Introduction to Econometrics (3rd Updated Edition)

by

James H. Stock and Mark W. Watson

Solutions to Odd-Numbered End-of-Chapter Exercises: Chapter 8

(This version August 17, 2014)

8.1. (a) The percentage increase in sales is $100\times\frac{198-196}{196}=1.0204\%$. The approximation is 100 × [ln (198) − ln (196)] = 1.0152%.

(b) When $Sales_{2010}=205$, the percentage increase is $100\times\frac{205-196}{196}=4.5918\%$ and the approximation is 100 × [ln (205) − ln (196)] = 4.4895%. When $Sales_{2010}=250$, the percentage increase is $100\times\frac{250-196}{196}=27.551\%$ and the approximation is 100 × [ln (250) − ln (196)] = 24.335%. When $Sales_{2010}=500$, the percentage increase is $100\times\frac{500-196}{196}=155.1\%$ and the approximation is 100 × [ln (500) − ln (196)] = 93.649%.

(c) The approximation works well when the change is small. The quality of the approximation deteriorates as the percentage change increases.

8.3. (a) The regression functions for hypothetical values of the regression coefficients that are consistent with the educator's statement are: $\beta_1 > 0$ and $\beta_2 < 0$. When $TestScore$ is plotted against STR the regression will show three horizontal segments. The first segment will be for values of $STR < 20$; the next segment for $20 \leq STR \leq 25$; the final segment for $STR > 25$. The first segment will be higher than the second, and the second segment will be higher than the third.

(b) It happens because of perfect multicollinearity. With all three class size binary variables included in the regression, it is impossible to compute the OLS estimates because the intercept is a perfect linear function of the three class size regressors.

8.5. (a) (1) The demand for older journals is less elastic than for younger journals because the interaction term between the log of journal age and price per citation is positive. (2) There is a linear relationship between log price and log of quantity follows because the estimated coefficients on log price squared and log price cubed are both insignificant. (3) The demand is greater for journals with more characters follows from the positive and statistically significant coefficient estimate on the log of characters.

(b) (i) The effect of $\ln(\textit{Price per citation})$ is given by $[-0.899 + 0.141 \times \ln(\textit{Age})] \times \ln(\textit{Price per citation})$. Using $\textit{Age} = 80$, the elasticity is $[-0.899 + 0.141 \times \ln(80)] = -0.28$.

(ii) As described in equation (8.8) and the footnote on page 263, the standard error can be found by dividing 0.28, the absolute value of the estimate, by the square root of the F-statistic testing $\beta_{\ln(\textit{Price per citation})} + \ln(80) \times \beta_{\ln(\textit{Age}) \times \ln(\textit{Price per citation})} = 0$.

(c) $\ln\left(\frac{\textit{Characters}}{a}\right) = \ln(\textit{Characters}) - \ln(a)$ for any constant a. Thus, estimated parameter on *Characters* will not change and the constant (intercept) will change.

8.7. (a) (i) ln(*Earnings*) for females are, on average, 0.44 lower for men than for women.

(ii) The error term has a standard deviation of 2.65 (measured in log-points).

(iii) Yes. But the regression does not control for many factors (size of firm, industry, profitability, experience and so forth).

(iv) No. In isolation, these results do not imply gender discrimination. Gender discrimination means that two workers, identical in every way but gender, are paid different wages. Thus, it is also important to control for characteristics of the workers that may affect their productivity (education, years of experience, etc.) If these characteristics are systematically different between men and women, then they may be responsible for the difference in mean wages. (If this were true, it would raise an interesting and important question of why women tend to have less education or less experience than men, but that is a question about something other than gender discrimination.) These are potentially important omitted variables in the regression that will lead to bias in the OLS coefficient estimator for *Female*. Since these characteristics were not controlled for in the statistical analysis, it is premature to reach a conclusion about gender discrimination.

(b) (i) If *MarketValue* increases by 1%, earnings increase by 0.37%

(ii) *Female* is correlated with the two new included variables and at least one of the variables is important for explaining ln(*Earnings*). Thus the regression in part (a) suffered from omitted variable bias.

(c) Forgetting about the effect or *Return*, whose effects seems small and statistically insignificant, the omitted variable bias formula (see equation (6.1)) suggests that *Female* is negatively correlated with ln(*MarketValue*).

8.9. Note that

$$\begin{aligned} Y &= \beta_0 + \beta_1 X + \beta_2 X^2 \\ &= \beta_0 + (\beta_1 + 21\beta_2)X + \beta_2(X^2 - 21X). \end{aligned}$$

Define a new independent variable $Z = X^2 - 21X$, and estimate

$$Y = \beta_0 + \gamma X + \beta_2 Z + u_i.$$

The confidence interval is $\hat{\gamma} \pm 1.96 \times \text{SE}(\hat{\gamma})$.

8.11. Linear model: $E(Y \mid X) = \beta_0 + \beta_1 X$, so that $\frac{dE(Y|X)}{dX} = \beta_1$ and the elasticity is

$$\beta_1 \frac{X}{E(Y|X)} = \frac{\beta_1 X}{\beta_0 + \beta_1 X}$$

Log-Log Model: $E(Y \mid X) = E\left(e^{\beta_0 + \beta_1 \ln(X) + u} \mid X\right) = e^{\beta_0 + \beta_1 \ln(X)} E(e^u \mid X) = ce^{\beta_0 + \beta_1 \ln(X)}$,

where $c = E(e^u \mid X)$, which does not depend on X because u and X are assumed to be independent.

Thus $\frac{dE(Y|X)}{dX} = \frac{\beta_1}{X} ce^{\beta_0 + \beta_1 \ln(X)} = \beta_1 \frac{E(Y|X)}{X}$ and the elasticity is β_1.