ from 1000 → 100 unit
$'1000	$	$100	② $w_1 = \frac{1}{100}$
1	1000	10	③ $w_1 = 100$
2	2000	20	
3	3000	30	
4	4000	40	
5	5000	50	

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

If X ↑ 1 unit on average Y increase $\hat{\beta}_2$ thousand dollars.

i	X $thousand	$	X $ hundred
1	2	2000	20
2	3	3000	30
3	1	1000	10
4	4	4000	40
5	5	5000	50

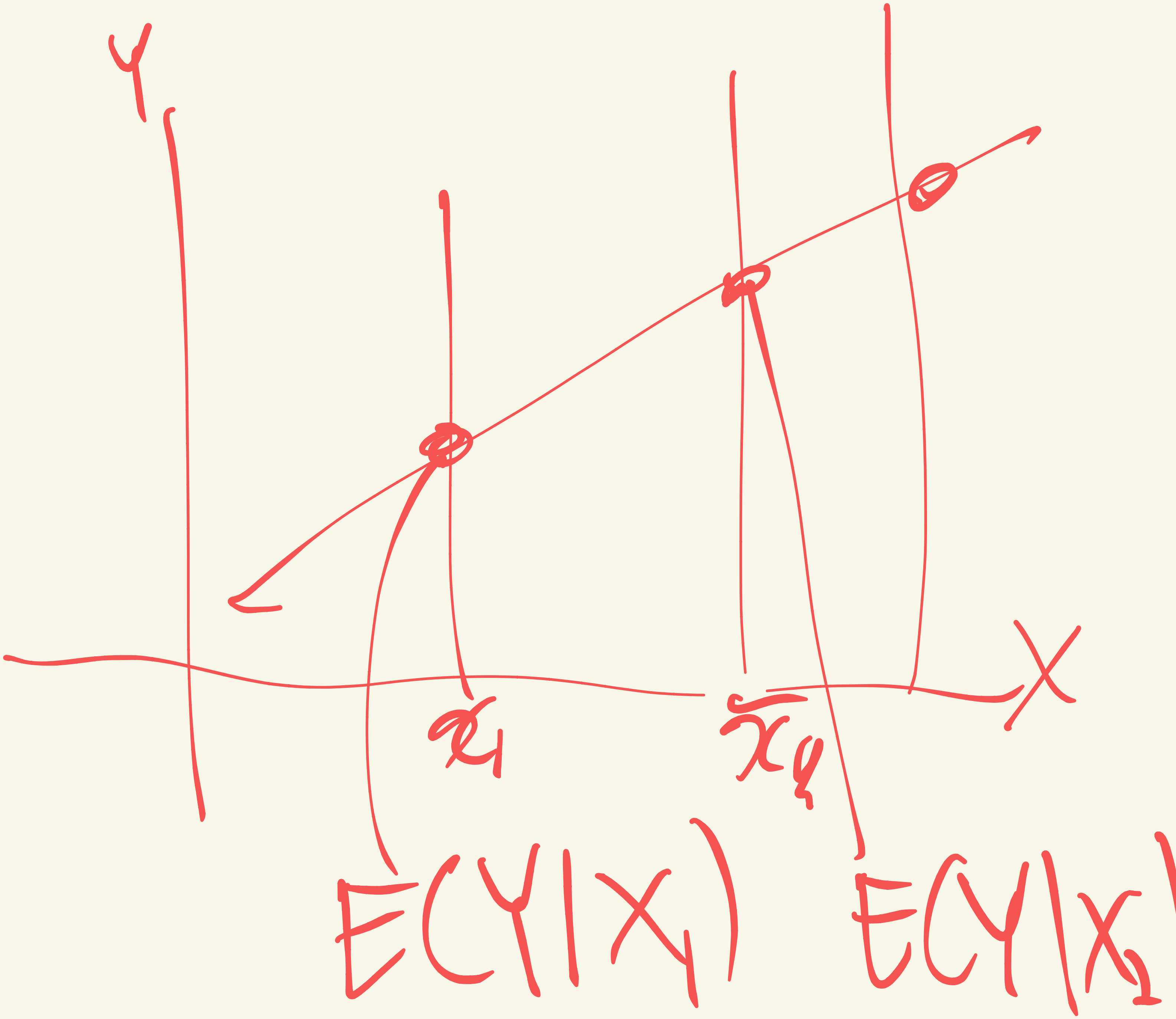
i	hundred	thousand	tenth
1	20	2	200
2	30	3	300
3	40	4	400
4	50	5	500
5	60	6	600

