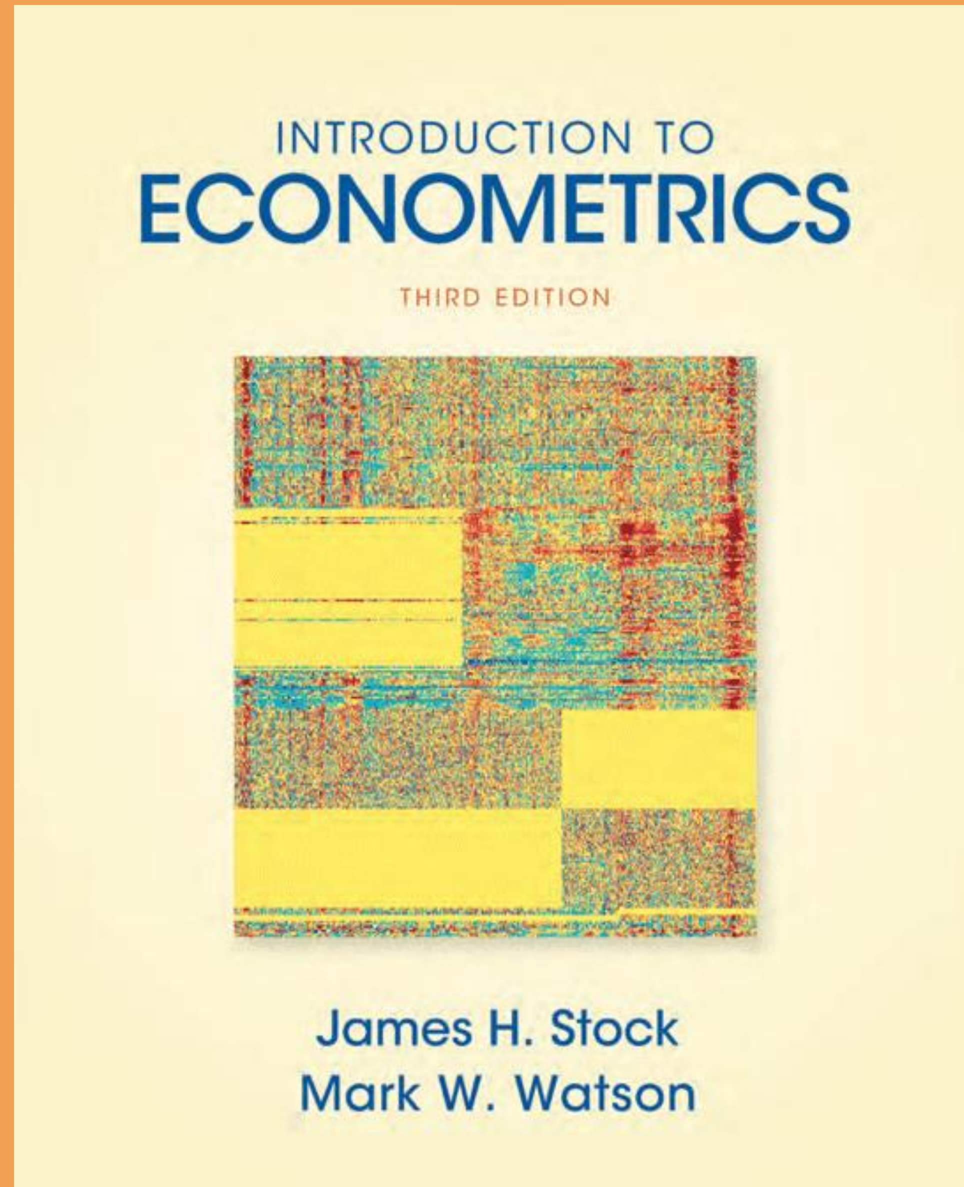
Chapter 10

- Regression with Panel Data
- 面板数据回归



Addison-Wesley
is an imprint of
PEARSON

Regression with Panel Data
(SW Chapter 10)

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

面板数据包含多个个体，每个个体在不同时间上有多个观测值

Hypothetical examples:

- Data on 420 California school districts in 1999 *and again* in 2000, for 840 observations total.
  1990 年和 2000 年加利福尼亚州的 420 个校区，共 840 个观测值
- Data on 50 U.S. states, each state is observed in 3 years, for a total of 150 observations.
  美国 50 个州，每个州有 3 年的观测值，共 150 个
- Data on 1000 individuals, in four different months, for 4000 observations total.
  1000 个个体，每个个体观测 4 个月，共 4000 个观测值

Notation for panel data
面板数据符号

A double subscript distinguishes entities (states) and time periods (years)
用两个下标来区分个体和时间

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:

有 k 个回归因子的面板数据

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

n = number of entities (states) 个体数量

T = number of time periods (years) 时间期数

Some jargon…

- Another term for panel data is ***longitudinal data***
- ***balanced panel***: no missing observations (all variables are observed for all entites [states] and all time periods [years])
  平衡面板:没有缺失值(所有变量在任何个体和时间上的数据都完整)

基本上，实证研究要考虑怎样确定面板

Why are panel data useful?

With panel data we can control for factors that:

利用面板数据我们可以控制以下因素:

- Vary across entities (states) but do not vary over time
  随个体变化，但不随时间变化的因素 不随时间变化的都被控制啦
- Could cause omitted variable bias if they are omitted
  可能导致遗漏变量偏误的因素
- are unobserved or unmeasured – and therefore cannot be included in the regression using multiple regression
  不可观察或不可测的因素——因此无法使用多元回归

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.

如果一个遗漏变量不随时间变化，那么 Y 随时间变化的成分就不可能由该因素引起。

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

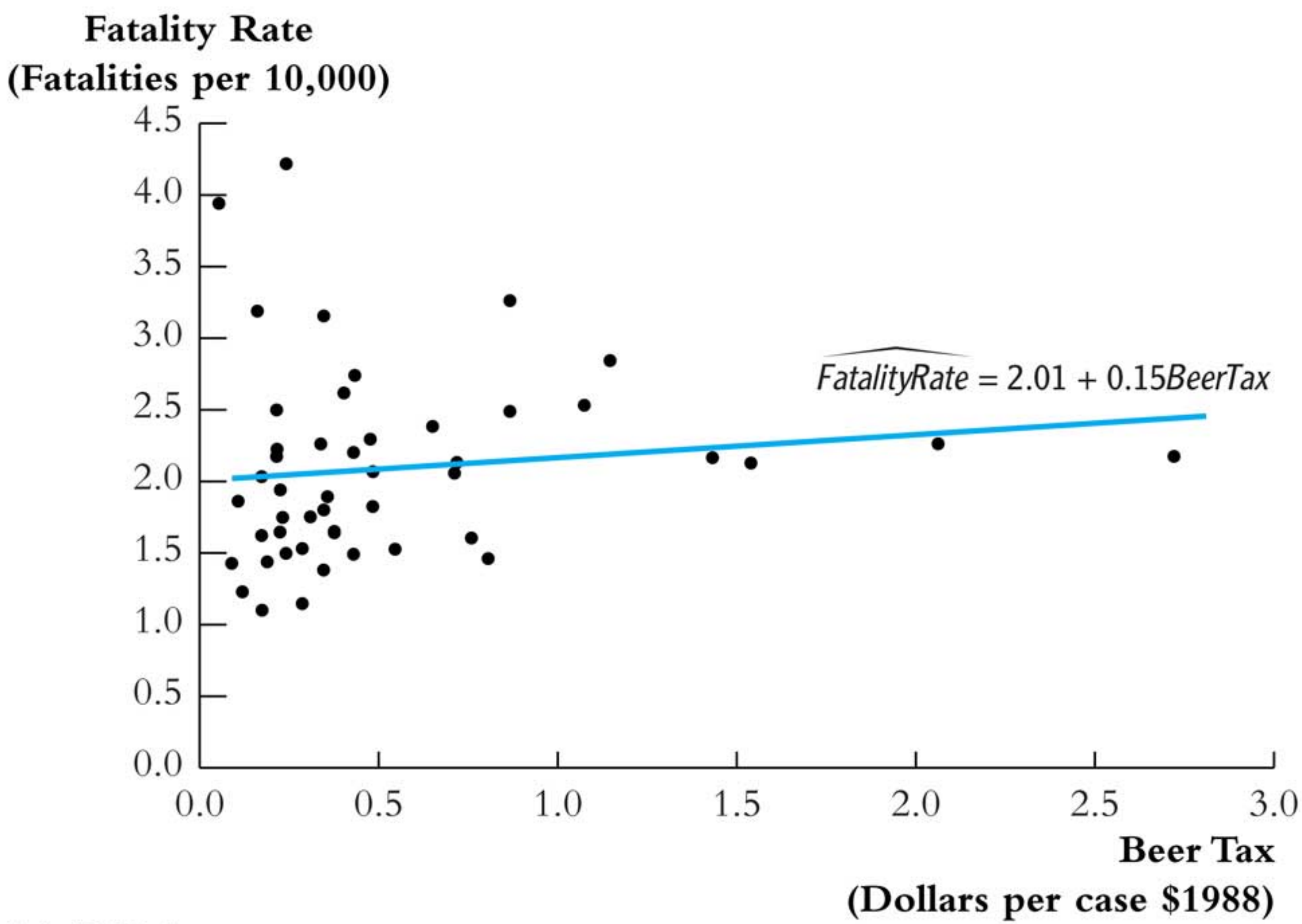
Observational unit: a year in a U.S. state

