Regression with Panel Data
(SW Ch. 10)

A ***panel dataset*** contains observations on multiple entities (individuals), where each entity is observed at two or more points in time.

Examples:

- Data on 420 California school districts in 1999 *and again* in 2000, for 840 observations total.
- Data on 50 U.S. states, each state is observed in 3 years, for a total of 150 observations.
- Data on 1000 individuals, in four different months, for 4000 observations total.

Notation for panel data

A double subscript distinguishes entities (states) and time periods (years)

i = entity (state), n = number of entities,

so $i = 1,\ldots,n$

t = time period (year), T = number of time periods

so $t = 1,\ldots,T$

Data: Suppose we have 1 regressor. The data are:

$$(X_{it}, Y_{it}),\ i = 1,\ldots,n,\ t = 1,\ldots,T$$

Panel data notation, ctd.

Panel data with k regressors:

$$(X_{1it}, X_{2it},\ldots,X_{kit}, Y_{it}),\ i = 1,\ldots,n,\ t = 1,\ldots,T$$

n = number of entities (states)

T = number of time periods (years)

Some jargon…

- Another term for panel data is ***longitudinal data***
- ***balanced panel***: no missing observations
- ***unbalanced panel***: some entities (states) are not observed for some time periods (years)

Why are panel data useful?

With panel data we can control for factors that:

- Vary across entities (states) but do not vary over time
- Could cause omitted variable bias if they are omitted
- are unobserved or unmeasured – and therefore cannot be included in the regression using multiple regression

Here's the key idea:

If an omitted variable does not change over time, then any *changes* in *Y* over time cannot be caused by the omitted variable.

Example of a panel data set: Traffic deaths and alcohol taxes

Observational unit: a year in a U.S. state

- 48 U.S. states, so n = of entities = 48
- 7 years (1982,…, 1988), so T = # of time periods = 7
- Balanced panel, so total # observations = 7×48 = 336

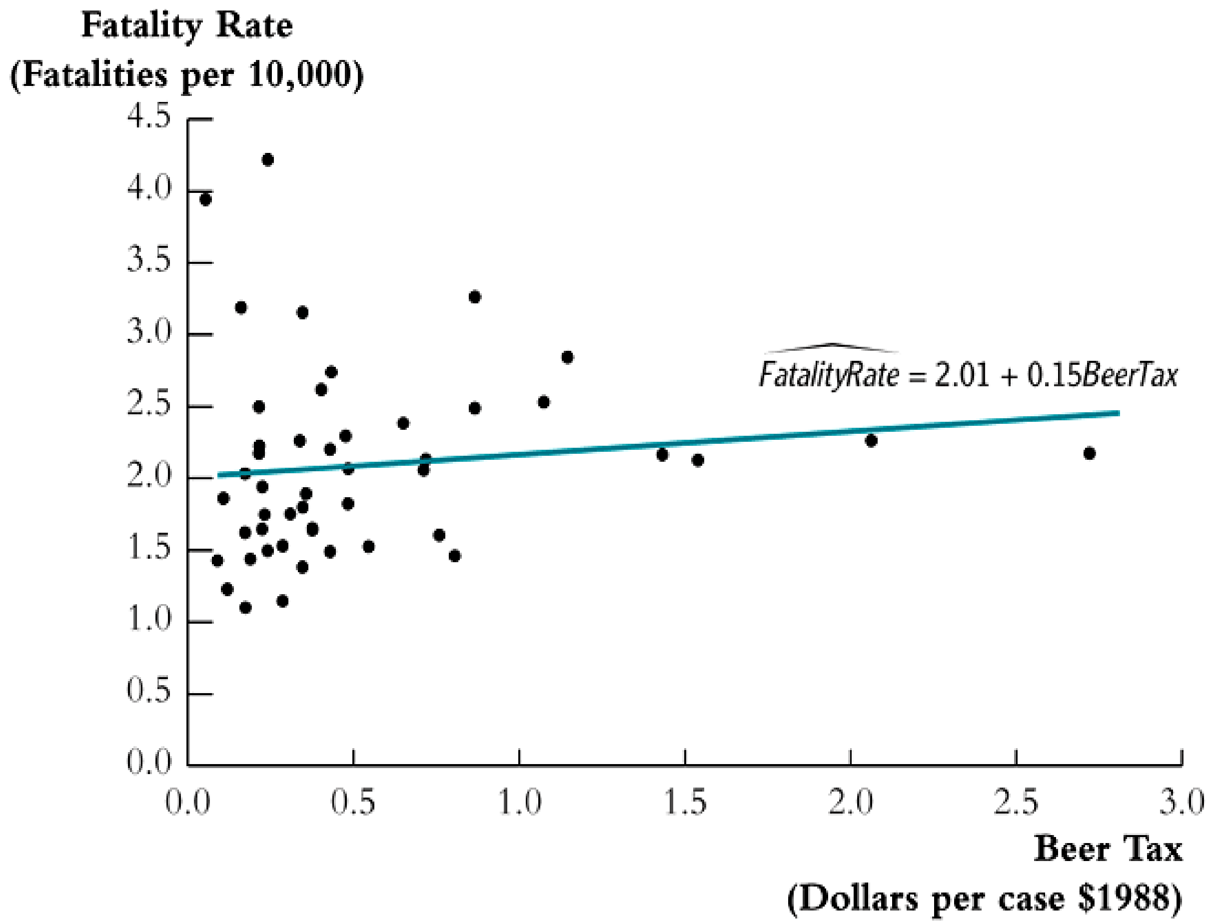
Variables:

- Traffic fatality rate (# traffic deaths in that state in that year, per 10,000 state residents)
- Tax on a case of beer
- Other (legal driving age, drunk driving laws, etc.)

