

La valutazione delle politiche pubbliche – valutare ex post con la valutazione d’impatto

Dal saggio di Trivellato U. (2009), La valutazione degli effetti di politiche pubbliche: paradigma e pratiche, IRVAPP Working Paper Series

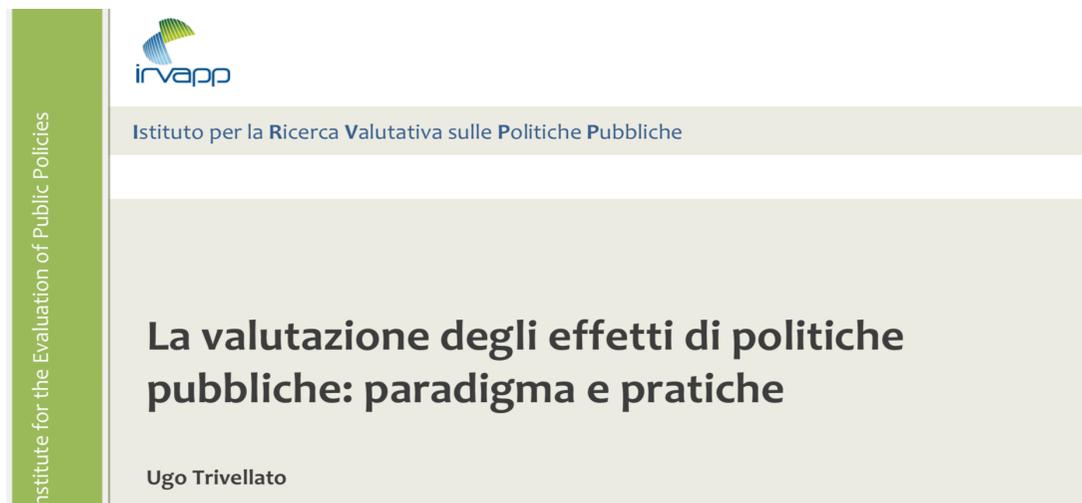
Corso Analisi e Valutazione delle Politiche APP, didattica a distanza aa 2019/2020,
III modulo Prof. Maria Tullia Galanti

Premessa

- Nelle slide di introduzione al tema della valutazione, abbiamo fornito una definizione di ricerca valutativa, ne abbiamo sottolineato gli scopi ultimi (learning e accountability), e abbiamo sottolineato come esistano diversi approcci e teorie alla valutazione
- Sempre nell'introduzione, abbiamo ripassato i concetti che costituiscono gli elementi costitutivi delle analisi valutativa: i diversi tipi di effetto, le diversi criteri (le 4 «e»)
- In queste slides, utilizzeremo lo working paper di Trivellato (2009) per illustrare gli elementi fondamentali di una delle possibili tecniche di analisi per svolgere una valutazione ex post degli effetti: la valutazione di impatto
- Il nostro obiettivo NON è padroneggiare la tecnica della valutazione di impatto per poi applicarla ad esempi empirici concreti (altri corsi trattano di questi aspetti metodologici)
- Piuttosto, il nostro obiettivo è comprendere in che modo la valutazione di impatto può essere usata come strumento per individuare l'effetto di un intervento di policy al netto di altre variabili che possono intervenire sul problema:
 - Conseguenza 1: ripercorreremo il testo di Trivellato per comprendere la logica del metodo controfattuale
 - Conseguenza 2: illustreremo in maniera generale alcune strategie per la soluzione dei problemi che si incontrano quando si vuole ricostruire un rapporto di causa-effetto tra l'erogazione di una politica e le conseguenze sul problema iniziale
 - Conseguenza 3: osserveremo come l'analisi di impatto possa essere utilizzata per studiare le politiche del lavoro in diversi paesi europei
 - Conseguenza 4: concluderemo con una riflessione sulla diffusione di questo tipo di analisi in Italia

Valutazione ex post e valutazione di impatto

- Ugo Trivellato (2009), **La valutazione degli effetti di politiche pubbliche: paradigma e pratiche**, Working Paper dell'Istituto per la Ricerca Valutativa sulle Politiche Pubbliche IRVAPP, [scaricabile qui](#)
 - Questo approccio è al centro dei lavori sviluppati dall'Ufficio Valutazione Impatto - UVI del Senato, che è orientato a produrre conoscenza valutativa sia per i policy-makers che per i cittadini: <https://www.senato.it/4779>



Cosa devo sapere? I contenuti fondamentali

- Definire una policy come programma
- Comprendere la logica dell'analisi controfattuale
- Comprendere i possibili problemi nell'individuare gli effetti del solo intervento di policy
- Comprendere le distorsioni dovute alla selezione (selection bias):
 - Distorsione da selezione
 - Distorsione da autoselezione
- Capire come gestire la distorsione da selezione tramite la creazione di gruppi di controllo
- Descrivere le linee generali delle principali strategie per risolvere il problema dell'inferenza causale ed individuare gli effetti medi
- Discutere le buone pratiche per la valutazione di impatto nei vari paesi

Definire una policy come un
programma di intervento (e
ridefinire gli effetti)

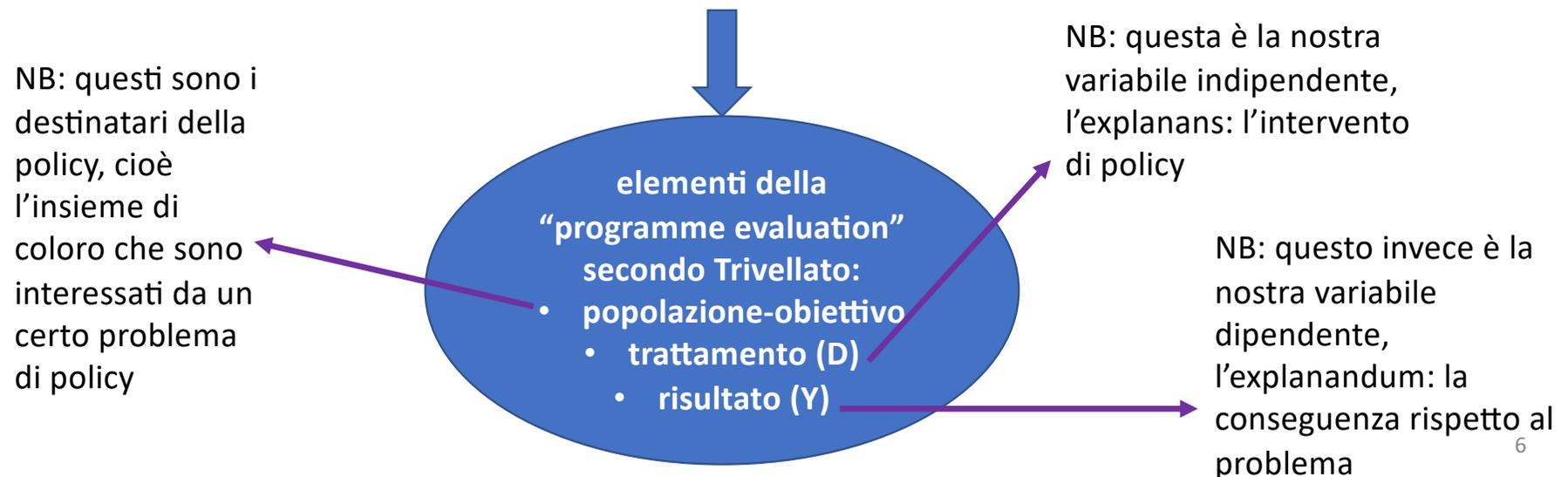
Introduzione e definizioni: la policy come programma

- Trivellato intende "policy" in senso secondo lui ampio – in realtà per noi in senso ristretto...

"un intervento

mirato ad una popolazione determinata

con l'intento di indurre un cambiamento in una condizione e/o in un comportamento"



Domanda di ricerca e oggetto di indagine

- Rispetto ad un programma di policy, Trivellato cerca di rispondere a questa domanda generale: Il programma è stato efficace? **Ha risolto il problema per il quale è stato pensato? Lo ha aggravato?**
- La difficoltà è stabilire con evidenze scientifiche e metodi rigorosi **una attribuzione causale di effetti ad una politica**, separando i cambiamenti derivati dalla politica da quelli che si sarebbero comunque verificati **anche senza la politica**.
- Il paper spiega **la logica della valutazione degli effetti delle politiche c.d. di impatto**, seguendo un disegno ispirato **alle teorie sperimentali e/o quasi sperimentali**
 - Ricordiamo dall'introduzione che queste teorie sperimentali sono solo uno dei molti approcci teorici alla ricerca valutativa (es. valutazione realista, etc.). Questo approccio alla valutazione è prevalente tra gli economisti.

