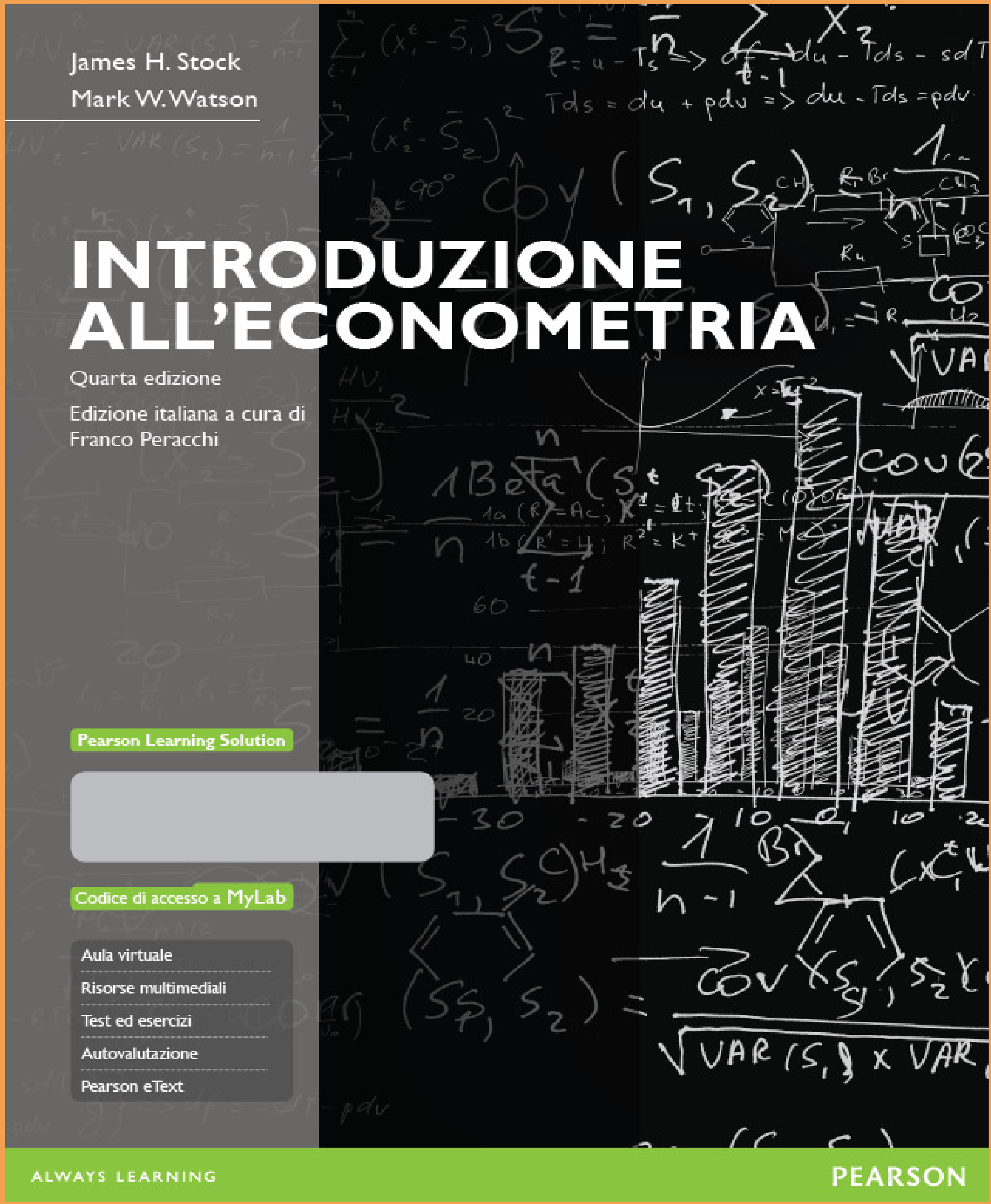
Capitolo 10

Regressione con dati panel



Sommario

1. Dati panel: cosa e perché
2. Dati panel con due periodi temporali
3. Regressione con effetti fissi
4. Regressione con effetti temporali
5. Errori standard per regressione con effetti fissi
6. Applicazione a guida in stato di ebbrezza e sicurezza stradale

Dati panel: cosa e perché (Paragrafo 10.1)

Un ***panel*** contiene osservazioni su più unità (individui, stati, imprese) in cui ogni entità è osservata in due o più istanti temporali diversi.

Esempi:

- Dati su 420 distretti scolastici della California nel 1999 *e ancora* nel 2000, per 840 osservazioni in totale.
- Dati su 50 stati USA, ognuno è osservato per 3 anni, per un totale di 150 osservazioni.
- Dati su 1000 individui, in quattro mesi diversi, per 4000 osservazioni in totale.

Notazione per dati panel

Un doppio pedice distingue unità (stati) e periodi temporali (anni)

i = unità (stato), n = numero di entità,
perciò $i = 1,\ldots,n$

t = periodo temporale (anno), T = numero di periodi temporali
perciò $t = 1,\ldots,T$

Dati: supponiamo di avere 1 regressore. I dati sono:

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

Notazione per dati panel (continua)

Dati panel con k regressori:

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

n = numero di unità (stati)
T = numero di periodi temporali (anni)

Un po' di gergo…

- I dati panel sono chiamati anche ***dati longitudinali***
- ***panel bilanciato***: non ci sono osservazioni che mancano, cioè tutte le variabili sono osservate per tutte le unità (stati) e tutti i periodi temporali (anni)

Perché sono utili i dati panel?

Con i dati panel possiamo controllare per fattori che:

- Variano tra le unità ma non nel tempo
- Potrebbero causare distorsione da variabili omesse se fossero omessi
- Sono inosservati o non misurati, e perciò non possono essere inclusi in una regressione multipla

Ecco l'idea chiave:

Se una variabile omessa non varia nel tempo, allora qualsiasi *variazione* in *Y* nel tempo non può essere causata dalla variabile omessa.

