

PAOLO PINOTTI

CORRELAZIONE



CAUSALITÀ

L'analisi empirica nell'ambito delle scienze sociali identifica relazioni sistematiche tra due o più fenomeni d'interesse sulla base della correlazione statistica tra gli stessi. D'altra parte, la semplice correlazione non rivela i nessi causali sottostanti a tali relazioni, che rivestono un ruolo fondamentale per la valutazione d'impatto delle politiche. Dopo aver chiarito la distinzione tra correlazione e causalità, il presente articolo discute recenti approcci sperimentali e quasi sperimentali all'identificazione degli effetti causali delle politiche.

L'analisi empirica nell'ambito delle scienze sociali identifica relazioni sistematiche tra due o più fenomeni d'interesse sulla base della correlazione statistica tra di essi. Una correlazione positiva tra gli eventi X e Y indica che una maggiore incidenza (o intensità) dell'evento X è associata, tipicamente, a una maggiore incidenza (o intensità) del fenomeno Y. Per esempio, esiste una correlazione positiva tra istruzione e reddito, in quanto individui più istruiti godono tipicamente di redditi più elevati. Ovviamente, esistono numerosi controesempi – individui molto istruiti con reddito basso e, per contro, individui poco istruiti con reddito alto. Tuttavia, tali combinazioni dovrebbero avere un'incidenza molto limitata in campioni sufficientemente ampi – quando si applica, cioè, la *legge dei grandi numeri*. La correlazione descrive dunque, in modo efficace e parsimonioso, la relazione 'tipica' tra istruzione e reddito. D'altra parte, l'analisi di correlazione non rivela, in generale, le determinanti della relazione empirica tra X e Y. In particolare, la correlazione tra i due fenomeni può riflettere l'effetto di X su Y, l'effetto di Y su X, e l'effetto

di fattori terzi che influenzano sia X che Y. Riprendendo l'esempio precedente, la correlazione positiva tra istruzione e reddito potrebbe essere dovuta al fatto che una maggiore istruzione migliori effettivamente le prospettive occupazionali. La stessa correlazione può altresì riflettere differenze lungo altre dimensioni, per esempio il *background* familiare. È ragionevole supporre, infatti, che i figli provenienti da famiglie più agiate abbiano accesso a migliori opportunità sia in ambito scolastico che lavorativo, indipendentemente dall'effetto causale dell'istruzione sul reddito. Distinguere tra queste spiegazioni alternative è di cruciale importanza per comprendere l'impatto causale dell'istruzione sul reddito e, quindi, le implicazioni di diverse politiche pubbliche in quest'ambito. In generale, l'analisi causale riveste un ruolo fondamentale ai fini della valutazione d'impatto degli effetti delle politiche pubbliche, un argomento al centro sia del dibattito politico che di quello scientifico. La valutazione d'impatto mira a quantificare l'effetto *causale* di una determinata politica X su uno o più *outcome* d'interesse Y. Dal punto di vista metodologico, la valutazione si basa sul confronto degli outcome osservati nelle unità statistiche – per esempio, individui, città o Paesi – rispettivamente esposte e non esposte alla politica. Tale confronto presuppone la correlazione tra esposizione alla politica X e la realizzazione del fenomeno Y. Tuttavia, come chiarito dal precedente esempio, la correlazione non è necessariamente informativa in merito all'effetto *causale* di X su Y.

La prima parte del presente articolo approfondisce l'importanza della distinzione tra causalità e correlazione, anche alla luce di un esempio tratto dal dibattito sul possibile effetto deterrente della pena di morte negli Stati Uniti. Si introduce quindi il metodo sperimentale come soluzione ideale all'identificazione degli effetti causali, evidenziando al contempo i vincoli di ordine economico, politico ed etico che ne limitano fortemente l'applicabilità nell'ambito delle scienze sociali. Si discutono dunque i recenti sviluppi di metodologie 'quasi-sperimentali' che mirano a replicare l'esperimento ideale superandone, al contempo, i limiti di fattibilità.

