



Dipartimento di
Economia e Statistica
Cognetti de Martiis

UNIVERSITÀ
DEGLI STUDI
DI TORINO



La valutazione di impatto delle politiche: modelli di studio

Roberto Leombruni

Università di Torino e Laboratorio Revelli

HEALTH EQUITY AUDIT

Bari, 7 ottobre 2015

Menù della lezione

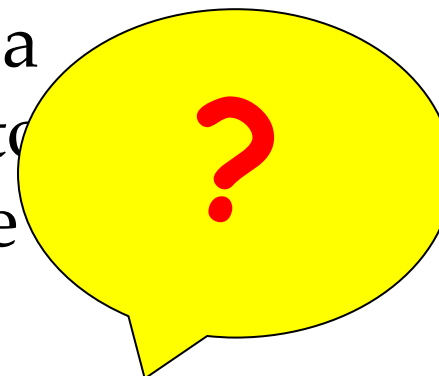
- ❖ dai *randomized clinical trial* agli esperimenti sociali
- ❖ gli esperimenti in 3 disegni
- ❖ dalle stime sperimentali alle stime... naïve
- ❖ una “mappa” dei modelli quasi-sperimentali
 - ❖ stimatori delle “differenze nelle differenze”
 - ❖ stimatori basati su matching
 - ❖ esperimenti naturali
 - ❖ *regression discontinuity design*

Dai *randomized clinical trial* agli esperimenti sociali

Gli esperimenti randomizzati e controllati sono da molti considerati il “gold standard” per la valutazione delle relazioni di causa effetto di fuori del campo medico e delle scienze

→ anche in ambito economico sociale, implementando i c.d. “**esperimenti sociali**”

→ lo scopo è quello di contribuire al disegno di **politiche basate sull’evidenza**, e poi di valutare a posteriori la loro efficacia.



Dai *randomized clinical trial* agli esperimenti sociali

Nei paesi anglosassoni è una prassi consolidata quella di organizzare dei *pilot* per valutare se una certa politica ha gli effetti ipotizzati.

➤ il **project STAR**, lanciato negli anni Ottanta per valutare l'effetto di alcune classi sull'apprendimento, ha rivelato che
→ esistono ormai evidenze del **basic income** in molti paesi sviluppati e in via di sviluppo.

in **CANADA**, durante gli anni 70, l'esperimento **mincome** ha rivelato effetti di disincentivo al lavoro trascurabili, ma effetti positivi su istruzione e salute

