

**LA VALUTAZIONE D'IMPATTO DI POLITICHE PUBBLICHE:  
PARADIGMA E BUONE PRATICHE**

**Ugo Trivellato**

*Università di Padova e IRVAPP*

Materiali della relazione invitata al

**Convegno dell'Associazione Italiana di Epidemiologia**

*Roma, 14-15 maggio 2009*

# **OUTLINE**

**1. INTRODUZIONE**

**2. L'ANALISI CONTROFATTUALE**

**3. UNO SCORCIO A ESPERIENZE DI ALTRI PAESI**

**4. IL SOSTRATO E LE CONDIZIONI PER BUONE PRATICHE DI VALUTAZIONE**

# 1. INTRODUZIONE

**POLITICA PUBBLICA** ( $\equiv$  programma, intervento, *programme*):

un **intervento** mirato a una **popolazione determinata** con l'intento di indurre un **cambiamento** in una condizione e/o in un comportamento.

Uno schema con tre dimensioni essenziali: “**Popolazione** – **trattamento** – **risultato**”  
*per chi*                      *di cosa*                      *su cosa*

**VALUTARE GLI EFFETTI DI UNA POLITICA PUBBLICA:**

$\equiv$  Affrontare (e risolvere) un problema cognitivo: **L'ATTRIBUZIONE CAUSALE.**

**NB.** Non tutta l'azione e la spesa pubblica rispondono alla caratteristica di **intervento mirato ad un cambiamento**. Ma **una parte rilevante certamente sì.**



## LA SITUAZIONE IN UN CONTESTO DI OSSERVAZIONE

$D (1,0)$	Risultati osservati e ipotetici per l'individuo $i$	
	Se esposto al trattamento	Se non esposto al trattamento
<b>Trattati</b>	$Y_i^T$	$Y_i^{NT}$
<b>Non-trattati</b>	$Y_i^T$	$Y_i^{NT}$

$Y_i^T$  e  $Y_i^{NT}$ : risultati osservati

$Y_i^T$  e  $Y_i^{NT}$ : risultati ipotetici ( $\equiv$  controfattuali).

**NB.** In generale  $Y_i^{NT} \neq Y_i^{NT}$ , per la presenza di “**distorsione da selezione**”: differenze nelle ‘condizioni di partenza’ e/o nella ‘dinamica spontanea’, cioè rispetto a caratteristiche che (i) rendono **trattati e non-trattati diversi in modo sistematico fin dall’origine** ( $\equiv$  prima del – o comunque a prescindere dal – l’intervento) e (ii) **influiscono sulla variabile-risultato**.

## CONDIZIONI PER “IMPARARE” SUGLI GLI EFFETTI CAUSALI DI UNA POLITICA

Che cosa è ragionevole richiedere per poter “imparare” in merito ad effetti causali?

### 1. **Sulle caratteristiche del trattamento:**

- (a) **il trattamento non deve essere “sfuocato”**, cioè a dire (a1) il trattamento non deve variare tra le unità trattate, e (a2) lo stesso deve valere per l’assenza di trattamento per le unità non trattate;
- (b) **non vi deve essere alcuna interazione ( $\equiv$  ‘contagio’) fra le unità della popolazione-obiettivo**, né rispetto alla (auto)assegnazione o meno al trattamento né rispetto al/i risultato/i. [Assunzione SUTVA]

### 2. **Sulla natura dei soggetti**, trattati e non-trattati: si devono **osservare sia unità trattate che unità non-trattate**. [Condizione del supporto comune]

### 3. **Sulla comparabilità**: trattati e non-trattati debbono essere in tutto e per tutto **equivalenti ( $\equiv$ non differire rispetto a caratteristiche, osservabili e non osservabili, che influenzano sia l’(auto)assegnazione a uno dei due stati sia la variabile risultato)**, **salvo che per la probabilità di essere trattati**. [Condizione *ceteris paribus*]

UN PRIMO ESEMPIO STILIZZATO:

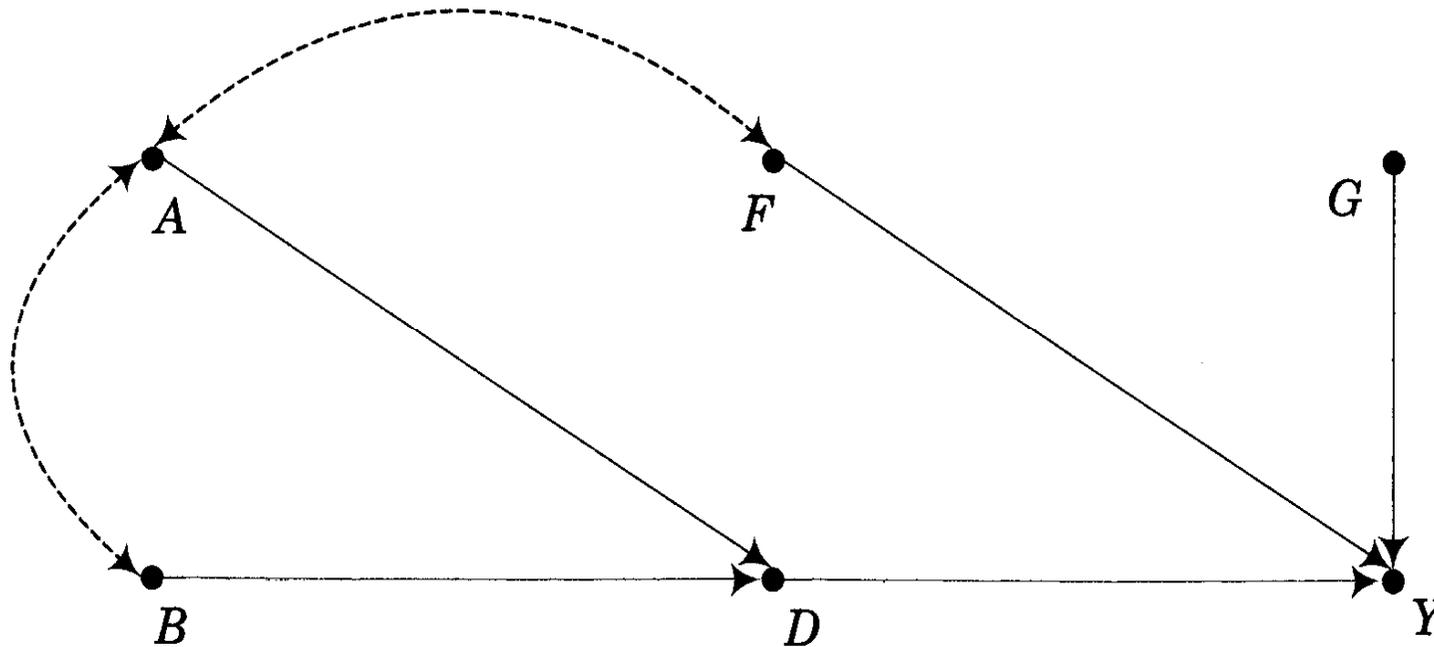
**SERVIZI DI ORIENTAMENTO ALLA RICERCA DI LAVORO**

