ECON4150 - Introductory Econometrics Seminar 6

Stock and Watson EE10.1

April 28, 2015

Guns data set

- Some U.S. states have enacted "shall-issue" laws which allow citizens to carry concealed weapons
- We are going to investigate the effect of shall-issue laws on violent crime rates.
- Two opposite hipotesis: "more guns less crime" and "more guns more crime"
- In this exercise we use the data set Guns.dta.
- Balanced panel of data on 50 US states, plus the District of Columbia for the years 1977- 1999
- There are a total of 51 states x 23 years = 1173 observations

Variables Definition

Variable Definitions

Variable	Definition
vio	violent crime rate (incidents per 100,000 members of the population)
rob	robbery rate (incidents per 100,000)
mur	murder rate (incidents per 100,000)
shall	= 1 if the state has a shall-carry law in effect in that year = 0 otherwise
incarc_rate	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
density	population per square mile of land area, divided by 1000
avginc	real per capita personal income in the state, in thousands of dollars
pop	state population, in millions of people
pm1029	percent of state population that is male, ages 10 to 29
pw1064	percent of state population that is white, ages 10 to 64
pb1064	percent of state population that is black, ages 10 to 64
stateid	ID number of states (Alabama = 1, Alaska = 2, etc.)
year	Year (1977-1999)

Panel data aspect

- So far cross-section data

	stateid	vio	mur	rob	incarc_rate	pb1064	pw1064	pm1029
1	1	414.4	14.2	96.8	83	8.384873	55.12291	18.17441
2	2	443.2	10.8	96.8	63	7.835672	62.1413	22.35269
3	4	494.2	9.5	138.2	125	3.436422	66.61958	18.72813
4	5	322.9	8.8	83.2	115	5.325915	59.49733	17.29585
5	6	706	11.5	287	85	5.011663	65.51936	18.78835
6	8	511.9	6.3	170.7	87	2.026637	72.37897	20.01816

- Panel data, long form

	stateid	year	vio	mur	rob	incarc_rate	pb1064	pw1064
1	1	77	414.4	14.2	96.8	83	8.384873	55.12291
2	1	78	419.1	13.3	99.1	94	8.352101	55.14367
3	1	79	413.3	13.2	109.5	144	8.329575	55.13586
4	1	80	448.5	13.2	132.1	141	8.408386	54.91259
5	2	77	443.2	10.8	96.8	63	7.835672	62.1413
6	2	78	441.9	12.9	91.3	75	7.945618	62.65009
7	2	79	491.1	13.3	109.6	127	8.057296	63.09599
8	2	80	436	8.9	81.8	133	8.123558	62.7206
9	4	77	494.2	9.5	138.2	125	3.436422	66.61958
10	4	78	552.1	9.4	162.9	129	3.457416	66.48782
11	4	79	593	8.9	175.7	146	3.473442	66.49958
12	4	80	650.9	10.3	193.6	139	3.592809	66.15736

a)

```
 gen ln_vio=ln(vio)
 regress ln_vio shall, robust

Linear regression                               Number of obs =    1173
                                                F(  1,  1171) =   86.86
                                                Prob > F      =  0.0000
                                                R-squared     =  0.0866
                                                Root MSE      =  .61735
------------------------------------------------------------------------------
             |               Robust
      ln_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       shall |  -.4429646   .0475283    -9.32   0.000    -.5362148   -.3497144
       _cons |   6.134919   .0193039   317.81   0.000     6.097045    6.172793
------------------------------------------------------------------------------
 est sto reg1
```

additional information for tables

```
 estadd local State "-"  , replace

added macro:
              e(State) : "-"

 /*
        adds a variable called state with the "value" "-" to display that we did not
        control for state fixed effects
*/
 estadd local Time   "-" , replace

added macro:
               e(Time) : "-"

 /*
         adds a variable called time with the "value" "-" to display that we did not
         control for time fixed effects
*/
 estadd local Control  "-" , replace

added macro:
            e(Control) : "-"

 /*
>         adds a variable called time with the "value" "-" to display that we did not
>         control for other control variables
```

