Lecture Module 7

Agenda

1. Standardizing Variables
2. Quadratics
3. Interactions
4. Dummies
5. Regression Discontinuity
6. Discrete Dependent Variables

Regressions

Standardizing variables

- Sometimes interpreting β is difficult.
- We can "standardize" to improve interpretation.

$$y^s = \frac{y - \widehat{\mu}_y}{\widehat{\sigma}_y}$$

- Suppose we start with the model:

$$y = \beta_0 + \beta_1 x + u \tag{1}$$

- Take the mean of the entire equation

$$\widehat{\mu}_y = \beta_0 + \beta_1 \widehat{\mu}_x \tag{2}$$

- Subtract (1) from (2):

$$\left(y - \widehat{\mu}_y\right) = \beta_1 \left(x - \widehat{\mu}_x\right) + u$$

- Estimating this will not change the coefficients.

Regressions

Standardizing variables

- Divide both sides by $\widehat{\sigma}_y$

$$\frac{\left(y-\widehat{\mu}_y\right)}{\widehat{\sigma}_y}=\frac{1}{\widehat{\sigma}_y}\beta_1\left(x-\widehat{\mu}_x\right)+\frac{1}{\widehat{\sigma}_y}u$$

- Multiply β_1 by $\frac{\widehat{\sigma}_x}{\widehat{\sigma}_x}$

$$\frac{\left(y-\widehat{\mu}_y\right)}{\widehat{\sigma}_y}=\frac{1}{\widehat{\sigma}_y}\frac{\widehat{\sigma}_x}{\widehat{\sigma}_x}\beta_1\left(x-\widehat{\mu}_x\right)+\frac{1}{\widehat{\sigma}_y}u$$

- Manipulating:

$$\frac{\left(y-\widehat{\mu}_y\right)}{\widehat{\sigma}_y}=\underbrace{\frac{\widehat{\sigma}_x}{\widehat{\sigma}_y}\beta_1}_{\widetilde{\beta}_1}\frac{\left(x-\widehat{\mu}_x\right)}{\widehat{\sigma}_x}+\underbrace{\frac{1}{\widehat{\sigma}_y}u}_{\widetilde{u}}$$

- Simplifying:

$$y^s=\widetilde{\beta}_1x^s+\widetilde{u}$$

Regressions

Standardizing variables

- Wages

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$$

- Standardize variables

$$wage^s = \widetilde{\beta}_1 educ^s + \widetilde{\beta}_2 exper^s + \widetilde{\beta}_3^s tenure + \widetilde{u}$$

- Coefficients and standard errors change
- T-statistics, R^2, P-values do not change!!!
- Which variable appears to be most important in determining wages?

Regressions

Quadratic terms

- The basic model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + u$$

- How should we interpret β_1 , β_2?
  - Together
- Take the derivative w.r.t x_1.

$$\frac{\partial y}{\partial x_1} = \beta_1 + 2\beta_2 x_1$$

- If $\beta_2 \neq 0$, then we can solve for the maximum or minimum

$$\begin{aligned} \beta_1 + 2\beta_2 x_1^* &= 0 \\ x_1^* &= -\frac{\beta_1}{2\beta_2} \end{aligned}$$

- What determines whether it is a maximum or minimum?

Example

Wages and Age

- Regression Equation

$$\log(wage) = \beta_0 + \beta_{age} age + \beta_{age^2} age^2 + \beta_{educ} educ + \beta_{exper} exper + u$$

- Conditional on education and experience, how does age affect wages?

- Marginal effect of age on education

$$\frac{\partial \log(wage)}{\partial age} = \beta_{age} + 2\beta_{age^2} age$$

- Estimate!

Regressions

Quadratic terms

- Determining wages
- Base model

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$$

- Squared model

$$\begin{aligned}\log(wage) &= \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure \\ &\quad + \beta_4 exper^2 + \beta_5 tenure^2\end{aligned}$$

- How do we interpret? Should we keep in the squared terms?
- $H_0: \ \beta_4 = 0, \beta_5 = 0$

Regressions

Dummy variables

- The basic model
$$y = \beta_0 + \beta_1 x_1 + \beta_2 d + u$$
- How do we interpret β_2 ?
  - Discrete effect of being in a group
- Type of dummy variables
  1. Race, gender
  2. State, County, City
  3. Industry, Year
- An equivalent regression:
$$y = \beta_1 x_1 + \beta_2 d + \beta_3 (1 - d) + u$$
- If you include d and $(1 - d)$ as regressors, you must get rid of β_0!
- Generally, remove β_0 if any combination of regressors can be combined to create a constant.

Regressions

Dummy variables

- Wage regression
- Add in, *urban*.