Effetto			Identificazione degli effetti: strategie di analisi e dati
Di cosa: variabile trattamento	Su cosa: variabile risultato	Per chi: popolazione obiettivo	
Servizio ricevuto	Durata della disoccupazione (dopo il trattamento)	<b>A.</b> Iscritti a liste di Disoccupazione	
	Condizione occupazionale dopo $t$ mesi	<b>B.</b> Esperimento (su un campione casuale di iscritti alle liste)	

UN SECONDO STILIZZATO:  
**L'INTRODUZIONE DELLA PATENTE A PUNTI**

<b>Effetto</b>			<b>Identificazione degli effetti: strategie di analisi e dati</b>
<b>Di cosa:</b>	<b>Su cosa:</b>	<b>Per chi:</b>	
<b>variabile trattamento</b>	<b>variabile risultato</b>	<b>popolazione obiettivo</b>	
<b>Entrata in vigore della normativa (trattamento ≡ ammissibilità)</b>	<b>Decessi per incidenti stradali in un dato intervallo di tempo (ad es., un anno)</b>	<b>Tutti i detentori di patente di guida</b>	<b>Confronto tra periodi senza (≡ immediatamente precedente) e con (≡ immediatamente dopo) l'entrata in vigore della nuova normativa</b>

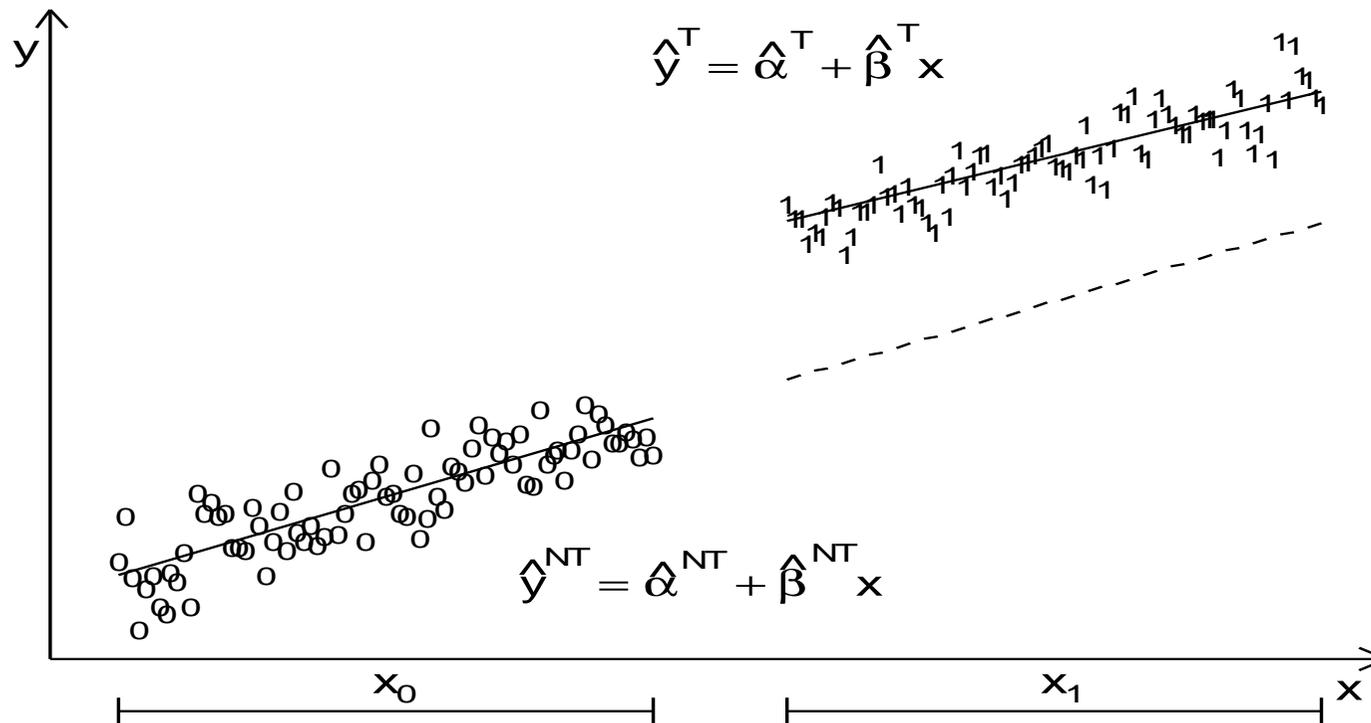
**UNA PRIMA STRATEGIA: specificazione di un **modello completo****



Tramite un **modello completo** si perviene alla  
**stima di tutte le cause dell'effetto totale su Y**, quindi anche alla  
**⇒ stima dell'effetto causale (marginale) di D su Y**, quello di interesse.

**IL RISCHIO DI MODELLI COMPLETI: il ricorso ad assunzioni arbitrarie (sulle variabili incluse, sulla loro misura, sulla forma funzionale, sull'estrapolazione, ecc.).**