Esempio di dati panel: morti sulle strade e imposte sugli alcolici

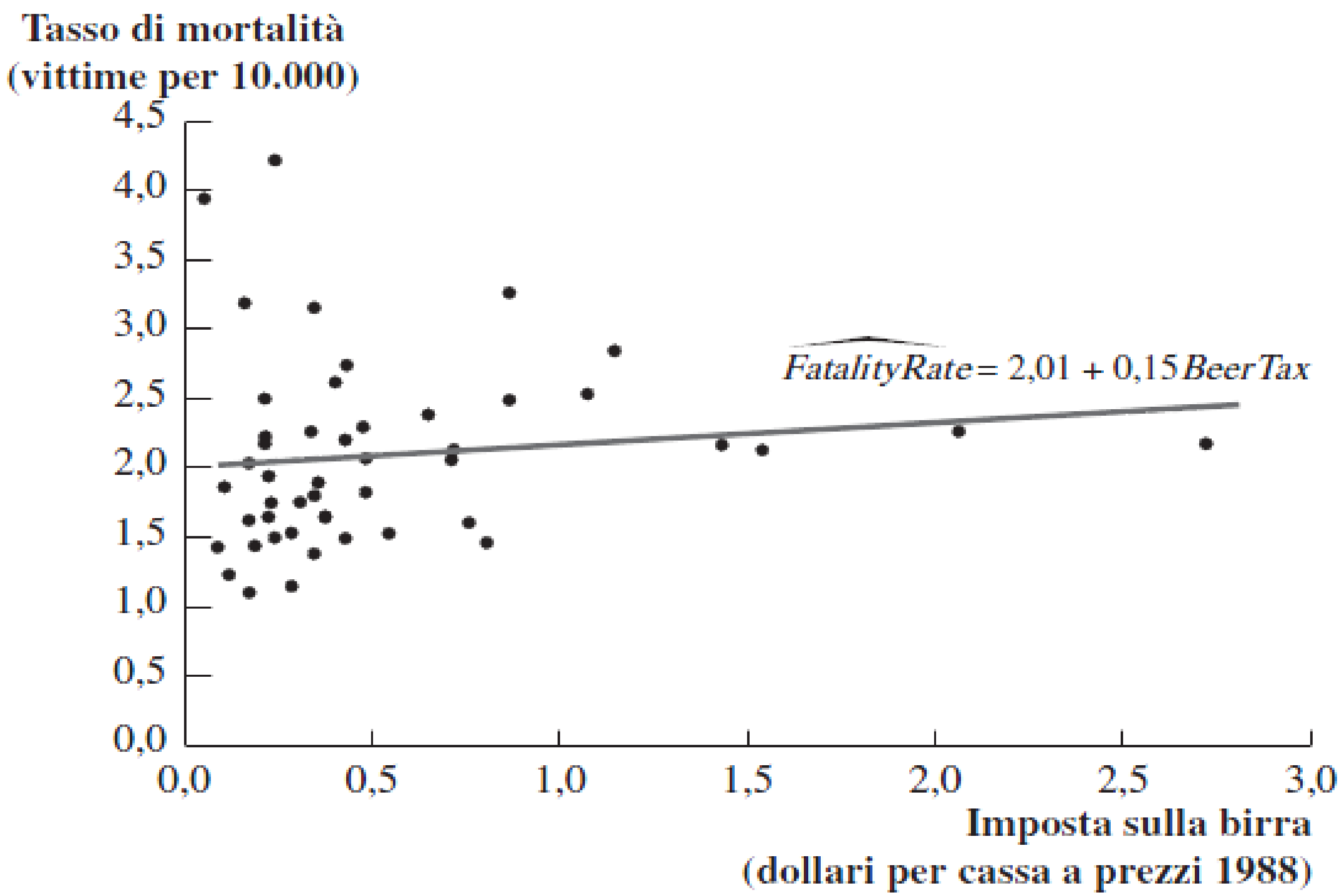
Unità di osservazione: un anno in uno stato USA

- 48 stati USA, perciò n = numero di unità = 48
- 7 anni (1982,…, 1988), perciò T = numeri di periodi temporali = 7
- Panel bilanciato, perciò numero totale di osservazioni = 7×48 = 336

Variabili:

- Tasso di mortalità stradale (numero di morti sulle strade in quello stato in quell’anno, per 10.000 residenti nello stato)
- Imposta su una cassa di birra
- Altre (età minima per guidare, leggi sulla guida in stato di ebbrezza, ecc.)

Mortalità stradale USA nel 1982:



Imposte sugli alcolici più elevate e maggiore mortalità?

Perché potrebbero esserci *più* morti sulle strade in stati in cui ci sono imposte più elevate sugli alcolici?

Altri fattori che influenzano il tasso di mortalità stradale:

- Qualità (età) delle automobili
- Qualità delle strade
- "Cultura" sul bere e guidare
- Densità di auto sulle strade

Questi fattori omessi potrebbero causare distorsione da variabili omesse.

Esempio 1: densità del traffico. Supponiamo:

I. Elevata densità del traffico significa più morti sulle strade
II. Gli stati con minore densità di traffico (all'ovest) hanno imposte sugli alcolici minori

- Allora le due condizioni per la distorsione da variabili omesse sono soddisfatte. Nello specifico, "imposte elevate" potrebbero riflettere "alta densità di traffico" (perciò il coefficiente OLS sarebbe distorto positivamente
  – imposte elevate, più morti)
- I dati panel ci consentono di eliminare la distorsione da variabili omesse quando le variabili omesse sono costanti nel tempo in un dato stato.

Esempio 2: attitudini culturali verso il bere e la guida:

(i) sono presumibilmente un determinante della mortalità stradale; e

(ii) sono potenzialmente correlate con le imposte sulla birra.

•Allora le due condizioni per la distorsione da variabili omesse sono soddisfatte. Nello specifico, "alte imposte" potrebbe captare l'effetto di "attitudini culturalli verso il bere", perciò il coefficiente OLS sarebbe distorto.

•I dati panel ci consentono di eliminare la distorsione da variabili omesse quando le variabili omesse sono costanti nel tempo in un dato stato.

Dati panel con due periodi temporali (Paragrafo 10.2)

Consideriamo il modello dei dati panel,

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

Z_i è un fattore che non cambia nel tempo (densità), almeno durante gli anni per cui abbiamo dati.

- Supponiamo che Z_i non sia osservato, perciò la sua omissione potrebbe comportare distorsione da variabili omesse.
- L'effetto di Z_i può essere eliminato usando T = 2 anni.

L'idea chiave:

Qualsiasi *variazione* nel tasso di mortalità dal 1982 al 1988 non può essere causata da Z_i, perché Z_i (per ipotesi) non varia tra il 1982 e il 1988.

Matematica: consideriamo i tassi di mortalità nel 1988 e nel 1982:

$$FatalityRate_{i1988} = \beta_0 + \beta_1 BeerTax_{i1988} + \beta_2 Z_i + u_{i1988}$$
$$FatalityRate_{i1982} = \beta_0 + \beta_1 BeerTax_{i1982} + \beta_2 Z_i + u_{i1982}$$

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

Sottraendo 1988 – 1982 (ovvero calcolando la variazione) si elimina l'effetto di Z_i…

$$FatalityRate_{i1988} = \beta_0 + \beta_1 BeerTax_{i1988} + \beta_2 Z_i + u_{i1988}$$
$$FatalityRate_{i1982} = \beta_0 + \beta_1 BeerTax_{i1982} + \beta_2 Z_i + u_{i1982}$$

perciò

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

- Il nuovo termine d'errore, $(u_{i1988} - u_{i1982})$, non è correlato con $BeerTax_{i1988}$ o $BeerTax_{i1982}$.
- Questa equazione "alle differenze" può essere stimata con OLS, anche se Z_i non è osservata.
- La variabile omessa Z_i non cambia, perciò non può essere una determinante della *variazione* in Y
- Questa regressione alle differenze non ha un'intercetta, che è stata eliminata dalla sottrazione

Esempio: mortalità stradale e imposte sulla birra

Dati del 1982:

$$\widehat{FatalityRate} = \underset{(0,15)}{2,01} + \underset{(0,13)}{0,15}BeerTax \qquad (n = 48)$$

Dati del 1988:

