

# **LA VALUTAZIONE D'IMPATTO DI POLITICHE: METODI E CRITERI , CON (SCARNI) RIFERIMENTI A POLITICHE DI RIORGANIZZAZIONE DEL SISTEMA SANITARIO**

Ugo Trivellato

FBK-IRVAPP e Università di Padova

# Per quali politiche pubbliche?

2

**Politica pubblica**  $\equiv$  **programma**  $\equiv$  **intervento**  $\equiv$  **trattamento**: una ben definita **misura** rivolta a un ben definito **insieme di destinatari** con lo scopo di **causare un cambiamento in condizioni o comportamenti dei partecipanti** ritenuti problematici.

Non tutta l'azione pubblica risponde a queste caratteristiche: ad es., **non le “macropolitiche”**,  
**non le prestazioni correnti erogate da PA.**

Ma **una parte rilevante sì.**

# Le politiche di riorganizzazione del servizio sanitario: scarni cenni

3

Fra tali politiche si collocano le **innovazioni nella organizzazione dei servizi sanitari**, motivate da preoccupazione di *efficacia* (analisi degli effetti) e di *efficienza* (analisi costi-efficacia).

**Tre esempi** come scarno riferimento:

- Profili F., *La valutazione delle politiche in ambito sanitario in Toscana*, Convegno ARS Toscana, Firenze, 14.06.2014.
- Arena V. et al., *La composizione delle misure nella valutazione delle politiche sanitarie*, inizialmente in AIE 2015.
- Fiorio C. et al., *Evaluating organizational change in health Care: The Patient-Centered Hospital Model*, 2015, mimeo.

# Uno studio di caso stilizzato

## “*Patient-Centered Hospital Model*” vs. configurazione di ospedale funzionale

- Il **trattamento**: organizzazione flessibile e orientata al processo/*multi-speciality* units/assistenza infermieristica modulare/aggregazione delle risorse/management centralizzato di posti letto/costruzione di H orientata a aggregazione di risorse, raggruppamento di pazienti, flessibilità e modularità degli spazi.
- vs. precedente configurazione: **H funzionale**.

# Un studio di caso stilizzato (segue)

- La **popolazione** di interesse: H pubblici non orientati alla ricerca della Lombardia: dati aggregati su **dimissioni** per ognuna delle **25 principali categorie diagnostiche (MDC)**, per 86 H, dei quali 3 con *P-C Model* dall'ultimo trimestre 2010, osservate al più per 9 anni (2004-12).  
Totale osservazioni: 12.120.
- Le **variabili risultato**:
  - (i) **efficacia**: numero medio di ri-ospedalizzazioni entro 30 gg. rispettivamente in H di regione e nello stesso H;
  - (ii) **efficienza** numero medio di giorni di ricovero in H; numero medio di trasferimenti di un paziente fra differenti dipartimenti dell'H.

# L'interrogativo cruciale

6

Il programma **funziona**? ha operato **in modo efficace**?  
**Efficacia**  $\equiv$  capacità del programma di produrre **effetti**,  
innanzitutto quelli attesi (ma anche quelli inattesi).

Tre sono gli snodi:

- quali sono gli **strumenti analitici appropriati**? Capaci cioè di produrre solida evidenza empirica sugli effetti delle politiche – sulla loro efficacia, sulla loro eterogeneità – allo scopo di **migliorare le politiche**?
- Come possiamo di **validare**, con “checks di **robustezza**” l’appropriatezza del metodo usato?
- Perché è cruciale una **valutazione** “prospettica”?

# Il seguito

7

1. La valutazione controfattuale degli effetti  
*Counterfactual impact evaluation* (CIE)
2. I test di sovraidentificazione
3. La valutazione prospettica

# 1. La valutazione controfattuale degli effetti

8

Occorre affrontare una questione di ordine generale, di natura cognitiva: l'**attribuzione causale**.