Traffic death data for 1982

FIGURE 8.1 The Traffic Fatality Rate and the Tax on Beer

Panel a is a scatterplot of traffic fatality rates and the real tax on a case of beer (in 1988 dollars) for 48 states in 1982. Panel b shows the data for 1988. Both plots show a positive relationship between the fatality rate and the real beer tax.



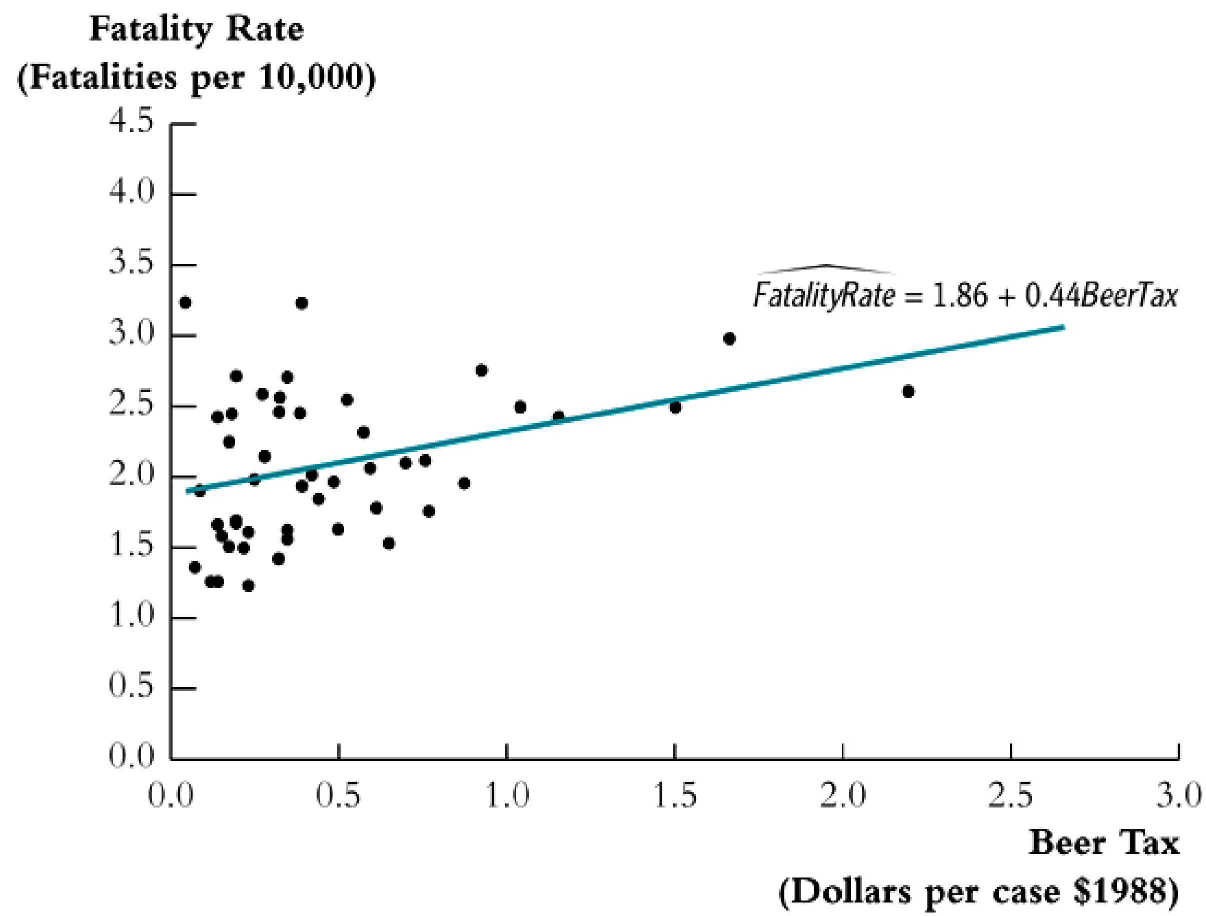
(a) 1982 data

Higher alcohol taxes, more traffic deaths?

Traffic death data for 1988

FIGURE 8.1 The Traffic Fatality Rate and the Tax on Beer

Panel a is a scatterplot of traffic fatality rates and the real tax on a case of beer (in 1988 dollars) for 48 states in 1982. Panel b shows the data for 1988. Both plots show a positive relationship between the fatality rate and the real beer tax.



(b) 1988 data

Higher alcohol taxes, more traffic deaths?

Why might there be *more* traffic deaths in states that have higher alcohol taxes?

Other factors that determine traffic fatality rate:

- Quality (age) of automobiles
- Quality of roads
- "Culture" around drinking and driving
- Density of cars on the road

These omitted factors could cause omitted variable bias.

Example #1: traffic density. Suppose:

(i) High traffic density means more traffic deaths

(ii) (Western) states with lower traffic density have lower alcohol taxes

- Then the two conditions for omitted variable bias are satisfied. Specifically, "high taxes" could reflect "high traffic density" (so the OLS coefficient would be biased positively – high taxes, more deaths)
- Panel data lets us eliminate omitted variable bias when the omitted variables are constant over time within a given state.

Example #2: cultural attitudes towards drinking and driving

(i) arguably are a determinant of traffic deaths; and

(ii) potentially are correlated with the beer tax, so beer taxes could be picking up cultural differences (omitted variable bias).

