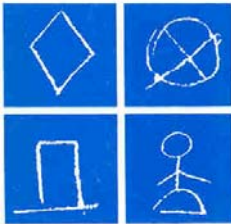


**VALUTAZIONE DELL'EFFICACIA  
DI INTERVENTI PUBBLICI  
CONTRO LA POVERTA': QUESTIONI  
DI METODO E STUDI DI CASI**



Commissione di  
indagine  
sulla povertà  
e sull'emarginazione



PRESIDENZA DEL CONSIGLIO DEI MINISTRI  
DIPARTIMENTO PER L'INFORMAZIONE E L'EDITORIA

**VALUTAZIONE DELL'EFFICACIA  
DI INTERVENTI PUBBLICI  
CONTRO LA POVERTÀ:  
QUESTIONI DI METODO  
E STUDI DI CASI**

**Commissione di indagine  
sulla povertà e sull'emarginazione**

**PRESIDENZA DEL CONSIGLIO DEI MINISTRI  
DIPARTIMENTO PER L'INFORMAZIONE E L'EDITORIA**

PRESIDENZA DEL CONSIGLIO DEI MINISTRI

*COMMISSIONE DI INDAGINE  
SULLA POVERTÀ E SULL'EMARGINAZIONE*

*Presidente*

Pierre Carniti

*Membri*

Andrea Brandolini, Giuliano Cazzola, Mario Colombo,  
Luigi Di Liegro, Paolo Garonna, Massimo Livi Bacci, Gio-  
vanni Moro, Armando Sanguini, Chiara Saraceno,  
Ugo Trivellato

*Segreteria*

Antonella Buono

*Collaborano*

Guido Bolaffi, Giuliana Coccia, Viviana Egidi, Sergio  
Lucaresi, Maurizio Manente, Mauro Maselli, Mario Maraz-  
ziti, Carlo Trevisan

## SOMMARIO

# VALUTAZIONE DELL'EFFICACIA DI INTERVENTI PUBBLICI CONTRO LA POVERTÀ: QUESTIONI DI METODO E STUDI DI CASI

<b>Introduzione</b>	5
<b>1. Alcune premesse di ordine metodologico</b>	7
1.1. <i>Tipologie di politiche contro la povertà e l'esclusione sociale</i>	7
1.2. <i>Di quale valutazione stiamo parlando?</i>	8
1.3. <i>La valutazione di efficacia delle politiche contro la povertà</i>	16
<b>2. Il disegno della valutazione di efficacia</b>	21
2.1. <i>La valutazione dell'efficacia (complessiva) di un intervento</i>	21
2.2. <i>Un disegno più analitico (e realistico) di valutazione</i>	26
<b>3. La valutazione dell'efficacia delle politiche attive contro la povertà</b>	33
3.1. <i>Aspetti metodologici della valutazione di impatto diretto</i>	34
3.1.1. <i>Le valutazioni sperimentali</i>	39
3.1.2. <i>Le valutazioni non sperimentali</i>	41
3.2. <i>Casi di valutazione di impatto condotta negli Stati Uniti</i>	43
3.3. <i>La valutazione delle altre componenti dell'efficacia complessiva</i>	55
3.3.1. <i>La scelta dei beneficiari</i>	55
3.3.2. <i>La scala dell'intervento</i>	58
3.3.3. <i>La presenza di effetti negativi</i>	59
<b>4. La valutazione dell'efficacia delle politiche passive contro la povertà</b>	61
4.1. <i>L'impatto diretto sui beneficiari</i>	61
4.2. <i>La scelta dei beneficiari</i>	63
4.3. <i>La scala dell'intervento</i>	65
4.4. <i>La presenza di effetti negativi</i>	66
<b>Riferimenti bibliografici</b>	69



## INTRODUZIONE

*I riferimenti all'importanza e all'utilità della valutazione sono ormai onnipresenti nel dibattito sulla riforma della pubblica amministrazione italiana, e con sempre maggior frequenza vengono recepiti in provvedimenti legislativi a livello nazionale e regionale. La ragione di tanta attenzione va probabilmente cercata nel diffondersi della consapevolezza che, in un clima di forte contenimento della spesa pubblica, le sempre più scarse risorse vadano usate in modo più razionale, e strumenti conoscitivi nuovi siano necessari per orientare tale processo di razionalizzazione. Tale risveglio di interesse nei confronti della valutazione non è una caratteristica solo italiana. Rileva un recente articolo sullo stato della valutazione in Francia: "Over the past 5-10 years in France, there has been a widespread infatuation with public policy evaluation ... evaluation constitutes a pragmatic attempt to find out what is going on, a seductive and reassuring idea at a time of ideological insecurity and when political actors appear to lack credible explanations" (Duran, Monnier e Smith, 1996, pagg. 45-46). L'articolo citato sottolinea la relazione tra lo sviluppo della valutazione in Francia e il processo di progressivo ridimensionamento dello stato sociale.*

*Il presente lavoro si occupa di un tipo particolare di valutazione, quella di efficacia, applicata ad un ambito relativamente ristretto di interventi pubblici, le politiche contro la povertà e l'esclusione sociale. La discussione inizia con alcuni chiarimenti di ordine generale. Innanzitutto, si offrono alcuni esempi di politiche contro la povertà, in modo da dare maggiore concretezza alla successiva discussione: importante è la distinzione tra politiche attive e politiche passive. In secondo luogo, si definisce più precisamente cosa si intende per valutazione, termine che viene usato in riferimento ad un ampio ventaglio di attività conoscitive, di cui solo alcune rilevanti nell'ambito di questo lavoro. Infine, si mette a fuoco il concetto stesso di efficacia, anche questo un termine a cui vengono attribuiti significati diversi a seconda del contesto in cui viene usato. La parte centrale del rapporto tratta più in concreto di come la valutazione di efficacia delle politiche contro la povertà possa essere condotta, con quali metodi analitici, con quali obiettivi conoscitivi e quali speranze di successo: la Sezione 2 illustra alcuni possibili disegni di valutazione di efficacia rilevanti per queste politiche, la Sezione 3 discute applicazioni della valutazione di efficacia alle politiche attive contro la povertà, e la Sezione 4 alle politiche passive.*



## 1. ALCUNE PREMESSE DI ORDINE METODOLOGICO

### 1.1 Tipologie di politiche contro la povertà e l'esclusione sociale

La classificazione che segue ha uno scopo esemplificativo e non ha pretese di completezza. In particolare, ci limiteremo ad una classificazione basata sull'esperienza italiana, in quanto il fine ultimo di questo lavoro è provvedere spunti di riflessione per l'applicazione della valutazione alle politiche attuate o progettate in Italia (per un'esauriente rassegna di tali politiche, si veda Negri e Saraceno, 1996.)

La classificazione parte dalla distinzione tra politiche attive e politiche passive: Nella loro essenza, le *politiche attive* mirano a rimuovere le cause della povertà e dell'esclusione sociale. Alla base sta l'idea che la povertà e l'esclusione sociale siano il risultato di comportamenti o condizioni che possono essere almeno potenzialmente manipolati mediante l'uso di risorse pubbliche o interventi di tipo regolativo (obblighi e divieti). Le *politiche passive* invece intendono curare i sintomi e alleviare le conseguenze della povertà. Esse quindi agiscono non sui comportamenti messi in atto dagli individui, bensì direttamente sulle loro condizioni di vita. La distinzione ha importanti implicazioni per come si può definire un'attività di valutazione di questi due tipi di politiche.

Va precisato che, mentre da un punto di vista espositivo una distinzione netta tra politiche passive e politiche attive è senz'altro utile, nella realtà si ritrova un *continuum* di interventi, lungo cui si ritrovano casi puri e casi intermedi, in cui si mescolano i caratteri di politica attiva e politica passiva. Ad un estremo dello spettro si situano gli interventi di puro sostegno del reddito, in cui l'intervento si pone l'obiettivo di supplire ad una situazione accertata (o almeno presunta) di carenza di risorse economiche. Esempi tratti dalla realtà italiana sono le pensioni sociali, le integrazioni al minimo delle pensioni, gli assegni al nucleo familiare e il minimo vitale erogato da alcuni Comuni italiani.

All'estremo opposto dello spettro si situano gli interventi che hanno l'obiettivo di incidere direttamente sui comportamenti che si ritiene *causino* le situazioni osservate di povertà o esclusione sociale. Esempi sono gli interventi di socializzazione a favore dei giovani a rischio, o la formazione professionale a favore di disoccupati di lungo periodo. Questi interventi assumono più spesso la forma di fornitura di servizi a soggetti ben determinati. Non sono tuttavia da escludere da questa categoria anche interventi di tipo regolativo, ad esempio la legislazione sulle pari opportunità o sull'azione positiva.

Tra la pura politica attiva e la pura politica passiva troviamo molteplici situazioni intermedie, in cui l'intervento ha una duplicità di obiettivi, alcuni dei quali di tipo passivo, altri di tipo attivo. Un caso tipico è quello della fornitura di servizi che hanno come obiettivo primario quello di alleviare una situazione di bisogno, ma come obiettivo ausiliario la modifica permanente di alcuni comportamenti. Un esempio è la fornitura di assistenza domiciliare agli anziani, che intende innanzitutto migliorare le condizioni di vita dei beneficiari, ma allo stesso tempo diminuire il rischio che essi ricorrono al ricovero in casa di riposo. Un altro esempio è la predisposizione di strutture di accoglienza per i senza-dimora, che all'obiettivo di migliorare le condizioni di vita dei beneficiari unisce in certi casi quello di favorire il loro reinserimento nel mercato del lavoro.

La presenza di casi intermedi, o misti, non sminuisce l'utilità della distinzione. Lo stesso intervento può essere infatti sottoposto a tipi diversi di valutazione, a seconda che si voglia cogliere la sua abilità di modificare i comportamenti o di alleviare le situazioni di disagio-povertà-esclusione.

## **1.2 Di quale valutazione stiamo parlando?**

Lo scopo della discussione seguente è di offrire al lettore dei punti di riferimento per orientarsi in una materia che negli ultimi anni anche in Italia è cresciuta in complessità e difficoltà di lettura. Questo inquadramento tenta un raccordo con il paradigma della valutazione prevalente nei Paesi di cultura anglosassone, e in particolare negli Stati Uniti. Per delineare tale inquadramento, partiamo dalle due domande che dovrebbero essere poste preliminarmente a qualsiasi progetto o iniziativa di valutazione: "perché si vuole valutare?" e "cosa si vuole valutare?"

La prima domanda si riferisce alle intenzioni che muovono il valutatore e soprattutto il suo committente. Perché si vuole valutare un intervento, un'attività o una politica pubblica? Cosa si spera di imparare, e a quale scopo? Soprattutto, a che livello del processo decisionale pubblico si vorrebbero utilizzare i risultati? La distinzione fondamentale da questo punto di vista è tra valutazione intesa come *strumento per il disegno delle politiche pubbliche* e valutazione intesa come *strumento di management pubblico*.

Le due prospettive si differenziano per il fatto di rispondere a due esigenze conoscitive fondamentalmente diverse. Al di là di una comunanza, spesso solo apparente, di terminologia e di strumenti, queste due prospettive guardano in direzioni diametralmente opposte. Nel primo caso si valuta perché esiste una lacuna nella conoscenza di chi è chiamato a decidere sul disegno di una politica o a giudicare della sua realizzazione complessiva: tale mancanza di conoscenza può riguardare la fattibilità dell'intervento, le caratteristiche della sua messa in opera, e gli effetti desiderati e indesiderati che esso ottiene. La valutazione ha l'obiettivo di *far apprendere* elementi nuovi che possono essere utili per decisioni di riforma-mantenimento-esten-

sione-eliminazione dell'intervento. Volendo usare un termine sintetico per definire questa prospettiva sulla valutazione, si può adottare quello di *valutazione come apprendimento*.

Nella prospettiva del *management* pubblico, la valutazione è soprattutto uno strumento per *controllare* l'operato di chi gestisce un intervento o un servizio. Lo scopo ultimo non è accumulare conoscenza sull'intervento in quanto tale, bensì identificare disfunzioni, imporre sanzioni, spostare risorse, e infine migliorare la *performance* delle unità operative che gestiscono l'intervento in questione. Per questa prospettiva si può usare la definizione, forse riduttiva ma crediamo efficace, di *valutazione come controllo*.

#### *La valutazione tra "apprendimento" e "controllo"*

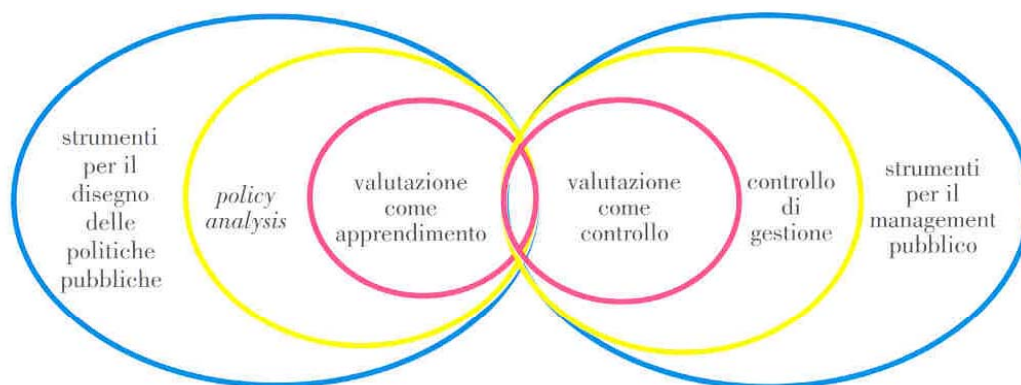
Come in ogni classificazione, è possibile trovare zone grigie e sovrapposizioni tra i diversi elementi che si tenta di classificare. Quindi non intendiamo assolutizzare questa distinzione, ma riteniamo che sia indispensabile per semplificare la complessità di un'attività conoscitiva multiforme quale la valutazione. La Figura 1 rappresenta schematicamente queste due diverse prospettive sulla valutazione.

Fig. 1

---

**DUE DIVERSE PROSPETTIVE SULLA VALUTAZIONE:  
"APPRENDIMENTO" E "CONTROLLO"**

---



La valutazione come apprendimento (VA) comprende quelle attività analitiche che hanno per oggetto l'intervento pubblico visto nel suo complesso, senza un immediato riferimento al funzionamento delle unità operative che gestiscono-producono-erogano l'intervento stesso. L'obiettivo della valutazione è migliorare la conoscenza di come l'intervento opera, di quali effetti produce, di come viene percepito dagli operatori e dagli utenti, e di come interagisce con altri interventi e politiche. Le tipiche "domande" a cui questo tipo di valutazione tenta di rispondere sono quelle che chiedono se l'intervento:

- ottiene gli effetti desiderati?
- evita gli effetti indesiderati?
- vale quello che costa?
- serve i beneficiari giusti?
- si scosta dal disegno originario?

Questa attività analitica non ha solitamente lo scopo di produrre un ricaduta immediata sulla gestione dell'intervento: più che diretta ai manager pubblici, questo tipo di valutazione è diretta ai *policy-makers*, che formulano le politiche pubbliche ma non le gestiscono direttamente, o ad altri *stakeholders* nel processo decisionale pubblico. Per inciso, va sottolineato che la valutazione è solo di uno dei molti inputs possibili alla formulazione delle politiche pubbliche, in quanto coesiste con una gran quantità di altri veicoli e strumenti con cui i decisori pubblici acquisiscono informazione per decidere. Una "famiglia" di tali *inputs* è quella che in ambito anglosassone è conosciuta come *policy analysis*, cioè l'uso pragmatico degli strumenti analitici delle scienze sociali per migliorare la conoscenza di presupposti, costi, vincoli ed effetti delle politiche pubbliche.

La valutazione come controllo (VC) invece comprende attività analitiche che hanno per oggetto il concreto operare delle unità che gestiscono un intervento. Lo scopo è di verificare sistematicamente come tali unità operative utilizzano le risorse loro assegnate, quali costi vanno ad esse imputati, cosa producono e con quali caratteristiche qualitative, e come gli utenti diretti di ciascuna unità sono serviti. L'enfasi sulle unità operative (cioè i singoli uffici, scuole, ospedali, ripartizioni, servizi, linee di attività, ecc.) va messa in evidenza, in quanto rappresenta la discriminante fondamentale tra i due tipi di valutazione.

Se la valutazione come apprendimento è da considerarsi come un sottoinsieme della *policy analysis*, la valutazione come controllo è da considerarsi parte del controllo di gestione, che è stato definito come "l'insieme delle procedure e delle condizioni che permettono ai dirigenti di un'organizzazione di controllare nel modo migliore il funzionamento dell'organizzazione stessa (Bondonio e Scacciati, 1991, pag. 43)." A sua volta, il controllo di gestione può essere considerato uno degli strumenti di *management* pubblico.

I due approcci alla valutazione si differenziano lungo varie dimensioni. La prima è il diverso *tipo di committenza* della valutazione stessa. Il controllo di gestione presuppone una gerarchia (all'interno di un'organizzazione o tra livelli diversi di governo), e la sua essenza è proprio quella di provvedere informazioni ad un livello superiore della gerarchia sugli aspetti quantitativi e qualitativi dell'operato di unità subordinate. La mancanza di conoscenza a cui il controllo di gestione vuole supplire è riconducibile ad una fondamentale asimmetria informativa che esiste tra controllori e controllati (Mancini, 1993). Quindi il committente delle attività di valutazione come controllo tende ad essere la struttura sovraordinata, che poi condividerà o meno l'informazione raccolta con le unità subordinate a seconda della strategia di *management* adottata.

La valutazione come apprendimento supplisce ad un diverso tipo di carenza conoscitiva, non riconducibile ad una asimmetria informativa tra i diversi livelli di un'organizzazione complessa. La carenza conoscitiva riguarda la natura stessa dell'intervento, le sue possibilità di successo, l'adeguatezza del suo disegno, le risposte che ottiene dall'ambiente circostante, la corrispondenza tra la missione assegnata all'intervento e la concreta capacità di implementazione degli apparati ad essa preposti. Il committente può quindi essere del tutto esterno alla struttura amministrativa. Il destinatario ultimo di questo tipo di valutazione è rappresentato dagli organi decisionali che hanno competenza per la formulazione della politica che si intende valutare. Inoltre, in quei Paesi in cui le decisioni di *policy* siano anche soggette ad un controllo di merito da parte di istituzioni espresse dalla società civile, i risultati della valutazione possono anche essere utilizzati da una pluralità di attori che agiscono al di fuori degli apparati istituzionali (ad esempio *i media*, *i networks* di esperti, *i think tanks*, i gruppi di interesse, le associazioni del volontariato.)

La seconda dimensione è rappresentata dal *grado di copertura nella raccolta dei dati* per la valutazione. Le rilevazioni di dati e ricognizioni sul campo condotte nell'ambito della VA tendono ad essere circoscritte ad alcuna unità operative, località, o situazioni particolari, a privilegiare i campioni, i "carotaggi", gli studi di casi, l'esame approfondito di alcuni aspetti dell'intervento o di alcune realtà o situazioni in cui viene realizzato. Per contrasto, la valutazione come controllo tende a modellarsi sulla struttura esistente di produzione-erogazione dell'intervento, talvolta diventandone parte integrante, normale, di *routine*. La rilevazione di dati e informazioni tende ad essere universalistica piuttosto che condotta su base campionaria, o quantomeno tende a coprire in modo omogeneo tutte le unità operative che si intendono controllare.

I due approcci alla valutazione tendono anche a differire in termini di *cadenza temporale*. E' nella natura stessa dell'attività di controllo quella di essere sistematica, ripetuta a scadenze regolari (anni fiscali, periodi di esercizio) o effettuata su base continua. E' proprio invece della VA quello di

essere episodica, occasionale, di protrarsi nel tempo andando spesso fuori sincronia con la cadenza "normale" di erogazione del servizio.

