**Dummy/Indicator Variables and Analysis of Covariance**

transform qualitative variables to incorporate into linear regression model, e.g. sex

set up a dummy variable for each class of the qualitative variable

qualitative variable with classes represented by binomial dummy variables

= speed that an insurance innovation is adopted

= total assets of the firm

= type of firm, stock or mutual company

mutual firms

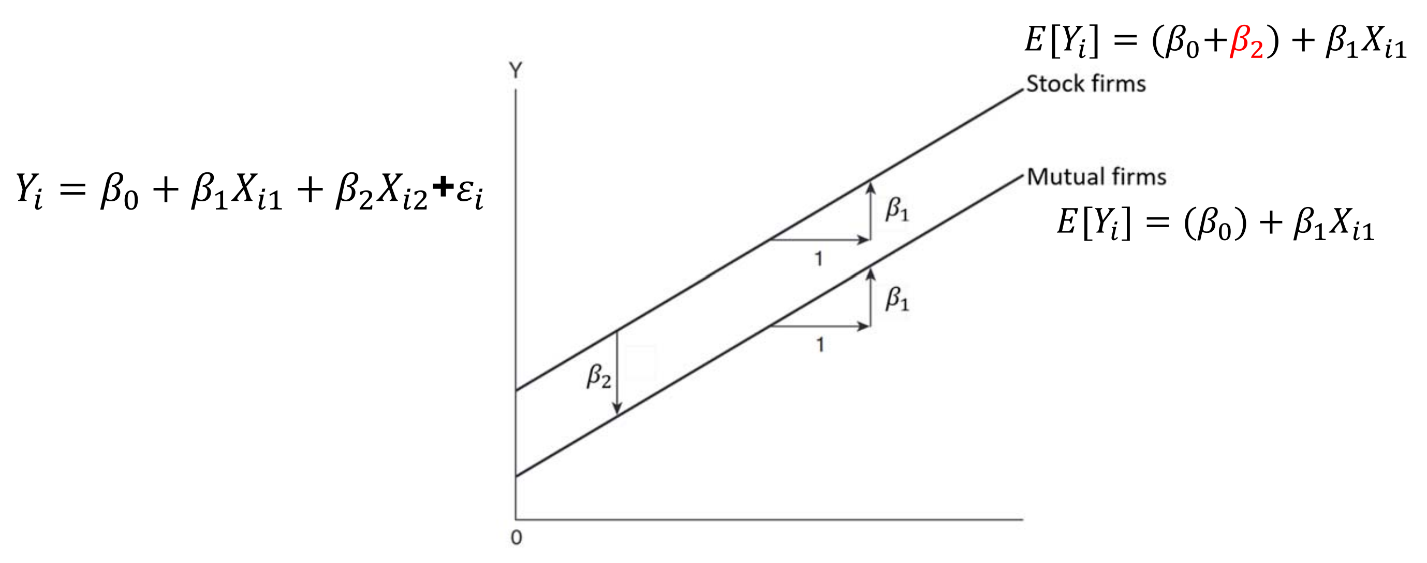
stock firms

= linear function of the size of the firm,

slope for both firms is the same for any given size of the firm

= mean time elapsed before the innovation is adopted

= difference in intercept between the response variable for stock firms than mutual firms for any given size of the firm