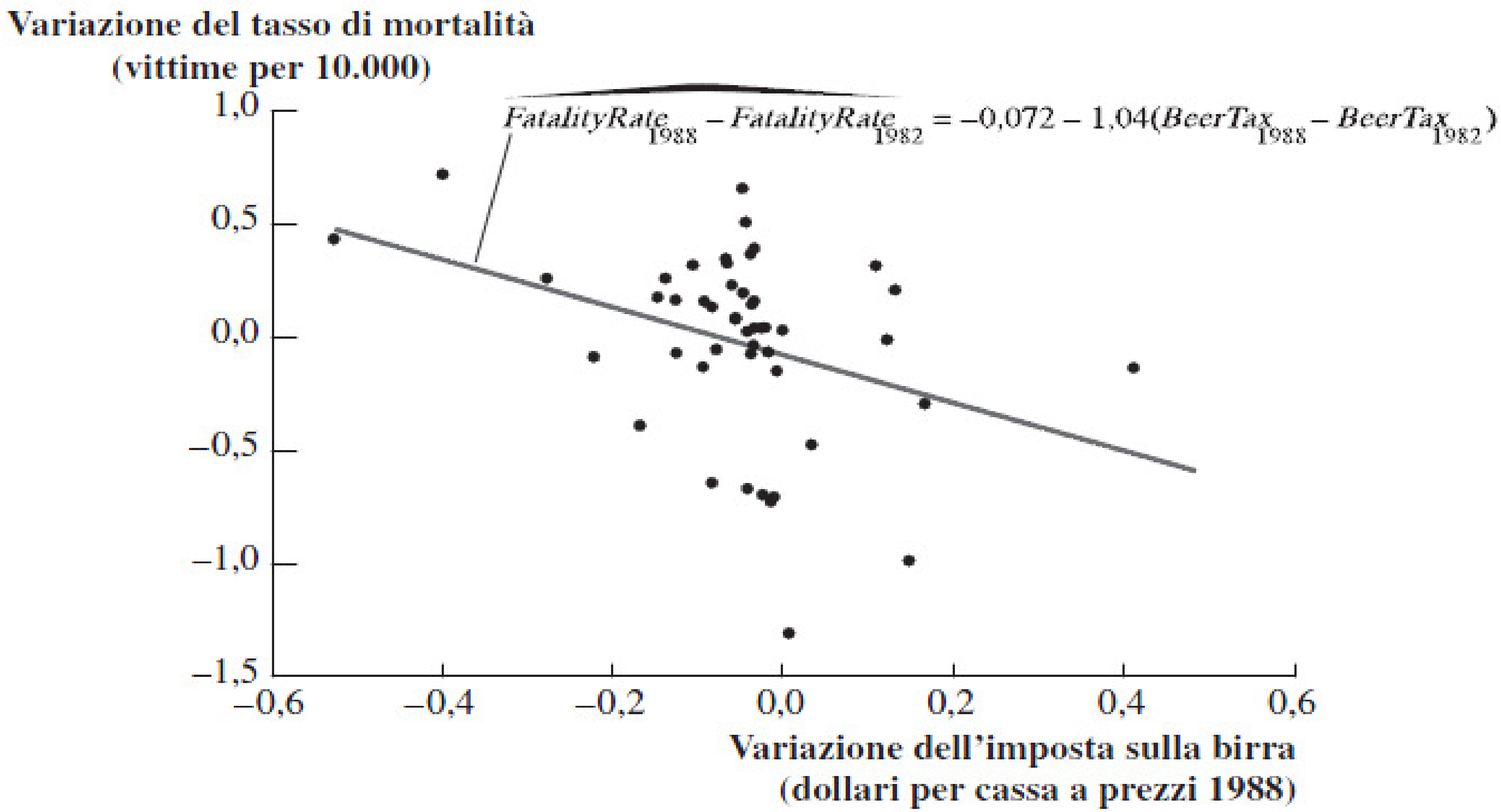
$$\widehat{FatalityRate} = \underset{(0,11)}{1,86} + \underset{(0,13)}{0,44}BeerTax \qquad (n = 48)$$

Regressione differenze (n = 48)

$$\widehat{FR_{1988} - FR_{1982}} = \underset{(0,065)}{-0,072} - \underset{(0,36)}{1,04}(BeerTax_{1988} - BeerTax_{1982})$$

Un'intercetta inclusa in questa regressione alle differenze consente che la variazione media in FR sia non nulla – riprenderemo questo punto più avanti…

Δ*FatalityRate* v. Δ*BeerTax*:



Si noti che l'intercetta è quasi a zero…

Regressione con effetti fissi (Paragrafo 10.3)

E se si hanno più di 2 periodi temporali ($T > 2$)?

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

Possiamo riscriverlo in due modi utili:

1. modello di regression " con *n*-1 regressori binari"
2. modello di regressione "con effetti fissi"

Prima lo riscriviamo nella forma "con effetti fissi". Supponiamo di avere $n = 3$ stati: California, Texas e Massachusetts.

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

Regressione per la California (i = CA):

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

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

O

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

- $\alpha_{CA} = \beta_0 + \beta_2 Z_{CA}$ non cambia nel tempo
- α_{CA} è l'intercetta per CA, e β_1 è la pendenza
- L'intercetta è specifica per CA, ma la pendenza è la stessa in tutti gli stati: rette parallele.

Per TX:

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

o

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

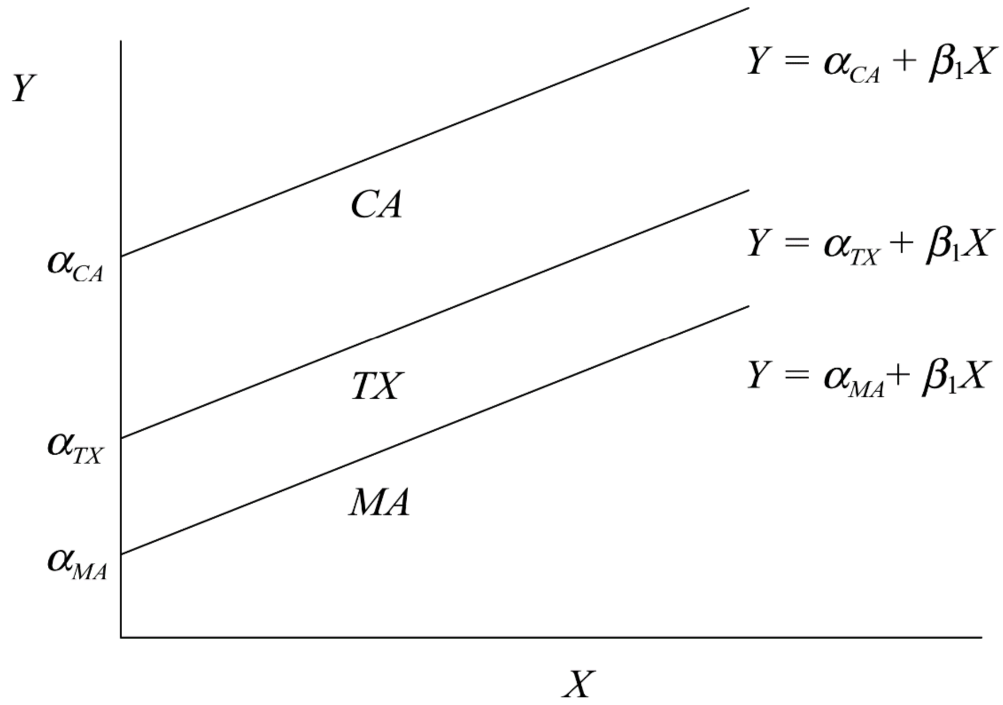
Mettendo insieme le rette dei tre stati:

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

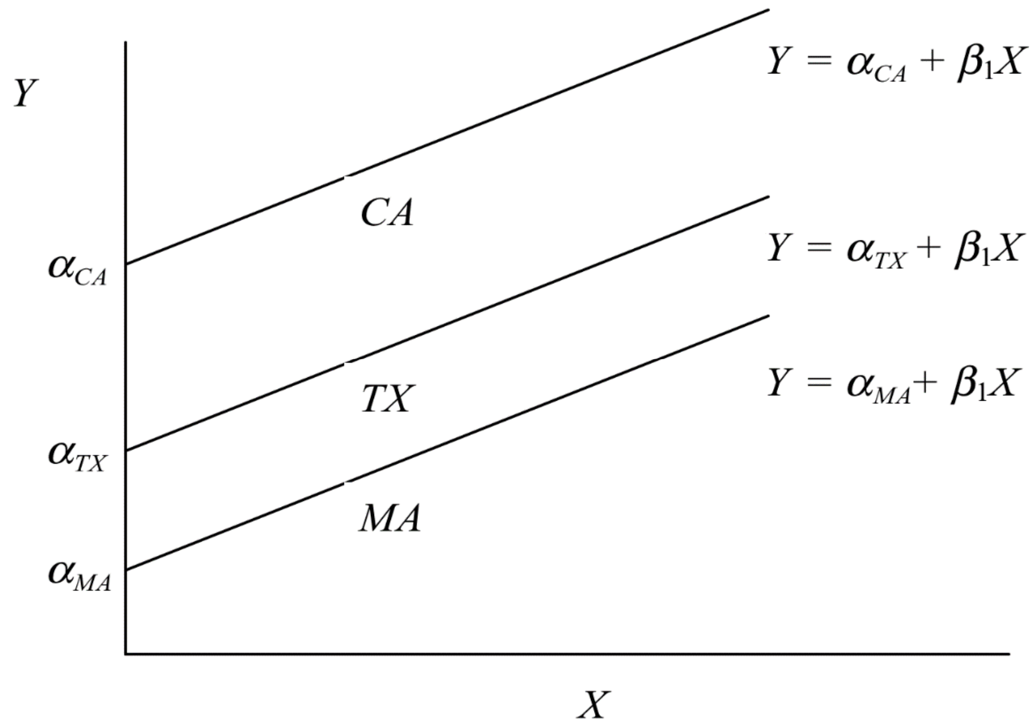
o

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

Le rette di regressione per ciascuno stato



Si ricordi che gli spostamenti nell'intercetta possono essere rappresentati mediante regressori binari…



Nella forma con regressori binari:

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

- DCA_i = 1 se lo stato è *CA*, = 0 altrimenti
- DTX_t = 1 se lo stato è *TX*, = 0 altrimenti
- si lascia fuori DMA_i (*perché*?)

Riepilogo: due modi per scrivere il modello con effetti fissi

1. Forma con "*n*-1 regressori binari" più costante

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

dove $D2_i = \begin{cases} 1 \text{ per } i\text{=2 (stato n. 2)} \\ 0 \text{ altrimenti} \end{cases}$, ecc.

2. Forma con "effetti fissi" :

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

- α_i è chiamato "effetto fisso dello stato" o "effetto dello stato" – è l'effetto costante (fisso) di trovarsi nello stato *i*

Regressione con effetti fissi: stima

Tre metodi di stima:

1. Regressione OLS con "n-1 regressori binari"
2. Regressione OLS con "unità in deviazioni dalle medie"
3. Specificazione "prima e dopo", senza un'intercetta (funziona solo per T = 2)

- Questi tre metodi producono identiche stime dei coefficienti di regressione e identici errori standard.
- Abbiamo già utilizzato la specificazione "prima e dopo" (1988 meno 1982) – che però funziona solo per T = 2 anni
- I metodi 1 e 2 funzionano per un generico T
- Il metodo 1 è praticabile solo quando n non è troppo grande

1. Regressione OLS con "*n*-1 regressori binari"

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

dove $D2_i = \begin{cases} 1 \text{ per } i=2 \text{ (stato n. 2)} \\ 0 \text{ altrimenti} \end{cases}$ etc.

- Prima si creano le variabili binarie $D2_i,\ldots,Dn_i$
- Poi si stima (1) mediante OLS
- L'inferenza (verifiche di ipotesi, intervalli di confidenza) è come di consueto (con errori standard robusti all'eteroschedasticità)
- Non è pratico quando *n* è molto grande (per esempio se *n* = 1000 lavoratori)

2. Regressione OLS con "unità in deviazioni dalle medie"

Modello di regressione con effetti fissi:

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

Le medie delle unità soddisfano:

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

Deviazioni dalle medie:

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

Regressione OLS con "unità in deviazioni dalle medie" (continua)

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

o

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

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

- $\tilde{X}_{it}$ e $\tilde{Y}_{it}$ sono dati "in deviazioni dalle medie"
- Per i=1 e t = 1982, $\tilde{Y}_{it}$ è la differenza tra il tasso di mortalità in Alabama nel 1982 e il suo valor medio in Alabama calcolato in media su tutti e sette gli anni.

Regressione OLS con "unità in deviazioni dalle medie" (continua)

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

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

- Prima si costruiscono le unità in deviazioni dalle medie $\tilde{Y}_{it}$ e $\tilde{X}_{it}$
- Poi si stima (2) con la regressione di $\tilde{Y}_{it}$ su $\tilde{X}_{it}$ usando OLS
- È simile all'approccio "prima e dopo", ma con Y_{it} in deviazione dalla media al posto di Y_{i1}.
- Gli errori standard vanno calcolati in un modo che tenga conto della natura "panel" dei dati (ne parleremo più avanti)
- Si può fare con un unico comando in STATA

Esempio: mortalità stradale e imposte sulla birra in STATA

Prima si indica a STATA che si lavora con dati panel definendo la variabile di unità (state) e quella temporale (year):

```
. xtset state year;
       panel variable:  state (strongly balanced)
        time variable:  year, 1982 to 1988
                delta:  1 unit
```

```
. xtreg vfrall beertax, fe vce(cluster state)

Fixed-effects (within) regression               Number of obs      =       336
Group variable: state                           Number of groups   =        48
R-sq:  within  = 0.0407                         Obs per group: min =         7
       between = 0.1101                                        avg =       7.0
       overall = 0.0934                                        max =         7
                                                F(1,47)            =      5.05
corr(u_i, Xb)  = -0.6885                        Prob > F           =    0.0294

                                 (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6558736   .2918556    -2.25   0.029    -1.243011   -.0687358
       _cons |   2.377075   .1497966    15.87   0.000     2.075723    2.678427
------------------------------------------------------------------------------
```

- Il comando xtreg con l'opzione fe esegue una regressione con effetti fissi. L'intercetta riportata è arbitraria, e i singoli effetti stimati non sono riportati nell'output di default.
- L'opzione fe indica l'uso di regressione con effetti fissi
- L'opzione vce(cluster state) indica a STATA di usare gli errori standard per dati raggruppati (clustered) – ne parleremo più avanti

Esempio (continua). Per *n* = 48, *T* = 7:

$$\widehat{FatalityRate} = -0{,}66BeerTax + \textit{State fixed effects}$$
$$(0{,}29)$$

- Va riportata l'intercetta?
- Quanti regressori binari includereste per una stima con il metodo del "regressore binario"?
- Si confrontino pendenza ed errore standard con la stima per il 1988 v. 1982 della specificazione "prima e dopo" (T = 2, n = 48) (*si noti che è inclusa un'intercetta - ci torniamo poi*):

$$\widehat{FR_{1988} - FR_{1982}} = -0{,}072 - 1{,}04(BeerTax_{1988} - BeerTax_{1982})$$
$$(0{,}065) \quad (0{,}36)$$

Tra l’altro… quanto variano le imposte sulla birra?