CAUSALITÀ E CORRELAZIONE

Utilizzando la definizione di causalità originariamente proposta dallo statistico polacco Jerzy Neyman nella sua tesi di dottorato presso l'Istituto Nazionale per l'Agricoltura di Bydgoszcz nel 1923 – e perfezionata, in anni più recenti, da statistici ed econometrici come Donald Rubin, Josh Angrist e James Heckman – la politica X causa l'outcome Y se, in assenza di X, la realizzazione di Y sarebbe stata diversa. Chiaramente, questa realizzazione alternativa di Y non è osservabile, e viene definita *controfattuale*. L'effetto causale di X è dunque la differenza tra la realizzazione 'fattuale' di Y e quella controfattuale.

Dal momento che possiamo osservare solo l'effettiva realizzazione di Y, ma non quella controfattuale, è impossibile identificare l'effetto di X per ogni singola unità di analisi. Si consideri, per esempio, l'effetto della pena di morte sul tasso di omicidi. Da un punto di vista teorico, è lecito ipotizzare che la pena di morte abbia un effetto deterrente e possa quindi causare una riduzione degli omicidi. Dal punto di vista empirico, tuttavia, non osserveremo mai il tasso di omicidi controfattuale nel caso in cui un determinato Paese adottasse politiche diverse riguardo alla pena di morte. Per esempio, negli Stati Uniti, dove è prevista la pena di morte – perlomeno nella maggior parte degli Stati – il tasso di omicidi nel 2016 è stato di 5,3 ogni 100.000 abitanti (fonte: Fbi, Uniform Crime Reports). Tuttavia, non sappiamo quale sarebbe stato il tasso di omicidi controfattuale se non ci fosse stata la pena di morte e, conseguentemente, non possiamo identificare l'effetto causale di tale politica sul tasso di omicidi negli Stati Uniti.

Per ovviare all'assenza di un controfattuale per ciascuna unità di analisi, è possibile confrontare gli outcome tra unità esposte e non esposte alla politica X. Tale analisi rivela la correlazione tra X e Y. In particolare, la correlazione sarà positiva se la politica X è associata mediamente a una maggiore incidenza (o intensità) del fenomeno Y, e sarà negativa nel caso contrario. In particolare, il *coefficiente di correlazione* riassume il segno e l'intensità della correlazione tra X e Y in un indice che varia tra 1 e -1, dove 0 indica l'assenza di correlazione.

Il coefficiente di correlazione descrive dunque, in modo efficace e parsimonioso, la relazione presente nei dati tra politica X e outcome d'interesse Y. Inoltre, se l'outcome osservato nelle unità non esposte approssima accuratamente l'outcome controfattuale che sarebbe osservato nelle unità esposte in assenza della politica, la correlazione tra X e Y fornisce una stima dell'effetto causale della politica X.

Ciò è vero, tuttavia, solo a condizione che unità esposte e non esposte siano simili in termini di altre determinanti dell'outcome Y. Questa condizione è generalmente violata nella valutazione d'impatto delle politiche pubbliche, in quanto queste ultime sono solitamente implementate in certe aree (e non in altre) in risposta a precedenti criticità. Tali criticità costituiscono dunque una differenza fondamentale tra aree esposte e non esposte in aggiunta, eventualmente, a differenze lungo altre dimensioni. In altri termini (e riprendendo la terminologia di cui sopra) l'outcome osservato nelle aree non esposte non fornisce una stima accurata del controfattuale nelle aree esposte, in quanto queste ultime erano già diverse in partenza.