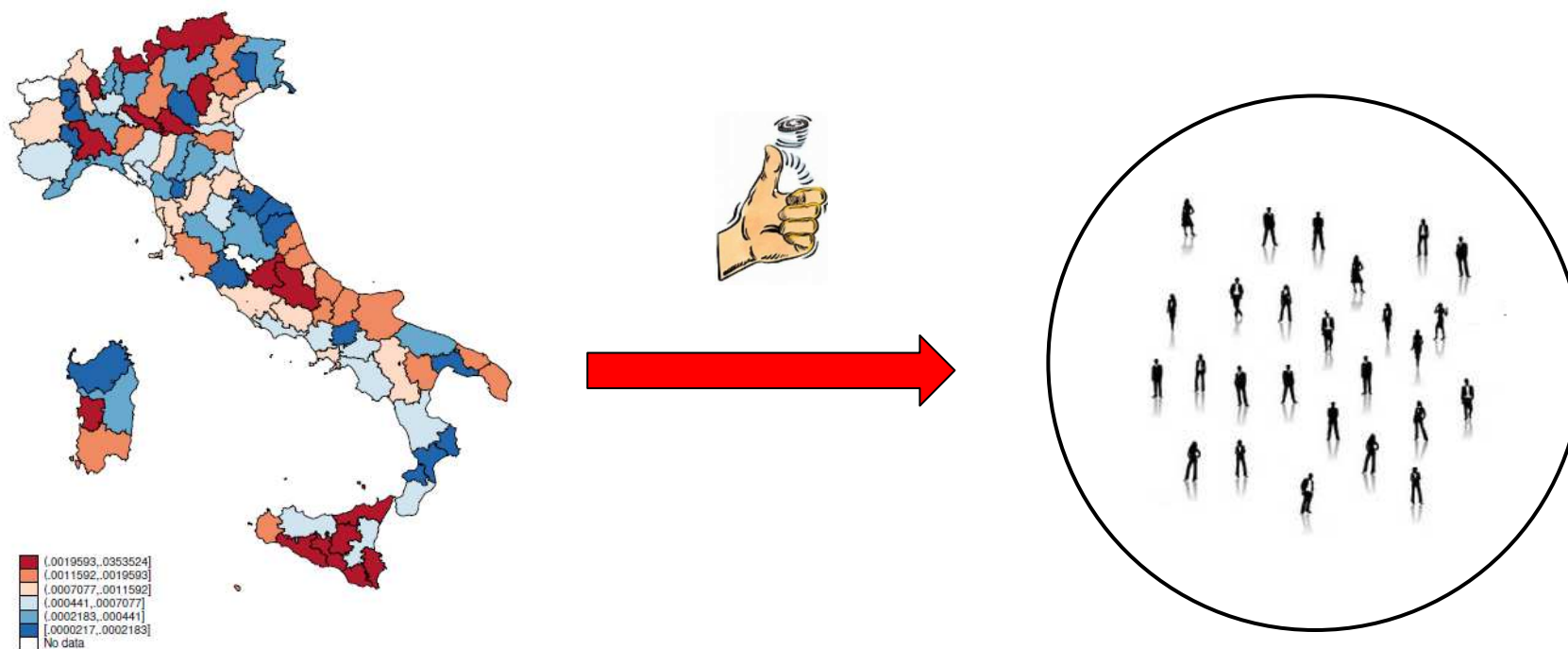
Dai *randomized clinical trial* agli esperimenti sociali

Non sempre però è possibile (anche avendo la buona volontà!), organizzare un esperimento sociale. I motivi possono riguardare la loro fattibilità **politica**, ma anche la loro fattibilità **pratica**.

Inoltre, può darsi che la nostra “curiosità” riguardi una politica o un altro evento che si è già verificato nel passato! Quindi non abbiamo altra possibilità che rivolgerci a dati *osservazionali*, vale a dire a quanto si è spontaneamente verificato, senza nessun intervento sperimentale.