- 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: 变量

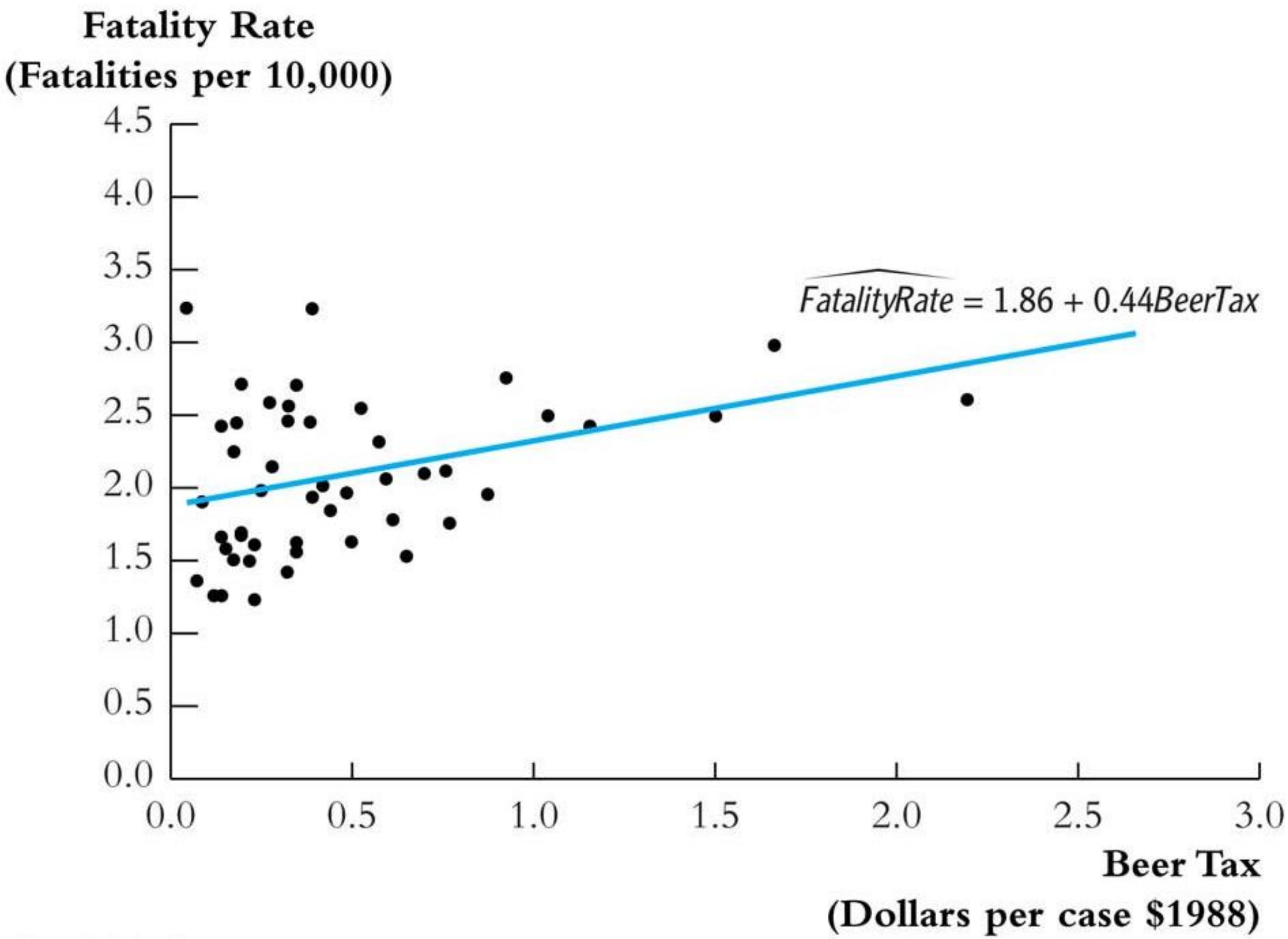
- Traffic fatality rate (# traffic deaths in that state in that year, per 10,000 state residents)
  交通事故死亡率(某个州某年每 10000 人中因交通事故死亡的人数
- Tax on a case of beer 啤酒税
- Other (legal driving age, drunk driving laws, etc.)
  其他因素(法律驾驶年龄，醉酒驾车法律等)

U.S. traffic death data for 1982:



Higher alcohol taxes, more traffic deaths?
酒精税越高，交通死亡率越高？

U.S. traffic death data for 1988



Higher alcohol taxes, more traffic deaths?

Why might there be higher *more* traffic deaths in states that have higher alcohol taxes?
为什么酒精税越高的州，交通死亡率可能越高?

Other factors that determine traffic fatality rate:

其他影响交通死亡率的因素

- Quality (age) of automobiles 汽车质量（使用年份）
- Quality of roads 道路状况
- “Culture” around drinking and driving 饮酒或驾车”文化”
- Density of cars on the road 道路上的汽车密度

These omitted factors could cause omitted variable bias.
这些遗漏变量可能导致遗漏变量偏误

Example #1: traffic density. Suppose:

例 1：交通密度。假设：

(i) High traffic density means more traffic deaths

更高的交通密度意味着更高的交通事故死亡率

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

（西部的）具有较低交通密度的州执行较低的酒精税

(cont.) These omitted factors could cause omitted variable bias.

Example #1: traffic density. Suppose:

- 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)
  从而遗漏变量偏误的两个条件都满足。具体而言，"高税率"可能导致"高交通密度"（因此 OLS 回归系数可能偏大，即高税率，高死亡率）
- Panel data lets us eliminate omitted variable bias when the omitted variables are constant over time within a given state.
  当这类遗漏变量在给定的个体中不随时间而变化时，面板数据让我们可以消除这类遗漏变量偏误。

Example #2: cultural attitudes towards drinking and driving:
例2: 饮酒和驾驶“文化”

(i) arguably are a determinant of traffic deaths; and
可论证饮酒和驾驶“文化”与交通事故死亡率相关

(ii) potentially are correlated with the beer tax, so beer taxes could be picking up cultural differences (omitted variable bias).
文化因素可能与啤酒税相关，因此啤酒税反映的可能是文化上的差异（遗漏变量偏误）

Example #2: cultural attitudes towards drinking and driving:
例2: 饮酒和驾驶“文化”

- 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)
  因此导致遗漏变量偏误的两个条件满足。具体而言，“高税率”可能反映了“饮酒的文化”（因此 OLS 回归系数是有偏的）
- Panel data lets us eliminate omitted variable bias when the omitted variables are constant over time within a given state.
  当这类遗漏变量在给定的个体中不随时间而变化时，面板数据可以消除这类遗漏变量偏误。

Panel Data with Two Time Periods (SW Section 10.2)
两期的面板数据

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.

Z_i是一个不随时间而改变的因素（密度），至少在我们所考虑的时间范围内

- Suppose Z_i is not observed, so its omission could result in omitted variable bias.

  假设 Z_i不可观察，那么遗漏它会导致遗漏变量偏误
- The effect of Z_i can be eliminated using $T = 2$ years.

  当使用两期的数据，Z_i的影响可以被消除

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.