La logica dell'analisi controfattuale

Cosa è l'effetto della policy? Perché non basta guardare a «Prima e dopo» l'intervento?

- è difficile rispondere a domande quali: il miglioramento osservato è merito della politica o si sarebbe comunque verificato? Il peggioramento osservato sarebbe stato più grave in assenza della politica?
 - Perché alcuni processi spontanei continuano a condizionare destinatari e non
 - Perché i destinatari (e non) del programma reagiscono alla policy stessa
 - Perché i programmi sono realizzati in contesti diversi e con modalità specifiche, cui corrispondono esiti diversi
- Per questo, è *ingenuo e fallace* definire l'effetto del programma come “la differenza tra **la situazione osservata *dopo*** l'esposizione ad un intervento e **la situazione osservata *prima*** dell'esposizione”, perché quella differenza dipende **da cause ulteriori rispetto al solo programma di policy**
- Se l'effetto non può esser definito come confronto tra “prima” e “dopo”, la soluzione è definirlo in maniera diversa (vedi slide successiva)

L'analisi controfattuale come soluzione

- Per isolare l'effetto dell'intervento, la soluzione è adottare una **definizione di effetto secondo la logica controfattuale**:
- “l'effetto dell'intervento è la differenza tra quanto si osserva in presenza dell'intervento e **quanto si sarebbe osservato in sua assenza**”
 - Il primo termine del confronto è **un risultato osservabile** (cioè è il risultato che posso rilevare empiricamente in un certo tempo e spazio, es. l'andamento della disoccupazione giovanile tra il 2012 e il 2019 nei giovani **che hanno beneficiato** di politiche di accompagnamento al lavoro)
 - il secondo termine, invece, è solo un **risultato ipotetico** (cioè il risultato che posso stimare, es. l'andamento della disoccupazione giovanile nei giovani con caratteristiche simili, ma che **NON hanno beneficiato** di quelle politiche) ed è non osservabile per definizione, perché non si può osservare contemporaneamente l'intervento e il non intervento sullo stesso individuo
- Il risultato ipotetico, non osservabile per definizione, è definito come **situazione o risultato controfattuale**

Il problema dell'inferenza causale

- Per isolare l'effetto della politica, la logica del controfattuale propone dunque di **confrontare un risultato osservabile (le conseguenze sul problema dell'intervento di policy) con un risultato ipotetico (le conseguenze sul problema senza l'intervento)**
- Tali diversi risultati NON possono esser rilevati basandoci sull'osservazione di uno stesso individuo: non posso avere una stessa persona che allo stesso tempo è esposta e non esposta al trattamento di policy
 - quindi *l'effetto di un intervento a livello individuale resta non osservabile per definizione*
- Questa impossibilità è definita in letteratura come “il problema fondamentale della valutazione degli effetti” o, più in generale, “il problema fondamentale dell'inferenza causale”

Come risolvere il problema dell'inferenza causale? Rilevare l'effetto medio

- Una possibile soluzione al “problema fondamentale dell'inferenza” è «sfruttare» l'esistenza di una pluralità di soggetti, alcuni dei quali sono esposti alla policy, e alcuni no.
- Sfruttando tale pluralità, la soluzione diventa considerare non gli effetti causali totali ma **gli effetti medi, osservando gli effetti nella popolazione o in specifici sottogruppi**
 - Per confrontare gli effetti in specifici sottogruppi, si considerano dunque gruppi di individui/soggetti esposti al trattamento, e gruppi di individui/soggetti NON esposti al trattamento
- Ma... come è possibile costruire i due gruppi da confrontare? (vedi slide successiva)

Come rilevare l'effetto medio? Ricostruire il controfattuale

- La logica del paradigma controfattuale implica per l'appunto **la costruzione di questi due gruppi di soggetti (trattati e non trattati) e la loro comparazione** per consentire di individuare l'effetto medio come la differenza tra gli effetti nei due gruppi
- Il paradigma controfattuale dunque è la logica che accomuna i metodi analitici che, per attribuire causalità ad un intervento, ***ricostruiscono credibilmente cosa si sarebbe osservato sul tipo di soggetti esposti all'intervento in assenza dell'intervento stesso.***
 - Origine del termine e simili definizioni: logica del What If della Cowles Commission, anni 40; il modello causale di Rubin

Illustrare la logica controfattuale con le politiche del lavoro: due esempi

- Trivellato utilizza nel WP due diversi interventi di policy per mostrare la logica della valutazione degli effetti con il paradigma controfattuale:
 - Una politica di offerta di servizi di orientamento al lavoro (es. Garanzia Giovani)
 - Una politica di offerta di incentivo monetario alle imprese per stabilizzazione lavoratori (es. incentivi per contratto a tutele crescenti nel Jobs act)
- Di seguito, commentiamo gli elementi concettuali usati da Trivellato per definire, per ogni esempio di intervento (vedi tabella 1, pag. 9):
 - quali sono i destinatari (c.d. popolazione obiettivo)
 - quali sono gli interventi messi in campo (c.d. variabili trattamento, vx)
 - quali sono gli effetti attesi (c.d. variabile risultato, vy)

Come leggere la tabella 1: l'esempio dei servizi di orientamento alla ricerca del lavoro

Esempio 1: intervento di politica tramite servizi di orientamento al lavoro utilizzando **gli iscritti alle liste di disoccupazione come destinatari (A)**

questo è il tipo di intervento di policy messo in campo, la variabile indipendente

questi sono i due indicatori che vengono utilizzati per rilevare la variabile dipendente, cioè le conseguenze dell'intervento di policy

Due esempi stilizzati di politica pubblica	Effetto		popolazione obiettivo	Identificazione degli effetti: strategie di analisi e dati
	Di cosa:	Su cosa:		
	variabile trattamento	Variabile risultato		
Servizi di orientamento nella ricerca di lavoro	Servizio ricevuto	Durata della disoccupazione (dopo il trattamento)	A. Iscritti a liste di disoccupazione	<p>A. Alto rischio di <i>autoselezione</i> dei disoccupati nell'accedere al servizio e di <i>selezione</i> da parte del personale dei Centri per l'impiego nell'assegnazione dei migliori ai servizi (migliori).</p> <p>Necessità di <i>ricche basi di dati individuali</i> sugli iscritti alle liste di disoccupazione (caratteristiche socio-demografiche, scolarità/formazione, storia lavorativa precedente, ecc.), nonché <i>sugli operatori</i>, per eliminare la distorsione da selezione.</p> <p>Caso particolare di <i>confronto attorno alla soglia</i>: il servizio è razioneato, e l'ammissibilità allo stesso è ristretta a disoccupati al di sopra di una data soglia di una variabile, fissata in maniera ben definita.</p>
		Condizione occupazionale dopo t mesi		

questo è il gruppo dei destinatari dell'intervento

questi i problemi nell'individuare gli effetti, dovuti alle caratteristiche dei destinatari

queste sono le due strategie per risolvere i problemi

Esempio 1: intervento di politica tramite servizi di orientamento al lavoro utilizzando un esperimento che mette a confronto due gruppi di iscritti alle liste (B)

questo è il tipo di intervento di policy messo in campo, la variabile indipendente

questi sono i due indicatori che vengono utilizzati per rilevare la variabile dipendente, cioè le conseguenze dell'intervento di policy

Due esempi stilizzati di politica pubblica	Effetto		popolazione obiettivo	Identificazione degli effetti: strategie di analisi e dati
	Di cosa:	Su cosa:		
	variabile trattamento	Variabile risultato		
Servizi di orientamento nella ricerca di lavoro	Servizio ricevuto	Durata della disoccupazione (dopo il trattamento)	B. Esperimento (su un campione casuale di iscritti alle liste)	B. Differenza fra medie di trattati e controlli.
		Condizione occupazionale dopo t mesi		

qui si identificano due gruppi, uno che riceve il trattamento, l'altro che non lo riceve

questa è la strategia per risolvere il problema dell'inferenza: eseguire un esperimento e calcolare la differenza tra medie

Tabella riadattata da Trivellato 2009, 9

Esempio 2: incentivi alle imprese per stabilizzare i precari Esercitemoci...

qual è il tipo di intervento ???

quali gli indicatori per rilevare le conseguenze attese??

