

VALUTAZIONE CONTROFATTUALE DELL'IMPATTO DELLE POLITICHE SULL'OCCUPAZIONE GIOVANILE

GUIDA METODOLOGICA

MÁRTON CSILLAG, JUDIT KREKÓ, ÁGOTA SCHARLE

contributi di:

MARCO CENTRA, MASSIMILIANO DEIDDA, LUCÍA GORJÓN,
VALENTINA GUALTIERI, IGA MAGDA, MARTA PALCZYŃSKA,
FRANCESCO TRENTINI, CLAUDIA VILLOSIO

dicembre 2020



Valutazione controfattuale dell'impatto delle politiche sull'occupazione giovanile

Guida metodologica

Il progetto Youth Employment PartnerSHIP²
è finanziato da Islanda, Lichtenstein e Norvegia tramite gli
EEA and Norway Grants Fund for Youth Employment

IMPLEMENTATO DA:



Il Collegio Carlo Alberto (CCA) è una fondazione costituita nel 2004 su iniziativa congiunta della Compagnia di San Paolo e dell'Università degli Studi di Torino. Ha la missione di: promuovere attività di ricerca interdisciplinare e di frontiera sia di base sia applicata nelle scienze economiche, giuridiche, politiche e sociali; svolgere programmi di formazione avanzata; condurre attività di divulgazione sui grandi temi sociopolitici contemporanei, fondata sulla migliore ricerca internazionale. Le principali aree di studio del Collegio Carlo Alberto riguardano l'economia, la sociologia, le scienze politiche e le scienze giuridiche, la finanza e le politiche pubbliche e sono portate avanti con un approccio interdisciplinare. L'attività di didattica si articola nel programma Allievi, due dottorati e quattro master universitari. Il Collegio condivide e fa propri i valori e i metodi della comunità scientifica internazionale.

www.carloalberto.org

L'Istituto nazionale per l'analisi delle politiche pubbliche (INAPP) è un ente pubblico di ricerca che si occupa di analisi, monitoraggio e valutazione delle politiche del lavoro, delle politiche dell'istruzione e della formazione, delle politiche sociali e, in generale, di tutte le politiche economiche che hanno effetti sul mercato del lavoro. Nato il 1° dicembre 2016 a seguito della trasformazione dell'Isfol e vigilato dal Ministero del Lavoro e delle politiche sociali, l'Ente ha un ruolo strategico - stabilito dal decreto legislativo 14 settembre 2015, n. 150 - nel nuovo sistema di governance delle politiche sociali e del lavoro del Paese. Inapp fa parte del Sistema statistico nazionale (SISTAN) e collabora con le istituzioni europee. Da gennaio 2018 è Organismo Intermedio del PON Sistemi di politiche attive per l'occupazione (SPAO) per svolgere attività di assistenza metodologica e scientifica per le azioni di sistema del Fondo sociale europeo ed è Agenzia nazionale del programma comunitario Erasmus+ per l'ambito istruzione e formazione professionale. È l'ente nazionale all'interno del consorzio europeo ERIC-ESS che conduce l'indagine European Social Survey.

www.inapp.org

CONTENUTI

1. INTRODUZIONE	5
2. I DATI AMMINISTRATIVI	9
3. IL PROCESSO DI VALUTAZIONE	15
3.1. Scelta delle misure specifiche da valutare	15
3.2. Scelta delle variabili risultato	16
3.3. Strategia di identificazione: scelta del metodo della valutazione d'impatto controfattuale	17
3.4. Effetti eterogenei	23
4. QUESTIONI RILEVANTI PER I DECISORI POLITICI	25
4.1. Interpretazione dei risultati	25
4.2. Relazione tra gli esiti del programma e la sua definizione e implementazione	26
4.3. Validità esterna, comparabilità con altri studi	26
4.4. Effetto di spiazzamento, perdite secche	26
4.5. Analisi costi-benefici e analisi costi-efficacia	27
Bibliografia	28
Appendice 1: Breve descrizione delle valutazioni effettuate nei quattro Paesi	30
Italia	30
Ungheria	31
Polonia (Valuazione 1)	31
Polonia (Valuazione 2)	32
Spagna	33
Appendice 2: Glossario e abbreviazioni	34

1. INTRODUZIONE¹

Con il progetto “Youth employment partnerSHIP - evaluation studies in Spain, Hungary, Italy and Poland”, i partner italiani, spagnoli, ungheresi e polacchi, che hanno vinto la call indetta dal Fondo EEA and Norway Grants 2014-2021, si sono riuniti in un consorzio internazionale di ricerca. Scopo del progetto è la valutazione con metodo controfattuale (*Couterfactual Impact Evaluation*) dell’impatto sull’occupazione giovanile delle politiche del lavoro in Spagna, Ungheria, Italia e Polonia. I paesi partecipanti condividono tassi di disoccupazione giovanile elevati e una scarsa capacità di effettuare valutazione controfattuale di qualità utilizzando le banche dati disponibili di fonte amministrativa.

La valutazione controfattuale o valutazione d’impatto o degli effetti di una politica è un processo attraverso cui stimare, con metodologie adeguate e dati oggettivi e verificabili, se un certo intervento pubblico riesce a produrre un cambiamento su un certo fenomeno diverso da quello che si sarebbe verificato in assenza di quell’intervento.

La Guida metodologica intende accompagnare il lettore lungo un percorso di valutazione dell’impatto delle politiche del lavoro sull’occupazione giovanile, evidenziando i problemi pratici incontrati nei quattro paesi che partecipano al Progetto Youth Employment PartnerSHIP (Spagna, Ungheria, Italia e Polonia). Tuttavia, la Guida è pensata anche per essere utilizzata nella valutazione di altri tipi di programmi e di politiche.

Scopo di questa Guida è quello di fornire una cassetta degli attrezzi a ricercatori e amministratori che intendono pianificare e realizzare valutazioni di misure per l’occupazione giovanile. I requisiti per la lettura della Guida sono una conoscenza di base della statistica e dei concetti di valutazione. In appendice è disponibile un glossario dei principali concetti e vocaboli tecnici. Dal momento che le misure analizzate si differenziano nei diversi Stati membri e i dati e le fonti disponibili per la valutazione sono diversi, la Guida non aspira a proporre un unico modello di valutazione²; piuttosto, riepiloga e descrive alcune tipiche questioni pratiche incontrate nel corso delle valutazioni degli effetti sui giovani di alcune misure di politica per il lavoro, svolte con metodo controfattuale³.

In particolare, il Programma Garanzia Giovani⁴, presente in tutta l’Unione europea, è l’oggetto di alcune valutazioni condotte con rigore metodologico in alcuni Stati membri. In due Paesi, Italia e Spagna, è stato valutato l’effetto sui giovani di altre tipologie di misure, finalizzate ad aumentare la stabilità lavorativa e ridurre la precarietà occupazionale attraverso incentivi alle imprese per l’assunzione con contratti di lavoro dipendente a tempo indeterminato.

L’attività di valutazione è indispensabile per conoscere l’impatto delle politiche. La valutazione quantitativa basata sull’analisi dei dati risulta cruciale per sapere se le politiche hanno raggiunto l’obiettivo, ma anche per migliorare la loro efficienza.

1 Questo documento è la traduzione in lingua italiana di Counterfactual Evaluation of Youth Employment Policies: Methodological Guide <<https://bit.ly/3bpNYMj>>

2 Questa Guida si focalizza sulla valutazione di tipo controfattuale effettuata tramite dati amministrativi. Esempi più ampi di guide di valutazione e monitoraggio che partano dallo sviluppo della teoria del cambiamento sono Leeuw and Vaessen (2009), UNDP (2009), Gertler et al. (2016.), European Commission (2013).

3 Per una descrizione dettagliata dei metodi di valutazione quantitativa controfattuale si veda ad esempio Angrist e Pischke (2009), Abadie e Cattaneo (2018), Gertler et al. (2016), European Commission (2013).

4 Questi sono sostenuti finanziariamente dall’Iniziativa per l’occupazione giovanile (YEI - Youth Employment Initiative) nelle regioni in cui il tasso di disoccupazione giovanile è superiore al 25%.

Scopo della valutazione delle politiche del lavoro con metodo controfattuale è fornire una stima dell'effetto causale delle misure previste su alcune variabili risultato (*outcome*), come ad esempio la probabilità e la durata dello stato di disoccupazione per gli individui. Significa stimare l'effetto della politica rispetto a quello che sarebbe stato il risultato se la politica non fosse stata introdotta, a parità di altre condizioni (*ceteris paribus*). In pratica, dal momento che il controfattuale non è osservabile, occorre trovare un "gruppo di controllo" credibile, che possa essere considerato simile in tutto al gruppo che partecipa al programma e che differisca unicamente per il fatto che i suoi membri non partecipano al programma. In tutte le fasi del processo di valutazione, aspetti teorici o limiti derivanti dalla disponibilità di dati giocano un ruolo decisivo sulle scelte dei ricercatori.

Il Progetto Youth Employment PartnerSHIP ha effettuato una valutazione di politiche del mercato del lavoro che agiscono principalmente dal lato della domanda, attraverso incentivi monetari alle imprese che assumono giovani⁵. In due paesi (Ungheria e Polonia), il Gruppo di ricerca ha valutato gli incentivi all'assunzione offerti da "Garanzia giovani", il programma dell'Unione europea (UE) che impegna tutti gli Stati membri a garantire ai giovani NEETs (*not in education, employment or training*) un'offerta di servizi di istruzione, tirocinio, formazione, di buona qualità ed erogata in tempi rapidi. In due Paesi (Italia e Spagna) le politiche sotto scrutinio avevano come obiettivo l'aumento della stabilità occupazionale tramite l'incentivazione di contratti a tempo indeterminato e la riduzione dei costi in caso di licenziamento anticipato. Le caratteristiche principali delle valutazioni sono schematizzate nei Riquadri 1 e 2 e sintetizzate in maggior dettaglio nell'Appendice 1.

Le valutazioni realizzate nei quattro Paesi partecipanti al progetto Youth Employment PartnerSHIP sono state condotte utilizzando dati amministrativi a livello individuale (liste di disoccupazione, registri previdenziali, comunicazioni obbligatorie dei datori di lavoro relative ai flussi degli avviamenti e delle cessazioni di contratti di lavoro dipendente ecc.), diversi in funzione della disponibilità dei dati nei vari Stati membri europei. Questa guida pone grande enfasi sulle sfide e le questioni aperte sull'uso dei dati amministrativi per la *policy evaluation*.

La guida è strutturata nel modo seguente. La prima parte riassume gli aspetti legati all'accesso e alla preparazione dei dati amministrativi per l'utilizzo a fini statistici, rispetto ai quali è importante ricordare che limiti legati alla possibilità di utilizzo dei dati (di qualunque natura) hanno un impatto in tutte le fasi del processo di valutazione. La seconda parte ripercorre il processo di valutazione quantitativa e si concentra sulle scelte che devono essere fatte perché i risultati ottenuti possano considerarsi validi. L'ultima parte illustra i punti di attenzione e le problematiche delle fasi successive, quelle che emergono quando si portano a conoscenza i *policy makers* dei risultati delle politiche da loro introdotte. Le diverse sezioni illustrano le esperienze e gli esempi specifici tratti dal lavoro svolto dai quattro Partner Europei del Progetto Youth Employment PartnerSHIP. Gli elementi principali dei quattro esercizi valutativi sono schematizzati nei Riquadri 1 e 2 e infine riassunti nell'Appendice 1 con maggiore dettaglio. L'Appendice 2 fornisce le definizioni dei termini tecnici più ricorrenti.

⁵ Molte delle questioni discusse, tuttavia, non riguardano esclusivamente i contributi all'assunzione.

Riquadro 1: dettagli delle valutazioni per paese partner: Italia

ITALIA	
Politica analizzata	Valutazione congiunta di: 1) Riduzione dei costi di licenziamento per i contratti a tempo indeterminato (contratto a tutele crescenti) 2) Sgravi contributivi per i nuovi contratti a tempo indeterminato e per le trasformazioni da tempo determinato a tempo indeterminato.
Criteri di idoneità	Persone non precedentemente occupate con contratto di lavoro dipendente a tempo indeterminato interrotto nei sei mesi antecedenti la data di nuova assunzione.
Durata e tipo	1) Riordino permanente della disciplina dei licenziamenti. 2) 36 mesi di sgravi contributivi
Incentivi	1) Riordino permanente della disciplina dei licenziamenti. 2) Riduzione del 100% del costo dei contributi.
Dati	SISCO (Sistema Informativo Statistico delle Comunicazioni Obbligatorie). SISCO è il database amministrativo pubblico che contiene le informazioni elementari su assunzioni, trasformazioni e cessazioni dei contratti di lavoro.
Campione	Contratti di lavoro sottoscritti nel periodo 2014-2015.
Metodologia	DiD, con correzione parametrica della selezione campionaria e interazioni tra trattamento e classi di età per ottenere gli effetti specifici per le coorti giovani.
Outcome	Quota di nuovi avviamenti a tempo indeterminato sul totale dei nuovi contratti di lavoro dipendente registrati nel 2015.