In termini di *metodi analitici* usati, la VA tende a fare uso di un ampio spettro di metodi messi a disposizione dalle scienze sociali. Ad un estremo di questo spettro troviamo l'insieme di metodi analitici che si prefiggono come scopo l'individuazione dei legami causa-effetto tra i fenomeni sociali. Questa parte è rilevante soprattutto per l'analisi degli effetti dell'intervento, cioè della sua efficacia. All'estremo opposto troviamo metodi di indagine prettamente qualitativi, che tendono a privilegiare l'osservazione non-sistematica, l'uso di testimoni privilegiati, fino all'impiego di metodi etnografici. La VC si colloca invece in posizione intermedia lungo questo spettro, rimanendo lontana sia dal problema spinoso della causalità sia dall'uso di metodi di indagine qualitativi e non-strutturati, e privilegiando invece la raccolta sistematica, a tappeto, di informazioni di tipo descrittivo. Lo strumento analitico più frequentemente citato nella letteratura sul controllo di gestione è quello delle "batterie di indicatori". Oltre alla costruzione di tali indicatori, il controllo di gestione fa largo uso di strumenti di tipo contabile, che esulano dal dominio delle scienze sociali, e tendono invece a prendere a prestito strumenti propri da un lato delle discipline giuridiche e dall'altro delle tecniche gestionali di impresa.

La VA tende ad occuparsi delle fasi "critiche" nel ciclo di vita di un \_ intervento, prima nella fase di implementazione (in cui maggiore è l'incertezza sulla fattibilità), e poi nelle fasi in cui sia in discussione il mantenimento dell'intervento stesso, e quando si tenda a metterne in dubbio alcune caratteristiche fondamentali. Per contrasto, la VC ha più spazio quando l'intervento si è normalizzato, è andato a regime, e la questione si sposta dalle sue caratteristiche e funzioni fondamentali a come viene gestito. La VC può migliorare lo *status quo*, mentre la VA può contribuire a crearlo o a metterlo in crisi.

Un'ultima distinzione tra i due tipi di valutazione sta nel diverso peso assegnato alla valutazione degli aspetti di *efficacia* di un'attività pubblica e a quelli di *efficienza*. La distinzione tra efficienza ed efficacia è molto enfatizzata nella letteratura italiana sulla valutazione, spesso rappresentando il principale se non l'unico criterio di classificazione tra tipi diversi di valutazione. In che relazione sta la dicotomia efficienza-efficacia rispetto a quella appena introdotta di apprendimento-controllo? Va innanzitutto chiarito che le due distinzioni non si equivalgono in quanto si collocano su piani diversi.

Efficienza ed efficacia sono caratteristiche desiderabili di ogni attività pubblica, mentre qui distinguiamo tra due approcci diversi alla valutazione, entrambi i quali possono occuparsi sia di efficienza sia di efficacia. Va però riconosciuto che il grado di efficienza è un attributo che è più facilmente riferibile a singole unità operative, anziché ad un intervento nel suo complesso. Rimane difficile concettualizzare l'efficienza di *un'intera* politica pubblica, mentre è più facile farlo con riferimento all'attività di singoli uffici o ripartizioni amministrative. D'altro canto, l'efficacia di un intervento è

più difficilmente rilevabile o persino definibile per singole unità, mentre tende ad essere una caratteristica complessiva di un intervento, e richiede per essere valutata un apparato concettuale più sofisticato, di tempi più lunghi e di approfondimenti specifici, caratteristiche spesso incompatibili con la sistematicità e regolarità proprie del controllo di gestione.

In conclusione, il baricentro della valutazione come apprendimento tende ad essere spostato verso l'efficacia degli interventi pubblici, mentre il controllo di gestione tende per sua natura a prestare attenzione maggiore all'efficienza della produzione pubblica. Inoltre, va riconosciuto che al termine stesso di "efficacia" vengono attribuiti significati diversi, alcuni dei quali sono più appropriati in una prospettiva di controllo di gestione. Nella letteratura italiana su questo tema, ritroviamo l'efficacia definita ad esempio come "differenza tra risultati programmati e risultati ottenuti" (Bondonio e Scacciati, 1991, pag.103), o come "grado in cui l'offerta risulta essere anche qualitativamente rispondente alla domanda espressa indirettamente dai cittadini o direttamente dai responsabili politici" (Mancini, 1993, pag. 74). In questo lavoro usiamo il termine efficacia in un senso diverso da quelli appena riportati, e precisamente nel senso di "*capacità di un intervento di produrre gli effetti desiderati*", cioè vicino all'uso del termine inglese *effective-ness*. La domanda al centro della valutazione di efficacia così intesa diventa "l'intervento serve allo scopo per cui è stato introdotto?", piuttosto che "è stato gestito come programmato?" oppure "risponde ai desideri dell'utente?" Tutte queste domande sono legittime e pertinenti, ma è importante tenere presente che si tratta di domande diverse.

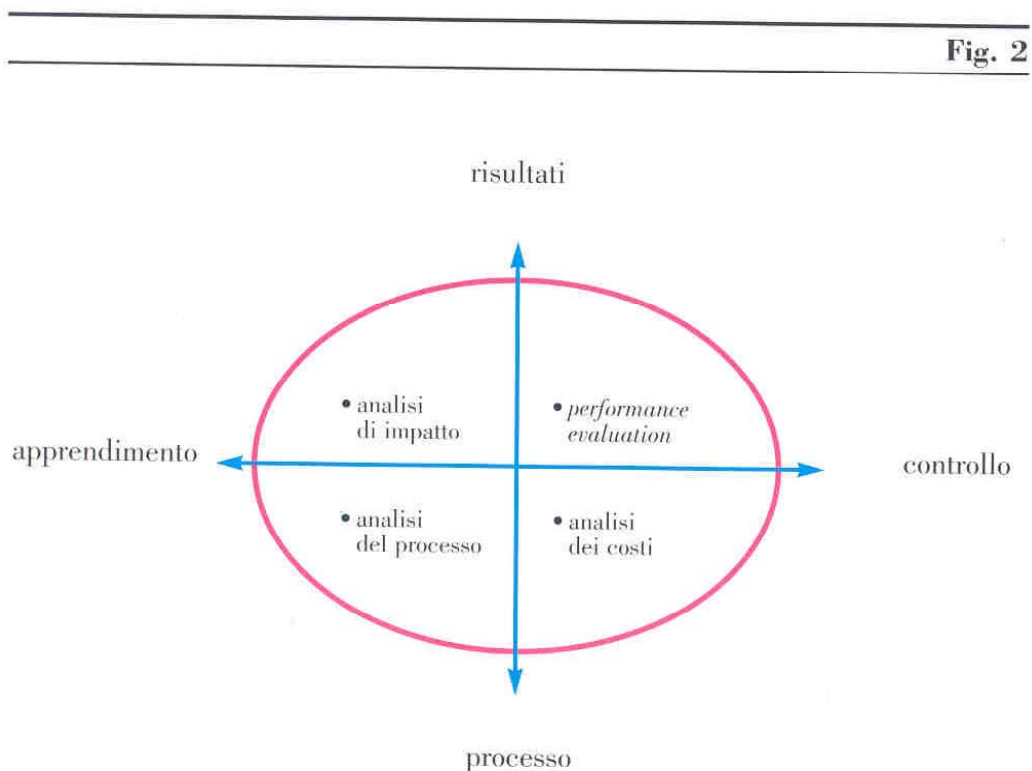
#### *La distinzione tra processo e risultati*

Alla seconda domanda preliminare alla valutazione posta all'inizio di questa sezione (*cosa si vuole valutare?*) molte risposte sono ovviamente possibili. Riteniamo che la distinzione più importante, propedeutica a tutte le altre, sia quella tra processo e risultati (*process e outcomes*). I risultati si riferiscono a tutti quegli aspetti di un'attività pubblica che si riflettono sinl'ambiente esterno, dai prodotti (*outputs*) dell'attività pubblica che sono utilizzati esternamente, alle modifiche dell'ambiente che l'attività pubblica intende produrre (*effects*). Il termine processo riassume invece tutti gli aspetti interni dell'attività pubblica, dall'acquisizione degli *inputs* produttivi, alle regole che determinano il loro utilizzo, fino al momento della trasformazione degli *inputs* in prestazioni, erogazioni, prodotti.

La Figura 2 è ottenuta incrociando la distinzione tra apprendimento e controllo con quella tra risultati e processo. Nei quattro quadranti che ne risultano è possibile posizionare un'ampia gamma di attività di valutazione, molto diverse tra loro per obiettivi e modalità di esecuzione. A titolo di esempio, nei quadranti della Figura 2 collochiamo quattro tipologie particolari, che peraltro non esauriscono l'intero spettro della valutazione, bensì rappresentano casi tipici o emblematici. Nel quadrante in alto a sinistra tro-

viamo l'analisi d'impatto, cioè quel tipo di valutazione che tenta di misurare l'efficacia intesa come capacità dell'intervento di ottenere gli effetti desiderati. A questo tipo di analisi dedicheremo una parte rilevante di questo rap-porto. Nel quadrante in basso a sinistra, collochiamo l'analisi del processo (*process analysis*), intesa come l'insieme di attività di rilevazione e analisi tese a stabilire se l'effettiva implementazione di un intervento corrisponde al disegno originario, ad identificare gli ostacoli che si frappongono ad una piena implementazione, e ad individuare i casi di eccellenza da usare come modelli (*best practice research*).

Spostandoci sulla parte destra, entriamo nell'ambito della valutazione come controllo. In alto a destra troviamo quella che nel mondo anglosassone è conosciuta come *performance evaluation*, cioè il monitoraggio dell'attività pubblica mediante una misurazione sistematica della quantità e qualità di ciò che viene prodotto e della corrispondenza degli indicatori *a standards* quantitativi e qualitativi. Infine, in basso a destra troviamo l'analisi dei costi, cioè quell'insieme di attività di misurazione che, utilizzando tecniche quali la contabilità analitica, tentano di imputare a ciascuna attività pubblica le risorse per essa utilizzate.



### *Interventi-pilota e interventi a regime*

Un'ulteriore distinzione applicabile al lato della valutazione come apprendimento, e messa in risalto nella Figura 3, è quella tra valutazione di interventi-pilota e valutazione di interventi a regime. Questa distinzione può essere interpretata come risposta alla domanda "a quale stadio nello sviluppo di una politica pubblica è utile condurre la valutazione?", ed è necessaria per comprendere l'esperienza di valutazione condotta negli Stati Uniti, dove ampio è il ricorso agli interventi-pilota, definiti col termine di *demonstrations*. -

Sul lato sinistro della Figura 3 collochiamo le valutazioni di interventi-pilota, in cui l'intervento stesso è concepito e realizzato con l'espresso (e principale) scopo di valutarne effetti e fattibilità. Si vuole verificare se l'intervento può funzionare, anticipare problemi nella sua implementazione, individuare opportunità e soluzioni, verificare se sia in grado di produrre gli effetti desiderati, almeno su una scala limitata.

Al centro della Figura 3 collochiamo invece la valutazione di interventi a regime. In questo caso la valutazione ha a che fare con un intervento che è stato istituito per l'intera popolazione o territorio. Siamo ancora in un

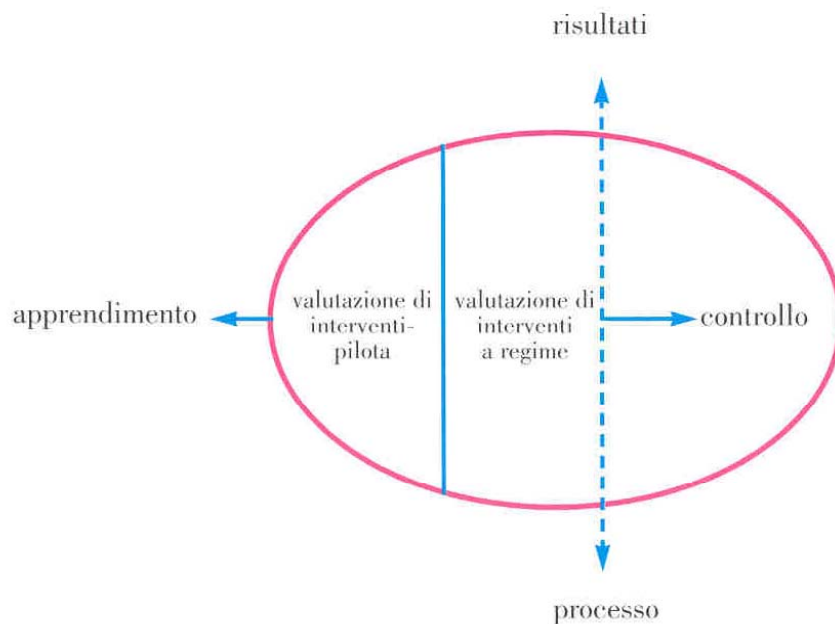


Fig. 3

ambito di valutazione come contributo al disegno delle politiche, ma l'enfasi è sul presente piuttosto che sul futuro. La politica è già in atto, la fase di progettazione è finita. Ora la valutazione vuole verificare se le strutture amministrative sono in grado di sopportarne l'onere, se l'implementazione risponde agli intenti originari, se l'intervento a regime è in grado di incidere sul fenomeno nel suo complesso. Il confine tra valutazione come apprendimento e come controllo in questo caso può diventare più sfocato, mentre nel caso degli interventi-pilota è chiaro che l'unica motivazione è quella dell'apprendimento.

Il ricorso a interventi-pilota preliminarmente alla realizzazione su larga scala presenta opportunità e limiti, alcuni dei quali verranno discussi nelle pagine seguenti. Il principale beneficio è quello di sperimentare con un costo limitato per la collettività soluzioni nuove, quindi evitando errori potenzialmente costosi derivanti dall'implementazione su vasta scala di soluzioni inefficaci. Il principale limite deriva invece dalla difficoltà di generalizzare ad un contesto globale quello che si è appreso in un contesto particolare. Su questo argomento ritorneremo più avanti. L'uso di interventi-pilota richiede poi molta lungimiranza e pazienza: possono passare anni prima che i risultati di una *demonstration* diventino disponibili, soprattutto se si vogliono osservare gli effetti dell'intervento. Questo ritardo mal si adatta a quei problemi percepiti (giustamente o meno) come emergenze, a quelle situazioni in cui i politici sentono di dover dare risposte immediate e visibili. Sul lato degli svantaggi del ricorso preliminare ad interventi-pilota va quindi incluso il rischio di ritardare l'azione su un problema già presente, e in taluni casi quello di offrire ai politici una giustificazione per rimandare l'ingrato (e costoso) compito di affrontare un problema nella sua globalità.

### **1.3 La valutazione di efficacia delle politiche contro la povertà**

Lo scopo della discussione precedente non era quello di individuare tipi di valutazioni più o meno importanti, bensì mettere in luce la loro diversità e chiarire la logica che sta dietro tale diversità.

Tutti i tipi di valutazione finora brevemente descritti (valutazione come controllo o come apprendimento, focalizzata sul processo o sui risultati, incorporata in interventi-pilota o applicata ad interventi a regime) sono in linea di principio rilevanti per valutare le politiche contro la povertà e l'esclusione sociale.

E' evidente che anche queste politiche comportano l'erogazione di servizi da parte di unità operative. L'aspetto di produzione pubblica può essere più o meno marginale rispetto alla portata complessiva dell'intervento, ma è pur sempre presente. Persino le politiche passive contro la povertà hanno una importante dimensione di produzione pubblica, in quanto anche l'erogazione di sussidi comporta operazioni quali l'accettazione di domande di sussidio, la verifica dei requisiti di ammissione, l'effettuazione di pagamen-

ti, operazioni che sono solitamente gestite da una pluralità di uffici dislocati sul territorio, e che possono essere oggetto di verifica sistematica mediante controllo di gestione. Ad esempio, è importante che i costi per unità servita siano rilevati, che l'esistenza di sovraccarichi di lavoro o sottoutilizzi del personale sia individuata, che differenze nel livello qualitativo nel servizio agli utenti siano rilevate e analizzate. Quindi tecniche di controllo di gestione sono rilevanti per l'implementazione delle politiche contro la povertà.

Tuttavia, l'attenzione di questo lavoro è sulla valutazione *dell'efficacia* di queste politiche, efficacia intesa nel senso ristretto di "capacità di produrre gli effetti desiderati." Questa scelta ci pone chiaramente nell'ambito della valutazione come apprendimento, e rende poco rilevante la filosofia complessiva del controllo di gestione. Inoltre, la nostra scelta pone l'enfasi sulla valutazione dei risultati, anziché del processo. Conoscere a fondo il processo con cui un intervento è realizzato è importante per capire perché è efficace oppure non lo è, ma resta un'attività analitica distinta.

Va ribadito che la scelta di focalizzare l'attenzione sulla valutazione di efficacia non implica sminuire l'importanza di altri tipi di valutazione. Riteniamo tuttavia che sia importante mettere in evidenza che tipi diversi di valutazione sono appropriati in situazioni diverse, e che possono soddisfare esigenze conoscitive tra loro differenti.

Un intervento è efficace se i cambiamenti osservati in un certo fenomeno sociale (la cui modifica costituisce la *ratio* dell'intervento) sono da attribuirsi all'esistenza e alle caratteristiche dell'intervento stesso, nel senso che esiste tra i due un legame di causa-effetto. L'identificazione di tale legame non è un esercizio fine a sé stesso, bensì rappresenta la base logica per successive elaborazioni, soprattutto per calcoli di costo-efficacia e costi-benefici. L'essenza di queste elaborazioni è quella di verificare se i benefici, cioè gli effetti positivi che l'intervento produce, giustificano i costi sopportati per realizzarlo. Se non si possono attribuire con ragionevole sicurezza gli effetti all'intervento, non è possibile condurre quest tipo di elaborazioni. Se l'intervento costa ma non produce effetti (e i cambiamenti osservati sono da attribuirsi ad altri fattori e processi) è improbabile concludere che esso sia costo-efficace.

Sulla base della definizione di efficacia come "capacità di ottenere gli effetti desiderati", la principale domanda a cui la valutazione vuole rispondere può essere formulata come segue:

D1. (Di quanto) l'intervento è stato in grado di incidere sul fenomeno o problema che intendeva correggere o eliminare?

Volendo dare maggiore concretezza a questa domanda, è possibile riformularla in relazione a casi particolari relativi alle politiche contro la povertà e l'esclusione sociale. Ad esempio:

D1.1 (Di quanto) si è ridotta l'incidenza dei casi di denutrizione tra gli anziani soli grazie all'istituzione del servizio di refezione a domicilio?

D1.2 (Di quanto) è diminuita l'incidenza (o l'intensità) della povertà *grazie* all'introduzione del minimo vitale?

D1.3 (Di quanto) è aumentato il tasso di occupazione tra i portatori di handicap *grazie* all'introduzione delle norme di assunzione preferenziale?

D1.4 (Di quanto) è diminuita la numerosità dei senza-dimora *grazie* alla creazione di nuove strutture di accoglienza?

D1.5 (Di quanto) si è ridotto il tasso di criminalità tra i minorenni *grazie* all'introduzione di iniziative di socializzazione sul territorio?

È utile a questo punto una breve digressione. Politici e amministratori assumono spesso un atteggiamento che rende il tipo di domande appena esemplificato del tutto irrilevanti: definiremo questo atteggiamento come "presunzione di efficacia". Tale atteggiamento non è dettato tanto dalla mancanza di strumenti di rilevazione e di analisi, quanto piuttosto da un misto di convincimenti aprioristici, quindi difficili da scalfire. Questo atteggiamento vede l'efficacia come un fatto scontato, una necessità logica, che sta alla base dell'introduzione dell'intervento stesso. Tornando agli esempi fatti, per chi adotta questo atteggiamento le iniziative di socializzazione sul territorio *non possono che* ridurre il tasso di criminalità tra i minorenni, per questo sono state introdotte; il minimo vitale *non può che* diminuire l'incidenza della povertà, per questo è stato istituito; le nuove strutture di accoglienza *non possono che* diminuire la numerosità dei senza-dimora, per questo sono state create.

