

LA VALUTAZIONE DEI PROGRAMMI PER IL CONTRASTO ALLA POVERTÀ EDUCATIVA

1. INTRODUZIONE
2. STRATEGIE
3. RIFERIMENTI

Premessa

Queste brevi note intendono descrivere che cosa si intende per *valutazione* e che tipo di *problemi* si incontrano nell'effettuarla. Esse sono rivolte a lettori che non abbiano familiarità con tali questioni, e hanno lo scopo di richiamare la loro attenzione sulla necessità di impostare il disegno di un intervento anticipando i requisiti di informazione e le analisi che consentiranno di valutarlo. Non sono invece un sostituto per una conoscenza approfondita dei problemi e delle tecniche di valutazione, per la quale si rimanda ai manuali citati nei riferimenti o alla collaborazione con esperti di valutazione.

1. INTRODUZIONE

L'azione pubblica spesso consiste in interventi che modificano le condizioni o gli incentivi di un certo gruppo di persone: possono essere programmi di formazione, trattamenti sanitari, agevolazioni fiscali, erogazione di servizi, concessione di permessi... Prendendo a prestito dalla terminologia medica, interventi di questo tipo vengono spesso denominati *trattamenti*, e sono chiamati *trattati* le persone oggetto dell'intervento, mentre sono detti *non trattati* o *controlli* coloro che invece sono esclusi dall'intervento.

Il trattamento è effettuato con l'intenzione di indurre una qualche modifica nel comportamento dei soggetti trattati. Come possiamo sapere che alla (buona) intenzione corrispondono gli effetti desiderati?

Spesso i politici e gli amministratori, a dimostrazione dello sforzo fatto, citano misure descrittive delle caratteristiche dell'intervento (soldi stanziati, numero di bandi emessi o autorizzazioni emanate, numero di soggetti effettivamente coinvolti o potenzialmente coinvolgibili); queste misure, per quanto informative e utili, non dicono nulla degli effetti ottenuti.

Più raramente, e meritoriamente, vengono richiamati i cambiamenti (aumenti, riduzioni) nelle variabili che l'intervento intendeva modificare; queste variazioni vengono però automaticamente attribuite all'esistenza stessa dell'intervento, interpretandole come prova della sua efficacia.

Ma se il monitoraggio dell'intervento effettuato e la misurazione dei risultati sono certamente indispensabili, e sono un importante contributo alla trasparenza dell'azione pubblica nei confronti dei contribuenti e dei cittadini, essi non sono strumenti sufficienti per stabilire se l'intervento ha avuto successo oppure no.

Infatti, l'osservazione di una modifica, prima e dopo l'intervento, in alcuni dei risultati per i soggetti trattati potrebbe derivare dall'operare di fattori indipendenti, che hanno indotto variazioni che si sarebbero comunque verificate, anche in assenza dell'intervento (si parla di fattori di *confondimento*). Un problema ancora più pervasivo è dato dal fatto che l'osservazione di una differenza tra i risultati per i soggetti trattati e quelli per i soggetti non trattati potrebbe dipendere dal modo con cui si sono scelti i soggetti raggiunti dal trattamento.

Ciò che vogliamo, per poter decidere dell'efficacia di un intervento, è stabilire se esso sia stato effettivamente la *causa* della modifica osservata nei comportamenti.

Mettiamo enfasi sul termine *causa* per richiamare l'attenzione sul fatto che non necessariamente una differenza tra il comportamento dei soggetti trattati prima e dopo il trattamento o tra quello dei soggetti trattati e di quelli non trattati è attribuibile al trattamento; è soltanto quando

possiamo ragionevolmente ritenere che il trattamento abbia causato la modifica che potremo ritenerlo efficace, ed eventualmente estenderne l'applicazione ad altri e più ampi gruppi di soggetti.

Per avere un'idea della difficoltà di stabilire l'effetto causale di un trattamento è utile riflettere su un esempio apparentemente banale e ovvio. Supponiamo di confrontare il tasso di mortalità di coloro che lo scorso anno sono stati ricoverati in ospedale (il trattamento) e coloro che non lo sono stati. Osserviamo che la mortalità dei primi è notevolmente superiore. Ne potremmo concludere che è l'ospedale a causare la morte dei pazienti? L'intuizione ci porta ovviamente a negare questa conclusione, ma cerchiamo di capire meglio perché la conclusione sarebbe infondata. Il problema sta nel fatto che la differenza tra il tasso di mortalità di coloro che sono stati in ospedale e coloro che non ci sono stati assomma e confonde due aspetti: la differenza, per coloro che sono andati in ospedale, tra il tasso effettivo di mortalità e quello che avrebbero avuto se non ci fossero andati—questo, e solo questo, è l'effetto causale dell'ospedale, poiché confronta gli stessi soggetti, con e senza il trattamento—e la differenza tra la mortalità di coloro che sono andati in ospedale, se non ci fossero andati, e la mortalità di coloro che non ci sono andati—e questa differenza non è in genere zero, perché va in ospedale chi già sta male!