a.

```
 regress ln_vio shall incarc_rate density avginc pop pb1064 pw1064 pm1029, robust

Linear regression                               Number of obs =    1173
                                                F(  8,  1164) =   95.67
                                                Prob > F      =  0.0000
                                                R-squared     =  0.5643
                                                Root MSE      =  .42769

------------------------------------------------------------------------------
             |               Robust
      ln_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       shall |  -.3683869   .0347879   -10.59   0.000     -.436641   -.3001329
 incarc_rate |   .0016126   .0001807     8.92   0.000     .0012581    .0019672
     density |   .0266885   .0143494     1.86   0.063    -.0014651     .054842
      avginc |   .0012051   .0072778     0.17   0.869     -.013074    .0154842
         pop |   .0427098   .0031466    13.57   0.000     .0365361    .0488836
      pb1064 |   .0808526   .0199924     4.04   0.000     .0416274    .1200778
      pw1064 |   .0312005   .0097271     3.21   0.001      .012116    .0502851
      pm1029 |   .0088709   .0120604     0.74   0.462    -.0147917    .0325334
       _cons |   2.981738   .6090198     4.90   0.000     1.786839    4.176638
------------------------------------------------------------------------------

est sto reg2
estadd local State "-"   , replace
estadd local Time   "-" , replace
estadd local Control  "Yes" , replace
```

Table: dependent variable is ln(vio)

	1	2
shall	-0.443^{***}	-0.368^{***}
	(0.0475)	(0.0348)
Control variables	–	*Yes*

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- The coefficient in regression (2) equals -0.368, which suggests that shall-issue law reduce the violent crime rate by 36%.
- The coefficient in (1) is -0.443; in (2) it is -0.369. Both are highly statistically significant.
- Adding the control variables results in a small drop in the estimated coefficient.
- The "real world" significance of the coefficient remains stable

a., iii.

- Possible omitted variables that vary between states but not over time Attitudes towards guns and crime, quality of police and other crime-prevention programs.

b.

```
 /*
xtset tells stata that we are dealing with panel data and that stateid is the panel
variable and year is the time variable
*/

 xtset stateid year
       panel variable:  stateid (strongly balanced)
        time variable:  year, 77 to 99
                delta:  1 unit



// Now we have two alternatives, add a dummy for each state or use xtreg
```

b.

```
xtreg ln_vio shall incarc_rate density avginc pop pb1064 pw1064 pm1029, fe robust
Fixed-effects (within) regression               Number of obs      =      1173
Group variable: stateid                         Number of groups   =        51
R-sq:  within  = 0.2178                         Obs per group: min =        23
       between = 0.0033                                        avg =      23.0
       overall = 0.0001                                        max =        23
                                                F(8,50)           =     34.10
corr(u_i, Xb)  = -0.3687                        Prob > F          =    0.0000
                                 (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
      ln_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       shall |  -.0461415   .0417616    -1.10   0.275    -.1300223    .0377392
 incarc_rate |   -.000071   .0002504    -0.28   0.778    -.0005739    .0004318
     density |  -.1722901   .1376129    -1.25   0.216    -.4486936    .1041135
      avginc |  -.0092037   .0129649    -0.71   0.481    -.0352445     .016837
         pop |   .0115247    .014224     0.81   0.422    -.0170452    .0400945
      pb1064 |   .1042804   .0326849     3.19   0.002     .0386308    .1699301
      pw1064 |   .0408611   .0134585     3.04   0.004     .0138289    .0678932
      pm1029 |  -.0502725   .0206949    -2.43   0.019    -.0918394   -.0087057
       _cons |   3.866017   .7701057     5.02   0.000     2.319214    5.412819
-------------+----------------------------------------------------------------
         rho |  .94712779   (fraction of variance due to u_i)
------------------------------------------------------------------------------
est sto reg3
estadd local State "Yes"  , replace
estadd local Time   "-" , replace
estadd local Control  "Yes" , replace
```

