# PB4A7 Seminar 4

## MT Week 4, 23 October 2024

## Overview:

During this seminar:

• We will go through a series of multivariate regression exercises to better understand the relationship between birth weight, maternal smoking, and infant health outcomes.

• Please conduct your analyses on your do-files and keep a log-file.

***About the Dataset:***

We will use a dataset on infants born in the United States, focusing on birth weight, maternal smoking behavior, and various infant health outcomes such as hospital costs and mortality rates. The dataset is designed to explore the effects of birth weight on health outcomes, controlling for other factors like maternal age and education.

\*\* Note: this is the same dataset you will use in Seminar 5.

***Exercise Questions: Multivariate Regression with Multiple Regressors***

**1. Basic Multivariate Regression**

* **Question:** Regress birth weight on maternal smoking behavior. What is the estimated effect of smoking during pregnancy on birth weight?
* **STATA Code:**

reg birthweight smoker

* **Interpretation:** Interpret the coefficient for smoking. What does this result tell us about the impact of smoking on birth weight?

**2. Interaction Terms**

* **Question:** Estimate how the effect of smoking on birth weight differs with different levels of birth weight by adding an interaction term. Does smoking have a larger effect on lower birth weights?
* **STATA Code:**

gen smoker\_birthweight = smoker \* birthweight

reg birthweight smoker smoker\_birthweight

* **Interpretation:** What is the effect of the interaction term? Explain whether smoking has a more significant impact on lower or higher birth weights.

**3. Controlling for Covariates**

* **Question:** Extend the previous model by controlling for maternal age and education. How do these variables influence birth weight? Has the effect of smoking changed?
* **STATA Code:**

reg birthweight smoker smoker\_birthweight age educ

* **Interpretation:** Assess whether the coefficients for smoking change after adding controls. What impact do maternal age and education have on birth weight?

**4. Logistic Regression: Binary Outcome**

* **Question:** Investigate the effect of birth weight on the likelihood of a mother being unmarried using logistic regression. What is the relationship between birth weight and the mother’s marital status?
* **STATA Code:**

logit unmarried birthweight smoker

* **Interpretation:** Interpret the odds ratio for birth weight and smoking. How do these variables affect the likelihood of the mother being unmarried?

**5. Nonlinearity in Birth Weight**

* **Question:** Investigate whether the relationship between birth weight and maternal smoking is nonlinear by including a quadratic term for birth weight.
* **STATA Code:**

gen birthweight\_sq = birthweight^2

reg birthweight birthweight\_sq smoker

* **Interpretation:** Is the relationship between birth weight and smoking nonlinear? Interpret the coefficients for birth weight and its square.

**Key Points:**

**Interpreting Coefficients:** Make sure to interpret both the magnitude and sign of the coefficients in each regression, paying attention to interaction and quadratic terms.

**Model Fit:** Consider the R-squared values and model significance.

**Real-World Implications:** Think about the policy implications of the results, particularly when considering maternal smoking and its impact on birth weight.