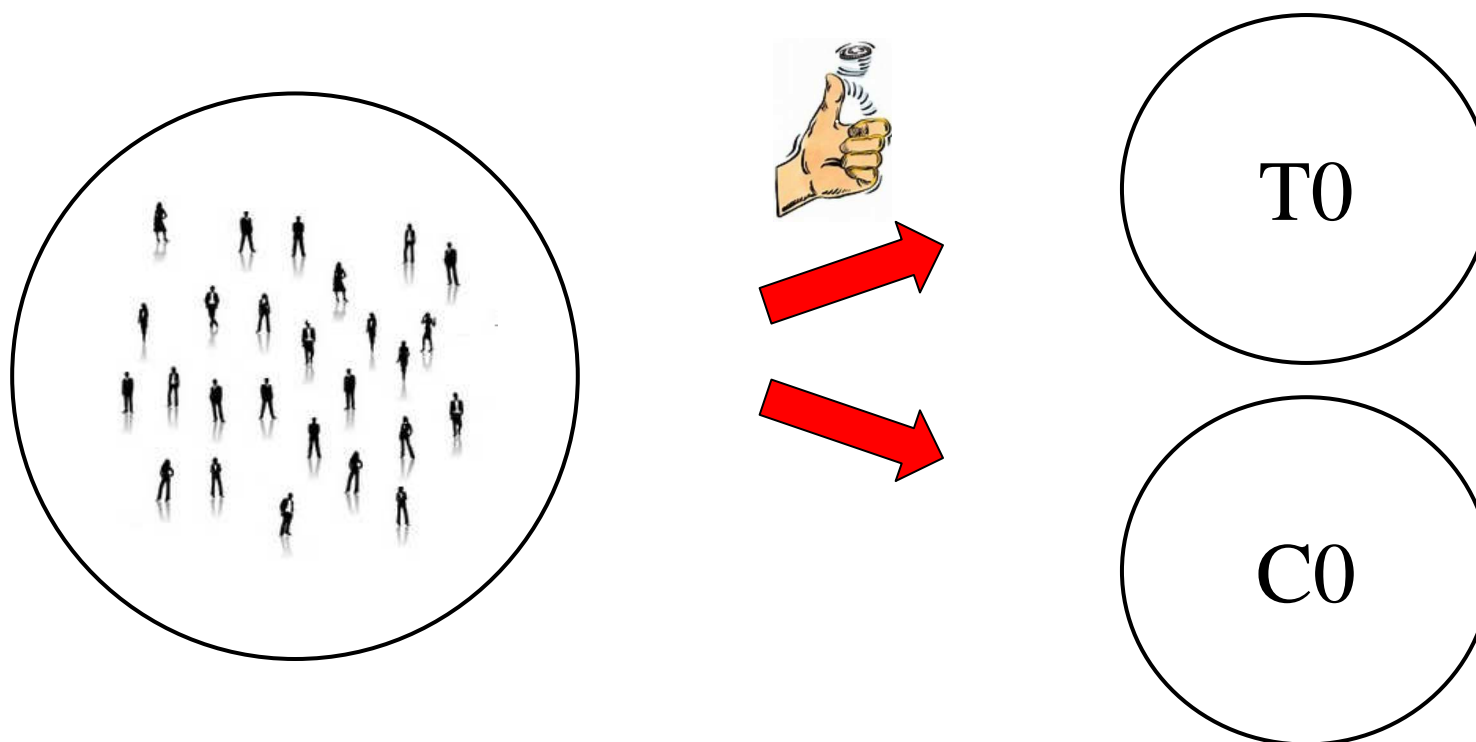
Gli esperimenti in 3 disegni

Negli esperimenti controllati randomizzati vi è, possibilmente (ma non sempre!) una prima randomizzazione, che riguarda la selezione della **popolazione in studio**.



Gli esperimenti in 3 disegni

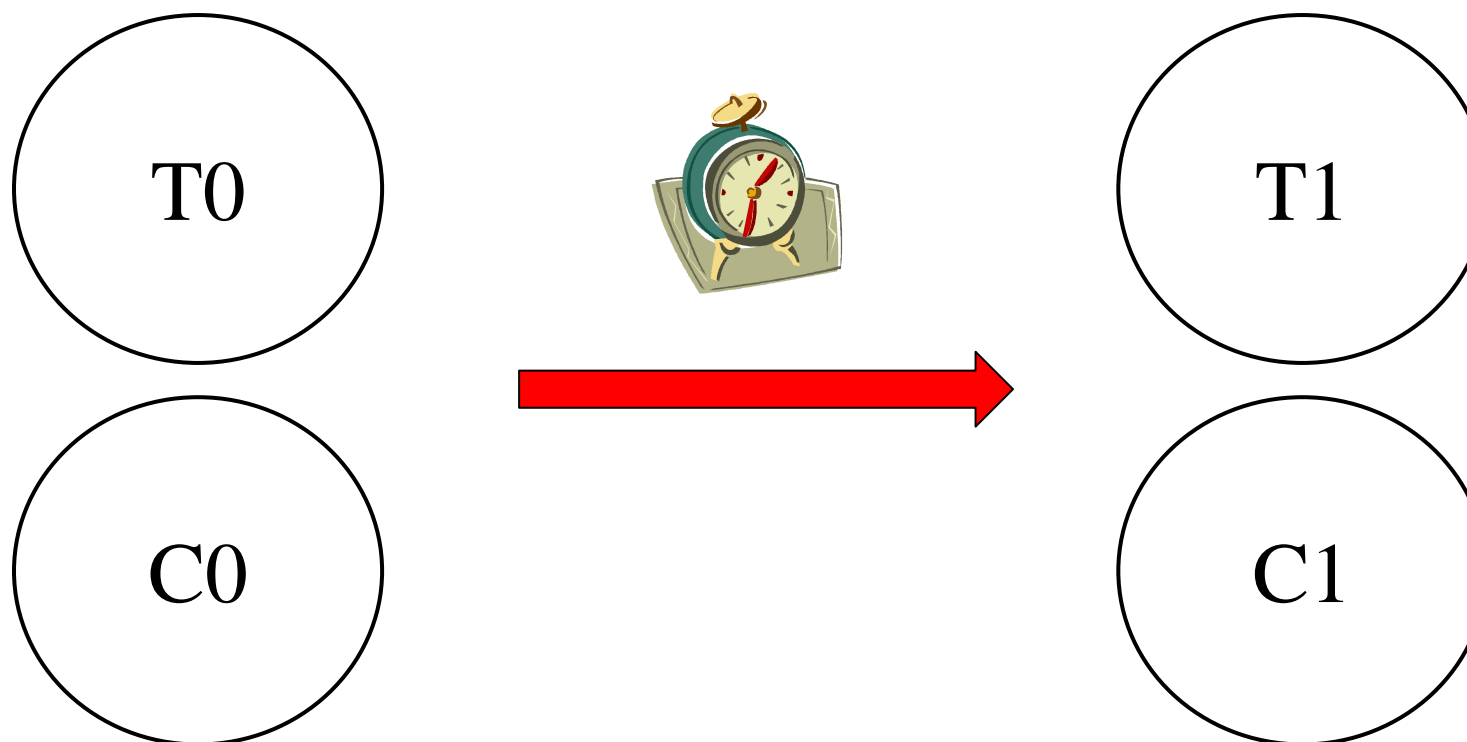
La seconda randomizzazione (questa c'è sempre!) riguarda la divisione della popolazione in studio in due gruppi, quello dei **trattati** e quello dei **controlli**.



