

Lezione 5

Valutazione delle Politiche Pubbliche (SA)

Il metodo sperimentale

L'idea fondamentale alla base del metodo sperimentale è la stessa della *sperimentazione clinica utilizzata per testare l'efficacia dei farmaci*.

Due gruppi divisi mediante **sorteggio**:

- gruppo sperimentale, che subirà il trattamento
- gruppo di controllo, non subirà il trattamento

Funzione del gruppo di controllo  *riprodurre il controfattuale*.

Qui, nessuno dei trattati/pazienti sceglie a quale gruppo appartenere.

La **RANDOMIZZAZIONE**, se applicata in modo corretto, elimina le differenze di partenza tra i due gruppi (DISTORSIONE DA SELEZIONE).

Equivalenza statistica

- I due gruppi (trattamento-controllo) non saranno identici. Ma devono essere statisticamente equivalenti... hanno la stessa distribuzione di tutte le caratteristiche (la stessa media, la stessa varianza, gli stessi quantili), eccezion fatta per l'errore casuale (piccolo, se il campione è sufficientemente "grande").
- L'equivalenza statistica implica l'eliminazione delle differenze di partenza tra i due gruppi, di conseguenza elimina la *distorsione da selezione*, anzi previene che si formi, in quanto nessuno dei "pazienti" sceglie a quale gruppo appartenere.
- In questo caso (e solo in questo caso) **l'effetto di una politica è identificato da una semplice differenza: quella tra la media (della variabile risultato) di trattati e non-trattati** (o gruppo sperimentale e gruppo di controllo).

Differenze di partenza

- Se applicata ad un campione sufficientemente grande, la randomizzazione produce due gruppi statisticamente equivalenti sia nelle caratteristiche osservabili che nelle caratteristiche non osservabili.
- Motivazione nella ricerca di lavoro. **Non osservabile.**
- Rapporti sociali. **Non osservabile.**
- Istruzione, età sesso... **Osservabile.**

Un esempio di randomizzazione (involontaria)

- Politica DRIL: una parte delle richieste escluse perché erroneamente considerate in ritardo. Tre gruppi (non più 2):

Tabella 9.1 Composizione e caratteristiche dei beneficiari e degli esclusi dalle borse-lavoro

	Domande di borsa-lavoro presentate nel 1999	Numero di interviste effettuate	Tasso di risposta	Caratteristiche socio-demografiche al momento della richiesta			
				Età media	% di donne	% con diploma	Punteggio al test
Beneficiari	728	627	86%	32,4	22,8%	27,9%	87,2
Esclusi per errore	656	535	82%	31,7	23,3%	27,4%	86,3
Esclusi per ritardo	770	563	73%	28,4	18,0%	21,1%	71,2
Totale esclusi	1.426	1.098	77%	30,0	20,6%	24,2%	78,1

Stima plausibile dell'effetto...

Tabella 9.2 Stima dell'effetto delle borse-lavoro e quantificazione della distorsione da selezione

	Ruolo rispetto al metodo sperimentale	Numero di interviste effettuate	Reddito annuale nel 2001	% di occupati stabili a fine 2001
Beneficiari della borsa	(gruppo sperimentale)	627	€9.835	32,4%
Esclusi per errore	(gruppo di controllo)	535	€9.343	27,4%
		<i>Stima dell'effetto</i>	€492	5,0
		<i>Effetto in % sul controfattuale</i>	5,3%	18,3%

In questo caso **l'effetto di una politica è identificato da una semplice differenza: *quella tra la media della variabile risultato di trattati e non-trattati*** (o gruppo sperimentale e gruppo di controllo).

...e della distorsione da autoselezione.

Tabella 9.2 Stima dell'effetto delle borse-lavoro e quantificazione della distorsione da selezione

	Ruolo rispetto al metodo sperimentale	Numero di interviste effettuate	Reddito annuale nel 2001	% di occupati stabili a fine 2001
Beneficiari della borsa	(gruppo sperimentale)	627	€9.835	32,4%
Esclusi per errore	(gruppo di controllo)	535	€9.343	27,4%
		<i>Stima dell'effetto</i>	€492	5,0
		<i>Effetto in % sul controfattuale</i>	5,3%	18,3%
Esclusi per ritardo	-	563	€5.755	20,0%
Totale degli esclusi	-	1.098	€7.503	23,6%
		Differenza "lorda"	€2.332	8,8%
Decomposizione della differenza tra beneficiari e totale degli esclusi		Effetto vero	€492	5,0%
		Distorsione da selezione	€1.840	3,8%

Il metodo sperimentale

*“As is the case with epidemiologists, the fundamental challenge faced by economists — and a root cause of many disagreements in the field — is our **limited ability to run experiments**.*

If we could randomize policy decisions and then observe what happens to the economy and people’s lives, we would be able to get a precise understanding of how the economy works and how to improve policy.