Imposte sulla birra nel 2005
Fonte: Federation of Tax Administrators
http://www.taxadmin.org/fta/rate/beer.html

	Accisa ($ per gallone)	Applicazione imposta vendita	Altre imposte
Alabama	$0,53	Sì	$0,52/gallone di imposta locale
Alaska	1,07	ND	$0,35/gallone per piccoli birrifici
Arizona	0,16	Sì	
Arkansas	0,23	Sì	sotto il 3,2% - $0,16/gallone, più imposta di $0,008/gallone e 3% per consumo altrove, 10% per consumo locale
California	0,20	Sì	
Colorado	0,08	Sì	
Connecticut	0,19	Sì	
Delaware	0,16	ND	
Florida	0,48	Sì	2,67¢/12 once di tassa locale al dettaglio

Georgia	0,48	Sì	$0,53/gallone di tassa locale
Hawaii	0,93	Sì	$0,54/gallone sulla birra alla spina
Idaho	0,15	Sì	oltre il 4% - $0,45/gallone
Illinois	0,185	Sì	$0,16/gallone a Chicago e $0,06/gallone a Cook County
Indiana	0,115	Sì	
Iowa	0,19	Sì	
Kansas	0,18	--	oltre il 3,2% - {8% fuori sede e 10% sul posto}, sotto il 3,2% - 4,25% imposta sulla vendita
Kentucky	0,08	Sì*	9% imposta sull'ingrosso
Louisiana	0,32	Sì	$0,048/gallone di imposta locale
Maine	0,35	Sì	addizionale del 5% imposta per consumo locale

Maryland	0,09	Sì	$0,2333/gallone a Garrett County
Massachusetts	0,11	Sì*	0,57% su vendite in club privati
Michigan	0,20	Sì	
Minnesota	0,15	--	sotto il 3,2% - $0,077/gallone, 9% di imposta sulle vendite
Mississippi	0,43	Sì	
Missouri	0,06	Sì	
Montana	0,14	ND	
Nebraska	0,31	Sì	
Nevada	0,16	Sì	
New Hampshire	0,30	ND	
New Jersey	0,12	Sì	
New Mexico	0,41	Sì	

New York	0,11	Sì	$0,12/gallone a New York City
North Carolina	0,53	Sì	$0,48/gallone birra sfusa
North Dakota	0,16	--	7% imposta vendite stato, birra sfusa $0,08/gallone
Ohio	0,18	Sì	
Oklahoma	0,40	Sì	sotto il 3,2% - $0,36/gallone; 13,5% consumo locale
Oregon	0,08	ND	
Pennsylvania	0,08	Sì	
Rhode Island	0,10	Sì	$0,04/cassa imposta sull'ingrosso
South Carolina	0,77	Sì	
South Dakota	0,28	Sì	
Tennessee	0,14	Sì	17% imposta sull'ingrosso
Texas	0,19	Sì	oltre il 4% - $0,198/gallone, 14% consumo locale e $0,05/drink in aeroporto

Utah	0,41	Sì	oltre 3,2% - vendite tramite negozi stato
Vermont	0,265	no	alcool dal 6% all’8% - $0,55; 10% imposta consumo locale
Virginia	0,26	Sì	
Washington	0,261	Sì	
West Virginia	0,18	Sì	
Wisconsin	0,06	Sì	
Wyoming	0,02	Sì	
Dist. of Columbia	0,09	Sì	imposte 8% fuori sede e 10% consumo locale
U.S. Median	$0,188		

Regressione con effetti temporali (Paragrafo 10.4)

Una variabile omessa potrebbe variare nel tempo ma non tra gli stati:

- auto più sicure (air bag, ecc.); modifiche nelle leggi nazionali
- producono intercette che variano nel tempo
- Sia S_t l'effetto combinato di variabili che cambiano nel tempo ma non tra gli stati ("auto più sicure").
- Il modello di regressione risultante è:

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

Soli effetti fissi temporali

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

Questo modello può essere ricomposto con un'intercetta che varia da un anno al successivo:

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

dove $\lambda_{1982} = \beta_0 + \beta_3 S_{1982}$ Similmente,

$$Y_{i,1983} = \lambda_{1983} + \beta_1 X_{i,1983} + u_{i,1983},$$

dove $\lambda_{1983} = \beta_0 + \beta 3 S_{1983}$, ecc.

Due formulazioni di regressione con effetti temporali

1. Formulazione con "T-1 regressori binari":

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

dove $B2_t = \begin{cases} 1 \text{ quando } t=2 \text{ (anno n. 2)} \\ 0 \text{ altrimenti} \end{cases}$, ecc.

2. Formulazione con "effetti fissi" temporali:

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

Effetti temporali: metodi di stima

1. Regressione OLS con "T-1 regressori binari"

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

   - Si creano variabili binarie $B2,\ldots,BT$
   - $B2$ = 1 se t = anno n. 2, = 0 altrimenti
   - Si esegue la regressione di Y su X, $B2,\ldots,BT$ con OLS
   - Dov'è $B1$?

2. Regressione OLS "in deviazione dalle medie dell'anno"
   - Si devia Y_{it}, X_{it} dalle medie dell'anno (non dello stato)
   - Si stima con OLS usando dati "in deviazione dalle medie dell'anno"

Stima con effetti fissi stato ed effetti fissi temporali

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

- Quando $T = 2$, calcolare la differenza prima e includere una costante è equivalente a (fornisce esattamente la stessa regressione di) includere effetti individuali e temporali.
- Quando $T > 2$, esistono vari modi equivalenti di incorporare effetti individuali e temporali:
  - deviazione dalle medie e $T - 1$ indicatori temporali (viene fatto nel seguente esempio con STATA)
  - deviazione dalle medie temporali e $n - 1$ indicatori individuali
  - $T - 1$ indicatori temporali e $n - 1$ indicatori individuali
  - deviazione dalle medie individuali e temporali

```
. gen y83=(year==1983);          Prima si generano tutte le variabili binarie temporali
. gen y84=(year==1984);
. gen y85=(year==1985);
. gen y86=(year==1986);
. gen y87=(year==1987);
. gen y88=(year==1988);
. global yeardum "y83 y84 y85 y86 y87 y88";
. xtreg vfrall beertax $yeardum, fe vce(cluster state);

Fixed-effects (within) regression               Number of obs      =       336
Group variable: state                           Number of groups   =        48
R-sq:  within  = 0.0803                         Obs per group: min =         7
       between = 0.1101                                        avg =       7.0
       overall = 0.0876                                        max =         7
corr(u_i, Xb)  = -0.6781                        Prob > F           =    0.0009
                                 (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6399799   .3570783    -1.79   0.080    -1.358329    .0783691
         y83 |  -.0799029   .0350861    -2.28   0.027    -.1504869   -.0093188
         y84 |  -.0724206   .0438809    -1.65   0.106    -.1606975    .0158564
         y85 |  -.1239763   .0460559    -2.69   0.010    -.2166288   -.0313238
         y86 |  -.0378645   .0570604    -0.66   0.510    -.1526552    .0769262
         y87 |  -.0509021   .0636084    -0.80   0.428    -.1788656    .0770615
         y88 |  -.0518038   .0644023    -0.80   0.425    -.1813645    .0777568
       _cons |    2.42847   .2016885    12.04   0.000     2.022725    2.834215
-------------+----------------------------------------------------------------
```

Gli effetti temporali sono congiuntamente significativi a livello statistico?

```
.   test $yeardum;

 ( 1)  y83 = 0
 ( 2)  y84 = 0
 ( 3)  y85 = 0
 ( 4)  y86 = 0
 ( 5)  y87 = 0
 ( 6)  y88 = 0

       F(  6,    47) =    4.22
            Prob > F =    0.0018
```

Sì

Le assunzioni e gli errori standard della regressione con effetti fissi
(Paragrafo 10.5 e Appendice 10.2)

Sotto le assunzioni dei minimi quadrati nella versione per dati panel, lo stimatore OLS con effetti fissi di β_1 ha distribuzione normale. Tuttavia, è necessario introdurre una nuova formula dell'errore standard, quella per dati raggruppati, o "clustered". Questa nuova formula è necessaria perché le osservazioni per la stessa unità non sono indipendenti (è la stessa unità!), anche se le osservazioni di unità diverse sono indipendenti se tali unità sono ottenute mediante campionamento casuale semplice.

Qui consideriamo il caso di effetti fissi individuali. Gli effetti temporali possono semplicemente essere inclusi quali regressori binari aggiuntivi.

Assunzioni dei minimi quadrati per dati panel

Si consideri una singola X:

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

1. $E(u_{it}|X_{i1},\ldots,X_{iT},\alpha_i) = 0$.
2. $(X_{i1},\ldots,X_{iT},u_{i1},\ldots,u_{iT})$, $i = 1,\ldots,n$, sono i.i.d. dalla distribuzione congiunta.
3. (X_{it}, u_{it}) hanno momenti quarti finiti.
4. Non vi è collinearità perfettà (molteplicità di X)

Le assunzioni 3 e 4 sono identiche al caso dei minimi quadrati, le assunzioni 1 e 2 sono diverse.

