Large-Sample Properties

- Let's take a few minutes to focus on the asymptotic properties of OLS.
  - These are often called large-sample properties.
  - Many of you will be working with huge datasets, and it's good to summarize how these work.
- We've listed a lot of assumptions that look daunting.
  - If we have a large sample size, we don't need many of the stronger assumptions in the classical model.
  - As long as we have a large sample and use heteroskedasticity-robust standard errors, we generally focus on MLR.1–3 and MLR.4'.

Crucial Assumptions for Large Samples

1. **Assumption MLR.1 (linear in parameters)**
   $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$
2. **Assumption MLR.2 (random sampling)**
   $\{(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i) : i = , \ldots n\}$
   Data points are independent draws from population
3. **Assumption MLR.3 (no perfect collinearity)**
4. **Assumption MLR.4' (exogeneity)**
   $\mathrm{Cov}(x_j, u) = 0$ for all j.

OLS Consistency

- We already know that under MLR.1–3 and MLR.4', OLS estimators are consistent.
  - $\text{plim}_{n \rightarrow} (\hat{\beta}_j) = \beta_j$
- This means that we can always get the right answer if we collect an infinite number of data points.

Asymptotic Normality

- What about the shape of the distribution?
  - The central limit theorem tells us that our coefficients have an asymptotically normal sampling distribution.
  - The proof is tough, so we'll skip it here, but there's a sketch in an appendix of Wooldridge. (The theorem is stated under MLR.1–MLR.5.)
- **Theorem 5.2 (Asymptotic normality of OLS):** $\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \overset{a}{\sim} N(0, 1)$ also $\text{plim}\hat{\sigma}^2 = \sigma^2$
- Note that if you use heteroskedasticity-robust standard errors, you can drop MLR.5.

Asymptotic Normality (cont.)

What does this mean from a practical standpoint?

- As n increases, sampling distributions become normal.
  - We can't see this because we only get one sample, but the math tells us it's happening.
- Since the t-distribution is asymptotically normal, it doesn't matter if we use a normal or t-distribution in stating our theorem.
  - This means that t-tests are valid for large samples; the same is true for confidence intervals and F-tests.
- For large samples, there are two key assumptions to focus on.
  - Random sampling: Are the observations correlated in some way, and is there clustering or a time dimension?
  - Exogeneity: Is any x correlated with the error, and is there some unmeasured factor that ends up in the error that's related to an x?

Most of the rest of this course is about what to do when we can't meet the key assumptions we just discussed.