Per capire quanto possa essere fuorviante attribuire valenza causale a semplici correlazioni, torniamo all'esempio precedente e consideriamo differenze territoriali nel tasso di omicidi all'interno degli Stati Uniti. Da sempre, il tasso di omicidi è considerevolmente maggiore negli Stati che applicano la pena di morte (5.6 ogni 100.000 abitanti, nel 2016) rispetto agli Stati che non applicano la pena di morte (4.5 ogni 100.000 abitanti). Differenze simili si riscontrano a livello internazionale: i Paesi che applicano la pena di morte (tra cui, appunto, gli Stati Uniti) si caratterizzano per un numero molto maggiore di crimini violenti, in rapporto alla popolazione, rispetto ai Paesi che non ap-

plicano la pena di morte. Per esempio, in tutti i Paesi dell'Europa occidentale il tasso di omicidi è largamente inferiore a 2 ogni 100.000 abitanti, in Italia, nel 2016, è stato pari a 0.67. Un'interpretazione di tali differenze in senso causale suggerirebbe che il ricorso alla pena di morte aumenti il numero di omicidi, una conclusione difficile da conciliare con i paradigmi di razionalità applicati alla decisione di commettere crimini. Plausibilmente, la correlazione positiva tra applicazione della pena di morte e omicidi riflette, invece, il fatto che la pena di morte sia adottata in risposta a un preesistente problema di criminalità.

L'APPROCCIO SPERIMENTALE ALL'IDENTIFICAZIONE DELLA CAUSALITÀ

L'esempio precedente chiarisce che, in generale, unità esposte e non esposte possono essere profondamente diverse ex ante e, di conseguenza, il confronto tra i due gruppi non rileva l'effetto causale della politica. La soluzione ideale a tale problema sarebbe un *esperimento randomizzato controllato* analogo a quelli condotti in laboratorio per testare l'effetto di un nuovo farmaco. In quest'ultimo caso, un numero relativamente ampio di pazienti viene suddiviso in modo del tutto casuale (per esempio, sulla base del lancio di una moneta) in due gruppi: 'trattati' e 'controlli'. Al primo gruppo viene somministrato il farmaco da testare, mentre al secondo gruppo viene somministrato un 'placebo' che non contiene il principio attivo. La legge dei grandi numeri garantisce che, su campioni sufficientemente ampi, due gruppi selezionati in modo casuale abbiano, in media, le stesse caratteristiche. Il gruppo di controllo approssima dunque il controfattuale del gruppo trattato in assenza del trattamento. Ne consegue che ogni differenza tra i due gruppi in termini di outcome – in questo caso, lo stato di salute – può essere ricondotta all'effetto causale del farmaco stesso¹.

La stessa logica si applica anche a esperimenti condotti in ambito sociale, che tuttavia presentano, per ovvi motivi, maggiori difficoltà di ordine pratico, economico e politico. Nonostante ciò, gli ultimi due decenni hanno visto una proliferazione di valutazioni basate su esperimenti randomizzati controllati, aventi come oggetto soprattutto interventi di contrasto alla povertà nei Paesi in via di sviluppo². Analogamente a quanto avviene in ambito medico, i potenziali beneficiari di tali interventi sono divisi casualmente in due gruppi, uno dei quali sarà effettivamente destinatario dell'intervento (i trattati) mentre l'altro fungerà da termine di confronto (i controlli). Questo approccio ha consentito una stima accurata degli effetti di microinterventi quali, per esempio, trasferimenti monetari alle famiglie, distribuzione di medicinali e programmi di microcredito.

