ØAMET4100 · Spring 2019
Worksheet 6

Instructor: Fenella Carpena

February 14, 2019

1 Review the Concepts

Exercise 1.1 (Stock & Watson, Review the Concepts 10.1) Why is it necessary to use two subscripts, i and t, to describe panel data? What does i refer to? What does t refer to?

Exercise 1.2 (Stock & Watson, Review the Concepts 10.2) A researcher is using a panel data set on $n = 1000$ workers over $T = 10$ years (from 2001 to 2010) that contains the workers' earnings, gender, education, and age. Assume that the education variable is changing over time for at least some individuals in the sample (for example, because some individuals return to school). The researcher is interested in the effect of education on earnings. Give some examples of unobserved person-specific variable that are correlated with both education and earnings. Can you think of examples of time-specific variables that might be correlated with education and earnings? How would you control for these person-specific and time-specific effects in a panel data regression?

Exercise 1.3 (Stock & Watson, Review the Concepts 10.3) Can the regression that you suggested in response to Exercise 1.2 above be used to estimate the effect of gender on an individual's earnings? Can that regression be used to estimate the effect of the national unemployment rate on an individual's earnings? Explain.

Exercise 1.4 (Stock & Watson, Review the Concepts 10.4) In the context of the regression you suggested in Exercise 1.2, explain why the regression error for a given individual might be serially correlated.

2 Traffic Deaths and Beer Taxes

Exercise 2.1 Drunk driving is a major cause of fatal car crashes in the US (e.g., as much as 25% of all fatal car crashes involve a driver who had been drinking). To address this issue, many states have implemented various policies, such as taxing beer. Our goal is to investigate whether beer taxes—which is designed to discourage drunk driving—are indeed effective in reducing traffic deaths. For the following exercise, refer to Table 10.1 of the textbook. This table shows regression results for an analysis for the relationship between beer taxes and traffic death, and the table is attached at the end of this worksheet.

These regressions use panel data on 48 states for the years 1982 to 1988. Some of the variables are:

- `Fatality rate`: the number of traffic deaths per 10,000 people in each state and year
- `Beer tax`: the tax in dollars per case of beer (adjusted for inflation)
- `Drinking age`: the legal drinking age in each state and year; in the regression table, we have binary variables indicating whether the legal drinking age is 18, 19, or 20.
- `Mandatory jail or community service`: a dummy variable equal to one if the state's laws require jail time or community service for a convicted drunk driver.

(a) Interpret each of the coefficients on beer taxes in columns (1) and (2). Why do you think the sign of the coefficient changed between these two columns?

(b) Suppose that the average beer tax in this data is equal to \$0.50 per case, and that the average fatality rate is 2 per 10,000. Using the results in column (4), what is the estimated effect of doubling the average beer tax on the traffic fatality rate? Do you think this effect is large or small? Is this effect precisely estimated?

(c) (Stock & Watson, Exercise 10.1a) New Jersey has a population of 8.1 million people. suppose that New Jersey increased the tax on a case of beer by $1. Use the results in column (4) to predict the number of lives that would be saved over the next year. Construct a 95% confidence interval for your answer.

(d) (Stock & Watson, Exercise 10.1b) The drinking age in New Jersey is 21. Suppose that New Jersey lowered the drinking age to 18. Use the results in column (4) to predict the change in the number of traffic fatalities in the next year. Construct a 95% confidence interval for your answer. (Note: In the set of dummy variables *Drinking age 18*, *Drinking age 19*, and *Drinking age 20* in the regression, the omitted category is for the case when drinking age is 21.)

(e) (Stock and Watson, Exercise 10.1c.) Suppose that real income per capita in New Jersey increases by 1% in the next year. Use the results in column (4) to produce the change in the number of traffic fatalities in the next year. Construct a 90% confidence interval for your answer.

(f) Using the results from column (4), does the legal drinking age affect traffic deaths? Does the punishment associated with the first conviction for driving under the influence of alcohol affect traffic deaths?

(g) Using the results from column (4), do economic conditions in the city affect traffic deaths? Interpret the sign and magnitude of the coefficients on unemployment rate and income. Provide an explanation for the relationship you find.

(h) Compare columns (4) and (5). On the basis of these results, would you decide to include or exclude economic variables in the regression?

(i) Compare columns (4) and (6). Is the coefficient on Beer Tax sensitive to the functional form of the Drinking age variable?

(j) (Stock & Watson, Exercise 10.1d) Should time effects be included in this regression? Why or why not?

(k) Discuss the potential threats to the validity of the results, even if we include state fixed effects and time fixed effects in our regressions.

(l) (Stock and Watson, Exercise 10.1e) A researcher conjectures that the unemployment rate has a different effect on traffic fatalities in the western states than in the other states. How would you test this hypothesis? Be specific about the specification of the regression and the statistical test you would use.

TABLE 10.1 Regression Analysis of the Effect of Drunk Driving Laws on Traffic Deaths

Dependent variable: Traffic fatality rate (deaths per 10,000).

Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Beer tax	0.36**	−0.66*	−0.64+	−0.45	−0.69*	−0.46	−0.93**
	(0.05)	(0.29)	(0.36)	(0.30)	(0.35)	(0.31)	(0.34)
Drinking age 18				0.028	−0.010		0.037
				(0.070)	(0.083)		(0.102)
Drinking age 19				−0.018	−0.076		−0.065
				(0.050)	(0.068)		(0.099)
Drinking age 20				0.032	−0.100+		−0.113
				(0.051)	(0.056)		(0.125)
Drinking age						−0.002	
						(0.021)	
Mandatory jail or community service?				0.038	0.085	0.039	0.089
				(0.103)	(0.112)	(0.103)	(0.164)
Average vehicle miles per driver				0.008	0.017	0.009	0.124
				(0.007)	(0.011)	(0.007)	(0.049)
Unemployment rate				−0.063**		−0.063**	−0.091**
				(0.013)		(0.013)	(0.021)
Real income per capita (logarithm)				1.82**		1.79**	1.00
				(0.64)		(0.64)	(0.68)
Years	1982–88	1982–88	1982–88	1982–88	1982–88	1982–88	1982 & 1988 only
State effects?	no	yes	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes	yes
Clustered standard errors?	no	yes	yes	yes	yes	yes	yes
***F*-Statistics and *p*-Values Testing Exclusion of Groups of Variables**							
Time effects = 0			4.22	10.12	3.48	10.28	37.49
			(0.002)	(< 0.001)	(0.006)	(< 0.001)	(< 0.001)
Drinking age coefficients = 0				0.35	1.41		0.42
				(0.786)	(0.253)		(0.738)
Unemployment rate, income per capita = 0				29.62		31.96	25.20
				(< 0.001)		(< 0.001)	(< 0.001)
$\bar{R}^2$	0.091	0.889	0.891	0.926	0.893	0.926	0.899

These regressions were estimated using panel data for 48 U.S. states. Regressions (1) through (6) use data for all years 1982 to 1988, and regression (7) uses data from 1982 and 1988 only. The data set is described in Appendix 10.1. Standard errors are given in parentheses under the coefficients, and *p*-values are given in parentheses under the *F*-statistics. The individual coefficient is statistically significant at the +10%, *5%, or **1% significance level.