**12.10** Consider the instrumental variable regression model $Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$, where $Z_i$ is an instrument. Suppose that data on $W_i$ are not available and the model is estimated omitting $W_i$ from the regression.

a. Suppose $Z_i$ and $W_i$ are uncorrelated. Is the IV estimator consistent?

b. Suppose $Z_i$ and $W_i$ are correlated. Is the IV estimator consistent?

## Empirical Exercises

**E12.1** During the 1880s, a cartel known as the Joint Executive Committee (JEC) controlled the rail transport of grain from the Midwest to eastern cities in the United States. The cartel preceded the Sherman Antitrust Act of 1890, and it legally operated to increase the price of grain above what would have been the competitive price. From time to time, cheating by members of the cartel brought about a temporary collapse of the collusive price-setting agreement. In this exercise, you will use variations in supply associated with the cartel's collapses to estimate the elasticity of demand for rail transport of grain. On the textbook Web site **www.aw-bc.com/stock_watson**, you will find a data file **JEC** that contains weekly observations on the rail shipping price and other factors from 1880 to 1886.[4] A detailed description of the data is contained in **JEC_Description** available on the Web site.

Suppose that the demand curve for rail transport of grain is specified as $\ln(Q_i) = \beta_0 + \beta_1 \ln(P_i) + \beta_2 Ice_i + \sum_{j=1}^{12} \beta_{2+j} Seas_{j,i} + u_i$, where $Q_i$ is the total tonnage of grain shipped in week $i$, $P_i$ is the price of shipping a ton of grain by rail, $Ice_i$ is a binary variable that is equal to 1 if the Great Lakes are not navigable because of ice, and $Seas_j$ is a binary variable that captures seasonal variation in demand. *Ice* is included because grain could also be transported by ship when the Great Lakes were navigable.

a. Estimate the demand equation by OLS. What is the estimated value of the demand elasticity and its standard error?

b. Explain why the interaction of supply and demand could make the OLS estimator of the elasticity biased.

c. Consider using the variable *cartel* as instrumental variable for $\ln(P)$. Use economic reasoning to argue whether *cartel* plausibly satisfies the two conditions for a valid instrument.

[4]These data were provided by Professor Robert Porter of Northwestern University and were used in his paper "A Study of Cartel Stability: The Joint Executive Committee, 1880–1886," *The Bell Journal of Economics* 1983; 14(2): 301–314.

d. Estimate the first-stage regression. Is *cartel* a weak instrument?

e. Estimate the demand equation by instrumental variable regression. What is the estimated demand elasticity and its standard error?

f. Does the evidence suggest that the cartel was charging the profit-maximizing monopoly price? Explain. (*Hint:* What should a monopolist do if the price elasticity is less than 1?)

**E12.2** How does fertility affect labor supply? That is, how much does a woman's labor supply fall when she has an additional child? In this exercise you will estimate this effect using data for married women from the 1980 U.S. Census.[5] The data are available on the textbook Web site **www.aw-bc.com/stock_watson** in the file **Fertility** and described in the file **Fertility_Description**. The data set contains information on married women aged 21–35 with two or more children.

a. Regress *weeksworked* on the indicator variable *morekids* using OLS. On average, do women with more than two children work less than women with two children? How much less?

b. Explain why the OLS regression estimated in (a) is inappropriate for estimating the causal effect of fertility (*morekids*) on labor supply (*weeksworked*).

c. The data set contains the variable *samesex*, which is equal to 1 if the first two children are of the same sex (boy–boy or girl–girl) and equal to 0 otherwise. Are couples whose first two children are of the same sex more likely to have a third child? Is the effect large? Is it statistically significant?

d. Explain why *samesex* is a valid instrument for the instrumental variable regression of *weeksworked* on *morekids*.

e. Is *samesex* a weak instrument?

f. Estimate the regression of *weeksworked* on *morekids* using *samesex* as an instrument. How large is the fertility effect on labor supply?

g. Do the results change when you include the variables *agem1*, *black*, *hispan*, and *othrace* in the labor supply regression (treating these variable as exogenous)? Explain why or why not.