Per concretezza, supponiamo che il tasso di mortalità di coloro che sono andati in ospedale sia del 20%, e che quello di coloro che non ci sono andati sia del 5%; supponiamo anche che la mortalità dei primi, se non ci fossero andati, sarebbe stata del 40%. Abbiamo quindi che la differenza tra 20% e 5%, cioè 15%, può essere scomposta nella somma di due differenze: la differenza tra 20% e 40%, cioè *meno* 20%, e la differenza tra 40% e 5%, cioè +35% (e infatti, $-20\% + 35\% = +15\%$). La prima differenza, -20%, cioè una *riduzione* di 20 punti percentuali del tasso di mortalità, è l'effetto *causale* dell'essere andati in ospedale: andare in ospedale riduce il tasso di mortalità. La seconda, +35%, è l'effetto della *selezione*: sceglie di andare in ospedale chi è già malato, ed è quindi naturale che la sua mortalità, se non ci fosse andato, sarebbe stata superiore a quella di chi ha scelto di non andarci perché non è malato. La conclusione che trarremmo dal confronto tra 20% e 5% sarebbe quindi del tutto fuorviante: ci porterebbe a ritenere che l'effetto dell'ospedale è di aumentare il tasso di mortalità di 15 punti percentuali, mentre in realtà è quello di ridurla di 20!

Si noti però che tanto il 20% quanto il 5% sono numeri facilmente osservabili: il primo è la mortalità media di coloro che sono andati in ospedale, il secondo la mortalità media di coloro che non ci sono andati. Il 40%, e cioè la mortalità media di coloro che sono andati in ospedale *se non ci fossero andati*, non è invece osservabile: poiché quelle persone sono andate in ospedale, non sapremo mai quale sarebbe stata la loro mortalità nel caso che non ci fossero andati. Siamo di fronte a quello che si chiama un *controfattuale*: un'osservazione che potremmo fare solo se potessimo tornare indietro nel tempo e far prendere ai soggetti in questione una decisione diversa.

La difficoltà nel misurare l'effetto causale di un trattamento sta proprio nella presenza di un elemento controfattuale nella misura, e vedremo nel seguito varie strategie per aggirare la fondamentale non osservabilità posta dal controfattuale.

Prima di passare a tali strategie, esaminiamo di nuovo il problema all'interno di un esempio meno ovvio e direttamente rilevante per il tipo di interventi oggetto del bando.

Esempio

Si consideri l'offerta di attività extra-curricolari a bambini in condizione di povertà di un determinato quartiere, volta a determinare un miglioramento del loro rendimento scolastico. La partecipazione al programma è stabilita su base volontaria: tutti i genitori che lo vogliono possono iscrivere i propri figli alle attività extra-curricolari.

Supponiamo ora di voler misurare l'effetto causale del programma, confrontando il rendimento scolastico a fine anno dei bambini che hanno partecipato al programma con quello dei bambini

che non hanno partecipato. Sarebbe questa una misura adeguata?

Probabilmente no. È infatti probabile che le famiglie che hanno deciso di iscrivere i propri figli al programma siano relativamente più istruite, più abbienti, o comunque più attente alle esigenze dei propri figli rispetto alle famiglie che non si sono preoccupate di iscrivere i propri figli al programma. In conseguenza di ciò, il rendimento scolastico dei bambini nel programma (i trattati) sarà verosimilmente migliore di quello dei bambini non nel programma (i non trattati), anche indipendentemente dal fatto che i primi hanno partecipato al programma: il rendimento scolastico dei primi sarebbe stato comunque migliore, quindi nel fare la differenza tra i rendimenti scolastici dei due gruppi di bambini si sta attribuendo al programma anche una differenza di rendimento che ci sarebbe stata comunque. Si sta cioè sovrastimando l'effetto del programma. In linea di principio, potrebbe anche essere che il programma ha un effetto causale negativo sul rendimento scolastico, mentre la differenza tra il rendimento di trattati e non trattati è positiva!