Riquadro 2: dettagli delle valutazioni per paese partner: Polonia

POLONIA	
Politica analizzata	1) Sussidi salariali 2) Formazione on-the-job, formazione in classe, programmi di lavoro pubblico, sussidi salariali, voucher per la formazione on-the-job e voucher per la formazione in classe.
Criteri di idoneità	Disoccupati fino a 30 anni di età per 1) e 2)
Durata e tipo	1) Incentivi per il datore di lavoro pari al valore corrispondente al salario minimo + contributi sociali (~500 + 100 € nel 2018) (2) sussidi ai programmi di lavoro pubblico e vouchers per la formazione on-the-job
Incentivi	1) Incentivi per il datore di lavoro pari al valore corrispondente al salario minimo + contributi sociali (~500 + 100 € nel 2018) (2) sussidi ai programmi di lavoro pubblico e vouchers per la formazione on-the-job
Dati	Registro amministrativo dei Servizi per l'Impiego
Campione	2015–2017
Metodologia	1) Combinazione di RDD e DiD 2) propensity score matching
Outcome	1) Assenza dalle liste di disoccupazione tra 12 e 36 mesi dopo la prima iscrizione e 2) assenza dalle liste e assenza dai programmi di Politiche attive del lavoro nei 12 - 36 mesi successivi all'iscrizione.

Riquadro 3: dettagli delle valutazioni per paese partner: Spagna

SPAGNA	
Politica analizzata	Tirocini
Criteri di idoneità	Laurea triennale o diploma professionale, per disoccupati fino a 30 anni di età
Durata e tipo	Minimo 6 mesi, massimo 2 anni. Full or part time.
Incentivi	50-75% di sgravio contributivo. Bonus per la conversione a tempo indeterminato di 500€ per anno per gli uomini o di 700€ per le donne, per i primi tre anni.
Dati	(1) Registro dei Servizi per l'Impiego (SPI), archivio completo dei contratti, collegato a: (2) campione tratto dagli archivi della Previdenza sociale: Campione continuo delle Storie Lavorative (CSWH)
Campione	Archivi dal 2002 al 2018
Metodologia	Modello logit multinomiale per la comparazione dei partecipanti ai tirocini con idonei non-partecipanti al tirocinio e assunti con regolare contratto di lavoro a termine.
Outcome	1) Probabilità che l'individuo rimanga presso l'impresa al termine del tirocinio. 2) Probabilità che l'individuo firmi un contratto permanente presso la stessa impresa. 3) Probabilità che l'individuo firmi un contratto permanente presso un'altra impresa.

Riquadro 4: dettagli delle valutazioni per paese partner: Ungheria

UNGHERIA	
Politica analizzata	90 giorni di incentivo salariale per il periodo di prova
Criteri di idoneità	Iscritti alle liste di disoccupazione fino a 25 anni di età
Durata e tipo	Periodo massimo di 90 giorni
Incentivi	100% del costo del lavoro (salario + versamenti previdenziali del datore di lavoro)
Dati	1) registro dei Servizi per l'Impiego e collegati agli: 2) archivi della previdenza sociale e del sistema d'istruzione
Campione	2015-2017 (serie storiche costruite a partire dal 2003)
Metodologia	1) Propensity score e nearest neighbor matching per confrontare il gruppo dei partecipanti ai "job trial" con i partecipanti ai tirocini. 2) Difference-in-differences, per sfruttare il ritardo di nove mesi nella data di avvio del programma in una delle regioni.
Outcome	1) Occupati a 6 e 12 mesi dall'inizio degli incentivi 2) Salari cumulati tra 6 e 12 mesi dall'inizio della sovvenzione

2. I DATI AMMINISTRATIVI

Le basi dati amministrative di solito si caratterizzano per una copertura elevata o totale (censuaria), per una dimensione ampia dei campioni e, in molti casi, per la possibilità di unire le banche dati (*data-linkage*) anche raccolte con indagini campionarie (*survey*⁶). L'affidabilità del dato amministrativo è, di solito, maggiore rispetto a quella dei dati di survey, tuttavia gli archivi amministrativi possono essere talvolta incompleti e in certi casi contenere errori. Nella valutazione delle politiche del lavoro i dataset di fonte amministrativa più utilizzati sono quelli degli istituti di previdenza sociale, i registri della disoccupazione e quelli dei Servizi per l'Impiego (SPI) - che possono includere informazioni su destinatari e sulla tipologia di misura di politica attiva del lavoro - nonché i registri delle imprese. Dal momento che gli archivi amministrativi contengono microdati a livello individuale o d'impresa, l'utilizzo è rigidamente disciplinato e l'accesso è regolato dalle norme sulla protezione dei dati. Sia la raccolta e l'archiviazione dei dati amministrativi sia l'accesso e l'utilizzo dei dataset sono soggetti a regimi in parte differenziati nei diversi paesi e per questa ragione il valutatore può incontrare difficoltà e affrontare una lunga procedura per accedere ai dati richiesti. Molte altre questioni sorgono infine quando si passa all'utilizzo dei dati⁷, come emerge da un sondaggio tra i partners del progetto.

- **Esiste un archivio amministrativo centrale a livello nazionale?**

La portata della valutazione è influenzata dalla disponibilità di registri centralizzati a livello nazionale o di registri decentrati a livello delle diverse amministrazioni territoriali.

In **Ungheria, Spagna e Polonia** i sistemi informativi dei Servizi per l'Impiego sono integrati a livello nazionale, mentre in **Italia** a livello centrale è disponibile un registro amministrativo che raccoglie le informazioni relative a tutti gli avviamenti, cessazioni e trasformazioni di contratti di lavoro dipendente e parasubordinato, relativi a tutti i settori economici, compresa la Pubblica Amministrazione (con esclusione dei lavoratori autonomi), denominato "Sistema Informativo Statistico delle Comunicazioni Obbligatorie del Ministero del Lavoro e delle Politiche Sociali", SISCO.

- **Esiste la possibilità di integrare il dataset con altri dati di fonte amministrativa al fine di arricchire l'insieme delle variabili di controllo e delle variabili risultato (*data-linkage*)?**

Le informazioni relative a ciascun individuo di solito sono registrate in archivi di competenza di autorità amministrative diverse, per esempio, istituti di previdenza sociale, servizi per l'impiego, enti di formazione, fisco. Il *data linkage* di *database* di natura e fonte diversa ha dunque un potenziale enorme per aumentare l'estensione dell'informazione disponibile. In **Ungheria**, ad esempio, esiste la possibilità di integrare gli archivi dei Servizi per l'Impiego (SPI) con quelli del sistema pensionistico, che includono i dati sulla vita lavorativa e sui salari relativi anche di individui in cerca di occupazione presenti nel *database* degli SPI. Questo consente l'uso di più variabili risultato (comprese informazioni sulla qualità del lavoro), così come un ricco insieme di altre caratteristiche osservabili, relative anche a periodi precedenti l'occupazione. In **Spagna**,

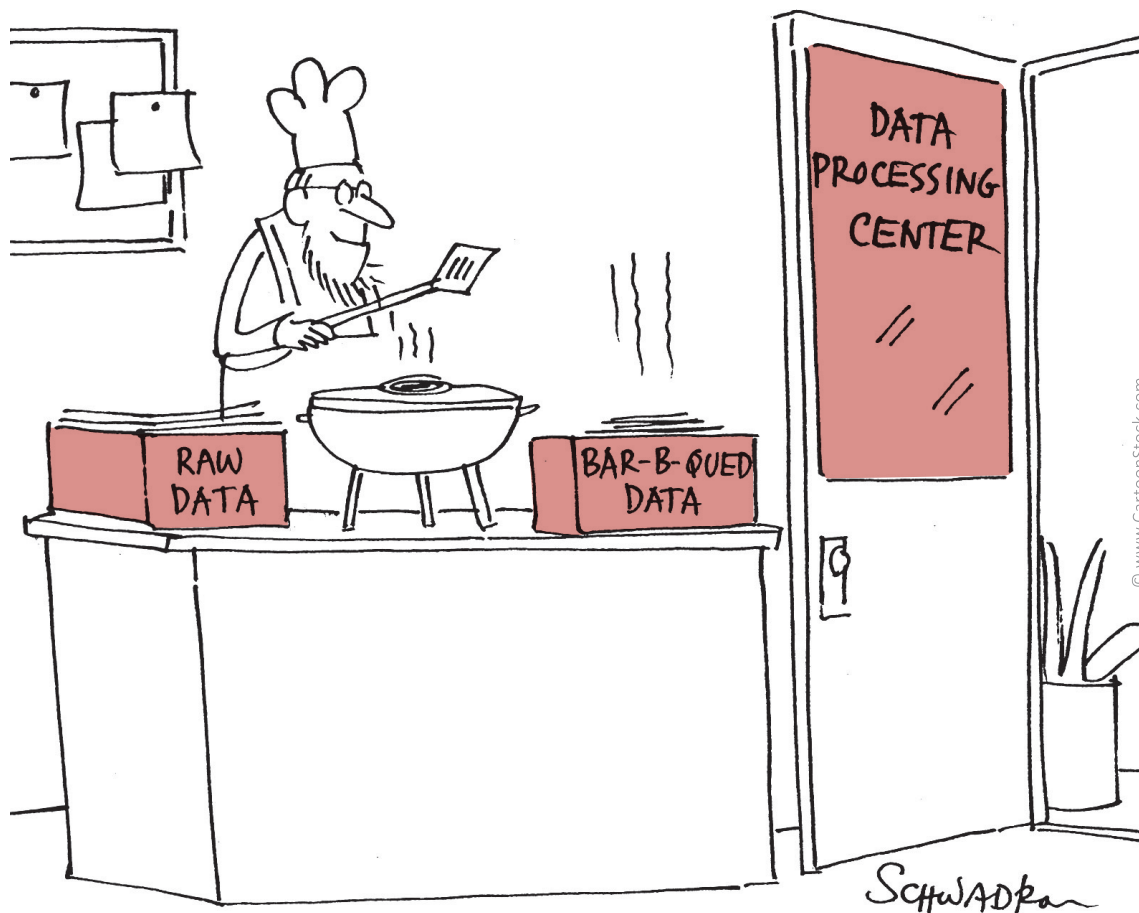
6 Si veda, ad esempio, Caliendo et al. (2017) sull'importanza dell'accesso a dati basati su rilevazioni statistiche (*survey*) nella valutazione delle politiche attive per il lavoro.

7 Per una guida dettagliata sull'uso di dati amministrativi per la valutazione degli interventi del Fondo Sociale Europeo si veda European Commission (2020).

la Fundación de Estudios de Economía Aplicada, FEDEA, ha accesso a due distinti database amministrativi integrabili tra loro. Il campione tratto dagli archivi del sistema di previdenza sociale (Continuous Sample of Working Histories, CSWH, in spagnolo “Muestra Continua de Vidas Laborales”) e dal registro della disoccupazione e dei contratti di lavoro raccolti mensilmente dai Servizi per l’Impiego spagnoli. Ciononostante, in molti casi, l’integrazione di diversi database è rigidamente regolamentata e limitata o proibita dalle leggi sulla privacy e sulla protezione dei dati. Ad esempio, in **Polonia**, i dati degli archivi SPI non possono essere oggetto di *data-linkage* con i dati della previdenza sociale su occupazione e salari, tranne nei casi specificamente previsti dalla norma, come ad esempio l’attività di rendicontazione alla Commissione europea.

- Qual è la procedura di accesso ai dati?

Non c’è una procedura comune a livello UE per l’accesso ai dati e i requisiti richiesti differiscono tra gli Stati membri. Ad esempio, in **Spagna**, il database contenente le storie lavorative (CSWH) può essere ottenuto facendo richiesta di accesso attraverso l’Istituto di previdenza sociale. I dati degli SPI non sono, in generale, immediatamente accessibili ai ricercatori. Tuttavia, FEDEA ha potuto accedere ai dati a seguito di un accordo di cooperazione firmato da entrambi le parti. Neppure in **Polonia** esiste una procedura formalizzata per l’accesso ai dati della previdenza e degli SPI, che rimane pertanto discrezionale⁸.



- **Caratteristiche istituzionali e uso amministrativo dei dati**

Una raccomandazione importante per i ricercatori è quella di prendere contatto con i titolari delle banche dati per chiarirne le caratteristiche e i vincoli istituzionali, la funzione originaria, i processi e le procedure utilizzate per la raccolta. L'uso e l'affidabilità dei dati, in termini di qualità, può differire anche all'interno dello stesso database a seconda che la singola informazione sia usata in fase di registrazione, valutazione o processo amministrativo. Una conoscenza dettagliata dei meccanismi di raccolta, conservazione ed elaborazione dei dati è essenziale per la preparazione della strategia di valutazione.