Incentivi alle imprese per la stabilizzazione dei precari	Ammissibilità all'incentivo	Trasformazioni in contratti a tempo indeterminato	Tutte le imprese da tempo t a $t+s$ + restrizione territoriale; + restrizione lavoratori precari età $\geq x$ anni	Confronto tra periodi con/senza incentivi + confronto tra aree territoriali con/senza incentivi + confronto attorno alla soglia di età. <i>Serie storiche di trasformazioni in contratti a tempo indeterminato, totale e con incentivi.</i>
		[Assunzioni immediate a tempo indeterminato]		Possibile <i>entry effect</i> complica l'analisi. - Utilizzo dei metodi segnalati – confronto tra periodi con/senza incentivi, confronto tra aree con/senza incentivi, confronto attorno alla soglia di età – <i>per la prima volta</i> in cui la politica viene introdotta. - Negli altri casi, servono notevoli affinamenti nei metodi e maggiori informazioni.

quali i destinatari dell'intervento???

Quali le problematiche che?

quali le strategie di soluzione??

Tabella riadattata da Trivellato 2009, 9

I problemi nell'individuare gli effetti della policy: ammissibilità e assegnazione del trattamento

Quali i problemi per individuare l'effetto della politica? esempio 1

- Es. Servizi di orientamento al lavoro
 - Un mix di misure attive (colloquio, matching domanda e offerta lavoro) e passive (sostegno al reddito, obbligo di accettare offerta congrua)
 - Rischi di distorsione: solo i disoccupati più capaci accedono ai servizi; i funzionari assegnano i servizi ai disoccupati migliori
- Da qui, emergono alcuni problemi che derivano sia dal disegno che dalla successiva implementazione della policy:
- Per il disegno: Importanza dei **criteri di ammissibilità** per definire popolazione-obiettivo -> **possono variare nel tempo**
- Per l'implementazione: Importanza dei meccanismi di **effettiva assegnazione**, stabiliti dalle pratiche dei funzionari -> **possono determinare di fatto chi sono i trattati e chi i non-trattati**

Quali i problemi per individuare l'effetto della politica? esempio 2

- Es. Incentivi a imprese per stabilizzazione dei precari
 - Incentivo come misura tesa a favorire comportamenti desiderabili, determina variazione negli esiti a partire da un certo momento nel tempo
 - Rischi di distorsione: gli attori economici possono non aderire all'incentivo; laddove aderiscano, è comunque difficile capire l'effetto, poiché il comportamento desiderato – l'assunzione – potrebbe essere stato determinato anche da altre variabili
- Anche qui contano i criteri di ammissibilità – per definire popolazione-obiettivo -> possono variare nel tempo
- **Per capire l'effetto reale degli incentivi bisogna chiedersi: quanti contratti precari sarebbero stati trasformati in contratti a tempo indeterminato in assenza di incentivo?**

Tipi di distorsione dovuta a
selezione o *selection biases*

Selection biases: due processi di distorsione dovuta a selezione

- Come evidente dalla precedente discussione sull'ammissibilità e sull'assegnazione del trattamento, nelle politiche pubbliche è frequente il caso in cui il gruppo degli ammissibili (cioè coloro che per legge possono ricevere il servizio o l'incentivo) NON coincide con il gruppo dei trattati (cioè coloro che effettivamente ricevono il servizio o l'incentivo): questo produce **due tipi di distorsione o selection bias**:
 - **Autoselezione**: - Caso 1: nelle loro caratteristiche generali, I soggetti trattati sono **sistematicamente diversi** dai soggetti trattati per ragioni che NON dipendono dall'intervento di policy
 - **Selezione** - Caso 2: se esiste qualche discrezionalità nell'assegnazione del trattamento agli ammissibili, i funzionari possono scegliere di erogare il trattamento sistematicamente ad alcuni soggetti e non ad altri

Selection bias: perché è problematica

- L'autoselezione da parte dei destinatari o la selezione da parte degli operatori configura una **distorsione da selezione o *selection bias***
- Infatti, entrambi i processi, di selezione e di autoselezione, producono **una differenza nelle condizioni di partenza della popolazione target** (indotta dai comportamenti degli ammissibili e dalle scelte degli erogatori)
- Tale differenza nelle condizioni di partenza IMPEDISCE di utilizzare come stima dell'effetto medio del trattamento la differenza tra trattati e non trattati

Un rimedio per il selection bias: assegnazione casuale e metodo sperimentale

- Eccezione: quando il processo di selezione avviene tramite assegnazione casuale dei soggetti al gruppo trattati o al gruppo non trattati (es. *pilot studies* su scala ridotta)
- In medicina: Randomized Clinical Trials
- **Ma** l'impiego del metodo sperimentale ha implicazioni sia comportamentali che etiche (il caso vs il bisogno), e soprattutto non è sempre disponibile (es. mancano i soldi per svolgere lo studio pilota) o desiderabile (es. come decidere a chi tocca il servizio e a chi no?)
- Quindi, per poter identificare gli effetti **in mancanza del metodo sperimentale**, è necessario di **riuscire a rendere nulla la distorsione da selezione attraverso una serie di strategie per ricostruire la situazione controfattuale (vedi slide seguente).**

Ricostruire la situazione controfattuale attraverso un modello

- Per gestire il problema della distorsione da selezione, **si crea un modello** che rappresenta i processi di assegnazione al trattamento
- Nel saggio di Trivellato, questo modello viene spiegato ricorrendo al cosiddetto diagramma di Pearl

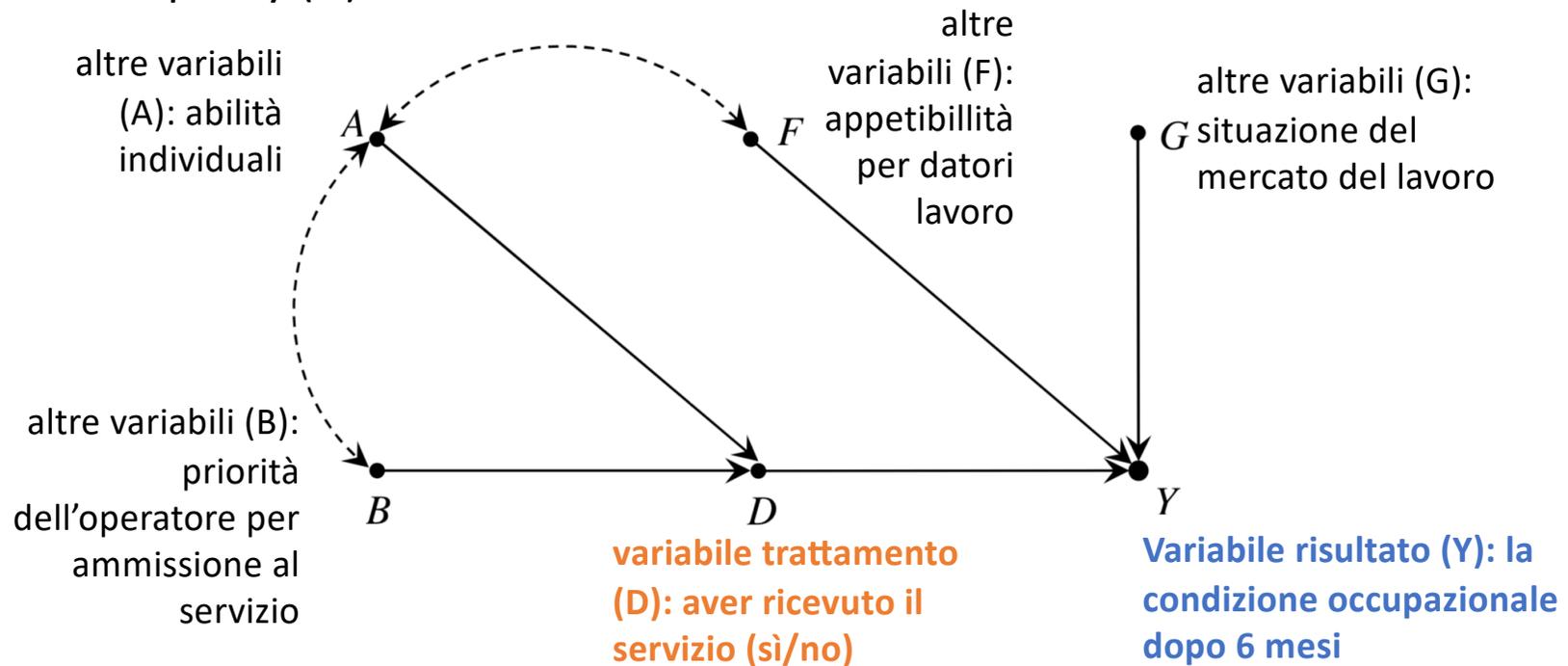
Un modello per gestire le distorsioni e ricostruire la situazione controfattuale