Both GPDI and GDP in billions of dollars:

$$\widehat{GPDI}_t = -926.090 + 0.2535GDP_t$$
$$se = (116.358) \quad (0.0129)$$
$$r^2 = 0.9648$$

GPDI in billions of dollars → millions of dollars

GDP in billions of dollars → millions of dollars



Y
X
x1
x4
E(Y|X)

Billion → Million

$$w_1 = 1000$$
$$w_2 = 1000$$

$$w_1\hat{\beta}_1 = 1000 * -926.090 = -926{,}090$$

$$\frac{w_1}{w_2}\hat{\beta}_2 = \frac{1000}{1000} * 0.2535 = 0.2535$$

$$\widehat{GPDI}_t = -926{,}090 + 0.2535 GDP_t$$
$$se = (116{,}358) \quad (0.0129)$$
$$r^2 = 0.9648$$

Million dollars.

If GDP increase by 1 million USD on average GPDI would increase 0.235 million

Both GPDI and GDP in millions of dollars:

$$\widehat{GPDI}_t = -926{,}090 + 0.2535 GDP_t$$
$$se = (116{,}358) \quad (0.0129)$$
$$r^2 = 0.9648$$

GDPI in millions of dollars ⟶ billions of dollars

GDP in millions of dollars ⟶ billions of dollars

$$w_1 = \frac{1}{1000}$$

$$w_2 = \frac{1}{1000}$$

$$w_1\hat{\beta}_1 = \frac{1}{1000} * -926{,}090 = -926.090$$

$$\frac{w_1}{w_2}\hat{\beta}_2 = \frac{\frac{1}{1000}}{\frac{1}{1000}} * 0.2535 = 0.2535$$

$$\widehat{GPDI}_t = -926.090 + 0.2535 GDP_t$$

$$se = (116.358) \quad (0.0129)$$

$$r^2 = 0.9648$$

Both GDPI and GDP in billions of dollars:

$$\widehat{GPDI}_t = -926.090 + 0.2535 GDP_t$$
$$se = (116.358) \quad (0.0129)$$
$$r^2 = 0.9648$$

$w_1 = 1$

GDPI in billions of dollars ⟶ billions of dollars

GDP in billions of dollars ⟶ millions of dollars

X unit change $w_2 = 1000$

$$w_1 = 1$$
$$w_2 = 1000$$

$$w_1\hat{\beta}_1 = 1 * -926.090 = -926.090$$

$$\frac{w_1}{w_2}\hat{\beta}_2 = \frac{1}{1000} * 0.2535 = 0.0002535$$

$$\widehat{GPDI}_t = -926.090 + 0.0002535GDP_t$$
$$se = (116.358) \quad (0.0000129)$$
$$r^2 = 0.9648$$

Both GPDI and GDP in millions of dollars:

$$\widehat{GPDI}_t = -926{,}090 + 0.2535 GDP_t$$
$$se \quad = (116{,}358) \qquad (0.0129)$$
$$r^2 = 0.9648$$

GDPI in millions of dollars → billions of dollars

GDP in millions of dollars → millions of dollars

$$w_1 = \frac{1}{1000}$$
$$w_2 = 1$$

$$w_1\hat{\beta}_1 = \frac{1}{1000} * -926{,}090 = -926.090$$

$$\frac{w_1}{w_2}\hat{\beta}_2 = \frac{\frac{1}{1000}}{1} * 0.2535 = 0.0002535$$

$$\widehat{GPDI}_t = -926.090 + 0.0002535 GDP_t$$
$$se \quad = (116.358) \quad (0.0000129)$$
$$r^2 = 0.9648$$