In alcuni casi, in particolar modo nei casi in cui la raccolta e la valutazione dei dati è gestita in cooperazione tra diversi enti, il valutatore potrebbe cercare di influenzare la metodologia utilizzata e i contenuti della raccolta dati per lo sfruttamento anche a fini statistici e aumentare la possibilità di utilizzo per la valutazione delle politiche. L'armonizzazione e reciproca calibrazione dei processi di valutazione e di raccolta è particolarmente utile per valutazioni condotte nel lungo periodo o a intervalli regolari di tempo.

- **Qualità e completezza dei dati**

La rappresentatività dei dati spesso non costituisce un problema nell'uso di dati amministrativi in quanto questi vengono generalmente raccolti per tutta la popolazione o per campioni rappresentativi. Ciononostante i dati amministrativi possono contenere errori o dati mancanti (*missing*), specialmente nei casi in cui la registrazione di un certo dato sia opzionale al momento della raccolta. Il ricercatore deve sempre verificare l'ampiezza e la sistematicità dei dati mancanti e applicare dei correttivi se necessario. Uno dei modi per condurre questo tipo di verifica è confrontare le consistenze e le caratteristiche aggregate dei dati con quelle di dati di fonte diversa.

- **I dati presenti negli archivi sono disponibili sia per il "gruppo dei trattati" sia per il "gruppo di controllo"?**

Un problema ricorrente quando si valutano le politiche pubbliche utilizzando, ad esempio, i dati dei Servizi per l'Impiego è la mancanza di integrazione tra i diversi archivi di dati che possono essere diversi: per struttura, numero e tipologia di variabili, provenienti da fonti diverse e diversi per i partecipanti ai vari programmi, nonché per il potenziale gruppo di controllo. Da qui deriva la difficoltà di definire variabili risultato e variabili di controllo. In molti casi, i registri della disoccupazione contengono solo poche informazioni sulla condizione occupazionale degli iscritti dopo l'uscita dalla condizione di disoccupazione o dopo il completamento dei programmi di inserimento lavorativo.

In **Ungheria**, i dati amministrativi degli SPI contengono informazioni sulla condizione occupazionale relativa ai soli partecipanti ai programmi di politica attiva del lavoro (PAL) e coprono unicamente un periodo di 180 giorni dal completamento del programma. Inoltre, non sono disponibili dati sulla condizione occupazionale degli iscritti ai registri che non partecipano a programmi PAL. Una delle possibili soluzioni per ovviare alla mancanza di dati consiste nell'utilizzare come variabile risultato la sola condizione occupato/in cerca di occupazione, disponibile anche per i non-partecipanti ai programmi che sono in cerca di lavoro. Una persona che risulti ancora iscritta nel registro può essere considerata "in cerca di occupazione", mentre si può ipotizzare che chi è uscito dal registro

risulti occupato. Ciononostante, compito della ricerca scientifica è non perdere di vista che l'uscita dal registro può avere molte cause diverse dall'averne un'occupazione, dal momento che vi sono molti inattivi tra coloro che non risultano più iscritti ai registri. Nell'esercizio di valutazione svolto in Ungheria, la questione è stata risolta utilizzando come variabile risultato i dati sull'occupazione, contenuti nel database degli SPI integrato con i dati della previdenza sociale. La situazione è analoga negli SPI in **Polonia** dove, per rilevare la condizione occupazionale, ci si affida al criterio della presenza/assenza nei registri della disoccupazione e/o partecipazione a programmi di politica attiva del lavoro.

- **È possibile identificare i partecipanti al programma o solo coloro che appartengono al gruppo degli eleggibili?**

In molti casi il database non contiene informazioni che consentano di identificare i partecipanti ai programmi, ma solo gli individui che appartengono al gruppo degli eleggibili a partecipare al programma. In questo caso, i ricercatori possono stimare l'impatto del programma solo sulla sottopopolazione degli eleggibili, ma non possono identificare l'effetto del programma sugli effettivi partecipanti. Usando la terminologia della valutazione, i ricercatori non sono nelle condizioni di poter stimare l'effetto medio del trattamento sulla popolazione né quello sui trattati – rispettivamente *Average treatment Effect (ATE)* e *Average Treatment Effect on Treated (ATT)* - ma unicamente l'effetto medio del trattamento sugli individui eleggibili, cioè l'Intention To Treat effect (ITT)⁹.

Si tratta di un problema minore se si vuole valutare una misura universalistica (ad es., voucher destinati alla totalità o alla quasi-totalità della popolazione dei giovani disoccupati oppure obbligatori per alcuni gruppi). Tuttavia, la distinzione può assumere rilievo maggiore laddove la partecipazione risulti parziale o le motivazioni, sottostanti la mancata partecipazione, poco conosciute. Ad esempio, nel caso in cui l'effetto stimato sulla popolazione eleggibile risulti basso, il ricercatore potrebbe non essere in grado di distinguere se i deboli risultati siano causati da una bassa partecipazione al programma o da un debole effetto di quest'ultimo. Distinguere tra bassa partecipazione al programma (*take-up rate*) e scarsi effetti sui partecipanti al programma può assumere una notevole rilevanza fin dal disegno della politica¹⁰. Il primo caso richiederebbe infatti una maggiore attività di sensibilizzazione in fase di implementazione, mentre il secondo potrebbe derivare da una cattiva progettazione del programma.

Gli archivi SPI contengono di solito informazioni di dettaglio sulle politiche attive del lavoro e consentono l'identificazione dei partecipanti ai vari programmi¹¹. In **Ungheria**, il database degli SPI contiene informazioni di dettaglio sulle misure di politica attiva del lavoro, inclusa la tipologia specifica e le date di inizio e conclusione dei programmi. Nel caso in cui le misure siano accompagnate da politiche passive, come sussidi salariali, l'importo dei sussidi e i salari non sono tuttavia contenuti nel database, ma possono essere resi disponibili attraverso il *data-linkage* con gli archivi della previdenza sociale.

9 Dopo approfondito dibattito interno ai partners Italiani del progetto si è deciso di non abbandonare il modello teorico di riferimento mainstream, sebbene il carattere universalistico delle misure introdotte in Italia nel 2015 porterebbero a considerare l'effetto stimato sull'intera popolazione eleggibile come un ATT.

10 Questo problema può essere attenuato utilizzando informazioni esterne per approssimare il tasso di assorbimento

11 Tuttavia, l'eleggibilità alla partecipazione a ciascun programma deve essere generalmente approssimata sulla base dei dati relativi alle caratteristiche del background delle persone in cerca di lavoro.

In **Spagna** entrambi gli archivi amministrativi consentono l'identificazione del contratto di tirocinio, oggetto della valutazione svolta dal Gruppo di ricerca spagnolo. In entrambi i database utilizzati per l'esercizio valutativo svolto, cioè sia nel Campione Continuo delle Storie Lavorative (in inglese, *Continuous Sample of Work Histories – CSWH*), sia nel registro della disoccupazione degli SPI, tutti gli individui che sottoscrivono un contratto di tirocinio sono perfettamente identificati, con un corredo informativo che comprende data di inizio e conclusione del contratto. Nei registri sono inoltre disponibili le informazioni relative alle altre tipologie contrattuali, che sono dunque utilizzabili per identificare un gruppo di controllo. Per quanto riguarda l'**Italia**, l'archivio amministrativo SISCO, "Sistema Informativo Statistico delle Comunicazioni Obbligatorie" del Ministero del Lavoro e delle Politiche Sociali, non è strutturato per raccogliere informazioni aggiuntive relative alla partecipazione degli individui censiti a programmi di politica attiva del lavoro. La strategia italiana di identificazione degli effetti utilizza, pertanto, come variabile risultato il contratto a tempo determinato o indeterminato per stimare l'effetto della presenza degli incentivi all'assunzione e della maggiore flessibilità in uscita nel mercato del lavoro sul totale dei nuovi contratti di lavoro dipendente eleggibili al trattamento. In **Polonia** i dati degli SPI contengono dettagli sui programmi di politica attiva del lavoro a cui l'individuo registrato ha partecipato, come la data di inizio e fine del programma e le informazioni personali.

- **Qual è il livello di dettaglio presente nel set di variabili che descrivono le caratteristiche osservabili?**

L'insieme delle variabili che descrivono le caratteristiche osservabili utilizzabili come covariate nel modello influisce sia nella definizione della strategia di identificazione sia per l'analisi dell'eterogeneità degli effetti sulle diverse sottopopolazioni. Di conseguenza, prima di elaborare la strategia di valutazione, è fondamentale mappare la disponibilità e l'accessibilità al set di variabili da utilizzare come covariate. Negli archivi amministrativi di microdati il genere, l'età e alcune variabili legate al territorio sono di solito disponibili; tuttavia, a eccezione delle informazioni di base, gli archivi amministrativi differiscono per disponibilità di dati utilizzabili come covariate.

Il livello di istruzione è solitamente un'informazione cruciale perché ha una grande influenza sui risultati in termini di condizione occupazionale (*outcome*) e di fatto rappresenta una delle variabili più importanti nella valutazione delle politiche del mercato del lavoro.

Il livello di istruzione è una delle variabili disponibili negli archivi SPI in **Ungheria**, **Polonia** e nel registro SISCO in **Italia**. Mentre in **Spagna**, il dato viene aggiornato ogni dieci anni e proviene dal censimento.

Alcune utili variabili aggiuntive sono quelle sulle caratteristiche dei nuclei familiari (ad esempio, numero di conviventi, in **Spagna**), le informazioni sull'accesso a diversi sussidi (come ad esempio, i contributi in caso di maternità, in **Ungheria**) e l'informazione sulla presenza di un figlio minore nella famiglia (**Polonia**).

I dati amministrativi sui contratti di lavoro tratti dagli archivi SISCO in **Italia** contengono informazioni sul settore economico di attività (NACE). Dati simili non sono disponibili, ad esempio, nell'archivio della previdenza sociale in **Spagna**.

Oltre alle variabili relative alle caratteristiche osservabili degli individui, ai contratti e alle caratteristiche dei datori di lavoro, può essere utile ricorrere anche a variabili macroeconomiche tratte da fonti esterne per “controllare” i risultati in funzione delle condizioni ambientali del segmento di mercato del lavoro. Ad esempio, in **Italia** è stato utilizzato il tasso di crescita del valore aggiunto, in ciascun settore economico di attività, del trimestre successivo a quello di riferimento. Questa variabile è stata usata per isolare l’effetto delle politiche dal ruolo giocato dalla performance economica settoriale, indicatore utilizzato come proxy delle aspettative dei datori di lavoro al momento delle decisioni di assunzione.

L’ampiezza dell’intervallo temporale coperto dai dati, comprese le finestre temporali prima e dopo l’inizio del programma, può essere rilevante. In alcuni casi può essere utile includere i dati relativi ai periodi precedenti l’introduzione o la riforma delle politiche in modo da poter confrontare le condizioni precedenti con quelle successive utilizzando un modello *Difference in differences (diff-in-diff)*. In altri casi le variabili provenienti da lunghe serie storiche possono essere usate come controlli. Per esempio, in **Ungheria** l’intera storia occupazionale degli individui è disponibile nelle basi di dati della previdenza sociale. Per questo, usando queste informazioni, il ricercatore può essere in grado di ricostruire l’intera storia occupazionale di un individuo per includere nell’analisi, ad esempio, la sua esperienza lavorativa o le caratteristiche motivazionali.

3. IL PROCESSO DI VALUTAZIONE

3.1. Scelta delle misure specifiche da valutare

Molte politiche del lavoro prevedono una pluralità di misure. Chi le valuta (così come chi le implementa) deve decidere quali delle misure previste possono costituire oggetto di valutazione. La valutazione può riguardare la totalità o solo alcune misure e sotto-misure, oppure focalizzarsi su un insieme definito di misure omogenee tra loro. La scelta è vincolata dall'insieme di regole e procedure previste per l'implementazione delle misure. Ad esempio, se la politica prevede diverse combinazioni possibili di misure da applicare per rispondere a specifiche esigenze dei disoccupati, non è possibile valutare separatamente le singole misure (a meno che non siano disponibili informazioni ricche di dettagli sui bisogni individuali degli individui in cerca di occupazione). In **Ungheria**, ad esempio, l'oggetto della valutazione sono: un incentivo salariale disponibile durante il "periodo di prova" di 90 giorni che è spesso seguito da un sussidio che integra il salario fino ad un periodo massimo di ulteriori 8 mesi. Nei casi in cui il programma preveda l'uso combinato dei due strumenti, non considerare questa combinazione di misure provocherebbe una distorsione per eccesso della stima del tasso di occupazione a 6 mesi dal completamento del programma di lavoro in prova, dato che per coloro che beneficiano del sussidio a lungo termine il tasso di occupazione risulta prossimo al 100%. Diversamente, in **Polonia**, le politiche del mercato del lavoro offerte dagli SPI sono costituite, nella maggioranza dei casi, da un'unica misura, cioè, alternativamente, o formazione o tirocini, o incentivi all'imprenditorialità. Questi interventi molto raramente sono offerti in combinazione tra loro, nonostante l'assistenza nella ricerca di lavoro sia una misura disponibile per tutti.

