Problem Set 1, Econ220C

Yixiao Sun

Due on April 16

1. Consider the causal model $y \leftarrow \sin(x_f) + (\cos x_f)\, x_o$. Suppose that the settings of x_f and x_o are iid sequences $(X_{fi}, X_{oi})_{i=1}^n$ satisfying $E(X_{oi}|X_{fi}) = 0$ and the response of the free variable y is given by

$$Y_i = \sin(X_{fi}) + X_{oi}\cos(X_{fi}).$$

Given the observations $\{(X_{fi}, Y_i), i = 1, 2, ..., n\}$, we define

$$(\hat{\alpha}, \hat{\beta}) = \arg\min_{\alpha,\beta} \sum_{i=1}^n (Y_i - \alpha - X_{fi}\beta)^2.$$

Note that $\{X_{oi}\}$ are not observed.

(a) Assume that X_f is standard normal. Can you find $\text{plim}_{n\to\infty}\hat{\alpha}$ and $\text{plim}_{n\to\infty}\hat{\beta}$ without imposing any distribution assumption on X_o? If yes, provide the numerical values for $\text{plim}_{n\to\infty}\hat{\alpha}$ and $\text{plim}_{n\to\infty}\hat{\beta}$.

(b) Do the probability limits $\text{plim}_{n\to\infty}\hat{\alpha}$ and $\text{plim}_{n\to\infty}\hat{\beta}$ depend on the distribution of X_f? To answer this, you can try to obtain the probability limits when X_f is $N(1, 1)$.

2. Consider the causal model

$$y \leftarrow d\,(x_f) + e\,(x_o)$$

for

$$d\,(x_f) = \alpha + x_f\beta + x_f^2\gamma$$

and

$$e\,(x_o) = x_o$$

where $\gamma \neq 0$. With iid settings X_{fi} and X_{oi} satisfying $E(X_{oi}|X_{fi}) = 0$, we obtain

$$Y_i = d(X_{fi}) + e(X_{oi}).$$

Given the observations $(X_{fi}, Y_i)_{i=1}^n$, we want to estimate the average cause effect

$$\delta = ED(X_{fi}) \text{ where } D\,(x_f) = \frac{\partial}{\partial x_f} d\,(x_f)\,.$$

Assume that $X_{fi} \sim iid$ with $EX_{fi} = \mu, E(X_{fi} - \mu)^2 = \sigma^2$, and $E(X_{fi} - \mu)^3 = \omega$.

(a) Find δ in terms of $\alpha, \beta, \gamma, \mu, \sigma^2$ and ω.

(b) Suppose the function form of $d(x_f)$ is not known to us, and we run the following regression by OLS:

$$Y_i = \hat{\delta}_0 + X_{fi}\hat{\delta}_1 + error_i$$

Is $\hat{\delta}_1$ consistent for δ as $n \to \infty$? If yes, explain. If no, under what additional condition(s) is $\hat{\delta}_1$ consistent for δ? (Express the condition(s) in terms of μ, σ^2 and ω.)

(c) Do you agree with the following often-heard statement: "The coefficient estimated by the linear OLS, by virtue of a Taylor series expansion, provides a reliable estimate of the derivative of the function of interest at the mean of X_{fi}"? Explain.

3. Consider the fixed-effects panel data model:

$$Y_{it} = \alpha_i + X_{it}\beta + u_{it}, \ i = 1, ..., N \text{ and } t = 1, ..., T.$$

Suppose there is a policy change at time $1 < \tau + 1 < T$ such that

$$X_{it} = \begin{cases} 1, & if \ t \geq \tau + 1, \\ 0, & if \ t \leq \tau. \end{cases}$$

Assume that $T > 2$.

(a) Derive the fixed-effects estimator. Please give an explicit expression.

(b) Derive the first-differenced estimator. Please give an explicit expression.

(c) How do the FE and FD estimators compare? Under what condition(s) on $\{u_{it}\}$ is the FE estimator consistent? Under what condition(s) is the FD estimator consistent?

4. Use your favorite package to answer this question. A sample Matlab program is posted on the Canvas. I encourage you to write your own program before reading the sample program. The sample program works for a scalar x_{it}. For vector cases, some modifications are required. Publish your codes and report the URL on Canvas. Group study is encouraged, but you have to write your own code and report your own analysis. Please upload your analysis directly to Canvas.