1. MATTHEWS 2006.

2. DUFLO – BANERJEE 2011.

Al contempo, gli esperimenti in ambito sociale sono spesso soggetti a vincoli di ordine economico, etico e politico. Nell'ordine, essi sono solitamente molto costosi, in quanto presuppongono il finanziamento dell'intervento stesso per un ampio numero di beneficiari – tale da garantire la comparabilità di trattati e controlli in base alla legge dei grandi numeri – in aggiunta alla gestione di complesse macchine organizzative. Inoltre, gli esperimenti incontrano spesso forti obiezioni dal punto di vista etico, in quanto l'allocatione di interventi assistenziali importanti viene decisa in modo puramente casuale. Rispetto a questo tipo di critiche è importante notare, tuttavia, che lo stesso approccio è da anni universalmente accettato come presupposto essenziale per test clinici di nuovi farmaci, in alcuni casi salvavita. In entrambi gli ambiti, medico e sociale, l'allocatione casuale nell'iniziale fase di valutazione trova giustificazione nel fatto che, in caso di esito positivo, l'intervento potrebbe (e dovrebbe) essere poi esteso anche ai soggetti inizialmente esclusi. In ogni caso, è innegabile che gli esperimenti pongano dilemmi etici di non facile soluzione, che determinano inevitabilmente forti resistenze alla diffusione di questo tipo di valutazioni.

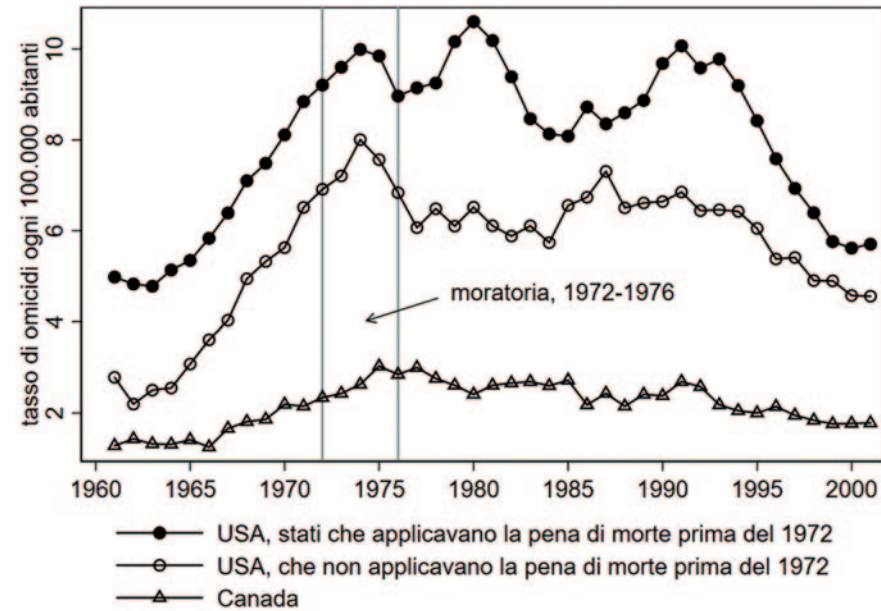
Tali resistenze sono comprensibilmente maggiori quando ci spostiamo dalla valutazione di microinterventi assistenziali come quelli descritti in precedenza alla valutazione di macropolitiche in materia, per esempio, di istruzione, fiscalità e lavoro, fino a diventare insormontabili nell'ambito delle politiche sulla giustizia e la sicurezza. In questi contesti, la sperimentazione di politiche e interventi allocati su base casuale è generalmente improponibile; si pensi al caso estremo della pena di morte, discusso in precedenza. In questi casi, un approccio alternativo consiste nello sfruttare eventi e situazioni che producono variazioni similmente casuale nell'esposizione a una determinata politica, anche in assenza di un vero e proprio esperimento.

GLI ESPERIMENTI NATURALI

Molte politiche incorporano elementi di arbitrarietà e discrezionalità che generano differenze di trattamento tra soggetti sostanzialmente simili. Per esempio, una riforma pensionistica può modificare parametri fondamentali del trattamento pensionistico, quali i requisiti contributivi e il coefficiente di trasformazione, solo per i nati da un certo anno in poi. In questo caso, la prima generazione di individui esposti alla riforma e l'ultima generazione non esposta sono plausibilmente simili da tutti i punti di vista a eccezione dell'esposizione alla riforma. Ogni differenza in termini di outcome – per esempio, offerta di lavoro – può essere ragionevolmente ricondotta all'effetto causale della riforma.