La scelta della politica da valutare influenza la scelta e la definizione del trattamento e l'identificazione del possibile gruppo di controllo. Se gli effettivi partecipanti non possono essere distinti nel database dagli individui eleggibili, la valutazione in modo controfattuale di una specifica misura non può stimare l'effetto medio del trattamento sui trattati (Average Treatment effect on Treated, ATT) ma unicamente sugli eleggibili (ITT, vedi sopra).

È cruciale esaminare e chiarire l'intero disegno e tutti i dettagli delle misure per individuare il perimetro di una eventuale valutazione. Le domande che emergono a questo proposito sono le seguenti: qual è il problema che la misura introdotta tenta di risolvere e quali sono a tal riguardo le informazioni che gli estensori del programma hanno a disposizione? Il programma è obbligatorio o volontario? Quali sono i criteri di eleggibilità alla misura? Qual è la durata e la copertura territoriale e temporale? Cosa sappiamo riguardo agli obiettivi, la sequenza e l'ampiezza del programma? Quali sono i dettagli da chiarire per capire i punti di forza e di debolezza della *policy*? Quali sono le combinazioni di misure previste e in che modo possono influenzare la valutazione e l'interpretazione dei risultati?

In **Italia** si è scelto di valutare, ad esempio, l'effetto combinato di due politiche: gli incentivi alle imprese, sotto forma di esonero contributivo totale per i datori di lavoro in caso di assunzione a tempo indeterminato, insieme all'incremento della flessibilità contrattuale "in uscita" del nuovo "contratto di lavoro a tempo indeterminato a tutele crescenti". Quest'ultimo, infatti, crea aspettative nei datori di lavoro di minori costi da affrontare in caso di licenziamento anticipato dei lavoratori. Le due politiche sono state valutate congiuntamente in quanto introdotte entrambe nel 2015, con un lasso di tempo tra la decorrenza delle due misure ritenuto trascurabile ai fini della valutazione. L'accesso agli sgravi contributivi avviene, inoltre, a richiesta, mentre il nuovo istituto contrattuale a tutele crescenti si applica a tutti i nuovi avviamenti di rapporti di lavoro a tempo indeterminato.

3.2. Scelta delle variabili risultato

Il punto di partenza per la scelta della variabile risultato da analizzare e del metodo di stima da utilizzare consiste nel mettere a fuoco il meccanismo causale e l'obiettivo del programma. Quali sono gli effetti attesi? Qual è l'orizzonte temporale degli impatti? Vi sono effetti collaterali? Dopo essersi fatti un'idea chiara in merito agli effetti del programma, sorgono le seguenti domande:

- Qual è il risultato d'interesse/variabile risultato/outcome: condizione occupazionale, tipologia contrattuale, occupato presso la stessa azienda in cui si è svolto il tirocinio, impiego sul mercato del lavoro primario, salario (cumulato) ecc.
- Quale orizzonte analizzare: a breve o lungo termine (oltre un anno dalla fine del programma)? In linea di principio il ricercatore potrebbe analizzare i risultati su diversi orizzonti temporali e valutare come cambia l'impatto a seconda della lunghezza dell'orizzonte considerato. Nella valutazione dei programmi di sussidio salariale, il ricercatore deve tener conto del fatto che tale sussidio può essere associato a un obbligo di assunzione successiva (non sovvenzionata), per cui deve considerare l'effetto solo al termine del periodo di obbligo di assunzione.

Le considerazioni teoriche potrebbero però essere fortemente vincolate dalla disponibilità dei dati. In molti casi, i dati sugli effetti a lungo termine non sono disponibili e i dataset SPI solitamente non contengono i dati sui salari.

I sussidi salariali di solito hanno lo scopo di aumentare la probabilità di trovare un (buon) lavoro, pertanto le variabili risultato più rilevanti sono la probabilità di essere assunti in un certo lasso di tempo dal completamento del programma ed il reddito da lavoro. Ad esempio, nell'esercizio di valutazione **ungherese**, l'obiettivo del programma di incentivi nei 90 giorni di prova sul lavoro, consiste nel favorire l'ingresso nel mondo del lavoro e ridurre il rischio associato alle performance di un giovane senza alcuna esperienza lavorativa. In linea di massima, ci si attende che la partecipazione al programma migliori le prospettive di impiego in quanto possa aumentare la probabilità di essere assunti dopo il completamento del programma e, idealmente, aumenti le probabilità di ottenere salari migliori. Ci si attende quindi che la partecipazione al programma possa accrescere le prospettive di trovare un lavoro meglio retribuito, arricchendo di esperienza e valorizzando il CV dei giovani. In questo caso, le principali variabili risultato sono lo stato occupazionale misurato a distanza di sei mesi, dopo uno, due o più anni dalla conclusione del programma e il differenziale salariale cumulativo misurato sullo stesso orizzonte temporale. Tuttavia, la disponibilità dei dati limita il panorama della scelta delle variabili risultato. In questo caso specifico non sono disponibili, ad esempio, i risultati a lungo termine, in quanto la politica è stata introdotta nel 2015 e il campione disponibile comprende solo il triennio 2015-2017.

L'utilizzo dei dati SPI per valutare le politiche attive del mercato del lavoro (PAML) in **Polonia** offre una scelta limitata di variabili risultato. È possibile osservare se un individuo è stato reinserito nel registro dei disoccupati dopo un certo numero di mesi. Tuttavia, non solo gli occupati, ma anche gli inattivi e i partecipanti al programma sono esclusi dal registro. La variabile risultato si limita all'alternativa: al "non figurare" nel registro e "non partecipare" ad alcuna PAML¹².

12 Anche l'orizzonte temporale del risultato valutato è importante. Ad esempio, il sussidio imprenditoriale in Polonia richiede il proseguimento dell'attività per almeno un anno, pena la restituzione del sussidio in caso di durata inferiore, in modo che la sua reale efficacia possa essere valutata solo dopo un anno.

L'obiettivo del contratto di tirocinio **spagnolo** e degli incentivi all'assunzione **italiani** è quello di aumentare la stabilità e la durata dell'occupazione, oltre all'incidenza complessiva dei contratti a tempo indeterminato sul totale del lavoro dipendente. Le variabili risultato utili sono quindi quelle identificano tipologia e durata dei contratti dei giovani coinvolti nel programma. La principale variabile risultato nel caso della valutazione **spagnola** è la probabilità di trovare occupazione nella stessa azienda al termine del contratto di tirocinio. La seconda variabile risultato è la probabilità di permanenza in quell'azienda con un contratto di lavoro dipendente a tempo indeterminato. La terza variabile risultato è la probabilità di trovare un contratto a tempo indeterminato in un'altra azienda dopo il contratto di tirocinio, come indicatore di un migliore posizionamento nel mercato del lavoro come effetto indiretto del tirocinio. L'obiettivo degli sgravi contributivi e della riduzione dei costi di licenziamento, proprio delle politiche oggetto della valutazione **italiana**, è simile: aumentare qualità e durata dei contratti di lavoro. La principale variabile risultato della valutazione delle due politiche italiane è la percentuale di nuovi contratti di lavoro dipendente a tempo indeterminato sul totale dei nuovi contratti di lavoro dipendente avviati nel 2015.

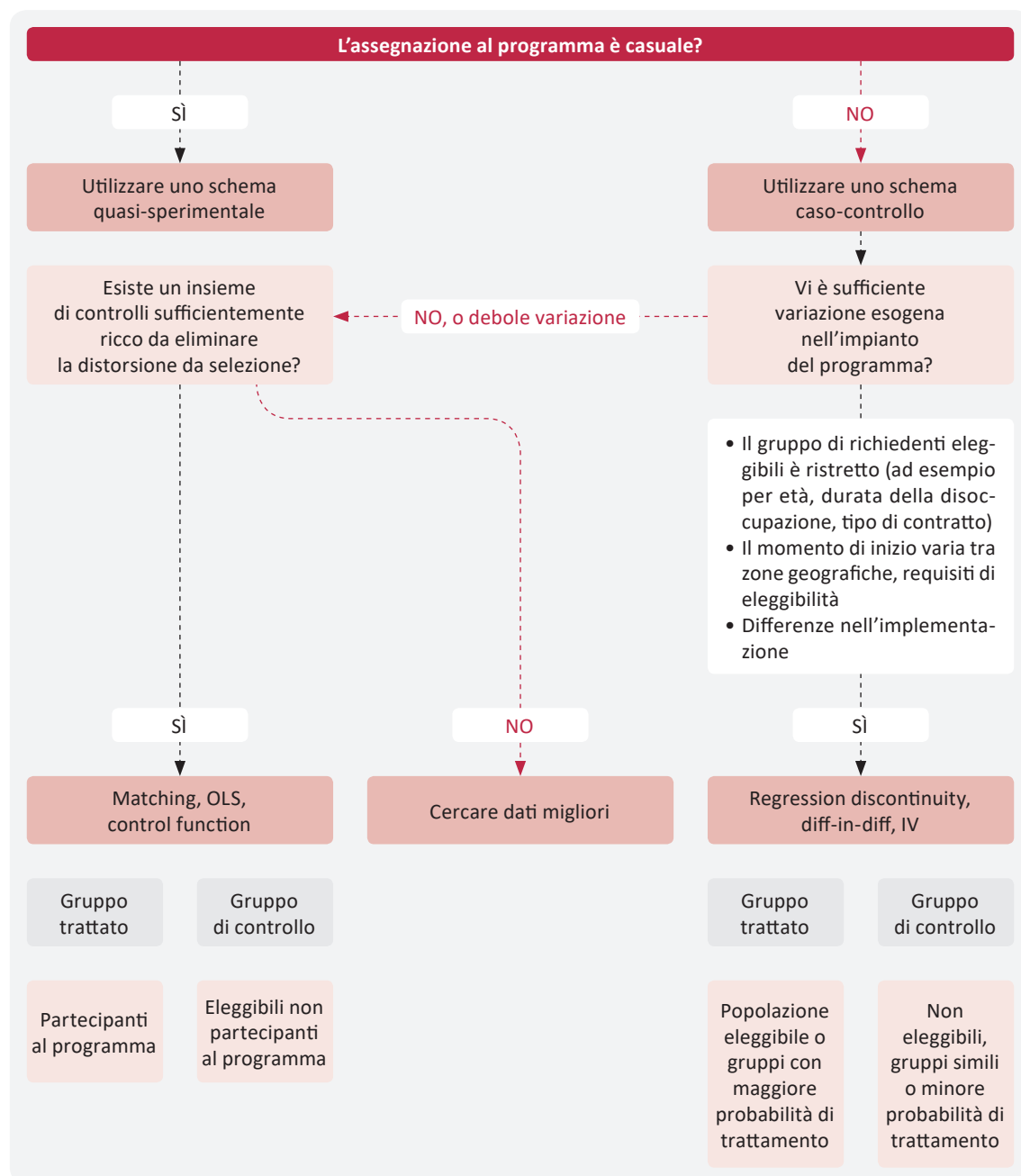
3.3. Strategia di identificazione: scelta del metodo della valutazione d'impatto controfattuale¹³

La difficoltà principale nel misurare l'impatto causale di qualsiasi politica è che il mondo controfattuale, in cui è lo stesso individuo che non riceve il trattamento, non è osservabile. Il ricercatore deve dedurre la condizione controfattuale dai soli dati osservazionali disponibili.

In un mondo ideale, la valutazione delle politiche del mercato del lavoro dovrebbe basarsi sulla simulazione di una situazione sperimentale, che prevede una selezione con metodo casuale dei partecipanti dal bacino di individui eleggibili a partecipare al programma. In tutte le situazioni in cui non è possibile garantire un quadro sperimentale, il processo di selezione non avviene con metodo casuale e i non-partecipanti possono risultare diversi dai partecipanti in termini di caratteristiche sia osservabili sia non osservabili. Da ciò scaturiscono due fonti di distorsione di base. In primo luogo, gli effetti delle politiche, ad esempio le prospettive occupazionali e salariali dei non partecipanti, avrebbero potute essere diverse rispetto a quelle dei non partecipanti anche in assenza della partecipazione al programma (condizioni di partenza, o *baseline*, in inglese). Il programma potrebbe inoltre avere un impatto diverso sui partecipanti rispetto ai non partecipanti, nel caso in cui questi ultimi partecipassero al programma (effetto di trattamento eterogeneo). In mancanza di condizioni sperimentali, il ricercatore deve applicare un quadro quasi-sperimentale per simulare le condizioni sperimentali, tenendo in considerazione che la selezione non è avvenuta con metodo casuale all'interno del programma (si veda la Figura 3-1).