L'approccio controfattuale

Come già detto, per misurare l'effetto causale del trattamento è necessario affrontare il problema del controfattuale: l'effetto di un intervento su un individuo è la differenza tra il risultato osservato in presenza del trattamento e il risultato che si sarebbe osservato, *per lo stesso individuo*, in sua assenza (quest'ultimo è controfattuale, poiché se quell'individuo è stato trattato non avremo modo di osservare il risultato se non lo fosse stato). Per ricostruire in maniera credibile il risultato controfattuale servono dati e metodi appropriati, *che spesso devono essere raccolti sin da una fase precedente l'avvio dell'intervento stesso*. Questo fa sì che occorra *progettare il disegno di valutazione e di raccolta dei dati necessari a questa di pari passo con il disegno degli interventi*. L'obiettivo è fornire strumenti conoscitivi per imparare dall'esperienza e migliorare, in una fase successiva, il disegno degli interventi, dopo aver capito se questi hanno funzionato e nei confronti di chi. La valutazione effettuata con l'approccio controfattuale ha finalità di apprendimento sugli effetti dell'intervento, e non costituisce un giudizio su chi ha promosso o realizzato l'intervento stesso.

È utile introdurre una semplice formalizzazione della situazione descritta. Indichiamo con Y il risultato a cui siamo interessati: lo stato di salute, il rendimento scolastico, lo stipendio guadagnato, il numero di libri letti... Vogliamo una notazione sintetica per indicare il risultato potenziale per un individuo che sia stato oggetto di un intervento (sia stato trattato), nel caso controfattuale in cui non lo fosse stato, e viceversa il risultato potenziale per un individuo che non sia stato trattato, nel caso controfattuale che lo fosse stato. Se conveniamo di indicare con il suffisso T l'essere stati potenzialmente trattati, con il suffisso NT l'essere stati potenzialmente non trattati, con il suffisso 1 l'essere stati effettivamente trattati, con il suffisso 0 il non esserlo stati, avremo che Y^{NT1} indica il risultato potenziale di un individuo effettivamente trattato, *se invece non fosse stato trattato*, e Y^{T0} indica il risultato potenziale di un individuo che non è stato trattato, *se invece fosse stato trattato*. La tabella seguente riporta i quattro casi possibili:

	Risultati potenziali per un individuo	
	Se trattato (T)	Se non trattato (NT)
Trattato (1)	Y^{T1}	Y^{NT1}
Non trattato (0)	Y^{T0}	Y^{NT0}

Gli unici dati osservabili sono Y^{T1} (il risultato per un individuo trattato) e Y^{NT0} (il risultato per un individuo non trattato). L'effetto causale del trattamento (su un individuo effettivamente trattato), è dato dalla differenza $Y^{T1} - Y^{NT1}$ (cioè, la differenza di risultato per lo stesso identico individuo nel caso in cui sia trattato e non); ma questa differenza non è osservabile, visto che in

un certo punto del tempo, all'interno di un dato programma, un singolo individuo o è trattato o non lo è.

Possiamo essere tentati di approssimare quella differenza con la differenza tra il risultato di un individuo trattato e il risultato di un individuo non trattato, $Y^{T1} - Y^{NT0}$. Ma, come si è visto nell'esempio sulla mortalità dei ricoverati in ospedale, in genere Y^{NT1} è diverso da Y^{NT0} , perché in genere le persone a cui viene somministrato il trattamento (le persone con il suffisso 1) sono diverse dalle persone a cui non viene somministrato (le persone con il suffisso 0), a prescindere dall'effetto del trattamento: dopotutto, l'intervento è indirizzato a certe categorie di persone e non ad altre, a persone che ne hanno più bisogno o che sono meglio in grado di apprezzarne i benefici o che si trovano in qualche situazione particolare. E, in generale, queste differenze nelle condizioni di partenza hanno, direttamente o indirettamente, un'influenza sul risultato, indipendentemente dal trattamento: nell'esempio, le condizioni di partenza dei trattati danno luogo a una mortalità maggiore rispetto ai non trattati.

Per queste ragioni, la differenza effettivamente osservata ($Y^{T1} - Y^{NT0}$) potrebbe essere ben diversa da quella che si sarebbe voluta osservare ($Y^{T1} - Y^{NT1}$), addirittura di segno opposto, come è nell'esempio dei ricoverati dove sembra che ricevere cure aumenti la mortalità, rendendo in tal caso impossibile valutare l'efficacia dell'intervento.