Questo tipo di eventi sono denominati *esperimenti naturali* o *quasi-esperimenti*, poiché riproducono due caratteristiche essenziali degli esperimenti descritti nella sezione precedente:

- la ripartizione dei soggetti d'interesse in due gruppi, trattati e controlli;
- l'allocatione casuale dei soggetti tra i due gruppi e, di conseguenza, la comparabilità di trattati e controlli.



Tasso di omicidi negli Stati Uniti e Canada, 1961-2001.

Diversamente da quanto accade negli esperimenti veri e propri, la casualità nella ripartizione non deriva da una decisione intenzionale ma è una conseguenza di una politica attuata a prescindere, e non solleva dunque obiezioni di ordine economico, etico e politico. Analisi basate su esperimenti naturali sono dunque praticabili anche in alcuni casi in cui esperimenti veri e propri sarebbero impraticabili.

Per illustrare l'utilità degli esperimenti naturali, si torni all'esempio precedente, ovvero la stima dell'effetto deterrente della pena di morte negli Stati Uniti. Per ovvi motivi, non è possibile svolgere l'analisi utilizzando un esperimento randomizzato controllato. Tuttavia, è possibile approfittare del fatto che l'effettiva applicazione della pena di morte è variata notevolmente nel corso degli anni. In particolare, le esecuzioni raggiunsero un minimo storico tra la fine degli anni Sessanta e l'inizio degli anni Settanta. Si giunse così a una moratoria totale, sancita dall'intervento della Corte Suprema, tra il 1972 e il 1976, prima che le esecuzioni capitali riprendessero nell'anno successivo.

La moratoria determinò dunque un drastico cambio nella politica sulla pena di morte nel periodo 1972-1976 rispetto sia agli anni precedenti che a quelli successivi, generando quindi un esperimento naturale utile a identificare l'impatto della stessa sul numero di omicidi. Si noti altresì che la moratoria incise sull'effettiva applicazione della pena di morte negli Stati in cui la stessa era effettivamente praticata: questo primo gruppo di Stati fungerà dunque da gruppo 'trattato'. Per contro, la moratoria fu irrilevante per gli Stati che avevano abolito la pena di morte già prima del 1972, che fungeranno dunque da gruppo di controllo³.


3. Si rimanda a DONOHUE – WOLFERS 2005 per un'analisi approfondita.

La figura nella pagina precedente mostra la dinamica degli omicidi nei due gruppi di Stati tra il 1961 e il 2001. Il tasso di uccisioni rimane parallelo durante l'intero periodo, inclusi gli anni della moratoria. La sospensione della pena di morte non sembra dunque aver inciso in nessun modo sul numero di delitti; in altri termini, non c'è nessuna evidenza che la pena di morte eserciti un significativo effetto deterrente. È importante notare, ancora una volta, che la semplice correlazione tra politica e outcome sarebbe fuorviante. Da un lato, la correlazione tra applicazione della pena di morte e tasso di omicidi tra diversi periodi è negativa, in quanto questi crescono marcatamente e raggiungono un picco proprio durante il periodo della moratoria. Questa correlazione sembrerebbe supportare la tesi che la pena di morte eserciti effettivamente un effetto deterrente. D'altra parte, come già notato in precedenza, la correlazione tra applicazione della pena di morte e tasso di omicidi tra i diversi Stati è di segno opposto, in quanto tale tasso è maggiore laddove è in vigore la pena di morte. In realtà, nessuna delle due correlazioni fornisce chiara evidenza in merito all'effetto causale della pena di morte sugli omicidi, in quanto l'applicazione della pena di morte può essere essa stessa causata dalla diffusione della criminalità e, inoltre, entrambe le variabili possono dipendere da altri fattori. Per isolare il nesso causale d'interesse è dunque necessario effettuare il 'doppio confronto' descritto nella figura: Stati con e senza la pena di morte, durante i periodi in cui la moratoria era e non era in vigore. I risultati di tale confronto sono dettagliati nella tabella sotto. Confrontando gli anni della moratoria con il periodo precedente, il tasso di omicidi aumenta da 6.3 a 9.5 negli Stati che applicavano la pena di morte e da 3.9 a 7.3 negli altri Stati. La differenza tra questi due incrementi è molto vicina a zero. Lo stesso accade quando confrontiamo il periodo della moratoria con quello successivo.

