

RECENSIONE

CRISTINA BERLIRI “**Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme**”, di James Heckman, Hidehiko Ichimura and Petra Todd, in *Review of Economic Studies* 64 (4) (1997), pagg. 605-654, recensione uscita in *Osservatorio Isfol* - N. 1-2, anno 2001.

Il lavoro verifica l'efficacia dei metodi di *matching* nel produrre stime sull'impatto di programmi formativi vicine a quelle ottenibili attraverso l'uso dei metodi sperimentali. Nella valutazione dell'impatto di un programma o di un trattamento, occorrerebbe poter mettere a confronto i risultati realizzati dai partecipanti con i risultati che costoro avrebbero sperimentato se non avessero partecipato al programma. Il problema di valutazione si configura pertanto come un problema di dati mancanti essendo le due condizioni di “partecipante” e “non partecipante” ad un programma mutuamente escludenti. Il confronto avviene dunque tra il gruppo dei trattati ed un gruppo di controllo che si vorrebbe idealmente il più possibile simile al primo. In molti casi per la valutazione dei programmi sono state implementate procedure sperimentali consistenti nella selezione casuale del gruppo dei trattati e del gruppo di controllo. Resta tuttavia il problema di valutare i risultati di programmi per i quali non siano disponibili dati sperimentali. I metodi di *matching* consistono nell'abbinare ai partecipanti al programma i membri di un gruppo di controllo non sperimentale che sia simile dal punto di vista delle caratteristiche osservabili degli individui. Gli impatti del trattamento vengono stimati sottraendo dai risultati medi dei partecipanti i risultati medi dei membri abbinati del gruppo di confronto. Nella pratica può risultare difficile realizzare il *matching* sulla base di un numero elevato di variabili. In tal caso può essere utilizzata una procedura di valutazione in due stadi consistente nello stimare la probabilità di un individuo di partecipare al programma e nell'abbinare i partecipanti con individui appartenenti al gruppo di controllo sulla base dei valori della probabilità stimata anziché delle singole caratteristiche individuali.

In questo saggio gli autori propongono di estendere i metodi tradizionali di *matching*, e in particolare la procedura a due stadi, in diverse direzioni: incorporando restrizioni di esclusione in riferimento alle variabili che appaiono nelle equazioni di partecipazione al programma e di esito del trattamento; presentando condizioni più deboli affinché il metodo di *matching* risulti giustificato come strumento per la valutazione; incorporando informazioni a priori sulla forma funzionale delle equazioni stimate ed in particolare l'ipotesi che le determinanti osservabili e inosservabili degli esiti siano additivamente separabili; proponendo nel contesto di dati longitudinali uno stimatore generalizzato con il metodo delle differenze nelle differenze; fornendo una teoria della distribuzione asintotica per gli stimatori *matching*. La performance di metodi alternativi di *matching* e delle loro estensioni proposte viene comparata empiricamente facendo uso di dati relativi ad un esperimento sociale su vasta scala che ha consentito la costruzione di un gruppo di controllo sperimentale, selezionato casualmente tra i richiedenti risultati ammissibili, e di un gruppo di confronto non sperimentale, composto da individui che non hanno fatto richiesta di partecipare al programma pur soddisfacendo le condizioni di ammissibilità. Ulteriori gruppi di confronto sono stati costruiti sulla base di altre fonti di dati non sperimentali. L'efficacia dei diversi stimatori non sperimentali viene misurata dalla loro capacità di ridurre le differenze in termini di esiti tra il gruppo di confronto non sperimentale e il gruppo di controllo selezionato casualmente. Scopo del lavoro è inoltre quello di sottoporre a verifica empirica le assunzioni che giustificano l'uso dei diversi metodi di *matching*.

Il confronto empirico tra stimatori di impatto fondati su metodi di *matching* si basa sui dati relativi alla realizzazione del programma JTPA (Job Training Partnership Act). Il programma JTPA costituisce un esperimento sociale su vasta scala appositamente realizzato su commissione del *Department of Labor* degli Stati Uniti al fine di valutare l'efficacia dei principali programmi

governativi di formazione professionale dei lavoratori svantaggiati. Il programma fornisce formazione sul luogo di lavoro, assistenza nella ricerca di un impiego e formazione in aula a giovani e adulti appartenenti a famiglie disagiate. Sotto la supervisione degli autori sono stati raccolti dati longitudinali sia sui partecipanti al programma sia sul gruppo di controllo selezionato casualmente e su un gruppo di confronto composto da non partecipanti eleggibili al programma.

Gli autori considerano un'estensione della metodologia *matching* al contesto di informazioni longitudinali utilizzando uno stimatore condizionale con il metodo delle differenze nelle differenze. Lo stimatore si basa sul confronto tra la condizione in termine di reddito prima e dopo la partecipazione al programma dei partecipanti e la condizione in termine di reddito dei non partecipanti alle stesse date. Tale stimatore a differenza degli altri stimatori *matching* consente una parziale correzione delle distorsioni connesse a differenze nelle variabili inosservabili, in particolare per quanto riguarda le caratteristiche invarianti nel tempo.