**Un esempio: trattare come confrontabili soggetti che in realtà confrontabili non sono**



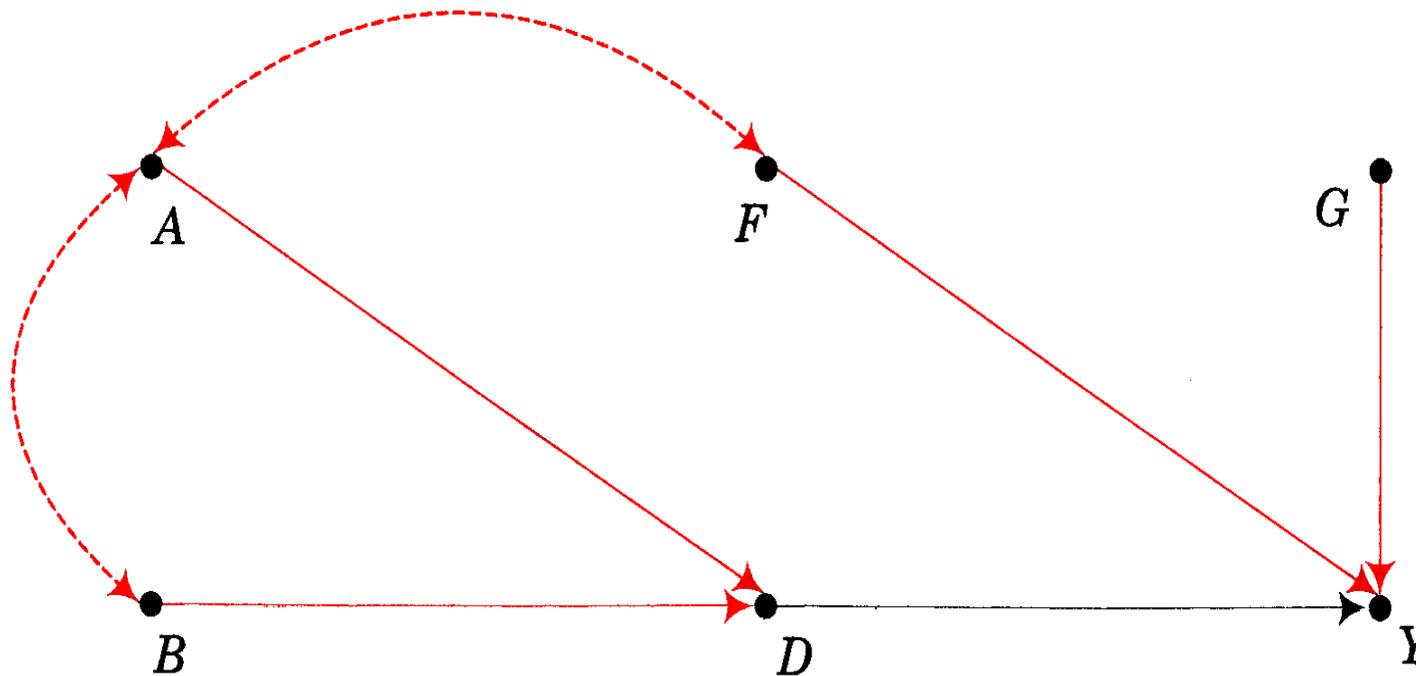
*Stima dell'effetto medio con un modello di regressione lineare: ipotetico risultato per una variabile obiettivo Y, nel caso di insiemi  $X_1$  e  $X_0$  disgiunti per trattati e non-trattati*

UNA SECONDA STRATEGIA: metodi **non parametrici** ( $\cong$  ‘senza modello’) per eliminare la **distorsione da selezione**.

- **Obiettivo 1:** individuare **tutte le variabili responsabili del processo di selezione** e reperire **informazioni affidabili** sulle stesse.
- **Obiettivo 2:** costruire un gruppo di **controllo**, formato da soggetti non-trattati **equivalenti** ai trattati rispetto a tutte le variabili responsabili della selezione:  
 $\Rightarrow$  trattati e controlli risultano **“pareggiati”** rispetto a alle caratteristiche rilevanti per la variabile-risultato, **salvo il fatto che non sono stati esposti all’intervento**.

$D (1,0)$	Risultati osservati e ipotetici per l’individuo $i$	
	Se esposto al trattamento	Se non esposto al trattamento
Trattati	$Y_i^T$	$Y_i^{NT}$
Controlli	$Y_i^T$	$Y_i^C$

Controllata la **distorsione da selezione**, e realizzata così la **condizione *ceteris paribus***, è:  
 $Y_i^{NT} \cong Y_i^C$ .



ANCORA SULLA SECONDA STRATEGIA: individuazione di **tutte le variabili** responsabili del processo di selezione e “**pareggiamento**” rispetto alle stesse:

⇒ **controlli** (≡ non-trattati “pareggiati”) sono **equivalenti** ai **trattati**, quindi

⇒ **stima dell’effetto di una sola causa: D su Y**, il trattamento sulla variabile-risultato.

## (A) L'ESPERIMENTO RANDOMIZZATO COME ..... STANDARD

- Il disegno è tipicamente adottato nei **Randomized Clinical Trials (RCTs)**. Un insieme di pazienti affetti da una patologia viene **suddiviso in due gruppi mediante sorteggio**: a un gruppo – ‘sperimentale’ – viene somministrato il (nuovo) farmaco del quale si vuole saggiare l'efficacia; all'altro – ‘di controllo’ – nulla (o il farmaco già in uso).
  - Se il processo di selezione ha luogo mediante randomizzazione dei soggetti all'uno o all'altro dei due gruppi, **le differenze di partenza risultano nulle per costruzione**. Pertanto, l'effetto medio dell'intervento viene stimato mediante **la differenza tra i risultati medi ottenuti, rispettivamente, dai trattati e dai non-trattati**.
  - Difficoltà nel realizzare l'esperimento ben note anche nei RCTs: vedi disegno “**cieco**” e “**doppio cieco**” (col senno di poi, per far fronte a problemi semplici; le condizioni per “imparare” in merito agli effetti causali potrebbero non essere violate).
  - Difficoltà acute in **esperimenti sociali**, per ragioni etiche e **pratiche**: rischio di *treatment group dropout*, di *control group substitution bias* (già in Fisher, 1926) e, in generale, di **interazioni sociali**. [NB: esperienze diffuse in *demonstrations* USA e in *pilots* in UK.]
- ⇒ L'**esperimento randomizzato** resta il *gold standard*, o meglio il *bronze standard* (Beck, 2001).
- ⇒ **Non c'è soluzione di continuità** fra **RCT**, **esperimento naturale**, metodi per ‘mimarlo’ *ex-post*.