Gestire il selection bias: il diagramma di Pearl

- Il diagramma di Pearl è una rappresentazione causale che mette in relazione i processi di assegnazione al trattamento (es. ricevere un servizio di accompagnamento al lavoro) e la determinazione della variabile-risultato (es. valutare l'andamento della occupazione)
 - **Ogni nodo** del grafico è una variabile osservabile;
 - **le frecce** indicano la direzione della relazione causale;
 - **gli archi tratteggiati** indicano variabili comuni non osservate che causano le variabili terminali

Gestire il selection bias: il diagramma di Pearl

- Quali variabili dobbiamo osservare e come dobbiamo usarle per stimare l'effetto causale di D (la politica) su Y (il livello di occupazione)? Per rappresentarlo si usa il diagramma causale di un modello di assegnazione al trattamento (D) e di determinazione della variabile-risultato (Y)
- In nero, sono riportate tutte le variabili del modello che possono influenzare il risultato occupazionale (Y), e che rendono difficile dunque individuare qual è l'effetto netto dell'intervento di policy (D)



I problemi di stima nel diagramma di Pearl

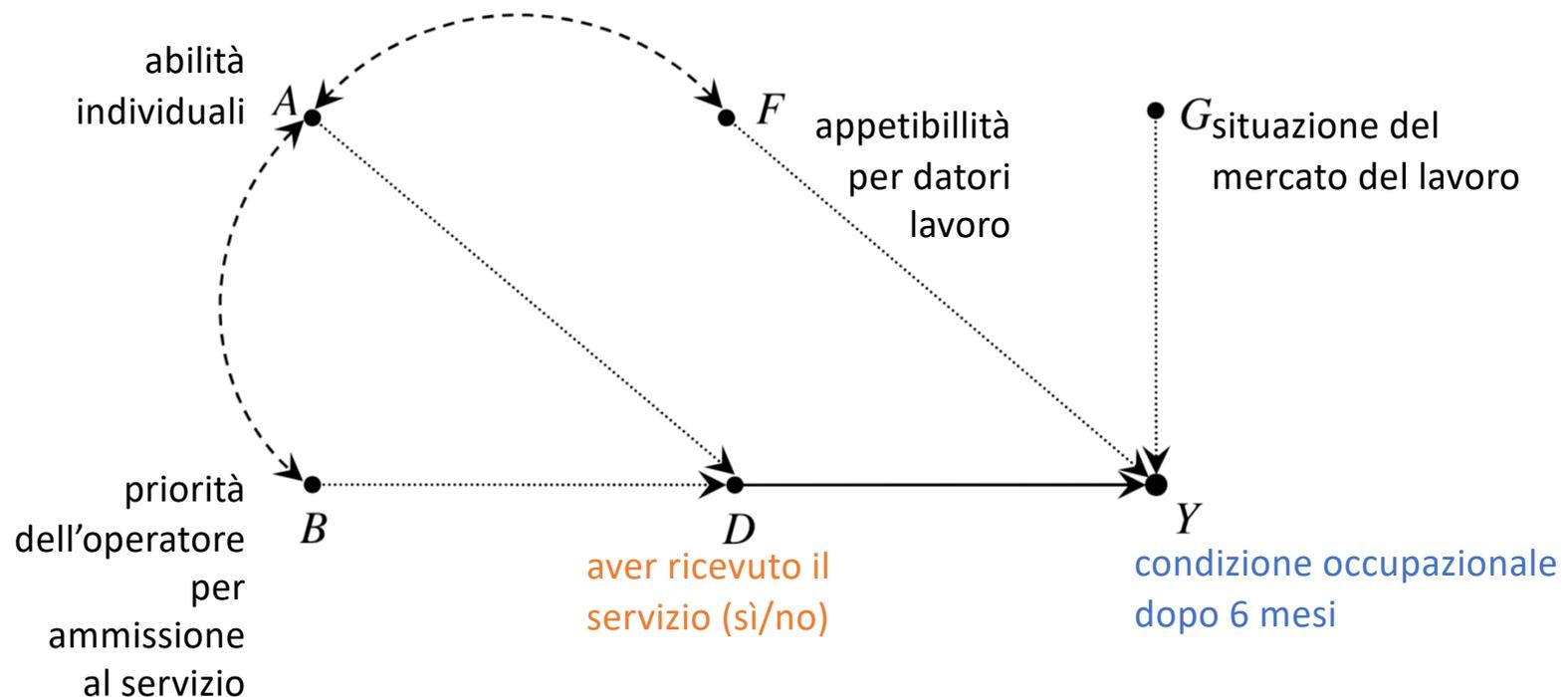
- Il diagramma di Pearl fornisce una utile rappresentazione del modello che riassume tutte le variabili che possono avere un effetto sulla nostra variabile risultato, i livelli occupazionali
 - attraverso il modello è possibile infatti individuare tutte le variabili che influenzano il risultato, considerando sia l'intervento di policy che le altre variabili
- Ora, il problema principale è che spesso le altre variabili (quelle esogene rispetto alla policy) **NON sono facili da stimare, per vari motivi**
- Ecco perché è necessario ingegnarsi: l'unica possibilità è costruire due gruppi di destinatari che sono simili rispetto a tutte le variabili del modello (tutte le variabili esogene) **tranne una**, che è proprio **l'aver ricevuto o meno l'intervento di policy** (la variabile trattamento, D).
- Costruire due gruppi di destinatari identici in tutte le variabili tranne aver ricevuto il trattamento consente di ricreare una situazione *ceteris paribus* e quindi di ricostruire una situazione controfattuale
- Avrò quindi bisogno di individuare due gruppi di destinatari simili: un gruppo che ha ricevuto il trattamento i policy, ed un gruppo che non ha ricevuto il trattamento, detto anche **gruppo di controllo (vedi slide successiva)**

Soluzione: determinare un plausibile gruppo di controllo

- Per eliminare le distorsioni e poter valutare l'effetto netto, è dunque necessario individuare un plausibile gruppo di confronto, detto **gruppo dei controlli**
- Il gruppo di controllo è formato da soggetti **non trattati** che sono però **equivalenti** ai trattati rispetto alle variabili rilevanti per Y (nel diagramma, A, B, F, G)
- Per costruire il gruppo di controllo, è necessario individuare **tutte le variabili responsabili del processo di selezione-assegnazione al trattamento** “servizio di orientamento”, reperendo informazioni affidabili su tali variabili
- **Costruire il gruppo di controllo consente di riprodurre la condizione ceteris paribus e eliminare la distorsione: trattati e non trattati sono equivalenti in tutte le variabili che condizionano Y, tranne che in D (il programma di policy)**
 - **In questo modo, la differenza tra le medie dei due gruppi può fornire una stima dell'effetto del programma di policy**

Il diagramma di Pearl con condizionamento delle variabili A, B, F, G

- I segmenti punteggiati indicano che rispetto alle variabili di partenza è stata realizzata la condizione di equivalenza. La freccia nera tra D e Y indica che la differenza tra le medie dei due gruppi (trattati e gruppo di controllo di non trattati) è una stima dell'effetto dell'intervento



Ricostruire i due gruppi: how to?