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

- u_{it} ha media zero, dato l'effetto fisso *e* l'intera storia delle X per l'unità corrispondente
- Questa è un'estensione della precedente assunzione 1 della regressione multipla
- Ciò significa che non vi sono effetti passati omessi (qualsiasi effetto passato di X deve essere incluso esplicitamente)
- Inoltre, non c'è feedback da u su X futuri:
  - il fatto che uno stato abbia un tasso di mortalità particolarmente alto quest'anno non influisce sull'aumento delle imposte sulla birra.
  - talvolta questa assunzione di "assenza di feedback" è plausibile, talvolta no. Ci torneremo quando affronteremo le serie temporali.

Assunzione 2: $(X_{i1},\ldots,X_{iT},u_{i1},\ldots,u_{iT})$, $i=1,\ldots,n$, sono i.i.d. dalla distribuzione congiunta.

- È un'estensione dell'assunzione 2 per la regressione multipla con dati sezionali
- È soddisfatta se le unità sono prese a caso dalla popolazione mediante campionamento casuale semplice.
- ***Non*** richiede che le osservazioni siano i.i.d. *nel tempo* per la stessa unità – sarebbe irrealistico. Il fatto che uno stato abbia un'imposta sulla birra elevata quest'anno è un buon predittore del (è correlato con) fatto che avrà un'imposta sulla birra elevata l'anno seguente. Similmente, il termine d'errore per un'unità in un anno è plausibilmente correlato con il suo valore l'anno dopo, cioè corr(u_{it}, u_{it+1}) è plausibilmente diverso da zero.

Autocorrelazione (correlazione seriale)

Supponiamo che una variabile Z sia osservata in diverse date t, perciò le osservazioni sono su Z_t, $t = 1,\ldots, T$. (consideriamo che vi sia una sola unità). Allora Z_t è detta ***autocorrelata*** o ***serialmente correlata*** se corr(Z_t, Z_{t+j}) ≠ 0 per date $j \neq 0$.

- "Autocorrelazione" significa correlazione con se stesso.
- cov(Z_t, Z_{t+j}) è detta ***j-esima autocovarianza*** di Z_t.
- Nell'esempio della guida in stato di ebbrezza, u_{it} include la variabile omessa delle condizioni meteo dell'anno per lo stato i. Se gli inverni nevosi si presentano in gruppi (uno segue l'altro), allora u_{it} sarà autocorrelata (*perché*?)
- In molte applicazioni con dati panel, u_{it} è plausibilmente autocorrelata.

Indipendenza e autocorrelazione in dati panel, un'immagine:

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

← Campionamento i.i.d. tra le unità →

- Se le unità sono ottenute per campionamento casuale semplice, allora $(u_{i1},\dots, u_{iT})$ è indipendente da $(u_{j1},\dots, u_{jT})$ per unità diverse con $i \neq j$.
- Ma se i fattori omessi compresi in u_{it} sono serialmente correlati, allora u_{it} è serialmente correlato.

Sotto le assunzioni dei minimi quadrati per dati panel:

- Lo stimatore OLS con effetto fisso $\hat{\beta}_1$ è non distorto, consistente e ha distribuzione asintotica normale
- Tuttavia, i consueti errori standard OLS (sia di omoschedasticità pura sia robusti all'eteroschedasticità) saranno in generale sbagliati perché assumono che u_{it} non sia serialmente correlata.
  - In pratica, gli errori standard OLS spesso sottostimano l'incertezza del campionamento reale: se u_{it} è correlato nel tempo, non si hanno molte informazioni (molta variazione casuale) come si avrebbero se u_{it} fosse incorrelata.
  - Il problema si risolve usando errori standard "clustered".

Errori standard per dati raggruppati

- Gli errori standard per dati raggruppati stimano la varianza di $\hat{\beta}_1$ quando le variabili sono i.i.d. tra le unità ma sono potenzialmente autocorrelate in una unità.
- È più facile comprenderli se si considera prima il problema più semplice di stimare la media di *Y* usando dati panel…

Errori standard clustered per la media stimata con dati panel

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

Lo stimatore della media μ è $\bar{Y} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} Y_{it}$.

È utile scrivere $\bar{Y}$ come media tra le unità del valore medio per ciascuna unità:

$$\bar{Y} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} Y_{it} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T} Y_{it}\right) = \frac{1}{n}\sum_{i=1}^{n}\bar{Y}_i \quad ,$$

dove $\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it}$ è la media campionaria per l'unità *i*.

Poiché le osservazioni sono i.i.d. tra le entità, $(\bar{Y}_1,\ldots,\bar{Y}_n)$ sono i.i.d. Quindi, se *n* è grande, vale il TLC e

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n}\bar{Y}_i \xrightarrow{d} N(0, \sigma^2_{\bar{Y}_i}/n),\ \text{dove}\ \sigma^2_{\bar{Y}_i} = \text{var}(\bar{Y}_i).$$

- L'errore standard di $\bar{Y}$ è la radice quadrata di uno stimatore di $\sigma^2_{\bar{Y}_i}/n$.
- Lo stimatore naturale di $\sigma^2_{\bar{Y}_i}$ è la varianza campionaria di $\bar{Y}_i$, $s^2_{\bar{Y}_i}$.

Questo ci fornisce la formula dell'errore standard per dati raggruppati per $\bar{Y}$ calcolata usando dati panel:

$$\textit{Errore standard Clustered di } \bar{Y} = \sqrt{\frac{s^2_{\bar{Y}_i}}{n}},\ \text{dove}\ s^2_{\bar{Y}_i} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\bar{Y}_i - \bar{Y}\right)^2$$

Che cos'hanno di speciale gli errori standard per dati raggruppati?

- Non molto, in verità – la procedura di derivazione vista in precedenza è la stessa usata nel Capitolo 3 per derivare l'errore standard della media campionaria, con la differenza che qui i "dati" sono le medie di unità i.i.d. ($\bar{Y}_1, \ldots, \bar{Y}_n$) anziché una singola osservazione i.i.d. per ciascuna unità.

- C'è una caratteristica importante: nella derivazione dell'errore standard per dati raggruppati non abbiamo mai assunto che le osservazioni siano i.i.d. *in* una unità. Quindi abbiamo implicitamente consentito la correlazione seriale in una unità.

- E la correlazione seriale, dov'è finita? Determina $\sigma^2_{\bar{Y}_i}$, la varianza di $\bar{Y}_i$ …

La correlazione seriale in Y_{it} entra in $\sigma^2_{\bar{Y}_i}$:

$$\sigma^2_{\bar{Y}_i} = \text{var}(\bar{Y}_i)$$
$$= \text{var}\left(\frac{1}{T}\sum_{t=1}^{T} Y_{it}\right) = \frac{1}{T^2}\text{var}\left(Y_{i1} + Y_{i2} \ldots + Y_{iT}\right)$$
$$= \frac{1}{T^2}\{\text{var}(Y_{i1}) + \text{var}(Y_{i2}) + \ldots + \text{var}(Y_{iT})$$
$$+ 2\text{cov}(Y_{i1},Y_{i2}) + 2\text{cov}(Y_{i1},Y_{i3}) + \ldots + 2\text{cov}(Y_{iT-1},Y_{iT})\}$$

- Se Y_{it} è serialmente incorrelata, tutte le autocovarianze = 0 e abbiamo la consueta derivazione del Capitolo 3.
- Se queste autocovarianze non sono zero, la formula consueta (che le pone a 0) sarà errata.
- Se queste autocovarianze sono positive, la formula consueta sottostimerà la varianza di $\bar{Y}_i$.