Both GDPI and GDP in billions of dollars:

$$\hat{GPDI}_t = -926.090 + 0.2535 GDP_t$$

$$se = (116.358) \quad (0.0129)$$

$$r^2 = 0.9648$$

GDPI in billions of dollars ⟶ millions of dollars

GDP in billions of dollars ⟶ billions of dollars

$$w_1 = 1000$$

$$w_2 = 1$$

$$\hat{\beta}_1^* = w_1 \hat{\beta}_1 = 1000 * -926.090 = -926,090$$

$$\hat{\beta}_2^* = \left(\frac{w_1}{w_2} \right) \hat{\beta}_2 = \left(\frac{1000}{1} \right) 0.2535 = 253.524$$

$$\widehat{GPDI}_t = -926,090 + 253.524 GDP_t$$

$$se = (116,358) \quad (12.9465)$$

$$r^2 = 0.9648$$

Both GPDI and GDP in millions of dollars:

$$\widehat{GPDI}_t = -926.090 + 0.2535 GDP_t$$
$$se = (116.358) \quad (0.0129)$$
$$r^2 = 0.9648$$

GDPI in millions of dollars ⟶ millions of dollars

GDP in millions of dollars ⟶ billions of dollars

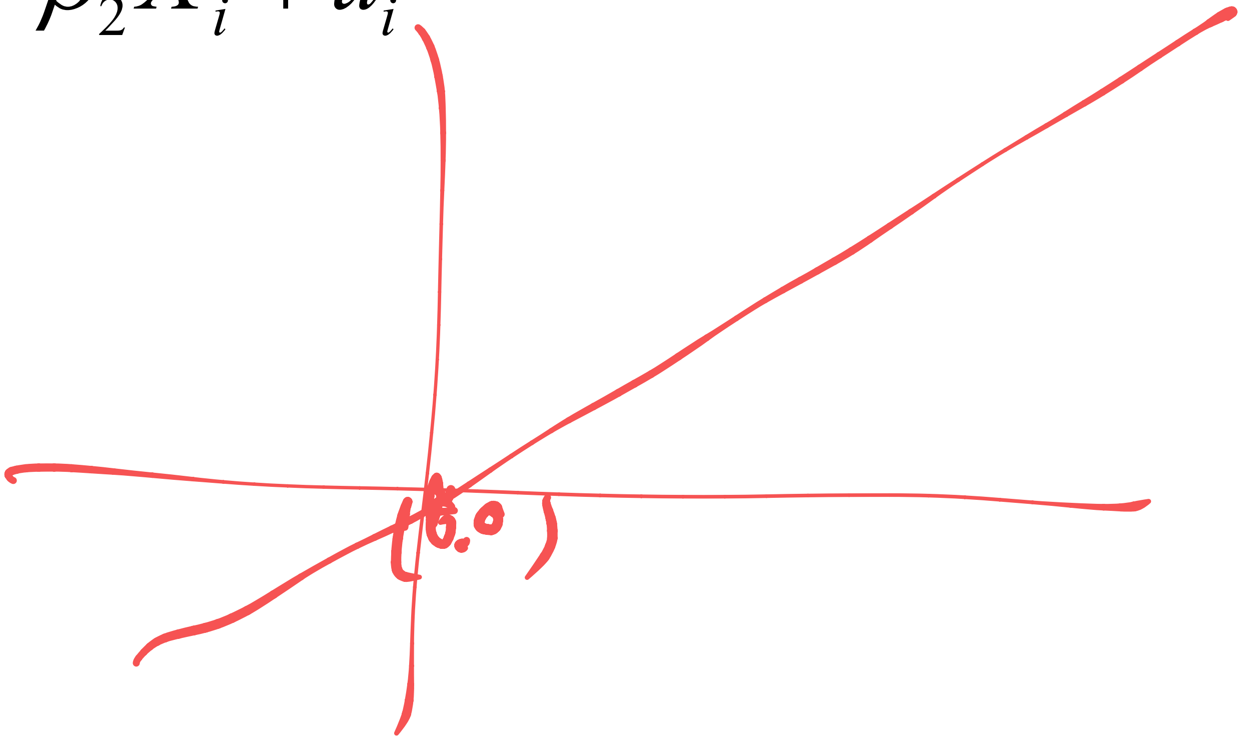
$$w_1 = 1$$
$$w_2 = \frac{1}{1000}$$
$$w_1\hat{\beta}_1 = 1 * -926{,}090 = -926{,}090$$