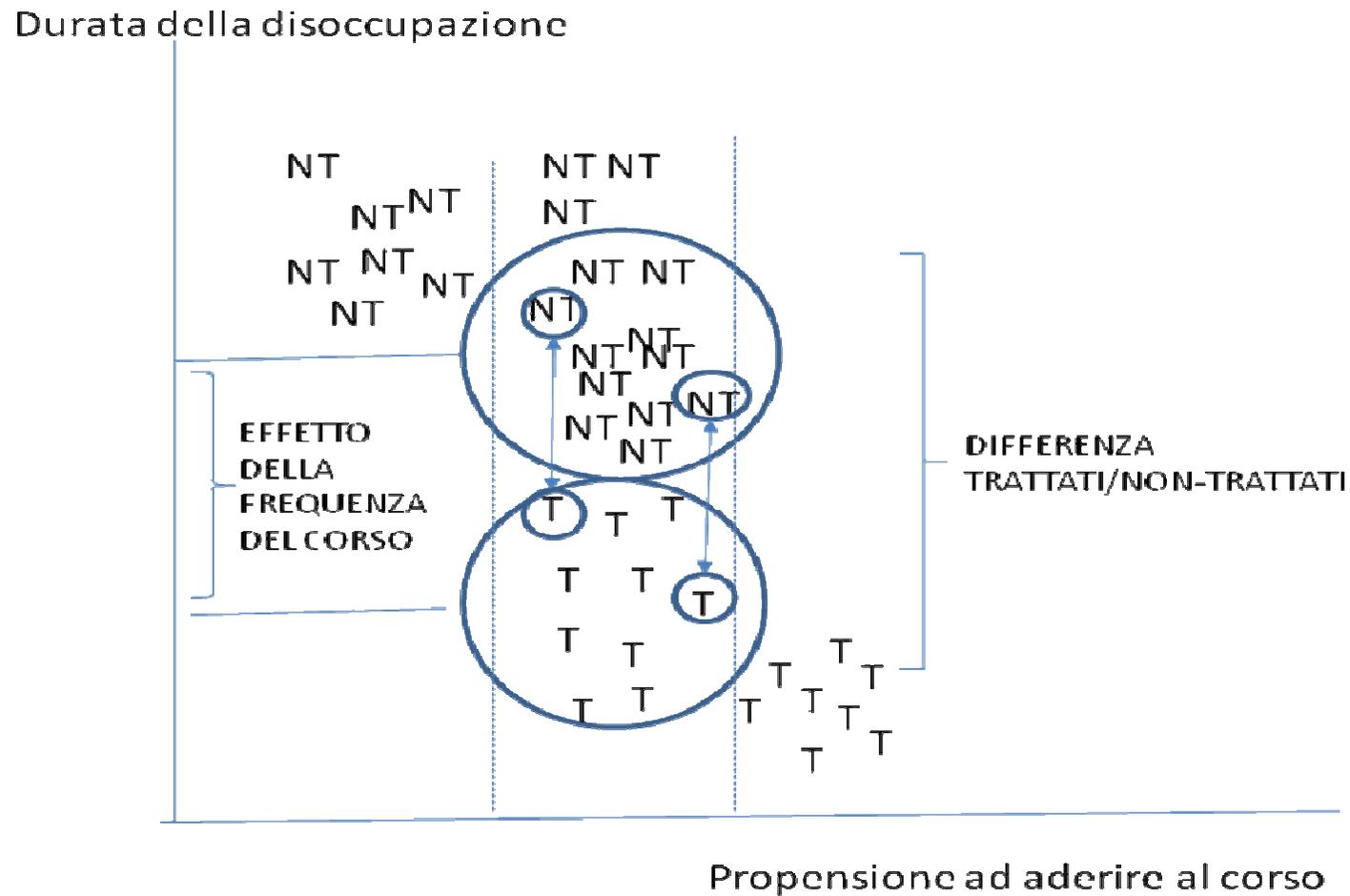
## (B) Ricostruire la situazione controfattuale in contesti osservazionali: L'ABBINAMENTO

- L'idea fondamentale: ricostruire la situazione controfattuale utilizzando l'informazione sui risultati conseguiti dai *non destinatari della politica più simili ai destinatari in termini di caratteristiche osservabili*.
- **L'abbinamento**: si scelgono i non destinatari più simili ai destinatari: fatto ciò, **l'effetto della politica viene ottenuto tramite confronto fra le rispettive medie della variabile-risultato**.

Tale interpretazione è corretta se vale l'**assunto che, condizionandosi alle osservabili, i due gruppi siano stati resi equivalenti**. Quante più sono le caratteristiche osservabili, tanto più tale assunto è credibile. **NB: l'assunto ha implicazioni testabili.**

- La somiglianza tra i due gruppi viene creata utilizzando **un indicatore detto “propensity score”** – nell'esempio che segue è la propensione ad accedere a un programma di orientamento al lavoro –, cioè la probabilità che il soggetto venga trattato in funzione delle sue caratteristiche osservabili. **A ciascun trattato viene abbinato il non-trattato più simile in termini di propensity score**, in generale entro un livello di somiglianza minima pre-definito.

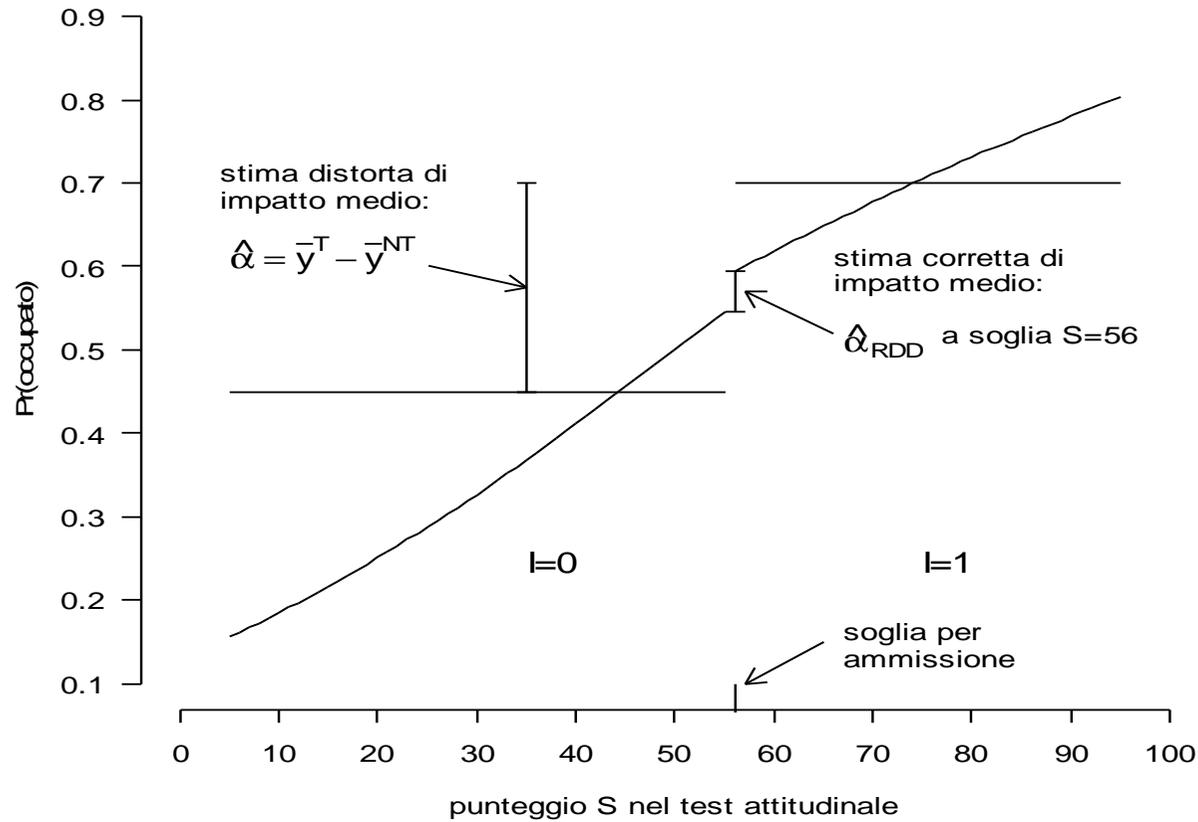
**Figura 1:** *L'abbinamento applicato a un programma di orientamento al lavoro*