从 1982 年到 1988 年间死亡率的任何变化都不可能由 Z_i 引起，因为根据假设，Z_i 在这两年间并没有发生改

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…

两年数据相减便可消除 Z_i 的影响

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

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

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 $BeerTax_{i1982}$.
  新的遗漏变量与 $BeerTax_{i1988}$ 和 $BeerTax_{i1982}$ 都不相关
- This “difference” equation can be estimated by OLS, even though Z_i isn’t observed.
  即使 Z_i 不可观察，“不同之处”仍可以用 OLS 估计
- The omitted variable Z_i doesn’t change, so it cannot be a determinant of the *change* in Y
  由于 Z_i 没有发生改变，因此它不可能导致 Y 的变化

Example: Traffic deaths and beer taxes

1982 data:

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

1988 data:

这两个方程估计出来都是有排偏的。

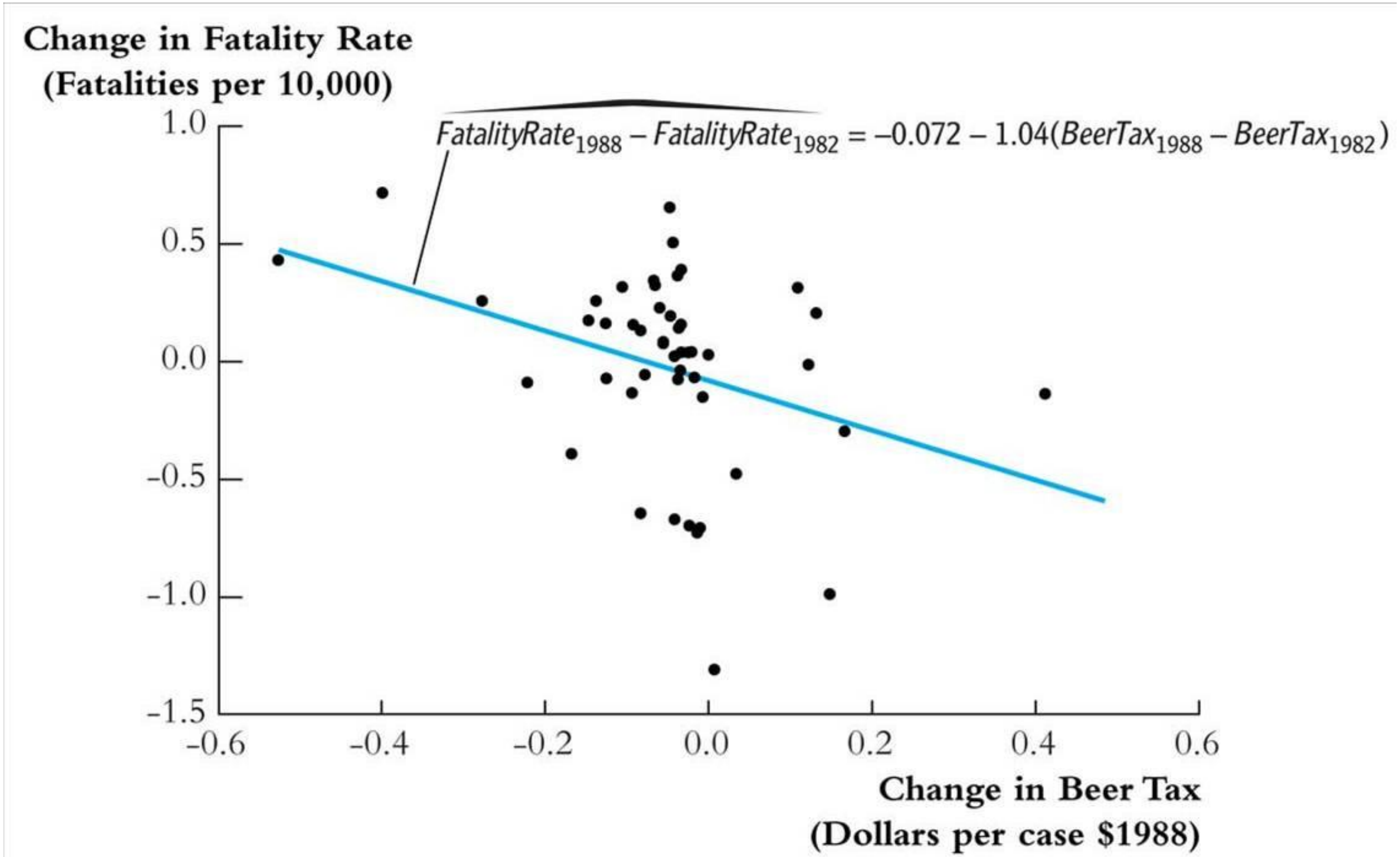
$$\widehat{FatalityRate} = 1.86 + 0.44BeerTax \qquad (n = 48)$$
$$\qquad (.11) \quad (.13)$$

Difference regression (n = 48)

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

Δ*FatalityRate* v. Δ*BeerTax*:

这就很棒了



Fixed Effects Regression 固定效应回归

(SW Section 10.3)

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

如果你拥有多于 2 期的数据?

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

We can rewrite this in two useful ways:

1. "*n*-1 binary regressor" regression model

   "n-1 个虚拟变量" 回归模型

2. "Fixed Effects" regression model

   "固定效应" 回归模型

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

我们首先重写为 "固定效应" 模型的形式

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

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

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

这部分和时间没有关系，标黄的，是扰动项

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

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.

  斜率对 CA 而言是特定的，但斜率对于所有州都是相同的：一组平行的直线

  重要

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

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

阿尔法就是把截距和扰动项合并在一起了

Collecting the lines for all three states:

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

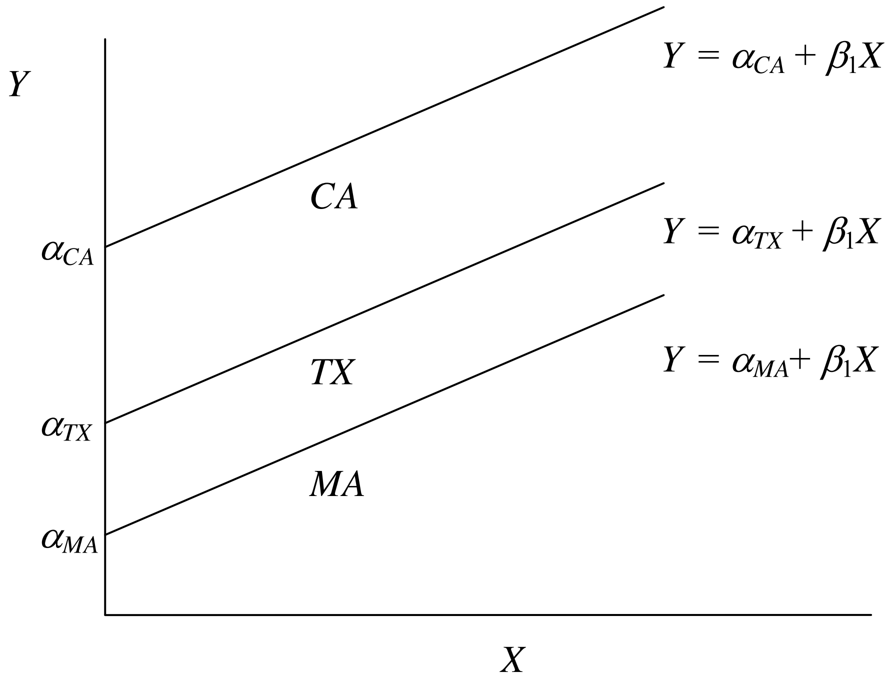
$$Y_{TX,t} = \alpha_{TX} + \beta_1 X_{TX,t} + u_{TX,t}$$