[5]These data were provided by Professor William Evans of the University of Maryland and were used in his paper with Joshua Angrist, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review* 1998, 88(3): 450–477.

**E12.3** (This requires Appendix 12.5) On the textbook Web site **www.aw-bc.com/stock_watson** you will find the data set **WeakInstrument** that contains 200 observations on $(Y_i, X_i, Z_i)$ for the instrumental regression $Y_i = \beta_0 + \beta_1 X_i + u_i$.

a. Construct $\hat{\beta}_1^{TSLS}$, its standard error, and the usual 95% confidence interval for $\beta_1$.

b. Compute the $F$-statstic for the regression of $X_i$ on $Z_i$. Is there evidence of a "weak instrument" problem?

c. Compute a 95% confidence interval for $\beta_1$ using the Anderson-Rubin procedure. (To implement the procedure, assume that $-5 \leq \beta_1 \leq 5$.)

d. Comment on the differences in the confidence intervals in (a) and (c). Which is more reliable?

## APPENDIX 12.1 The Cigarette Consumption Panel Data Set

The data set consists of annual data for the 48 continental U.S. states from 1985 to 1995. Quantity consumed is measured by annual per capita cigarette sales in packs per fiscal year, as derived from state tax collection data. The price is the real (that is, inflation-adjusted) average retail cigarette price per pack during the fiscal year, including taxes. Income is real per capita income. The general sales tax is the average tax, in cents per pack, due to the broad-based state sales tax applied to all consumption goods. The cigarette-specific tax is the tax applied to cigarettes only. All prices, income, and taxes used in the regressions in this chapter are deflated by the Consumer Price Index and thus are in constant (real) dollars. We are grateful to Professor Jonathan Gruber of MIT for providing us with these data.

## APPENDIX 12.2 Derivation of the Formula for the TSLS Estimator in Equation (12.4)

The first stage of TSLS is to regress $X_i$ on the instrument $Z_i$ by OLS, and to compute the OLS predicted value $\hat{X}_i$, and the second stage is to regress $Y_i$ on $\hat{X}_i$ by OLS. Accordingly,

**d.** Suppose that outliers are rare, so that $(u_i, X_i)$ have finite fourth moments. Is it appropriate to use OLS and the methods of Chapters 4 and 5 to estimate and carry out inference about the average values of $\beta_{0i}$ and $\beta_{1i}$?

**e.** Suppose that $\beta_{1i}$ and $X_i$ are positively correlated, so that observations with larger than average values of $X_i$ tend to have larger than average values of $\beta_{1i}$. Are the assumptions in Key Concept 4.3 satisfied? If not, which assumption(s) is (are) violated? Is it appropriate to use OLS and the methods of Chapters 4 and 5 to estimate and carry out inference about the average value of $\beta_{0i}$ and $\beta_{1i}$?

**13.11** In Chapter 12, state-level panel data were used to estimate the price elasticity of demand for cigarettes, using the state sales tax as an instrumental variable. Consider in particular regression (1) in Table 12.1. In this case, in your judgment does the local average treatment effect differ from the average treatment effect? Explain.

## Empirical Exercises

**E13.1** A prospective employer receives two resumes: a resume from a white job applicant and a similar resume from an African American applicant. Is the employer more likely to call back the white applicant to arrange an interview? Marianne Bertrand and Sendhil Mullainathan carried out a randomized controlled experiment to answer this question. Because race is not typically included on a resume, they differentiated resumes on the basis of "white-sounding names" (such as Emily Walsh or Gregory Baker) and "African American–sounding names" (such as Lakisha Washington or Jamal Jones). A large collection of fictitious resumes was created, and the presupposed "race" (based on the "sound" of the name) was randomly assigned to each resume. These resumes were sent to prospective employers to see which resumes generated a phone call (a "call back") from the prospective employer. Data from the experiment and a detailed data description are on the textbook Web site **http://www.aw-bc.com/stock_watson** in the files **Names** and **Names_Description**.[6]

**a.** Define the "call-back rate" as the fraction of resumes that generate a phone call from the prospective employer. What was the call-back rate

[6]These data were provided by Professor Marianne Bertrand of the University of Chicago and were used in her paper with Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, 94(4).

for whites? For African Americans? Construct a 95% confidence interval for the difference in the call-back rates. Is the difference statistically significant? Is it large in a real-world sense?