A mo' di prologo:

*«Nei primi anni '60 il nostro medico di famiglia a Blackpool si chiamava dottor Piggott. [...] Se avevi il morbillo, o la varicella, o la bronchite, ti prescriveva una medicina e guarivi. Se guarivi **grazie** alla medicina era un'altra storia.»*

Tim Parks, *Insegnaci la quiete*, 2010



# La questione cognitiva a fronte di agenti

9

L'attenzione è su:

- interventi che si applicano ad **agenti** (persone, famiglie, imprese, scuole, ospedali, ecc.), che hanno:
  - irriducibile componente di **variabilità individuale**;
  - **reazioni comportamentali**;
- analisi della causalità **probabilistica**  $\cong$  la politica (la “causa”) non produce necessariamente l'effetto per tutti i trattati (o, all'opposto, per nessuno); tipicamente, cambia la probabilità che l'effetto si manifesti, e questa probabilità può variare per sottogruppi di trattati ( $\equiv$  eterogeneità degli effetti).

# Attenzione a processi di apprendimento meccanici (effetto $\neq$ cambiamento)

10

In generale, l'effetto **non è la differenza fra le medie osservate:**

- **di trattati e non-trattati**, perché in generale i due gruppi differiscono per caratteristiche (i) presenti dall'origine, prima del programma e (ii) che influiscono sul risultato ( $\equiv$  **distorsione da selezione**).

Il detto *cum hoc ergo propter hoc* è in generale falso.

- in un programma universale, **di trattati dopo e prima (quando erano non-trattati)**, perché in generale in assenza del programma il fenomeno avrebbe avuto una sua dinamica spontanea.

Il detto *post hoc ergo propter hoc* è in generale falso.

**Osserviamo associazione; ci interessa l'attribuzione causale.**

# Il controfattuale

11

**La logica controfattuale:** per un dato soggetto in un dato tempo **l'effetto** di un intervento è la **differenza** tra quanto si osserva **in presenza** dell'intervento (dopo che il soggetto è stato trattato) e quanto si sarebbe osservato **in sua assenza** (se, cioè, il soggetto non fosse stato trattato)

Il primo risultato è **fattuale**.

Il secondo risultato è ipotetico, non osservabile per definizione, ed è detto il **controfattuale**.

# Il problema fondamentale

12

- I. Il **problema fondamentale dell'attribuzione causale**: per un dato soggetto è **impossibile osservare** entrambi i risultati potenziali:  $\Rightarrow$  è impossibile **osservare** l'effetto dell'intervento a livello **individuale**.

Nella ricerca scientifica – e nelle politiche – interessano **effetti medi** per collettività: sui trattati (oppure, se vi è eterogeneità, su sottogruppi entro i quali l'effetto sia omogeneo).

- II. **Identificare effetti causali medi** è possibile sulla base di **assunzioni**, che consentono di **(ri)costruire in maniera credibile il controfattuale** a partire da dati osservati: sui *non-trattati* e/o sui *trattati pre-intervento*.

# Strategie per costruire il controfattuale (1)

13

Se il decisore è in grado di “manipolare” il processo di assegnazione, l'**esperimento randomizzato** è il modo “ideale” per costruire il controfattuale, perché elimina la distorsione da selezione **per costruzione**.

- Un insieme di potenziali destinatari dell'intervento è suddiviso in due gruppi – “sperimentale” e “di controllo” – **mediante sorteggio**: ⇒ fra i due gruppi non vi sono differenze sistematiche di partenza, sicché il risultato medio dei controlli “approssima bene” quello che sarebbe stato il risultato medio dei trattati se non lo fossero stati.
- In un esperimento randomizzato (in ambito medico, RCT) è cruciale assicurare la **randomizzazione** e preservarne l'**integrità**: ⇒ rigorosi **protocolli** e uso di disegni “**cieco**” e “**doppio cieco**”.

