MULTIPLE REGRESSION SPECIAL TOPICS

Chapter 4

1. Dummy regression

- When we use categorical (nominal) level variables in regression, we refer to them as dummy variables.
- Dummy variables have only two values (0 and 1), so for variables with more than two categories we must create new variables.
  - One less variable than there are categories.
  - e.g. The four seasons of the year can be captured in one variable with four categories, otherwise, they can be represented by 3 dummy variables.

Dummy Example

Using the pay level data, we estimate the equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

- Where:
  - Y = daily earnings,
  - β_0 = the Y intercept,
  - X_1 = seniority
  - X_2 = gender (dummy)

Results with Gender

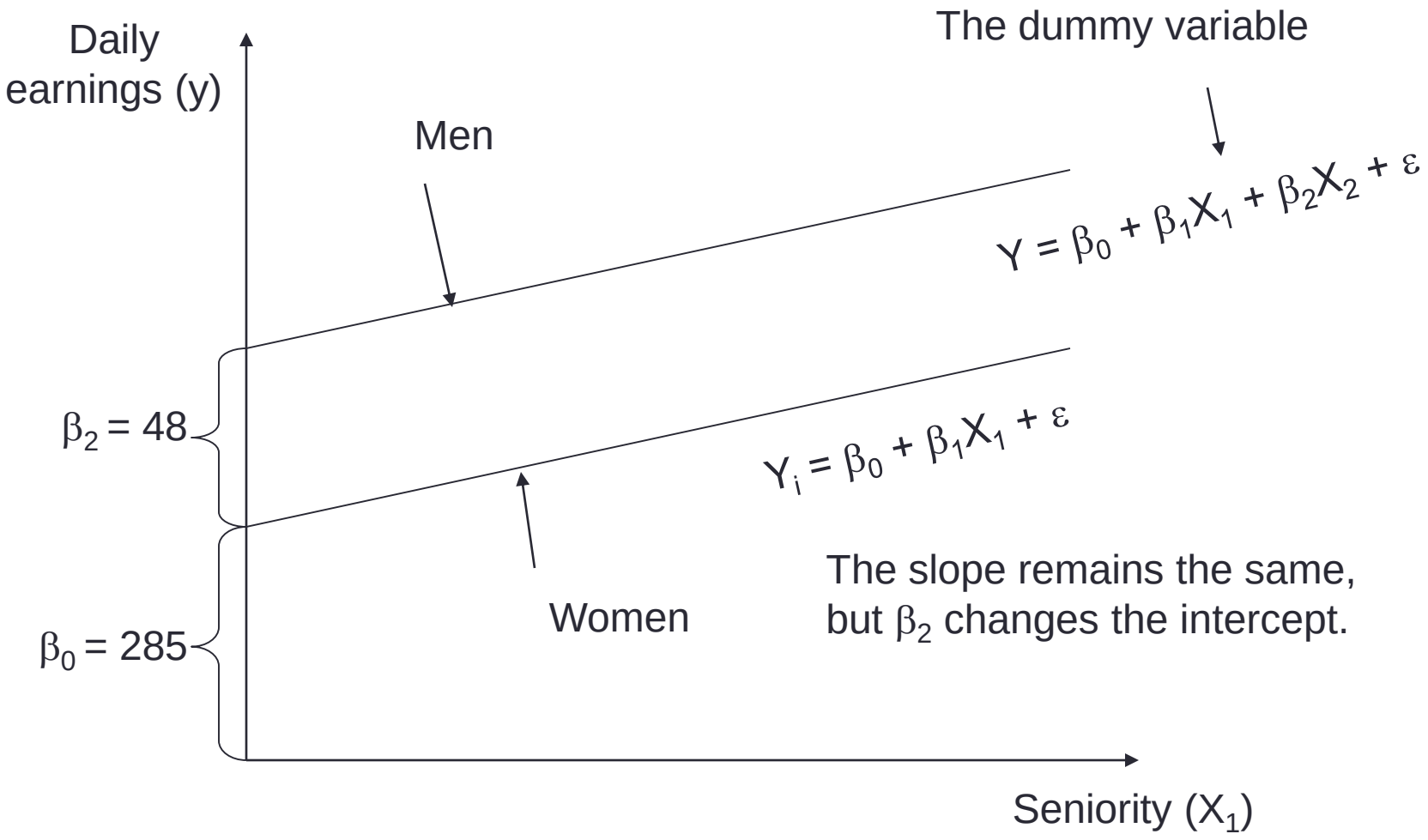
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	285.181	12.799		22.281	.000
	Seniority (in months)	.848	.064	.917	13.171	.000
	Gender	47.756	14.160	.235	3.373	.002

a. Dependent Variable: Daily earnings

Earnings = 285 + .85 Seniority + 48 Gender + ε

Graphically (only showing evaluation)



2. Interaction Effects

- We often want to include interaction effects in multiple regression.
- An interaction term in regression is an independent variable that is the multiple of two or more other independent variables. For example:
  1. Slope dummies interact with the regression model to change the slope.
  2. Product variables (e.g. X_1 x X_2) affect the linearity of the line to make it curvilinear.