In definitiva, affinché attraverso il disegno della valutazione si possa identificare un effetto causale è importante che gli individui trattati e quelli non trattati *siano il più possibile comparabili*, salvo il fatto che i primi sono trattati e i secondi non lo sono; in particolare, bisogna preoccuparsi che non ci siano caratteristiche rispetto alle quali trattati e controlli differiscano e che influenzino, direttamente o indirettamente, il risultato che si vuole ottenere con il trattamento.

2. STRATEGIE

Ci sono vari modi per ottenere la comparabilità indicata in precedenza come il requisito essenziale per identificare l'effetto causale di un trattamento. Vediamo i principali.

2.1 La randomizzazione

Come abbiamo detto, il motivo principale per cui il gruppo dei trattati è diverso da quello dei non trattati deriva dal fatto che il trattamento è solitamente somministrato in risposta a qualche condizione che differenzia i due gruppi. Il modo più diretto di aggirare questo problema è quello di somministrare il trattamento in modo casuale (naturalmente, una volta che si sia appurata la sua efficacia, lo si potrà estendere all'intera popolazione). Se il processo di selezione ha luogo mediante assegnazione casuale dei soggetti all'uno o all'altro dei due gruppi, non c'è motivo di ritenere che, salvo ovviamente che per il trattamento, essi (in media) differiscano. Questo di fatto è proprio il significato di una scelta casuale dei due gruppi: sia l'uno sia l'altro sono, prima del trattamento, campioni rappresentativi della stessa popolazione, e non presentano differenze di partenza significative. Quindi, la media di Y^{NT1} è uguale alla media di Y^{NT0} , e l'effetto dell'intervento può essere stimato mediante la differenza tra i risultati ottenuti, rispettivamente, dai trattati e dai non trattati ($Y^{T1} - Y^{NT0}$).

La randomizzazione del trattamento è il disegno tipicamente adottato negli studi clinici: un insieme di pazienti affetti da una patologia viene suddiviso in due gruppi mediante sorteggio (effettuato tramite software appositi): a un gruppo (trattati) viene somministrato il farmaco del quale si vuole testare l'efficacia, all'altro (controllo) nulla, o il farmaco già in uso. Questo disegno sperimentale ormai rappresenta la *best practice* per una corretta valutazione anche nell'ambito delle scienze sociali.

Esempio

Nel caso dell'esempio riportato nell'Introduzione, sarà necessario eliminare la volontarietà nella partecipazione all'intervento: tra tutte le famiglie che si dichiarano disposte a iscrivere i propri figli al programma, si estrae casualmente un sottoinsieme e solo i figli di quelle estratte possono partecipare. In questo caso, un confronto tra il rendimento scolastico dei bambini che sono stati selezionati e di quelli che non sono stati selezionati (ma che nelle intenzioni delle famiglie avrebbero dovuto farlo) identificherà l'effetto causale dell'intervento.

Problemi e soluzioni

La randomizzazione della partecipazione pone talvolta problemi di natura etica. Nel caso di studi clinici la somministrazione di un *placebo* (un farmaco dagli effetti nulli) ad alcuni pazienti potrebbe essere eticamente insostenibile, vista la difficoltà di escludere questi pazienti da un trattamento che potrebbe rivelarsi efficace. Difficoltà analoghe si pongono spesso negli esperimenti sociali; a queste si aggiungono difficoltà pratiche, legate a una corretta individuazione delle popolazioni di riferimento. Nonostante queste difficoltà, l'esperimento randomizzato incarna il prototipo della valutazione, poiché affronta nel modo più diretto il problema del controfattuale. E d'altra parte, l'identificazione degli effetti causali di una politica dovrebbe essere alla base delle decisioni che ne dispongono il proseguimento, l'estensione, o l'eventuale interruzione.

Un modo per attenuare il problema etico posto dalla randomizzazione è quello di procedere con esperimenti ripetuti nel tempo: coloro che non sono stati inizialmente inclusi tra i trattati avranno l'opportunità di esserlo in un secondo momento. Questo approccio ha due vantaggi. Da un lato garantisce che tutti gli individui nella popolazione siano prima o poi trattati, permettendo tuttavia di utilizzare gli individui ancora non trattati come campione di controllo per gli individui trattati nelle prime fasi. Dall'altro permette di correggere l'intervento laddove si siano delineati dei problemi nel corso delle prime sperimentazioni.