	Media	Media	Differenza tra periodi	Media	Differenza tra periodi
Anni	1961-71	1972-76	1972-76 e 1961-71	1977-2001	1977-2001 e 1972-76
Stati con pena di morte	6.26	9.52	3.26	8.51	-1.01
Stati senza pena di morte	3.92	7.31	3.38	6.05	-1.26
Differenza tra i due gruppi	2.33	2.22		2.45	
Differenze-in-differenze			-0.12		0.24

Tasso di omicidi negli Stati Uniti, diversi gruppi di Stati e periodi.

Questo approccio, comunemente denominato *differenza-in-differenze*, si fonda sull'idea che il gruppo di controllo fornisce un adeguato controfattuale per l'evoluzione dell'outcome nel gruppo trattato nel caso in cui quest'ultimo non fosse stato esposto al trattamento⁴. Nel caso in esame, tale confronto ci porta a concludere che il tasso di omicidi sarebbe aumentato comunque, durante gli anni Settanta, indipendentemente dall'applicazione della pena di morte – e, dunque, l'effetto deterrente di quest'ultima rimane trascurabile. Per rimarcare ulteriormente questo punto, la figura mostra che anche il tasso di omicidi in Canada – ancora meno influenzato dalla moratoria implementata negli Stati Uniti – evidenzia una dinamica simile.

L'ultimo decennio ha visto una proliferazione di valutazioni d'impatto che utilizzano simili approcci quasi-sperimentali. Un'altra metodologia che ha visto una crescente diffusione è stata, per esempio, il *regression discontinuity design*⁵. Il successo di tali metodi si fonda sulla combinazione di esperimenti naturali atti a isolare il nesso causale d'interesse e la disponibilità di dati sufficientemente dettagliati e disaggregati sui fenomeni d'interesse (per esempio, dati amministrativi a livello individuale). Se queste due condizioni sono soddisfatte, è possibile condurre valutazioni d'impatto che isolino l'effetto causale della politica dalla correlazione prodotta da altri fattori utilizzando metodi statistici estremamente semplici in alcuni casi, come l'esempio appena esposto, il semplice confronto tra medie 

4. CARD – KRUEGER 1994.

5. Per un compendio metodologico, cfr. LEE E LEMIEUX 2010; per un'applicazione alla valutazione d'impatto delle politiche migratorie in Italia, cfr. PINOTTI 2017.

BIBLIOGRAFIA

- D. CARD – A.B. KRUEGER, *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, «American Economic Review» LXXXIV (1994) 4, pp. 772-793.
- J.J. DONOHUE – J. WOLFERS, *Uses and Abuses of Empirical Evidence in the Death Penalty Debate*, «Stanford Law Review» LVIII (2005), pp. 791-845.
- E. DUFLO – A. BANERJEE, *Poor economics*, Public Affairs, New York 2011.
- D.S. LEE – T. LEMIEUX, *Regression discontinuity designs in economics*, «Journal of economic literature» XLVIII (2010) 2, pp. 281-355
- J.N. MATTHEWS, *Introduction to randomized controlled clinical trials Chapman and Hall/CRC* (2006).
- P. PINOTTI, *Clicking on heaven's door: The effect of immigrant legalization on crime*, American «Economic Review» CVII (2017) 1, pp. 138-168.
- J. NEYMAN, *Sur les applications de la théorie des probabilités aux expériences agricoles: Essai des principes*, Tesi di Dottorato (1923).