# Ed esperimenti sociali (“*field trials*”)? ...

14

- **Difficoltà nel realizzarli**, per ragioni:
  - **giuridiche**: partecipazione non può essere imposta;
  - **etiche**: se ci attendiamo che l'intervento abbia effetti positivi, in nome di che cosa escludiamo qualcuno dal beneficiarne?
  - **basso consenso** (di operatori, opinione pubblica);
  - **problemi di non obbedienza** all'assegnazione.
- **Opportunità di realizzarli** tanto maggiore quanto più:
  - l'intervento è su **piccola scala**, è un *pilot*,
  - la sua **efficacia è incerta**, sicché servono per imparare (e rivelano **l'intenzione di imparare**),
  - Si devono fronteggiare problemi di **razionamento**.

# Quando l'esperimento non è possibile (2)

15

... non siamo in un *cul de sac*.

Gli ultimi quarant'anni hanno visto una forte crescita di metodi per situazioni **non sperimentali**.

- Il valutatore non può “manipolare” l'esposizione alla politica ( $\equiv$  il processo di selezione).
- **Osserva** ciò che avviene e cerca di ricondursi alla condizione in cui confronta trattati e non-trattati **a parità delle altre condizioni**.

L'idea-guida: **comparare soggetti comparabili**,  
cioè a dire comparare ai trattati **i non-trattati resi equivalenti**, **salvo il fatto che sono non-trattati**.

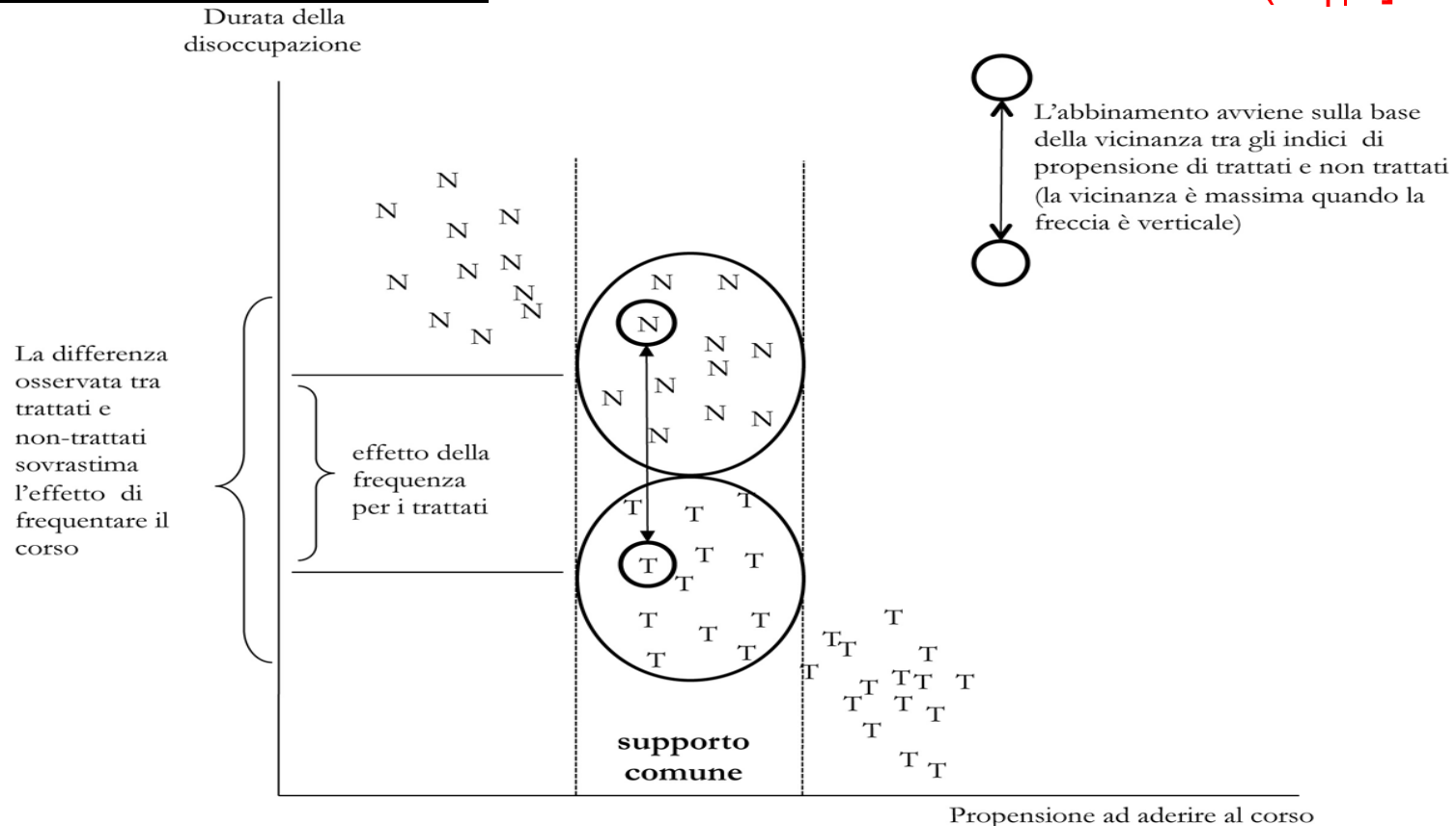
# Alcune classi di metodi per ricostruire il controfattuale: (i) *statistical matching*

