PSC 400
SYRACUSE UNIVERSITY

DATA ANALYTICS FOR POLITICAL SCIENCE

EXTENSIONS TO REGRESSION

ASSIGNMENT

- Data Analysis Memo 3 due on Friday

CATEGORICAL VARIABLE

- cces19.csv
- DV: Registered to vote (voters)
  - 1 if registered, 0 if not
- IV: Gender (female)
  - 1 if female, 0 if male
  - categorical independent variable

CATEGORICAL VARIABLE

- Pr(Registered) = 0.94 - 0.035 * female
  - What is the predicted probability that a woman is registered?

CATEGORICAL VARIABLE

- Pr(Registered) = 0.94 - 0.035 * female
  - What is the predicted probability that a woman is registered?
  - 0.94 - 0.035 * 1 = 0.905

CATEGORICAL VARIABLE

- Pr(Registered) = 0.94 - 0.035 * female
  - What is the predicted probability that a man is registered?

CATEGORICAL VARIABLE

- Pr(Registered) = 0.94 - 0.035 * female
  - What is the predicted probability that a woman is registered?
  - 0.94 - 0.035 * 0 = 0.94

CATEGORICAL VARIABLE

- DV: Registered to vote or not
- IV: Partisanship
  - Democrat, Republican, Independent
  - Categorical variable

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that a Republican is registered?

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that a Republican is registered?
  - 0.9425 - 0.0617 * 0 - 0.0004 * 1 = 0.9421

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that an Independent is registered?

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that an Independent is registered?
  - 0.9425 - 0.0617 * 1 - 0.0004 * 0 = 0.8808

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that a Democrat is registered?

CATEGORICAL VARIABLE

- Pr(Registered) = 0.9425 - 0.0617 * Independent - 0.0004 * Republican
  - What is the predicted probability that a Democrat is registered?
  - 0.9425 - 0.0617 * 0 - 0.0004 * 0 = 0.9425

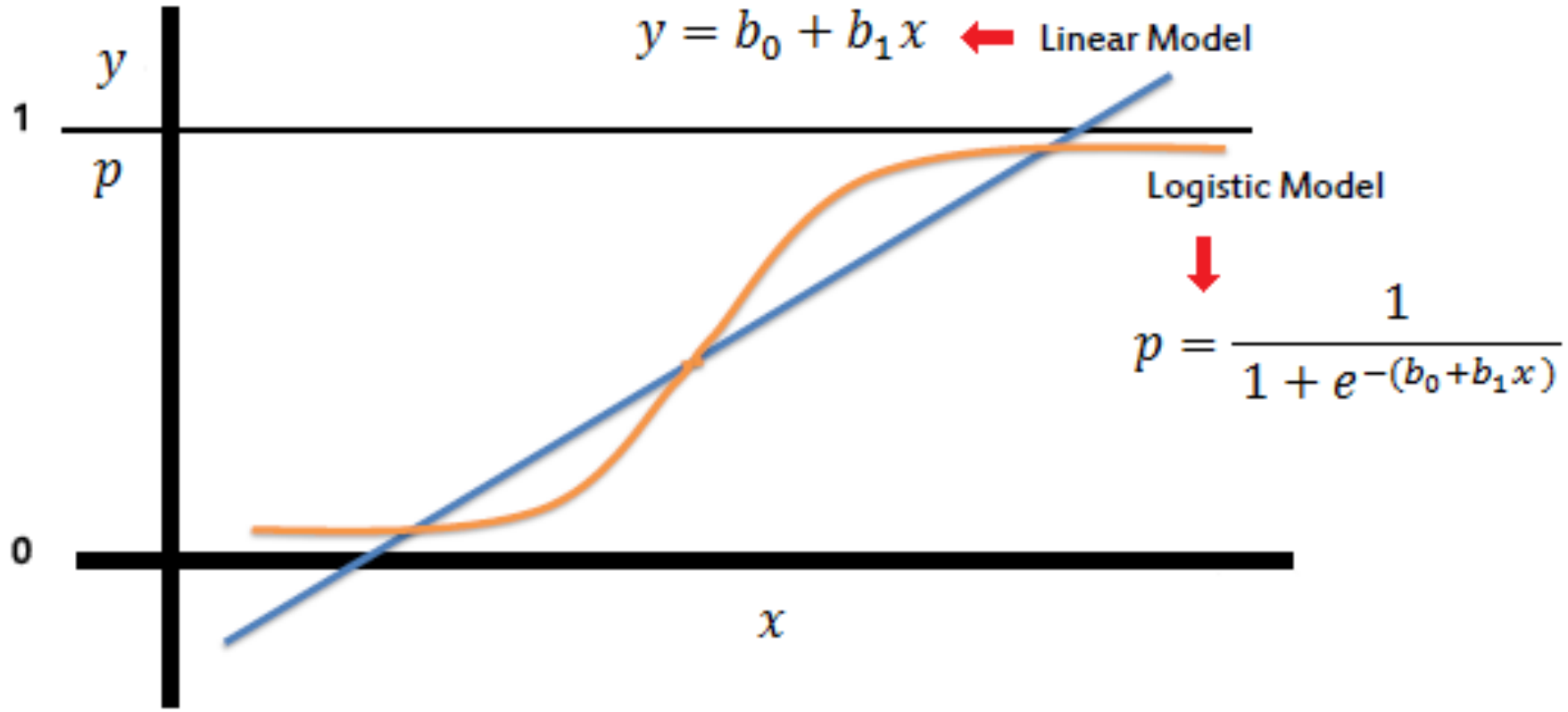
CATEGORICAL VARIABLE

- Key insight: If a categorical variable has x categories, the regression will estimate (x-1) regression coefficients
- Category that is left out: baseline category
  - The other categories give the effect of being in a certain category relative to the baseline
    - e.g. if baseline: male
    - then coefficient gives effect of being female vs. being male

EXAMPLE

- What is the effect of age on whether respondents are registered to vote or not?
  - Control for: female, partisanship, nohighschool, collegeorhigher, nonwhite, married, employed

LINEAR VS. LOGIT



EXPERIMENT ANALYSIS

Dear Registered Voter:

WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach.
We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

DO YOUR CIVIC DUTY – VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	Voted	Voted	_____
995 JENNIFER KAY SMITH		Voted	_____
9997 RICHARD B JACKSON		Voted	_____
9999 KATHY MARIE JACKSON		Voted	_____

EXPERIMENT ANALYSIS

- social.csv
  - primary2006: 1 if voted, 0 if abstained
  - neighbors: 1 if received treatment, 0 if not
  - age: voter age in years
- We go back to OLS
  - What is the treatment effect?