Gli esperimenti in 3 disegni

Una volta assegnato il trattamento, dobbiamo solo aspettare per vedere cosa succede.

→ **Come misuriamo l'effetto del trattamento?**



Gli esperimenti in 3 disegni

Verrebbe naturale dire che l'effetto del trattamento è la differenza al tempo 1 nel risultato per i trattati e i controlli, che potremmo scrivere come:

$$(1) \quad \beta = Y_{T1} - Y_{C1}$$

Una formulazione più precisa dice che **SE** la selezione della popolazione in studio è stata corretta, e **SE** la randomizzazione tra trattati e controlli è stata corretta, **ALLORA** il β della (1) è una **STIMA** dell'effetto del trattamento.

Dalle stime sperimentali alle stime... naïve

La precisazione è dovuta perché (ahimè) nell'ambito delle politiche economiche e del lavoro sono molto comuni le due seguenti definizioni naïve di effetto, utilizzate **al di fuori** di contesti sperimentali:

$$(1) \quad \beta = Y_{T1} - Y_{C1}$$

$$(2) \quad \beta = Y_{T1} - Y_{T0}$$

Dalle stime sperimentali alle stime... naïve

Per chiarire l'errore retrostante, e impostare correttamente il tema della valutazione delle politiche, è prassi partire innanzitutto da una definizione più rigorosa di effetto del trattamento, basata sul concetto di **outcome potenziale**:

$Y^1 \rightarrow$ il risultato che si otterrebbe se l'individuo ricevesse il trattamento.

$Y^0 \rightarrow$ il risultato che si otterrebbe se l'individuo NON ricevesse il trattamento.

Dalle stime sperimentali alle stime... naïve

Risulta così definito in modo concettualmente chiaro cosa si intende per **effetto del trattamento**:

$$\beta = Y^1 - Y^0$$

Il problema è che si verifica solo una delle due situazioni (il c.d. **fattuale**), e cioè o l'individuo riceve, oppure NON riceve il trattamento.

Ma per calcolare l'effetto del trattamento ci servono entrambi gli addendi: dobbiamo sapere anche cosa succederebbe nella situazione **controfattuale**.

Dalle stime sperimentali alle stime... naïve

È quello che Holland definì il “**fundamental problem of causal inference**”. E che, come è stato fatto notare, si riduce di fatto a un problema di dati mancanti... non si può calcolare su base individuale l’effetto di un trattamento, perché ci manca il controfattuale.

Riscriviamo i nostri stimatori naïve secondo la notazione degli outcome potenziali...