Come si intuisce dagli esempi elencati, la plausibilità di tale "presunzione di efficacia" varia da caso a caso. Quello che qui preme sottolineare è che la presenza di questo atteggiamento, giustificato o meno che sia, elimina la motivazione fondamentale che sta alla base di ogni valutazione di efficacia, cioè *l'esistenza dell'incertezza* sulla capacità dell'intervento di ottenere gli effetti desiderati. In assenza questa incertezza, la valutazione di efficacia non ha ragione di essere. Si tratta quindi di un problema essenzialmente culturale. Quando prevale lo scetticismo nei confronti dell'intervento pubblico, o quantomeno quando si diffonde la consapevolezza che l'intervento pubblico, per quanto "giusto" e "desiderabile" a priori, possa produrre effetti non voluti, o semplicemente possa fallire nell'ottenere gli effetti desiderati, si creano le premesse "culturali" per la valutazione di efficacia.

Va anche ribadito che, per motivare la valutazione di efficacia nel senso usato in questo lavoro, non è sufficiente l'incertezza sulla capacità, da parte di particolari unità produttive, di svolgere in modo adeguato la propria attività. Questo tipo di incertezza motiva essenzialmente l'uso di tecniche di controllo di gestione. *L'incertezza a cui facciamo riferimento va al di là della semplice inadeguatezza organizzativa e investe la natura stessa dell'intervento e la sua adeguatezza strutturale a risolvere il problema o alleviare il fenomeno che ne ha motivato l'introduzione.* In questo senso la valutazione di efficacia, come intesa in questo lavoro, si colloca a pieno titolo nella prospettiva di valutazione come apprendimento.

Nel caso delle politiche contro la povertà e l'esclusione sociale, la presunzione di efficacia sembra particolarmente poco plausibile, soprattutto nel caso delle politiche attive. Il tratto comune di queste politiche è il tentativo di modificare comportamenti e condizioni che sono il risultato di processi sociali molto complessi, rispetto ai quali poco si conosce in termini di cause, interazioni e dinamiche. Ci pare quindi che porsi la domanda "l'intervento è capace di ottenere gli effetti desiderati?" sia in gran parte dei casi plausibile quando si tratti di valutare queste politiche pubbliche.



## 2. IL DISEGNO DELLA VALUTAZIONE DI EFFICACIA

Le fasi fondamentali di un disegno di valutazione sono le seguenti: (i) la definizione delle domande a cui la valutazione vuole rispondere; (ii) l'individuazione della metodologia di analisi più appropriata per dare una risposta alle domande poste; (iii) la definizione del fabbisogno informativo richiesto dall'analisi; (iv) la ricognizione delle fonti disponibili e dei metodi (e costi) per la raccolta dei dati non disponibili; (v) un piano per la disseminazione e l'utilizzo dei risultati.

In questo lavoro ci soffermiamo sulle prime tre fasi. Nella sezione 2.1 iniziamo offrendo un esempio altamente stilizzato di un disegno di valutazione di efficacia. Si tratta di un approccio utile per chiarire alcune questioni metodologiche di fondo, ma non adatto come base per una concreta valutazione. Le ragioni di questa limitazione verranno chiarite in chiusura di sezione. Nella sezione 2.2 presentiamo invece un disegno di valutazione più realistico, pensato più specificamente per valutare le politiche contro la povertà e l'esclusione sociale.

### 2.1 La valutazione dell'efficacia (complessiva) di un intervento

Valutare l'efficacia di un intervento significa essenzialmente stabilire un legame di causa-effetto tra la realizzazione dell'intervento e le modifiche osservate nel fenomeno che l'intervento intende modificare. Si tratta cioè di identificare l'*impatto* dell'intervento sul fenomeno in questione. Per questo la valutazione di efficacia, soprattutto nel contesto americano, viene comunemente definita come *analisi di impatto* (Mohr, 1988; Rossi e Freeman, 1989; Blalock, 1990). Dal punto di vista terminologico, l'uso del termine "impatto" ha il vantaggio di comprendere anche possibili effetti non voluti, mentre il termine "efficacia" tende ad avere sempre una connotazione positiva.

In generale, un intervento può avere sul fenomeno molteplici tipi di effetti, anche di segno opposto. Questa molteplicità non rappresenta un problema metodologico dallo stretto punto di vista dell'analisi di impatto: la stessa metodologia può essere applicata alle diverse variabili, o indicatori, che corrispondono ai diversi effetti. La molteplicità rappresenta invece un problema in sede di utilizzo dei risultati della valutazione: se gli effetti dell'intervento sono di segno diverso, il *policy-maker* dovrà trovare un metodo per "pesare" gli effetti sulle diverse dimensioni del fenomeno, e dare un giu-

dizio complessivo sull'intervento. Questa problematica, per quanto fondamentale, fuoriesce tuttavia da quella affrontata in sede di analisi di impatto, e non verrà qui trattata.

*Post hoc ergo propter hoc?*

È utile iniziare la discussione dell'analisi di impatto mettendo in guardia il lettore contro la principale fallacia da evitare in questo tipo di analisi, lo scambiare la causalità (*propter hoc*) con la sequenzialità temporale (*post hoc*) tra intervento e modifica osservata nel fenomeno. Questa fallacia può essere intuitivamente illustrata mediante la sostituzione di "grazie a" con "in seguito a" nelle domande D1.1-D1.5: ad esempio, di quanto si è ridotta l'incidenza di casi di denutrizione tra gli anziani soli *in seguito* all'istituzione del servizio di refezione a domicilio? Di quanto è aumentato il tasso di occupazione tra i portatori di handicap *in seguito* all'introduzione delle norme di assunzione preferenziale? Se la causalità viene ridotta a semplice sequenzialità temporale, la valutazione di impatto tende a ridursi alla misurazione della variazione del fenomeno tra un momento precedente all'introduzione dell'intervento e un momento successivo. Per mostrare la struttura logica di questa fallacia, è utile introdurre un minimo di formalizzazione. L'impatto come differenza prima-dopo può essere espresso come:

(T) impatto = indicatore nel periodo T - indicatore nel periodo T-K

dove l'intervento si pone sia stato introdotto nel tempo intercorrente tra il periodo T-K (pre-intervento) e T (post-intervento). In questo contesto, "indicatore" significa semplicemente "rappresentazione quantitativa di un aspetto rilevante del fenomeno che l'intervento vuole modificare." Negli esempi utilizzati più sopra, gli indicatori sono l'incidenza della povertà in una certa popolazione, l'incidenza di casi di denutrizione tra gli anziani, il tasso di occupazione tra i portatori di handicap, la numerosità dei senza-dimora.

Secondo la (1), la differenza "prima-dopo" fornirebbe direttamente una misura dell'impatto dell'intervento: La fallacia è quella di attribuire all'intervento tutto il cambiamento osservato nell'indicatore, o di interpretare l'eventuale mancanza di cambiamento come prova del fallimento dell'intervento. La ragione di tale fallacia è che si ignora la presenza di altri possibili fattori e processi che determinano l'evoluzione del fenomeno nel tempo. Ridurre la valutazione di efficacia ad un semplice confronto prima-dopo significa presumere che il fenomeno in questione sia privo di una *dinamica propria* che esiste a prescindere dall'intervento. In altre parole, usare la (1) come misura di impatto significa presumere a priori che il fenomeno non possa cambiare in assenza di intervento. Questa presunzione contrasta con l'evidenza empirica e l'esperienza quotidiana, le quali offrono infiniti esempi di come i fenomeni sociali si evolvano di continuo, e siano soggetti a molteplici influenze, tra cui l'intervento pubblico è solo una delle determinanti, e sovente del tutto marginale.

### *L'impatto come differenza tra osservato e controfattuale*

La definizione corretta di impatto è la seguente:

$$(2) \text{ impatto} = \text{indicatore nel periodo } T - \text{indicatore nel periodo } T \text{ se l'intervento non fosse stato attuato}$$

Il secondo termine sul lato destro della (2) è ovviamente una quantità ipotetica, cioè non direttamente osservabile (qui e nel seguito utilizziamo il corsivo per indicare quantità ipotetiche.) Nella letteratura sulla valutazione questa quantità viene definita come "situazione controfattuale", e rappresenta quello che, ipoteticamente, sarebbe successo se l'intervento non fosse stato attuato. Quindi l'impatto di un intervento è *definito* come differenza tra ciò che si osserva e ciò che presumibilmente si sarebbe osservato in assenza di intervento.

Poiché la situazione controfattuale non è osservabile (per definizione), il compito primario della valutazione di impatto è quello di trovare per essa un'approssimazione plausibile. La semplice individuazione di indicatori e la loro misurazione prima e dopo l'intervento non è quindi sufficiente, in quanto produce soltanto:

$$(3) \text{ cambiamento osservato} = \text{indicatore nel periodo } T - \text{indicatore nel periodo } T-K$$

Per illustrare l'essenza della valutazione di efficacia, è utile aggiungere e sottrarre dal lato destro della (3) il termine che rappresenta la situazione controfattuale, ottenendo:

$$(4) \text{ cambiamento osservato} = [\text{indicatore nel periodo } T - \text{indicatore nel periodo } T \text{ se l'intervento non fosse stato attuato}] + [\text{indicatore nel periodo } T \text{ se l'intervento non fosse stato attuato} - \text{indicatore nel periodo } T-K]$$

Sostituendo dalla (2) otteniamo:

$$(5) \text{ cambiamento osservato} = \text{impatto} + [\text{indicatore nel periodo } T \text{ se l'intervento non fosse stato attuato} - \text{indicatore nel periodo } T-K]$$

La differenza contenuta nella parentesi sul lato destro della (5) rappresenta di quanto il fenomeno si sarebbe modificato in assenza di intervento.

Q uindi rappresenta ciò che la dinamica del fenomeno avrebbe comunque prodotto, a prescindere dall'esistenza di quell'intervento pubblico. Se questa dinamica non produce alcun cambiamento, il termine in parentesi sul lato destro della (5) si annulla, e il cambiamento osservato nell'indicatore rap-

presenta davvero l'impatto dell'intervento. Ma a prescindere da questo caso estremo (che, come si notava in precedenza, è difficile da immaginare per gran parte dei fenomeni sociali) *il cambiamento osservato non è una misura accettabile dell'impatto dell'intervento.*

Ignorare l'esistenza di una dinamica autonoma del fenomeno può portare a grossolani errori di interpretazione, che illustriamo mediante i grafici contenuti nella Figura 4 e 5. Il tempo è rappresentato sull'asse orizzontale, e l'indicatore sull'asse verticale. Nella Figura 4, la dinamica autonoma del fenomeno porta l'indicatore a livelli più alti del livello di partenza A: in assenza di intervento, produrrebbe un miglioramento pari ad AB. Con l'intervento, invece, il cambiamento osservato è pari ad AC. Attribuire all'intervento tutto il cambiamento osservato AC significa però *sovrastimare* l'impatto dell'intervento, che corrisponde soltanto alla quantità BC (la differenza tra osservato e controfattuale.) All'intervento verrebbero *attribuiti* meriti che non ha.

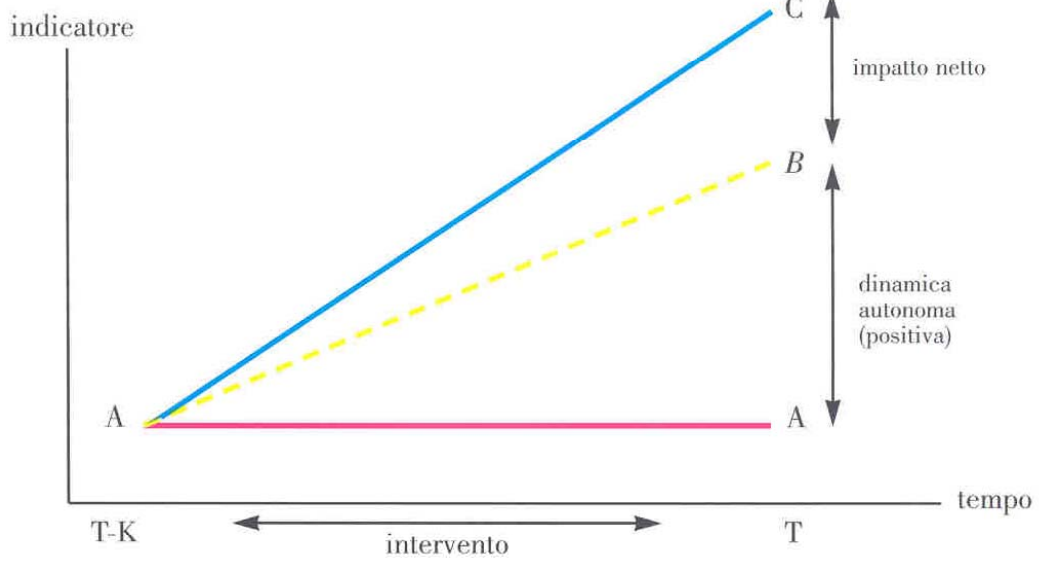
Nella Figura 5 l'errore è di segno opposto, e l'impatto viene *sottostimato* se si usa il cambiamento osservato. Qui il fenomeno ha una dinamica autonoma negativa, che lo porta ad un peggioramento uguale ad AB. Usare la differenza osservata AC porta in questo caso ad una sottostima del vero impatto BC. In altre parole, all'intervento vengono *negati* meriti che invece ha.

- Evitare queste fallacie richiede l'uso di metodi analitici che, sfruttando l'informazione sulla "storia" del fenomeno osservata nel periodo precedente all'introduzione dell'intervento, siano in grado di proiettare questa storia fino al momento in cui si conduce la valutazione, generando una misura della situazione controfattuale B, di quello che sarebbe successo senza l'intervento. Questi metodi dovrebbero cioè ricostruire la linea tratteggiata nelle Figure 4 e 5.

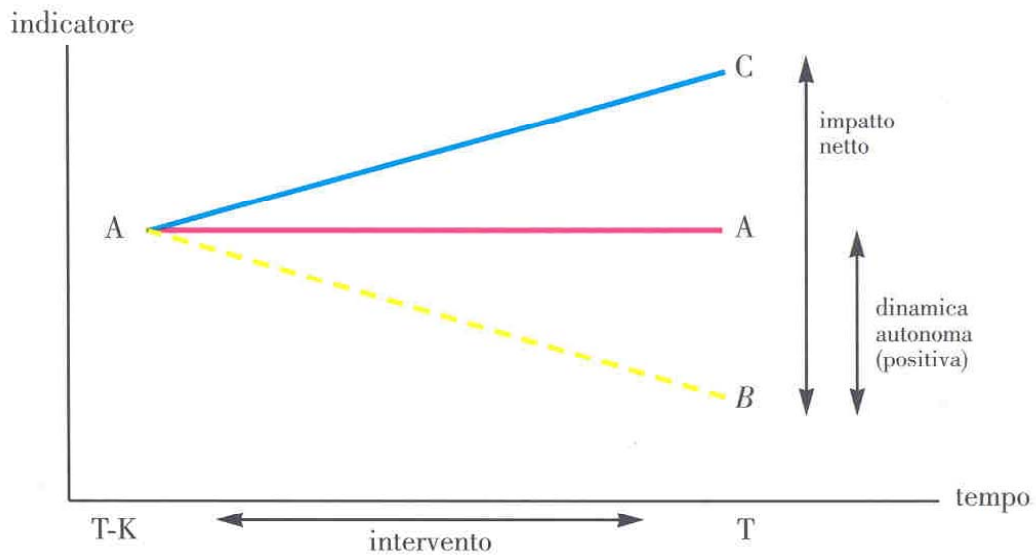
Esistono metodologie statistiche ed econometriche che consentono, almeno in linea di principio, di condurre l'operazione appena descritta. Non è tuttavia utile soffermarsi su tali metodologie, in quanto il disegno di valutazione che esse sottendono è piuttosto semplicistico, ed è stato introdotto per pure finalità espositive, per mettere in evidenza alcuni concetti generali. Quanto esposto finora nella Sezione 2.1 intendeva attrarre l'attenzione del lettore su due concetti fondamentali: (a) per valutare l'impatto di un intervento non è sufficiente trovare indicatori che *descrivono* l'evoluzione del fenomeno nel tempo; (b) occorre di più, occorre cioè ricostruire quello che sarebbe successo se l'intervento non fosse stato attuato.

Questo è il problema conoscitivo fondamentale per qualsiasi valutazione di efficacia, nel senso inteso in questo lavoro. Tuttavia è prematuro approfondirlo in questo contesto, in quanto la domanda D1 posta all'inizio della sezione è troppo astratta, troppo generica per rappresentare un utile disegno di valutazione. Questo per tre ragioni, su cui è utile soffermarsi prima di passare ad illustrare un disegno di valutazione più realistico e plausibile.

**Fig. 4**



**Fig. 5**



Innanzitutto, i metodi statistici ed econometrici a cui si è appena accennato richiedono la disponibilità di dati (ad esempio, serie storiche molto "lunghe") che sono raramente disponibili su fenomeni di interesse per le politiche sociali. Il rischio è quindi che la disponibilità di dati guidi la formulazione delle domande valutative, e non viceversa.

In secondo luogo, così come è formulata, la domanda D1 ("l'intervento è stato in grado di incidere sul fenomeno o problema che intendeva correggere o eliminare?") è posta ad un livello troppo aggregato, cioè fa riferimento all'intervento come ad un unico processo compatto ed indistinto. Il rischio è che il risultato della valutazione diventi la proverbiale "*black box*", impedendo di far luce sui meccanismi concreti che possono o meno rendere efficace l'intervento. Una domanda formulata a questo livello di generalità difficilmente offre utili elementi informativi ai *policy-makers*. La risposta alla D1 rischia di diventare una mera curiosità, un esercizio accademico, per quanto condotto con professionalità e rigore.

Infine, la domanda D1 è formulata con riferimento a come l'intervento è stato realizzato nel passato. Tuttavia, le domande conoscitive espresse dal processo decisionale pubblico tendono per loro natura a guardare al futuro, non al passato. Una domanda alternativa che ha più probabilità di sollevare l'interesse dei *policy-makers* è:

D2. Quali caratteristiche deve avere l'intervento per ottenere gli effetti desiderati sul fenomeno o problema che si intende correggere o eliminare?

Così formulata, la D2 non è una domanda da cui si può partire direttamente per impostare un disegno di valutazione. Il modo corretto per affrontare la D2 è quello di analizzare separatamente i vari aspetti dell'intervento che possono contribuire, positivamente e negativamente, alla sua efficacia complessiva. Questo è l'oggetto della prossima sezione.

## **2.2 Un disegno più analitico (e realistico) di valutazione**

Il punto di partenza per il disegno di valutazione qui proposto è la modellazione dell'intervento e delle sue interazioni con il fenomeno che vuole modificare. Si parte dall'individuazione delle diverse determinanti che possono influenzare (positivamente o negativamente) l'efficacia complessiva dell'intervento, mentre ad una valutazione complessiva si arriva solo per gradi, mettendo insieme spezzoni di evidenza empirica ottenuti con i metodi analitici che meglio si adattano a ciascuna determinante dell'efficacia.

La discussione che segue dovrà scontare un certo grado di astrattezza, in quanto è applicata ad una generica politica pubblica che si ponga l'obiettivo di modificare comportamenti o condizioni in una specifica popolazione. Il vantaggio è quello di fornire una metodologia generale, che poi adatteremo successivamente a singole tipologie di politiche contro la

povertà e l'esclusione sociale, distinguendo in particolare tra politiche attive e politiche passive.

