EC200: Econometrics and Applications

Empirical Specifications

A helpful, five-step template![1]

(1) First, I state what in broad terms type of model I am using (probably OLS!) and what I am going to do with it. If I'm using a fancy speciifcation like a difference-in-differences or panel data, now is a good time to mention this.

> I use ordinary least squares to estimate the relationship between ice cream flavor consumed and risk aversion, using state-level fixed effects to control for time-invariant determinants of risk aversion that may vary by state.

(2) Then, I'm going to write the population model I am estimating in equation form. I'm going to use appropriate subscripts:

$$outcome_{is} = \beta_0 + \beta_1 VariableName_{is} + \beta_2 AnotherVariableName_{2,is} + f_s + u_{is}$$

- Use descriptive names (these don't actually have to be your variable names)
- If you use fixed effects, you can notate these without having to expand them:
  - adding state fixed effects could be written as $ ... + f_s + ...$, for example
  - or, you could add them this way: $\sum_{s=1}^{50} D_s$ for a set of 50 state-level fixed effects (with one omitted)
- List covariates you include. If you have a ton of these, you could include a vector of individual-specific covariates ($X'\gamma$, for example) and then list them in paragraph form. However, this isn't likely for most papers.

(3) Then, I'm going to define my independent and dependent variables in a list or paragraph form. I'm going to also describe what the subscripts are

> Where $outcome_{is}$ is measured risk aversion for individual i living in state s, following the scale described above. $VariableName_{is}$ is a measure of flavor intensity, standardized around Edy's French Vanilla ...

I'll want to include any controls and fixed effects.

(4) I'll mention any special things I do when coding (missing value flags, etc), and what types of standard errors I'm using

(5) Now, I write how I will interpret my coefficient of interest (causal, correlational, etc.) - what will the coefficients on my key indicator variables tell me? Are there any key identifying assumptions at play?
If you are using multiple specifications, then you have two options:

- If it's a matter of adding additional controls, then mention that you will also add them in a second model in part (3).
- If it's a fundamentally different population model, then include a second model, and define any terms that were not previously defined. Make sure to discuss how the interpretation of results would differ.

[1] Use the template, but please do not plagarize the template.