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

or

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

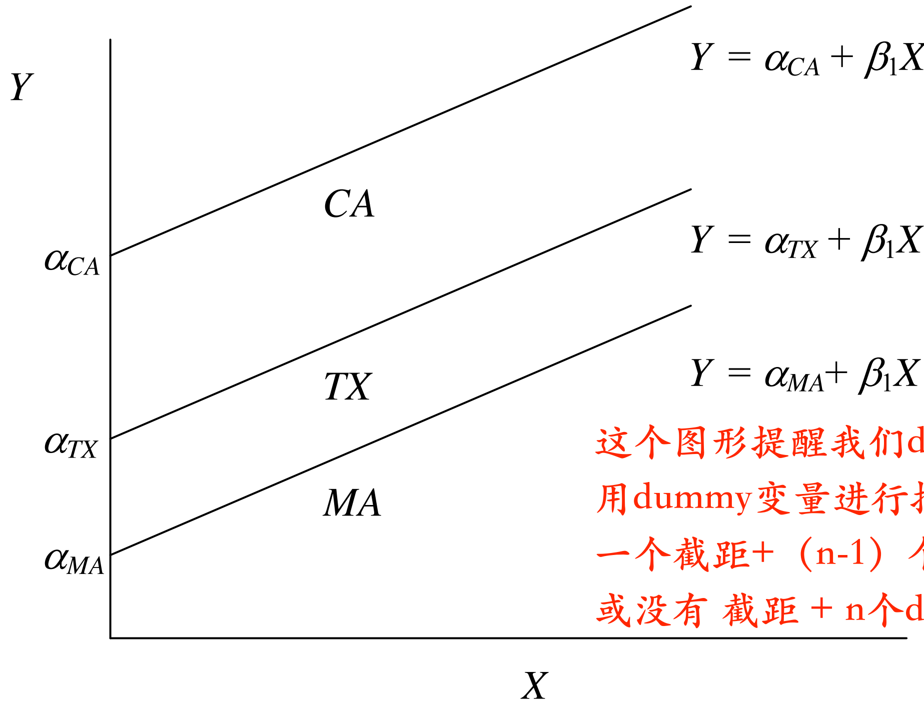
The regression lines for each state in a picture



截距反应不同的州的一些特定的东西

Recall that shifts in the intercept can be represented using binary regressors…

回忆：采用虚拟变量可以表示截距之间的差异



这个图形提醒我们dummy 变量的特点，这里也可以用dummy变量进行描述

一个截距+（n-1）个dummy

或没有 截距 + n个dummy

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

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

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

"Fixed effects" form:

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

- α_i is called a "state fixed effect" or "state effect" – it is the constant (fixed) effect of being in state *i*
  α_i称作"州固定效应"或者"州效应"，即它是在州 i 特有的固定效应

Fixed Effects Regression: Estimation
固定效应回归模型的估计

Three estimation methods:
有 3 种估计方法：

1. "*n*-1 binary regressors" OLS regression
   "n-1 个二元变量"的 OLS 回归
2. "Entity-demeaned" OLS regression
   "个体中心化"OLS 回归
3. "Changes" specification, without an intercept (only works for $T = 2$)
   "前后变化"模型，不包含截距项（只适合于 T=2）

(cont.)Fixed Effects Regression: Estimation

- These three methods produce identical estimates of the regression coefficients, and identical standard errors.
  这三种方法得到的回归系数和标准误都是相同的

- We already did the “changes” specification (1988 minus 1982) – but this only works for $T = 2$ years
  我们已经采用了“前后变化”模型（用 1988 年的数据减去 1982 年的数据）；但这种方法只适合于 T=2 的情况

- Methods #1 and #2 work for general T
  方法 1 和方法 2 适合更一般化的 T

- Method #1 is only practical when n isn’t too big
  方法 1 在实际运用中适合 N 不是很大的情况

1. "n-1 binary regressors" OLS regression

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

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

- First create the binary variables $D2_i,\ldots,Dn_i$
  首先创建虚拟变量
- Then estimate (1) by OLS 然后用 OLS 方法估计方程(1)
- Inference (hypothesis tests, confidence intervals) is as usual (using heteroskedasticity-robust standard errors)
  统计推断（假设检验，置信区间）仍适用（使用异方差标准误）
- This is impractical when n is very large (for example if n = 1000 workers)
  当 n 很大时，这种方法不适用 要定义太多dummy

2. "Entity-demeaned" OLS regression

The fixed effects regression model:

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

The state averages satisfy:

每个州时间内的均值满足

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

Deviation from state averages:

偏离均值

这样就把阿尔法减掉了，而且没有drop任何一年的数据

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

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.
  $\tilde{Y}_{it}$是 1982 年 Alabama 和死亡率和其在整个 7 年间平均水平之间的差异

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)
  统计推断（假设检验和致信区间）仍适用（使用异方差标准误）
- This is like the “changes” approach, but instead Y_{it} is deviated from the state average instead of Y_{i1}.
  这与“前后变化”方法类似，但采用的是 Y_{it} 与均值的偏离
- This can be done in a single command in STATA
  在 STATA 中一条命令即可完成

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 Y X　用state做ID来回归

- "areg" automatically de-means the data
  "areg"自动将数据减去均值
- this is especially useful when *n* is large
- the reported intercept is arbitrary

Example, ctd. For n = 48, T = 7:

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

- Should you report the intercept?
  是否应该报告截距项？
- How many binary regressors would you include to estimate this using the "binary regressor" method?
  需要采用多少个虚拟变量来估计?
- Compare slope, standard error to the estimate for the 1988 v. 1982 "changes" specification (T = 2, n = 48) (*note that this includes an intercept – return to this below*):
  比较斜率、标准误与"前后变化"模型中的差异（注意，后一个模型有截距项）

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

By the way… how much do beer taxes vary?

Beer Taxes in 2005

Source: Federation of Tax Administrators

http://www.taxadmin.org/fta/rate/beer.html

	EXCISE TAX RATES ($ per gallon)	SALES TAXES APPLIED	OTHER TAXES
Alabama	$0.53	Yes	$0.52/gallon local tax
Alaska	1.07	n.a.	$0.35/gallon small breweries
Arizona	0.16	Yes	
Arkansas	0.23	Yes	under 3.2% - $0.16/gallon; $0.008/gallon and 3% off- 10% on-premise tax
California	0.20	Yes	
Colorado	0.08	Yes	
Connecticut	0.19	Yes	
Delaware	0.16	n.a.	
Florida	0.48	Yes	2.67¢/12 ounces on-premise retail tax

Georgia	0.48	Yes	$0.53/gallon local tax
Hawaii	0.93	Yes	$0.54/gallon draft beer
Idaho	0.15	Yes	over 4% - $0.45/gallon
Illinois	0.185	Yes	$0.16/gallon in Chicago and $0.06/gallon in Cook County
Indiana	0.115	Yes	
Iowa	0.19	Yes	
Kansas	0.18	--	over 3.2% - {8% off- and 10% on-premise}, under 3.2% - 4.25% sales tax.
Kentucky	0.08	Yes*	9% wholesale tax
Louisiana	0.32	Yes	$0.048/gallon local tax
Maine	0.35	Yes	additional 5% on-premise tax
Maryland	0.09	Yes	$0.2333/gallon in Garrett County
Massachusetts	0.11	Yes*	0.57% on private club sales
Michigan	0.20	Yes	
Minnesota	0.15	--	under 3.2% - $0.077/gallon. 9% sales tax
Mississippi	0.43	Yes	
Missouri	0.06	Yes	
Montana	0.14	n.a.	
Nebraska	0.31	Yes	
Nevada	0.16	Yes	
New Hampshire	0.30	n.a.	
New Jersey	0.12	Yes	
New Mexico	0.41	Yes	

New York	0.11	Yes $0.12/gallon in New York City
North Carolina	0.53	Yes $0.48/gallon bulk beer
North Dakota	0.16	-- 7% state sales tax, bulk beer $0.08/gal.
Ohio	0.18	Yes
Oklahoma	0.40	Yes under 3.2% - $0.36/gallon; 13.5% on-premise
Oregon	0.08	n.a.
Pennsylvania	0.08	Yes
Rhode Island	0.10	Yes $0.04/case wholesale tax
South Carolina	0.77	Yes
South Dakota	0.28	Yes
Tennessee	0.14	Yes 17% wholesale tax
Texas	0.19	Yes over 4% - $0.198/gallon, 14% on-premise and $0.05/drink on airline sales
Utah	0.41	Yes over 3.2% - sold through state store
Vermont	0.265	no 6% to 8% alcohol - $0.55; 10% on-premise sales tax
Virginia	0.26	Yes
Washington	0.261	Yes
West Virginia	0.18	Yes
Wisconsin	0.06	Yes
Wyoming	0.02	Yes
Dist. of Columbia	0.09	Yes 8% off- and 10% on-premise sales tax
U.S. Median	$0.188	

Regression with Time Fixed Effects (SW Section 10.4)
时间固定效应回归

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

遗漏变量可能随时间变化，但不随州变化

- Safer cars (air bags, etc.); changes in national laws

  更安全的汽车；国家法律的变化

- These produce intercepts that change over time

  这些会导致截距项随时间变化

- Let these changes ("safer cars") be denoted by the variable S_t, which changes over time but not states.