**(C1) Ricostruire la situazione controfattuale in contesti osservazionali:  
LA DISCONTINUITÀ RISPETTO A UNA CARATTERISTICA**

- Quando esiste *razionamento* nell'accesso a una prestazione pubblica e l'ammissibilità è determinata dalla *soglia* di una **data caratteristica**, è possibile confrontare i soggetti che stanno da un lato della soglia (es. sopra) e sono ammessi alla prestazione con i soggetti che stanno dall'altro lato della soglia (es. sotto), e non vi sono ammessi.
- Il punteggio che definisce la soglia di ammissibilità all'intervento rappresenta un **'punto di discontinuità'**, e attorno ad esso si determina una situazione simile alla randomizzazione. I **trattati immediatamente sopra la soglia** sono sostanzialmente *equivalenti* ai **non-trattati immediatamente sotto la soglia**. NB. Alta **validità interna** vs. bassa validità esterna.  
**Assunto: *smoothness* in assenza del trattamento: ⇒ con implicazioni testabili.**
- Il **confronto** tra trattati e non-trattati condotto **attorno alla soglia** di ammissibilità al servizio (noto come **RDD: *Regression Discontinuity Design***) rivela **l'effetto dell'intervento**.
- In questo confronto attorno alla soglia l'effetto non è ottenuto confrontando soggetti simili in tutte le caratteristiche osservabili, bensì sfruttando la **differenza 'al margine'** nella **sola variabile che determina l'ammissibilità**. NB: controllo anche rispetto a **non osservabili**.

**Figura 2: La discontinuità applicata a un programma di orientamento al lavoro, con ammissione dei migliori in base alla graduatoria in un test attitudinale. Variabile-risultato: probabilità di essere occupato a 6 mesi dalla conclusione del programma**



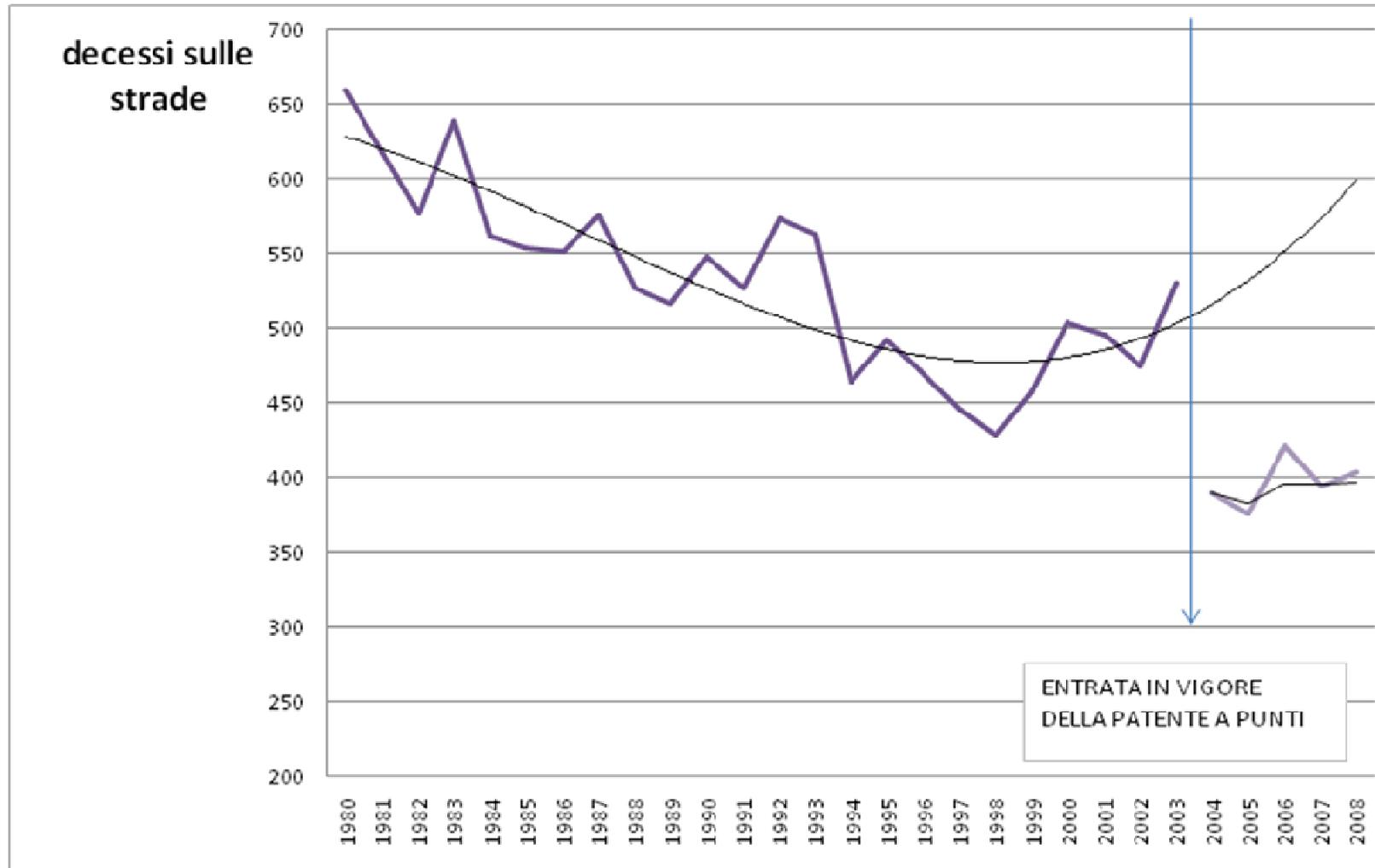
**Riferimento al paper di *Eugenio Paci*, presentato ieri:**

- **screening per patologie oncologiche** su popolazione obiettivo definita da **limiti di età**;
- **RDD** come “naturale” **strategia di identificazione** dell'effetto causale dello *screening*?