Dalle stime sperimentali alle stime... naïve

Partiamo dal primo stimatore naïve, che riscriviamo utilizzando la notazione degli OP:

$$(1) \quad \beta = Y^1_{T1} - Y^0_{C1}$$

aggiungiamo e togliamo il termine Y^0_{T1} :

$$(1) \quad \beta = (Y^1_{T1} - Y^0_{T1}) + \\ (Y^0_{T1} - Y^0_{C1})$$

Dalle stime sperimentali alle stime... naïve

Partiamo dal primo stimatore naïve, che riscriviamo utilizzando la notazione degli OP:

$$(1) \quad \beta = Y^1_{T1} - Y^0_{C1}$$

aggiungiamo e togliamo il termine Y^0_{T1} :

$$(1) \quad \beta = (Y^1_{T1} - Y^0_{T1}) + \rightarrow ATT$$

$$(Y^0_{T1} - Y^0_{C1}) \rightarrow \textit{selection bias}$$

Dalle stime sperimentali alle stime... naïve

Analogamente con il secondo stimatore naïve:

$$(2) \quad \beta = Y^1_{T1} - Y^0_{T0}$$

aggiungiamo e togliamo il termine Y^0_{T1} :

$$(2) \quad \beta = (Y^1_{T1} - Y^0_{T1}) + \\ (Y^0_{T1} - Y^0_{T0})$$

Dalle stime sperimentali alle stime... naïve

Analogamente con il secondo stimatore naïve:

$$(2) \quad \beta = Y^1_{T1} - Y^0_{T0}$$

aggiungiamo e togliamo il termine Y^0_{T1} :

$$(2) \quad \beta = (Y^1_{T1} - Y^0_{T1}) + \rightarrow ATT$$

$$(Y^0_{T1} - Y^0_{T0}) \rightarrow \textit{trend naturale}$$

Una “mappa” dei modelli quasi-sperimentali

I modelli per la valutazione di effetto che facciamo rientrare nella famiglia dei “quasi”-esperimenti si possono dividere in due categorie:

- Modelli in cui non vi è alcuna randomizzazione, ma viene “mimato” un disegno sperimentale
 - il modello **diff-in-diff**
 - le serie temporali interrotte
 - stime basate su **matching**
- Modelli di studio in cui vi è una parziale randomizzazione:
 - eventi naturali
 - cambiamenti non naturali ma esogeni
 - discontinuità nella politica

Il modello delle “differenze nelle differenze”