$$\frac{w_1}{w_2}\hat{\beta}_2 = \frac{1}{\frac{1}{1000}} * 0.2535 = 253.524$$

$$\widehat{GPDI}_t = -926{,}090 + 253.524 GDP_t$$
$$se = (116{,}358) \quad (12.9465)$$
$$r^2 = 0.9648$$

Regression through the origin

$$Y_i = \beta_2 X_i + u_i$$



$Y_i = \hat{\beta}_1 - \hat{\beta}_2 X_i + \hat{\beta}_3 X_2$ (n-3) → $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + u_i$

Regression through the origin

$Y_i = \hat{\beta}_2 X_i + \hat{u}_i$

$$\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})}$$

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$\mathrm{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2}, \textit{ where } \sigma^2 = \frac{\sum \hat{u}_i^2}{n-1}$$

$\hat{u} = Y_i - \hat{Y}_i$

$$\sigma^2 = \frac{\sum \hat{u}_i^2}{n-2}$$

R-squared for Regression through Origin Model

$$raw\ r^2 = \frac{(\sum X_i Y_i)^2}{\sum X_i^2 \sum Y_i^2}$$

The capital asset pricing model (CAPM)

- Risk premium form may be expressed as

$$\left(ER_i - r_f\right) = \beta_i \left(ER_m - r_f\right)$$

Where

ER_i=expected rate of return on security i

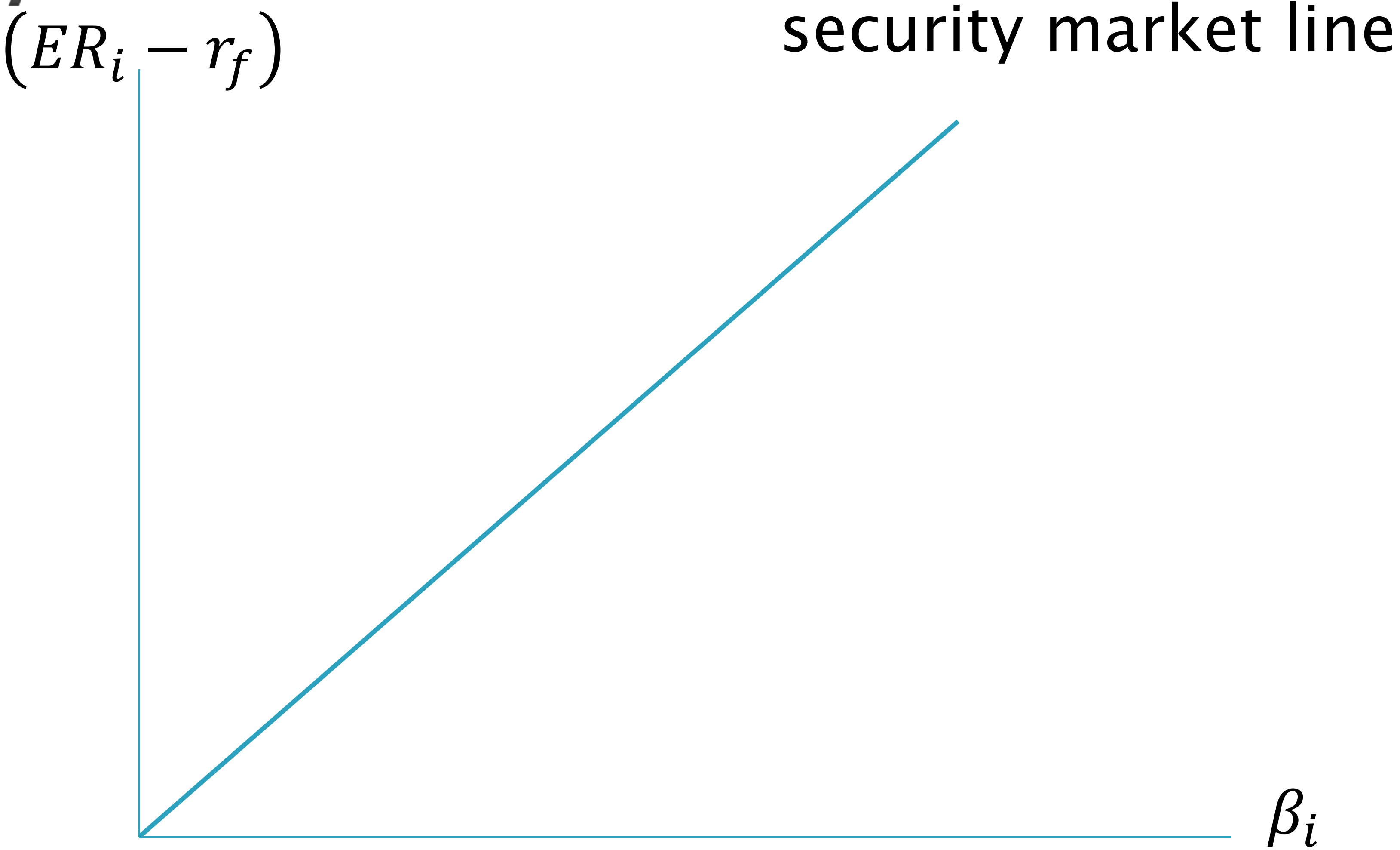
ER_m=expected rate of return on the market portfolio as represented by, say, the S&P 500 companies stock index