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 \text{exper} + \beta_3 tenure + \beta_4 urban + u$$

  - $urban = 1$ if the respondent lives in a SMSA, 0 otherwise.

- How do we interpret these results?
- Do they make sense?
- Should I keep in the *urban* dummy?

Regressions

Interaction terms

- The basic model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1 x_2 + \beta_3 x_2 + u$$

- How do we interpret β_1 , β_2?

- Take the derivative w.r.t x_1.

$$\frac{\partial y}{\partial x_1} = \beta_1 + \beta_2 x_2$$

- How do we interpret $\frac{\partial y}{\partial x_1}$?

$\Rightarrow$ If $\beta_2 > 0$, then higher x_2 increases $\frac{\partial y}{\partial x_1}$

$\Rightarrow$ If $\beta_2 < 0$, then higher x_2 decreases $\frac{\partial y}{\partial x_1}$

Regressions

Interaction terms

- Interact *age* with *educ*

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 educ * age + \beta_3 age + \beta_4 exper + u$$

- How do we interpret these results?
- Do they make sense?
- Should we keep in the interaction term?

Regressions

Multiple Slopes

- The basic model

$$y = \beta_0 + \beta_1 x_1 * d + \beta_2 x_1 (1 - d) + \beta_3 d + u$$

- How do we interpret β_1 and β_2?
  - $\beta_1 = \frac{\partial y}{\partial x_1}$ for $d = 1$
  - $\beta_2 = \frac{\partial y}{\partial x_1}$ for $d = 0$
- Test for differences in slopes?
- $\theta = \beta_1 - \beta_2$
- Insert $\beta_1 = \theta + \beta_2$ and simplify:

$$\begin{aligned}
y &= \beta_0 + \beta_1 x_1 * d + \beta_2 x_1 (1 - d) + \beta_3 d + u \\
y &= \beta_0 + (\theta + \beta_2) x_1 d + \beta_2 x_1 (1 - d) + \beta_3 d + u \\
y &= \beta_0 + \theta x_1 d + \beta_2 x_1 (d + 1 - d) + \beta_3 d + u \\
y &= \beta_0 + \theta x_1 d + \beta_2 x_1 + \beta_3 d + u
\end{aligned}$$

Regressions

Multiple Slopes

- Wage regression

- Are the returns to education different for urban residents?

- Interact *urban* and educ.

$$\begin{aligned}\log(wage) &= \beta_0 + \beta_1 educ + \beta_2 \text{exper} + \beta_3 tenure \\ &+\beta_4 urban + \beta_4 urban{*}educ + u\end{aligned}$$

- Returns to education

$$\frac{\partial \widehat{\log(wage)}}{\partial educ} = \widehat{\beta}_1 + \widehat{\beta}_4 urban$$

- How do we interpret these results?

- Do they make sense?

Endogeneity

Regression Discontinuity

- Use a policy discontinuity that is quasi-exogenous
- Equation of interest:

$$y = \beta_0 + \beta_1 \cdot (x - d) + \beta_2 \cdot I(x > d) + \beta_3 \cdot I(x > d) \cdot (x - d)$$

    - y is the outcome
    - x is a "continuous" input
    - $I(x > d)$ is an dummy variable identifying if x is above some threshold.
- Coefficient of interest is β_2
- Example: Dobkin, Gil, and Marion (2010)

$$FinalExam = \beta_0 + \beta_1 \cdot MT + \beta_2 \cdot I(MT > d) + \beta_3 \cdot I(MT > d) \cdot MT$$

- Attendance was made compulsory for MT scores below d.
- What was the effect of compulsory attendance on exam scores?