Consider the following data generating process:

$$Y_{it} = X_{it}\beta + u_{it}, i = 1, 2, ..., N \text{ and } t = 1, 2, ..., T \tag{1}$$

where $\beta = 1$, $X_{it} \sim iidN(0, 1)$ across i and t, and conditional on X_i, $\{u_{it}\}$ are independent across t with distribution $N(0, |X_{it}|^2)$. u_{it} is independent of u_{js} for any $i \neq j$. Now suppose we use the FE estimator to estimate β :

$$\hat{\beta} = \beta + \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right)^2 \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right) \left(u_{it} - \bar{u}_{i,\cdot} \right) \right). \tag{2}$$

(a) Let $N = 500$ and $T = 5$. Simulate the sampling distribution of $\hat{\beta}$ using 1000 simulation replications. For each simulated sample, compute the robust standard errors $\hat{\sigma}_\beta$ and $\tilde{\sigma}_\beta$ according to

$$\begin{aligned}\hat{\sigma}_\beta^2 &= S_{XX}^{-2} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right) \left(\hat{u}_{it} - \overline{\hat{u}_{i,\cdot}} \right) \right]^2 \\ &= S_{XX}^{-2} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right) \left(\hat{u}_{it} - \overline{\hat{u}_{i,\cdot}} \right) \left(\hat{u}_{is} - \overline{\hat{u}_{i,\cdot}} \right) \left(X_{is} - \bar{X}_{i,\cdot} \right) \qquad (3)\end{aligned}$$

and

$$\tilde{\sigma}_\beta^2 = S_{XX}^{-2} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right)^2 \left(\hat{u}_{it} - \overline{\hat{u}_{i,\cdot}} \right)^2, \qquad (4)$$

where

$$S_{XX} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \bar{X}_{i,\cdot} \right)^2 \right)$$

and $\hat{u}_{it}$ is the estimated residual.

(b) Compute the standard deviation $sd(\hat{\beta})$ of the finite sample distribution of $\hat{\beta}$.

(c) Compute the bias, std and rmse (root mean squared error) of $\hat{\sigma}_\beta$ and $\tilde{\sigma}_\beta$. According to the rmse criterion, which estimator is better?

(d) Repeat (a)–(c) for $T = 10, 20$. Does the relative rmse advantage of the two estimators remain valid for different values of T?

(e) Can you provide a theoretical explanation of your findings?

(f) Finally, do you think your results will change if you change the value of β, say let $\beta = 314.15926$?

5. The debate regarding crime and guns is of course long running. The book 'More Guns, Less Crime: Understanding Crime and Gun Control Laws' by Lott (American Enterprise Institute) loudly made the claim that 'shall' laws reduce crime based on correlation analysis. In this question, we will evaluate the claim and see whether we can shoot down the 'More Guns, Less Crime' hypothesis (Ayres and Donohue III in the Stanford Law Review (2003)). The book received 4.5 out of five stars at Amazon.com and there are 175 customer reviews. Everybody has something to say about this issue. Let's see what we can conclude from econometric analysis.

The questions are based on the dataset handguns.dta which you can download from the Ted. The data consists of data from 50 States plus DC for each year from 1977 to 1999. The data we will be analyzing are crimes rates for various crime definitions provided by the Bureau of Justice Statistics. The variables are described in the STATA data set. The main regressor we will be focussing on is a dummy variable for whether or not the state allows widespread carrying of concealed weapons. The variable shall is one for states which have 'shall issue' laws, which means that licenses must be given to all applicants that are citizens, mentally competent and have not been convicted of a felony.

For additional background,you may want to read
http://en.wikipedia.org/wiki/More_Guns,_Less_Crime#Shall_issue_laws
or http://islandia.law.yale.edu/ayers/Ayres_Donohue_article.pdf

Note: you do not need to submit your STATA output. However, please submit your Stata do file.

I. We will examine the effect of shall on rates of violent crime, murder rates and robberies. To this end, run regressions of the logs of each of these variables on *shall* (including an intercept) with the robust option. Report the results in a table with a column for each regression and the values and their standard errors in rows. That is, fill in the following table:

Dependent Var =	ln(vio)	ln(mur)	ln(rob)
$\hat{\beta}_0$	6.13		
	(0.02)	( )	( )
$\hat{\beta}_1$ (shall)	-0.443		
	(0.048)	( )	( )
R^2	0.09		

(a) What is the effect of 'shall' laws on each of the crime rates. Are the effects large statistically? Explain.