  用 S_t 表示这些改变，它会随时间变化，但不随州发生变化

- 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:
事实上，截距项每年都在发生变化

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

or

不同年份，不同的表达式

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

Similarly,

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

etc.

Two formulations for time fixed effects 时间固定效应的两种建模

1. “*T*-1 binary regressor” formulation:

   “T-1 个虚拟变量”

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

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

2. “Time effects” formulation:

   “时间效应”模型

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

Time fixed effects: estimation methods
时间效应：估计方法

虚拟变量和areg两种方法要求掌握，其余不用

1. "*T*-1 binary regressor" OLS regression
   "T-1 个虚拟变量"

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

   - 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 "年中心化"
   - Deviate Y_{it}, X_{it} from *year* (not state) averages
     计算 Y_{it}和 X_{it}与其年均值的差异
   - Estimate by OLS using "year-demeaned" data

These two methods can be combined…

```
. 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)
```

n-1个虚拟变量，用82年的做benchmark

(there are other ways to do this in STATA)

Combined entity and time fixed effects

- When $T = 2$, computing the first difference and including an intercept is equivalent (gives exactly the same regression) as the previous STATA command.
  当 T=2 时，计算一阶差分并包括截距项，这与之前 STATA 命令是等价的
- So there are various equivalent ways to allow for both entity and time fixed effects:
  因此有各种等价的方法可以处理状态和时间固定效应
  - differences & intercept ($T = 2$ only) – this is what we did initially
    差分并包括截距项（只适合于 T=2）
  - entity demeaning & $T – 1$ time indicators
    状态去均值化和 T-1 个时间虚拟变量
  - time demeaning & $n – 1$ entity indicators
    时间去均值化和 n-1 个状态虚拟变量
  - $T – 1$ time indicators & $n – 1$ entity indicators
    T-1 个时间虚拟变量和 n-1 个状态虚拟变量
  - entity & time demeaning 状态和时间都去均值化

The Fixed Effects Regression Assumptions and Standard Errors for Fixed Effects Regression (SW Section 10.5 and App. 10.2)
固定效应回归假设和标准误

Under assumptions that are basically extensions of the least squares assumptions, the OLS fixed effects estimator of β_1 is normally distributed. However, there are some subtleties associated with computing standard errors that don't come up with cross-sectional data.

在最小二乘扩展假设下，OLS 固定效应回归估计量β_1满足正态分布。然而，在计算标准误时，会遇到一些横截面数据下不会遇到的问题。

<u>Outline</u>

A. The fixed effects regression assumptions 固定效应回归假设

B. Standard errors for fixed effects regression – in two cases, one of which is new. 固定效应回归标准误 – 两种情况，其中一种是全新的

A. Extension of LS Assumptions to Panel Data 面板数据最小二乘扩展架设

Consider a single X: 重点掌握第一个假设，其余定性了解

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

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,…,n, are i.i.d. draws from their joint distribution. 允许存在一些相关性，因为只规定了i，没有规定T
3. $(X_{it},\ Y_{it})$ have finite fourth moments. 有限四阶矩
4. There is no perfect multicollinearity (multiple X’s) 不存在完全共线性
5. $\text{corr}(u_{it},u_{is}|X_{it},X_{is},\alpha_i) = 0$ for $t \neq\ s$.

Assumptions 3&4 are least squares LS assumptions 3&4
Assumptions 1&2 differ
Assumption 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
  给定状态固定效应和该状态下 X 所有的历史值，u_{it} 的均值为 0
- This is an extension of the previous multiple regression Assumption #1
  这是先前多元回归假设 1 的扩展
- This means there are no omitted lagged effects (any lagged effects of X must enter explicitly)
  这意味着不存在被遗漏的滞后效应（任何 X 的滞后效应都显性地进入条件中

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

- Also, there is not feedback from *u* to future *X*:
  另外，也不存在 u 对未来 X 的反馈
  - Whether a state has a particularly high fatality rate this year doesn't subsequently affect whether it increases the beer tax.
    无论一个州今年是否具有很高的死亡率，都不会影响未来它是否会提高啤酒税
  - We'll return to this when we take up time series data.
    未来学习时间序列的时候我们还会提及这个问题

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
  这是横截面数据多元回归假设 2 的扩展

- 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.
  如果个体（州，个人）是从总体中经过随机简单抽样得到的，且这些数据在每个时间段中都是如此采集的，那么就满足独立同分布这个条件。

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

- 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 beer tax this year is strongly related to whether it will have a high tax next year).
  这并不要求观察值对于同个个体，在整个时间跨度内都满足独立同分布——这样的要求是不现实的（事实上，一个州今年是否有啤酒税与明年是否会征收啤酒税是极度相关的）

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

- We haven't seen this before.
  先前我们并没有学过这个假设
- This says that (given *X*), the error terms are uncorrelated over time within a state.
  它说的是（给定 X），在一个州内，误差项在不同时间内是不相关的
- For example, $u_{CA,1982}$ and $u_{CA,1983}$ are uncorrelated
  例如，$u_{CA,1982}$和$u_{CA,1983}$是不相关的

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

- Is this plausible? What enters the error term?
  这个假设可行吗?误差项中包括了哪些内容
  - Especially snowy winter
    恶劣的大雪天
  - Opening major new divided highway
    新的高速公路投入使用
  - Fluctuations in traffic density from local economic conditions
    当地经济状况影响交通密度的变化
- Assumption #5 requires these omitted factors entering u_{it} to be uncorrelated over time, within a state.
  假设 5 要求这些进入 u_{it} 的遗漏变量在一个州内，在不同时间之间是不相关的

Assumption #5 in a picture:

列与列之间是独立的（假设2）
行与行不相关（假设5）

$$
\begin{array}{cccccc}
 & i=1 & i=2 & i=3 & \cdots & i=n \\
t=1 & u_{11} & u_{21} & u_{31} & \cdots & u_{n1} \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
t=T & u_{1T} & u_{2T} & u_{3T} & \cdots & u_{nT}
\end{array}
$$

← Sampling is i.i.d. across entities →

(by Assumption #2)

Assumption #5: u's are uncorrelated over time, same entity

- Is this plausible? 这可行吗?
- The u's consist of omitted factors – are they uncorrelated over time? 误差项包含了各种遗漏变量——它们是否在时间跨度上是不相关的?

What if Assumption #5 fails: so corr(u_{it},u_{is}|X_{it},X_{is},αi) ≠0? 如果假设5不满足会怎样?

- A useful analogy is heteroskedasticity.
  一个有用的类比就是异方差
- OLS panel data estimators of β_1 are unbiased, consistent
  OLS 估计量β_1依旧是无偏且一致的
- The OLS standard errors will be wrong – usually the OLS standard errors understate the true uncertainty
  但 OLS 标准误是错误的——一般，OLS 标准误会低估真实的不确定性

(cont.)What if Assumption #5 fails: so corr(u_{it},u_{is}|X_{it},X_{is},αi) ≠0?

同一个ID的不同时间的数据存在相关性

- 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.
  直觉：如果 u_{it} 在时间上是相关的，那么你并没有像在 u_{it} 是不相关时拥有那么多的信息（即：允许那么多的随机扰动）
  因此方差会变大
- This problem is solved by using "heteroskedasticity and autocorrelation-consistent standard errors"
  这个问题的解决可以依靠"异方差自相关一致性标准误"

B. Standard Errors 标准误

怎样处理。

B.1 First get the large-*n* approximation to the sampling distribution of the FE estimator

首先在大样本下得到固定效应估计量的样本分布

Fixed effects regression model: $\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$

这里demean了

这个推导不用掌握

OLS fixed effects estimator: $$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{Y}_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^2}$$

so: $$\hat{\beta}_1 - \beta_1 = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}}{\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^2}$$

这部分就是bias

Sampling distribution of fixed effects estimator, 固定效应估计量的样本分布

Fact:

$$\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it} = \sum_{t=1}^{T}\tilde{X}_{it}u_{it} - \left[\sum_{t=1}^{T}\left(X_{it} - \bar{X}_{i}\right)\right]\bar{u}_{i} = \sum_{t=1}^{T}\tilde{X}_{it}u_{it}$$

so

$$\sqrt{nT}\,(\hat{\beta}_1 - \beta_1) = \frac{\sqrt{\frac{1}{nT}}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{v}_{it}}{\hat{Q}_{\tilde{X}}^{2}} = \frac{\sqrt{\frac{1}{n}}\sum_{i=1}^{n}\eta_i}{\hat{Q}_{\tilde{X}}^{2}}$$

对ID的sample average

where $\eta_i = \sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{v}_{it}$, $\tilde{v}_{it} = \tilde{X}_{it}u_{it}$, and $\hat{Q}_{\tilde{X}}^{2} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{X}_{it}^{2}$.