**Interaction Effects**

= interaction term between size of firm and type of firm

= mean time elapsed before the innovation is adopted

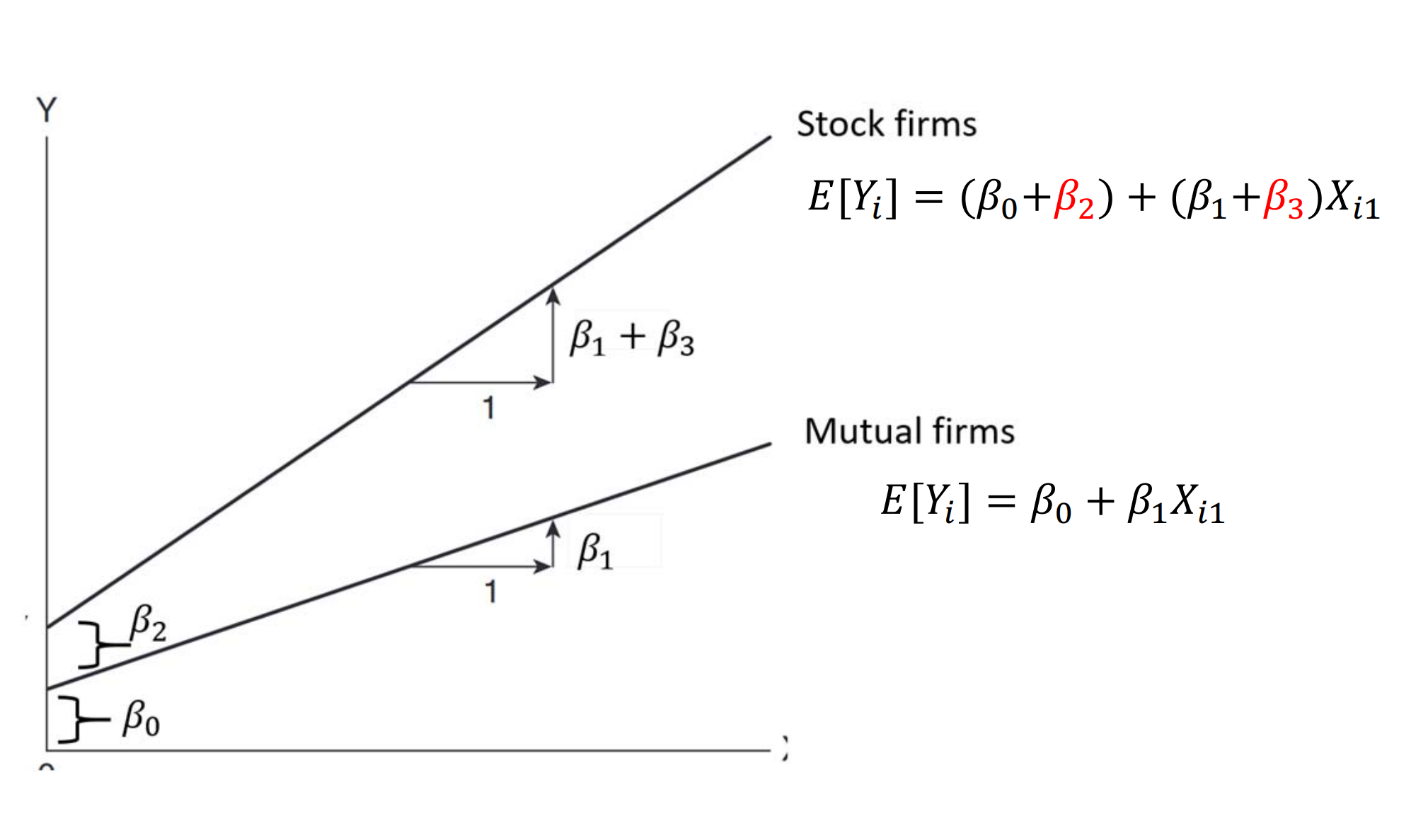
= difference in intercept between the response variable for stock firms than mutual firms

= difference in slope/effect size between the response variable for stock firms than mutual firms for any given size of the firm

if there is no interaction between the two variables, the slope for both firms will remain the same

mutual firms

stock firms



interaction term is a second order polynomial

drop higher order terms if they’re not significant

**Least Square Means/Adjusted Means**

raw mean = an average of the observations without considering other covariates

least square means = mean estimated from a linear regression adjusted for other covariates

least square means are theoretical estimates of the true population mean

**Qualitative Variable with > 2 Classes**

= test score

= number of hours spent studying

= level of education: high school, undergraduate, master’s, doctorate

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Matrix** | | | | | |
| **Level of Education** |  |  |  |  |  |
| high school | 1 |  | 1 | 0 | 0 |
| undergraduate | 1 |  | 0 | 1 | 0 |
| master’s | 1 |  | 0 | 0 | 1 |
| doctorate | 1 |  | 0 | 0 | 0 |

doctorate

high school

undergraduate

master’s

= linear function of the number of hours spend studying,

slope for all levels of education is the same for any given number of hours spent studying

= mean test score

= difference in intercept between the response variable for high school than doctorate for any given number of hours spent studying

= difference in intercept between the response variable for undergraduate than doctorate for any given number of hours spent studying

= difference in intercept between the response variable for master’s than doctorate for any given number of hours spent studying

**ANOVA and ANCOVA**

ANOVA models association between a continuous response variable and continuous quantitative predictor variables

ANCOVA models association between a continuous response variable and a categorical qualitative variable, adjusting for a continuous covariate

e.g. modeling how long two types of companies take to adopt a new innovation, adjusting for size of the company

covariance models = chief independent variables of interest are qualitative and quantitative independent variables are introduced primarily to reduce the variance of error terms