1. Exercise 12.2.

2. Exercise 12.8.

3. Exercise 12.10.

4. E12.2.

5. Consider the following model to estimate the effects of several variables, including cigaratte smoking, on the weight of newborns:

$$\ln(bwght)_i = \beta_0 + \beta_1 \, packs_i + \beta_2 \, male_i + \beta_3 \, parity_i + \beta_4 \, \ln(faminc)_i + u_i \quad (1)$$

where *male* is a binary indicator equal to one if the child is male; *parity* is the birth order of this child; *faminc* is family income; and *packs* is the average number of packs of cigarattes smoked per day during pregnancy.

   (a) Why might we expect *packs* to be correlated with u?

   (b) Suppose that you have data on average cigaratte price in each womam's state of residence in the United States. Discuss whether this information is likely to satisfy the properties of a good instrumental variables for *packs*?

   (c) Use the data in **bwght.dta** to estimate equation (1). First, run an OLS estimation. Then, run the 2SLS estimation **without** correcting the standard errors in the second stage, where *cigprice* is an instrument for packs. Discuss any important differences in the OLS and 2SLS estimates.

   (d) Do the 2SLS estimation by the stata command **ivreg**, compare the standard errors with those from 2SLS in (c). Where does the difference in standard error come from?