13 Gli elementi basilari dei principali concetti e metodi di valutazione sono descritti nell'Appendice 2.

Figura 3-1: Diagramma della scelta della strategia di identificazione sulla base delle caratteristiche del programma e della sua implementazione



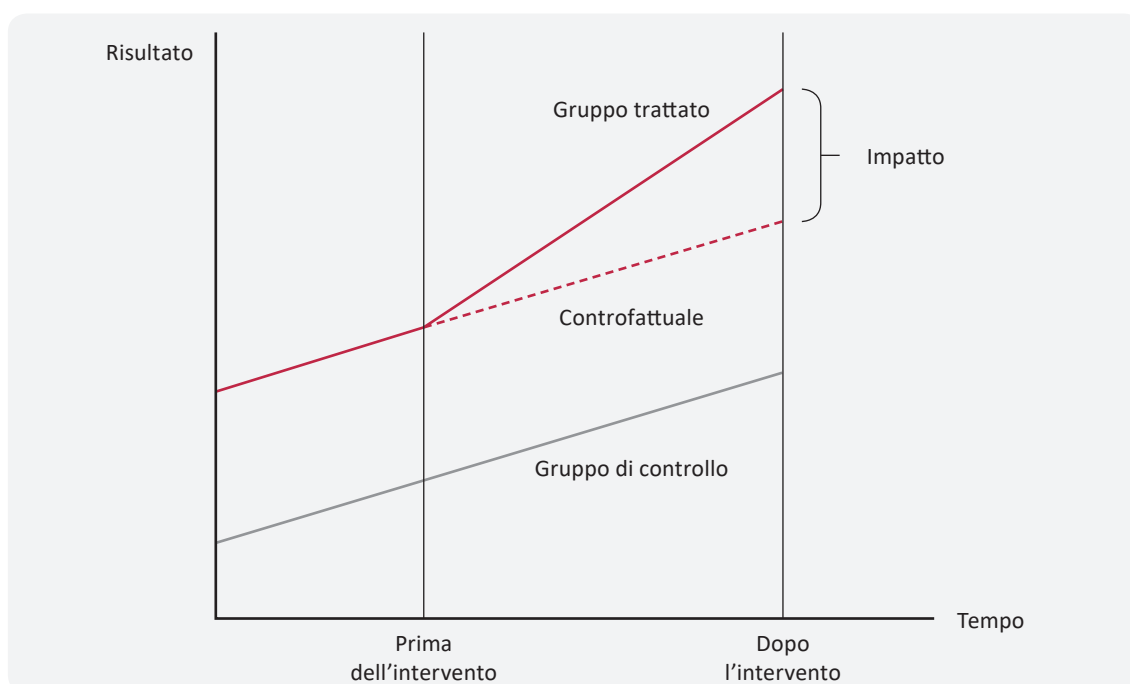
Ci sono due direzioni possibili che possono essere seguite per garantire un quadro quasi sperimentale.

Un'opzione consiste nel trovare una qualche **variazione esogena rispetto alla variabile risultato nel quadro del programma (le "regole interne")** e sfruttare questa esogeneità. L'obiettivo di tale esercizio è di differenziare gli individui che per motivi esogeni sono esclusi dal programma sulla base di una o più restrizioni alla partecipazione definite nella regolamentazione del programma, da quelli che vengono per motivi esogeni selezionati per partecipare al programma o che sono eleggibili a partecipare. Di seguito vengono riportati gli esempi più comuni di restrizioni di natura esogena da sfruttare per identificare l'effetto della politica:

- *Bacino di candidati eleggibili (ad es. limite d'età, interruzione del periodo di disoccupazione o presenza/assenza di un precedente contratto di lavoro dipendente a tempo indeterminato)*

La strategia **italiana** di identificazione degli effetti sfrutta la possibilità offerta da uno dei criteri di selezione degli eleggibili, secondo cui potevano beneficiare dell'incentivo all'assunzione solo coloro che nei 6 mesi precedenti all'assunzione non risultavano occupati con contratto di lavoro dipendente a tempo indeterminato. La definizione dei due gruppi è stata realizzata utilizzando questo criterio di eleggibilità. Il gruppo degli eleggibili comprende tutti i contratti di lavoro di giovani assunti nel 2015 e non precedentemente occupati con un contratto a tempo indeterminato nei sei mesi precedenti. Mentre il gruppo di controllo (non eleggibili) comprende i contratti di lavoro di giovani assunti nel corso del 2015 che nei 6 mesi precedenti l'inizio del nuovo contratto risultavano occupati con contratto di lavoro a tempo indeterminato e che per tale ragione non possono beneficiare degli sgravi contributivi previsti (per una rappresentazione grafica del modello *Difference-in-differences* si veda la Figura 3-2). La strategia di identificazione si basa su un modello *Difference-in-differences* applicato ai contratti registrati nell'archivio delle comunicazioni obbligatorie (SISCO) tra il 1° gennaio 2015 e il 31 dicembre 2015 rispetto a quelli registrati nel 2014, tra gli eleggibili e i non eleggibili, per stimare la percentuale aggiuntiva dei nuovi contratti di lavoro a tempo indeterminato registrati nel corso del 2015 che non sarebbero stati firmati in assenza delle due politiche.

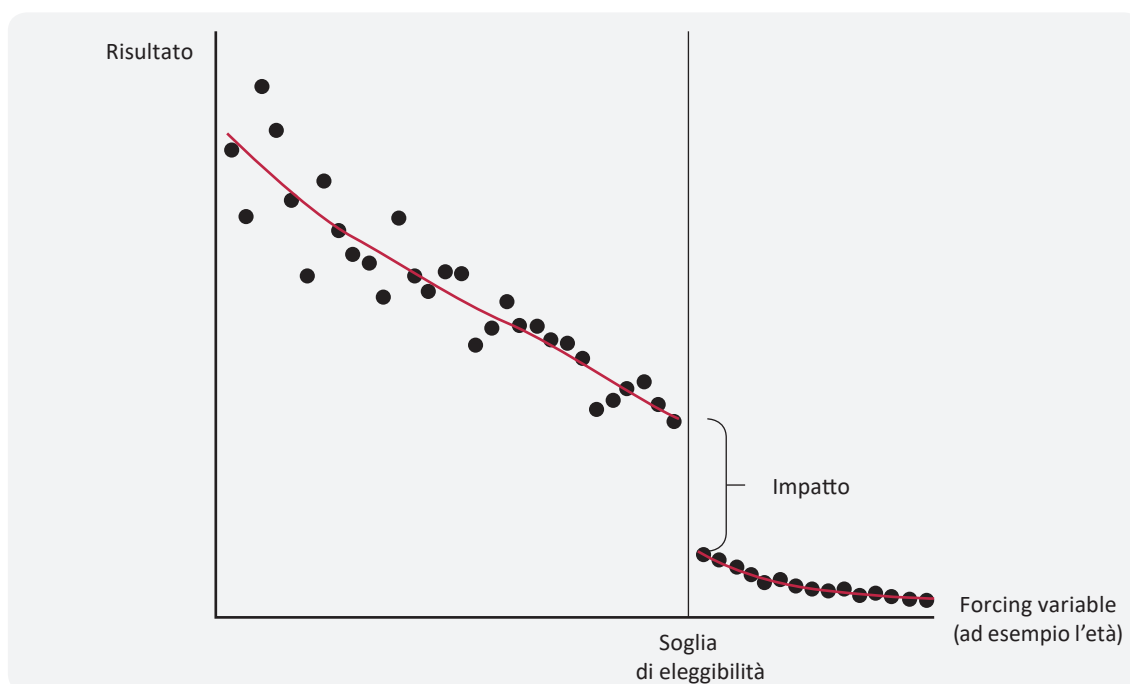
Figura 3-2: Illustrazione: metodo *difference-in-differences* (DiD)



Poiché il database non riporta informazioni dirette sugli effettivi beneficiari degli incentivi e consente di identificare solo i soggetti eleggibili, la valutazione stima l'effetto della politica sul gruppo degli eleggibili (Intention To Treat effect, ITT).

La valutazione dei programmi di Garanzia Giovani, essendo destinati ai giovani di età inferiore ad una soglia determinata, può sfruttare il limite di età in una cornice di *Regression Discontinuity Design (RDD)*, come nel caso della valutazione **polacca**. In Polonia il programma di sussidi salariali della durata di 12 mesi era disponibile solo per i disoccupati di età inferiore ai 30 anni, il che ha consentito di sfruttare quel limite di età per applicare un modello “sharp” di *Regression Discontinuity Design*. Anche se successive norme hanno previsto soglie di accesso in prossimità dei 30 anni di età, in primo luogo, la gamma di programmi a disposizione (voucher di formazione, voucher di riassegnazione e voucher di job trial) per i lavoratori disoccupati sotto i 30 anni di età è più ampia e in alcune Politiche attive nel mercato del lavoro (PAML) le regole possono essere diverse (ad es. i job trial possono durare fino a 12 mesi in caso di lavoratori più giovani e fino a 6 mesi in caso di lavoratori più anziani). In secondo luogo, le fonti di finanziamento in parte differiscono per i disoccupati più giovani e per quelli più anziani. Nel caso dei lavoratori più giovani, l'erogazione avviene sulla base di un “Programma Operativo” Garanzia Giovani gestito a livello centrale, mentre nel caso dei lavoratori più anziani è finanziata attraverso “Programmi Operativi Regionali” gestiti a livello di autorità amministrative locali (i Voivodati). Ciò può comportare differenze nell'attuazione e nella disponibilità dei programmi tra i lavoratori disoccupati al di sotto e al di sopra dei 30 anni di età. La strategia di identificazione ricorre al fatto che, prima dell'introduzione del programma di sussidi salariali, tutte le differenze a livello di istituti normativi tra i destinatari del trattamento e il gruppo di controllo, erano preesistenti con l'eccezione del programma di sussidi salariali che è oggetto della valutazione. È stato dunque possibile combinare un approccio di *Regression discontinuity Design (RDD)* con un modello *Difference-in-differences (DiD)*. Una premessa importante di questo approccio è che i risultati si basano sul confronto tra il gruppo dei trattati e il gruppo di controllo solo in prossimità della soglia di discontinuità che separa i due gruppi (*cut-off*). Ciò comporta, al contempo, dei limiti alla validità esterna dei risultati ottenuti e richiede una grande cautela nel generalizzare gli effetti stimati ad altre fasce d'età.

Figura 3-3: Illustrazione: regression discontinuity design (RDD)

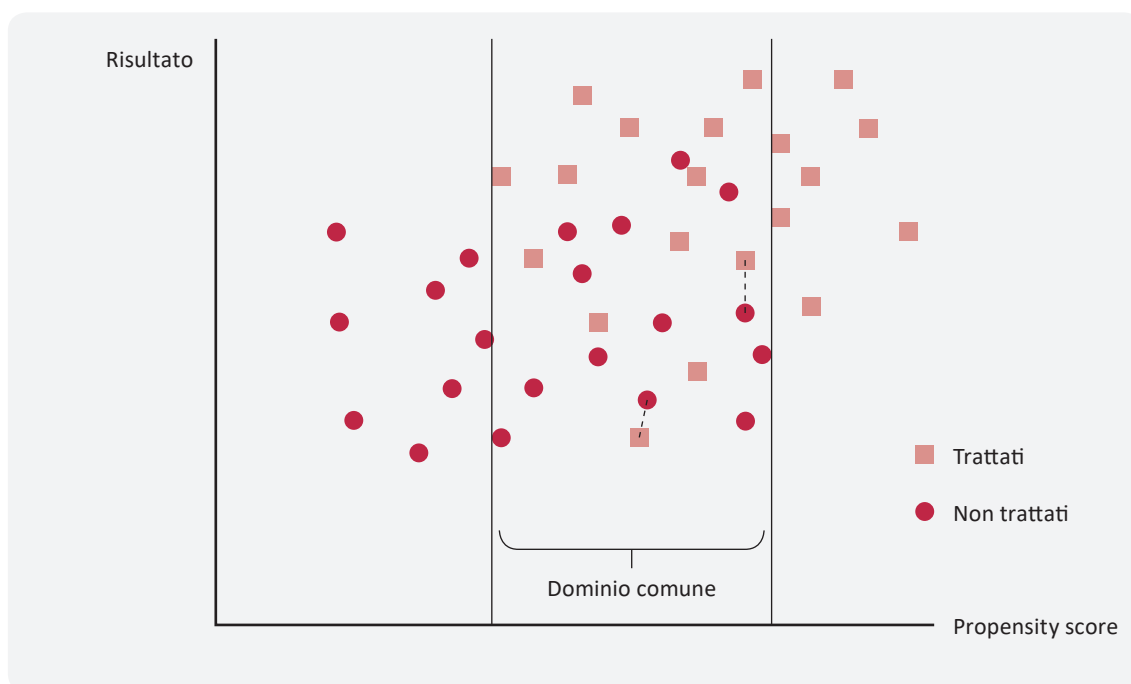


- *Eterogeneità regionale nell'intensità del programma, data di inizio, numeri obiettivo ecc.*

In **Ungheria**, il programma Garanzia Giovani nella Regione Centrale è stato lanciato nell'ottobre 2015, con un ritardo di 9 mesi rispetto alle altre regioni. La ragione di questo ritardo è che la Regione Centrale dell'Ungheria, a differenza delle altre sei regioni ungheresi, non è una Regione dell'obiettivo "Convergenza" dei Fondi Strutturali dell'UE, pertanto le fonti finanziarie del programma sono alquanto diverse. Tale ritardo è dovuto a motivi amministrativi e può essere considerato esogeno. Il programma avviato nella "Regione centrale" dell'Ungheria presenta elementi e requisiti di eleggibilità identici al resto del Paese. L'eterogeneità esogena dovuta alla data di avvio consente di applicare il metodo della *Difference-in-differences* che viene di seguito descritto. Il gruppo dei trattati è composto da individui eleggibili al programma in cerca di lavoro nelle "Regioni Convergenza", il gruppo di controllo è composto da individui in cerca di lavoro che si trovano nella "Regione Centrale" dell'Ungheria. Il trattamento e i gruppi di controllo vengono messi a confronto nei 9 mesi prima e dopo l'inizio del programma nelle Regioni convergenza, cioè nel gennaio 2015. Un ovvio limite di questa strategia è che, sfruttando l'implementazione graduale, si confrontano sostanzialmente il gruppo dei trattati e il gruppi di controllo per un breve lasso di tempo, in questo caso, di 9 mesi.