- Then the two conditions for omitted variable bias are satisfied. Specifically, "high taxes" could reflect "cultural attitudes towards drinking" (so the OLS coefficient would be biased)
- Panel data lets us eliminate omitted variable bias when the omitted variables are constant over time within a given state.

Panel Data with Two Time Periods
(SW Section 10.2)

Consider the panel data model,

$$FatalityRate_{it} = \beta_0 + \beta_1 BeerTax_{it} + \beta_2 Z_i + u_{it}$$

Z_i is a factor that does not change over time (density), at least during the years on which we have data.

- Suppose Z_i is not observed, so its omission could result in omitted variable bias.
- The effect of Z_i can be eliminated using $T = 2$ years.

The key idea:

Any *change* in the fatality rate from 1982 to 1988 cannot be caused by Z_i, because Z_i (by assumption) does not change between 1982 and 1988.

The math: consider fatality rates in 1988 and 1982:

$$FatalityRate_{i1988} = \beta_0 + \beta_1 BeerTax_{i1988} + \beta_2 Z_i + u_{i1988}$$

$$FatalityRate_{i1982} = \beta_0 + \beta_1 BeerTax_{i1982} + \beta_2 Z_i + u_{i1982}$$

Suppose $E(u_{it}|BeerTax_{it}, Z_i) = 0$.

Subtracting 1988 – 1982 (that is, calculating the change), eliminates the effect of Z_i…

$$FatalityRate_{i1988} = \beta_0 + \beta_1 BeerTax_{i1988} + \beta_2 Z_i + u_{i1988}$$

$$FatalityRate_{i1982} = \beta_0 + \beta_1 BeerTax_{i1982} + \beta_2 Z_i + u_{i1982}$$

so

$$FatalityRate_{i1988} - FatalityRate_{i1982} =$$
$$\beta_1(BeerTax_{i1988} - BeerTax_{i1982}) + (u_{i1988} - u_{i1982})$$

- The new error term, $(u_{i1988} - u_{i1982})$, is uncorrelated with either $BeerTax_{i1988}$ or $BeerTax_{i1982}$.
- This "difference" equation can be estimated by OLS, even though Z_i isn't observed.
- The omitted variable Z_i doesn't change, so it cannot be a determinant of the *change* in Y

Example: Traffic deaths and beer taxes

1982 data:

$$\widehat{FatalityRate} = 2.01 + 0.15BeerTax \qquad (n = 48)$$
$$(.15) \quad (.13)$$

1988 data:

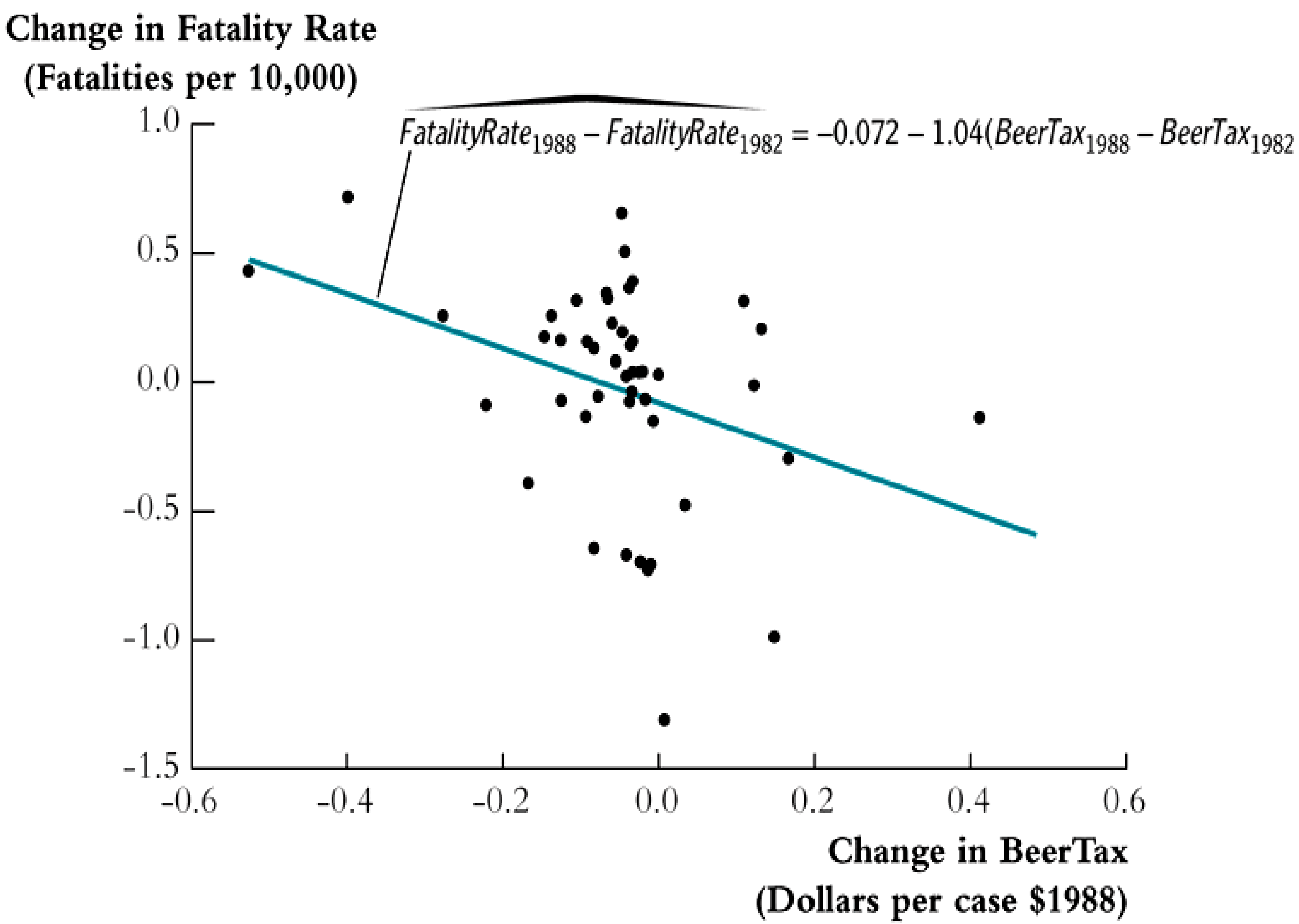
$$\widehat{FatalityRate} = 1.86 + 0.44BeerTax \qquad (n = 48)$$
$$(.11) \quad (.13)$$