La disponibilità di dati sperimentali consente di stimare la distorsione nella valutazione degli impatti che deriva dall'uso della media dei risultati realizzati dal gruppo di confronto non sperimentale come *proxy* dei risultati medi che avrebbero conseguito i partecipanti al programma se non avessero partecipato. La stima si basa sulla ipotesi che questi ultimi siano ben stimati dai risultati medi del gruppo di controllo sperimentale e viene misurata convenzionalmente dalla differenza tra i risultati medi conseguiti dal gruppo di confronto e quelli conseguiti dal gruppo sperimentale. Per ciascun gruppo di confronto utilizzato questa misura della distorsione è stata scomposta in tre componenti: la componente di distorsione dovuta alla selezione; la componente attribuibile a differenze nel supporto rispetto al quale viene stimato l'impatto nei due casi di procedura sperimentale e non sperimentale; differenze attribuibili al *mismatching* o alla diversa ponderazione dei dati. Similmente le stime di impatto realizzate con ciascuno stimatore *matching* sono state messe a confronto con il valore stimato attraverso il metodo sperimentale.

La validità dei metodi di *matching* si basa su specifiche assunzioni sulla indipendenza degli esiti dall'appartenenza al gruppo dei partecipanti piuttosto che al gruppo di confronto. I diversi metodi di *matching* possono richiedere specificazioni più o meno restrittive della condizione di indipendenza. Gli autori applicano metodi sviluppati in Heckman, Ichimura, Smith e Todd (1996) e una loro generalizzazione nel contesto di dati longitudinali per testare empiricamente le condizioni di validità dei diversi stimatori.

Dall'analisi empirica realizzata emerge come la distorsione delle stime per le valutazioni di impatto realizzate con i metodi di *matching*, sebbene sia ridotta rispetto a quella derivante dalla semplice differenza tra medie, costituisca una frazione sostanziale dell'impatto totale del programma. Anche lo stimatore condizionale con il metodo delle differenze nelle differenze non è esente da questo limite. L'impiego di procedure di *matching* equivale, in generale, ad assumere che la selezione sia riconducibile a variabili osservabili. Sebbene il metodo delle differenze nelle differenze consenta di correggere la distorsione connessa a componenti inosservabili invarianti nel tempo, la distorsione residua è risultata comunque elevata.

La scomposizione nelle sue diverse componenti della misura convenzionale di distorsione mostra come le differenze nel supporto e le differenze nella ponderazione dei dati rappresentino le componenti principali. I metodi di *matching* applicati consentono di eliminare entrambe le fonti di distorsione ma non eliminano una terza causa riconducibile a differenze tra gruppi nella distribuzione di variabili inosservabili.

L'efficacia degli stimatori dipende dalle caratteristiche dei dati utilizzati nella costruzione del gruppo di confronto. Metodi di *matching* applicati a gruppi di confronto situati nello stesso mercato

locale del lavoro dei partecipanti, cui sia stato somministrato lo stesso questionario e pertanto le stesse definizioni di esito e variabili esplicative, eliminano molta della distorsione misurata convenzionalmente dalla differenza tra medie.

Le ipotesi generalmente invocate in letteratura per giustificare l'uso di metodi di *matching* non hanno passato i test, mentre le assunzioni più deboli che giustificano l'estensione delle procedure di *matching* al metodo delle differenze nelle differenze hanno trovato conferma nei dati.

Il metodo generalizzato delle differenze nelle differenze riduce la distorsione, specialmente nel caso in cui questa dipenda da componenti inosservabili costanti nel tempo quali gli effetti della somministrazione di questionari diversi e della localizzazione dei due gruppi in mercati del lavoro non coincidenti. Nel complesso il metodo proposto dagli autori è risultato in alcuni casi più efficace di altri stimatori di *matching*, ma non è risultato lo stimatore dominante malgrado sia il solo le cui condizioni di validità non siano state rifiutate dai test. La distorsione stimata, sebbene ridotta rispetto a quella risultante dalla semplice differenza nelle medie, rappresenta una frazione sostanziale dell'impatto totale del programma.

Il lavoro fornisce utile evidenza empirica sulla performance dei metodi non sperimentali nel realizzare stime di impatto dei programmi di formazione al lavoro e nel consentire analisi inferenziali sui programmi. Esso mette in luce le caratteristiche desiderabili dei dati utilizzati nella costruzione dei gruppi di confronto. Nel caso risultino disponibili dati longitudinali viene proposta un'estensione della procedura di *matching* al metodo delle differenze nelle differenze la cui performance risulta generalmente migliore rispetto a quella degli altri stimatori *matching*.