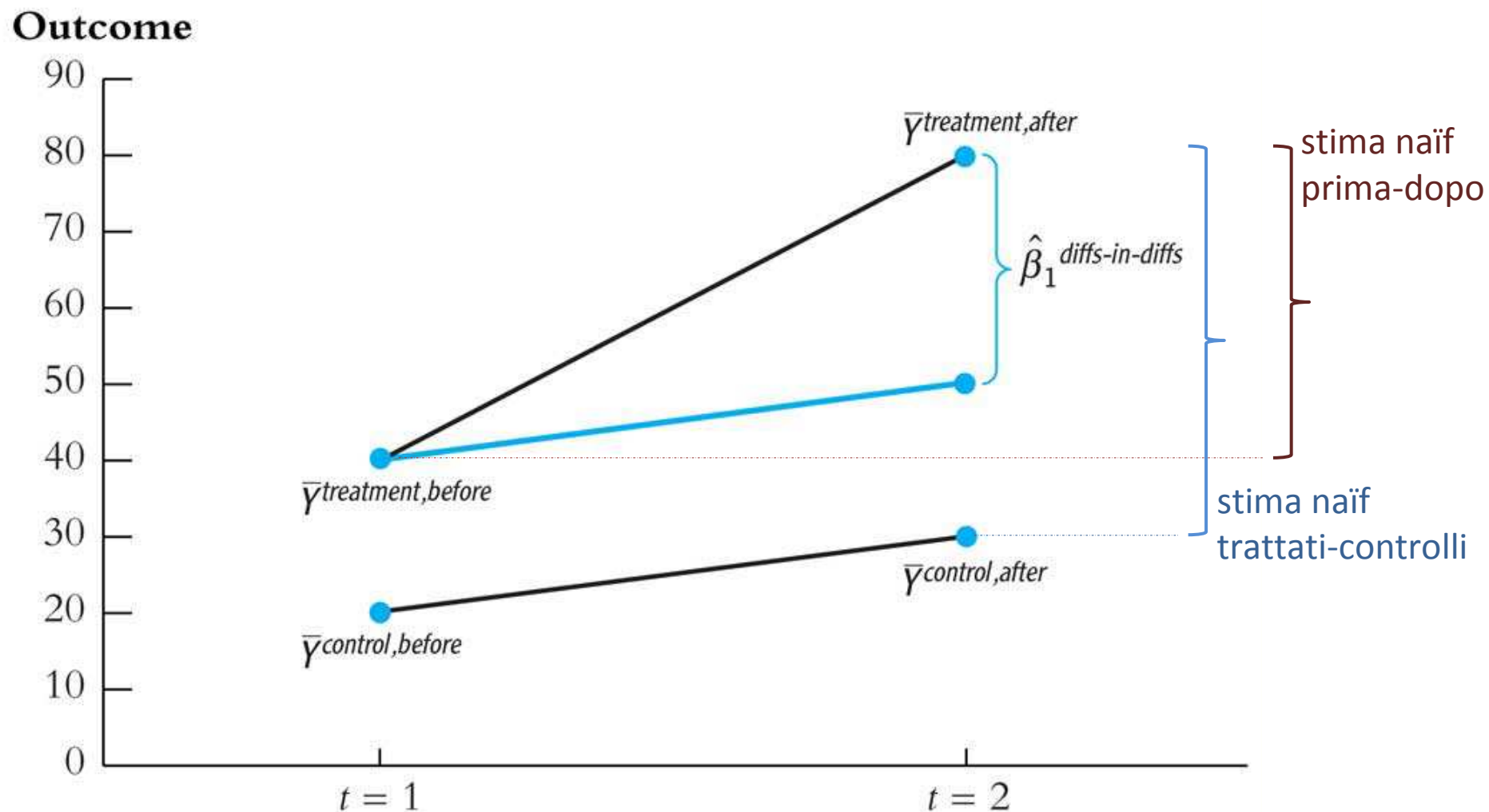
Ripartiamo dalla stima naïve “prima-dopo”:

$$(2) \quad \beta = \frac{(Y^1_{T1} - Y^0_{T1}) - (Y^0_{T1} - Y^0_{T0})}{1} \rightarrow \text{ATT}$$

$\rightarrow \text{trend naturale}$

L’idea del “diff-in-diff” è quella di correggere la formula, utilizzando, come stima per il trend naturale che avrebbe caratterizzato il gruppo dei trattati, quanto si osserva in un gruppo di controlli scelto a posteriori.

Il modello delle “differenze nelle differenze”



Il modello delle “differenze nelle differenze”

Il gruppo dei controlli non deve necessariamente essere il più possibile simile al gruppo dei trattati: è sufficiente che si possa ipotizzare che la dinamica naturale che si osserva tra loro sia rappresentativa di quanto si sarebbe osservato tra i trattati.

$$\begin{aligned}
 (3) \quad \beta = & (Y_{T1} - Y_{T0}) - (Y_{C1} - Y_{C0}) = \\
 & (Y^1_{T1} - Y^0_{T1}) + \\
 & (Y^0_{T1} - Y^0_{T0}) - \quad \rightarrow \text{trend nei } T \\
 & (Y^0_{C1} - Y^0_{C0}) \quad \rightarrow \text{trend nei } C
 \end{aligned}$$

Il modello delle “differenze nelle differenze”

Un esempio classico in ambito economico è in Card e Krueger (1994), in cui si studia se l'imposizione di un salario minimo ha l'effetto di diminuire l'occupazione.

Gli autori considerano un cambio nella normativa del New Jersey nel 1992, che portò il salario minimo da \$4.25 a \$5.05.

Per “mimare” un esperimento, si focalizzarono solo sui lavoratori della catena MacDonald, e come caso di controllo considerarono i lavoratori nello stato confinante della Pennsylvania.