Difference regression (n = 48)

$$\widehat{FR_{1988} - FR_{1982}} = -.072 - 1.04(BeerTax_{1988} - BeerTax_{1982})$$
$$(.065) \quad (.36)$$

FIGURE 8.2 **Changes in Fatality Rates and Beer Taxes, 1982–1988**

This is a scatterplot of the *change* in the traffic fatality rate and the *change* in real beer taxes between 1982 and 1988 for 48 states. There is a negative relationship between changes in the fatality rate and changes in the beer tax.



Fixed Effects Regression
(SW Section 10.3)

What if you have more than 2 time periods ($T > 2$)?

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + u_i,\ i = 1,\ldots,n,\ T = 1,\ldots,T$$

We can rewrite this in two useful ways:

1. "n-1 binary regressor" regression model
2. "Fixed Effects" regression model

We first rewrite this in "fixed effects" form. Suppose we have n = 3 states: California, Texas, Massachusetts.

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + u_i,\ i = 1,\ldots,n,\ T = 1,\ldots,T$$

Population regression for California (that is, i = CA):

$$Y_{CA,t} = \beta_0 + \beta_1 X_{CA,t} + \beta_2 Z_{CA} + u_{CA,t}$$

$$= (\beta_0 + \beta_2 Z_{CA}) + \beta_1 X_{CA,t} + u_{CA,t}$$

or

$$Y_{CA,t} = \alpha_{CA} + \beta_1 X_{CA,t} + u_{CA,t}$$

- $\alpha_{CA} = \beta_0 + \beta_2 Z_{CA}$ doesn't change over time
- α_{CA} is the intercept for CA, and β_1 is the slope
- The intercept is unique to CA, but the slope is the same in all the states: parallel lines.

For TX:

$$Y_{TX,t} = \beta_0 + \beta_1 X_{TX,t} + \beta_2 Z_{TX} + u_{TX,t}$$
$$= (\beta_0 + \beta_2 Z_{TX}) + \beta_1 X_{TX,t} + u_{TX,t}$$

or

$$Y_{TX,t} = \alpha_{TX} + \beta_1 X_{TX,t} + u_{TX,t}, \text{ where } \alpha_{TX} = \beta_0 + \beta_2 Z_{TX}$$

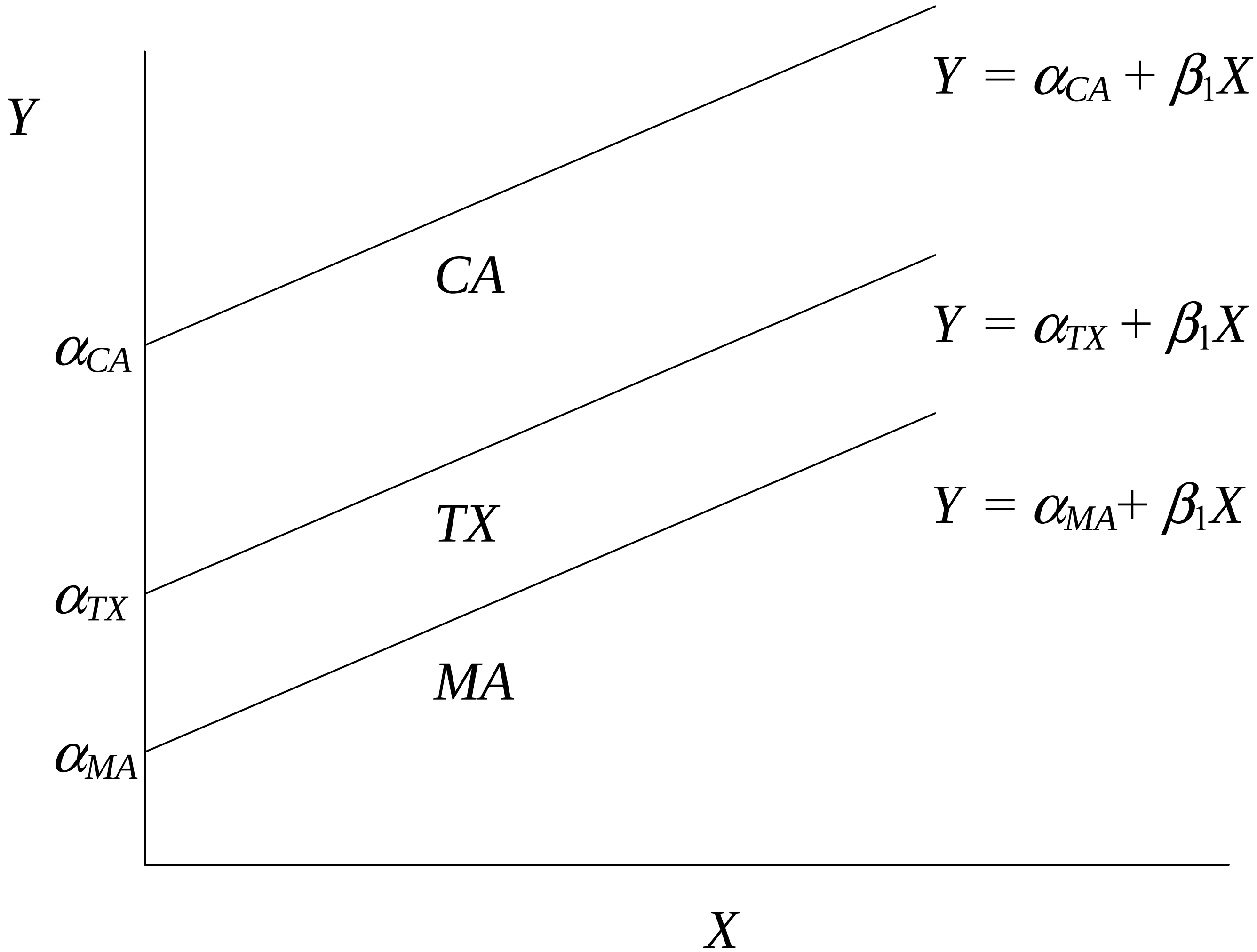
Collecting the lines for all three states:

$$Y_{CA,t} = \alpha_{CA} + \beta_1 X_{CA,t} + u_{CA,t}$$
$$Y_{TX,t} = \alpha_{TX} + \beta_1 X_{TX,t} + u_{TX,t}$$
$$Y_{MA,t} = \alpha_{MA} + \beta_1 X_{MA,t} + u_{MA,t}$$

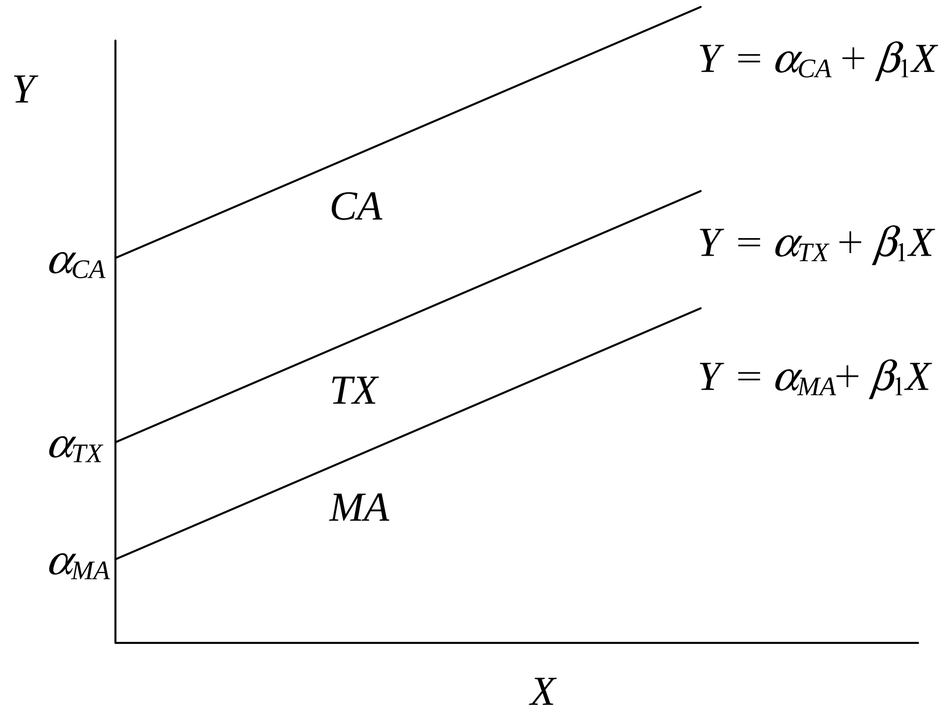
or

$$Y_{it} = \alpha_i + \beta_1 X_{it} + u_{it},\ i = \text{CA, TX, MA},\ T = 1,\ldots,T$$

The regression lines for each state in a picture



Recall (Fig. 6.8a) that shifts in the intercept can be represented using binary regressors…



In binary regressor form:

$$Y_{it} = \beta_0 + \gamma_{CA}DCA_i + \gamma_{TX}DTX_i + \beta_1 X_{it} + u_{it}$$

- DCA_i = 1 if state is *CA*, = 0 otherwise
- DTX_t = 1 if state is *TX*, = 0 otherwise
- leave out DMA_i (*why*?)

Summary: Two ways to write the fixed effects model

"*n*-1 binary regressor" form

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \ldots + \gamma_n Dn_i + u_i$$

where $D2_i = \begin{cases} 1 \text{ for } i\text{=2 (state \#2)} \\ 0 \text{ otherwise} \end{cases}$, etc.

"Fixed effects" form:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_i$$

- α_i is called a "state fixed effect" or "state effect" – it is the constant (fixed) effect of being in state *i*

Fixed Effects Regression: Estimation

Three estimation methods:

1. "*n*-1 binary regressors" OLS regression
2. "Entity-demeaned" OLS regression
3. "Changes" specification (only works for $T = 2$)

- These three methods produce identical estimates of the regression coefficients, and identical standard errors.
- We already did the "changes" specification (1988 minus 1982) – but this only works for $T = 2$ years
- Methods #1 and #2 work for general T
- Method #1 is only practical when n isn't too big

1. "*n*-1 binary regressors" OLS regression

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \ldots + \gamma_n Dn_i + u_i \qquad (1)$$

where $D2_i = \begin{cases} 1 \text{ for } i\text{=2 (state \#2)} \\ 0 \text{ otherwise} \end{cases}$ etc.