Slope Dummies

- Do you believe that, all else being equal, men earn more than women?
- Do you believe that the wage gap grows as women and men gain seniority?

- If you do, you can test your hypothesis with a slope dummy.

Worsening Wage Inequality

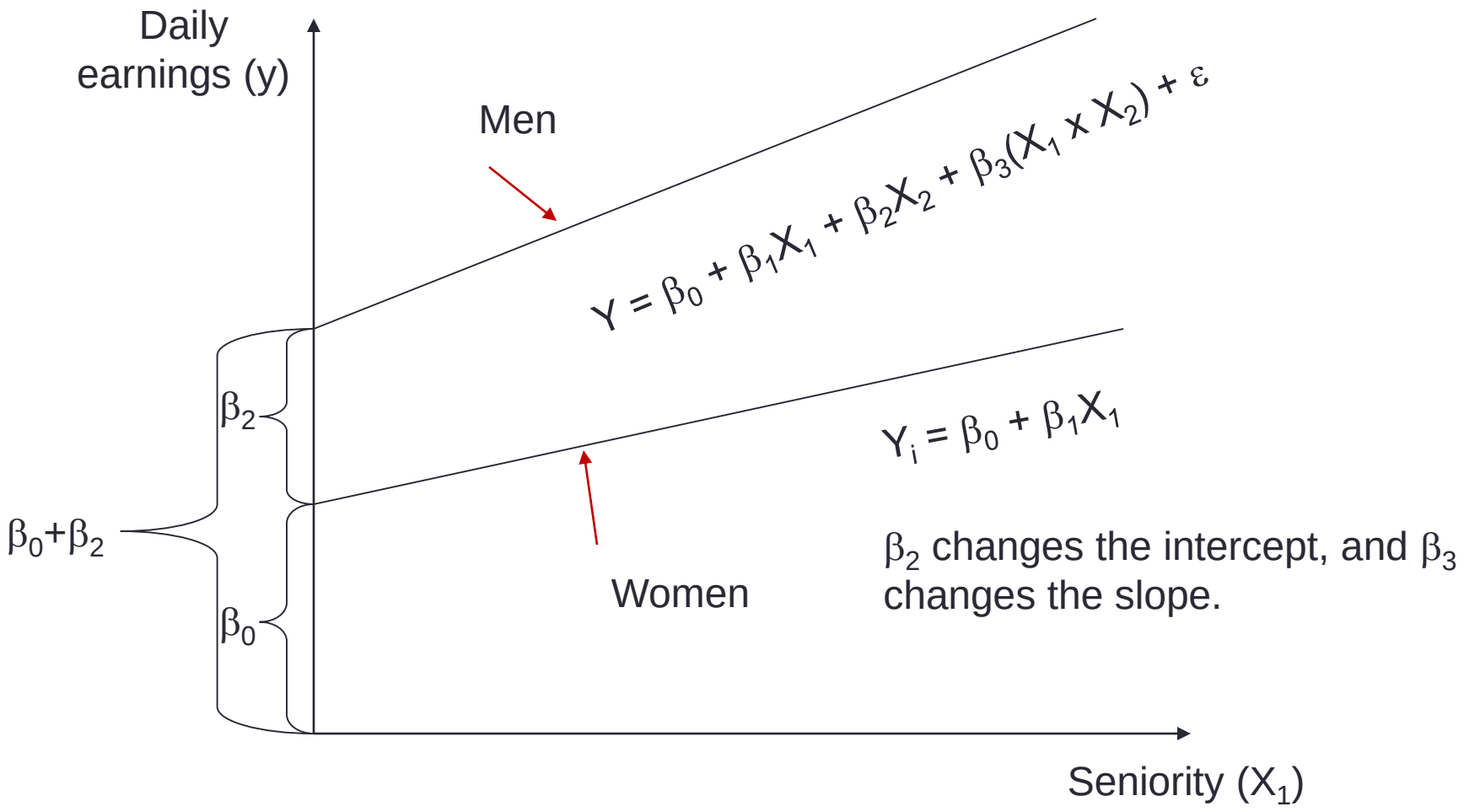
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 \times X_2) + \varepsilon$$

Seniority
Determines slope of line.