Nel caso di una generica politica pubblica di questo tipo, le componenti che possono contribuire (positivamente o negativamente) all'efficacia complessiva di un intervento sono:

- l'impatto diretto sui beneficiari dell'intervento;
- l'adeguatezza della scelta dei beneficiari;
- l'adeguatezza della scala dell'intervento rispetto alla dimensione del fenomeno;
- il verificarsi di effetti negativi, tra cui soprattutto:
  - effetti negativi di spiazzamento sui soggetti che non beneficiano direttamente dell'intervento;
  - effetti di attrazione (*entry effects*) che inducono alcuni soggetti a modificare il proprio comportamento per poter usufruire dell'intervento.

Nella Figura 6, l'efficacia complessiva dell'intervento è rappresentata come la "somma" delle quattro distinte componenti. Tale somma non è da interpretarsi in modo letterale, in quanto l'efficacia complessiva non può essere veramente scomposta in una somma algebrica. Tuttavia, il segno con cui i vari termini entrano nell'"equazione" rappresenta la direzione in cui ciascuna componente influenza l'efficacia complessiva dell'intervento.

Fig. 6

### LE DETERMINANTI DELL'EFFICACIA GLOBALE

efficacia globale	=	impatto diretto sui beneficiari	+	sceita corretta dei beneficiari	+	scala adeguata dell'intervento	-	effetti non voluti (spiazzamento) ( <i>entry effects</i> )
l'intervento è stato in grado ridurre o eliminare il fenomeno?		l'intervento ha un impatto positivo sui diretti beneficiari?		la scelta dei beneficiari diretti dell'intervento è corretta?		la scala con cui l'intervento è realizzato è adeguata rispetto alla dimensione del fenomeno che vuole correggere?		l'impatto positivo sui beneficiari è ottenuto al prezzo di danneggiare altri soggetti in condizioni analoghe di bisogno? si verificano effetti perversi di "attrazione" verso la condizione di povertà/esclusione che l'intervento vuole correggere?

### *L'impatto diretto sui beneficiari*

L'efficacia complessiva di un intervento dipende in prima istanza dalla misura in cui *i beneficiari* (utenti, fruitori, destinatari) sono stati "influenzati" dalla loro diretta esposizione all'intervento. Se l'intervento non ha avuto un effetto positivo sui suoi diretti beneficiari, diventa improbabile che lo abbia sul fenomeno preso complessivamente. L'incertezza riguardo all'impatto sui diretti beneficiari è particolarmente rilevante nel caso delle politiche attive, come avremo modo di sottolineare più avanti, ma può esserlo anche per le politiche passive. Nel caso delle politiche attive, la valutazione di impatto tenta di stabilire se l'intervento abbia effettivamente modificato comportamenti o condizioni tra i suoi diretti beneficiari. Ad esempio, per valutare l'efficacia di un intervento per ridurre l'incidenza di comportamenti devianti tra giovani a rischio, si tratta innanzitutto di stabilire se i giovani che hanno direttamente beneficiato dell'intervento hanno modificato i loro comportamenti rispetto a quello che si sarebbe osservato in assenza di intervento. Nel caso delle politiche passive, l'impatto diretto sui beneficiari, che a prima vista può apparire scontato, può essere ridotto o eliminato se l'intervento innesca tra i beneficiari stessi comportamenti controproducenti.

### *L'adeguatezza della scelta dei beneficiari (targeting)*

La presenza di un impatto diretto sui beneficiari è condizione *necessaria ma non sufficiente* perché si osservi un impatto sul fenomeno complessivo che l'intervento vuole modificare. Una ragione di mancata efficacia è la scelta "errata" dei beneficiari. Il fenomeno nel suo complesso può essere misurato con riferimento ad una popolazione che coincide solo parzialmente con la platea dei beneficiari effettivi dell'intervento. In questo caso, il processo di selezione dei beneficiari non è congruente rispetto agli obiettivi che l'intervento si pone.

Questa mancata coincidenza può avere cause molto diverse, a seconda del contesto e della natura dell'intervento. Ad esempio, l'aumento del salario minimo negli Stati Uniti è tradizionalmente considerato una misura con una efficacia limitata nel ridurre la povertà (anche se può essere considerata desiderabile per altri motivi, ad esempio per ridurre la disuguaglianza salariale o i disincentivi al lavoro) in quanto la gran parte dei lavoratori che percepiscono il salario minimo non vivono in famiglie povere (Burkhauser e Finegan, 1988). L'aumento del salario minimo ha senza dubbio un impatto positivo sul reddito da lavoro dei beneficiari dell'aumento, quantomeno se non si verificano effetti negativi sulla loro probabilità di restare occupati. Tuttavia, se l'impatto è misurato usando il tasso di povertà come indicatore di successo, tale impatto sarà decisamente ridotto a causa della limitata presenza di lavoratori a salario minimo tra i poveri.

La scelta errata dei beneficiari come motivo di mancata o ridotta efficacia complessiva di un intervento è rilevante sia per le politiche attive sia

per le politiche passive. Nel caso delle politiche attive, particolare attenzione ha ricevuto quello che la letteratura anglosassone definisce con il termine "scrematura" (*creaming*): quando gli operatori del servizio hanno la possibilità di scegliere con margini discrezionali i beneficiari, questi ultimi tenderanno ad essere scelti tra coloro che hanno maggior probabilità di far fare "bella figura" agli operatori, indipendentemente dalla loro effettiva condizione di bisogno o dalla possibilità che la loro condizione venga migliorata *per effetto* dell'intervento. Nel caso delle politiche passive il fenomeno più analizzato è invece quello del *leakage*, termine con cui si definisce la situazione in cui una parte del trasferimento di risorse viene perso a favore di soggetti che non si trovano in una situazione di bisogno, a causa di una insoddisfacente definizione o messa in opera delle condizioni di ammissibilità al sussidio.

#### *L'adeguatezza della scala dell'intervento*

Anche in presenza di un impatto positivo sui beneficiari e di scelta adeguata degli stessi, è possibile che l'efficacia complessiva dell'intervento sia ridotta o nulla perché la scala dell'intervento è gravemente sotto dimensionata rispetto alla dimensione del fenomeno che si vuole modificare, cioè quando i beneficiari serviti sono solo una frazione della platea dei beneficiari potenziali.

Due principali tipi di situazioni possono dare luogo ad una scala inadeguata dell'intervento. Da un lato, la limitatezza delle risorse messe a disposizione per finanziare l'intervento fa sì che solo una parte dei potenziali beneficiari possa essere servita (razionamento dell'offerta). Dall'altro, si possono verificare situazioni in cui sono i beneficiari stessi a rinunciare ad usufruire dell'intervento pur avendone diritto, per motivi che variano da situazione a situazione (mancanza di informazione, "stigma", difficoltà di accesso, valori e preferenze individuali in contrasto con gli obiettivi dell'intervento.)

L'inadeguatezza della scala può verificarsi sia nel caso delle politiche attive sia di quelle passive. Nel caso delle politiche attive è più frequente che siano i beneficiari stessi a rinunciare ad usufruire dell'intervento, ad esempio perché non è loro chiaro quale beneficio ne possano trarre, mentre la partecipazione impone loro dei costi, psichici o materiali, o perché l'intervento tenta di imporre una modifica del loro comportamento in contrasto con le loro preferenze. Nell'ambito delle politiche passive, è più frequente il caso di razionamento delle risorse sul lato dell'offerta, anche se si osservano situazioni di semplice non-fruizione di un sussidio a cui si ha diritto. Ad esempio, negli Stati Uniti il *take-up rate*, cioè la percentuale di aventi diritto che usufruiscono dei sussidi di povertà, è sempre inferiore al 100 per cento, scendendo in alcuni casi fino al 50-60 per cento. Questo ovviamente riduce gli effetti dell'intervento sul fenomeno in esame, quando questo sia definito rispetto all'intera platea degli aventi diritto.

### *Il verificarsi di effetti negativi*

Sotto questa rubrica si può far rientrare una varietà di situazioni in cui l'intervento genera effetti non voluti tali da ridurre o eliminare l'efficacia complessiva dell'intervento (anche in presenza di un impatto diretto positivo sui beneficiari, di scelta adeguata degli stessi e di scala adeguata dell'intervento.) Due sono i tipi di effetti negativi di cui ci pare più rilevante discutere nel contesto delle politiche contro la povertà: gli effetti di spiazzamento e quelli di attrazione.

La possibilità di *spiazzamento* di soggetti diversi dai diretti beneficiari va considerata soprattutto nel caso delle politiche attive. L'effetto positivo a favore di un beneficiario può venire "pagato" con un danno arrecato ad un non-beneficiario che si trovi in un'analoga situazione di bisogno. Ad esempio, la preferenza nell'accesso al lavoro accordata ad un soggetto portatore di *handicap* può essere pagata spiazzando altri soggetti che si trovano in un'analoga situazione di esclusione dal mercato del lavoro, ma per motivi diversi dall'essere portatori di *handicap*. Alla possibilità di spiazzamento viene dato ampio risalto nella valutazione delle politiche attive *del lavoro*, soprattutto quando l'efficacia di queste politiche è misurata in termini di aumento netto dell'occupazione, nell'intera economia, in un certo comparto o zona geografica. Nel caso invece delle politiche contro la povertà, la rilevanza dello spiazzamento è più limitata, soprattutto quando entrino in gioco considerazioni di equità. Quanto lo spiazzamento vada "contato" nel misurare l'efficacia complessiva dipende in modo cruciale dal peso relativo assegnato ai due tipi di soggetti, cioè coloro che beneficiano dell'intervento e coloro che ne vengono spiazzati. Quanto più l'intervento è mirato a particolari popolazioni, tanto meno è probabile che gli spiazzati rientrino nella stessa popolazione considerata per definire il fenomeno oggetto di intervento. Ad esempio, è possibile che gli interventi a favore di gruppi ristretti di disoccupati di lungo periodo abbiano effetti negativi diffusi sulla massa dei disoccupati, ma molto diluiti a causa della diversa dimensione dei due gruppi. Se l'obiettivo dell'intervento è quello di ridurre la disoccupazione di lungo periodo, tale spiazzamento può essere considerato un effetto accettabile dal punto di vista della redistribuzione delle opportunità di lavoro.

Il secondo tipo di effetto negativo, che abbiamo definito come "effetto di attrazione", è più rilevante per le politiche passive. Si tratta di quello che la letteratura anglosassone definisce come *entry effect*: alcuni individui sarebbero indotti a modificare strategicamente il proprio comportamento per poter essere ammessi a beneficiare dell'intervento. Questa possibilità è ovvia nel caso dei trasferimenti, ed un caso speciale di *entry effect* è quello dei disincentivi al lavoro: l'esistenza di un sussidio superiore o uguale al salario di riserva induce alcuni individui a restare o diventare disoccupati. L'esistenza di *entry effects* riduce l'efficacia dell'intervento in quanto tende ad alimentare il fenomeno che l'intervento stesso tenta di eliminare o ridurre. Si crea così un circolo vizioso: quanto più l'intervento è efficace per i

beneficiari, quanto più la platea dei beneficiari tende ad allargarsi. È quindi possibile che l'intervento sia adeguato sotto tutti gli altri punti di vista (impatto diretto, scelta dei beneficiari, scala), ma il fenomeno non si riduca in quanto è l'intervento stesso ad alimentarlo.

#### *Sovrapposizioni tra le quattro determinanti dell'efficacia*

Prima di procedere, è opportuno notare che, in pratica, la distinzione tra le quattro determinanti appena discusse non è sempre così netta, ed esistono sovrapposizioni e zone grigie. Per cominciare, l'inadeguatezza nella scelta dei beneficiari e l'inadeguatezza nella scala dell'intervento possono in alcuni casi essere visti come due facce dello stesso problema: ad esempio quando la scarsità di risorse che determina la scala limitata di un intervento è causata dal fatto che le risorse sono state utilizzate per interventi analoghi a favore di soggetti in condizione di minor bisogno. Questa situazione può essere letta sia come scala inadeguata dell'intervento sia come scelta errata dei beneficiari.

Un altro caso di sovrapposizione è quello in cui un intervento, a causa della difficoltà di individuare correttamente i beneficiari, venga messo in opera in modo limitato, così da favorire un'autoselezione dei beneficiari stessi. Questa situazione si presenta spesso nel caso di sussidi in denaro a famiglie in stato di povertà, quando esista una correlazione positiva tra l'ammontare del sussidio e la probabilità che individui *non* in stato di bisogno economico richiedano e ottengano il sussidio. Un sussidio limitato o difficile da ottenere può essere inadeguato a chiudere il *poverty gap* per le famiglie realmente povere, riducendo quindi l'efficacia dell'intervento; allo stesso tempo un sussidio limitato e difficilmente accessibile risulta meno attraente per chi non ne ha estremo bisogno, in quanto il costo psichico e materiale necessario a ottenerlo può superare l'utilità del sussidio stesso. Un *trade-off* esiste quindi tra l'auto-selezione da parte dei richiedenti (che aumenta l'efficacia in quanto migliora il *targeting*) e l'inadeguatezza o l'inaccessibilità del sussidio (che diminuisce la sua efficacia nell'alleviare la condizione di bisogno dei beneficiari.)

Un caso diverso di sovrapposizione è quello tra impatto diretto sui beneficiari e presenza di effetti negativi, quando l'assenza di impatto sui beneficiari è dovuta al fatto che si innescano da parte di questi stessi soggetti comportamenti controproducenti. Un esempio di tale situazione è quello di un sussidio in denaro che non riesce ad aumentare il reddito disponibile delle famiglie beneficiarie perché la sua esistenza induce i membri adulti di queste famiglie a ridurre la loro partecipazione al lavoro (ignorando gli effetti positivi in termini di benessere che la ridotta partecipazione al lavoro può avere, ad esempio per la maggior disponibilità di tempo per la cura dei figli.) Nelle pagine seguenti usiamo la seguente convenzione per distinguere le due situazioni: quando si tratta di comportamenti controproducenti messi in atto dai destinatari stessi dell'intervento, parleremo di ridu-

zione dell'impatto diretto; quando invece i comportamenti controproducenti sono messi in atto da altri (o a danno di altri) soggetti, parleremo di effetti negativi, come nel caso dello spiazzamento e degli effetti di attrazione.

Per concludere, la distinzione tra le quattro determinanti ha soprattutto un valore pedagogico: ci consente di mettere in luce meccanismi diversi che possono aumentare o diminuire l'efficacia complessiva di un intervento. Quello che è importante ai fini della discussione che segue è che le diverse determinanti dell'efficacia necessitano di strumenti tendenzialmente per poter essere valutate adeguatamente. Per questa ragione manterremo tale distinzione nel resto del lavoro, senza però la pretesa di assolutizzarla.

## 2. LA VALUTAZIONE DELL'EFFICACIA DELLE POLITICHE ATTIVE CONTRO LA POVERTÀ

In questa sezione ci proponiamo di illustrare i problemi e le opportunità che emergono nell'applicare lo schema proposto precedentemente alla valutazione delle *politiche attive* contro la povertà e l'esclusione sociale. Nel caso delle politiche attive, la maggior fonte di incertezza sull'efficacia complessiva proviene dal primo stadio, cioè dall'impatto diretto sui beneficiari. Per loro natura, le politiche attive tentano di intervenire su dinamiche complesse, con risultati necessariamente incerti. Quindi la prima domanda che ci si dovrebbe porre nella valutazione di una politica attiva è: "così come è realizzato, l'intervento ha l'effetto desiderato sui suoi diretti beneficiari?"

Negli Stati Uniti, dove maggiore è l'attenzione rivolta alla valutazione di efficacia, l'enfasi è infatti sull'analisi dell'impatto diretto. Gran parte delle valutazioni condotte negli ultimi trent'anni è stata rivolta a capire l'impatto diretto sui beneficiari, prestando relativamente meno attenzione agli altri aspetti, quali l'adeguatezza della scala dell'intervento o la scelta dei beneficiari. Il risultato è che una mole notevole di consapevolezze metodo-logiche e di esperienze pratiche sono state accumulate nell'analisi dell'impatto diretto. Nell'avvicinarsi a questo complesso di esperienze (e nel giudicare la loro applicabilità al contesto italiano), è però importante tenere presente che si tratta di un approccio parziale alla valutazione complessiva. I metodi sviluppati in questo ambito non sono necessariamente estendibili alla valutazione delle altre componenti dell'efficacia complessiva di un intervento.

La minore attenzione prestata negli Stati Uniti per le altre determinanti dell'efficacia complessiva è in parte legata al fatto che la parte più significativa delle valutazioni di impatto diretto delle politiche attive è stata effettuata in un ambito di interventi-pilota (*demonstrations*). In quest'ambito, gli aspetti di *targeting*, di scala, nonché effetti negativi come spiazzamento ed *entry effects*, diventano più difficili o addirittura impossibili da valutare. Il *targeting* dell'intervento è parte del disegno stesso dell'intervento-pilota, quindi è tenuto strettamente sotto controllo. Per contrasto, in un intervento a regime i meccanismi che selezionano la platea dei beneficiari effettivi non sono completamente controllabili, e quindi la congruenza tra beneficiari effettivi e beneficiari desiderati va valutata. Inoltre, l'adeguatezza della scala, per definizione, può solo essere valutata in un intervento a regime, in quanto un intervento-pilota per sua natura è effettuato su scala limitata. La stessa limitatezza della scala previene il verificarsi di spiazzamento ed *entry effects*, quindi la loro importanza e presenza non è valutabile.

L'uso di interventi-pilota quindi rende più "miope" la valutazione, in quanto consente solo di esaminare le determinanti più immediate dell'efficacia di un intervento, ignorando altri aspetti importanti. Questo senza dubbio limita l'utilità dei risultati ottenuti da questo tipo di approccio alla valutazione. Soprattutto, non consente di dare un giudizio definitivo sulla capacità dell'intervento di modificare il fenomeno nel suo complesso, in quanto l'efficacia complessiva di un intervento non dipende solamente dall'impatto diretto sui beneficiari.

Ma se si mettono temporaneamente da parte queste innegabili limitazioni, vanno riconosciuti all'uso di interventi-pilota notevoli vantaggi. Innanzitutto, facilitano l'uso di metodi di tipo sperimentale mediante casualizzazione nella selezione dei beneficiari. (Vantaggi e svantaggi connessi all'uso della casualizzazione come strumento di valutazione verranno discussi nella sezione 3.1.) Va notato come interventi-pilota e uso della casualizzazione non coincidono interamente, anche se il grado di sovrapposizione notevole. Si trovano casi di *demonstrations* dove non è fattibile fare uso di casualizzazione, e viceversa si trovano casi in cui la casualizzazione stessa è stata usata (seppure con molte difficoltà) per valutare interventi a regime.

Al di là del loro uso per l'analisi di impatto, gli interventi-pilota consentono di esaminare i problemi connessi alla messa in opera di un intervento, di anticipare i problemi che potrebbero sorgere quando l'intervento verrà generalizzato al complesso della popolazione. Quindi gli interventi-pilota facilitano un ampio spettro di valutazioni, che copre i risultati e il processo, gli effetti dell'intervento e la sua fattibilità. Come sottolineato alla Sezione 1.2, la valutazione basata su interventi-pilota è un tipico strumento di valutazione intesa come contributo al disegno delle politiche, mentre è scarsamente compatibile con gli obiettivi conoscitivi propri dal controllo di gestione.

Il resto della Sezione 3 è organizzato come segue. La Sezione 3.1 discute gli aspetti metodologici della valutazione di impatto diretto. La Sezione 3.2 presenta due esempi di valutazione di impatto condotta negli Stati Uniti nell'ambito di *demonstrations*. La sezione 3.3 illustra i principali aspetti metodologici relativi alle altre componenti dell'efficacia complessiva delle politiche attive, in particolare la scelta dei beneficiari e l'adeguatezza della scala dell'intervento.