2.2 La regressione alla discontinuità

Spesso, per via di risorse limitate, non tutti i soggetti potenzialmente interessati a un trattamento possono essere effettivamente inclusi nel programma di intervento. In tali casi è comune fissare una soglia nel livello di una variabile che caratterizza i vari individui (per esempio il reddito, o lo stato di salute, o il rendimento scolastico precedente), al di sopra o al di sotto della quale l'individuo è incluso o escluso dal trattamento. È allora possibile confrontare i soggetti che stanno poco sopra la soglia e sono trattati con i soggetti che stanno poco sotto la soglia e non sono trattati. Trattandosi di differenze molto piccole nella variabile che determina l'assegnazione o meno al trattamento, è ragionevole ritenere che le differenze nelle altre caratteristiche siano ugualmente modeste; in altri termini, i soggetti non trattati appena sotto la soglia sono confrontabili con i soggetti trattati, appena sopra la soglia, a meno dell'effetto del trattamento.

In termini tecnici il punteggio R che definisce la soglia di ammissibilità all'intervento rappresenta un punto di discontinuità, e attorno a esso si determina una situazione simile alla randomizzazione. I trattati immediatamente poco sopra la soglia (che indichiamo con R^+) sono sostanzialmente equivalenti ai controlli immediatamente sotto la soglia (R^-), cioè $Y^{NT1}_{R^+}$ è uguale a $Y^{NT0}_{R^-}$. Il confronto tra trattati e controlli condotto attorno alla soglia di ammissibilità all'intervento ($Y^{T1}_{R^+} - Y^{NT0}_{R^-}$), noto come *Regression Discontinuity Design*, identifica appunto

l'effetto dell'intervento.

Esempio

Nel caso dell'esempio riportato nell'Introduzione, si potrebbe procedere a individuare il livello di una caratteristica (o la combinazione di più caratteristiche), come per esempio il reddito, al di sotto della quale è possibile partecipare all'intervento. In questo caso, un confronto tra il rendimento scolastico dei bambini il cui reddito era di poco sotto e di poco sopra la soglia prestabilita identificherebbe l'effetto dell'intervento.

Problemi e soluzioni

Trattati e controlli sono tanto più simili quanto più ci si restringe a un intervallo molto ravvicinato intorno alla soglia. Sarà perciò importante assicurarsi che la numerosità campionaria intorno alla soglia sia elevata (per esempio, tramite l'individuazione di una soglia a un livello altamente frequente nella popolazione), così da permettere un confronto statistico significativo.

Qualora gli individui siano in grado di manipolare la caratteristica che determina l'ammissibilità, è altresì importante non comunicare anticipatamente il valore della soglia: se così fosse, trovarsi sopra o sotto la soglia potrebbe essere il risultato di una scelta da parte degli individui, e questo potrebbe essere correlato ad altre loro caratteristiche che hanno un'influenza sul risultato che si vuole ottenere. Verrebbe così alterata la comparabilità intorno alla soglia. Naturalmente, questo precetto può scontrarsi con ragioni di fattibilità dell'intervento.

2.3 L'analisi longitudinale

Quando il risultato è misurabile in momenti ripetuti nel tempo (di cui almeno uno precedente all'implementazione del progetto), è possibile procedere al metodo della doppia differenza (nella terminologia anglosassone si parla di *Difference-in-Differences*), anche in assenza di assegnazione randomizzata del trattamento.

Indichiamo con P un momento nel tempo prima dell'intervento e con D un momento nel tempo successivo. Consideriamo poi la differenza nel risultato per gli individui trattati, prima e dopo l'intervento (differenza nel tempo), $Y^T_D - Y^T_P$. Questa differenza elimina, per così dire, ogni eventuale diversità nelle condizioni di partenza dei trattati, trattandosi dello stesso identico individuo osservato prima e dopo l'intervento. Essa potrebbe tuttavia essere determinata da effetti temporali che nulla hanno a che vedere con l'efficacia dell'intervento, per la presenza di qualche fenomeno che evolve nel tempo e influenza tutti gli individui, trattati e non trattati. Per questo motivo è necessario sottrarre da questa prima differenza anche la differenza nel tempo per il gruppo di controllo $Y^{NT}_D - Y^{NT}_P$. Questa rappresenta il controfattuale dell'effetto temporale in assenza di trattamento. L'effetto causale del trattamento può quindi essere misurato dalla doppia differenza $(Y^T_D - Y^T_P) - (Y^{NT}_D - Y^{NT}_P)$.