r_f= risk-free rate of return, say, the return on 90 day Treasury bills

β_i = the Beta coefficient, a measure of systematic risk, risk that cannot be eliminated through diversification. A measure of the extent to which the ith security rate of return moves with the market. A $\beta_i > 1$ implies a volatile or aggressive security, whereas $\beta_i < 1$ suggests a defensive security

- If capital market works efficiently, then CAPM postulates that security i's expected risk premium $(ER_i - r_f)$ is equal to that security's β coefficient times the expected market risk premium $(ER_m - r_f)$
- The line shown in the figure is known as the security market line (SML)

Systematic risk



- For empirical purposes, the equation is often expressed as

$$(R_i - r_f) = \beta_i (R_m - r_f) + u_i$$

$Y_i = \gamma_1 + \gamma_2 X_i$

Or

$Y_i = \beta_1 + \beta_2 X_i$

$Y_i = \alpha_1 + \alpha_2 X_i$

$$(R_i - r_f) = \alpha_i + \beta_i (R_m - r_f) + u_i \quad (*)$$

The latter model is known as the Market Model. If CAPM holds, α_i is expected to be zero.

$H_0: \beta_1 = 0$ intercept

$H_1: \beta_1 \neq 0$

$H_0: \beta_2 = 0$ Slope

$H_1: \beta_2 \neq 0$

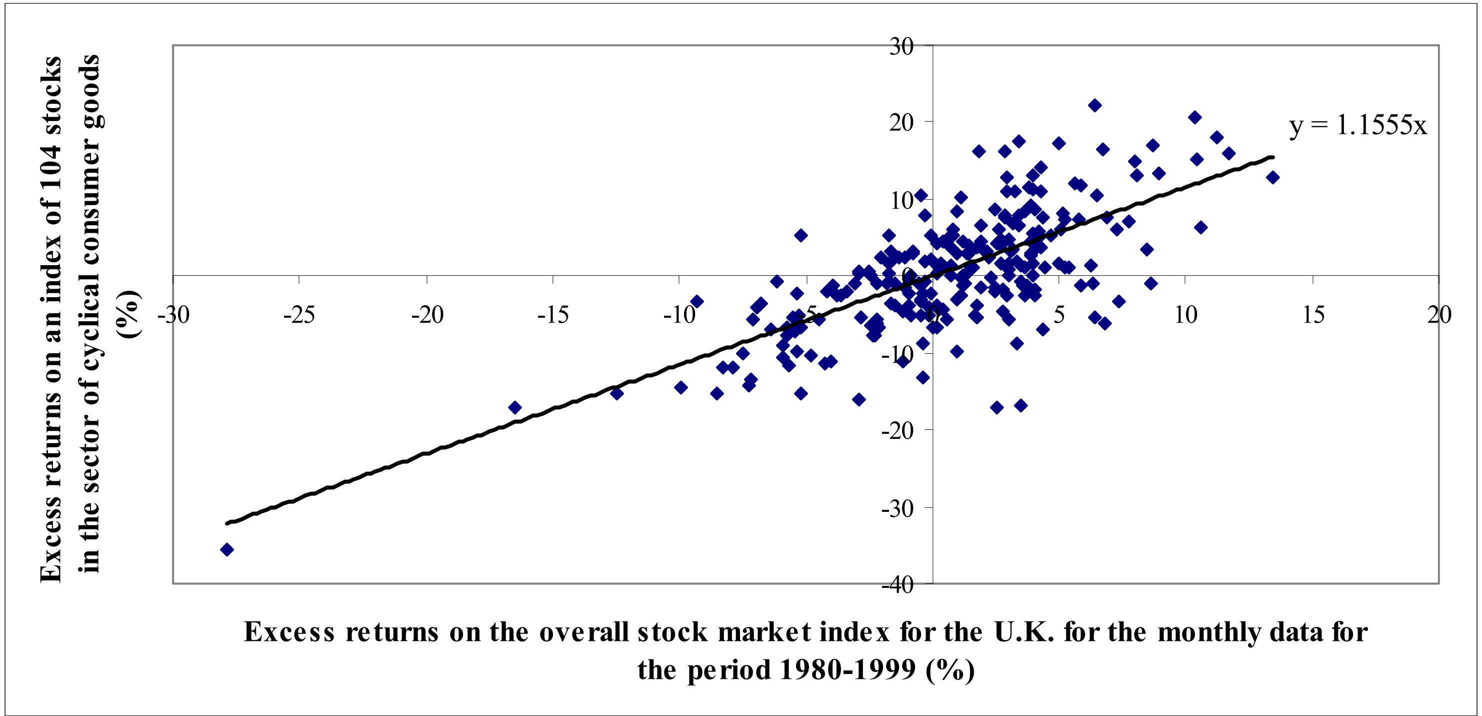
- The dependent variable, Y is $(R_i - r_f)$

- The explanatory variable X is β_i the volatile coefficient

Example

Table 6.1 (P.151) gives data on excess returns Y_t(%)on an index on 104 stocks in the sector of cyclical consumer goods and excess returns X_t (%) on the overall stock market index for the U.K. for the monthly data for the period 1980-1999, for a total of 240 observations.

Excess returns refers to return in excess of return on a riskless asset. (Capital asset pricing model, CAPM)


Excess returns on an index of 104 stocks in the sector of cyclical consumer goods (%)
y = 1.1555x
30
20
10
0
-10
-20
-30
-40
-30
-25
-20
-15
-10
-5
0
5
10
15
20
Excess returns on the overall stock market index for the U.K. for the monthly data for the period 1980-1999 (%)

Dependent variable: Y
Method: OLS
Sample: 1980 M01 1999 M12
N = 240

	Coefficient	Std. Error	T-stat	p-value
X	1.155512	0.074396	15.5320	0.0000

$R^2 = 0.500309$

	Coefficient	Std. Error	T-stat	p-value
X	1.171128	0.075386	15.5350	0.0000
Constant	-0.447481	0.362943	-1.232924	0.2188