In letteratura sono riportati diversi esercizi di valutazione che tengono conto sia della presenza di cosiddette "regole interne" esogene sia di fattori legati all'eterogeneità del rigore metodologico applicato da coloro che operano la selezione dei partecipanti. Esistono situazioni in cui le differenze nell'eleggibilità o nella probabilità di essere trattati si basano su differenze nelle regole interne informali tra agenzie SPI, province, esaminatori ecc. (Maestas Mueller, 2013). Le differenze nella probabilità di essere trattati in base alle regole interne possono essere sfruttate in un approccio che ricorre alle cosiddette "variabili strumentali".

Figura 3-4: Illustrazione: propensity score matching



Un'altra indicazione importante per la strategia di identificazione degli effetti consiste nel **rilevare le eventuali distorsioni da selezione esistenti controllando attraverso una ricca serie di caratteristiche osservabili** (covariate), la cronologia occupazionale e le caratteristiche personali. La distorsione da selezione (*selection bias*) può essere ridotta utilizzando una serie di variabili che descrivono le caratteristiche individuali sia dei partecipanti al programma sia dei non eleggibili, formando così un gruppo di non partecipanti che assomigliano il più possibile ai partecipanti. In questo caso, la strategia di identificazione si basa sul presupposto che, a seconda delle covariate osservate, sia garantito il rispetto dell'indipendenza condizionata delle variabili (*Conditional Independence Assumption - CIA - o unconfoundedness condition*), ovvero la rettifica per differenze nelle variabili osservate rimuove le distorsioni nel confronto tra le unità trattate e quelle di controllo. Tuttavia, la preoccupazione principale è che alcune caratteristiche non osservabili, in primo luogo la motivazione e la capacità, possano essere correlate alle variabili risultato. Ad esempio, sono maggiori le probabilità che un giovane più motivato si candidi a partecipare a un programma di politica attiva del mercato del lavoro, ma presumibilmente le sue prospettive occupazionali sono migliori anche in assenza della politica del mercato del lavoro. Una simile argomentazione potrebbe valere anche per altre competenze e abilità che non vengono "catturate" dal livello di istruzione. Il ricercatore può tentare di individuare variabili che rilevino abilità e motivazioni non osservate. Ad esempio, i punteggi più alti nei test di competenza o nei programmi di formazione e la storia occupazionale potrebbero indicare una maggiore motivazione.

Tuttavia, come regola di base, tali fattori non osservati determinano una distorsione da selezione verso l'alto dell'effetto del trattamento, per cui l'effetto stimato deve essere considerato come un limite superiore.

In **Ungheria**, gli effetti del programma "90 giorni di prova" vengono valutati rispetto ai risultati dei partecipanti di altri programmi, quali i programmi di formazione e di lavoro pubblico. Si assume che la distorsione derivante da caratteristiche non osservate sia inferiore, se il gruppo di controllo è composto da partecipanti ad altri programmi invece che da non partecipanti, tuttavia, anche la validità esterna dei risultati è inferiore. I risultati mostrano l'efficacia relativa del programma rispetto ad altri programmi. Ciononostante, non è possibile trarre alcuna conclusione sull'effetto del programma rispetto ad una situazione in cui l'individuo in cerca di lavoro non partecipa ad alcun programma, perché i Servizi per l'impiego cercano di assegnare ad un programma ciascun individuo in cerca di lavoro presente nei loro registri. Ciò implica che il gruppo di giovani che non partecipano ad alcun programma e che rimangono iscritti nei registri non rappresentano probabilmente un gruppo di controllo valido, in quanto si tratta o di coloro che hanno minore motivazione o, al contrario, di coloro con le maggiori possibilità di reinserimento lavorativo. Si prevede che persino coloro che partecipano a programmi di "pubblico impiego" nascondano comunque una certa eterogeneità in termini di prospettive del mercato del lavoro e di motivazione. Il livello di istruzione e la storia lavorativa completa di tutti gli individui presenti nel database permette di attenuare l'eterogeneità non osservata. Il secondo esercizio di valutazione **polacco** utilizza la medesima strategia per valutare l'efficacia relativa di altre misure quali: formazione *on-the-job*, vouchers per la formazione *on-the-job*, formazione in classe, voucher per la formazione in classe, integrazioni al reddito e programmi di lavoro pubblico.

La valutazione dei contratti di tirocinio in **Spagna** mette invece a confronto i risultati di un gruppo ben definito di partecipanti al programma e di una platea di eleggibili non partecipanti con modelli statistici basati su tecniche di regressione. Il gruppo dei trattati è composto da giovani al di sotto

dei 30 anni che hanno firmato un contratto di tirocinio, mentre il gruppo di controllo è composto da giovani disoccupati eleggibili ma che hanno firmato regolari contratti di lavoro temporanei. Il campione tratto dai registri della previdenza sociale consente di costruire le serie storiche delle variabili legate all'occupazione, come: la permanenza nel mercato del lavoro, il numero di precedenti esperienze lavorative, il numero dei precedenti datori di lavoro e la durata degli episodi di disoccupazione. Tali variabili potrebbero cogliere una parte significativa delle differenze tra il gruppo di controllo e il gruppo dei trattati. I gruppi trattati e di controllo vengono inoltre definiti in modo tale da consentire di ottenere una base informativa pulita, eliminando dalla stessa i casi particolari. Ad esempio, l'attenzione si concentra solo su coloro che sono appena entrati nel mercato del lavoro, poiché l'obiettivo del contratto di tirocinio (IC) è quello di fungere da prima esperienza di lavoro e aumentare le probabilità di permanenza all'interno della stessa organizzazione. A ciò si aggiunge il confronto tra gli individui al primo ingresso nel mercato del lavoro con un contratto di tirocinio e coloro che firmano un contratto di tirocinio avendo già alle spalle precedenti esperienze di lavoro. Questo permette di stabilire se l'impatto del contratto di tirocinio dipende o meno da precedenti esperienze nel mercato del lavoro.

In **Italia** la valutazione d'impatto riguarda le differenze nella probabilità di essere assunti con contratto a tempo indeterminato tra lavoratori con contratti di lavoro dipendente eleggibili e non eleggibili agli incentivi, controllando per le caratteristiche di base dei due gruppi. Il presupposto principale alla base dell'identificazione di tale effetto tramite una metodologia *diff-in-diff* è che l'incentivo non influisca sull'outcome degli individui appartenenti al gruppo di controllo, ovvero che la policy, volta ad aumentare la probabilità di assumere con un contratto di lavoro a tempo indeterminato un lavoratore eleggibile non influisca sulla probabilità di occupazione con contratto a tempo indeterminato di un lavoratore non eleggibile. Le statistiche descrittive mostrano che nel caso italiano questa ipotesi non è valida, poiché è stata evidenziata una discontinuità (riduzione) nel 2015 nelle serie storiche della variabile risultato per il gruppo di controllo. Una spiegazione possibile di tale discontinuità è che alcuni datori di lavoro avrebbero spostato le loro preferenze di assunzione dai lavoratori non eleggibili ai lavoratori eleggibili, a causa del nuovo costo opportunità tra i due gruppi causato dalla presenza degli incentivi¹⁴. In altre parole, non si può escludere che il gruppo di controllo non sia stato condizionato dal trattamento stesso. Tale distorsione potenziale è corretta con una procedura in due fasi. Nella prima fase sono stati stimati e imputati al gruppo di controllo i valori teorici della variabile risultato - i valori che sarebbero stati registrati per il gruppo di controllo in assenza delle modifiche normative. Nella seconda fase si è stimato l'effetto dei sussidi alle assunzioni sul data base corretto utilizzando un modello Ordinary Least Squares OLS con interazioni tra il trattamento e le diverse classi di età, al fine di quantificare l'impatto sulle coorti più giovani.

3.4. Effetti eterogenei

Un aspetto importante della valutazione della politica è l'eterogeneità dell'effetto della stessa. Uno stesso programma ha un impatto diverso sui diversi gruppi, ad esempio per sesso, età, ripartizione regionale, in anni diversi, in condizioni macroeconomiche diverse, con diverse combinazioni di programmi? Comprendere le eterogeneità nell'effetto del trattamento aiuta a comprendere e valutare:

14 Come verrà spiegato nella sezione successiva, la questione è ulteriormente complicata dal fatto che l'introduzione della nuova normativa ha modificato il costo opportunità in modo diseguale all'interno dei diversi gruppi di età.

- la validità esterna dei risultati e la comparabilità con altri risultati di valutazione;
- l'efficienza del programma in termini di costi-benefici;
- i punti deboli del programma e le possibilità di miglioramento attraverso l'adeguamento dello stesso;
- la dimensione di genere del programma.

L'effetto del programma può variare a seconda del livello di istruzione dei suoi partecipanti. I decisori politici accolgono probabilmente con favore la constatazione che un sussidio salariale ha un impatto più forte su coloro che dispongono di un'istruzione inferiore, in quanto tale gruppo presenta presumibilmente esiti occupazionali meno favorevoli sul mercato del lavoro. Di conseguenza, se dalla valutazione emerge che coloro che hanno un'istruzione secondaria o inferiore beneficiano del sussidio, senza alcun impatto sui laureati, è al primo gruppo che il sussidio deve essere destinato. Poiché, inoltre, i costi di un sussidio salariale per un lavoro ben istruito e retribuito sono presumibilmente più elevati, il programma potrebbe migliorare l'efficienza dei costi per gli individui meno istruiti in cerca di lavoro. È un risultato da cui si evince che è al primo gruppo che i sussidi dovrebbero essere destinati. In **Ungheria**, il programma di “90 giorni di prova”, rispetto sia all’ “impiego pubblico” sia alla formazione, ha avuto un impatto più forte sui partecipanti con un'istruzione di base rispetto agli individui con un'istruzione secondaria o superiore. Ciò si è rivelato in linea con le aspettative, poiché dal punto di vista delle prospettive occupazionali, i giovani con un basso livello di istruzione sono il gruppo più vulnerabile. Al contempo, i partecipanti al programma di “90 giorni di prova” risultano in media più istruiti rispetto ai partecipanti ai programmi di “impiego pubblico”, il che indica che nel programma vengono selezionati i giovani disoccupati maggiormente occupabili. Ciò suggerisce che una maggiore priorità ai giovani disoccupati con un livello di istruzione più basso potrebbe aumentare l'effetto medio del trattamento, riducendo al contempo la perdita secca del programma (*dead weight loss*), che si registra quando il governo finanzia assunzioni che sarebbero state fatte comunque. I valutatori possono inoltre constatare che la durata stessa del programma influenza gli effetti. Tale constatazione può aiutare a determinare la durata ottimale di un programma simile in futuro.

Lo studio **spagnolo** copre l'analisi della dimensione di genere, giungendo alla conclusione che le donne hanno meno probabilità degli uomini di trovarsi disoccupate dopo il contratto di tirocinio. Tuttavia, per le altre questioni oggetto dell'analisi spagnola non si riscontrano differenze statistiche tra uomini e donne.

La valutazione **italiana** ha rilevato differenze significative sull'effetto dei programmi tra i giovani uomini e le giovani donne, a svantaggio di queste ultime, insieme ad un forte dualismo Nord-Sud nello sviluppo territoriale del Paese che è all'origine dell'eterogeneità dell'impatto delle due politiche sulla probabilità di essere assunti con contratto di lavoro a tempo indeterminato. La valutazione in **Polonia** ha rilevato che i corsi di formazione e i voucher per la formazione funzionano relativamente meglio per i maschi rispetto alle femmine. Gli impieghi tramite lavoro pubblico sono meno efficaci per persone con un grado di istruzione inferiore e per coloro che vivono in regioni con alta incidenza di disoccupazione. I voucher per la formazione sul posto di lavoro sono risultati più efficaci della formazione diretta sul posto di lavoro, indipendentemente dal genere, dal livello di istruzione o dal luogo di residenza del disoccupato.