16

Es: Corso di f.p. per disoccupati, **con partecipazione volontaria**.

Variabile risultato: durata della ulteriore disoccupazione.

Assunto di identificazione: selezione nelle osservabili:  $M(C_T|x) = M(NT|x)$





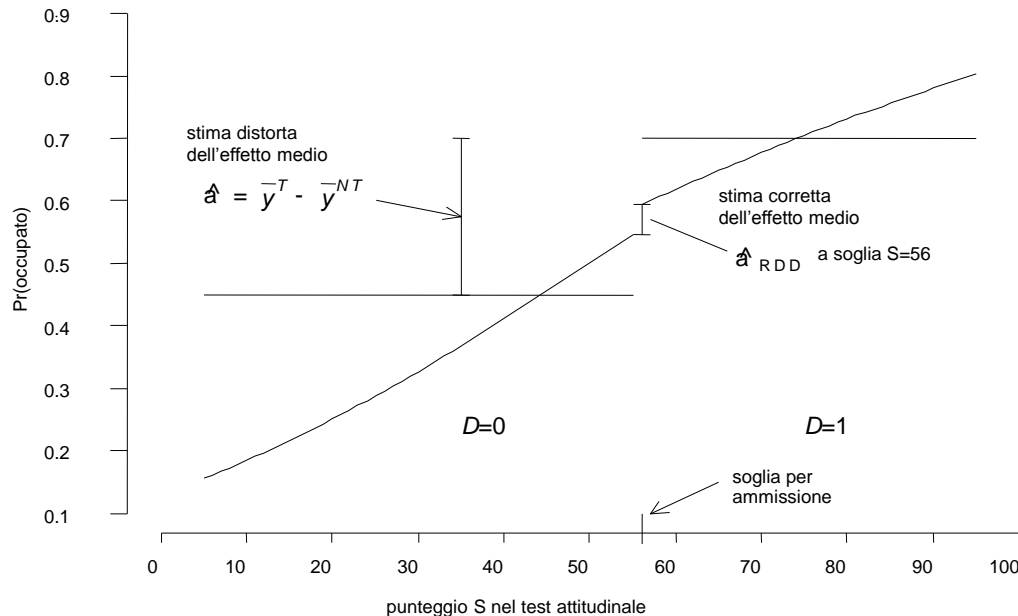
# Alcune classi di metodi per ricostruire il controfattuale: (ii) *Sharp RDD*

17

Es.: Concorrenti a un corso per infermieri ordinati secondo un indicatore  $S$  (punteggio a un test attitudinale); lo stato di trattato/ non-trattato segue deterministicamente la regola  $D = 1(S \geq s^*)$ , dove  $s^*$  è un “punto di discontinuità” noto nel supporto di  $S$ .