By the CLT,

$$\boxed{\sqrt{nT}\,(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N(0,\ \sigma_{\eta}^{2}/Q_{\tilde{X}}^{4})}$$

左右同乘

where $\xrightarrow{d}$ means converges in distribution and $\hat{Q}_{\tilde{X}}^{2} \xrightarrow{p} Q_{\tilde{X}}^{2}$.

(cont.)Sampling distribution of fixed effects estimator

$$\sqrt{nT}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N(0, \sigma_\eta^2/Q_{\tilde{X}}^4), \text{ where } \sigma_\eta^2 = \operatorname{var}\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{v}_{it}\right)$$

B.2 Obtain Standard Error: 得到标准误

- Standard error of $\hat{\beta}_1$: $SE(\hat{\beta}_1) = \sqrt{\dfrac{1}{nT}\dfrac{\hat{\sigma}_\eta^2}{\hat{Q}_{\tilde{X}}^4}}$
- Only part we don't have: what is $\hat{\sigma}_\eta^2$?
  - Case I: u_{it}, u_{is} uncorrelated u_{it}, u_{is} 不相关
  - Case II: u_{it}, u_{is} correlated u_{it}, u_{is} 相关

Case I: $\hat{\sigma}_B^2$ when u_{it}, u_{is} are uncorrelated

$$\sigma_\eta^2 = \text{var}\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{v}_{it}\right) = \text{var}\left(\frac{\tilde{v}_{i1}+\tilde{v}_{i2}+\ldots+\tilde{v}_{iT}}{\sqrt{T}}\right)$$

- Recall var($X + Y$) = var(X) + var(Y) + 2cov(X,Y).
- When u_{it} and u_{is} are uncorrelated, $\text{cov}(\tilde{v}_{it},\tilde{v}_{is}) = 0$ unless $t = s$, so all the covariance terms are zero and
  当 u_{it} 和 u_{is} 不相关时，除非 t=s，否则 $\text{cov}(\tilde{v}_{it},\tilde{v}_{is}) = 0$。因此所有的协方差项目都是 0

$$\sigma_\eta^2 = \frac{1}{T}\times T\,\text{var}\left(\tilde{v}_{it}\right) = \text{var}(\tilde{v}_{it})$$

- You can use the usual (hetero-robust) SE formula for standard errors if T isn't too small. This works because the usual hetero-robust formula is for uncorrelated errors – which is the case here.
  当 T 不是很小时，你可以采用一般的异方差稳健标准误。这种方法是可行的，因为一般的异方差稳健标准误适合于不相关的误差项

Case II: u_{it} and u_{is} are correlated – so Assumption 5 fails

$$\sigma_\eta^2 = \mathrm{var}\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{v}_{it}\right)$$
$$= \mathrm{var}\left(\frac{\tilde{v}_{i1} + \tilde{v}_{i2} + \ldots + \tilde{v}_{iT}}{\sqrt{T}}\right)$$
$$\neq \mathrm{var}(\tilde{v}_{it})$$

- Recall var(X + Y) = var(X) + var(Y) + 2cov(X,Y)
- If u_{it} and u_{is} are correlated, we have some nonzero covariances!! So in general we don't get any further simplifications.
  如果 u_{it} 和 u_{is} 相关，协方差项不再为 0。所以这个式子不可以进一步化简。
- However, we can still compute standard errors – but using a different method: "clustered" standard errors.
  尽管如此，我们仍旧可以计算标准误——但采用另一种方法：群标准误

Case II: Clustered Standard Errors

Variance:

$$\sigma_\eta^2 = \operatorname{var}\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\tilde{v}_{it}\right)$$

用估计里带进去，这个具体的表达式补考

Variance estimator:

这里把残差性带进去了

$$\hat{\sigma}^2_{\eta,clustered} = \frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\hat{\tilde{v}}_{it}\right)^2, \text{ where } \hat{\tilde{v}}_{it} = \tilde{X}_{it}\hat{u}_{it}.$$

知道cluster SE，是干什么用的，能处理什么问题，要掌握
具体做法不用掌握

Clustered standard error:

$$SE(\hat{\beta}_1) = \sqrt{\frac{1}{nT}\frac{\hat{\sigma}^2_{\eta,clustered}}{\hat{Q}^4_{\tilde{X}}}}$$

就是需要一个特殊的SE公式，不用背

Comments on clustered standard errors: 评述群标准误

- The clustered *SE* formula is NOT the usual (hetero-robust) SE formula!
  群标准误的表达式与一般的异方差稳健的标准误表达式并不相同！
- OK this is messy – but you get something for it – you can have correlation of the error for an entity from one time period to the next. This would arise if the omitted variables that make up u_{it} are correlated over time.
  你可以允许某个个体的误差项在时间跨度上具有相关性。事实上，当 u_{it} 中包含的遗漏变量在时间上是相关时，就会产生这种情况。

(cont.) Comments on clustered standard errors

- The Case II standard error formula goes under various names:
  第二个例子中的误差项表达式有多种名称
  - ***Clustered standard errors***, because there is a grouping, or "cluster," within which the error term is possibly correlated, but outside of which (across groups) it is not.
    群标准误，这是因为，那些可能具有相关性的误差项可以成群或成组，但不同组之间的误差项并不存在相关性
  - ***Heteroskedasticity- and autocorrelation-consistent standard errors*** (autocorrelation is correlation with other time periods – u_{it} and u_{is} correlated) 简称HAC
    异方差-自相关一致性标准误（自相关是指同个个体不同时期数据具有相关性，如 u_{it} 和 u_{is} 相关）

(cont.) Comments on clustered standard errors

- Extensions:
  - The clusters can be other groupings, not necessarily time
    群的划分标准可以是其他归类，不一定是“时间”
  - For example, you can allow for correlation of u_{it} between individuals within a given group, as long as there is independence across groups – for example i runs over individuals, the clusters can be families (correlation of u_{it} for i within same family, not between families).
    例如，在给定的组内，你可以允许 u_{it} 在不同个体间是相关的，只要不同组之间是独立的。例如，群可以是家庭（同个家庭中，不同个体的 u_{it} 是相关的，但不同家庭之间是不相关的）

Implementation in STATA

Case I: treat u_{it} and u_{is} as uncorrelated

这里没考虑时间上的相关性

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

Linear regression, absorbing indicators             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)
```

Case II: treat *uit* and *uis* as possibly correlated

这个命令要掌握

```
. areg vfrall beertax, absorb(state) robust cluster(state);

Linear regression, absorbing indicators                Number of obs =      336
                                                       F(  1,    47) =     4.34
                                                       Prob > F      =   0.0427
                                                       R-squared     =   0.9050
                                                       Adj R-squared =   0.8891
                                                       Root MSE      =   .18986

                                   (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6558736   .3148476    -2.08   0.043    -1.289265    -.022482
       _cons |   2.377075   .1615974    14.71   0.000     2.051983    2.702167
-------------+----------------------------------------------------------------
       state |   absorbed                                      (48 categories)
```

这里的cluster是用来控制时间上的相关性，每个ID内部

Coefficients are identical (*why*?)

Pretty big difference in the standard errors!

Try adding year effects:

```
. 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)

.   test $yeardum;

       F(  6,   281) =    2.47
            Prob > F =    0.0243
```

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

Linear regression, absorbing indicators                Number of obs =     336
                                                       F(  7,    47) =    3.74
                                                       Prob > F      =  0.0027
                                                       R-squared     =  0.9089
                                                       Adj R-squared =  0.8914
                                                       Root MSE      =  .18788

                                  (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6399799   .3857867    -1.66   0.104    -1.416083    .1361229
         y83 |  -.0799029   .0379069    -2.11   0.040    -.1561617   -.003644
         y84 |  -.0724206   .0474088    -1.53   0.133    -.1677948    .0229537
         y85 |  -.1239763   .0497587    -2.49   0.016    -.2240779   -.0238747
         y86 |  -.0378645   .0616479    -0.61   0.542    -.1618841    .0861552
         y87 |  -.0509021   .0687224    -0.74   0.463    -.1891536    .0873495
         y88 |  -.0518038   .0695801    -0.74   0.460    -.1917809    .0881732
       _cons |    2.42847   .2179038    11.14   0.000     1.990104    2.866836
-------------+----------------------------------------------------------------
       state |   absorbed                                     (48 categories)

    .   test $yeardum;

           F(  6,    47) =    3.61
                Prob > F =    0.0050
```

Fixed Effects Regression Results Dependent variable: *Fatality rate*

总结

	(1)	(2)	(3)	(4)
BeerTax	-.656** (.203)	-.656^{+} (.315)	-.640* (.255)	-.640^{++} (.386)
State effects?	Yes	Yes	Yes	Yes
Time effects?	No	No	Yes	Yes
F testing time effects = 0	–		2.47 (.024)	3.61 (.005)
Clustered SEs?	No	Yes	No	Yes

Significant at the **1% *5% $^{+}$10% level

$^{++}$ Significant at the 10% level using normal but not Student *t* critical values

This is a hard call – what would you conclude?

