Complex Modeling Issues

A thorough analysis of the Current Population Survey requires several additional tools that are useful in other general modeling situations, such as with nonlinear modeling or discrete data.

The tools demonstrated in this chapter include confounding, an important concern in any observational study, multi-colinearity, pooled tests used to test the significance of several variables simultaneously, categorical ("dummy" variables) used to describe discrete independent variables, and finally, interactions, which deserve a separate chapter.

Confounding

Confounding is the masking of the effect of one variable through its association with another variable. An example is the effect of work experience on wages. When controlling for gender alone, the effect of experience on wages remains the same since experience and gender are uncorrelated (r= 0.04). However, when controlling for education, the effect of experience increases from a factor of 1.09 (1.05, 1.12) to 1.14 (1.10, 1.17) per 10 years of experience, a 57% increase in effect. Education and experience are negatively correlated (r=-.186). Since people with more work experience tend to have less education, ceteris paribus, some of the effect of work experience is explained by lower levels of education. When both variables are in the model, the effect of work experience is "corrected" for education in that it measures the effect of experience <u>for constant education levels</u>. In statistical terms, including confounders in a model removes their effect from the independent variable of interest to obtain an unbiased measure of effect of that variable.
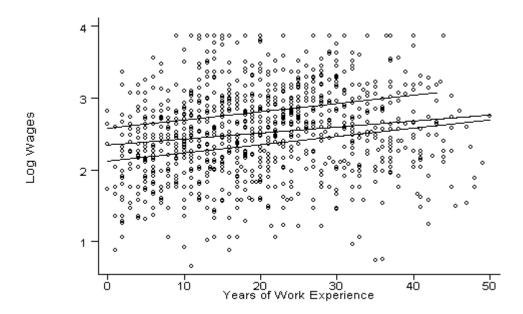
Figure 7.1 Plot of log hourly wages against years of work experience, superimposing fitted regression lines, not controlling (middle line) and controlling for education (lower line, <= 13 years; uppermost line, >13 years).
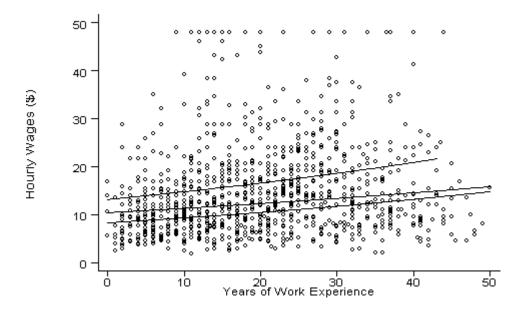

Figure 7.2 Plot of hourly wages against years of work experience, superimposing back-transformed log regression lines, not controlling (middle line) and controlling for education (lower line, <= 13 years; uppermost line, >13 years).