Variabile risultato: probabilità di occupazione.

Assunto di indentificazione:  $E[Y_0 | s]$  è una funzione continua di  $s$  a  $s^*$ .

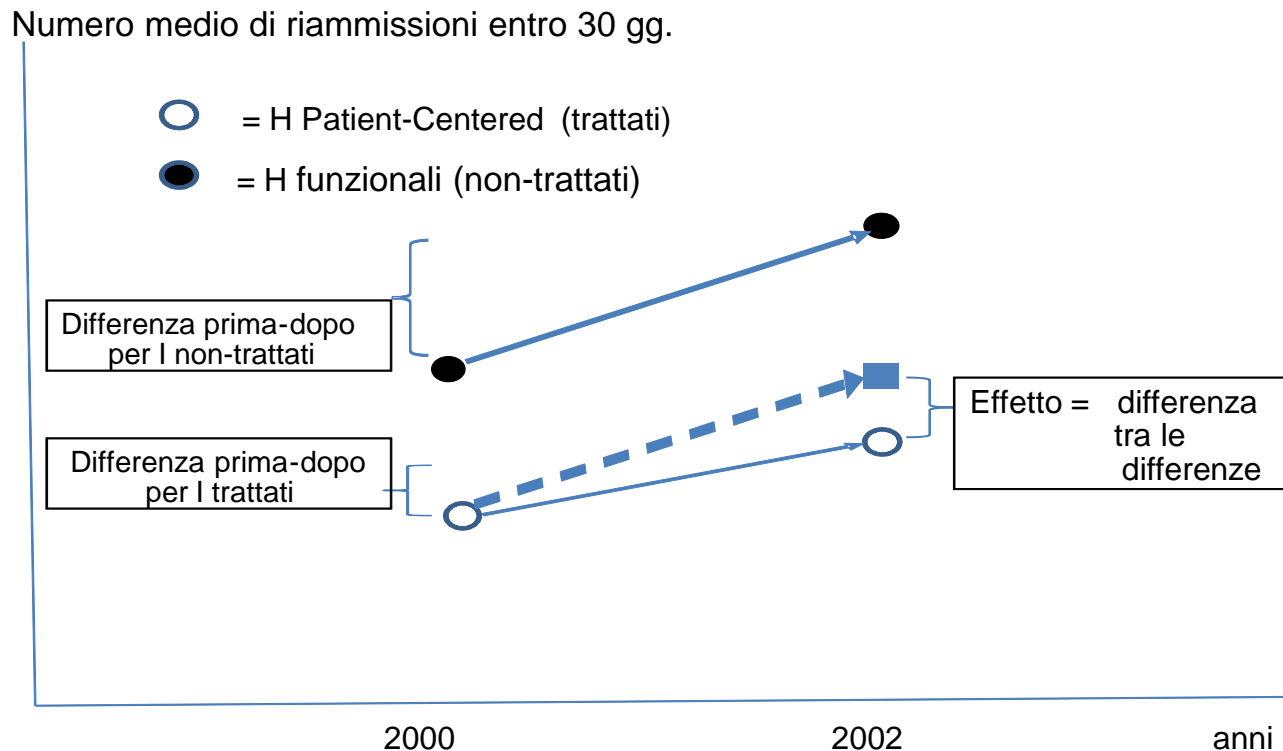


# Alcune classi di metodi per ricostruire il controfattuale: (iii) *Diff-in-Diffs*

18

Es.: Disponibili dati panel (T=2, almeno) di H trattati e non-trattati a t=2.  
Variabile risultato: numero medio di riammissioni entro 30 gg.