La "magia" degli errori standard per dati raggruppati è che, operando al livello delle unità e delle loro medie $\bar{Y}_i$, non occorre preoccuparsi di stimare le autocovarianze sottostanti, che sono stimate automaticamente dalla formula dell'errore standard. Ecco i calcoli:

Errore standard clustered di $\bar{Y} = \sqrt{s^2_{\bar{Y}_i} / n}$, dove

$$s^2_{\bar{Y}_i} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\bar{Y}_i - \bar{Y}\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}Y_{it} - \bar{Y}\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\left(Y_{it} - \bar{Y}\right)\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\left(Y_{it}-\bar{Y}\right)\right)\left(\frac{1}{T}\sum_{s=1}^{T}\left(Y_{is}-\bar{Y}\right)\right)$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(Y_{is}-\bar{Y}\right)\left(Y_{it}-\bar{Y}\right)$$

$$= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_{is}-\bar{Y}\right)\left(Y_{it}-\bar{Y}\right)\right]$$

- Il termine finale tra quadre, $\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_{is}-\bar{Y}\right)\left(Y_{it}-\bar{Y}\right)$, stima l'autocovarianza tra Y_{is} e Y_{it}. Quindi la formula dell'errore standard clustered implicitamente stima tutte le autocovarianze, usandole per stimare $\sigma^2_{\bar{Y}_i}$!
- Per contrasto, la formula "consueta" pone a zero queste autocovarianze omettendo tutti i termini misti – il che è valido solo se queste autocovarianze sono tutte zero.

Errori standard clustered per lo stimatore con effetti fissi nella regressione con dati panel

- Il concetto di errori standard clustered per dati panel è del tutto analogo al precedente caso della media per dati panel – solo più complesso per notazione e formule. Si veda l'Appendice 10.2.
- Gli errori standard clustered per dati panel sono l'estensione logica di quelli robusti all'eteroschedasticità per dati sezionali. Nella regressione con dati sezionali, gli errori standard robusti all'eteroschedasticità sono validi indipendentemente dal fatto che vi sia eteroschedasticità. Nella regressione con dati panel, gli errori standard clustered sono validi indipendentemente dal fatto che vi sia eteroschedasticità e/o correlazione seriale.
- *Tra l'altro…* Il termine "clustered" deriva dal fatto che si consente correlazione in un "cluster" (o "gruppo") di osservazioni (in una entità) ma non tra cluster.

Errori standard clustered: implementazione in STATA

```
. xtreg vfrall beertax, fe vce(cluster state)

Fixed-effects (within) regression               Number of obs      =       336
Group variable: state                           Number of groups   =        48
R-sq:  within  = 0.0407                         Obs per group: min =         7
       between = 0.1101                                        avg =       7.0
       overall = 0.0934                                        max =         7
                                                F(1,47)            =      5.05
corr(u_i, Xb)  = -0.6885                        Prob > F           =    0.0294

                                 (Std. Err. adjusted for 48 clusters in state)
------------------------------------------------------------------------------
             |               Robust
      vfrall |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.6558736   .2918556    -2.25   0.029    -1.243011   -.0687358
       _cons |   2.377075   .1497966    15.87   0.000     2.075723    2.678427
------------------------------------------------------------------------------
```

- vce(cluster state) indica di usare errori standard clustered, dove il raggruppamento è a livello di stato (osservazioni che hanno lo stesso valore della variabile "state" possono essere correlati, ma si assume che siano incorrelati se il valore di "state" è diverso)

Applicazione: leggi sulla guida in stato di ebbrezza e mortalità stradale (Paragrafo 10.6)

Alcuni fatti

- Circa 40.000 morti sulle strade ogni anno negli USA
- 1/3 degli incidenti mortali coinvolge un guidatore ubriaco
- 25% dei guidatori sulle strade tra l'1 e le 3 del mattino ha bevuto (stima)
- Un guidatore ubriaco ha 13 volte più probabilità di causare un incidente mortale rispetto a un guidatore sobrio (stima)

Leggi sulla guida in stato di ebbrezza e mortalità stradale (continua)

Aspetti di politica publica

- La guida in stato di ebbrezza causa importanti esternalità (guidatori sobri vengono uccisi, la società sostiene costi medici, ecc.) – vi è ampia giustificazione per un intervento del governo
- Esistono modi efficaci per ridurre la guida in stato di ebbrezza? Se sì, quali?
- Quali sono gli effetti di leggi specifiche:
  - pene obbligatorie
  - età minima legale per bere alcolici
  - interventi economici (imposte sugli alcolici)

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

Il Governatore Mitt Romney ha firmato oggi la più severa legge contro la guida in stato di ebbrezza nella storia del Commonwealth.

La nuova legge, così chiamata in onore della tredicenne Melanie Powell, fisserà pene più severe per incidenti dovuti a guida in stato di ebbrezza in Massachusetts e chiuderà qualsiasi spazio nel sistema legislativo che possa consentire ai guidatori ubriachi recidivi di tornare al volante.

"Oggi rendiamo onore a coloro che hanno perso la vita in insensati incidenti dovuti a guida in stato di ebbrezza, e agiamo per salvare le vite che altrimenti rischieremmo di perdere il prossimo anno", ha detto Romney. "Oggi abbiamo la Melanie's Law perché i cittadini del Commonwealth hanno fatto in modo che ciò accadesse".
La nuova misura fornisce al pubblico ministero il potere di presentare documenti per provare che un recidivo è già stato condannato per guida in stato di ebbrezza. Inoltre, la pena minima obbligatoria per qualsiasi persona ritenuta colpevole di omicidio con mezzo motorizzato sarà aumentata da 2 anni e ½ a cinque anni.
I recidivi dovranno installare un dispositivo di blocco su qualsiasi veicolo che possiedano o usino. Questi dispositivi misurano il tasso alcolico e impediscono l'avvio dell'auto se il guidatore supera il limite. Chiunque alteri il dispositivo di blocco rischia una condanna penale.

Per la prima volta il Massachusetts rispetterà gli standard federali per la legislazione sulla guida in stato di ebbrezza.
Romney è stato raggiunto da Tod e Nancy Powell, i genitori di Melanie Powell, e da suo nonno, Ron Bersani, per celebrare l'entrata in vigore della nuova legge sulla guida in stato di ebbrezza.

"Oggi dobbiamo ringraziare tutti coloro che hanno lavorato duramente per rendere possibile questo giorno", ha detto Bersani. "Il governatore Romney e l'assemblea legislativa hanno fatto progredire la lotta contro la guida in stato di ebbrezza fino a un punto che sembrava irraggiungibile soltanto sei mesi fa.
La legge inasprisce le pene per chi guida in stato di ebbrezza con in auto un bambino minore di 14 anni e per chi guida con un tasso alcolico di 0,20 o superiore, più del doppio del limite di legge.
Romney ha ringraziato l'assemblea per aver approvato una legge severa che combatte la guida in stato di ebbrezza in Massachusetts.
"La sicurezza pubblica è una delle principali priorità e la Melanie's Law renderà noi cittadini e automobilisti più sicuri", ha detto lo speaker Salvatore F. DiMasi. "Voglio elogiare i miei colleghi dell'assemblea e il governatore per aver intrapreso un'azione rapida ed efficace su questo tema importantissimo".

"Oggi comunichiamo un messaggio forte: che il Massachusetts prende sul serio la lotta contro la guida in stato di ebbrezza", ha detto il leader della minoranza Bradley H. Jones Jr. "Sono orgoglioso del Governatore, del Vicegovernatore e dei miei colleghi dell'assemblea per aver promosso leggi che rendono le nostre strade più sicure".