- Ma... come è possibile tecnicamente individuare il gruppo dei trattati e il gruppo dei controlli stante la distorsione da selezione o selection bias?
- In altri termini, come me la cavo quando non posso utilizzare il metodo sperimentale, e devo ricostruire due gruppi di destinatari simili in tutto tranne che ne trattamento?
- Di seguito, illustreremo le principali strategie per ricreare la situazione controfattuale e consentire l'identificazione degli effetti medi
- **ATTENZIONE:** non dovete spiegare il funzionamento delle diverse strategie, ma solo illustrare la logica che rende i due gruppi individuati comparabili

Strategie per individuare gli effetti medi attraverso la logica del controfattuale

Principali strategie di identificazione degli effetti medi

- Per riprodurre una situazione quanto più possibile simile a quella dei modelli sperimentali, si deve rendere equivalenti i trattati e i non trattati **RISPETTO A tutti gli aspetti rilevanti** (schematizzati nel diagramma di Pearl), ricostruendo ex post la condizione ceteris paribus
- A questo scopo, **l'analisi di impatto ricorre a 4 principali strategie:**
 1. **La selezione delle osservabili e il metodo dell'abbinamento statistico (nel wp, l'esempio è fatto sui servizi di inserimento lavorativo)**
 2. **Il confronto attorno alla soglia di discontinuità, o RDD (es. su inserimento)**
 3. **Distorsione da dinamica spontanea (es. su incentivi)**
 4. **Il doppio confronto spazio temporale, o difference-in-difference (es. su incentivi)**

1. La selezione delle osservabili e il metodo dell'abbinamento statistico

Per calcolare l'effetto netto nella policy attraverso la strategia del metodo dell'abbinamento statistico, che presuppone la disponibilità di ampie basi di dati, occorrono **due passaggi**:

- I. **Abbinare a ogni soggetto trattato un soggetto NON trattato che presenta le stesse caratteristiche -> il risultato Y del soggetto non trattato è una stima del risultato controfattuale**
 - II. **Calcolare la media delle differenze tra i risultati per osservati per le coppie dei soggetti abbinati**
- Nel caso dei servizi di orientamento, ad ogni soggetto trattato viene abbinato un non trattato che ha una simile propensione al trattamento (stimata attraverso i valori di abilità individuali, selezione nell'assegnazione, etc.)
 - Nel grafico seguente, la propensione ad accedere al servizio è rappresentata da un indicatore sintetico che riassume tutte le informazioni sulle variabili rilevanti e diverse dalla variabile trattamento. Tale indicatore composito di propensione varia da 0 a 1

es. il metodo dell'abbinamento statistico rispetto alla propensione al trattamento

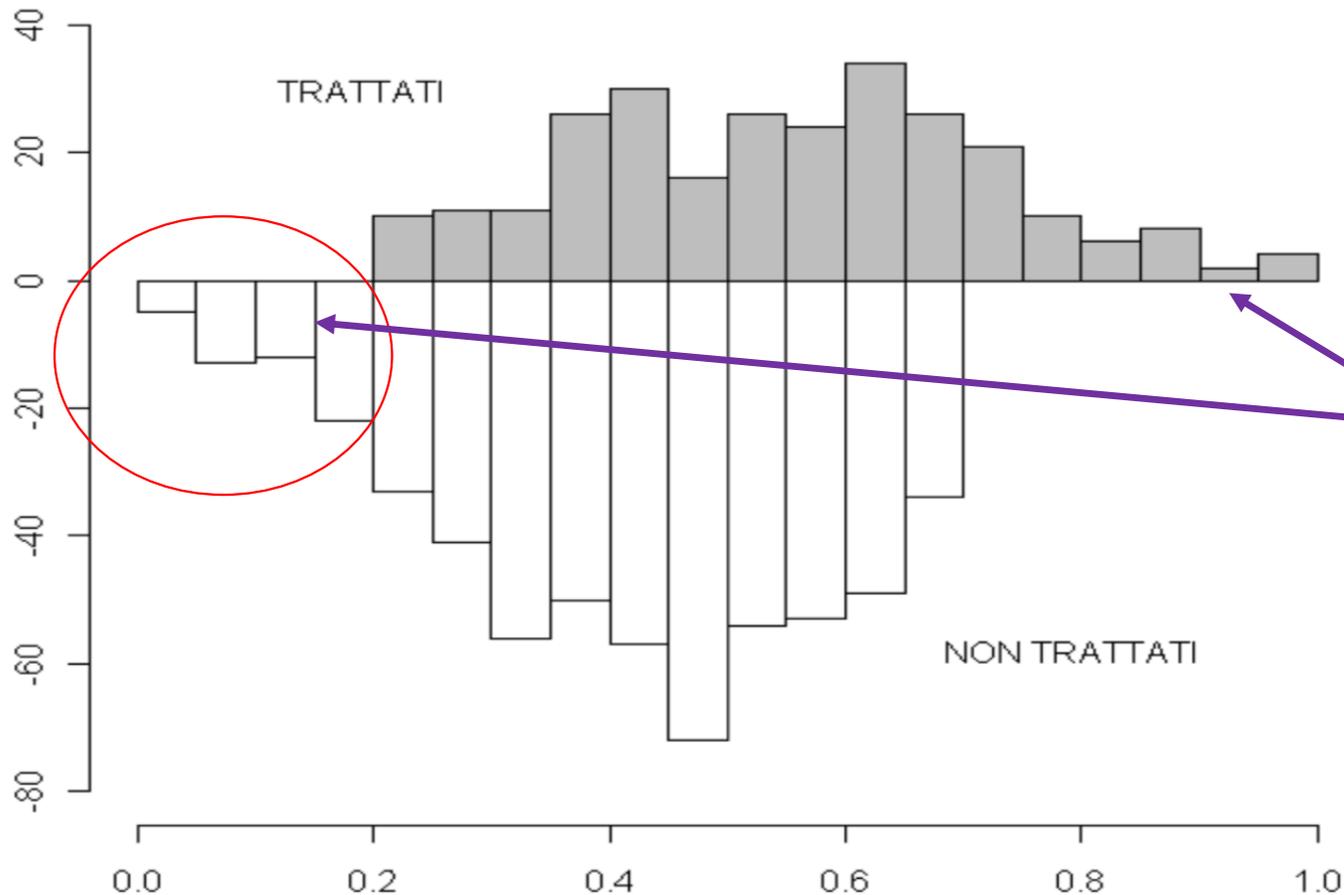


Grafico: ipotetico diagramma di distribuzione di trattati e non trattati per un servizio di orientamento al lavoro sulla base di indicatore di propensione che varia da 0 a 1 (Trivellato 2009, 19)

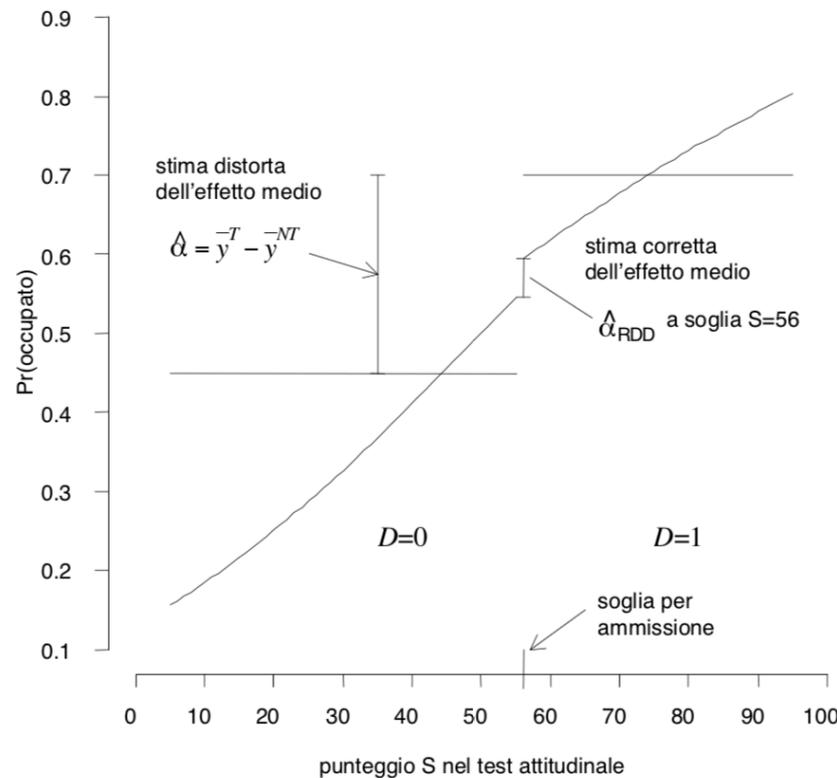
- sono necessarie ricche basi dati individuali sugli iscritti alle liste di disoccupazione e sugli operatori per fare gli abbinamenti
- in questo caso, molti dei trattati possono essere abbinati a simili non trattati
- NB: dal grafico si vede che ai trattati $>0,7$ e $<0,2$ non corrispondono non trattati
- considerando gruppo non trattati $<0,2$, si nota che **i disoccupati che avrebbero più bisogno di servizi di sostegno al lavoro non accedono al trattamento**