### **3.1 Aspetti metodologici della valutazione di impatto diretto**

L'obiettivo conoscitivo fondamentale della valutazione di impatto è stabilire se il "miglioramento" osservato tra i beneficiari sia da attribuirsi alla loro esposizione all'intervento, e non sia invece da attribuirsi ad altre determinanti che influiscono su di loro indipendentemente dall'intervento. Per "miglioramento" intendiamo il cambiamento nella direzione desiderata

in quei comportamenti o in quelle condizioni dei beneficiari su cui l'intervento intende incidere per raggiungere i propri obiettivi.

Poniamo ad esempio che l'intervento in questione, un corso di recupero dell'obbligo scolastico, abbia come obiettivo quello di aumentare la percentuale di occupati in una popolazione di soggetti socialmente svantaggiati e relegati ai margini del mercato del lavoro. Poniamo che un aumento di occupazione (che può essere misurato con una molteplicità di indicatori) sia effettivamente osservato tra i beneficiari (coloro che hanno frequentato il corso) *dopo* la loro esposizione all'intervento (cioè dopo la frequenza del corso). L'obiettivo dell'analisi di impatto in questo caso è di stabilire se questo cambiamento osservato (l'aumento di occupazione) sia effettivamente da attribuirsi all'intervento, cioè esista tra i due un legame di causa-effetto.

È utile iniziare questa discussione mettendo in evidenza quale tipo di errori un'analisi di impatto impostata correttamente consente di evitare. La prima fallacia è quella di scambiare l'impatto con il semplice "grado di diffusione" dell'intervento tra i potenziali beneficiari. In altre parole, il valore del rapporto:

$$(6) \quad \frac{\text{beneficiari}}{\text{potenziali beneficiari}}$$

non fornisce alcuna informazione rispetto all'impatto diretto dell'intervento. Tale rapporto è invece rilevante per evidenziare un'altra delle determinanti dell'efficacia complessiva dell'intervento, secondo la decomposizione presentata alla Sezione-2: è un modo di misurare l'adeguatezza della scala dell'intervento. Tuttavia, un valore più alto del rapporto (6) non indica necessariamente una maggiore efficacia dell'intervento. Solo nel caso in cui l'analisi di impatto abbia dimostrato che l'intervento ha un impatto positivo tra i beneficiari diretti (in questo caso i frequentanti del corso), allora un alto grado di diffusione suggerisce una maggior efficacia complessiva dell'intervento. Riteniamo quindi fonte di potenziale confusione definire il rapporto (6) come "indicatore di efficacia" dell'intervento: termini quali "gradi di diffusione" o "grado di penetrazione" sono più appropriati a definirne la funzione.

La seconda (e più insidiosa) fallacia che l'analisi di impatto ha il compito di evitare è di attribuire all'intervento tutto il cambiamento favorevole osservato tra i beneficiari—cioè il miglioramento osservato successivamente all'intervento. In altre parole, il rapporto:

$$(7) \quad \frac{\text{beneficiari per cui si è osservato un miglioramento}}{\text{totale dei beneficiari dell'intervento}}$$

non fornisce di per sé una misura dell'impatto diretto dell'intervento. Piuttosto, può essere definito come "tasso lordo di successo". Nell'ambito della letteratura sul controllo di gestione, rapporti simili a quelli rappresen-

tati nella (7) sono spesso definiti come "indicatori di efficacia". Ad esempio, la "percentuale dei frequentanti un corso di recupero dell'obbligo che trova lavoro trascorso un certo tempo dalla fine del corso" verrebbe definito come indicatore di efficacia. Tuttavia, questo uso del termine non è legittimo se la nozione di efficacia che sta alla base della valutazione è quella di "capacità dell'intervento di ottenere un risultato migliore di quello che si sarebbe ottenuto in assenza di intervento". Definire il rapporto (7) come indicatore di efficacia significa presumere che, in assenza di intervento, non si sarebbe potuto verificare un cambiamento favorevole per nessuno dei beneficiari dell'intervento. Il che significa presumere che l'intervento può essere l'unico fattore in grado di influenzare la condizione o il comportamento in questione. Nell'esempio fatto, significa presumere che *nessuno* dei frequentanti avrebbe trovato lavoro senza frequentare il corso.

Quello appena discusso è, nella sua essenza, lo stesso problema metodologico che si deve affrontare per stabilire l'efficacia complessiva dell'intervento, discussa alla Sezione 2.1. Anche nel caso dell'analisi di impatto diretto è necessario ricostruire la "situazione controfattuale", cioè *ricostruire cosa sarebbe successo ai beneficiari se non fossero stati esposti all'intervento*. La differenza fondamentale tra la valutazione dell'efficacia globale a cui si è accennato alla Sezione 2.1 e l'analisi dell'impatto diretto sui beneficiari sta nel fatto che la valutazione di impatto usa come unità di osservazione i beneficiari dell'intervento, non indicatori aggregati che rappresentano l'intervento nel suo complesso. Ci si sposta cioè da un livello macro ad un livello micro.

Dal punto di vista analitico, questo spostamento ha vantaggi e svantaggi. Da un lato, come si è detto, l'analisi di impatto diretto tende ad essere "miope", nel senso che spesso non è in grado di dire molto sulla capacità dell'intervento di incidere sul fenomeno preso nel suo complesso. Lo spostamento alla dimensione micro ha però anche un indubbio vantaggio: la possibilità di ricostruire la "situazione controfattuale" in questo contesto è resa possibile (almeno in linea di principio) dalla presenza di soggetti che, pur appartenendo alla platea dei potenziali beneficiari dell'intervento, non sono stati esposti all'intervento stesso. I non-beneficiari dell'intervento possono cioè fornire il "gruppo di controllo" per stabilire l'impatto dell'intervento.

*Il cuore dell'analisi di impatto diretto è quindi la ricerca di metodi analitici capaci di sfruttare l'informazione messa a disposizione dai non-beneficiari dell'intervento per ricostruire quello che sarebbe successo ai beneficiari se l'intervento non fosse stato attuato.*

Per spiegare in modo intuitivo come ciò possa avvenire, è utile introdurre il seguente ragionamento. Definiamo per brevità con termine di "migliorati" quei soggetti per cui si è osservato un "risultato favorevole", cioè un miglioramento nella condizione o comportamento che l'intervento intende modificare, rispetto ad un periodo precedente all'intervento. Ad esempio, il successo nel trovare lavoro, la decisione di smettere di drogarsi,

il completamento dell'obbligo scolastico, il successo nel trovare un'abitazione. Possiamo quindi riscrivere l'espressione (7) nel modo seguente:

$$(8) \frac{\text{migliorati}}{\text{beneficiari}} = \frac{\text{migliorati grazie all'intervento}}{\text{beneficiari}} + \frac{\text{migliorati in assenza di intervento}}{\text{beneficiari}}$$

L'impatto dell'intervento è rappresentato dal primo termine sul lato destro della (8). Il secondo termine sul lato destro rappresenta invece l'effetto di altri fattori che influenzano comportamenti o condizioni dei beneficiari

indipendentemente dall'intervento. Una percentuale di coloro per cui si è osservato un cambiamento favorevole avrebbero potuto avere lo stesso cambiamento anche in assenza di intervento. Ad esempio, una certa percentuale dei frequentanti il corso di recupero dell'obbligo avrebbe trovato lavoro anche senza la frequenza del corso; una certa percentuale di senza-dimora avrebbe trovato un'abitazione senza l'aiuto del Comune, ecc. Un modo alternativo per comunicare il contenuto della (8) è il seguente:

$$(9) \text{miglioramento osservato} = \text{impatto dell'intervento} + \text{situazione controfattuale}$$

Il fatto che un miglioramento sia osservato tra i beneficiari (cioè che il lato sinistro della (9) sia positivo) non implica di per sé che l'impatto sia positivo. Nel caso estremo in cui tutto il miglioramento sia dovuto all'effetto di altri fattori, l'impatto dell'intervento è zero. Ad esempio, il corso di recupero dell'obbligo risulterebbe essere poco utile (quantomeno se giudicato per la sua capacità di aumentare la probabilità di occupazione dei frequentanti) se tutti costoro avessero egualmente trovato lavoro. Stabilire quanta parte del miglioramento osservato tra i beneficiari sia imputabile direttamente all'intervento è quindi il compito fondamentale dell'analisi di impatto. La semplice misurazione del miglioramento tra i beneficiari di un intervento—il lato sinistro della (9)—non è sufficiente per condurre un'analisi di impatto, per quanto ne rappresenti una precondizione.

I due termini sul lato sinistro della (9) non sono osservabili direttamente, sono entrambi quantità ipotetiche. La strategia analitica usata dall'analisi di impatto, nella sua essenza, è quella di trovare un'approssimazione plausibile per la situazione controfattuale, di modo da ottenere, per differenza, una stima dell'impatto. Questo si può vedere mutando l'ordine dei termini della (9):

$$(10) \text{impatto dell'intervento} = \text{miglioramento osservato} - \text{situazione controfattuale}$$

Nel caso dell'analisi di impatto diretto, la strategia per ottenere una stima della situazione controfattuale è quella di utilizzare comportamenti o condizioni osservati tra i *non-beneficiari* dell'intervento, utilizzando cioè un gruppo di controllo. L'idea è che, se beneficiari e non-beneficiari sono simili nelle condizioni di partenza (cioè in assenza di intervento), ciò che si

osserva tra i non beneficiari approssima ciò che si sarebbe osservato tra i beneficiari se l'intervento non fosse stato attuato. Questa strategia è possibile cioè solo se vale la seguente eguaglianza:

$$(11) \quad \frac{\text{migliorati}}{\text{non-beneficiari}} = \frac{\text{migliorati in assenza di intervento}}{\text{beneficiari}}$$

Se la (11) è valida (e questo rappresenta un grosso "se") si può riscrivere la (8) come segue:

$$(12) \quad \frac{\text{migliorati}}{\text{beneficiari}} = \frac{\text{migliorati grazie all'intervento}}{\text{beneficiari}} + \frac{\text{migliorati}}{\text{non-beneficiari}}$$

A questo punto due dei tre termini contenuti nella (12) sono osservati, e il terzo termine, cioè l'impatto dell'intervento, è "identificato", cioè può essere ottenuto per differenza.

$$(13) \quad \text{impatto dell'intervento} = \frac{\text{migliorati}}{\text{beneficiari}} - \frac{\text{migliorati}}{\text{non-beneficiari}}$$

La strategia della valutazione di impatto è quindi quella di misurare il cambiamento negli indicatori di interesse sia per i beneficiari sia per un gruppo di non-beneficiari, che serve da gruppo di controllo. La possibilità di stimare l'impatto diretto di un intervento dipende da tre condizioni, elencate in ordine (crescente) di difficoltà:

(i) osservazioni sul cambiamento siano disponibili per i beneficiari dell'intervento;

(ii) osservazioni sul cambiamento siano disponibili per i non-beneficiari (gruppo di controllo);

(iii) beneficiari e non-beneficiari siano "simili" in assenza di intervento, cioè valga l'eguaglianza (11).

La condizione (i) è relativamente facile da soddisfare. I beneficiari dell'intervento, per definizione, sono individuabili, in quanto hanno ricevuto qualche tipo di prestazione o servizio. Che il cambiamento nei comportamenti o nelle condizioni rilevanti per la valutazione sia misurabile non è scontato: dipende dal tipo di fenomeno sociale, dagli strumenti di misurazione disponibili, e dalle risorse dedicate alla valutazione. Si tratta comunque di problemi pratici più che concettuali. La condizione (ii) è pure un problema largamente pratico, per quanto di più difficile soluzione. In molte situazioni può essere difficile entrare in contatto con i non-beneficiari di un intervento e raccogliere per loro le stesse informazioni che si hanno per i beneficiari. Ma la condizione di gran lunga più problematica da soddisfare è la (iii), stabilire cioè se beneficiari e non-beneficiari siano equivalenti in assenza di intervento (o meglio, selezionare i non-beneficiari in modo che siano equivalenti.)

Questa è una condizione cruciale per poter interpretare la differenza osservata tra i due gruppi come una stima dell'impatto dell'intervento.

Se i due gruppi sono diversi nelle condizioni di partenza (sarebbero cioè diversi anche in assenza di intervento), il confronto tra i due gruppi produce una stima *distorta* dell'impatto dell'intervento. Tale distorsione può andare in entrambe le direzioni, cioè può portare sia a sottostimare l'impatto dell'intervento sui beneficiari sia a sovrastimarli. Se i non-beneficiari tendono ad essere in una situazione peggiore di come sarebbero stati i beneficiari in assenza di intervento, l'impatto sarà sovrastimato, attribuendogli meriti che non ha. Viceversa, se i non-beneficiari tendono ad essere in una situazione di partenza migliore, all'intervento verranno negati meriti che ha.

Questa (possibile) distorsione è comunemente definita come *selection bias*, o distorsione da selezione. Ciò perché la differenza di partenza tra beneficiari e non-beneficiari che causa la distorsione dipende in modo cruciale da come avviene il *processo di selezione* attraverso cui alcuni individui usufruiscono o sono esposti all'intervento mentre altri non lo sono.

Una vastissima letteratura si è sviluppata sull'argomento del *selection bias*, e di come possa essere eliminato o ridotto in modo da poter stimare correttamente l'impatto dell'intervento sui beneficiari (ad esempio, Heckman and Robb, 1985; Mohr, 1988; Manski e Garfinkel, 1992; Manski, 1995). Parallelamente al crescere di questa letteratura, negli Stati Uniti si è accumulato, a partire dagli anni settanta, un notevole patrimonio di esperienze su come in concreto il problema del *selection bias* possa essere risolto.

Il valutatore dispone di due fondamentali tipi di strategie per eliminare il *selection bias*. Una è quella di eliminare il problema alla radice, manipolando direttamente il processo di selezione che determina chi usufruisce dell'intervento e chi no. Se tale manipolazione è possibile, ed è effettuata assegnando a caso gli individui al gruppo dei beneficiari e non-beneficiari, si ha la garanzia (almeno teorica) che i due gruppi siano equivalenti in termini di condizioni di partenza. La differenza nel miglioramento osservato tra i due gruppi sarà quindi interarpena dovuto all'esposizione all'intervento da parte dei beneficiari. Quella appena descritta è definita come strategia sperimentale mediante casualizzazione. La seconda strategia è invece quella di correggere mediante l'uso di tecniche statistiche le differenze di partenza che emergono "naturalmente", cioè per decisione degli individui. Questa è la strategia non-sperimentale, che presenta numerosissime varianti, a seconda del tipo di informazione che è disponibile sui beneficiari e i non-beneficiari.

### 3.1.1 Le valutazioni sperimentali

La strategia di casualizzare l'esposizione all'intervento è disponibile in relativamente poche situazioni. Per essere più precisi, tale approccio è raramente applicabile agli interventi a regime, in quanto il tentativo di escludere alcuni individui dalla fruizione di un intervento esistente si scontra con-

tro una lunga serie di ostacoli di natura etica, legale nonché organizzativa. Questo limita in gran parte l'applicazione della strategia sperimentale alla valutazione di interventi-pilota, che sono creati con l'intento esplicito di valutarne effetti e fattibilità. Tuttavia, anche in questo ambito ristretto, l'approccio sperimentale può incontrare difficoltà, in quanto le condizioni ideali in cui un esperimento dovrebbe essere condotto per fornire risultati utili non sono sempre soddisfatte (Hausman and Wise, 1985; Manski e Garfinkel, 1992). La correttezza delle stime dell'impatto può essere minacciata da complicazioni che possono sorgere durante la realizzazione dell'esperimento, quali la possibilità che il gruppo di controllo sia contaminato da prestazioni simili offerte da altri enti pubblici. Inoltre, l'integrità del processo di casualizzazione è spesso difficile da mantenere sul campo, e può essere compromessa dal comportamento degli operatori del servizio, che raramente condividono con i valutatori la stessa fede nella casualizzazione. Ad esempio, gli operatori possono tentare di alterare il risultato dell'assegnazione casuale sulla base delle loro preferenze e opinioni su chi debba usufruire dell'intervento e chi no. In questo modo si ricrea una forma di *selection bias*. In queste situazioni, quando l'equivalenza tra beneficiari e gruppo di controllo viene compromessa, si parla di mancanza di *validità interna* dell'esperimento.

Quando ad essere compromessa è invece la rappresentatività e generalizzabilità dei risultati della valutazione sperimentale, si parla di mancanza di *validità esterna*. In questo caso le stime sono corrette, ma possono egualmente risultare di utilità limitata per il successivo disegno delle politiche. Per una serie di ragioni, i risultati dell'esperimento possono essere poco rappresentativi della situazione in cui l'intervento opererebbe una volta a regime: possono contribuire fattori quali, per citarne solo alcuni, la scala ridotta degli interventi-pilota, la loro durata necessariamente limitata nel tempo, gli effetti psicologici indotti sul comportamento degli operatori e dei beneficiari dal fatto di sentirsi sotto osservazione; la dipendenza dei risultati ottenuti dal particolare contesto socio-economico in cui l'esperimento è stato realizzato. Quindi lo strumento della valutazione sperimentale va applicato solo dopo un'attenta verifica delle sue probabilità di successo. Si tratta di uno strumento assai potente, ma anche estremamente delicato nel suo uso.

Quando può venire applicata con successo, la strategia sperimentale ha senza dubbio notevoli vantaggi. Oltre a quello principale di garantire l'equivalenza dei due gruppi e quindi la correttezza delle stime di impatto, l'approccio sperimentale ha il vantaggio di produrre automaticamente un campione di non-beneficiari, cioè coloro che sono stati esclusi dall'intervento. Questo risolve un problema pratico che affligge molte valutazioni non-sperimentali: come selezionare in concreto un campione di non-beneficiari. Spesso il costo di tale selezione è elevato, e ciò può costituire un ostacolo non indifferente alla realizzazione di valutazioni non-sperimentali, a prescindere dalla problematica di tipo statistico relativa all'eliminazione del *selection bias*

### 3.1.2 Le valutazioni non-sperimentali

Il principale vantaggio della strategia non-sperimentale è quella di poter essere applicata ad una gamma di situazioni molto più ampia di quella sperimentale. Anche interventi-pilota possono venire valutati in questo modo, se la casualizzazione per qualche ragione non è applicabile. Ad esempio, quando l'intervento-pilota per sua natura deve essere applicato alla totalità di una popolazione per poterne esaminare gli effetti (ad esempio, se si vogliono individuare effetti sul complesso del mercato del lavoro) la casualizzazione a livello individuale non può essere utilizzata. Ma il vantaggio maggiore è ovviamente l'applicabilità agli interventi a regime, il che rende l'approccio non-sperimentale un ingrediente importante di una complessiva strategia di valutazione di impatto diretto.

L'idea di fondo della valutazione non-sperimentale è sempre quella di ricostruire la situazione controfattuale utilizzando l'esperienza di un gruppo di individui che non è stato coinvolto nell'intervento. La differenza è che in questo caso il processo di selezione non è manipolabile dal valutatore, bensì è spontaneo, cioè frutto delle decisioni degli individui e degli operatori. L'impossibilità di manipolare il processo di selezione pone innanzitutto il problema di come trovare il gruppo di controllo, e quello di verificare se l'esperienza di tale gruppo approssimi correttamente la situazione controfattuale. I due principali approcci per risolvere il secondo problema sono: (i) la modellazione del processo di selezione; e (ii) l'uso di informazioni longitudinali.