4. QUESTIONI RILEVANTI PER I DECISORI POLITICI



© www.CartoonStock.com

"Mi domando: stiamo avendo alcun impatto?"

Finora abbiamo esaminato gli elementi necessari per garantire che i ricercatori possano stimare l'effetto causale delle politiche (validità interna). Tuttavia, esistono diverse questioni che i ricercatori devono affrontare quando discutono i risultati delle valutazioni con i decisori politici. Il nostro scopo qui è di segnalare tali questioni, di menzionare le esigenze di disponibilità e qualità delle banche dati per poter offrire potenziali soluzioni senza aver la pretesa di offrire una guida esaustiva su ciascuno di questi aspetti.

4.1. Interpretazione dei risultati

È molto importante che i ricercatori illustrino con precisione l'effetto che sono in grado di stimare nel proprio esercizio di valutazione. Molto spesso, il progetto di valutazione porta ad una stima con una "validità interna" dell'effetto medio del trattamento o del "trattamento medio locale", quindi su coloro che in quel momento beneficiano delle misure o un (piccolo) sottogruppo degli stessi; tuttavia, quando si presentano questi risultati ai decisori politici, è fondamentale sottolineare che questi possono essere generalizzati ad una popolazione più ampia solo sotto ulteriori e determinati presupposti. È quindi opportuno considerare anche la raccolta di dati sulle caratteristiche della popolazione allargata, per poter procedere a questa generalizzazione. Per contro, le valutazioni che mirano a stimare l'intention to treat, quindi l'effetto medio del trattamento dovuto alla possibilità di accedere ai benefici della misura, ovvero l'effetto del trattamento su tutti gli eleggibili, non solo su coloro che lo richiedono e ottengono l'accesso, potrebbero produrre risultati più utili per i decisori politici, se integrati con informazioni sui motivi della mancata partecipazione (per l'ITT) o se prevedessero la possibilità di stimare l'effetto causale sui non partecipanti alla misura (nel caso dell'effetto medio del trattamento sull'intera popolazione, *Average Treatment Effect*, ATE)

4.2. Relazione tra gli esiti del programma e la sua definizione e implementazione

Oltre a fornire una stima quantitativa degli effetti del programma, la valutazione dovrebbe fornire una descrizione dei fattori che hanno contribuito a determinare questi esiti. Per esempio, se la valutazione trova che il programma non ha avuto effetti significativi sull'occupazione è importante indagare se il fallimento della misura è stato causato da una cattiva definizione del programma (come la definizione erronea del gruppo di beneficiari) oppure da problemi di implementazione (come la mancanza di informazione o la selezione arbitraria dei partecipanti).

4.3. Validità esterna, comparabilità con altri studi

I lettori degli studi di valutazione, interessati ai processi di *policymaking*, vorranno confrontare tra loro i risultati dei diversi studi sulle misure oggetto della valutazione. Tra questi vi è probabilmente l'efficacia di misure simili per i giovani di regioni e paesi diversi o di altre misure (alternative) per i giovani e l'efficacia (potenziale) delle misure valutate su gruppi di beneficiari alternativi. A questi fini, i ricercatori devono esaminare una serie di questioni. Innanzitutto, come presentare i risultati in un modo che consenta il confronto con altri studi. In secondo luogo, devono descrivere in che modo i risultati possano essere generalizzabili in altre situazioni, al fine di limitare le estrapolazioni semplicistiche e attirare l'attenzione sui limiti di tali operazioni. In terzo luogo, i ricercatori devono anche documentare chiaramente gli elementi più rilevanti del disegno di *policy* delle misure valutate, nonché le condizioni economiche prevalenti, in modo tale che altri soggetti interessati possano comprendere come queste possano aver influito sul successo delle politiche. Ad esempio, le politiche valutate in Ungheria, Italia e Polonia sono state tutte avviate in un periodo di ripresa economica dopo un lungo periodo di crisi, mentre la misura studiata in Spagna è stata osservata sia in un periodo di crescita economica sia durante la "Grande Recessione". Quando si interpretano e presentano i risultati in termini di occupazione e redditi, i decisori politici dovrebbero comparare le stime con statistiche aggregate ben conosciute. Per esempio, l'analisi potrebbe focalizzarsi sul numero di disoccupati registrati nelle liste di disoccupazione, dato che può differire largamente da quanto rilevato tramite le rilevazioni delle forze di lavoro, che si basano su una diversa definizione di disoccupato. L'interpretazione dovrebbe esporre chiaramente queste differenze.

4.4. Effetto di spiazzamento, perdite secche

Affinché i decisori politici possano avere un quadro più completo dell'efficacia delle misure, è necessario sottolineare due aspetti. In primo luogo, si verifica un effetto spiazzamento, quando la misura aiuta l'impiego dei giovani, ma danneggia le prospettive degli individui molto simili a loro in termini di competenze, ma che non hanno accesso alla misura. In secondo luogo, il programma genera perdite secche (*dead weight loss*) se una misura aumenta l'occupazione dei giovani, ma una parte non trascurabile del sostegno fornito potrebbe risultare non necessaria, nel senso che i datori di lavoro avrebbero comunque assunto quei particolari individui anche in assenza delle misure. Ciò porta a sottili questioni di scelte sociali (politiche), nel senso che occorre considerare l'importanza relativa degli obiettivi alternativi. Tuttavia, non è affatto semplice fornire evidenze su tali aspetti o, quantomeno, sono necessari ulteriori dati su altri lavoratori della stessa organizzazione e su altre organizzazioni, dati in cui auspicabilmente i dipendenti sono associati all'organizzazione che li impiega (*employer-employee database*).

4.5. Analisi costi-benefici e analisi costi-efficacia

Le valutazioni dell'impatto di cui abbiamo discusso finora definiscono l'effetto della politica in termini di risultati (*outcome*) nel mercato del lavoro, ma è necessario prendere in considerazione anche la valutazione dei costi totali del programma. Questo diventa ancora più importante nel momento in cui i decisori politici desiderano confrontare politiche alternative in termini di rendimenti relativi (monetari).

Esistono due tipi di analisi che mettono in relazione i costi e i benefici. L'analisi costo-efficacia confronta i costi del programma con i risultati espressi in unità fisiche, ad esempio il programma ha generato un contratto di lavoro per una certa quantità di euro spesi. Nell'analisi costi-benefici, invece, gli effetti del programma vengono espressi anche in unità monetarie, il che consente di valutare se il beneficio totale supera i costi totali. Chiaramente quest'ultima analisi è più impegnativa, poiché la contabilizzazione e la quantificazione dei benefici richiede maggiori informazioni.

Anche l'analisi costi-efficacia può essere piuttosto impegnativa se il database amministrativo non contiene i costi del programma, come spesso accade. Il ricercatore potrebbe in tal caso utilizzare dati aggregati provenienti da altre fonti per calcolare i costi per unità di programma. In alternativa, in caso di sussidi salariali o programmi di apprendistato, i dati salariali provenienti dagli archivi della previdenza sociale uniti con il database con le informazioni relative al programma, potrebbero fornire una stima dei costi. In ogni caso, i ricercatori dovrebbero già pianificare l'acquisizione di questo tipo di dati all'inizio della valutazione.

BIBLIOGRAFIA

Abadie, Alberto and Cattaneo, Matias D. (2018). Econometric Methods for Program Evaluation (August 2018). Annual Review of Economics, Vol. 10, pp. 465-503, 2018
<http://dx.doi.org/10.1146/annurev-economics-080217-053402>

Angrist J.D., Pischke J.S. (2015). Mastering Metrics: The Path from Cause to Effect, Princeton University Press, Princeton.
<https://doi.org/10.1111/j.1475-4932.2011.00742.x>

Angrist, Joshua and Pischke, Jörn-Steffen (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Caliendo M., Mahlstedt R., Mitnik O. (2017). Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. Labour Economics, 46, 14-25.
<https://doi.org/10.1016/j.labeco.2017.02.001>

Card D., Krueger A.B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania, American Economic Review, 84, n. 4, pp. 772–793.
<http://davidcard.berkeley.edu/papers/njmin-aer.pdf>

European Commission (2013). Evalsed Sourcebook, Methods and Techniques
https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf

European Commission (2020). How to use administrative data for European Social Funds counterfactual impact evaluations, Directorate-General for Employment Social Affairs and Inclusion, Directorate G.
<https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8303>

Gertler, Paul J.; Martinez, Sebastian; Premand, Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J. (2016). Impact evaluation in practice (English). Washington, D.C. : World Bank Group.
<http://documents.worldbank.org/curated/en/698441474029568469/Impact-evaluation-in-practice>

Imbens, Guido, and Thomas Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice, Journal of Econometrics, 142, 615-635.
<https://doi.org/10.1016/j.jeconom.2007.05.001>

Leeuw, Frans & Vaessen, Jos (2009). Impact Evaluations and Development: NONIE Guidance on Impact Evaluation. Independent Evaluation Group, World Bank.
<https://www.oecd.org/development/evaluation/dcdndep/47466906.pdf>

Morris, Stephen & Tödtling-Schönhofer, Herta & Wiseman, Michael (2013). Design and Commissioning of Counterfactual Impact Evaluations - A Practical Guidance for ESF Managing Authorities. Publisher: European Commission, Directorate-General for Employment, Social Affairs and Inclusion. DOI:10.2767/94454

UNDP (2009). Handbook on Planning Monitoring and Evaluating for Development Result.

<http://web.undp.org/evaluation/handbook/documents/english/pme-handbook.pdf>

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

<https://doi.org/10.1093/biomet/70.1.41>

APPENDICE 1: BREVE DESCRIZIONE DELLE VALUTAZIONI EFFETTUATE NEI QUATTRO PAESI

ITALIA



Programma valutato:

- 1) 100% di sgravi contributivi per nuovi contratti di lavoro dipendente a tempo indeterminato e per le trasformazioni dei contratti da tempo determinato a tempo indeterminato
- 2) Riduzione dei costi di licenziamento per i contratti a tempo indeterminato (contratto a tutele crescenti).

Dati:

Tutti i contratti registrati tra il 1° gennaio e il 31 dicembre 2015 in SISCO (Sistema Informativo Statistico delle Comunicazioni Obbligatorie) a confronto con quelli registrati nell'anno precedente. SISCO contiene le informazioni elementari sulle comunicazioni inviate dai datori di lavoro (Comunicazioni Obbligatorie – COB) su assunzioni, trasformazioni e cessazioni dei contratti di lavoro dipendente o parasubordinato.

Risultato analizzato, Outcome:

Quota di nuovi avviamenti con contratto di lavoro dipendente a tempo indeterminato sul totale dei nuovi contratti di lavoro dipendente registrati nel 2015.

Strategia di identificazione:

Modello Diff-in-diffs con correzione per violazione del principio SUTVA.

Gruppo eleggibile:

- Giovani assunti con contratto di lavoro dipendente a tempo indeterminato (apprendistato escluso) nel periodo 2015, che nei sei mesi precedenti non risultavano occupati con contratto di lavoro dipendente a tempo indeterminato.
- Tutti i datori di lavoro senza distinzioni settoriali o geografiche. Sono comprese le associazioni e le imprese pubbliche, ma sono escluse la pubblica amministrazione e l'agricoltura.
- I datori di lavoro che non hanno in sospeso arretrati contributivi o qualsiasi altra situazione irregolare in termini di contratti collettivi o accordi territoriali.

Gruppo non eleggibile:

- Giovani assunti nel corso del 2015 che, nei 6 mesi precedenti l'inizio del nuovo contratto a tempo indeterminato, risultavano assunti con un contratto di lavoro dipendente a tempo indeterminato;
- I giovani assunti nel corso del 2015 che nei 3 mesi precedenti la riforma (01.10.2014-31.12.2014) non sono stati licenziati dallo stesso datore di lavoro che li aveva assunti o da un altro datore di lavoro collegato a quest'ultimo (per evitare comportamenti opportunistici).

Definizione di variabile risultato til risultato è una Dummy, Y, con due possibili valori Y=1, per contratto di lavoro dipendente a tempo indeterminato; Y=0, per contratto di lavoro dipendente a tempo determinato.

Definizione di variabile di trattamento: T=1, per contratto eleggibile al trattamento; T=0, per contratto non eleggibile al trattamento.

Definizione di periodo: P=0, 2014; P=1, 2015.

UNGHERIA



Programma:

La valutazione ungherese è incentrata sul programma di “90 giorni di prova” (job trial) progettato per i giovani al di sotto dei 25 anni. Il job trial prevede un sussidio al 100% dei costi salariali per 90 giorni senza alcun obbligo di ulteriore assunzione. Il programma è iniziato nel 2015 come parte del programma Garanzia Giovani.