- First create the binary variables $D2_i,\ldots,Dn_i$
- Then estimate (1) by OLS
- Inference (hypothesis tests, confidence intervals) is as usual (using heteroskedasticity-robust standard errors)
- This is impractical when n is very large (for example if n = 1000 workers)

2. "Entity-demeaned" OLS regression

The fixed effects regression model:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_i$$

The state averages satisfy:

$$\frac{1}{T}\sum_{t=1}^{T} Y_{it} = \alpha_i + \beta_1 \frac{1}{T}\sum_{t=1}^{T} X_{it} + \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

Deviation from state averages:

$$Y_{it} - \frac{1}{T}\sum_{t=1}^{T} Y_{it} = \beta_1 \left(X_{it} - \frac{1}{T}\sum_{t=1}^{T} X_{it} \right) + \left(u_{it} - \frac{1}{T}\sum_{t=1}^{T} u_{it} \right)$$

Entity-demeaned OLS regression, ctd.

$$Y_{it} - \frac{1}{T}\sum_{t=1}^{T} Y_{it} = \beta_1\left(X_{it} - \frac{1}{T}\sum_{t=1}^{T} X_{it} \right) + \left(u_{it} - \frac{1}{T}\sum_{t=1}^{T} u_{it} \right)$$

or

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$$

where $\tilde{Y}_{it} = Y_{it} - \frac{1}{T}\sum_{t=1}^{T} Y_{it}$ and $\tilde{X}_{it} = X_{it} - \frac{1}{T}\sum_{t=1}^{T} X_{it}$

- For i=1 and t = 1982, $\tilde{Y}_{it}$ is the difference between the fatality rate in Alabama in 1982, and its average value in Alabama averaged over all 7 years.

Entity-demeaned OLS regression, ctd.

$$\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it} \qquad (2)$$

where $\tilde{Y}_{it} = Y_{it} - \frac{1}{T}\sum_{t=1}^{T} Y_{it}$, etc.

- First construct the demeaned variables $\tilde{Y}_{it}$ and $\tilde{X}_{it}$
- Then estimate (2) by regressing $\tilde{Y}_{it}$ on $\tilde{X}_{it}$ using OLS
- Inference (hypothesis tests, confidence intervals) is as usual (using heteroskedasticity-robust standard errors)
- This is like the "changes" approach, but instead Y_{it} is deviated from the state average instead of Y_{i1}.
- This can be done in a single command in STATA

Example: Traffic deaths and beer taxes in STATA

```
. areg vfrall beertax, absorb(state) r;

Regression with robust standard errors                 Number of obs =     336
                                                       F(  1,    287) =   10.41
                                                       Prob > F      =  0.0014
                                                       R-squared     =  0.9050
                                                       Adj R-squared =  0.8891
                                                       Root MSE      =  .18986

------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6558736   .2032797    -3.23   0.001    -1.055982   -.2557655
       _cons |   2.377075   .1051515    22.61   0.000     2.170109    2.584041
-------------+----------------------------------------------------------------
       state |   absorbed                                      (48 categories)
```

- "areg" automatically de-means the data
- this is especially useful when *n* is large
- the reported intercept is arbitrary

***Example*, ctd.**

For $n = 48$, $T = 7$:

$$\widehat{FatalityRate} = -.66BeerTax + State\ fixed\ effects$$
$$(.20)$$

- Should you report the intercept?
- How many binary regressors would you include to estimate this using the "binary regressor" method?
- Compare slope, standard error to the estimate for the 1988 v. 1982 "changes" specification ($T = 2$, $n = 48$):

$$\widehat{FR_{1988} - FR_{1982}} = -.072 - 1.04(BeerTax_{1988} - BeerTax_{1982})$$
$$(.065) \quad (.36)$$

Regression with Time Fixed Effects
(SW Section 10.4)

An omitted variable might vary over time but not across states:

- Safer cars (air bags, etc.); changes in national laws
- These produce intercepts that change over time
- Let these changes ("safer cars") be denoted by the variable S_t, which changes over time but not states.
- The resulting population regression model is:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 S_t + u_{it}$$

Time fixed effects only

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_3 S_t + u_{it}$$

In effect, the intercept varies from one year to the next:

$$\begin{aligned} Y_{i,1982} &= \beta_0 + \beta_1 X_{i,1982} + \beta_3 S_{1982} + u_{i,1982} \\ &= (\beta_0 + \beta_3 S_{1982}) + \beta_1 X_{i,1982} + u_{i,1982} \end{aligned}$$

or

$$Y_{i,1982} = \mu_{1982} + \beta_1 X_{i,1982} + u_{i,1982}, \quad \mu_{1982} = \beta_0 + \beta_3 S_{1982}$$

Similarly,

$$Y_{i,1983} = \mu_{1983} + \beta_1 X_{i,1983} + u_{i,1983}, \quad \mu_{1983} = \beta_0 + \beta_3 S_{1983}$$

etc.

Two formulations for time fixed effects

1. "Binary regressor" formulation:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \delta_2 B2_t + \ldots\ \delta_T BT_t + u_{it}$$

where $B2_t = \begin{cases} 1 \text{ when } t\text{=2 (year \#2)} \\ 0 \text{ otherwise} \end{cases}$, etc.

