What Is Specification?

- Process of selecting a population model
  - Assume population model is true, then estimate parameters.
- Choosing which variables to include, which to omit
- Choosing what transformation to apply to each variable, or what form they should appear in

Goals When Specifying a Model

- Be realistic.
  - Must test assumptions, or justify where appropriate.
- Be understandable.
  - Have an intuitive understanding of what our parameters mean.
- Make accurate predictions.
  - There's usually a trade-off between prediction and understandability.
- Highlight a specific relationship, test a specific hypothesis, measure a specific effect, etc.

Specifying a Linear Model

- Linear population model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + u$
- Flexibility within linearity assumption
  - Design x and y.
  - Transform a variable, or create entirely new ones from existing variables.
  - Include same variable in multiple forms, modeling data with parabolas or other curves.
  - Allow the value of one variable to alter the effect of another.
- Assumption of linearity only about how terms in regression combine
  - Cannot raise x_j to the power of β_j
  - βs: only able to scale terms we're adding together