Empirical Exercise:

1. On the text website, http://www.pearsonhighered.com/stock_watson/, you will find the data file **Birthweight_Smoking**, which contains data for a random sample of babies born in Pennsylvania in 1989. The data include the baby's birth weight together with various characteristics of the mother, including whether she smoked during the pregnancy. A detailed description is given in **Birthweight_Smoking_Description**, also available on the website. In this exercise you will investigate the relationship between birth weight and smoking during pregnancy. To begin, run three regressions:

   (1) *Birthweight* on *Smoker*

   (2) *Birthweight* on *Smoker*, *Alcohol*, and *Nprevist*

   (3) *Birthweight* on *Smoker*, *Alcohol*, *Nprevist*, and *Unmarried*

   (a) What is the value of the estimated effect of smoking on birth weight in each of the regressions?

   (b) Construct a 95% confidence interval for the effect of smoking on birth weight, using each of the regressions.

   (c) Does the coefficient on *Smoker* in regression (1) suffer from omitted variable bias? Explain.

   (d) Does the coefficient on *Smoker* in regression (2) suffer from omitted variable bias? Explain.

   (e) Consider the coefficient on *Unmarried* in regression (3).

      i. Construct a 95% confidence interval for the coefficient.

      ii. Is the coefficient statistically significant? Explain.

      iii. Is the magnitude of the coefficient large? Explain.

      iv. A family advocacy group notes that the large coefficient suggests that public policies that encourage marriage will lead, on average, to healthier babies. Do you agree? (*Hint:* Discuss some of the various factors that *Unmarried* may be controlling for and how this affects the interpretation of its coefficient.)

   (f) Consider the various other control variables in the data set. Which do you think should be included in the regression? Using a table like the one in question 5 above, examine the robustness of the confidence interval you constructed in (b). What is a reasonable 95% confidence interval for the effect of smoking on birth weight?