**(C2) Ricostruire la situazione controfattuale in contesti osservazionali:  
LA DISCONTINUITÀ RISPETTO AL TEMPO**

- Un caso particolare del metodo RDD si ha quando la discontinuità è *rispetto al tempo*, cioè **quando una politica entra in vigore**.
  - ⇒ È rilevante per **politiche universali** (o per le quali trattamento  $\equiv$  ammissibilità).
  - ⇒ Il metodo serve per controllare una specifica forma di distorsione da selezione: la **distorsione da dinamica spontanea**, ovvero la **fallacia del confronto pre- post-intervento** per stimarne l'effetto.
- A differenza del caso (C1), si considera una serie storica di dati aggregati, e si hanno quindi poche unità di osservazione attorno alla soglia.
  - ⇒ Il confronto è esteso a tutto il periodo disponibile precedente l'intervento. Si utilizza la **storia pre-intervento** della variabile-risultato per **modellarla** e per **predire quale sarebbe stato il suo valore nel periodo immediatamente post-intervento *se l'intervento non fosse stato attuato***.
- **Effetto della politica** = differenza tra l'andamento osservato post-intervento, **fattuale**, e l'andamento predetto, **controfattuale**.

**Figura 3: *La discontinuità rispetto al tempo applicata alla patente a punti***

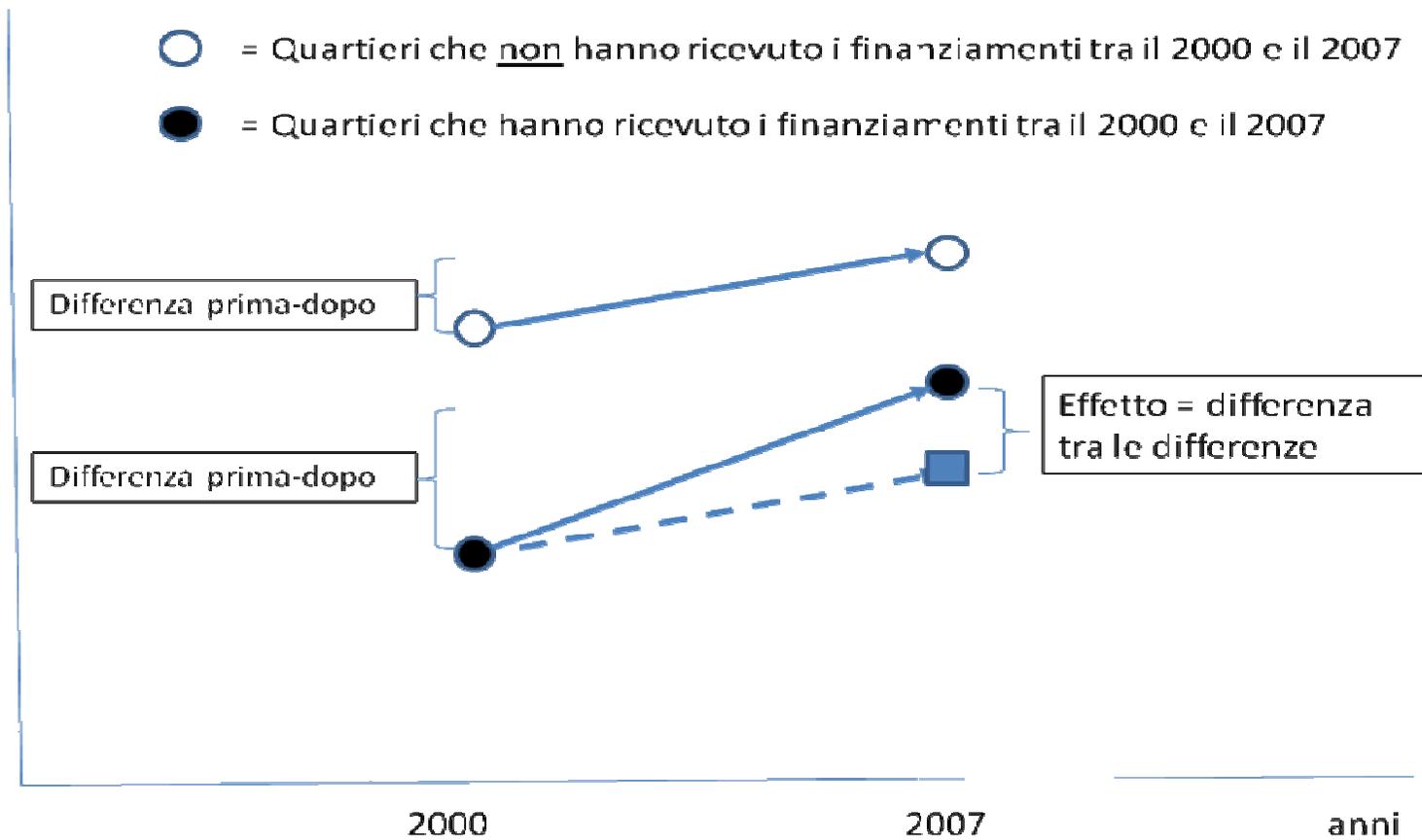


**(D) Ricostruire la situazione controfattuale in contesti osservazionali:  
IL DOPPIO CONFRONTO SPAZIO-TEMPORALE**

- Quando la variabile-risultato è misurabile in **momenti ripetuti nel tempo**, è appropriato il metodo del **doppio confronto**: l'effetto della politica è rivelato confrontando (i) la differenza tra prima e dopo l'attuazione della politica (**1<sup>^</sup> confronto, nel tempo**) (ii) tra trattati e non-trattati (**2<sup>^</sup> confronto, tra soggetti**).
  - **Illustrazione intuitiva del metodo nel caso di una politica di recupero urbano.**
    - \* **I quartieri trattati sono pochi, e selezionati in maniera tutt'altro che casuale.**
    - \* **Il policy-maker, affiancato dall'analista-valutatore, si è preoccupato di:**
      - (i) **osservare anche un campione di quartieri non-trattati** (tra quelli che avevano fatto domanda per accedervi) e **solo grossolanamente simili a quelli trattati**;
      - (ii) per ciascuno dei quartieri, trattati e non-trattati, **raccogliere dati sulla variabile-risultato** – ad es., il numero medio delle PMI insediate per 1.000 ab. – per **almeno un periodo precedente** e per **almeno un periodo successivo** all'intervento.
    - \* **Avendo disponibili questi dati, l'effetto medio dell'intervento di recupero urbano è ottenuto sottraendo al cambiamento osservato nei quartieri trattati il cambiamento osservato nei quartieri non-trattati.**
- Vedi Figura 4.**

**Figura 4: Il doppio confronto applicato alle politiche di recupero urbano**

N. nuove PMI insediate (per 1000 abitanti)



- Questo metodo di stima dell'effetto di una politica (noto come *difference-in-differences*) è valido sotto il seguente **assunto**:

**l'evoluzione della variabile-risultato che avrebbero avuto i trattati se non lo fossero stati è la stessa che hanno avuto i non-trattati: ⇒ con implicazioni testabili.**

Detto altrimenti: possono sì esserci **differenze sistematiche, anche non osservabili, tra trattati e non-trattati nel livello della variabile-risultato** prima dell'intervento, ma queste differenze devono mantenersi **stabili nel tempo**.

- Di conseguenza:

⇒ **la differenza prima-dopo l'attuazione della politica osservata per i non-trattati riproduce** perfettamente quella che sarebbe stata la dinamica per i trattati in assenza di trattamento, cioè il **controtrofattuale**;

⇒ **la differenza fra le differenze prima-dopo rispettivamente di trattati e non-trattati identifica l'effetto causale della politica.**

**(E) Stimare l'effetto del trattamento per i *compliers*:  
IL METODO DELLE VARIABILI STRUMENTALI**