Dati:

La valutazione si basa su due dataset amministrativi. I dati SPI includono tutti gli individui in cerca di lavoro registrati e i partecipanti al programma. I dati SPI sono collegati a un dataset amministrativo dell'istituto di previdenza, che riporta la condizione occupazionale e la cronologia occupazionale di circa la metà della popolazione totale.

Risultati:

La principale variabile risultato è la condizione occupazionale a distanza di sei e dodici mesi dal completamento del programma e del ricorso ai sussidi salari cumulati previsti da due specifiche misure.

Strategia di identificazione:

La valutazione applica due diverse strategie di identificazione. La prima identificazione confronta i risultati per i partecipanti al programma di 90 giorni con i partecipanti di altri programmi (lavoro pubblico e formazione), applicando il propensity score matching.

La seconda stima è una specificazione del difference-in-differences e sfrutta il fatto che nella “Regione centrale” dell'Ungheria il programma è iniziato con 9 di ritardo rispetto alle altre Regioni del Paese. Il gruppo di trattamento comprende i giovani eleggibili di tre “Regioni convergenza” e il gruppo di controllo è composto da giovani eleggibili della “Regione centrale” del paese. La stima confronta i risultati sul gruppo degli eleggibili e sui gruppi di controllo: coloro che si sono registrati nell'intervallo compreso nei 9 mesi precedenti e successivi alla data del 1° gennaio 2015, ovvero la data di introduzione del programma Garanzia Giovani nella “Regione centrale” dell'Ungheria. La specifica permette l'identificazione dell'ITT; mostra dunque l'effetto dell'intero programma Garanzia Giovani sulla sottopopolazione eleggibile.

POLONIA (VALUTAZIONE 1)



Programma:

In Polonia viene valutato un programma di sussidi salariali per i disoccupati al di sotto dei 30 anni di età. È stato operativo negli anni 2016-2018. Il sussidio veniva erogato al datore di lavoro a integrazione del reddito, con tetto massimo pari al salario minimo, più la previdenza sociale, per una durata di 12 mesi con obbligo per i datori di lavoro di prolungare l'impiego per 12 mesi dalla scadenza del sussidio.

Gruppo eliggibile:

Disoccupati al di sotto dei 30 anni registrati con il profilo I o II (il profilo III non è eleggibile). Inoltre, i datori di lavoro che hanno ridotto l'occupazione di propria iniziativa negli ultimi 6 mesi non erano eleggibili.

Dati:

Dati del registro dei servizi pubblici per l'impiego (SPI) con la cronologia completa delle iscrizioni alla disoccupazione e della partecipazione ai programmi, non collegati ai dati della previdenza sociale.

Campione:

Individui che sono entrati in disoccupazione tra il 2015 e il 2016, con l'esclusione degli individui disoccupati o partecipanti ad un programma di PAML nei precedenti 6 mesi, in modo da garantire che l'analisi di una nuova registrazione non facesse parte di un periodo di disoccupazione più lungo. Per poter osservare il risultato 36 mesi dopo la registrazione, il campione è stato limitato alle registrazioni effettuate fino all'aprile del 2016.

Risultati:

Sono analizzati due tipi di risultati: l'assenza dal registro della disoccupazione tra il 12° e il 36° mese dopo la l'ultima registrazione e il numero cumulativo di giorni di assenza dal registro da tra il 12° e il 36° mese dopo l'ultima registrazione.

Strategia di identificazione:

Regression Discontinuity Design che utilizza la soglia di età combinata con una Diff in Diffs (2016 rispetto al 2015: altre discontinuità nelle norme istituzionali in corrispondenza della soglia dei 30 anni erano presenti nel 2015).

POLONIA (VALUTAZIONE 2)



Programma:

Programma composto da più misure di Politica Attiva del Lavoro offerti ai giovani: Formazione *on-the-job*, formazione in aula, programmi di "impiego pubblico", sussidi salariali, voucher per la formazione *on-the-job* e voucher per la formazione in aula.

Gruppo eliggibile:

Disoccupati fino a 30 anni di età

Dati:

Dati dal Registro amministrativo dei Servizi per l'Impiego (SPI) con informazioni complete sulle registrazioni dello stato di disoccupazione e sulla partecipazione ai programmi, non collegato a dati di previdenza sociale.

Campione:

Campione ristretto ai programmi di politica attiva del lavoro del 2015-2017, di durata superiore ad un giorno. Solo I partecipanti di età 18-29 al momento della partenza del programma anni sono stati inclusi. I detenuti sono stati esclusi. Il campione totale comprende 319,610 unità statistiche.

Risultati:

Assenza dalle liste di disoccupazione e dai programmi di Politiche attive del lavoro nel periodo compreso tra il 12° e il 36° mese successivo all'iscrizione.

Strategia di identificazione:

Confronto tra coppie di partecipanti a ciascun programma abbinati tramite tecnica di *propensity score matching*.

SPAGNA



Programma:

Contratto di tirocinio. È stato implementato in Spagna nel 1998 e, secondo la norma, l'obiettivo principale consiste nell'aumentare la stabilità lavorativa aumentando l'esperienza professionale e offrendo incentivi all'azienda per l'assunzione di giovani lavoratori, al termine del periodo di tirocinio.

Eligibilità:

Le restrizioni dei criteri di eleggibilità si riferiscono all'istruzione e all'età: devono aver terminato il corso di laurea o di formazione professionale ed essere di età inferiore ai 30 anni o, meglio, aver terminato il corso di laurea meno di 5 anni prima.

Dati:

Continuous Sample of Work Histories (CSWH). Si tratta dei dati della previdenza sociale relativi al 5% della popolazione spagnola. Questo dataset include la cronologia completa della storia lavorativa di un campione rappresentativo dei lavoratori spagnoli.

Campione:

Individui che sono entrati nel mondo del lavoro a partire dal 2002, di età inferiore ai 30 anni, con un livello di istruzione elevato: potenziali utenti del contratto di tirocinio.

Risultati:

Vengono analizzati due tipi di risultati: i) la probabilità di rimanere in azienda dopo il contratto di tirocinio e ii) la probabilità di trovare un contratto a tempo indeterminato dopo il contratto di tirocinio, in particolare, nella stessa azienda (per quelli che rimangono) o in un'altra azienda (per quelli che cambiano).

Strategia di identificazione:

Per la prima analisi, si sviluppa una regressione *logit* multinomiale in cui la variabile dipendente assume tre diversi valori, nello specifico, i) passaggio ad altra azienda, ii) permanenza in azienda e iii) passaggio nello stato di disoccupazione. Nella seconda fase, si stima separatamente la probabilità di stipulare un contratto di lavoro dipendente a tempo indeterminato (rispetto ad un contratto a termine) a seconda che rimangano in azienda o che passino ad un'altra azienda con un modello di valutazione *logit*.

APPENDICE 2: GLOSSARIO E ABBREVIAZIONI

ATE, ATT *Average Treatment Effect* (Effetto Medio del Trattamento) e *Average Treatment effect on the Treated* (Effetto Medio del Trattamento sui Trattati). Un effetto del trattamento è l'effetto causale del trattamento (una variabile binaria, 0-1) su una variabile risultato. Stima la differenza tra il risultato potenziale di un'unità di popolazione con e senza il trattamento (esposizione alla politica, partecipazione ad un programma specifico ecc.). Esistono due definizioni principali dell'effetto medio del trattamento. L'**ATE** stima il valore atteso nella popolazione della differenza media nella coppia di risultati potenziali calcolati in media sull'intera popolazione di interesse. Si tratta della misura più rilevante, quando l'intera popolazione è esposta alla politica in esame.

$$ATE = E(Y_i(1) - Y_i(0))$$

Dove $Y_i(1)$ è il risultato dell'unità i quando riceve il trattamento e $Y_i(0)$ è il risultato dell'unità i quando non riceve il trattamento.

L'**ATT** mostra l'effetto medio del trattamento sulla sottopopolazione trattata:

$$ATE = E(Y_i(1) - Y_i(0) | D_i = 1)$$

CIE *Counterfactual impact evaluation* (Valutazione di impatto controfattuale) La valutazione controfattuale o valutazione d'impatto o degli effetti di una politica è un processo attraverso cui stimare, con metodologie adeguate e dati oggettivi e verificabili, se un certo intervento pubblico riesce a produrre un cambiamento su un certo fenomeno diverso da quello che si sarebbe verificato in assenza di quell'intervento.

CO, COB *Comunicazioni Obbligatorie on-line*. Si tratta del registro delle comunicazioni obbligatorie gestito dal Ministero del Lavoro e delle Politiche Sociali (Legge 27 dicembre 2006, n. 296; Legge Finanziaria 2007).

CSWH *Continuous Sample or Work History*, in spagnolo "Muestra Continua de Vidas Laborales", rilevazione continua delle storie lavorative, dell'Istituto nazionale di statistica spagnolo, Instituto Nacional de Estadística (INE).

Diff in Diffs/DiD *Difference in Differences* (Differenza di differenze). Metodo di stima quasi sperimentale in cui l'effetto del programma viene calcolato confrontando il gruppo di trattamento e un gruppo di controllo prima e dopo il trattamento. L'effetto del trattamento è la variazione media nel tempo della variabile risultato per il gruppo di trattamento, meno la variazione media per il gruppo di controllo. Il metodo DiD si basa sul presupposto che in assenza del trattamento, l'esito del gruppo di trattamento avrebbe seguito l'andamento delle unità non trattate; cioè, i due gruppi avrebbero avuto tendenze parallele (per maggiori dettagli, vedere, ad esempio, Card e Krueger, 1994, Angrist e Pischke, 2009).

ISCED *International Standard Classification of Education* (Classificazione internazionale standard dell'istruzione). Sistema di classificazione internazionale di riferimento per organizzare i programmi di istruzione e le relative qualifiche per livello e settore. L'ISCED 2011 (livelli di istruzione) è stato attuato in tutte le raccolte di dati dell'UE dal 2014. L'ISCED-F 2013 (settori dell'istruzione e della formazione) è stato attuato dal 2016 <https://bit.ly/3rtpg39>

ITT *Intention To Treat* (Effetto della sola presenza del trattamento sugli individui della popolazione eleggibile). Mostra l'effetto della politica sulla popolazione eleggibile. In altri termini, mostra l'effetto causale dell'offerta di trattamento. Se non tutti i membri della popolazione eleggibile ricevono il trattamento, perché alcuni lo rifiutano, l'ITT sarà diverso dall'ATE e dall'ATT.

OLS *Ordinary Least Squares* (Metodo dei minimi quadrati). Metodo per definire una funzione matematica. A partire da un insieme limitato di dati osservati si approssima una funzione matematica minimizzando la somma dei quadrati delle distanze dai punti osservati, rappresentati su un diagramma cartesiano.

PAM/PAL=Politiche attive del mercato del lavoro o Politiche attive del lavoro

RDD *Regression Discontinuity Design*. Metodo quasi sperimentale che consente di stimare l'effetto del trattamento del programma quando l'eleggibilità o la probabilità di partecipare al programma dipende da una certa caratteristica osservabile, come l'età (*forcing variable* o *running variable*). L'effetto del programma viene misurato confrontando i risultati al di sotto e al di sopra della soglia nelle immediate vicinanze della soglia stessa. L'RDD calcola l'effetto medio locale del trattamento adattando un polinomio (locale) su entrambi i lati della soglia e si basa sul presupposto che la selezione tra il gruppo di trattamento e il gruppo di controllo sia casuale. L'RDD può essere *fuzzy* (sfocato) o *sharp* (nitido) (per maggiori dettagli, vedere, ad esempio, Imbens e Lemieux, 2008; Angrist e Pischke, 2009).

SISCO=Sistema Statistico delle Comunicazioni Obbligatorie on-line, comunicazione obbligatoria raccolta in Italia dal Ministero del Lavoro e delle Politiche Sociali (MLSP).

SPI=Servizio pubblico per l'impiego

SUTVA *Stable unit treatment value assumption*. Ipotesi di invarianza dell'effetto del trattamento. Questa è un'ipotesi che viene solitamente formulata in inferenza causale: ovvero un trattamento applicato a un'unità non influisce sul risultato per un'altra unità (nessuna interferenza tra le unità).

Unconfoundedness: (Selezione sulle osservabili). Distorsioni nel confronto tra le unità trattate e quelle di controllo. Il termine si riferisce al caso in cui il controllo per differenze in un pre-trattamento assegnato alle covariate rimuove le distorsioni nel confronto tra le unità trattate e quelle di controllo, in modo che, a seconda di tali controlli, l'assegnazione del trattamento può essere considerata come randomizzata.

INAPP

Corso d'Italia, 33
00141, Roma (RM)
Italia
+39 06 854471
www.inapp.org

Collegio Carlo Alberto

Piazza Arbarello, 8
10122, Torino (TO)
Italia
+39 011 6705000
www.carloalberto.org



Quest' opera è pubblicata sotto [licenza CC BY-NC 4.0.](https://creativecommons.org/licenses/by-nc/4.0/)