Figure 4: Class Attendance Over Support of Midterm Score

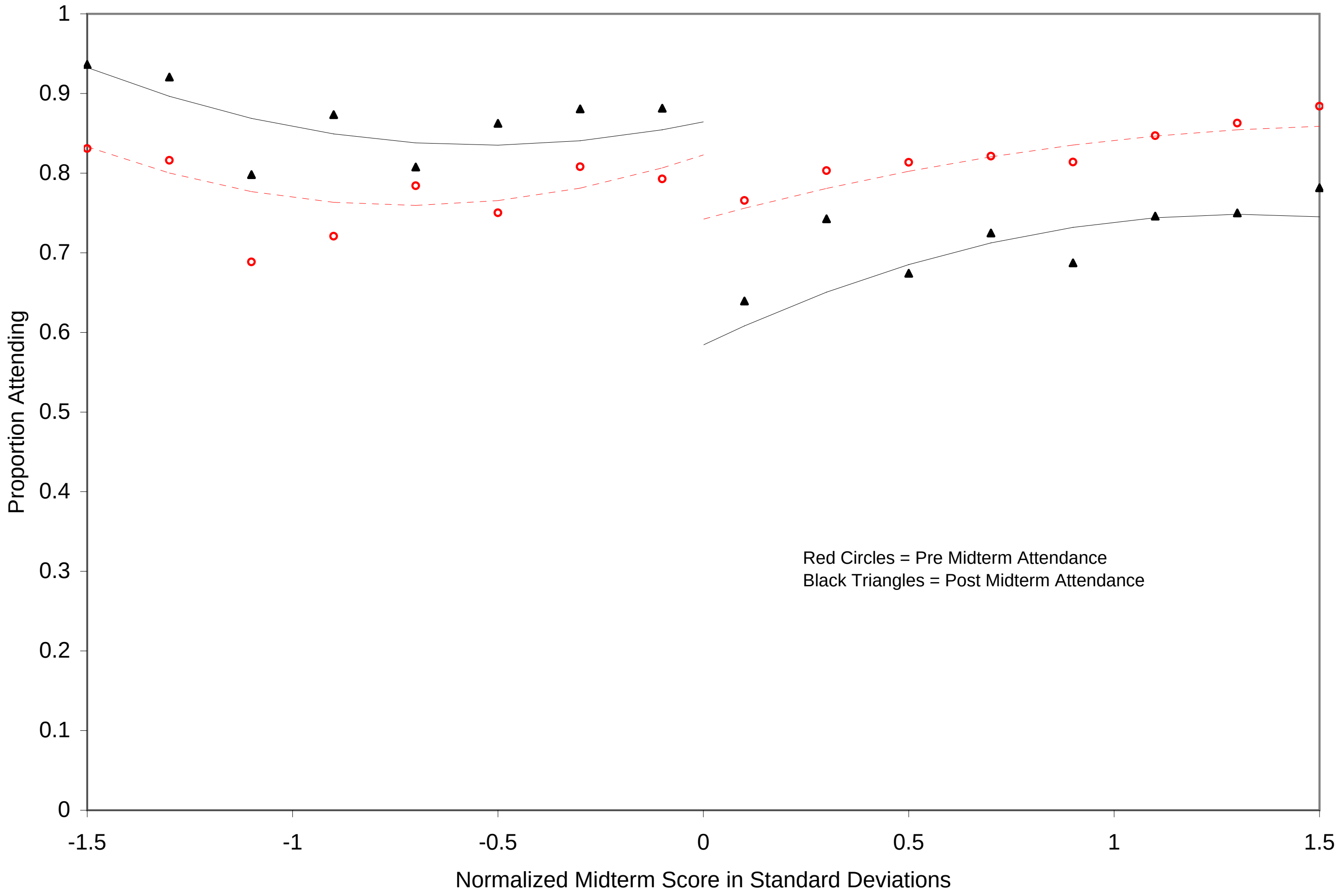

Figure 5: Final Exam Over Support of Midterm Score

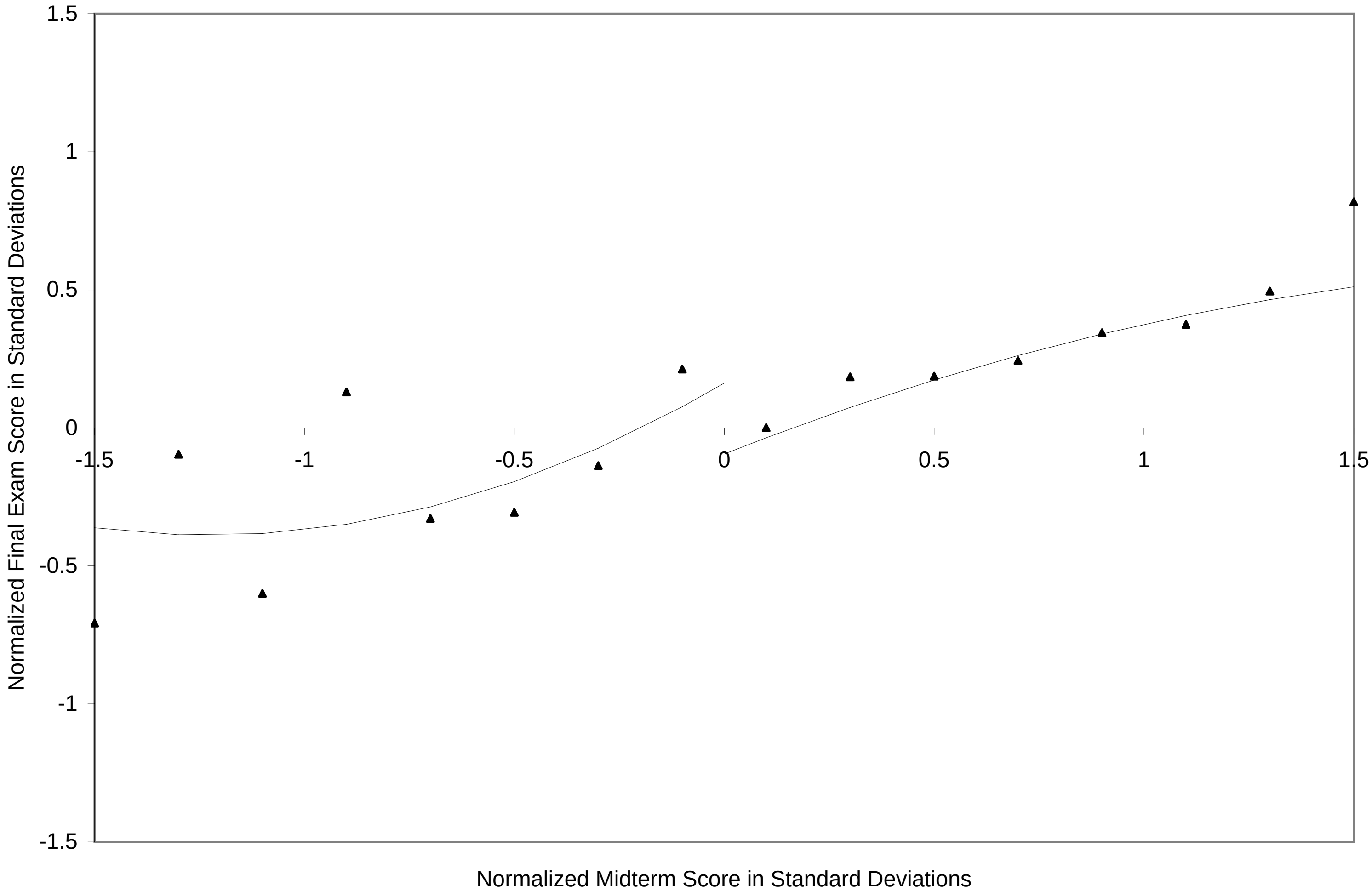

Table 8: Determinants of Attendance Across Classes

Dependent variable: Class attendance rate			
Male	-0.018	(0.026)	F-value 1.42
Instructor age	-0.001	(0.001)	p-value 0.22
Not native english	-0.003	(0.027)	
Instr. appearance	0.026	(0.016)	
Poor writing	-0.028	(0.038)	
Call on people in class	0.018	(0.022)	F-value 0.53
Power point	-0.007	(0.029)	p-value 0.75
Chalk board	-0.023	(0.026)	
Overhead projector	0.011	(0.032)	
Camera projector	-0.042	(0.036)	
Lecture available online	-0.045	(0.026)*	F-value 2.09
Lecture follows book	-0.028	(0.032)	p-value 0.13
Class mandatory	0.071	(0.032)**	
Attendance monitored	0.115	(0.034)***	
Week 3	0.003	(0.043)	
Week 4	-0.008	(0.044)	
Week 5	-0.062	(0.041)	
Week 6	-0.073	(0.039)*	
Week 7	-0.097	(0.040)**	
Week 8	-0.068	(0.040)*	
Week 9	-0.120	(0.043)***	
Quizzes in section	-0.131	(0.047)***	
Quizzes held in class	-0.083	(0.048)*	
Quizzes online	-0.167	(0.093)*	