- Caso (semplificato) di **assegnazione randomizzata**, ma di **eterogeneità dell'effetto** e di **non perfetta osservanza dell'assegnazione** (Angrist, Imbens e Rubin, 1996).

La popolazione consta logicamente di **quattro tipi** di soggetti, esaustivi e mutuamente esclusivi, ma **non osservabili**:

- **Compliers (CO)**
- **Always Takers (AT)**
- **Never Takers (NT)**
- **Defiers (DF)**, ovvero “bastian contrari”.

Z (1,0)	D (0,1)	
	D=1	D=0
Z=1	(CO + AT)	(NT + DF)
Z=0	(AT + DF)	(CO + NT)

- **Primo risultato:** in presenza di eterogeneità degli effetti del trattamento e di imperfetta *compliance*, il confronto dei risultati medi di assegnati (Z=1) e non assegnati (Z=0) al trattamento **non identifica un parametro causale!**
- **Secondo risultato:** se **non ci sono soggetti di tipo DF**, l'usuale metodo IV (dove lo **strumento Z** è il risultato di un'estrazione casuale, o assimilabile a tale, è correlato con D, ma non influenza *direttamente* la variabile risultato Y) **identifica l'effetto medio del trattamento per la sotto-popolazione composta dai soggetti di tipo CO.**

$$E\{Y_1 - Y_0 | CO\} = [E\{Y|Z=1\} - E\{Y|Z=0\}] / [E\{D|Z=1\} - E\{D|Z=0\}].$$

- Un esempio classico (Angrist, 1990): soldati statunitensi che hanno servito nella guerra del Vietnam: effetto dell'essere stato **complier** (*veteran status*, essendo stato estratto nella “lotteria” per reclutare i soldati) sui successivi redditi:
  - nato nel 1950: lotteria effettuata nel 1969, servizio inizia nel 1970.
  - nato nel 1951: lotteria effettuata nel 1970, servizio inizia nel 1971.

Vi sono NT fra gli estratti (si sottraggono al servizio) e AT fra i non estratti (volontari).

Netta evidenza di un **effetto causale dell'essere stato estratto** sia *sul veteran status* sia sui **successivi redditi!**

Ma: effetto causale si riferisce alla **sotto-popolazione dei CO, pari soltanto al 15%**.

La validità esterna è piuttosto limitata.

- In pratica, esclusi USA, UK e paesi in via di sviluppo, è raro il caso di esperimenti sociali ( $\equiv$  che comportano un'assegnazione causale allo stato di trattato/non-trattato). Si può cercare di usare **discontinuità temporali in regolazioni, confini amministrativi, ...** per definire **due gruppi di soggetti assegnati a trattamenti alternativi e plausibilmente simili rispetto a ogni altra caratteristica rilevante per la variabile risultato.**
  - ⇒ **Esperimenti naturali: come se si soggetti fossero stati assegnati a caso.**

### 3. UNO SCORCIO A ESPERIENZE DI ALTRI PAESI

- La *What Works Clearing House* sull'istruzione negli USA [+ **Campbell Collaboration**]
- Le “**riforme Hartz**” del mercato del lavoro e della previdenza sociale in Germania
- Una recente **innovazione nelle politiche di avviamento al lavoro** in Francia
- “*No One Written off*”: il disegno di riforma del *welfare* inglese.

### 3.1. LA WHAT WORKS CLEARING HOUSE NELL'ISTRUZIONE

Welcome to **WWC**

Established in 2002 by U.S. Department of Education's Institute of Education Sciences, the **What Works Clearinghouse (WWC)** is a central and trusted source of scientific evidence for **what works in education**. The WWC:

- Produces user-friendly practice guides for educators that address instructional challenges with research-based recommendations for schools and classrooms;
- **Assesses the rigor of research evidence on the effectiveness of interventions (programs, products, practices, and policies)**, giving educators the tools to make informed decisions;
- **Develops and implements standards for reviewing and synthesizing education research;** and
- Provides a public and easily accessible registry of education evaluation researchers to assist schools, school districts, and program developers with designing and carrying out rigorous evaluations.