Assunto di identificazione: “parallelismo”:  $\Delta_{D-P, C(T)} = \Delta_{D-P, NT}$



# Alcune classi di metodi per ricostruire il controfattuale: (iv) *Variabili strumentali (IV)*

19

Il modo più semplice per introdurre il **metodo IV** è far riferimento a un *field experiment* con *non compliance*.

- Stato **assegnato casualmente**:  $Z(1,0)$ .
- Stato **assunto realmente**:  $D(1,0)$ .
- $D$  e  $Z$  sono correlati (relazione testabile) e  $Z$  influenza la variabile risultato solo per il tramite di  $D$  (assunto di identificazione).

La popolazione si ripartisce in quattro gruppi mutuamente esclusivi ed esaustivi, **non osservabili**: **Compliers**, **Always Takers**, **Never Takers**, **Defiers**.

Sotto l'**assunto di assenza di Defiers**, lo **stimatore IV** (con  $Z$  strumento per  $D$ ) fornisce il **LATE**, riferito ai **Compliers**.

# Tornando allo studio di caso iniziale

20

## □ Come valutare l'effetto del *P-C Model* vs. H funzionale?

- La scelta di realizzare il trattamento è **volontaria**. Vi è dunque un processo di (auto)selezione.  
Se confrontiamo i due gruppi, troviamo che gli MDC degli H **trattati** hanno **caratteristiche osservabili**, potenzialmente influenti sulla variabile risultato, distribuite diversamente rispetto agli MDC degli H **non- trattati**:  
Numero di dimissioni, Età media dei pazienti, % di pazienti maschi, Effetti fissi di MDC  $j$  in H  $h$ , Effetti fissi di anno  $t$ ).
- **Controllo con modello di regressione con interazioni del 1<sup>o</sup> ordine fra tutte le v. osservabili ( $\cong$  *statistical matching*).**

# Tornando allo studio di caso iniziale (segue)

21

- Inserendo nella precedente regressione la dummy **PC** (1 se *P-C Model*, 0 altrimenti), l'associato parametro stimato dà la **stima diff-in-diffs**.
- Gli assunti di identificazione: (i) **il processo di selezione** dipende dalle 5 **caratteristiche osservabili**; (ii) **lo stimatore diff-in-diffs** poggia sull'assunto di “**parallelismo**”
- Risultati salienti: il ***P-C Model*** ha effetti significativi su **efficacia**: -0,7% e -0,4% in numero medio di ri-ospedalizzazioni rispettivamente in H di regione e nello stesso H; **efficienza**: -6% in durata del ricovero in H; -7,6% in numero medio di trasferimenti entro l'H.

## 2. Disegni di valutazione, implicazioni testabili e tests di sovra-identificazione

22

- **Validazione dei risultati:** test di sovra-identificazione.

La domanda: **l'assunto di identificazione è vero?**

L'assunzione **non** è saggiabile direttamente, perché coinvolge controfattuali.

È possibile derivarne **implicazioni** che possono essere sottoposte a verifica. Forniscono **condizioni necessarie** per la validità dell'inferenza causale; usabili per falsificarla o corroborarla.

Per farlo servono **informazioni supplementari**, tipicamente servono ulteriori dati (ad es., molteplici gruppi di confronto; serie di risultati pre-intervento).

# Qualche nota sull'analisi degli effetti in studi epidemiologici

23

- Osservazione appena ovvia, forse superflua: **non vi sono indicatori di effetti** che siano direttamente osservabili. Osserviamo *indicatori di risultato*. Gli *indicatori di effetti* discendono da assunti sul processo generatore dei dati (≡ sul processo di selezione) e da conseguenti metodi per ricostruire il controfattuale.
- Spesso un **trattamento consiste in un mix di misure coordinate e coerenti** (vedi lo studio di caso). La possibilità di scomporre il trattamento nelle **singole componenti** – che consideriamo come “trattamenti” – poggia sul fatto che vi sia **variabilità** nell’accesso a tali componenti nella popolazione ammissibile.