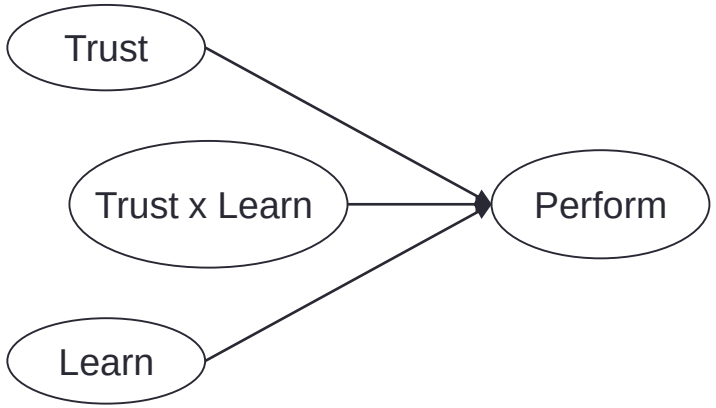
Male/Female Dummy
Changes intercept of line.

Seniority x (Male/Female Dummy)
Changes slope of line.

Slope Dummy Graph

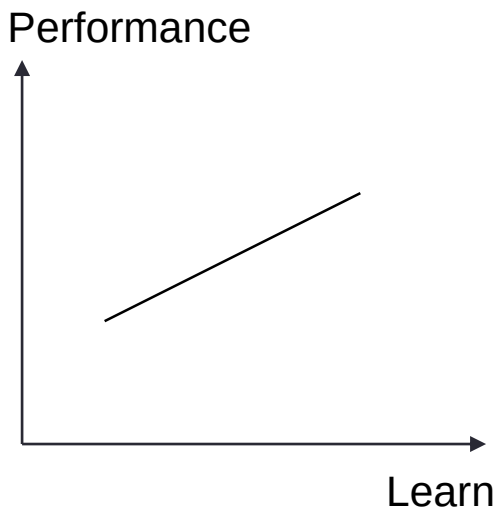


Product Interaction Terms

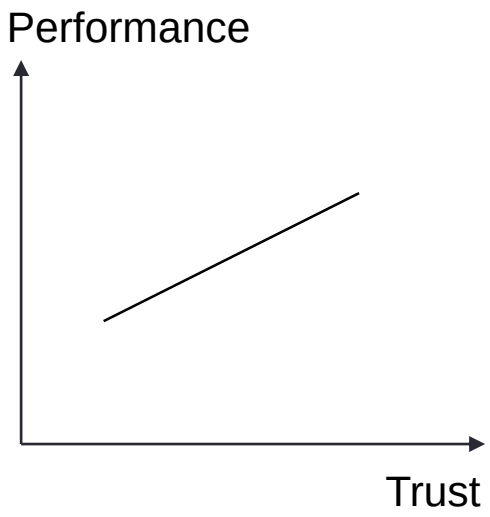


$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 \times X_2) + \varepsilon$$

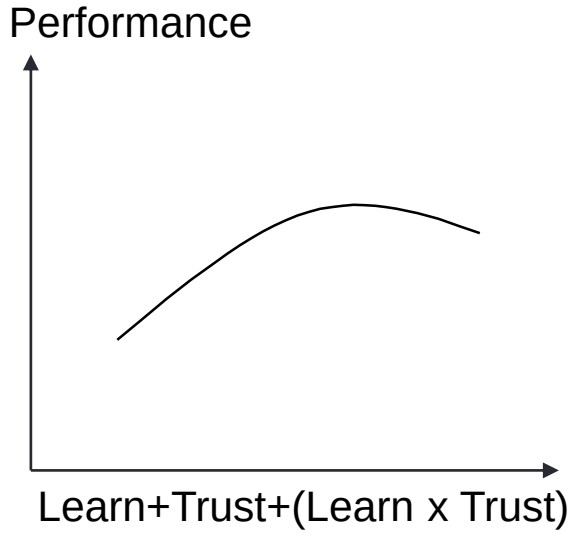
Product Interaction Effect



No Interaction



No Interaction



Interaction

3. Nominal Dependent Variables

We may have a research question that requires a nominal dependent variable.

- E.g. Foreign market entry mode of service firms can be divided between high control (subsidiary) and low control (agents etc.).

There are two main approaches to handling nominal dependent variables:

1. Discriminant analysis.

   Recommended when the dependent variable has more than two groups.

2. Logistic regression.

   Recommended when the dependent variable is dichotomous, there are qualitative and quantitative independent variables, and the underlying assumptions of multivariate normality may not be met.