Summary: *SE*s for Panel Data in a picture:

$$
\begin{array}{cccccc}
 & i=1 & i=2 & i=3 & \cdots & i=n \\
t=1 & u_{11} & u_{21} & u_{31} & \cdots & u_{n1} \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
t=T & u_{1T} & u_{2T} & u_{3T} & \cdots & u_{nT}
\end{array}
$$

$$\leftarrow \textit{i.i.d. sampling across entities} \rightarrow$$

- Intuition #1: This is similar to heteroskedasticity – you make an assumption about the error, derive *SEs* under that assumption, if the assumption is wrong, so are the *Ses*
直觉 1：这与异方差类似——你对标准误提出假设，得到在该假设下的标准误。如果假设是错误的，那么你得到的标准误自然也是错误的

(cont.) Summary: *SE*s for Panel Data in a picture:

$$
\begin{array}{cccccc}
 & i=1 & i=2 & i=3 & \cdots & i=n \\
t=1 & u_{11} & u_{21} & u_{31} & \cdots & u_{n1} \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
t=T & u_{1T} & u_{2T} & u_{3T} & \cdots & u_{nT}
\end{array}
$$

$$\leftarrow \textit{i.i.d. sampling across entities} \rightarrow$$

- Intuition #2: If u_{21} and u_{22} are correlated, there is less information in the sample than if they are not – and *SEs* need to account for this (usual *SEs* are typically too small)
  直觉 2： 如果 u_{21} 和 u_{22} 相关，那么比起它们是不相关时所拥有的信息量会小很多——因此标准误的计算应该考虑到这一点（一般的标准误会偏小）
- Hetero-robust (or homosk-only) *SE*s don't allow for this correlation, but clustered *SE*s do.
  异方差稳健（或是同方差）标准误都没有考虑相关性，但群标准误考虑了

Application: Drunk Driving Laws and Traffic Deaths (SW Section 10.6)

Some facts

- Approx. 40,000 traffic fatalities annually in the U.S.

  在美国每年大约有 40000 死于交通事故

- 1/3 of traffic fatalities involve a drinking driver

  1/3 的交通事故与酒驾有关

- 25% of drivers on the road between 1am and 3am have been drinking (estimate)

  在清晨 1 点到 3 点之间，25%的司机都有饮酒

- A drunk driver is 13 times as likely to cause a fatal crash as a non-drinking driver (estimate)

  一个酒驾的司机发生交通事故的可能性大概是非酒驾司机的 13 倍

Drunk driving laws and traffic deaths, ctd.

Public policy issues

公共政策事件

- Drunk driving causes massive externalities (sober drivers are killed, society bears medical costs, etc. etc.) – there is ample justification for governmental intervention
  酒驾会具有一系列的外部性（冷静的司机会因此遭殃，社会医疗成本增加等）——因此，政府有足够的理由干涉这一事件
- Are there any effective ways to reduce drunk driving? If so, what?
  是否存在有效降低酒驾的途径？如果有的话，那会是什么？

Drunk driving laws and traffic deaths, ctd.

Public policy issues

- What are effects of specific laws:
  具体法律的效果如何？
  - mandatory punishment u_{21} and u_{22}
    强制性惩罚
  - minimum legal drinking age
    法律最小饮酒年龄
  - economic interventions (alcohol taxes)
    经济干涉（酒精税）

The Commonwealth of Massachusetts
Executive Department
State House • Boston, MA 02133
(617) 725-4000

MITT ROMNEY
GOVERNOR

KERRY HEALEY
LIEUTENANT GOVERNOR

FOR IMMEDIATE RELEASE:
October 28, 2005

CONTACT:
Julie Teer
Laura Nicoll
(617) 725-4025

ROMNEY CELEBRATES THE PASSAGE OF MELANIE'S BILL

Legislation puts Massachusetts in line with federal standards for drunk driving

Governor Mitt Romney today signed into law the toughest drunk driving legislation in the Commonwealth's history.

Named in honor of 13-year-old Melanie Powell, the new law will stiffen penalties for drunk driving offenses in Massachusetts and close loopholes in the legal system that allow repeat drunk drivers to get back behind the wheel.

以纪念13岁Melanie Powell而命名的新法律将严惩麻省境内的醉驾违规行为，并消除了允许重复醉驾的司机继续开车的法律漏洞。

"Today we honor those who have lost their lives in senseless drunk driving tragedies and act to save the lives we could otherwise lose next year," said Romney. "We have Melanie's Law today because the citizens of the Commonwealth cared enough to make it happen."

"今天我们向那些在麻木醉驾中失去生命的人们，并采取行动去拯救那些可能在下一年中失去生命的个体。" Romney说道，"我们今天之所以有Melanie 法案是因为联邦公民对此足够关心。"

The new measure gives prosecutors the power to introduce certified court documents to prove that a repeat offender has been previously convicted of drunk driving. In addition, the mandatory minimum jail sentence for any individual found guilty of manslaughter by motor vehicle will be increased from 2 ½ to five years.

新法案赋予检察官新的权力，允许他们提出经法院认可的证明醉驾惯犯曾因醉驾而定罪的证据。此外，因车祸致受害人死亡的个人所获的最小量刑将由两年半增至五年。

Repeat offenders will be required to install an interlock device on any vehicle they own or operate. These devices measure the driver's Blood Alcohol Content (BAC) and prevent the car from starting if the driver is intoxicated. Any individual who tampers with the interlock device could face a jail sentence.

醉驾惯犯将被要求在他们驾驶的机动车内安装联锁装置。这些装置可测量驾驶员血液的酒精含量，若其酒精含量超标，则禁止车辆发动。任何毁损这一联锁装置的个人都将面临监禁。

For the first time, Massachusetts will be in compliance with federal standards for drunk driving laws.

Romney was joined by Tod and Nancy Powell, the parents of Melanie Powell, and her grandfather, Ron Bersani to celebrate the passage of the new drunk driving measure.

这将是有史以来第一次，麻省遵循联邦醉驾法律。Melanie Powell的父母Tod和Nancy Powell，她的祖父Ron Bersani同Romney一起庆祝了关于醉驾量刑的新法案。

"Today we should give thanks to all of those who have worked so hard to make this day possible," said Bersani. "Governor Romney and the Legislative leadership have advanced the fight against repeat drunk driving to heights that seemed unattainable just six months ago.

"今天我们应当感谢那些努力使这一切成为可能的人们，"Bersani说 道，"州长Romney先生和立法领导者已将打击反复醉驾的行动提升到了六个月前不能企及的高度

Under the law, stiff penalties will be established for individuals who drive while drunk with a child under the age of 14 in the vehicle and those who drive with a BAC of .20 or higher, more than twice the legal limit.

Romney thanked the Legislature for enacting a tough bill that cracks down on repeat drunk driving offenders in Massachusetts.

在这项法律下，车内载有14岁以下未成年人的醉酒司机和在驾驶中血液酒精含量超过0.2的个体将受到比之前重两倍的严厉惩罚。

Romney感谢麻省立法委员会通过这项严厉打击反复醉驾的法案。

"Public safety is one of our top priorities and Melanie's Law will go a long way towards making our citizens and roadways safer," said Speaker Salvatore F. DiMasi. "I commend the my colleagues in the Legislature and the Governor for taking comprehensive and quick action on this very important issue."