b

```
\\  i.state create a dummy for each state in the regression
 reg ln_vio shall incarc_rate density avginc pop pb1064 pw1064 pm1029 i.state, robust

Linear regression                                      Number of obs =    1173
                                                       F( 58,  1114) =  364.90
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9411
                                                       Root MSE      =  .16072

------------------------------------------------------------------------------
             |               Robust
      ln_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       shall |  -.0461415   .0199433    -2.31   0.021    -.0852721    -.007011
 incarc_rate |   -.000071   .0000973    -0.73   0.466    -.0002619    .0001199
     density |  -.1722901   .1048789    -1.64   0.101    -.3780725    .0334923
      avginc |  -.0092037   .0067335    -1.37   0.172    -.0224155     .004008
         pop |   .0115247   .0097044     1.19   0.235    -.0075162    .0305655
      pb1064 |   .1042804   .0165552     6.30   0.000     .0717976    .1367633
      pw1064 |   .0408611   .0053859     7.59   0.000     .0302935    .0514287
      pm1029 |  -.0502725   .0077908    -6.45   0.000    -.0655588   -.0349863
             |
     stateid |
           2 |   .0559649   .0788371     0.71   0.478     -.098721    .2106508
         ...
         ...
          56 |  -.4804004   .1293103    -3.72   0.000    -.7341196   -.2266813
             |
       _cons |   4.036775   .3845839    10.50   0.000     3.282185    4.791366
------------------------------------------------------------------------------
\\ same result on shall and estimate for each dummy, but we are not interested on those estimates
```

Table: dependent variable is ln(vio)

	1	2	3
shall	-0.443^{***}	-0.368^{***}	-0.0461
	(0.0475)	(0.0348)	(0.0418)
Control variables	—	*Yes*	*Yes*
State fixed effects	—	—	*Yes*

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- The results change when we add state fixed effects.
- The absolute value of the coefficient on shall falls to 0.046, a large reduction in the coefficient from 0.369 without fixed effects.
- Evidently there was important omitted variable bias in the specification without fixed effects.
- The estimate of the effect of shall issue laws on the violent crime rate is no longer statistically significantly different from zero
- The regression model with fixed effects is more credible because this controls for unobserved characteristics that vary between states but that are constant over time

```
xi: xtreg ln_vio shall i.year incarc_rate density avginc pop pb1064 pw1064 pm1029, fe robust
i.year            _Iyear_77-99        (naturally coded; _Iyear_77 omitted)

Fixed-effects (within) regression               Number of obs      =      1173
Group variable: stateid                         Number of groups   =        51
R-sq:  within  = 0.4180                         Obs per group: min =        23
       between = 0.0419                                        avg =      23.0
       overall = 0.0009                                        max =        23
                                                F(30,50)           =     56.86
corr(u_i, Xb)  = -0.2929                        Prob > F           =    0.0000

                              (Std. Err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
      ln_vio |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       shall |  -.0279935   .0407168    -0.69   0.495    -.1097757    .0537886
   _Iyear_78 |   .0585261   .0161556     3.62   0.001     .0260767    .0909755
   ...
   ...
   _Iyear_99 |   .4328776   .2862197     1.51   0.137    -.1420117    1.007767
 incarc_rate |    .000076   .0002079     0.37   0.716    -.0003416    .0004935
     density |   -.091555   .1238622    -0.74   0.463    -.3403396    .1572296
      avginc |   .0009587   .0164931     0.06   0.954    -.0321688    .0340861
         pop |  -.0047544   .0152294    -0.31   0.756    -.0353436    .0258347
      pb1064 |   .0291862   .0495407     0.59   0.558    -.0703192    .1286916
      pw1064 |   .0092501   .0237564     0.39   0.699    -.0384659    .0569662
      pm1029 |   .0733254   .0524733     1.40   0.168    -.0320704    .1787211
       _cons |   3.765525   1.152108     3.27   0.002     1.451448    6.079603
-------------+----------------------------------------------------------------
     sigma_u |   .6663043
     sigma_e |   .1400264
         rho |  .95770338   (fraction of variance due to u_i)
------------------------------------------------------------------------------
est sto reg4
estadd local State "Yes"  , replace
estadd local Time   "Yes" , replace
estadd local Control  "Yes" , replace
```

```
 /*
 Testing the joint significance of the year dummies: _Iyear*,
 with * you consider all the variables starting with _Iyear
*/
  testparm _Iyear*

 ( 1)  _Iyear_78 = 0
 ( 2)  _Iyear_79 = 0
 ( 3)  _Iyear_80 = 0
 ( 4)  _Iyear_81 = 0
...
...
 (18)  _Iyear_95 = 0
 (19)  _Iyear_96 = 0
 (20)  _Iyear_97 = 0
 (21)  _Iyear_98 = 0
 (22)  _Iyear_99 = 0

       F( 22,    50) =   21.62
            Prob > F =    0.0000
estadd scalar F_year = r(F)

added scalar:
            e(F_year) =  21.621583
```