DUNQUE: l'analisi valutativa mostra che **il programma ha funzionato relativamente male**, quindi va ridisegnato

2. Il confronto attorno alla soglia di discontinuità

- Qui si considerano i programmi di policy che prevedono un razionamento, dato da una **soglia di ammissibilità** (che corrisponde ad un certo valore di una variabile o a una posizione in una graduatoria)
- In un programma di orientamento al lavoro, il trattamento viene erogato ai disoccupati che hanno ottenuto un punteggio superiore ad una certa soglia S in un test attitudinale
- La soglia di ammissibilità al trattamento è un **punto di discontinuità** intorno al quale si manifesta una situazione simile alla randomizzazione -> **in altre parole, i trattati subito al di sopra della soglia sono equivalenti ai trattati subito al di sotto della soglia**
- **L'effetto della politica qui è dato dal confronto tra trattati e non trattati intorno alla soglia di ammissibilità**
 - cioè l'effetto della policy è la differenza nella media di Y per il gruppo dei trattati subito sopra soglia e il gruppo dei non trattati subito sotto soglia

- Nei disegni di ricerca che utilizzano il confronto attorno alla soglia, detti anche *regression discontinuity designs*, è **importante sottolineare un caveat**: la stima dell'effetto che si ottiene è circoscritta solo ai soggetti poco sopra e poco sotto soglia – non si può dire nulla degli altri soggetti più lontani dalla soglia (senza stimare altri parametri) -> **In altre parole non si può usare come stima dell'effetto la differenza tra i risultati medi di tutti i trattati e di tutti i non trattati**



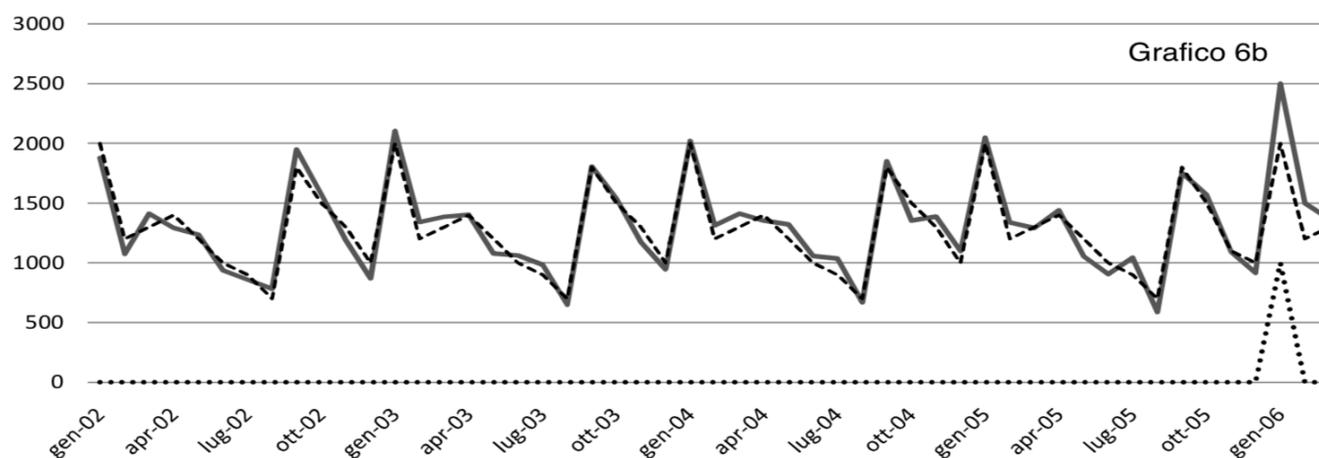
Graf. 5a: *Stima dell'effetto medio in corrispondenza della soglia per l'ammissione*

3. La distorsione da dinamica spontanea

- *Se destinatari ammissibili e trattati coincidono* (ad esempio perché la policy è universalistica), la questione dirimente diventa confrontare gli stessi soggetti ***prima e dopo*** aver ricevuto il trattamento -> es. incentivi alle imprese per la stabilizzazione precari
- Dati utilizzati: non a livello micro, ma a livello aggregato = serie storiche di trasformazioni di contratti TD in contratti TI
- In una certa area, si osserva la variabile risultato nei periodi prima e dopo l'erogazione del trattamento: nei valori di Y prima e dopo sono espresse in realtà due grandezze:
 - l'effetto del trattamento
 - la naturale evoluzione del fenomeno, detta anche **distorsione da dinamica spontanea**
- **Se si hanno a disposizione serie storiche del valore Y, il controfattuale si può ottenere elaborando una proiezione della storia pre-trattamento**

- **Dunque se si hanno a disposizione serie storiche del valore Y, il controfattuale è dato dalla proiezione della storia pre-trattamento**
→ si utilizza la storia della variabile risultato Y per predire quale sarebbe stata la sua evoluzione *se non fosse stato attuato l'intervento D*
- L'effetto del trattamento è qui dato dalla **differenza tra l'andamento osservato** post-intervento (quello che di fatto è avvenuto) **con l'andamento predetto grazie a una serie storica sufficientemente lunga (il controfattuale)**

Un esempio: programma di incentivi per stabilizzazione precari realizzato a gennaio 2006



Graf. 6b: *Serie storiche lunghe*

Grafico 6: *Ipotetico esempio di un programma di incentivi per la stabilizzazione dei precari realizzato nel gennaio 2006*

Legenda: spezzata continua: serie osservata; spezzata tratteggiata: serie prevista; spezzata punteggiata: trasformazioni con incentivi.

- La serie spezzata punteggiata mostra che l'effetto stimato a gennaio 2006 è di soli 500 contratti trasformati.
- Le imprese si sono comportate per lo più nello stesso modo in cui si sarebbero comportate se l'incentivo NON ci fosse stato.
- Possibili comportamenti adattivi: non assumo aspettando l'incentivo

4. Il doppio confronto spazio temporale o metodo «difference-in-difference»

- Se ho una serie storica della variabile risultato **Y sia per il gruppo dei trattati sia per il gruppo dei non trattati**, rilevata su gruppi di destinatari simili-> si può eliminare la distorsione con il metodo del doppio confronto
- Nel caso degli incentivi alle imprese, **si confrontano due aree industriali A e B con caratteristiche simili, una esposta al trattamento e l'altra no.**
- Si raccolgono dati su contratti TD e trasformazione in contratti TI sia per area A che per area B, sia nel periodo pre incentivi (erogati solo a A) che nel periodo post incentivi
- **L'effetto della politica è individuato confrontando:**
 - **La differenza tra prima e dopo l'attuazione dell'intevento (confronto nel tempo)**
 - **La differenza tra trattati e non trattati (confronto tra soggetti)**
 - Caveat: L'assunto di base è che si possa escludere l'esistenza di differenze sistematiche nei due gruppi, trattati e non trattati, nel trend della variabile Y in assenza della politica

- l'effetto medio dell'intervento del programma di incentivi per la stabilizzazione dei precari può essere ottenuto *sottraendo al cambiamento osservato nella frazione delle trasformazioni dei contratti temporanei in contratti a tempo indeterminato nell'area A l'analogo cambiamento osservato nell'area B*