To get started, you can first download the file 'handguns.dta' from the course webpage and then use the following commands in your STATA do file.

```
 clear
- clearn matrix
 #delimit ;
 set more off;
 set matsize 300;
 capture log close;
 cd "D:\Teaching\";
 log using shall.log, replace;
 use handguns.dta;
 desc;
 summarize;
 gen log_vio=log(vio);
 gen log_mur=log(mur);
 gen log_rob=log(rob);
/************ Question 1 *******/
 reg log_vio shall, r;
```

```
reg log_mur shall, r;
reg log_rob shall, r;
```

II. Now we will control for a number of variables. First, it is well understood that demographic variables play a role. Many have argued socioeconomic variables also play a part. Most also would at least hope that jail is a deterrent. Run the above regressions but now add the variables *incarc_ rate, density, pop*, *pm1029*, and *avginc* to the regression. Report the results in a table given below.

(a) What is the effect of the 'shall' laws now?

(b) Is the difference between the results here and in the results from Question (I) large in a practical sense?

Dependent Var =	ln(vio)	ln(mur)	ln(rob)
$\hat{\beta}_0$		-0.17	
	( )	(0.29)	( )
$\hat{\beta}_1$ (shall)		-0.309	
	( )	(0.037)	( )
R^2		0.55	

Note: `incarc_rate, density, pop, pm1029`, and `avginc` should be included in the regression but you do not have to report their coefficients.

III. One omitted variable from the above analysis is differences in laws and law enforcement across states and time. We want to understand how this might affect results to provide more foundation for the interval validity of the results. Recall the omitted variable bias formula:

$$\hat{\beta}_1 \to \beta_1 + \frac{cov(X_{1i}, u_i)}{var(X_{1i})}.$$

Stronger laws would hopefully deter crime, especially crimes that are more rational in nature like roberries, and perhaps violence. In this sense we would expect that stronger laws would be associated with less crime and hence lower values for u_i.

(a) Typically 'shall' laws are pushed using law and order arguments. States with a larger 'law and order' constituency would have stronger laws and would be more likely to have 'shall' laws. What does this suggest the sign of $cov(X_{1i}, u_i)$ where X_{1i} is the dummy variable for 'shall'.

(b) If there is a bias in $\hat{\beta}_1$ (the coefficient on shall), which direction is it?

IV. Since we have a panel data set, we are able to control for omitted variables that are constant over time. We want to run the same regressions (i.e. use the same control variables) as in QII, but now add state fixed effects. Do this for each of the three dependent variables we have examined, and construct three tables (one for each dependent variable). In each table report the coefficient on 'shall' along with its standard error, test for the inclusion of state effects if included.

Each table should look like the following (with the entries added instead of the XX's, of course).

Dep Var=ln(vio)	1	2	3	cluster option
$\hat{\beta}_1(shall)$ (std. error)	XX (XX)	XX (XX)	XX (XX)	XX (XX)
State Fixed Effect	No	Yes	Yes	Yes
Year Fixed Effect	No	No	Yes	Yes
F test for State Effect	-	Statistic (p-value)	Statistic (p-value)	-
F test for Year Effect	-	-	Statistic (p-value)	-

(a) Describe the effect of controlling for state effects on the coefficient estimate for the effect of 'shall' laws on crime.

(b) What does this tell us about omitted variables in the specification without state or time effects?

(c) What is the statistical evidence that state dummy variables should be included?

(d) Do these results suggest that the arguments in QIII are correct?

(e) What is the year fixed effects supposed to capture? Is there any statistical evidence that year fixed effects should be included?

(f) The standard errors of the estimate jump quite a bit when we use clustered standard errors. What does this suggest?

Stata issues:

The command **tab state, generate(statedummy)** will take a variable in your data set called state which has a number for each state and construct dummy variables named **statedummy1** through to the highest number **statedummy51** where **statedummy1=1** for **state** equal to 1 and zero otherwise, **statedummy2**=1 for **state** equal to 2 and zero otherwise, etc.

The following code can be used to generate state dummies and test their effects.

```
tab state, gen(statedummy);
 /* column 1 in the table */
 reg log_vio shall incarc_rate density pop pm1029 avginc, r;

 /* column 2 in the table */
 reg log_vio shall incarc_rate density pop pm1029 avginc statedummy*,
r;
 testparm statedummy*;
```

```
/* if you want to compute standard error that is robust to the time series
correlation in u_it, you can use the following commands. */
 reg log_vio shall incarc_rate density pop pm1029 avginc statedummy*,
cluster(state) r ;
/* you can also use xtreg here */
 testparm statedummy*;
```

Note: `testparm` provides a useful alternative to `test` that permits varlist rather than a list of coefficients (which is often nothing more than a list of variables), allowing use of standard Stata notation, including the wild card '*'.