**b.** Is the African American/white call-back rate differential different for men than for women?

**c.** What is the difference in call-back rates for high-quality versus low-quality resumes? What is the high quality/low quality difference for white applicants? For African American applicants? Is there a significant difference in this high quality/low quality difference for whites versus African Americans?

**d.** The authors of the study claim that race was assigned randomly to the resumes. Is there any evidence of nonrandom assignment?

**E13.2** A consumer is given the chance to buy a baseball card for \$1, but he declines the trade. If the consumer is now given the baseball card, will he be willing to sell it for \$1? Standard consumer theory suggests yes, but behavioral economists have found that "ownership" tends to increase the value of goods to consumers. That is, the consumer may hold out for some amount more than \$1 (for example, \$1.20) when selling the card, even though he was willing to pay only some amount less than \$1 (for example, \$0.88) when buying it. Behavioral economists call this phenomenon the "endowment effect." John List investigated the endowment effect in a randomized experiment involving sports memorabilia traders at a sports-card show. Traders were randomly given one of two sports collectibles, say good A or good B, that had approximately equal market value.[7] Those receiving good A were then given the option of trading good A for good B with the experimenter; those receiving good B were given the option of trading good B for good A with the experimenter. Data from the experiment and a detailed description can be found on the textbook Web site **http://www.aw-bc.com/stock_watson** in the files **Sportscards** and **Sportscards_Description**.[8]

**a.** i. Suppose that, absent any endowment effect, all of the subjects prefer good A to good B. What fraction of the experiment's subjects would you expect to trade the good that they were given for the

[7]Good A was a ticket stub from the game that Cal Ripken, Jr., set the record for consecutive games played, and Good B was a souvenir from the game that Nolan Ryan won his 300th game.

[8]These data were provided by Professor John List of the University of Chicago and were used in his paper "Does Market Experience Eliminate Market Anomalies," *Quarterly Journal of Economics*, 2003: 118(1): 41–71.

other good? (*Hint:* Random assignment means that approximately 50% of the subjects received good A and 50% received good B.)

ii. Suppose that, absent any endowment effect, 50% of the subjects prefer good A to good B, and the other 50% prefer good B to good A. What fraction of the subjects would you expect to trade the good that they were given for the other good?

iii. Suppose that, absent any endowment effect, $X$% of the subjects prefer good A to good B, and the other $(1 - X)$% prefer good B to good A. Show that you would expect 50% of the subjects to trade the good that they were given for the other good.

b. Using the sports-card data, what fraction of the subjects traded the good they were given? Is the fraction significantly different from 50%? What fraction of the subjects who received good A traded for good B? What fraction of the subjects who received good B traded for good A? Is there evidence of an endowment effect?

c. Some have argued that the endowment effect may be present, but that it is likely to disappear as traders gain more trading experience. Half of the experimental subjects were dealers and the other half were nondealers. Dealers have more experience than nondealers. Repeat (b) for dealers and nondealers. Is there a significant difference in their behavior? Is the evidence consistent with the hypothesis that the endowment effect disappears as traders gain more experience?

d. The data set contains two additional measures of experience: number of trades per month and number of years trading. Is there evidence that for nondealers the endowment effect decreases as their trading experience increases?

## APPENDIX 13.1 The Project Star Data Set

The Project STAR public access data set contains data on test scores, treatment groups, and student and teacher characteristics for the four years of the experiment, from academic year 1985–1986 to academic year 1988–1989. The test score data analyzed in this chapter are the sum of the scores on the math and reading portions of the Stanford Achievement Test. The