Quanta più conoscenza il valutatore ha del processo di selezione, tanto più sarà in grado di modellare esplicitamente tale processo, tentando di eliminare la distorsione che la selezione causa nel confronto tra i due gruppi. Un caso favorevole è quello in cui la selezione dei beneficiari sia effettuata dagli operatori del servizio interamente sulla base delle caratteristiche osservabili dei richiedenti, ad esempio mediante la compilazione di graduatorie. Ciò spesso accade quando esiste razionamento, cioè quando c'è un limite alla capacità di soddisfare la domanda. Questa è una situazione ideale dal punto di vista della valutazione, e per alcuni versi simile a quella della casualizzazione, in quanto il processo di selezione può essere modellato dal valutatore con un ricorso minimo ad assunti non verificabili. L'uso dell'analisi di regressione consente di eliminare le differenze osservabili tra i due gruppi, e di pervenire a stime corrette dell'impatto dell'intervento.

Un caso particolare è quello in cui siano state utilizzate graduatorie, e i beneficiari siano stati scelti come coloro che si situano al di sopra (o al di sotto) di una certa soglia o punteggio nella graduatoria. In questo caso una semplice stima dell'impatto può essere ottenuta confrontando la situazione post-intervento di coloro immediatamente al di sotto e immediatamente al di sopra della soglia (quindi molto simili tra loro, in quanto inclusi o esclusi dall'intervento in un modo *quasi* casuale: inoltre, quanto più la graduatoria

è imperfetta, tanto gli ultimi inclusi e i primi esclusi saranno simili). Se la dimensione campionaria di tali casi marginali non è sufficiente, tecniche di regressione possono essere usate. Si parla in questo caso di *regression-discontinuity design*, in quanto la discontinuità creata dalla graduatoria viene usata per identificare l'impatto dell'intervento.

Una situazione molto più complessa emerge quando la selezione sia frutto delle decisioni dei potenziali beneficiari del servizio, e quindi dipenda in larga misura da caratteristiche individuali (abilità, motivazione, preferenze, storia passata, aspettative) che il valutatore non è in grado di osservare ma che possono essere correlate con la condizione o il comportamento che l'intervento vuole modificare. Si parla qui di *self-selection*, o auto selezione. In questo caso è necessario il ricorso a tecniche statistiche, che necessariamente si basano su assunti non verificabili che il valutatore deve fare sul processo di selezione o sulla distribuzione nella popolazione di tali caratteristiche non osservabili. Quando c'è non-osservabilità, il ricorso ad assunti non verificabili è inevitabile. Quanta più informazione è disponibile, tanto più deboli saranno gli assunti mantenuti. Un'ampia letteratura si è sviluppata su questo argomento, raggiungendo livelli di sofisticazione molto elevati. Non ci addenteremo qui in un'esposizione di questi metodi. Per una breve rassegna, si rimanda a Martini (1992).

Il secondo approccio per la valutazione non sperimentale è quello di utilizzare dati longitudinali, cioè osservazioni ripetute sui medesimi individui. La disponibilità di osservazioni su beneficiari e non beneficiari per periodi *precedenti e seguenti* all'intervento consente l'applicazione di particolari metodi per eliminare (o verificare la presenza di) *selection bias*. Anche su questo punto l'elaborazione di tecniche statistiche è molto avanzata. Qui ci limiteremo a fornire un minimo di intuizione. Se il processo di auto-selezione genera diverse condizioni di partenza, queste si traducono in differenze negli indicatori tra beneficiari e non-beneficiari nel periodo *precedente* all'intervento. Se questa pre-differenza può essere osservata, essa può essere sottratta dalla post-differenza, eliminando almeno in parte la distorsione che il *selection bias* genera nelle post-differenze. Questo metodo è infatti definito come "differenza-nelle-differenze."

La disponibilità di dati longitudinali consente anche di verificare la validità dei metodi per l'eliminazione del *selection bias* basati sulla modellazione del processo di selezione. L'idea è quella di applicare questi metodi ai dati pre-intervento. Prima dell'intervento beneficiari e non-beneficiari differiscono non a causa dell'intervento (che non è ancora avvenuto) ma a causa di tutto ciò che li indurrà a fare una scelta diversa (motivazione, storia passata, abilità, ecc.). Se il metodo statistico proposto davvero elimina il *selection bias*, dovrebbe eliminare ogni differenza tra i due gruppi nell'indicatore osservato prima dell'intervento (Heckman e Hotz, 1989).

Due limitazioni all'applicabilità dei metodi basati sui dati longitudinali vanno notate. La prima è alquanto banale: i dati longitudinali sono costosi e difficili da raccogliere, e la decisione di raccogliergli richiede una dose

notevole di lungimiranza. La seconda limitazione è che la condizione o il comportamento che l'intervento vuole modificare può non essere "ripetibile" nel tempo, cioè può non produrre osservazioni ripetute e confrontabili. Ad esempio, mentre il reddito di una famiglia consente osservazioni ripetute nel tempo, il peso di un bambino alla nascita (un importante indicatore per giudicare il successo di interventi a sostegno delle gestanti) non è un'osservazione ripetibile nel tempo: per ogni gestante (esposta o meno all'intervento) si ha una sola osservazione. L'impatto di un intervento a favore delle gestanti può essere valutato solo confrontando una singola osservazione per ciascun beneficiario e una per ciascun non-beneficiario.

#### *Conclusioni sull'analisi di impatto diretto*

Si delinea quindi il seguente quadro di massima per la realizzazione di valutazioni di impatto diretto. Quando sia possibile realizzare interventi-pilota, la casualizzazione diventa la strategia preferibile, perché agli innegabili vantaggi statistici si oppongono limitati svantaggi sul lato etico ed organizzativo. Più che dalla casualizzazione stessa, il limite di questa strategia deriva dai limiti connaturati alla natura pilota dell'intervento. Quando invece la valutazione di impatto sia da condursi su interventi a regime, la casualizzazione può essere applicata molto raramente. La strategia è quella di estrarre gruppi di controllo dalla popolazione dei non-beneficiari, e di utilizzare metodi statistici per (tentare di) eliminare le differenze indotte dal processo di selezione.

### **3.2 Casi di valutazione di impatto condotta negli Stati Uniti**

In questa sezione esaminiamo le valutazioni di due interventi di politica attiva contro la povertà e l'esclusione sociale condotti negli Stati Uniti negli anni settanta e ottanta. Si tratta in entrambi i casi di interventi-pilota, creati con lo scopo specifico di valutare l'efficacia e altri aspetti dell'intervento, e in entrambi i casi la valutazione di impatto è stata condotta mediante un approccio sperimentale. Le due valutazioni si differenziano invece sotto altri profili, per la diversa popolazione a cui l'intervento è rivolto, e per il tipo di politica attiva la cui efficacia si intende valutare. Il primo di questi interventi è mirato a gruppi socialmente emarginati in età relativamente giovane, e la valutazione intende verificare la capacità dell'intervento di migliorare la loro posizione economica e ridurre l'incidenza di comportamenti socialmente distruttivi. Il secondo è invece rivolto alla popolazione di anziani soli ad alto rischio di istituzionalizzazione, e l'obiettivo è quello di ridurre tale rischio mediante l'offerta di una più intensa assistenza domiciliare. Va notato che queste due valutazioni sono state scelte tra alcune centinaia di valutazioni di interventi-pilota condotte negli Stati Uniti a partire dalla fine degli anni sessanta..

### *La National Supported Work Demonstration (SWD)*

Lo scopo di questo intervento-pilota, realizzato nella seconda metà degli anni settanta, è quello di valutare l'efficacia di un particolare approccio all'inserimento lavorativo di soggetti che soffrono di forme estreme di emarginazione sociale. Quattro gruppi distinti vengono scelti come *target* per questo intervento: madri non sposate che da più di tre anni ricevono il sussidio di povertà; giovani *drop-outs* della scuola dell'obbligo, molti con precedenti penali; ex-tossicodipendenti; ex-carcerati. (Il disegno e i risultati di questo esperimento sono illustrati in *Manpower Demonstration Research Corporation* (1980), su cui è basata la descrizione seguente).

Il *supported work* consiste nel fornire a tali soggetti un'esperienza lavorativa "vera" ma in un ambiente protetto e altamente strutturato, con un aumento graduale dell'intensità dello sforzo di lavoro, sotto stretta supervisione e con un inserimento in squadre di lavoro composte di individui con simili problemi, e con una retribuzione uguale o di poco superiore al salario minimo. Il periodo di *supported work* è destinato a durare circa un anno, dopodiché l'obiettivo è quello di un inserimento nel mercato del lavoro regolare. L'idea del *supported work* non è nuova, anzi è mutuata da varie esperienze di "lavoro protetto" condotte in Paesi europei quali il Regno Unito, la Svezia e i Paesi Bassi, dove sono state usate soprattutto per l'inserimento lavorativo di portatori di *handicap*. Tuttavia, si ritiene necessaria una valutazione in quanto esistono tra i *policy-makers* dubbi su quale possa essere l'efficacia di tale tipo di intervento quando sia applicato alle forme di esclusione sociale diffusa nei grandi centri urbani americani.

Così come realizzato nell'ambito della SWD, si tratta infatti di un intervento che richiede un investimento notevole di risorse per persona servita, tra i 5.000 e gli 8.000 dollari per partecipante, che alla metà degli anni '70 rappresenta uno dei tipi più costosi di politica attiva di reinserimento sociale. Invece che lanciare un simile intervento su larga scala, il governo federale americano decide di valutarne l'efficacia su scala ridotta, per verificare se l'alto costo è giustificabile quando sia rapportato ai benefici che produce (misurati in termini di aumento di partecipazione al lavoro, riduzione della dipendenza dai sussidi pubblici, e diminuzione di comportamenti anti-sociali, quali l'uso di droga e l'attività criminale.)

Per il disegno e la realizzazione della SWD viene costituita una nuova organizzazione *non-profit*, la *Manpower Demonstration Research Corporation* (MDRC), in quanto nella metà degli anni settanta manca ancora negli Stati Uniti un'organizzazione con le caratteristiche adeguate per svolgere una valutazione di questo tipo. (MDRC resta tutt'oggi un'organizzazione *leader* nel settore della valutazioni di interventi-pilota nel campo delle politiche sociali.) Il finanziamento per la SWD viene fornito da un *pool* di Ministeri (del lavoro, della giustizia, dell'istruzione, della sanità e servizi sociali, e dello sviluppo urbano) e da alcune organizzazioni private, tra cui soprattutto la *Ford Foundation*, che rimane oggi uno dei principali finanziatori pri-

vati di valutazioni di questo tipo. Il costo totale della *demonstration* è di quasi 82 milioni di dollari, di cui 11 milioni per la valutazione, e i restanti 70 milioni per la realizzazione dell'intervento vero e proprio (locali, equipaggiamento, retribuzione dei partecipanti e dei supervisori). Circa 50 milioni vengono forniti dai finanziatori nazionali, mentre la differenza viene coperta mediante fondi messi a disposizione dai governi locali, con la vendita dei beni e servizi prodotti dai partecipanti e con lo storno di parte dei sussidi di povertà che molti dei partecipanti avrebbero ottenuto se non fossero stati coinvolti nell'esperimento.

La *SWD* si svolge in 14 centri urbani, localizzati in altrettanti Stati. In ciascun sito viene costituita un'organizzazione *non profit* con il compito di gestire l'intervento. La *MDRC* mantiene il coordinamento delle operazioni a livello nazionale e la gestione della valutazione sperimentale. Le attività lavorative in cui i partecipanti sono coinvolti sono per la metà nei servizi, dalla manutenzione di edifici all'assistenza negli asili-nido, un quarto nel settore edilizio, e quasi un 10 per cento nel settore manifatturiero. L'intervento dura quasi quattro anni, dal 1975 al 1978, e coinvolge circa 10.000 individui, che partecipano al *supported work* per una media di circa sette mesi ciascuno.

L'aspetto che più ci interessa di questa *demonstration* è il fatto di essere accompagnata, fin dalla fase iniziale di progettazione, da una valutazione di impatto mediante casualizzazione. Lo scopo dell'intero intervento è infatti quello di fornire una risposta ai seguenti interrogativi:

- qual'è l'efficacia del *supported work* nell'aumentare la partecipazione al lavoro e nel ridurre la dipendenza dai sussidi pubblici?
- quale tipo di persone svantaggiate trae maggior beneficio da questo intervento?
- quanto costa operare un simile tipo di intervento?
- qual è il rapporto tra costi sostenuti e benefici ottenuti?
- quali strategie di implementazione funzionano meglio?
- quali caratteristiche dell'intervento hanno un impatto maggiore sul comportamento e sulla *performance* dei partecipanti?

Come si vede, si tratta di un ampio spettro di interrogativi, che va al di là del semplice impatto diretto dell'intervento, e copre anche questioni legate all'implementazione, cioè a come l'intervento viene concretamente realizzato. Ma in tutti i casi si tratta di interrogativi che scaturiscono da una genuina incertezza, e che sono destinati ad informare decisioni successive, qualora si tratti di realizzare l'intervento su larga scala.

L'analisi di impatto riguarda soprattutto il primo degli interrogativi elencati sopra. Per rispondere a questo interrogativo, in 10 dei 14 siti in cui la *demonstration* è condotta, una persona su due che fanno domanda per partecipare a questa esperienza viene esclusa dalla partecipazione mediante sorteggio. Coloro che sono esclusi dal *supported work* rappresentano il

gruppo di controllo. Da un punto di vista etico e legale questa esclusione non è un problema insormontabile, in quanto la partecipazione al *supported work* non rappresenta alcun tipo di diritto per coloro che fanno domanda: senza l'intenzione esplicita di condurre l'esperimento, l'intervento stesso non sarebbe stato realizzato (se non altro perché sarebbero mancati i finanziamenti stati messi a disposizione dal governo federale e delle fondazioni.) Resistenze alla casualizzazione ovviamente non mancano durante la messa

- in opera dell'esperimento, soprattutto da parte degli operatori locali, i quali non traggono alcun vantaggio dall'essere coinvolti in un esperimento, mentre hanno l'onere di reclutare un numero doppio di partecipanti e poi annunciare alla metà di questi che sono stati esclusi mediante sorteggio.

Dal punto di vista della valutazione di impatto, i vantaggi della casualizzazione sono già stati illustrati in precedenza: il fatto di essere assegnati casualmente fa sì che i membri del gruppo di controllo siano del tutto simili, in termini di caratteristiche osservabili come di quelle non osservabili, ai membri del gruppo che partecipa all'intervento. Le differenze osservate successivamente negli indicatori sono da attribuirsi interamente all'esposizione all'intervento.

Circa 6.600 individui vengono coinvolti nel processo di casualizzazione, e di essi circa 3.200 assegnati al gruppo di trattamento e 3.400 al gruppo di controllo. Ai primi viene consentito di partecipare al *supported work*, mentre i secondi vengono esclusi. Solo un terzo dei 10.000 partecipanti al *supported work* viene quindi utilizzato direttamente per la valutazione di impatto. Questo perché i fondi di ricerca disponibili consentono solamente di sottoporre ad osservazione circa 6.600 individui. Per tutti costoro, indipendentemente dal fatto che siano stati assegnati al trattamento o al gruppo di controllo, vengono condotte cinque ondate successive di interviste, cominciando con un'intervista immediatamente prima dell'assegnazione casuale, seguita da un'intervista periodica ogni nove mesi: l'ultima si svolge a circa tre anni di distanza dall'inizio dell'esperimento. Queste interviste vengono utilizzate per rilevare come variano nel tempo le condizioni e i comportamenti che l'intervento vuole modificare, cioè la partecipazione al lavoro, lo stato di povertà, la percezione di sussidi, l'uso di stupefacenti, i problemi con la giustizia penale. La dimensione del campione è comunque sufficiente per fornire stime attendibili dell'impatto dell'intervento.

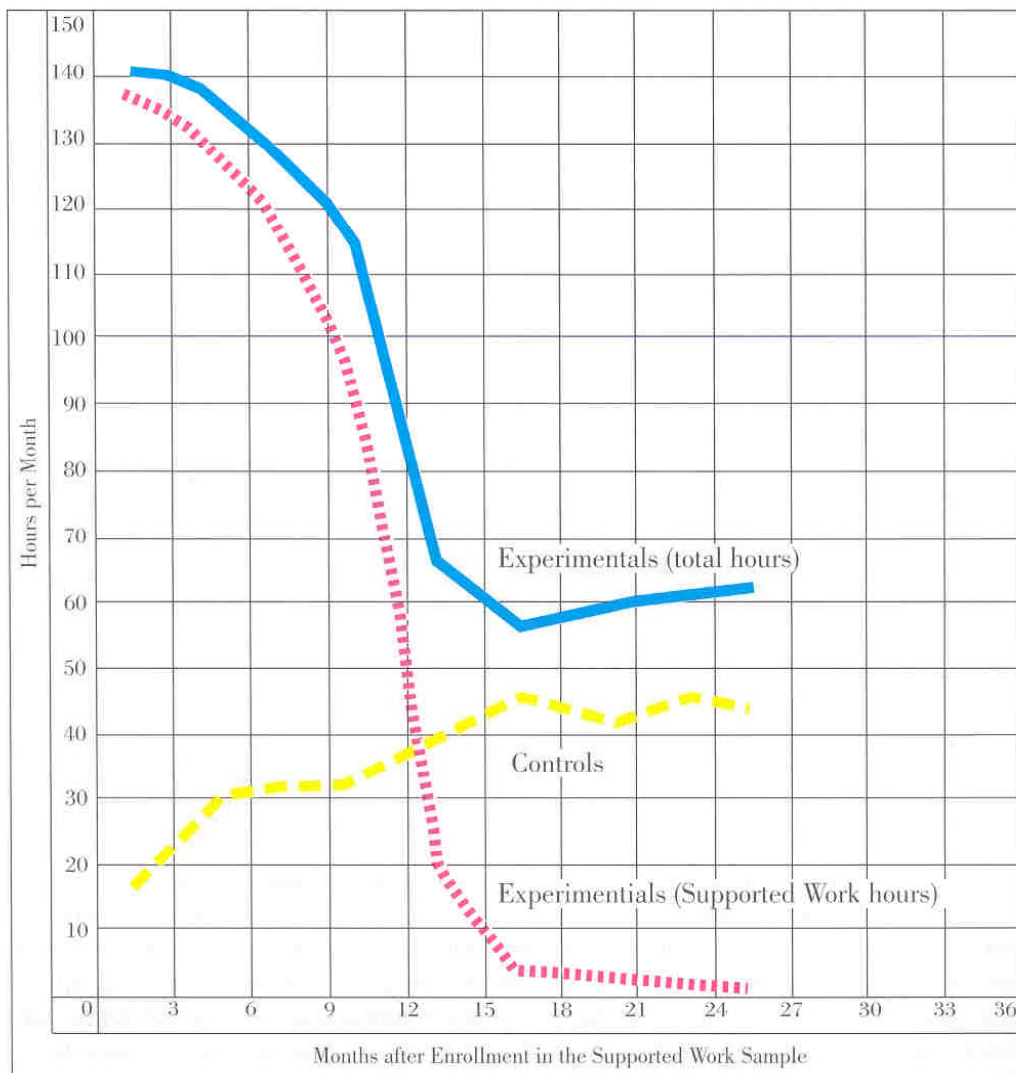
È interessante esaminare l'impatto sulla partecipazione al lavoro per due delle popolazioni coinvolte nell'esperimento, le madri non sposate che vivono del sussidio di povertà e i giovani *drop-outs* dell'obbligo. Questi risultati sono interessanti non solo per le stime di impatto che presentano, ma anche perché mostrano chiaramente che senza un gruppo di controllo è facile attribuire all'intervento miglioramenti che sono invece dovuti a processi di cambiamento che non hanno nulla a che vedere con l'intervento.