2. "Time effects" formulation:

$$Y_{it} = \beta_1 X_{it} + \mu_t + u_{it}$$

Time fixed effects: estimation methods

1. "*T*-1 binary regressors" OLS regression

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \delta_2 B2_{it} + \ldots \; \delta_T BT_{it} + u_{it}$$

- Create binary variables *B2*,…,*BT*
- *B2* = 1 if *t* = year #2, = 0 otherwise
- Regress *Y* on *X*, *B2*,…,*BT* using OLS
- Where's *B1*?

2. "Year-demeaned" OLS regression
   - Deviate Y_{it}, X_{it} from *year* (not state) averages
   - Estimate by OLS using "year-demeaned" data

State *and* Time Fixed Effects

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_i + \beta_3 S_t + u_{it}$$

1. "Binary regressor" formulation:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_2 D2_i + \ldots + \gamma_n Dn_i$$
$$+ \delta_2 B2_t + \ldots \delta_T BT_t + u_{it}$$

2. "State and time effects" formulation:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \mu_t + u_{it}$$

State *and* time effects: estimation methods

1. "*n*-1 and *T*-1 binary regressors" OLS regression
   - Create binary variables *D2*,…,*Dn*
   - Create binary variables *B2*,…,*BT*
   - Regress *Y* on *X*, *D2*,…,*Dn*, *B2*,…,*BT* using OLS
   - What about *D1* and *B1*?
2. "State- and year-demeaned" OLS regression
   - Deviate Y_{it}, X_{it} from year *and* state averages
   - Estimate by OLS using "year- and state-demeaned" data

These two methods can be combined too.

STATA example: Traffic deaths…

```
. gen y83=(year==1983);
. gen y84=(year==1984);
. gen y85=(year==1985);
. gen y86=(year==1986);
. gen y87=(year==1987);
. gen y88=(year==1988);
. areg vfrall beertax y83 y84 y85 y86 y87 y88, absorb(state) r;

Regression with robust standard errors                 Number of obs =      336
                                                       F(  7,    281) =     3.70
                                                       Prob > F      =   0.0008
                                                       R-squared     =   0.9089
                                                       Adj R-squared =   0.8914
                                                       Root MSE      =   .18788
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6399799   .2547149    -2.51   0.013    -1.141371   -.1385884
         y83 |  -.0799029   .0502708    -1.59   0.113    -.1788579    .0190522
         y84 |  -.0724206   .0452466    -1.60   0.111     -.161486    .0166448
         y85 |  -.1239763   .0460017    -2.70   0.007     -.214528   -.0334246
         y86 |  -.0378645   .0486527    -0.78   0.437    -.1336344    .0579055
         y87 |  -.0509021   .0516113    -0.99   0.325    -.1524958    .0506917
         y88 |  -.0518038     .05387    -0.96   0.337    -.1578438    .0542361
       _cons |    2.42847   .1468565    16.54   0.000     2.139392    2.717549
-------------+----------------------------------------------------------------
       state |   absorbed                                     (48 categories)
```

Go to section for other ways to do this in STATA!

Some Theory: The Fixed Effects Regression Assumptions (SW App. 10.2)

For a single X:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it},\ \ i = 1,\ldots,n,\ t = 1,\ldots,\ T$$

1. $E(u_{it}|X_{i1},\ldots,X_{iT},\alpha_i) = 0$.
2. $(X_{i1},\ldots,X_{iT},Y_{i1},\ldots,Y_{iT})$, $i = 1,\ldots,n$, are i.i.d. draws from their joint distribution.
3. (X_{it}, u_{it}) have finite fourth moments.
4. There is no perfect multicollinearity (multiple X's)
5. $\text{corr}(u_{it},u_{is}|X_{it},X_{is},\alpha_i) = 0$ for $t \neq s$.

Assumptions 3&4 are identical; 1, 2, differ; 5 is new

Assumption #1: $E(u_{it}|X_{i1},\ldots,X_{iT},\alpha_i) = 0$

- u_{it} has mean zero, given the state fixed effect *and* the entire history of the X's for that state
- This is an extension of the previous multiple regression Assumption #1
- This means there are no omitted lagged effects (any lagged effects of X must enter explicitly)
- Also, there is not feedback from u to future X:
  - Whether a state has a particularly high fatality rate this year doesn't subsequently affect whether it increases the beer tax.
  - We'll return to this when we take up time series data.

Assumption #2: $(X_{i1},\ldots,X_{iT},Y_{i1},\ldots,Y_{iT})$, $i = 1,\ldots,n$, are i.i.d. draws from their joint distribution.

- This is an extension of Assumption #2 for multiple regression with cross-section data
- This is satisfied if entities (states, individuals) are randomly sampled from their population by simple random sampling, then data for those entities are collected over time.
- This does ***not*** require observations to be i.i.d. over *time* for the same entity – that would be unrealistic (whether a state has a mandatory DWI sentencing law this year is strongly related to whether it will have that law next year).

Assumption #5: $\text{corr}(u_{it},u_{is}|X_{it},X_{is},\alpha_i) = 0$ for $t \neq s$

- This is new.
- This says that (given X), the error terms are uncorrelated over time within a state.
- For example, $u_{CA,1982}$ and $u_{CA,1983}$ are uncorrelated
- Is this plausible? What enters the error term?
  - Especially snowy winter
  - Opening major new divided highway
  - Fluctuations in traffic density from local economic conditions
- Assumption #5 requires these omitted factors entering u_{it} to be uncorrelated over time, within a state.

What if Assumption #5 fails: corr(u_{it},u_{is}|X_{it},X_{is},α_i) $\neq$0?