*But the **practical and ethical costs** of such experiments preclude this sort of approach. (Surely we don’t want to create more financial crises just to understand how they work.)”*

RAJ CHETTY, New York Times, 20/10/2013

Il metodo sperimentale: limiti

*L'esperimento con randomizzazione è il più solido disegno per valutare l'impatto degli interventi.
Ma...problemi di implementazione!*

Politiche a regime e interventi dimostrativi

1. Interventi dimostrativi: interventi che sono concepiti e attuati con lo scopo specifico di valutarne la fattibilità e l'efficacia.
 2. Politiche a regime: programmi applicati e con routine (anche amministrative...) consolidate. Qui è spesso difficile ottenere la cooperazione dello staff e degli operatori dei servizi. Vedono minacciato il loro ruolo dal fatto che si decida in base al caso chi può essere ammesso e chi no? Difficile modificare le routine già instaurate?
- (1) è più facilmente gestibile e attuabile. Ma le valutazioni di interventi dimostrativi hanno uno svantaggio importante: quello di riferirsi ad una situazione diversa dalla realtà operativa di un intervento a regime. *Problema della generalizzabilità delle stime di impatto.*

Politiche a regime: la politica JTPA

Job Training Partnership Act (JTPA). Politica USA di servizi all'impiego e formazione professionale realizzato negli anni '80 per aiutare le persone economicamente svantaggiate.

La sperimentazione (1987-1992) viene fatta quando l'intervento è a regime da alcuni anni, coinvolge 16 centri di supporto all'impiego, rappresentativi dei circa **600 esistenti a livello nazionale**.

Circa 20.000 dei potenziali utenti sottoposti a randomizzazione e 6.000 formano il gruppo di controllo (i.e. niente servizi per l'impiego per un periodo di 18 mesi).

Stime: effetto positivo della fruizione dei servizi sul reddito da lavoro degli adulti ma effetto nullo per i giovani (in particolare per i maschi).

Disegni sperimentali complessi

- Il Trattamento A è più efficace del Trattamento B nel modificare una certa situazione? Randomizzazione a tre vie: un terzo circa dei soggetti aventi diritto è assegnato al gruppo di controllo, un terzo al Trattamento A e un terzo al Trattamento B.
- Capire l'effetto di ciascuna componente e l'effetto della compresenza, o interazione in caso di trattamenti multipli. Esempio: i disoccupati presi in carico ricevono sia 1) l'assistenza personalizzata prestata da un navigator sia 2) la concessione di un incentivo monetario. Qui dovrei creare tre gruppi sperimentali + il gruppo di controllo (con randomizzazione):
 - solo il navigator, solo l'incentivo monetario, e la combinazione dei due per il terzo gruppo. Implementare ciascuna soluzione in una località diversa? Diventa difficile separare l'effetto del contesto dall'effetto del particolare trattamento (a meno di non avere un numero grandissimo di località a disposizione). In questi casi si rinuncia spesso ad ottenere una stima separata degli effetti delle due componenti (e si usano analisi con metodi qualitativi per approfondimenti).

Ostacoli tecnici

- L'applicabilità del metodo sperimentale è limitata alle *politiche non universali, cioè a copertura parziale*.
- *Tutti* quegli interventi in cui è impossibile definire una fruizione individuale (l'istruzione elementare, regole sulla circolazione stradale...), non si prestano all'individuazione di un gruppo di controllo e quindi alla randomizzazione.
- Questo vincolo può in alcuni casi essere attenuato nel caso vengano randomizzati gruppi o aree invece che individui.

Ostacoli di tipo etico

- E' eticamente accettabile escludere alcuni individui dalla fruizione di un servizio, da cui potrebbero (forse) trarre qualche beneficio?


Ostacoli di tipo etico

- E' un sacrificio individuale accettabile, alla luce del fatto che l'utilità stessa del servizio è in dubbio?

Se non esistesse questo dubbio sull'efficacia del servizio, non ci sarebbe ragione di proporre la valutazione degli effetti.

Ma questo problema è enorme nel caso dei clinical trials: con le dovute cautele, in questo ambito il dilemma è stato però ampiamente risolto in favore della desiderabilità della randomizzazione.

Avversione alla randomizzazione

- Avversione alla randomizzazione da parte degli operatori dei servizi coinvolti.
- Infatti, si chiede loro di escludere una quota dei loro utenti sulla base di un sorteggio: i malumori e lamentele degli esclusi ricadono *in primis sugli operatori stessi*  costo immediato notevole (almeno in termini di stress psicologico).
- Nota che la valutazione viene effettuata quando sussiste il *dubbio circa l'efficacia del servizio, dubbio che gli operatori raramente condividono in quanto spesso identificano la propria professionalità con il successo stesso del servizio.*

Ostacoli di tipo legale e politico

- Vincoli di tipo legale sorgono quando l'erogazione del servizio (o del sussidio o dell'agevolazione) è già prevista dalla legislazione vigente. Negli USA hanno emanato apposite esenzioni legislative (*waiver*) che consentono di differenziare l'erogazione dei servizi se ciò è fatto a scopo di valutazione.
- Problema "politico": è però politicamente difficile difendere l'esclusione di utenti dalla fruizione di servizi. L'incertezza sulla loro utilità ed efficacia, se da un lato motiva la valutazione, dall'altro non serve a placare il cittadino escluso dal servizio.

Il trade-off tra validità esterna e interna delle stime

- Se ottenute per un campione ridotto e con dati raccolti in un intervallo di tempo ridotto le stime ottenute con il metodo sperimentale sono corrette. Ma sono generalizzabili?

1. validità interna delle stime: esse riflettano realmente il contributo netto dell'intervento alla modifica della situazione che si sarebbe verificata in assenza di intervento.

2. validità esterna delle stime: se 1) periodo limitato di tempo e 2) scala ridotta (come avviene negli esperimenti), spesso si ottengono risultati difficili da generalizzare alla politica pubblica una volta andata a regime e implementata su scala vasta.