Esempio

Nel caso dell'esempio riportato nell'Introduzione, sarà possibile lasciare la volontarietà nella partecipazione all'intervento. Tuttavia, sarà necessario raccogliere le informazioni sui rendimenti scolastici dei bambini nei periodi precedenti l'inizio dell'intervento.

Problemi e soluzioni

Affinché questo metodo di stima dell'effetto causale di un intervento sia valido deve essere vero (almeno approssimativamente) che la variazione nel tempo del risultato che si sarebbe osservata per i trattati in assenza di trattamento è la stessa che si è osservata per i controlli. Detto altrimenti: possono esserci differenze sistematiche, anche non osservabili, tra trattati e controlli nel *livello* del risultato prima dell'intervento, ma queste differenze devono mantenersi *stabili* nel tempo. Di conseguenza, la differenza prima-dopo l'attuazione dell'intervento osservata per i controlli riproduce quella che sarebbe stata la dinamica per i trattati in assenza di trattamento, cioè il controfattuale.

Per via dell'assenza di assegnazione randomizzata del trattamento, in molti contesti potrebbe essere difficile stabilire a priori se questa condizione sia soddisfatta o meno. Sarà dunque opportuno, laddove si abbiano a disposizione le informazioni relative a più di un periodo antecedente l'intervento, verificare che i risultati passati abbiano seguito effettivamente una dinamica temporale simile nel gruppo dei trattati e dei controlli.

2.4 Requisiti aggiuntivi

Nelle precedenti sezioni abbiamo descritto alcuni delle principali strategie utili a garantire la comparabilità tra individui trattati e non trattati. E' bene qui ricordare altri requisiti, comuni a tutte e tre le strategie, che si rendono ulteriormente necessari affinché la valutazione produca risultati apprezzabili:

1. Il trattamento deve essere lo stesso per tutti gli individui trattati (o, se si vogliono misurare "dosi" diverse di un trattamento, si deve tenere traccia di chi è stato trattato con dosi differenti); ugualmente, i non trattati non devono ricevere (anche indirettamente) il trattamento; non deve cioè esserci "contagio". Nell'esempio dell'offerta extra-curricolare, è possibile che i bambini (o i genitori) che partecipano al programma riportino la loro esperienza ai compagni di classe che non vi partecipano. Se così fosse, non potremmo più considerare come non trattati i bambini che non partecipano al programma. Una possibile soluzione è quella di procedere all'assegnazione dell'intervento non su base individuale ma per gruppi (ad esempio, tra le classi di una scuola).
2. Gli individui non dovranno essere sottoposti ad altri trattamenti, pena l'impossibilità di distinguere quale sia stato l'intervento che ha generato un cambiamento nei risultati.
3. Si dovrà controllare che i trattati lo siano effettivamente, cioè che non adottino comportamenti che eludono il trattamento. Nell'esempio considerato, alcuni dei bambini iscritti al programma potrebbero in realtà non presentarsi, o frequentare solo alcune delle attività extra-curricolari proposte. Se trascurassimo questa partecipazione incompleta avremmo una sottostima dell'effetto del trattamento. D'altra parte, una partecipazione incompleta è spesso il massimo che si può ottenere, e quindi potrebbe essere più informativo misurare l'effetto causale della *intenzione a trattare*, piuttosto che del trattamento in sé.

3. RIFERIMENTI

3.1 Esperienze internazionali

- The Abdul Latif Jameel Poverty Action Lab, *Massachusetts Institute of Technology* (USA): <https://www.povertyactionlab.org/>
- What Works Clearinghouse, *Institute of Education Sciences* (USA): <http://ies.ed.gov/ncee/wwc/>

3.2 Bibliografia

- Joshua D. Angrist e Jorn-steffen Pischke (2015). *Mastering 'Metrics: The Path from Cause to Effect*, Princeton University Press.
- Paul J. Gertler, Sebastian Martinez, Patrick Premand, Laura B. Rawlings e Christel M. J. Vermeersch (2011). *Impact Evaluation in Practice*, The World Bank Publications.
- Shahidur R. Khandker, Gayatri B. Koolwal e Hussain A. Samad (2010). *Handbook on Impact Evaluation. Quantitative Methods and Practices*, The World Bank Publications.
- Antonio La Spina e Efisio Espa (2011). *Analisi e Valutazione delle Politiche Pubbliche*, Il Mulino.
- Alberto Martini e Marco Sisti (2009). *Valutare il successo delle politiche pubbliche*, Il Mulino.