La Figura 7 presenta questo tipo di risultati per il gruppo delle madri non sposate, definito come "*AFDC sample*", in quanto il sussidio di povertà che queste donne ricevono è l' *Aid to Families with Dependent Children*

Fig.7

**ANALISI DI IMPATTO DELLA SUPPORTED WORK DEMONSTRATION  
(Ore lavorate per mese, campione delle madri non-sposate)**

**TREND IN HORS WORKED PER MONTH: AFDC SAMPLE**



Note: All experimental-control differentials are significant at the 5 percent level.

Fonte: Manpower Demonstration Research Corporation (1980)

(AFDC). L'indicatore rispetto a cui si misura l'impatto è rappresentata dalle ore lavorate mensili, riportate sull'asse verticale. Sull'asse orizzontale troviamo invece il numero di mesi trascorsi dal reclutamento nel programma (a cui è immediatamente seguita l'assegnazione ai due gruppi).

La linea solida rappresenta le ore lavorate in media al mese da coloro che partecipano al *supported work* (definiti nella figura come "*experimentals*"). La linea a puntini rappresenta le ore che questo stesso gruppo lavora all'interno del *supported work*. All'inizio del periodo le due linee quasi coincidono al livello di 140 ore mensili, che significa che la maggior parte dei partecipanti lavora a tempo pieno e svolge solo una minima attività lavorativa al di fuori della struttura protetta. Dopo un anno le ore di *supported work* tendono a zero; le ore lavorate totali si stabilizzano sulle 60 ore mensili (cioè circa 15 ore settimanali) dopo due anni dall'inizio dell'esperimento. Rispetto alle 140 ore iniziali si tratta di una forte caduta. Però va ricordato che le 60 ore sono interamente ore lavorate al di fuori della struttura protetta, il che rappresenta l'obiettivo di fondo dell'esperimento: facilitare l'inserimento di queste persone in una situazione lavorativa "normale".

A questo punto va notato che, se l'informazione sulle ore lavorate fosse disponibile solo per il gruppo dei partecipanti, si potrebbero ricavare indicazioni poco attendibili riguardo all'efficacia dell'intervento, e tali indicazioni sarebbero scarsamente utilizzabili per decidere se i benefici dell'intervento giustificano i costi. La prima fallacia sarebbe quella di concludere che l'intervento è stato un fallimento, a causa della caduta di partecipazione a conclusione del periodo di lavoro protetto. Questo però significherebbe ignorare le condizioni di partenza di questo gruppo, che è stato selezionato proprio perché inizialmente ai margini del mercato del lavoro. Quindi anche 60 ore mensili di lavoro non protetto possono rappresentare un successo.

Un passo avanti da un punto di vista analitico è il confronto con la condizione pre-intervento. Questo confronto non rivela ancora l'impatto, in quanto ignora cosa sarebbe successo ai partecipanti in assenza di intervento. Conclusioni più attendibili in questo senso sono invece possibili grazie alla presenza del gruppo di controllo. L'esperienza del gruppo di controllo rappresenta la situazione controfattuale, cioè quello che presumibilmente sarebbe successo ai beneficiari dell'intervento se questo non fosse stato effettuato. Tale esperienza è rappresentata dalla linea tratteggiata nella Figura 7. In media i membri del gruppo di controllo lavorano all'inizio del periodo circa 20 ore al mese, ma dopo due anni la loro partecipazione al lavoro è più che raddoppiata, superando le 40 ore mensili. Quindi, se misurato come differenza tra risultato osservato e situazione controfattuale, l'impatto netto della partecipazione al *supported work* ad un anno circa dalla sua conclusione è in media di circa 20 ore mensili, cioè un aumento netto del 50 per cento.

Se invece si confrontassero le 60 ore mensili degli *experimentals* con le 20 ore mensili che costoro presumibilmente avrebbero lavorato all'inizio del periodo (essendo in tutto simili al gruppo di controllo), si concluderebbe che

L'impatto dell'intervento è un aumento da 20 a 60 ore, cioè un aumento del 200 per cento, sovrastimandolo notevolmente. La disponibilità di un gruppo di controllo evita di sovrastimare l'efficacia dell'intervento (che comunque in questo caso resta positiva.) Ad un anno dalla conclusione del periodo di lavoro protetto, i partecipanti lavorano circa il 50 per cento in più di quanto avrebbero lavorato se non avessero partecipato a questa esperienza.

La Figura 8 illustra i risultati ottenuti per un altro tipo di gruppo coinvolto nell'esperimento, i giovani tra i 17 e i 20 anni di età senza il diploma dell'obbligo e in gran parte con precedenti penali.

Le ore lavorate dai *controls* all'inizio del periodo di osservazione sono di poco superiori alle 30 ore mensili, mentre tre anni dopo sono salite a quasi 80, un aumento "spontaneo" anche in questo caso superiore al 100 per cento. Quello che è interessante notare è che l'esperienza dei partecipanti al *supported work*, una volta che questo si è concluso, segue quasi esattamente lo stesso percorso di coloro che non vi hanno partecipato. In assenza di gruppo di controllo, si sarebbe portati a concludere che, grazie all'esperienza di lavoro protetto, la partecipazione al lavoro di questi giovani è due volte superiore al livello iniziale. La presenza del gruppo di controllo invece consente di evitare tale conclusione sbagliata: se giudicato sulla base dell'aumento netto nella partecipazione al lavoro, l'efficacia del *supported work* è essenzialmente nulla.

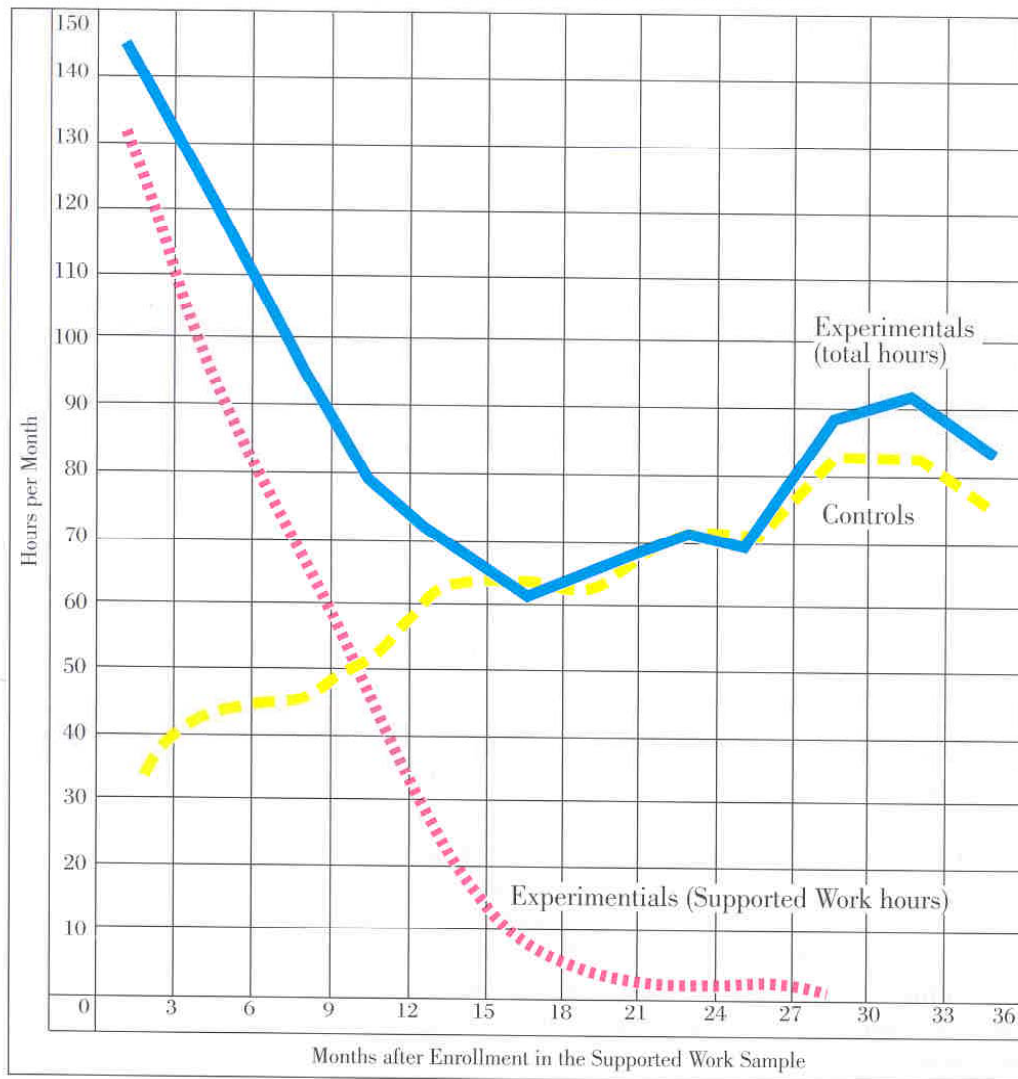
L'impatto di un intervento di questo tipo può anche essere giudicato sulla base di altri criteri, utilizzando cioè altri indicatori. Alcune di essi sono misurabili, come il coinvolgimento in attività criminali o l'uso di stupefacenti. Vi sono poi dimensioni non facilmente quantificabili, quali l'auto-stima e la capacità di interagire socialmente, che è possibile siano positivamente influenzate dalla partecipazione ad un'esperienza di lavoro pro-tetto.

Partendo dai risultati dell'analisi di impatto, i ricercatori di *MDRC* hanno condotto un'analisi costi-benefici della partecipazione al *supported work*, separatamente per le quattro tipologie di soggetti coinvolti. Va sottolineato che tale analisi costi-benefici può essere condotta solo dopo che è stata effettuata l'analisi di impatto, in quanto il principale elemento che entra nel calcolo sul lato dei benefici è proprio l'impatto netto dell'intervento, in termini di aumento di reddito, diminuzione di dipendenza dai sussidi pubblici, diminuzione di attività criminale. Tutte queste dimensioni vanno monetizzate, in modo da poter essere aggregate in un'unica misura, attualizzate al presente mediante un opportuno tasso di sconto, e confrontate ai costi dell'intervento. Queste operazioni richiedono una molteplicità di assunti non verificabili, che rendono i risultati dell'analisi costi-benefici meno precisi rispetto a quelli dell'analisi di impatto.

Per i giovani *drop-outs*, i ricercatori concludono che, anche sotto diversi scenari, i benefici prodotti dal *supported work* non sono in grado di giustificarne il costo. Per questo gruppo, il valore attualizzato dei benefici netti è negativo, variando tra - 4.000 e - 250 dollari a seconda degli assunti fatti, e

**Fig. 8**  
**ANALISI DI IMPATTO DELLA SUPPORTED WORK DEMONSTRATION**  
**(Ore lavorate per mese, campione dei giovani drop-outs)**

**TREND IN HORS WORKED PER MONTH: YOUTH SAMPLE**



Note: Experimental-control differentials are significant at the 5 percent level for months 1 to 12.

Fonte: Manpower Demonstration Research Corporation (1980)

in nessun caso diventa positivo. Questo risultato è determinato in larga parte da quanto illustrato nella Figura 8, cioè dal fatto che la partecipazione al *supported work* non *causa* alcun aumento della partecipazione al lavoro di questi giovani.

Un risultato diverso si ottiene dall'analisi costi-benefici applicata al gruppo delle madri non sposate. In questo caso il valore attualizzato dei benefici netti varia da 2.700 a 9.000 dollari. La realizzazione del *supported work* per questo particolare gruppo si rivela quindi un buon investimento dal punto di vista della società. Questa conclusione ovviamente è difficile da generalizzare dall'ambito ristretto di un intervento-pilota a quello di un intervento realizzato su vasta scala. È possibile che l'estensione di un intervento del genere ad una scala più ampia incontri difficoltà organizzative che ne riducano l'efficacia; è possibile che la sua capacità di penetrazione nella platea dei potenziali beneficiari si riveli molto limitata; è possibile infine che il successo ottenuto nell'aumentare la partecipazione al lavoro di un piccolo gruppo di individui non sia replicabile su larga scala, a causa della limitata capacità da parte del mercato del lavoro di assorbire questo tipo di lavoratori svantaggiati. Restano comunque lezioni importanti da questo esperimento: ad esempio, il fatto che, per i giovani emarginati, il coinvolgimento in un'esperienza di questo tipo non pare produrre benefici sufficienti a giustificarne i costi.

I risultati positivi della SWD spingono molti alla fine degli anni settanta a raccomandare l'applicazione di questo tipo di intervento su larga scala, come tentativo di soluzione al cronico problema dell'emarginazione dal mercato del lavoro da parte delle madri che dipendono per anni dal sussidio di povertà. Queste raccomandazioni si scontrano però con un brusco cambiamento di clima, seguito all'elezione di Reagan nel 1980. Uno dei cavalli di battaglia della piattaforma di Reagan è l'abolizione di ogni forma di *subsidized employment* (cioè di lavoro il cui costo è almeno in parte coperto da un contributo pubblico), con l'obiettivo di porre fine a forme di abuso e di spreco che sono emerse durante l'amministrazione Carter, ricevendo vasta eco sulla stampa e durante la campagna elettorale. Il *supported work* non è parte di tali forme di abuso, ma diventa vittima innocente di tale pubblicità negativa e viene sostanzialmente abbandonato.

#### *La National Long Term Care Demonstration*

Questo intervento-pilota viene realizzato nella prima metà degli anni ottanta, sponsorizzato dal Ministero della sanità e dei servizi sociali. La motivazione principale per questo intervento si ricollega all'opinione, largamente diffusa negli Stati Uniti negli anni settanta, che il sistema di assistenza agli anziani non-autosufficienti sia troppo sbilanciato verso il ricorso alla casa di riposo, ed invece presti troppo poca attenzione alla possibilità di mantenere l'anziano nella comunità mediante la fornitura di servizi di assistenza domiciliare. La *National Long Term Care Demonstration*, definita

nella letteratura sul tema come *channeling demonstration*, intende valutare gli effetti di un aumento della fornitura di servizi domiciliari e sul territorio su una varietà di possibili indicatori, dal ricorso alla casa di cura allo stato di salute e qualità della vita dell'anziano. (Il disegno e i risultati della *channeling demonstration* sono descritti in Kemper e altri, 1988.)

Due tipi diversi di modelli di intervento vengono realizzati nell'ambito della *channeling demonstration*. Il primo, definito come *basic case management model*, intende testare l'ipotesi che la maggiore difficoltà da parte dell'anziano non-autosufficiente nell'ottenere servizi adeguati sul territorio sia la mancanza di informazione e la difficoltà di accesso ai servizi già esistenti. Il *case management* vuole essere una risposta a questo mancato incontro tra bisogno e offerta di servizi. Il *case manager*, essenzialmente un assistente sociale specializzato, dedica la propria attenzione ad un numero limitato di anziani, cercando di utilizzare nel modo più efficiente possibile i servizi esistenti. Quindi l'unico uso addizionale di risorse da parte di questo intervento è rappresentato dal costo stesso dei *case managers*.

Il secondo modello di intervento della *channeling demonstration*, definito come *financial control model*, è destinato a testare l'ipotesi che il ricorso eccessivo al ricovero in casa di cura sia dovuto all'insufficienza dei fondi disponibili per finanziare servizi di assistenza sul territorio. Quindi in questo modello l'assistente sociale ha una disponibilità di risorse addizionali con cui acquistare i servizi che ritiene necessari. Parte di queste risorse deriva da un uso più "libero" che il *case manager* può fare dei programmi federali di assistenza sanitaria per i poveri (*Medicaid*) e per gli anziani (*Medicare*). Tuttavia, questo modello pone anche dei limiti complessivi sull'uso di risorse addizionali, per evitare di valutare l'efficacia di un intervento senza limiti di bilancio, il che sarebbe realizzabile in un ambito limitato ma poi non replicabile su ampia scala. Rispetto alla situazione esistente, il *case manager* ha più libertà di trasferire la disponibilità di risorse all'interno del *pool* di utenti serviti.

La valutazione della *channeling demonstration* quindi consiste nello stimare gli effetti dell'adozione di questi modelli di intervento rispetto alla *status quo*, in cui già esiste una gamma di servizi disponibili sul territorio per gli anziani non-autosufficienti. Quindi la situazione controfattuale è rappresentata non dalla totale assenza di servizi, bensì dai servizi normalmente esistenti. Questo rende legalmente ed eticamente possibile l'uso della casualizzazione: gli anziani esclusi casualmente dall'accesso ai servizi addizionali conservano il pieno accesso ai servizi esistenti. L'esclusione può rappresentare una delusione, o motivo di sconcerto o frustrazione, ma non costituisce la violazione di alcun diritto legittimo.

La *channeling demonstration* viene condotta in dieci stati, e all'interno di ciascun stato in alcune città o gruppi di contee. In ciascun sito l'intervento viene gestito o da agenzie appositamente costituite o più spesso da strutture pubbliche esistenti, che ricevono fondi addizionali per realizzare i due modelli di intervento che si intendono testare. Il disegno della valuta-

zione viene affidato mediante bando di gara a due organizzazioni di ricerca private: *Mathematica Policy Research* si aggiudica la gestione dell'esperimento vero e proprio - dalla casualizzazione dei partecipanti all'analisi dei risultati - mentre l'*Institute on Aging* della *Temple University* si occupa dell'implementazione delle procedure e del loro monitoraggio, per controllare che gli interventi realizzati corrispondano al disegno originario.

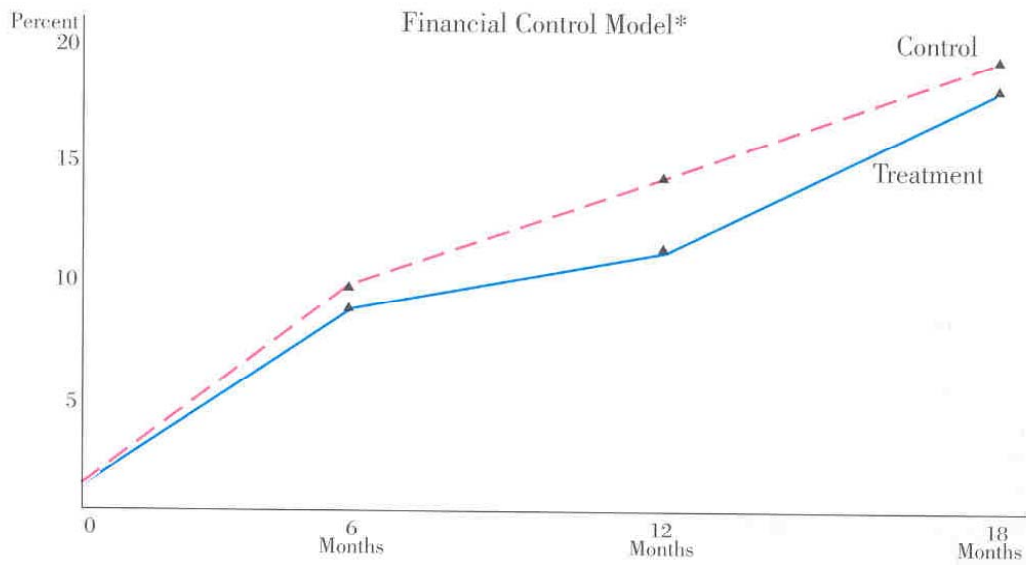
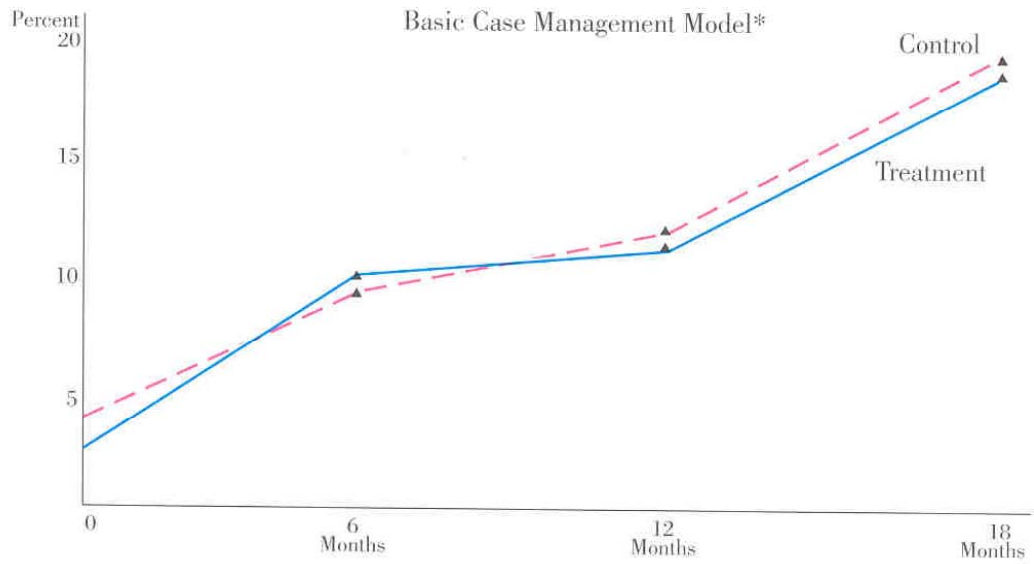
Circa 11.000 anziani oltre i 65 anni vengono contattati nei 10 siti tra il 1982 e il 1984. Tra questi solo 6.300 sono selezionati per partecipare all'esperimento, sulla base della presenza di determinate caratteristiche, soprattutto la presenza di forme di disabilità e limitazioni nel funzionamento, e la scarsità del supporto prestato da parte dei familiari. L'obiettivo (si scoprirà più avanti solo parzialmente raggiunto) è quello di selezionare anziani che siano ad alto rischio di istituzionalizzazione, perché è questa la popolazione per cui la disponibilità di servizi addizionali può veramente "fare differenza". Una metà del campione viene assegnata a ciascuno dei due modelli di intervento, e a sua volta all'interno di ciascun sottocampione una metà degli anziani viene inclusa nel gruppo di controllo, cioè non riceve alcun servizio addizionale. Quindi esistono due gruppi di controllo, uno per ciascun modello di intervento (questo perché i due modelli non sono applicati in modo omogeneo nei 10 siti, quindi non è possibile avere un unico gruppo di controllo.)