Table: dependent variable is ln(vio)

	1	2	3	4
shall	-0.443^{***}	-0.368^{***}	-0.0461	-0.0280
	(0.0475)	(0.0348)	(0.0418)	(0.0407)
Control variables	–	*Yes*	*Yes*	*Yes*
State fixed effects	–	–	*Yes*	*Yes*
Time fixed effects	–	–	–	*Yes*
F-time dummies				21.62

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- The absolute value of the coefficient on shall falls further to 0.028, the coefficient is not significantly different from zero.
- The time effects are jointly statistically significant, so this regression seems better specified than the regression in part (b).

d

```
// replicate with ln_rob and ln_mur
gen ln_rob=ln(rob)
regress ln_rob shall, robust
est sto reg1
estadd local State "-"  , replace
estadd local Time   "-" , replace
estadd local Control  "-" , replace

regress ln_rob shall incarc_rate density avginc pop pb1064 pw1064 pm1029, robust
est sto reg2
estadd local State "-"  , replace
estadd local Time   "-" , replace
estadd local Control  "Yes" , replace

xtreg ln_rob shall incarc_rate density avginc pop pb1064 pw1064 pm1029, fe robust
est sto reg3
estadd local State "Yes"  , replace
estadd local Time   "-" , replace
estadd local Control  "Yes" , replace

xi: xtreg ln_rob shall i.year incarc_rate density avginc pop pb1064 pw1064 pm1029, fe robust
est sto reg4
estadd local State "Yes"  , replace
estadd local Time   "Yes" , replace
estadd local Control  "Yes" , replace
testparm _Iyear*
estadd scalar F_year = r(F)

esttab reg* ,  label   se beta(2)  keep(shall) ///
  s(Control State Time  F_year , label("Control variables" "State fixed effects" "Time fixed effects"  "F-time dummies")) ///
  mtitles("1" "2" "3" "4") nonumbers title(Dependent variable is ln(rob))

log close
```

dependent variable is ln(rob)

	1	2	3	4
shall	−0.773***	−0.529***	−0.00782	0.0268
	(0.0693)	(0.0510)	(0.0552)	(0.0522)
Control variables	–	*Yes*	*Yes*	*Yes*
State fixed effects	–	–	*Yes*	*Yes*
Time fixed effects	–	–	–	*Yes*
F-time dummies				25.86

dependent variable is ln(dmur)

	1	2	3	4
shall	−0.473***	−0.313***	−0.0608	−0.0150
	(0.0485)	(0.0357)	(0.0370)	(0.0382)
Control variables	–	*Yes*	*Yes*	*Yes*
State fixed effects	–	–	*Yes*	*Yes*
Time fixed effects	–	–	–	*Yes*
F-time dummies				19.61

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results are similar to the results using violent crimes:

- There is a large estimated effect of concealed weapons laws in specifications (1) and (2).
- This effect is however due to omitted variable bias because the effect disappears when state and time effects are added.

e: Remaining threats to internal validity

Definition A statistical analysis is said to have internal validity if the statistical inferences about causal effects are valid for the population and setting being studied. The estimator should be unbiased and consistent and the standard errors are computed in a way that makes confidence intervals have the desired confidence levels. Threats to internal validity

- Functional form misspecification
- Omitted variable bias
- Measurement error
- Sample selection
- Simultaneous causality

$\Big\} \; E(u_i|x_i) \neq 0$

- Heteroskedasticity and/or correlated error terms $\rightarrow$ violation i.d.d.

- Simultaneous causality: violent crimes this may induce policy makers to change concealed weapons laws
- Omitted variables: There might be important variables that vary between states and over time that are omitted from the regression model. For example other policy measures that are related to the implementation of shall issue laws and that affect crime rates

- The most credible results include both state fixed effects and time fixed effects.
- These results indicate that there is no significant effect of concealed weapon laws on crime rates