“保障公共安全是我们最重要的目标之一，而Melanie法案将在使我们的公民和道路更加安全的征途中前进一大步。”Salvatore F. DimMsi评论道。 我对我立法委员会的同事和州长在这一重要问题上所采取的全面而迅速的行动表示赞赏。

"Today we are sending a powerful message that Massachusetts is serious about keeping repeat drunken drivers off the road," said House Minority Leader Bradley H. Jones Jr. "I am proud of the Governor, Lieutenant Governor, and my legislative colleagues for joining together to pass tough laws to make our roadways safer."

今天我们向公众传达了这样一个信息：麻省在打击反复醉驾上态度坚决。众议院少数党领袖Bradley H.Jones Jr说道，“我为州长、副州长和我立法委员会的同事们以联手通过这一严厉法案来确保公众道路更加安全的行为深感骄傲。

"I am pleased and proud that the Legislature did the right thing in the end and supported a Bill worthy of Melanie's name and the sacrifices made by the Powell family and all victims of drunk drivers," said Senator Robert L. Hedlund. "Melanie's Law will save lives and it would not have been accomplished if not for the tireless efforts and advocacy of the families."

我为立法委员会这一最终正确决策感到满意和骄傲。该法案是对Melanie及其家人所作出的牺牲和所有醉驾受害者的告慰。Rober L. Hedlund表示"Melanie法案将会拯救许多生命。而没有Melanie家族的不懈努力和大力倡导，这一切都不可能实现。"

Representative Frank Hynes added, “I’d like to commend Ron, Tod, and Nancy for their tireless work in support of Melanie’s bill. As a family, they were able to turn the horrific tragedy in their lives into a greater measure of safety for all families on Massachusetts roadways.”

议员代表Frank Hynes补充道“我想就Ron，Tod和Nancy为支持Melanie 法案所做的不懈努力表示赞扬。作为一个家庭，他们将生命中的这一可怕悲剧转变成了确保麻省所有家庭公路安全的法案。”

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 最低法制喝酒年龄
- Minimum sentencing laws for first DWI violation:
  第一次酒后驾车获罪的处罚
  - *Mandatory Jail* 强制监禁
  - *Manditory Community Service 强制参与社区服务*
  - otherwise, sentence will just be a monetary fine
- Vehicle miles per driver (US DOT) 每个司机拥有的汽车数 这个应该是里程数
- State economic data (real per capita income, etc.) 州经济数据

Why might panel data help?

- Potential OV bias from variables that vary across states but are constant over time:
  存在随州变化但不随时间变化的遗漏变量
  - culture of drinking and driving 饮酒和驾驶文化
  - quality of roads 道路质量
  - vintage of autos on the road 道路上汽车的制造年份

    $\Rightarrow$ use state fixed effects 采用州固定效应
- Potential OV bias from variables that vary over time but are constant across states:
  存在随时间变化但不随州变化的遗漏变量
  - improvements in auto safety over time 汽车安全性的提高
  - changing national attitudes towards drunk driving
    全国对酒驾态度的转变

    $\Rightarrow$ use time fixed effects 采用时间固定效应

这里就有七个模型

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** (0.05)	−0.66** (0.20)	−0.64* (0.25)	−0.45* (0.22)	−0.70** (0.25)	−0.46* (0.22)	−0.45 (0.32)
Drinking age 18				0.028 (0.066)	−0.011 (0.064)		0.028 (0.076)
Drinking age 19				−0.019 (0.040)	−0.078 (0.049)		−0.019 (0.054)
Drinking age 20				0.031 (0.046)	−0.102* (0.046)		0.031 (0.055)
Drinking age						−0.002 (0.017)	
Mandatory jail?				0.013 (0.032)	−0.026 (0.065)		0.013 (0.018)
Mandatory community service?				0.033 (0.115)	0.147 (0.137)		0.033 (0.144)
Mandatory jail or community service?						0.039 (0.084)	
Average vehicle miles per driver				0.008 (0.008)	0.017 (0.010)	0.009 (0.008)	0.008 (0.007)
Unemployment rate				−0.063** (0.012)		−0.063** (0.012)	−0.063** (0.014)
Real income per capita (logarithm)				1.81** (0.47)		1.79** (0.45)	1.81* (0.69)
State effects?	no	yes	yes	yes	yes	yes	yes
Time effects?	no	no	yes	yes	yes	yes	yes
Clustered standard errors?	no	no	no	no	no	no	yes

第七个模型，经济上显著，但是统计上不显著，很tricky，要解读

自由度为6的F检验

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)	11.62 (< 0.001)	8.64 (< 0.001)
Drinking age coefficients = 0				0.48 (0.696)	2.09 (0.102)		0.30 (0.825)
Jail, community service coefficients = 0				0.17 (0.845)	0.59 (0.557)		0.28 (0.758)
Unemployment rate, income per capita = 0				38.29 (< 0.001)		40.15 (< 0.001)	25.88 (< 0.001)
$\bar{R}^2$	0.090	0.889	0.891	0.926	0.893	0.926	0.926

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

Empirical Analysis: Main Results
实证结果: 主要结论

- Sign of beer tax coefficient changes when fixed state effects are included
  当引入州固定效应时，啤酒税的符号发生了改变

- Fixed time effects are statistically significant but do not have big impact on the estimated coefficients
  时间固定效应是显著的，但对其他估计量影响不大

- Estimated effect of beer tax drops when other laws are included as regressor
  当其他法律被作为解释变量时，啤酒税的估计影响降低了

(cont.)Empirical Analysis: Main Results

- The only policy variable that seems to have an impact is the tax on beer – not minimum drinking age, not mandatory sentencing, etc. – *however the beer tax is not significant even at the 10% level using clustered SEs.*
  看上去所有的政策变量中只有啤酒税是有明显作用的——而不是最低法定饮酒年龄或是强制性惩罚等。然而，即使在 10%显著水平下，采用群标准误，啤酒税这一变量仍就是不显著的
- The other economic variables have plausibly large coefficients: more income, more driving, more deaths
  其他经济变量的系数较大，更多的收入意味着更多的司机，更高的死亡率 但是不具备可操纵的意义

Digression: extensions of the "*n*-1 binary regressor" idea

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.

采用更多的虚拟变量来消除遗漏变量误差的方法并不局限于面板数据——关键在于，遗漏变量在同组之内是相同的，从而每组数据拥有其各自的截距项

(cont.)Digression: extensions of the "*n*-1 binary regressor" idea

Example**: Class size problem.**

实例：教室大小问题

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).

假设赞助和课程设置取决于每个县各自的状况，且每个县都有一些社区。那么，遗漏变量问题可以通过引入虚拟变量来解决，这些变量用于指示社区属于哪个县（为了避免完全共线性问题，N 个县只需要引入 N-1 个虚拟变量）

Summary: Regression with Panel Data

Advantages and limitations of fixed effects regression

固定效应回归模型的优势和不足

Advantages 优势

- You can control for unobserved variables that:
- 你可以控制这些不可观察的变量
  - vary across states but not over time, and/or
  - 随州变化但不随时间变化
  - vary over time but not across states
    随时间变化但不随州变化
- More observations give you more information
- 更多的观察值会提供更多的信息
- Estimation involves relatively straightforward extensions of multiple regression 估计方法只需在多元回归下进行一定的扩展即可

- **Fixed effects regression can be done three ways:**
  固定效应回归可以采用三种方法估计

  1. "Changes" method when $T = 2$
     当只有两期时可采用"前后"回归方法，
  2. "n-1 binary regressors" method when n is small
     当 n 比较小时采用"n-1 个二元变量"方法
  3. "Entity-demeaned" regression
     "个体中心化"估计

- Similar methods apply to regression with time fixed effects and to both time and state fixed effects
  对于时间固定效应或是同时考虑了时间和州固定效应的模型，都可采用类似的估计方法
- Statistical inference: like multiple regression.
  统计推断与多元回归类似

Limitations/challenges

局限/挑战

- Need variation in *X* over time within states

  在州内，要求 X 有时间上的变化 没有变化的，都在阿尔法中有体现，放进去没意义

- Time lag effects can be important

  时间滞后效应是很重要的

- You should use heteroskedasticity- and autocorrelation-consistent (clustered) standard errors if you think u_{it} could be correlated over time

  如果你认为 u_{it} 在时间上是自相关的，那么就必须采用异方差-自相关一致性（群）标准误