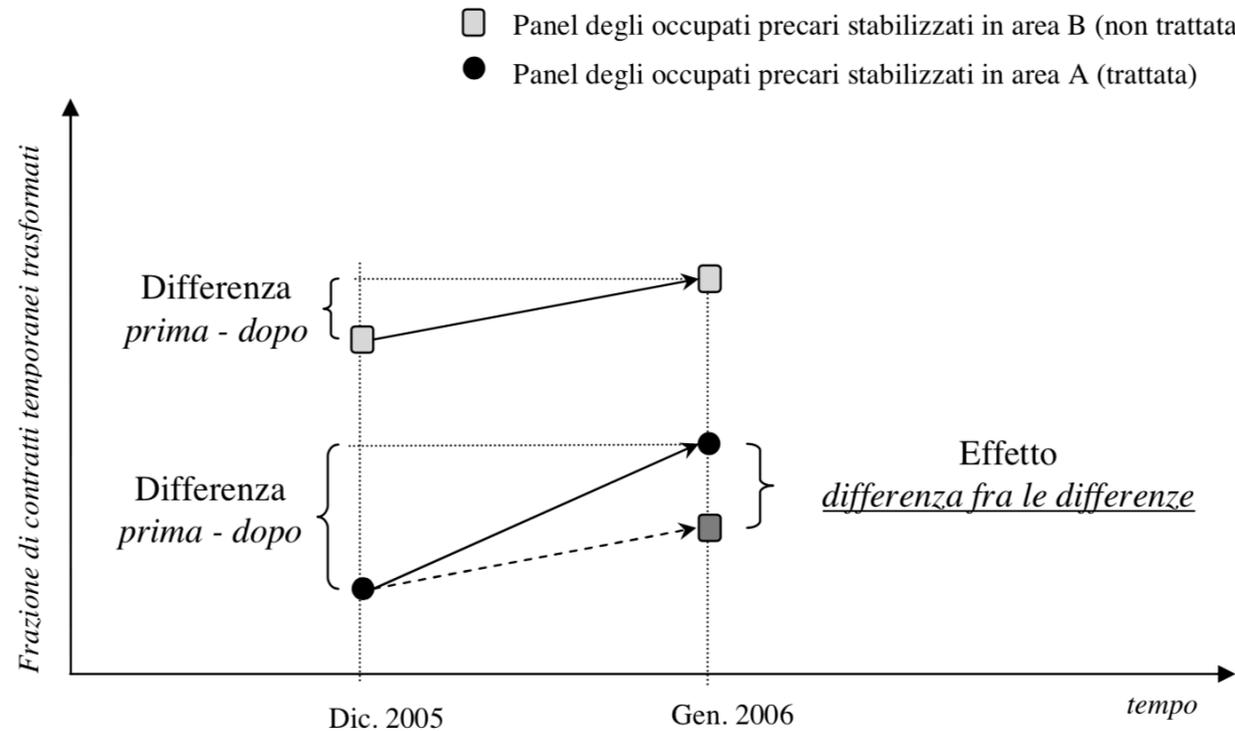


Grafico 7: *Il doppio confronto spazio-temporale applicato all'esempio del programma di incentivi per la stabilizzazione dei precari nell'area A nel gennaio 2006*

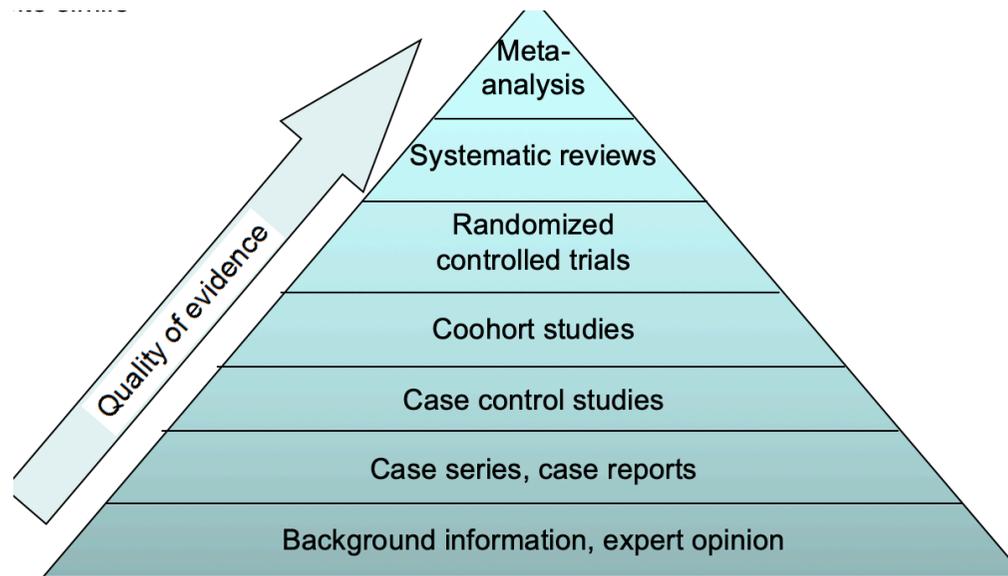
Discutere le buone pratiche
nell'utilizzo della valutazione di
impatto in alcuni paesi

Uno scorcio alle buone pratiche negli altri paesi

- Le pratiche di valutazione di impatto nascono negli Stati Uniti, metà anni 60, per valutare i programmi di welfare
- Tali pratiche si diffondono poi anche oltreoceano: nei paesi Westminster, nei paesi scandinavi, nei paesi continentali (Germania), ma anche nei paesi in via di sviluppo sotto la spinta della Banca Mondiale
- Le buone pratiche nei paesi stranieri mostrano tutte che inserire **nella fase di disegno e nella fase realizzazione della policy la necessità di una vera e propria analisi di impatto** è un aspetto fondamentale per poter avere a disposizione i dati necessari a stimare correttamente gli effetti

Esempio: la What Works Clearing House: US

- La What Works Clearing House del Education Department (USA) costruisce conoscenze solide sulle molte evidenze dalle politiche educative
- Dal un ampio numero di ricerche valutative sull'educazione, la WWC: accerta il rigore dell'evidenza scientifica sull'effettività degli effetti; sviluppa standard per rivedere e sintetizzare la ricerca valutativa
 - Sviluppa procedure per strutturare rigorosamente il processo di comparazione e sintesi



Un altro esempio: le riforme Hartz in Germania

- Dopo l'unificazione, il sistema di welfare tedesco fu sottoposto a forti tensioni -> affronta i problemi con una serie di riforme, le riforme Hartz, note a partire dal nome della commissione che le ha elaborate, realizzate in quattro fasi dal 2002 al 2005
- Per tutti i principali programmi di riforma era prevista una valutazione sistematica sia dell'implementazione che degli effetti
- Investimento nel miglioramento del sistema informativo
- Possibilità di accesso ai micro dati da parte dei ricercatori
- Produce una diffusa attività di valutazione da parte di istituti di ricerca e gruppi di ricercatori
- Crescente attenzione ai risultati delle valutazioni nel dibattito pubblico

Ancora: le politiche di avviamento al lavoro in Francia

- Anche in un paese la cui tradizione amministrativa non gioca a favore della ricerca valutativa, questo approccio si afferma la valutazione grazie all'interazione tra Ministero del Lavoro e accademia
- A partire dal 2007 viene avviato un esperimento sociale su ampia scala su 200.000 disoccupati di lunga durata, per valutare tre diverse modalità di avviamento al lavoro

”Nuove” politiche di welfare: UK

- Nel 2008, il Department for Work and Pensions, propone un Nuovo programma di riforma del welfare sul lavoro attraverso il **Green paper** **“No one written off: reforming welfare to reward responsibility”**
- Qui la buona pratica riguarda come disegnare un nuovo e ambizioso programma di riforme, utilizzando evidenze empiriche (evidenze based) e promuovendo una consultazione pubblica
- Per promuovere il green paper
 - Background research and discussion paper – discute precedenti esperienze di riforma a partire da studi di valutazione
 - Consultazione strutturata
 - Prevede una rigorosa analisi dell’impatto della regolazione (AIR – ex ante): costi, benefici, effetti attesi

Cosa suggeriscono queste esperienze di valutazione degli effetti?

