

Introduction to Econometrics (4th Edition)

by

James H. Stock and Mark W. Watson

**Solutions to Odd-Numbered End-of-Chapter Exercises:
Chapter 10**

(This version September 18, 2018)

- 10.1. (a) With a \$1 increase in the beer tax, the expected number of lives that would be saved is 0.45 per 10,000 people. Since New Jersey has a population of 8.1 million, the expected number of lives saved is $0.45 \times 810 = 364.5$. The 95% confidence interval is $(0.45 \pm 1.96 \times 0.30) \times 810 = [-111.78, 840.78]$.
- (b) When New Jersey lowers its drinking age from 21 to 18, the expected fatality rate increases by 0.028 deaths per 10,000. The 95% confidence interval for the change in death rate is $0.028 \pm 1.96 \times 0.070 = [-0.1092, 0.1652]$. With a population of 8.1 million, the number of fatalities will increase by $0.028 \times 810 = 22.68$ with a 95% confidence interval $[-0.1092, 0.1652] \times 810 = [-88.45, 133.81]$.
- (c) The low p -value (or high F -statistic) associated with the F -test for the null hypothesis that the time effects are zero suggests that time effects should be included in the regression.
- (d) Define a binary variable $west$ which equals 1 for the western states and 0 for the other states. Include the interaction term between the binary variable $west$ and the unemployment rate, $west \times (\text{unemployment rate})$, in the regression equation corresponding to column (4). Suppose the coefficient associated with unemployment rate is denoted by β and the coefficient associated with $west \times (\text{unemployment rate})$ is denoted by γ . Then β captures the effect of the unemployment rate in the eastern states, and $\beta + \gamma$ captures the effect of the unemployment rate in the western states. The difference in the effect of the unemployment rate in the western and eastern states is γ . Using the coefficient estimate ($\hat{\gamma}$) and the standard error $SE(\hat{\gamma})$, you can calculate the t -statistic to test whether the difference is statistically significant at a given significance level.

10.3. The five potential threats to the internal validity of a regression study are: omitted variables, misspecification of the functional form, imprecise measurement of the independent variables, sample selection, and simultaneous causality. You should think about these threats one-by-one. Are there important omitted variables that affect traffic fatalities and that may be correlated with the other variables included in the regression? The most obvious candidates are the safety of roads, average weather patterns, and so forth. These variables are essentially constant over the sample period, so their effect is captured by the state fixed effects. You may think of something that we missed. Since most of the variables are binary variables, the largest functional form choice involves the *Beer Tax* variable. A linear specification is used in the text, which seems generally consistent with the data in Figure 8.2. To check the reliability of the linear specification, it would be useful to consider a log specification or a quadratic. Measurement error does not appear to be a problem, as variables like traffic fatalities and taxes are accurately measured. Similarly, sample selection is not a problem because data were used from all of the 48 continental states. Simultaneous causality could be a potential problem. That is, states with high fatality rates might decide to increase taxes to reduce consumption. Expert knowledge is required to determine if this is a problem.

10.5. Let $D2_i = 1$ if $i = 2$ and 0 otherwise; $D3_i = 1$ if $i = 3$ and 0 otherwise ... $Dn_i = 1$ if $i = n$ and 0 otherwise. Let $B2_t = 1$ if $t = 2$ and 0 otherwise; $B3_t = 1$ if $t = 3$ and 0 otherwise ... $BT_t = 1$ if $t = T$ and 0 otherwise. Then $\beta_0 = \alpha_1 + \lambda_1$; $\gamma_i = \alpha_i - \alpha_1$ and $\delta_t = \lambda_t - \beta_1$.

- 10.7. (a) Average snow fall does not vary over time, and thus will be perfectly collinear with (and will be controlled for by) the state fixed effect.
- (b) $Snow_{it}$ does vary with time, and so this method can be used along with state fixed effects.

10.9. (a) $\hat{\alpha}_i = \frac{1}{T} \sum_{t=1}^T Y_{it}$ which has variance $\frac{\sigma_u^2}{T}$. Because T is not growing, the variance is not getting small. $\hat{\alpha}_i$ is not consistent.

(b) The average in (a) is computed over T observations. In this case T is small ($T = 4$), so the normal approximation from the CLT is not likely to be very good.

10.11 Using the hint, equation (10.22) can be written as

$$\begin{aligned}\hat{\beta}_1^{DM} &= \frac{\sum_{i=1}^n \left(\frac{1}{4}(X_{i2} - X_{i1})(Y_{i2} - Y_{i1}) + \frac{1}{4}(X_{i2} - X_{i1})(Y_{i2} - Y_{i1}) \right)}{\sum_{i=1}^n \left(\frac{1}{4}(X_{i2} - X_{i1})^2 + \frac{1}{4}(X_{i2} - X_{i1})^2 \right)} \\ &= \frac{\sum_{i=1}^n (X_{i2} - X_{i1})(Y_{i2} - Y_{i1})}{\sum_{i=1}^n (X_{i2} - X_{i1})^2} = \hat{\beta}_1^{BA}\end{aligned}$$