"Sono felice e orgoglioso che l'assemblea alla fine abbia fatto la cosa giusta sostenendo una legge degna del nome di Melanie e dei sacrifici fatti dalla famiglia Powell e da tutte le vittime di guidatori ubriachi", ha detto il senatore Robert L. Hedlund. "La Melanie's Law salverà vite e non sarebbe stata possibile senza gli sforzi incessanti dei familiari".

Il delegato Frank Hynes ha aggiunto: "Desidero elogiare Ron, Tod e Nancy per il loro grande lavoro in sostegno della legge. Come famiglia hanno saputo trasformare quell'orribile tragedia in una grande misura di sicurezza per tutte le famiglie sulle strade del Massachusetts".

###

Dati panel per la guida in stato di ebbrezza

n = 48 stati USA, T = 7 anni (1982,…,1988) (bilanciato)

Variabili

- Tasso di mortalità stradale (morti per 10.000 residenti)
- Imposta su una cassa di birra (*Beertax*)
- Età minima di legge per bere alcolici
- Pene minime per la prima violazione:
  - *Pena obbligatoria*
  - *Servizio sociale obbligatorio*
  - altrimenti, la sentenza sarà soltanto pecuniaria
- Miglia per veicolo per guidatore (US DOT)
- Dati economici sullo stato (reddito pro capite, ecc.)

Perché i dati panel potrebbero aiutare?

- Potenziale distorsione da variabili omesse per variabili che variano tra stati ma sono costanti nel tempo:
  - cultura del bere e del guidare
  - qualità delle strade
  - età delle automobili sulle strade
    - usa effetti fissi di stato
- Potenziale distorsione da variabili omesse per variabili che variano nel tempo ma sono costanti tra stati:
  - miglioramenti nella sicurezza delle auto nel tempo
  - mutamento atteggiamenti verso la guida in stato di ebbrezza a livello nazionale
    - usa effetti temporali

Tabella 10.1 Analisi degli effetti della legislazione in materia di guida in stato d'ebbrezza sulla mortalità stradale.

Variabile dipendente: tasso di mortalità stradale (morti su 10.000 abitanti)

Regressore	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Imposta sulla birra	0,36** (0,05)	– 0,66* (0,29)	– 0,64+ (0,36)	– 0,45 (0,30)	– 0,69* (0,35)	– 0,46 (0,31)	– 0,93** (0,34)
Età minima legale 18				0,028 (0,070)	– 0,010 (0,083)		0,037 (0,102)
Età minima legale 19				– 0,018 (0,050)	– 0,076 (0,068)		– 0,065 (0,099)
Età minima legale 20				0,032 (0,050)	– 0,100+ (0,056)		– 0,113 (0,125)
Età minima legale						– 0,002 (0,021)	
Pena detentiva o servizi per la comunità				0,038 (0,103)	0,085 (0,111)	0,039 (0,103)	0,089 (0,164)
Miglia medie per guidatore				0,008 (0,007)	0,017) (0,011)	0,009 (0,007)	0,124 (0,049)
Tasso di disoccupazione				– 0,063** (0,013)		– 0,063** (0,013)	– 0,091** (0,021)
Reddito reale pro capite (logaritmo)				1,82** (0,64)		1,79** (0,64)	1.00 (0,68)
Anni	1982-88	1982-88	1982-88	1982-88	1982-88	1982-88	solo 1982 e 1988
Effetti fissi di stato?	no	sì	sì	sì	sì	sì	sì
Effetti temporali?	no	no	sì	sì	sì	sì	sì
Errori standard per i dati raggruppati?	no	sì	sì	sì	sì	sì	sì

Statistiche F e valori-p per l'esclusione di gruppi di variabili							
Effetti temporali = 0			4,22 (0,002)	10,12 (< 0,001)	3,48 (0,006)	10,28 (< 0,001)	37,49 (< 0,001)
Coefficienti età minime legali = 0				0,35 (0,786)	1,41 (0,253)		0,42 (0,738)
Tasso di disoccupazione e reddito pro capite = 0				29,62 (< 0,001)	31,96 (< 0,001)		25,20 (< 0,001)
$\bar{R}^2$	0,091	0,889	0,891	0,926	0,893	0,926	0,899

Queste regressioni sono state stimate utilizzando i dati panel per 48 stati USA. Le regressioni da (1) a (6) utilizzano dati per tutti gli anni dal 1982 al 1988, mentre la regressione (7) utilizza solo dati del 1982 e del 1988. I dati sono descritti nell'Appendice 10.1. Gli errori standard sono riportati tra parentesi sotto i coefficienti, e i valori-p sono riportati tra parentesi sotto le statistiche F. I coefficienti sono statisticamente significativi al livello del $^{+}$10%, *5% o **1%.

Analisi empirica: risultati principali

- Il segno del coefficiente dell’imposta sulla birra cambia quando sono inclusi gli effetti fissi dello stato
- Gli effetti temporali sono statisticamente significativi ma la loro inclusione non ha un grande impatto sui coefficienti stimati
- L’effetto stimato dell’imposta sulla birra cala quando si includono altre leggi.
- L’unica variabile politica che sembra avere un impatto è l’imposta sulla birra – non l’età legale minima per bere alcolici, non la pena minima obbligatoria ecc. – tuttavia l’imposta sulla birra non è significativa anche al livello del 10% usando errori standard clustered nelle specifiche che controllano per le condizioni economiche dello stato (tasso di disoccupazione, reddito personale)

Risultati empirici (continua)

- In particolare, l'età legale minima per bere alcolici ha un coefficiente piccolo che è stimato con precisione – riducendola non pare si abbia un grande effetto sulla mortalità stradale complessiva.
- Quali sono le minacce alla validità interna? Cosa si può dire su:

1. Distorsione da variabili omesse
2. Errata forma funzionale
3. Distorsione da errori nelle variabili
4. Distorsione da selezione del campione
5. Distorsione da causalità simultanea

Che cosa ne pensate?

Digressione: estensioni del concetto di "n-1 regressori binari"

L'idea di utilizzare molti indicatori binari per eliminare la distorsione da variabili omesse può essere estesa a dati non panel – la chiave è che la variabile omessa sia costante per un gruppo di osservazioni, il che in effetti significa che ciascun gruppo ha la propria intercetta.

Esempio: effetto della dimensione delle classi.

Supponiamo che livelli di finanziamento e di istruzione siano determinati a livello della contea, e che ogni contea abbia diversi distretti. Se si è preoccupati della distorsione da variabili omesse risultante da variabili non osservate a livello di contea, si possono includere gli effetti di contea (indicatori binari, uno per ciascuna contea, omettendo una sola contea per evitare la collinearità perfetta).

Riepilogo: regressione con dati panel (Paragrafo 10.7)

Vantaggi e limitazioni della regressione con effetti fissi

Vantaggi

- Si può controllare per variabili non osservate che:
  - variano tra stati ma non nel tempo e/o
  - variano nel tempo ma non tra stati
- Più osservazioni forniscono più informazioni
- La stima coinvolge estensioni relativamente semplici della regressione multipla

- La regressione con effetti fissi si può eseguire in tre modi:

1. Metodo "prima e dopo" quando $T = 2$
2. "n-1 regressori binari" quando n è piccolo
3. Regressione "in deviazione dalle medie"

- Metodi simili si applicano alla regressione con effetti temporali e a quella con effetti fissi e temporali
- Inferenza statistica: come nella regressione multipla.

Limitazioni/problemi aperti

- Necessaria la variazione in X nel tempo nelle entità
- Gli effetti di ritardo temporale possono essere importanti – anche se non ne abbiamo tenuto conto nel modello dell'imposta sulla birra
- È necessario usare errori standard clustered per evitare la possibilità che u_{it} sia autocorrelato