- Le buone pratiche suggeriscono alcuni punti rilevanti per il disegno delle politiche e per la valutazione dei loro effetti
- Presupposto per avere buone pratiche di valutazione, **e quindi una istituzionalizzazione delle pratiche di valutazione (vedi slides Regonini sui capitoli 3 e 5 del manuale)**
- **Importanza fondamentale degli assetti istituzionali:** ad esempio, è utile **inserire clausole valutative all'interno dei testi di legge**. Tali clausole valutative devono contenere gli obiettivi della valutazione, le indicazioni sulle modalità di svolgimento, gli stanziamenti per lo svolgimento

- Le clausole valutative funzionano però in presenza di certe **pre-condizioni**:
 - a) Il processo di ridefinizione delle politiche deve essere partire dalla disponibilità di dati, e quindi essere evidence-based, basato sulle evidenze
 - b) Il modello anglo-americano di accountability democratica: le funzioni di controllo dell'esecutivo e valutazione delle politiche sono svolte da agenzie indipendenti che agiscono come watchdogs: specializzazione, autonomia, operano *per conto* del legislativo

Nb: Al contrario di quanto accade nel modello Westminster o anglo americano, **il modello continentale è fondato sul controllo delle magistrature indipendenti ed equidistanti da esecutivo e legislativo (es. Corte dei Conti) -> controllo di conformità di legge e di bilancio**

- anche in questo contesto si può affermare una tendenza a potenziare le funzioni di controllo del legislativo – vedi riforma costituzionale 2008 in Francia: il parlamento controlla il governo e valuta le politiche pubbliche
- c) Se la valutazione è istituzionalizzata, la ridefinizione delle politiche avviene in un contesto di dibattito pubblico focalizzato sulle policy e sulle evidenze
 - Sviluppo del dibattito attraverso università, think tanks, media

I cinque aspetti che caratterizzano le buone pratiche per la valutazione secondo Trivellato

- 1) Occorre iniziare la valutazione il prima possibile, addirittura prima di attuare la riforma
 - Importanza del disegno della policy e del disegno della valutazione dei suoi effetti
 - Implica una interazione tra policy-maker e valutatore
 - Sforzo di concretezza nella discussione pubblica
 - Si individuano subito le info necessarie per ricostruire la situazione controfattuale
- 2) Data la complessità, le buone pratiche combinano analisi di implementazione e valutazione degli effetti
- 3) Necessità di buoni sistemi informativi per avere base dati adeguata

I cinque aspetti che caratterizzano le buone pratiche per la valutazione

- 4) L'ideale sarebbe poter svolgere piccoli progetti pilota prima di attuare la politica su larga scala (social experiment)
- 5) Considerare che la valutazione non è una panacea o la pietra filosofale -> importanza di iniziative per cumulare e sintetizzare la conoscenza disponibile
 - Anche la valutazione è un processo che si sviluppa e rafforza nel tempo:
 - conoscenza cumulativa
 - accessibilità delle informazioni valutative ad un pluralità di analisti
 - replicabilità degli studi

Un bilancio al 2009 delle conoscenze valutative in tema di lavoro

- Nel merito, gli studi di valutazione degli effetti delle politiche del lavoro mostrano che queste politiche, nei paesi più capaci, sono spesso inefficaci; anche se mostrano eterogeneità dei risultati nei diversi gruppi
- Es. Heckman et al. 1999 notano che le spese pro capite sono troppo basse per superare i deficit di abilità
- Più in generale, gli studi di valutazione mettono in evidenza che un targeting adeguato migliora gli effetti
- Alcune evidenze, secondo Trivellato, sono «conclusive»: le politiche di occupazione sussidiata sono inefficaci; politiche che combinano formazione e obblighi reciproci hanno effetti positivi

L'esperienza italiana in tema di valutazione (al 2009)

- I rapporti dei Ministri al Parlamento sullo stato di attuazione di una politica hanno una lunga storia, ma **non sono finalizzati** alla valutazione degli effetti – al più danno info sul processo di implementazione (attività organizzative, risorse, soggetti, risultati “lordi” dati dalle differenze pre-post intervento, difficoltà incontrate etc.)
- Un primo tentativo di innovazione viene fatto dalla Commissione Onofri per la riforma del welfare alla fine degli anni 90
- Altre innovazioni su strumenti e attività di valutazione: avvio di esempi di sperimentazione ed esercizi di valutazione che coinvolgono i policy makers
 - Studi di valutazione commissionati dal Ministero del Lavoro e Sviluppo economic
 - Unità di valutazione investimenti pubblici (UVAL)
 - Studi commissionati dalle Regioni e introduzione delle clausole valutative es. CAPIRE

Gli sviluppi più recenti della valutazione in Italia

- La decennale sperimentazione di clausole valutative a livello regionale: il progetto CAPIRE
- L'Ufficio di Valutazione di impatto: una rilevante novità istituzionale
 - Alcuni aspetti critici: disegno, livelli di istituzionalizzazione; «regole di ingaggio»
- Il tentativo (fallito) di inserire la valutazione nel processo decisionale tramite la riforma Costituzionale
- Lo sviluppo di un maggior numero di think tank che adottano anche tecniche di valutazione di impatto (Es. IRVAPP)
- L'impegno nel policy advice (la consulenza esperta sui problemi di policy) da parte di enti pubblici di ricerca (es. INAPP) e agenzie governative (ANPAL, INPS, altri)
- Nonostante i segnali di cambiamento positivo, il livello di istituzionalizzazione delle pratiche di valutazione nel sistema italiano è ancora basso (vedi articolo Regonini 2018)
 - **Allo stesso tempo, anche il sapere prodotto dalle ricerche valutative e dalla evidence-based policy making è sempre più contestato all'interno della società – considerate ad esempio la polemica di questi giorni tra il Premier UK Boris Johnson e gli epidemiologi inglesi**

IMMUNITÀ DI GREGGE

LONDRA, EPIDEMIOLOGI BRITANNICI CONTRO LE DECISIONI DI BORIS JOHNSON

Gli scienziati britannici, in due lettere aperte, dichiarano la loro preoccupazione sull'approccio al Coronavirus da parte del governo inglese

di Tiziana Di Giovannandrea

14 marzo 2020 - Centinaia di membri della comunità scientifica britannica hanno inviato due lettere aperte al governo del Regno Unito, esprimendo forte preoccupazione per l'approccio scelto contro la diffusione del Coronavirus nel Paese.

In una missiva, **198 accademici** nell'ambito della matematica e della scienza hanno esortato ad adottare misure urgenti di distanziamento sociale, definendo la strategia governativa che punta allo sviluppo di un'immunità di gregge un'opzione che "non sembra praticabile".

In un'altra lettera, **164 scienziati** comportamentali hanno messo in dubbio la tesi di Downing Street secondo la quale se all'opinione pubblica vengono imposte misure drastiche troppo presto, alla fine torneranno al comportamento precedente.

Finora il governo britannico si è rifiutato di adottare iniziative sociali, come la chiusura delle scuole o addirittura la chiusura della nazione come altri Paesi europei, consigliando l'auto-isolamento per chiunque mostri sintomi del Coronavirus. Per il premier britannico, Boris Johnson, provvedimenti contro la diffusione del Coronavirus sono inutili. Per evitare il contagio sarebbe sufficiente lo sviluppo dell' 'immunità di gregge'.

L'immunità di gregge si ha quando un gran numero di persone sono immunizzate e o vaccinate, nel caso ci sia un vaccino, contro un'infezione. Nel caso dell'attuale virus ancora non c'è un vaccino e quando ci sarà bisognerà valutare l'efficacia del vaccino sulle persone. Se il vaccino funzionerà tutti i vaccinati faranno da barriera alla diffusione del contagio.

Considerazioni conclusive (i)

- Lo studio del saggio di Trivellato consente di entrare nelle difficoltà della ricerca valutativa che ha come obiettivo l'identificazione degli effetti netti di un intervento di politica pubblica
- Questa tecnica, la valutazione di impatto, è però solo una delle tecniche che attraverso la quale è possibile valutare l'efficacia di un intervento di policy in fase ex post
- Approfondendo la logica del controfattuale e le strategie per ricreare i gruppi di controllo, abbiamo potuto apprezzare come **la disponibilità di dati completi e affidabili sia un elemento imprescindibile**, non solo per l'analisi di impatto, ma per tutta la valutazione
- Inoltre, è emerso come tale disponibilità di dati è facilitata dall'**istituzionalizzazione delle pratiche valutative**
- La vicenda italiana della valutazione dimostra come tale istituzionalizzazione sia complessa, **ma anche che con il tempo il mestiere del valutatore di politiche sta diventando più rilevante**

Considerazioni conclusive (ii)

- Alla fine di questa piccola maratona sulla valutazione di impatto, lo studente dovrebbe aver acquisito le conoscenze necessarie a rispondere a queste domande
 - Esistono dei metodi per isolare gli effetti di un intervento di policy sul problema?
 - Perché la logica del controfattuale ci consente di rilevare gli effetti correttamente?
 - Con quali strategie è possibile ricreare la situazione controfattuale?
 - Quali sono le buone pratiche da seguire per consentire lo sviluppo della ricerca valutativa?
 - Quanto conta l'istituzionalizzazione delle pratiche valutative per lo sviluppo di questo tipo di conoscenza sulle politiche?
 - In una situazione come quella della attuale crisi sanitaria da COVID-19, quali possono essere i principali ostacoli per l'elaborazione di politiche basate sull'evidenza, e per il disegno di ricerche di valutazione dei reali effetti delle misure adottate dai diversi governi europei?