# ***WHAT WORKS CLEARING HOUSE***

[segue]

## TOPICS



**Beginning Reading**



**Elementary School Math**



**English Language Learners**



**Middle School Math**



**Early Childhood Education**



**Dropout Prevention**



**Character Education**

## 3.2. LE “RIFORME HARTZ” TEDESCHE

Una decina d’anni fa lo stato dell’arte sulla valutazione degli effetti di politiche pubbliche in Germania non era molto migliore che in Italia.

⇒ Tra 2002 e 2005 “**riforme Hartz I-IV**” del **sistema di protezione sociale** e delle **politiche del lavoro**:

- Per tutti i principali programmi di riforma: **obbligatoria la valutazione della implementazione e degli effetti.**
- **Miglioramento delle pre-condizioni:**
  - **sistema informativo**, basato su **archivi amministrativi**;
  - **accesso ai microdati** da parte dei ricercatori secondo regole liberali.
- **Diffusa attività di valutazione** da parte di **istituti di ricerca** (IZA, ZEW, DIW, IAB, ...) e di **team di ricercatori universitari**: ⇒ **netto miglioramento della qualità.**
- **Attenzione ai risultati delle valutazioni** nei **media** e nel **dibattito pubblico.**

### 3.3. UNA RECENTE INNOVAZIONE IN POLITICHE DI AVVIAMENTO AL LAVORO IN FRANCIA

**Premessa:**

*International Workshop on Impact Evaluation Practice and Prospects* (CSS, aprile 2007).

**Mediocre interesse** per studi di valutazione di effetti di politiche pubbliche in Francia, e loro **scarso sviluppo** (eccezione: CREST-INSEE, con taglio prettamente accademico).

⇒ **Nell'arco di un anno e mezzo:**

- **Maggiore interazione fra Ministero del Lavoro e Ministero dell'Economia.**
- **Decollo di Paris School of Economics, con:**
  - **Master “Politiche pubbliche e sviluppo”;**
  - **Partenariato con J-PAL** (J-Poverty Action Laboratory, MIT), e sede di **J-PAL Europe.**
- **Avvio di esperimento sociale su larga scala** (200.000 soggetti): **valutazione dell'effetto di tre diverse modalità di avviamento al lavoro per disoccupati di lunga durata** (Behagel *et al.*, 2008).

### 3.4. “NO ONE WRITTEN OFF”: VERSO LA RIFORMA DEL WELFARE IN UK

Alle spalle: esperienza decennale del *New Deal*, i cui programmi sono stato oggetti di sistematica valutazione degli effetti.

“NO ONE WRITTEN OFF”: disegno di un nuovo programma di riforma del *welfare*.

- Preliminare: *Background research and discussion paper*:
  - (i) rivisitazione del *New Deal* (e delle sue valutazioni);
  - (ii) proposta della nuova filosofia, incentrata sulla “condizionalità”: *More support, higher expectations*.
- *Green Paper*: “*No one written off: Reforming welfare to reward responsibility*”, per consultazione pubblica.
- Consultazione sul *Green Paper*: in maniera strutturata e pubblica, sulla base di una lista di domande che focalizza la discussione fra attori sociali, media e cittadini.
- *Green Paper* è accompagnato dall’*Impact Assessment* (AIR).

## 4. SOSTRATO E CONDIZIONI PER BUONE PRATICHE DI VALUTAZIONE

### 4.1. Il SOSTRATO CULTURALE E ISTITUZIONALE:

- Il processo di (ri)definizione delle politiche è *evidence-based*.
- Il contesto istituzionale si ispira al “*modello anglo-americano di ‘accountability’ democratica*”:
  - controllo dell’esecutivo e valutazione degli effetti di politiche pubbliche sono svolte da *agenzie indipendenti (“watching dogs”)*, con elevata specializzazione, notevoli professionalità e reputazione di autonomia, che *operano per conto della/e assemblea/e legislativa/e*.
  - NB: Decisamente meno efficace è il cosiddetto “*modello continentale*”, che affida la funzione di controllo a magistrature indipendenti: le Corti dei Conti.
- Dibattito pubblico approfondito e *focalizzato sulle policy issues*: bene alimentato da *università, think tanks, media* e associazioni professionali e di interessi.

**4.2. Le “CLAUSOLE VALUTATIVE”:** stabilire nella legge istitutiva di una politica:

- gli **obiettivi** con riferimento ai quali avrà luogo la valutazione;
- le indicazioni circa il **modo** con cui sarà condotta;
- lo **stanziamento** per lo svolgimento della valutazione;
- il termine per il **riesame** della politica da parte del *policy maker* [talvolta una vera e propria *sunset clause*].

### 4.3. CINQUE SUGGERIMENTI DALLE ESPERIENZE STRANIERE MATURE:

- (I) Quando occorre iniziare la valutazione degli effetti di una politica? **Prima possibile, contestualmente all'avvio del disegno della politica.**  
Poco utile la contrapposizione valutazione ex-ante vs. ex-post.  
Rilevante la contrapposizione **valutazione prospettica vs. retrospettiva.**
- (II) Serve **combinazione** di *analisi di implementazione* e di *valutazione degli effetti*, condotte in modo rigoroso e coordinato, ma senza confusione di ruoli.
- (III) È essenziale la disponibilità di un'**adeguata base informativa.**
- (IV) È opportuno mettere inizialmente in atto **progetti-pilota a scala ridotta**, con lo specifico scopo di “imparare” sulle modalità di funzionamento dell'intervento e i suoi effetti.
- (V) **Accessibilità delle informazioni** a una pluralità di analisti, condizione imprescindibile per la necessaria **replicazione degli studi.**

## **A MO' DI CONCLUSIONE, DUE TAKEAWAY MESSAGES**

*“It is important to compare comparable people. Many non-experimental evaluations identify the parameter of interest by comparing observationally different persons using extrapolations based on inappropriate functional forms imposed to make incomparable people comparable. A major advantage of non-parametric methods [≡potential outcome framework ≡ counterfactual analysis] for solving the problem of selection is that, rigorously applied, they force the analyst to compare only comparable people.”*

(Heckman, LaLonde and Smith, 1999, p. 1869)

*“Evaluations build on cumulative knowledge.”*

(Heckman e Smith, 1995, p. 93)

**NB. Rispetto alle esperienze in campo clinico ed epidemiologico, le valutazioni degli effetti di politiche economiche e sociali registrano un ritardo nella meta-analisi.**