- A useful analogy is heteroskedasticity.
- OLS panel data estimators of β_1 are unbiased, consistent
- The OLS standard errors will be wrong – usually the OLS standard errors understate the true uncertainty
- Intuition: if u_{it} is correlated over time, you don't have as much information (as much random variation) as you would were u_{it} uncorrelated.
- This problem is solved by using "heteroskedasticity and autocorrelation-consistent standard errors" – we return to this when we focus on time series regression

Application: Drunk Driving Laws and Traffic Deaths (SW Section 10.5)

Some facts

- Approx. 40,000 traffic fatalities annually in the U.S.
- 1/3 of traffic fatalities involve a drinking driver
- 25% of drivers on the road between 1am and 3am have been drinking (estimate)
- A drunk driver is 13 times as likely to cause a fatal crash as a non-drinking driver (estimate)

Drunk driving laws and traffic deaths, ctd.

Public policy issues

- Drunk driving causes massive externalities (sober drivers are killed, etc. etc.) – there is ample justification for governmental intervention
- Are there any effective ways to reduce drunk driving? If so, what?
- What are effects of specific laws:
  - mandatory punishment
  - minimum legal drinking age
  - economic interventions (alcohol taxes)

The drunk driving panel data set

n = 48 U.S. states, T = 7 years (1982,…,1988) (balanced)

Variables

- Traffic fatality rate (deaths per 10,000 residents)
- Tax on a case of beer (*Beertax*)
- Minimum legal drinking age
- Minimum sentencing laws for first DWI violation:
  - *Mandatory Jail*
  - *Manditory Community Service*
  - otherwise, sentence will just be a monetary fine
- Vehicle miles per driver (US DOT)
- State economic data (real per capita income, etc.)

Why might panel data help?

- Potential OV bias from variables that vary across states but are constant over time:
  - culture of drinking and driving
  - quality of roads
  - vintage of autos on the road
    - use state fixed effects
- Potential OV bias from variables that vary over time but are constant across states:
  - improvements in auto safety over time
  - changing national attitudes towards drunk driving
    - use time fixed effects

TABLE 8.1 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)
Beer tax	0.36** (0.05)	−0.66** (0.20)	−0.64* (0.25)	−0.45* (0.22)	−0.70** (0.25)	−0.46* (0.22)
Drinking age 18				0.028 (0.066)	−0.011 (0.064)	
Drinking age 19				−0.019 (0.040)	−0.078 (0.049)	
Drinking age 20				0.031 (0.046)	−0.102* (0.046)	
Drinking age						−0.002 (0.017)
Mandatory jail?				0.013 (0.032)	−0.026 (0.065)	
Mandatory community service?				0.033 (0.115)	0.147 (0.137)	
Mandatory jail or community service?						0.031 (0.076)
Average vehicle miles per driver				0.008 (0.008)	0.017 (0.010)	0.009 (0.008)
Unemployment rate				−0.063** (0.012)		−0.063** (0.012)
Real income per capita (logarithm)				1.81** (0.47)		1.79** (0.45)
State effects?	no	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes

These regressions were estimated using panel data for 48 U.S. states from 1982 to 1988 (336 observations total), described in Appendix 8.1. 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 *5% level or **1% significance level.

TABLE 8.1 Regression Analysis of the Effect of Drunk Driving Laws on Traffic Deaths

Dependent Variable: Traffic Fatality Rate (Deaths Per 10,000).

	(1)	(2)	(3)	(4)	(5)	(6)
***F*-statistics and *p*-values Testing Exclusion of Groups of Variables:**						
Time effects = 0			2.47 (0.024)	11.44 (<0.001)	2.28 (0.037)	11.59 (<0.001)
Drinking age coefficients = 0				0.48 (0.696)	2.09 (0.102)	
Jail, community service coefficients = 0				0.17 (0.845)	0.59 (0.557)	
Unemployment rate, income per capita = 0				38.29 (<0.001)		40.12 (<0.001)
$\overline{R}^2$	0.090	0.889	0.891	0.926	0.893	0.926

These regressions were estimated using panel data for 48 U.S. states from 1982 to 1988 (336 observations total), described in Appendix 8.1. 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 *5% level or **1% significance level.

Empirical Analysis: Main Results

- Sign of beer tax coefficient changes when fixed state effects are included
- Fixed time effects are statistically significant but do not have big impact on the estimated coefficients
- Estimated effect of beer tax drops when other laws are included as regressor
- The only policy variable that seems to have an impact is the tax on beer – not minimum drinking age, not mandatory sentencing, etc.
- The other economic variables have plausibly large coefficients: more income, more driving, more deaths

Extensions of the "*n*-1 binary regressor" approach

The idea of using many binary indicators to eliminate omitted variable bias can be extended to non-panel data – the key is that the omitted variable is constant for a group of observations, so that in effect it means that each group has its own intercept.

Example: Class size problem.

Suppose funding and curricular issues are determined at the county level, and each county has several districts. Resulting omitted variable bias could be addressed by including binary indicators, one for each county (omit one to avoid perfect multicollinearity).

Summary: Regression with Panel Data
(SW Section 10.6)

Advantages and limitations of fixed effects regression

Advantages

- You can control for unobserved variables that:
  - vary across states but not over time, and/or
  - vary over time but not across states
- More observations give you more information
- Estimation involves relatively straightforward extensions of multiple regression

- Fixed effects estimation can be done three ways:
  1. "Changes" method when $T = 2$
  2. "n-1 binary regressors" method when n is small
  3. "Entity-demeaned" regression
- Similar methods apply to regression with time fixed effects and to both time and state fixed effects
- Statistical inference: like multiple regression.

Limitations/challenges

- Need variation in X over time within states
- Time lag effects can be important
- Standard errors might be too low (errors might be correlated over time)