Il modello delle “differenze nelle differenze”

Per “mimare” un esperimento, si focalizzarono solo sui lavoratori della catena MacDonald, e come caso di controllo considerarono i lavoratori nello stato confinante della Pennsylvania.

Occupati prima e dopo la riforma in New Jersey

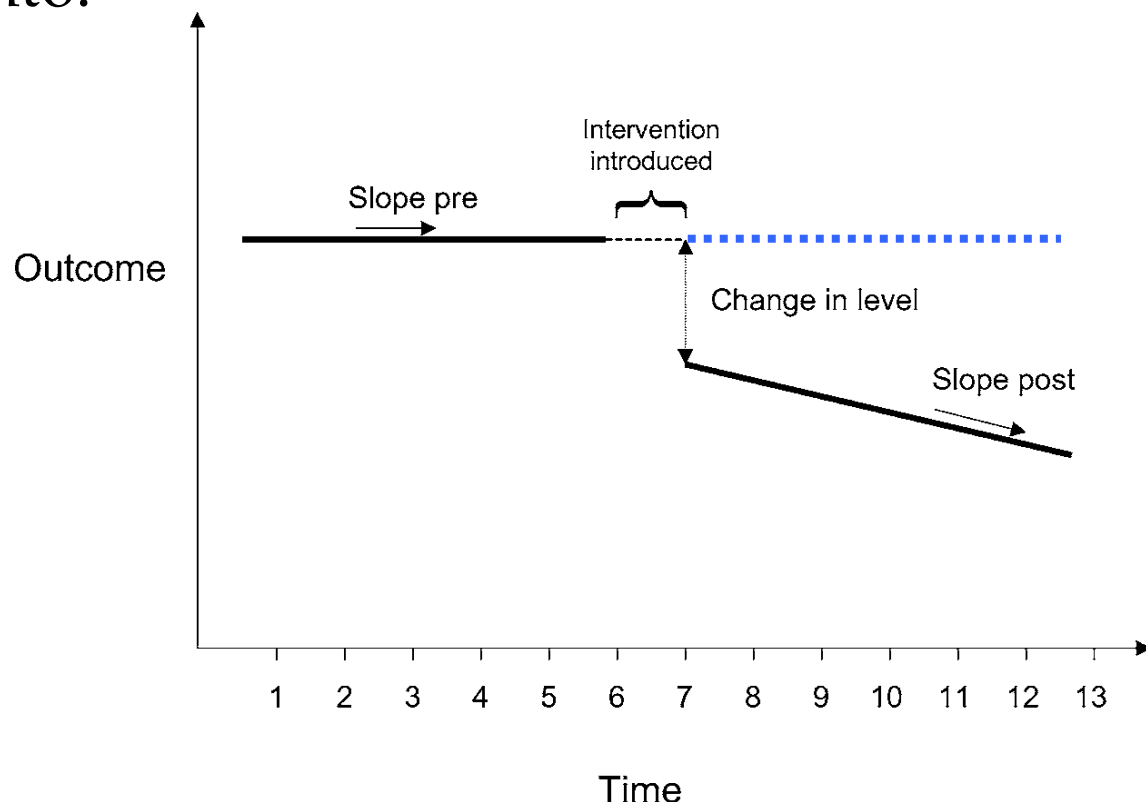
	New Jersey	Pennsylvania	differenza
T0	20,44	23,33	-2,89
T1	21,03	21,17	-0,14
differenza	0,59	-2,26	2,75



Il modello delle “differenze nelle differenze”

In questo modello consideriamo possibili cambi sia nel livello che nel trend...

...e “mimiamo” il controfattuale estrapolando la serie pre-trattamento.



Le stime basate su matching

Le stime basate su matching portano all'estremo l'idea di ricostruire a posteriori un gruppo di controllo, cercando, per ogni individuo trattato, un individuo che sia il più possibile simile a lui: una sorta di sosia.

Tecnicamente, l'ipotesi necessaria è che devono essere osservabili tutte le caratteristiche che hanno un ruolo nel *selection bias*.