Class size (X100)	-0.035	(0.016)**	
Upper division	-0.026	(0.027)	
Two meetings/wk	0.015	(0.084)	
Three meetings/wk	0.038	(0.084)	
Start time 8-10	0.020	(0.042)	
Start time 11-5	0.071	(0.041)*	
Monday	0.137	(0.048)***	
Tuesday	0.097	(0.059)*	
Wednesday	0.100	(0.036)***	

Discrete Dependent Variables

Linear Probability

- "Linear probability model": y takes on a value of 1 or 0
- The model:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$$

- Note that:

$$\Pr(y = 1|\mathbf{x}) = E(y|\mathbf{x})$$

- The estimates:

$$\Pr(y = 1|\mathbf{x}) = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \cdots + \widehat{\beta}_k x_k$$

- Are there conditions on $\widehat{y}$?

$\longrightarrow$ No! $\widehat{y}$ could be negative.

$\longrightarrow$ Predicted probabilities could be negative

Discrete Dependent Variables

Linear Probability

- Estimate

$$Smoke = \beta_0 + \beta_1 faminc + \beta_2 meduc + \beta_3 feduc + u$$

  - Smoke: Whether a mother smokes
  - Faminc: Family Income
  - meduc: mother's education
  - feduc: father's education

- Do the estimates make sense?
- Are any predictions negative?