$R^2 = 0.503480$

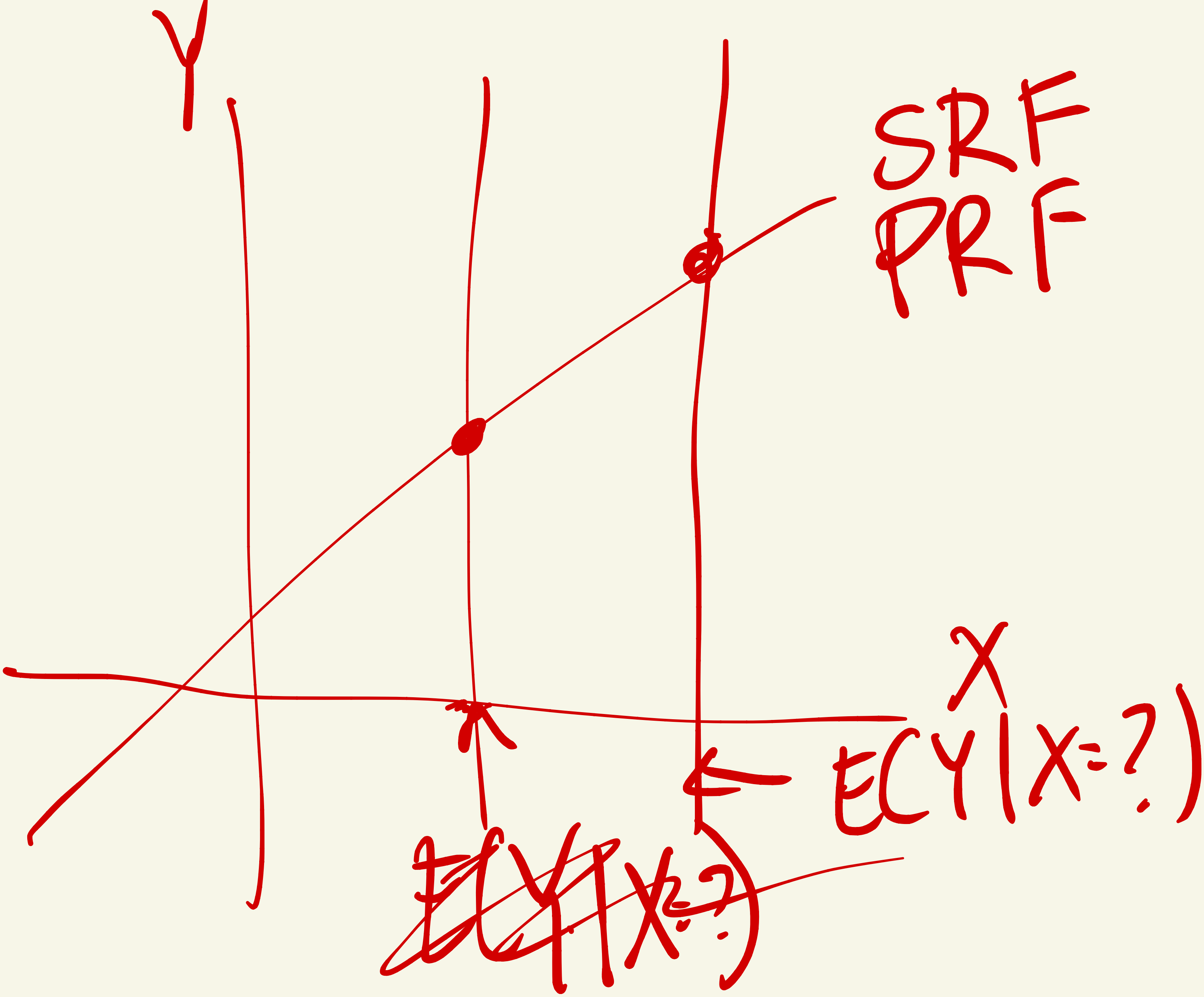
The slope coefficient is **highly significant**

If the excess market rate goes up by **1 percentage point**, the excess return on the index of consumer goods sector goes up by about **1.15 percentage points.**

If a Beta coefficient is greater than 1, such a security is said to be volatile; it moves more than proportionately with the overall stock market index

Source

Gujarati, D.N. (2009) Basic Econometrics. 5th ed. Singapore, McGraw-Hill. (G)



Y
SRF
PRF
X
E(Y|X=?)