Ne risultano due gruppi di trattati e controlli che non hanno differenze sistematiche tra loro, non **a priori**, grazie alla randomizzazione, ma **a posteriori**, grazie all'introduzione nel bacino dei controlli di tutti e soli gli individui il più possibile simili ai trattati.

La quasi-randomizzazione: eventi naturali

Cipollone e Rosolia, in due studi differenti (2007 e 2011), considerano la capacità di un evento naturale – il terremoto dell'Irpinia nel 1980 – di influenzare le scelte scolastiche, e per studiare quindi **l'effetto dell'istruzione** su vari outcome, tra cui la mortalità.

A seguito del terremoto venne concessa l'esenzione della leva; gli autori dimostrano come questo fattore **esogeno** avesse aumentato del 2% tra i maschi la percentuale di individui che raggiunsero il diploma superiore (e, con un *peer effect*, anche tra le donne!).

Grazie a questa variazione esogena, riescono a dimostrare che un livello di istruzione più elevato riduce la mortalità in età giovanile. Un aumento del 10% nella quota di diplomati riduce nelle loro stime dell'1% la quota di coloro che complessivamente muoiono tra i 25 e i 35 anni.

La quasi-randomizzazione: altri eventi esogeni

Consideriamo l'effetto del numero di poliziotti sul tasso di criminalità: in particolare ci aspettiamo che **riduca** la criminalità. Nei dati, tipicamente, la relazione è negativa! → se c'è tanta criminalità viene aumentata la consistenza delle forze dell'ordine.

Anche in questo caso è difficile da un punto di vista etico/politico giustificare un esperimento sociale. Ma anche in questo caso è possibile trovare variazioni **esogene** con le quali è possibile ricostruire un contesto sperimentale.

La variazione sfruttata in un classico lavoro di Levitt (1997) è quella tra ciclo elettorale e numero di poliziotti.

→ Il numero di poliziotti nella città di New York non varia in anni "normali", aumenta del 2.1% negli anni di elezione del sindaco, del 2.0% negli anni di elezioni politiche!

→ Riesce ad identificare un effetto positivo sia per i crimini violenti che contro la proprietà.

La quasi-randomizzazione: soglie nelle politiche

Un modello che viene considerato il più vicino agli esperimenti controllati è quello del *regression discontinuity design*, applicato per la prima volta da Thistlewaite e Campbell (1960), per valutare l'efficacia delle borse di studio per gli studenti meritevoli.

In questo caso la selezione nel trattamento è fortissima: Se a prendere le borse di studio sono gli studenti più bravi, è naturale che su di loro si misurino gli outcome migliori!

Di nuovo, randomizzare l'assegnazione delle borse di studio non è molto praticabile da un punto di vista etico/politico. Però si può lavorare sulla soglia che definisce l'eleggibilità alla borsa, confrontando gli individui “appena non eleggibili” con gli individui “appena eleggibili”.

La quasi-randomizzazione: soglie nelle politiche

Le politiche sono piene di soglie!! Oltre alle borse di studio, sono molti gli interventi che lavorano su graduatorie. Ma, se dovunque c'è una graduatoria c'è sicuramente anche un *bias* da selezione, allo stesso tempo c'è quasi sempre anche una soglia che separa individui simili, tra i quali solo quelli appena sopra la soglia ricevono l'intervento.

Le soglie, anziché sull'ammissibilità alla politica, ne variano l'intensità:

- durata dell'indennità di disoccupazione (età e area)
- durata della cassa integrazione (età)
- norme sui contratti di lavoro (dimensione d'impresa)

SUMMING UP!!

Il primo e in molti casi fattibile (basta la buona volontà!) modello per la valutazione delle politiche è il “caro vecchio” *randomized trial*. Anche se, ovviamente, in ambito sociale sono maggiori le difficoltà di implementazione del protocollo sperimentale.

Anche in assenza di randomizzazione, sono a disposizione però diversi modelli di studio che ne mimano i principi: diff-in-diff, matching estimators, ITS. In molti casi, è possibile però rintracciare o in variazioni esogene, o in discontinuità nella normativa, una fonte almeno parziale di randomizzazione, che garantisce una solidità agli studi paragonabile a quella dei RT.