# Qualche nota sull'analisi degli effetti in studi epidemiologici (segue)

24

Un assunto cruciale per fare inferenza sugli **effetti di un trattamento** ( $\equiv$  una causa) è la cosiddetta **Stable Unit Treatment Value Assumption (SUTVA)**.

Essa incorpora due condizioni:

- (a) c'è soltanto una forma di trattamento e una forma di non-trattamento per ciascuna delle unità;
- (b) **non ci sono interazioni fra i membri della popolazione-obiettivo**, nel senso che lo stato di trattamento e il risultato del membro  $i$  non è influenzato dallo stato di trattamento né dal risultato di ogni altro membro della popolazione.

**NB.** Questo assunto può essere implausibile in vari interventi epidemiologici!



### 3. La valutazione prospettica

In generale una CIE può essere condotta solo su un intervento che è stato attuato; in tal senso è *ex-post*. Apparentemente sono dunque partito dalla coda. Ma quanto detto ci consegna tre insegnamenti:

- Problema comune a tutti i contesti di valutazione degli effetti è di disporre di **dati adeguati** e **metodi appropriati**, per ricostruire (quindi neutralizzare) il processo di selezione. Più presto ci si attrezza a farlo, meglio è.
- Dal punto di vista dell'**utilizzo**, ogni valutazione degli effetti è *ex-ante*, perché ha la sua ragion d'essere nell'informare il (ri)disegno di politiche che devono ancora essere attuate.
- Dal punto di vista della **generazione dell'informazione**, tutta la valutazione è *ex-post*, compresa quella detta *ex-ante*, nel senso che si basa sull'osservazione di comportamenti passati.

# Perché “valutazione prospettica”?

Più che l'usuale distinzione fra valutazione *ex-ante* ed *ex-post*, torna utile sottolineare la distinzione fra valutazione condotta **retrospettivamente**, dopo che l'intervento è stato attuato per un significativo periodo di tempo, e valutazione condotta **prospettivamente**, in parallelo all'attuazione dell'intervento.

La **valutazione prospettica** prende **avvio contestualmente all'inizio della riflessione sull'introduzione/modifica di una politica**. Essa ha vantaggi cruciali per **il disegno e lo svolgimento della valutazione controfattuale** e, insieme, un **ruolo disciplinatore sullo stesso processo di *policy***.

# I vantaggi della valutazione prospettica

27

- **Funzione disciplinatrice sul processo di *policy*:**
  - induce concretezza e confronto ragionato fra attori;
  - sollecita un “esperimento mentale”, che agevola la definizione chiara e convincente della politica (esempi: *targeting*, condizionalità, non anticipazione di annuncio);
  - usa modelli di microsimulazione, o comunque modelli previsivi, che tipicamente impiegano stime e parametri forniti da precedenti CIE.
- **Vantaggi per il processo di valutazione:**
  - porta a scegliere il metodo di valutazione appropriato;
  - induce a un attento monitoraggio della sua implementazione (tipicamente ad opera di operatori decentrati);
  - favorisce la raccolta tempestiva dei dati necessari.

# Due commenti conclusivi sulla valutazione prospettica

28

- La valutazione prospettica configura un **processo iterativo** fra disegno/microsimulazione/previsione e valutazione degli effetti controfattuale, che si alimentano a vicenda:
  - svolgendo alternativamente, l'una per l'altra, l'usuale ruolo *ex-ante* ed *ex-post*,
  - avendo entrambe un'ottica di utilizzo orientata a fornire informazioni al decisore per migliorare le politiche (il disegno e/o il targeting).
- La scelta di condurre una valutazione prospettica è una cartina di tornasole di **motivazioni di apprendimento in vista di decisioni.**

**GRAZIE PER L'ATTENZIONE**