Gli effetti attesi della *channeling demonstration*, che la valutazione intende misurare, coprono un'ampia gamma di indicatori: si va dagli effetti diretti, quali l'aumento nell'uso di servizi nella comunità e la riduzione nei ricoveri in casa di riposo e in ospedale, ad effetti meno diretti, quali la (temuta) riduzione dell'aiuto prestato da parenti e vicini, al miglioramento nella qualità della vita, alla riduzione della mortalità. Una varietà di fonti informative vengono utilizzate per ottenere dati su questa gamma di indicatori: da interviste periodiche con gli anziani ancora in vita all'uso di basi dati amministrative ai documenti dello stato civile.

La valutazione della *channeling demonstration* produce una gran quantità di risultati, che vengono analizzati a fondo negli anni successivi. I rapporti finali della valutazione vengono pubblicati nel 1986, cioè quasi sei anni dopo l'inizio ufficiale dell'intero processo, e ancora oggi, a dieci anni di distanza, rappresentano una preziosa fonte di informazione per gli esperti del settore. Non è qui possibile neppure tentare di rendere conto di una gamma così ampia di risultati. A scopo puramente esemplificativo, ci limitiamo ad illustrare i risultati relativi all'effetto sui ricoveri in casa di riposo, mostrati nella Figura 9. Il grafico mostra chiaramente che l'effetto atteso di riduzione dei tassi di istituzionalizzazione non si è verificato. Nel caso del modello di *basic case management*, non si osserva alcuna differenza tra partecipanti all'intervento e membri del gruppo di controllo. Una piccola riduzione appare nel caso del modello di *financial control*, ma la differenza non risulta essere statisticamente significativa. Per quanto deludente, questo è un risultato importante: mostra come, per il tipo di popolazione selezionata,

Fig. 9

ANALISI DI IMPATTO DELLA *NATIONAL LONG TERM CARE DEMONSTRATION*  
(Percentuale dei partecipanti entrati in casa di riposo)



Source: Wooldridge and chore. 1960, table C.1.  
None of the treatment/control differences is statistically significant.

la disponibilità dei particolari servizi addizionali offerti non porta ad alcun beneficio in termini di accresciuta permanenza nella comunità. E possibile che per questo tipo di popolazione solo eventi catastrofici (peggioramento radicale della salute, perdita di un familiare) portino al ricovero in casa di cura: in questi casi, la disponibilità dei servizi addizionali offerti sostanzialmente non fa differenza.

Un'osservazione importante è che il tasso di istituzionalizzazione osservato in questa popolazione è relativamente basso: a 18 mesi dal momento della selezione nel campione, meno del 20 per cento degli anziani ancora in vita è stato ricoverato in casa di cura. Questo significa che il campione selezionato per l'intervento non era in origine ad alto rischio di istituzionalizzazione. Ma questo rappresenta un problema diverso da quello della valutazione di impatto di un intervento sui diretti beneficiari: rappresenta piuttosto un problema di scelta adeguata dei beneficiari. E possibile che, per una popolazione a più alto rischio, l'intervento si riveli più efficace. Ma questa è un interrogativo a cui non si può dare risposta sulla base dei dati generati dalla *channeling demonstration*.

### **3.3 La valutazione delle altre componenti dell'efficacia complessiva**

In questa sezione presentiamo alcune considerazioni sugli aspetti metodologici più rilevanti per valutare le altre componenti dell'efficacia complessiva delle politiche attive.

#### *3.3.1 La scelta dei beneficiari (targeting)*

La metodologia per l'analisi dell'impatto diretto discussa nella sezione precedente tende a trattare i beneficiari di un intervento come se fossero omogenei dal punto di vista della possibilità di modificarne comportamenti e condizioni. I risultati delle valutazioni di impatto diretto sono infatti spesso presentati in termini di effetti medi per il complesso dei beneficiari. Il passo successivo è quello di considerare esplicitamente l'eterogeneità dei beneficiari e le sue implicazioni per l'efficacia complessiva dell'intervento. Tre aspetti diversi di tale eterogeneità vanno presi in considerazione.

Il primo è *l'eterogeneità nell'impatto*: per alcuni beneficiari l'intervento può essere più efficace che per altri. E importante sottolineare che qui intendiamo eterogeneità nell'impatto netto, non tanto di differenze nei risultati osservati tra i beneficiari dopo l'intervento. Il problema sta nel fatto che alcuni beneficiari sono più "suscettibili di miglioramento" di altri. Quanto più un intervento è in grado di individuare i casi più suscettibili di miglioramento e di dare loro priorità nella fruizione delle prestazioni, tanto più alta sarà la sua efficacia, a parità di risorse utilizzate.

Nella sua forma più semplice, l'individuazione dei casi più "migliorabili" richiede un'estensione dell'analisi di impatto diretto: se la dimensione

del campione lo consente, l'analisi di impatto può venire replicata per sottogruppi identificati sulla base di caratteristiche osservabili. Questo tipo di analisi può rivelare se l'intervento è più efficace per alcuni sottoinsiemi della platea dei beneficiari che per altri. L'informazione può quindi essere utilizzata per usare le risorse disponibili in modo da massimizzare, date certe risorse, l'efficacia complessiva dell'intervento. Ad esempio, una delle indicazioni che chiaramente emergono dalla valutazione della *supported work demonstration* discussa precedentemente è che l'impatto di questo tipo di intervento è positivo per alcune categorie di beneficiari (ad esempio, le madri non sposate), quasi nullo per altre (*i drop-outs*). Se la quantità di risorse che la collettività decide di dedicare a questo tipo di intervento è fissata a priori, l'efficacia complessiva sarà massimizzata mirando l'intervento alle madri non sposate.

Un secondo aspetto dell'eterogeneità dei beneficiari è dovuta alle *differenze nel loro stato di "bisogno"*. Un caso particolare di bisogno, rilevante soprattutto per le politiche attive, è il fatto di essere più o meno esposti al rischio di particolari eventi o situazioni che possano portare alla povertà o all'esclusione sociale. Nella semplice analisi di impatto diretto, nel valutare l'impatto non viene dato peso alcuno al diverso stato di bisogno o di "rischiosità" dei beneficiari: quello che conta è *quanto* il gruppo migliora in media, non *chi* migliora. Considerazioni di equità sono tipicamente assenti dall'analisi di impatto. Quando però la prospettiva si allarga, e si vuole valutare la capacità dell'intervento di incidere su un certo fenomeno nel suo complesso, lo stato di bisogno o rischiosità dei beneficiari diventa più rilevante.

È possibile che si crei una tensione tra i due obiettivi, massimizzare l'impatto e rispondere a situazioni di maggiore bisogno. Ad esempio, un intervento teso a migliorare le possibilità di inserimento lavorativo può funzionare *meno* bene per gli individui che vivono in situazioni di estrema povertà. Quindi l'impatto sui beneficiari può essere migliorato se i più disagiati socialmente sono *esclusi* dall'intervento. Questo però riduce l'efficacia globale dell'intervento, se questa è espressa ad esempio in termini di riduzione del tasso di povertà.

Un terzo aspetto legato all'eterogeneità dei beneficiari, effettivi e potenziali, è quello che si manifesta con il cosiddetto *creaming*, o scrematura, che si verifica quando un intervento tende a servire beneficiari che hanno una maggior probabilità di mostrare un buon risultato finale, indipendentemente dalla loro capacità di migliorare la loro condizione iniziale, e indipendentemente dal loro bisogno.

Il *creaming* può venir favorito dall'uso di sistemi di misurazione della *performance* che pongono troppa enfasi su aspetti quali il numero di beneficiari serviti, o su misure di risultato *lordo* finale. Questo problema è stato studiato negli Stati Uniti nell'ambito degli interventi di formazione professionale di recupero per giovani socialmente svantaggiati (Dickinson e altri, 1988). Se questi interventi vengono giudicati sulla base, ad esempio,

della percentuale dei partecipanti che trovano lavoro, si crea un incentivo da parte degli operatori del servizio ad attrarre persone che a priori hanno migliori prospettive lavorative, invece che attrarre coloro che hanno maggior bisogno di essere aiutati o coloro che possono migliorare di più ma partendo da condizioni di partenza particolarmente sfavorevoli (e quindi con prospettive finali non ottimali). Ad esempio, per massimizzare il risultato lordo di un corso di formazione professionale può essere più vantaggioso attrarre disoccupati con un qualche passato lavorativo alle spalle, piuttosto che persone che siano state sempre al di fuori del mercato del lavoro.

Alla radice del *creaming* sta il fatto che il *livello* del risultato lordo finale e il *miglioramento netto* ottenuto (cioè l'impatto dell'intervento) non sono necessariamente correlati positivamente, e possono anche esserlo negativamente. L'unica soluzione a questo problema è un maggior ricorso, per quanto possibile, a valutazioni di impatto netto, che mettano cioè in evidenza dove il miglioramento è avvenuto a causa dell'intervento, non grazie ad una selezione strategica (e perversa) dei beneficiari da servire.

In termini di metodi di valutazione, i diversi problemi posti dall'eterogeneità dei beneficiari richiedono approcci diversi. Come si è detto, l'eterogeneità dell'impatto può essere valutata come sottoprodotto dell'analisi di impatto diretto. L'evidenza ottenuta sul grado di migliorabilità di diversi tipi di beneficiari può quindi essere usata in un contesto più ampio, sia per giudicare a posteriori perché un certo intervento sia più o meno efficace, sia per dare indicazioni a priori su come gli interventi possono essere mirati meglio. Valutare l'eterogeneità del bisogno non richiede una previa analisi di impatto, bensì il semplice confronto di caratteristiche tra campioni estratti dalla popolazione dei beneficiari e di quella dei non-beneficiari. Questo tipo di analisi è più utile nell'ambito di interventi a regime, in quanto è più utilmente svolta dopo che si sia verificato un "naturale" processo di selezione dei beneficiari. Lo stesso vale per l'analisi del *creaming*: si tratta di un fenomeno che va analizzato una volta che l'intervento sia andato a regime, mentre avrebbe poco senso nell'ambito di interventi-pilota.

Valutare la rischiosità dei beneficiari, nel senso della probabilità di esposizione ad eventi che portano a situazioni di esclusione o povertà, richiede la disponibilità di dati longitudinali. Un campione di individui va seguito nel tempo, in modo da osservare il verificarsi di tali eventi e da stimare quanto la probabilità dell'evento sia correlata con le caratteristiche osservabili dell'individuo. Un esempio di questo tipo di analisi è offerto da uno studio compiuto negli Stati Uniti per individuare, tra coloro che hanno di recente perso il lavoro, quali caratteristiche individuali siano in grado di predire che il lavoratore diventi un disoccupato di lunga durata, con il conseguente esaurimento del sussidio di disoccupazione e l'aumentato rischio di povertà (Decker e Dickinson, 1996). La motivazione dell'analisi è quella di offrire gli interventi di riqualificazione e assistenza alla ricerca del lavoro con priorità a quegli individui che manifestino un alto rischio di restare disoccupati per un lungo periodo.

### 3.3.2 La scala dell'intervento

La scala dell'intervento è una determinante fondamentale dell'efficacia globale, anche se può venire totalmente ignorata quando si voglia valutare l'impatto diretto di un intervento. In particolare, la valutazione degli interventi-pilota per sua natura non può dire nulla sull'adeguatezza della scala dell'intervento: non solo, ma le stesse stime di impatto ottenute in un ambito ristretto possono non essere generalizzabili ad una scala più ampia.

Con riferimento alla scala dell'intervento, emergono due tipi di domande, una di tipo descrittivo e una di tipo comportamentale. La prima è: quale frazione della platea dei beneficiari è effettivamente servita dall'intervento? Una metodologia che può essere usata per rispondere a questa domanda è quella di stimare separatamente la numerosità dei beneficiari effettivi e quella dei beneficiari potenziali. Il rapporto tra le due grandezze fornisce una stima del grado di penetrazione o diffusione dell'intervento. Strumenti di rilevazione diversi possono essere usati per stimare le due grandezze. Il numeratore del rapporto, cioè la numerosità dei beneficiari, può essere ricavato da dati di tipo amministrativo, purché le strutture preposte all'erogazione dell'intervento siano dotate di un sistema informativo adeguato. Stimare il denominatore (la numerosità dei beneficiari potenziali) invece solitamente richiede la disponibilità di un campione rappresentativo della popolazione complessiva, all'interno del quale sia possibile identificare il sottocampione dei potenziali beneficiari. Il veicolo più comune è rappresentato da indagini campionarie sulle famiglie. Quello che si richiede è che le caratteristiche che definiscono i potenziali beneficiari di un intervento (ad esempio, l'età, il reddito, la struttura familiare, la presenza di condizioni problematiche) siano state raccolte nel corso dell'indagine. Tale condizione non è sempre soddisfatta, soprattutto quanto la platea dei beneficiari è definita dalla presenza di condizioni particolari e non facili da rilevare, quali ad esempio l'alcolismo, l'uso di droga, la violenza domestica, la disabilità. L'alternativa è di realizzare indagini campionarie *ad hoc*, localizzate sulla rilevazione di queste particolari condizioni. Il rischio di indagini specializzate e mirate è però quello di non ottenere stime rappresentative della popolazione di interesse.

Quindi, per quanto concettualmente facile da definire, il grado di penetrazione di un intervento può essere difficile da stimare in modo attendibile. Questa difficoltà cresce ancora quando si voglia stimare il grado di penetrazione di un intervento realizzato a livello locale. Quanto più ci si sposta a livello locale tanto più cresce la difficoltà ad ottenere stime attendibili della popolazione dei beneficiari potenziali, in quanto raramente indagini sul complesso delle famiglie producono campioni rappresentativi per piccole aree. L'alternativa può essere quella di usare delle *proxies* per il denominatore: ad esempio, invece del numero di anziani soli, non autosufficienti e bisognosi di aiuto domestico, si può usare il numero di anziani che vivono soli, una grandezza che può essere ricavata dall'anagrafe della popo-

lazione. Inevitabilmente, i rapporti così calcolati tenderanno a produrre stime molto basse del grado di penetrazione dell'intervento, in quanto il denominatore tende ad essere sovrastimato.

Un caso particolare del grado di penetrazione di un intervento è rappresentato dal grado di soddisfacimento della domanda: la particolarità sta nel fatto che la platea dei potenziali beneficiari è definita dal fatto di aver in qualche modo espresso interesse o volontà di usufruire di un intervento. Le ragioni per cui solo parte della domanda viene soddisfatta possono essere le più diverse, e possono essere di per sé oggetto di indagine. (Questo rappresenta un tipico caso di analisi del processo di implementazione di un intervento). Dal punto di vista della misurazione, il vantaggio di questa situazione è che il contatto che entrambi beneficiari serviti e non serviti hanno avuto con la struttura amministrativa può consentire di stimare sia il numeratore sia il denominatore del rapporto utilizzando la stessa fonte informativa. Condurre questo tipo di valutazione a livello locale non rappresenta un ostacolo, bensì un vantaggio, in quanto la fonte informativa può esistere in alcune località ma non in altre.

La seconda domanda rilevante per valutare il ruolo della scala dell'intervento è di tipo comportamentale: quanto sono generalizzabili ad un intervento realizzato su scala vasta le stime di impatto diretto sui beneficiari ottenute mediante interventi pilota? Secondo la terminologia introdotta in precedenza, qual è la validità esterna di queste stime? Vi sono almeno due meccanismi attraverso cui la scala può influire sull'efficacia dell'intervento. Uno è l'esistenza di vincoli di tipo macroeconomico o macrosociale, che diventano rilevanti solo quando l'intervento entra a regime. L'esempio classico è quello degli interventi sul mercato del lavoro, che possono essere molto efficaci quando limitati a pochi individui, ma perdono via via di efficacia quando il numero dei beneficiari diventa una frazione rilevante della forza lavoro. L'altro meccanismo può dipendere dal fatto che, quando la scala aumenta, cambia la natura stessa dell'intervento. Cambiano le caratteristiche e i comportamenti degli operatori, cambiano le caratteristiche dei beneficiari, si possono modificare le norme sociali che influiscono sul comportamento dei beneficiari, cambiano le aspettative che la collettività ha nei confronti dell'intervento. La validità esterna delle stime di impatto diretto dipende quindi in modo cruciale dal particolare contesto, e non possono essere stabiliti criteri generali.

### 3.3.3 La presenza di effetti negativi

Nel caso delle politiche attive, gli *entry effects* (discussi nella Sezione 2.2) sono un aspetto relativamente poco rilevante. Per poter ipotizzare l'esistenza di un'attrazione dell'intervento sui non-beneficiari, è necessario che essi percepiscano l'intervento come efficace, o in qualche modo ad essi utile. In un certo senso, l'esistenza di un effetto di attrazione sarebbe un indizio che l'intervento sta avendo qualche effetto (positivo) sui beneficiari.

Più rilevante nel caso delle politiche attive ci pare il problema dello spiazzamento: gli effetti positivi ottenuti a favore dei beneficiari possono andare a scapito di altri soggetti, in eguale stato di bisogno, che non sono beneficiari dell'intervento. E possibile che una ridotta efficacia complessiva di un intervento sia dovuta a questo meccanismo di compensazione, che ovviamente non può essere colto quando la valutazione si concentra solo sui beneficiari diretti dell'intervento. Interrogarsi sullo spiazzamento è senza dubbio rilevante: tuttavia, le probabilità di ottenere una risposta empiricamente fondata sono scarse. La difficoltà maggiore è quella di verificare l'esistenza di un peggioramento da parte dei non-beneficiari (che di per sé sono più difficili da identificare), e soprattutto quello di stabilire un legame di causa-effetto tra l'esistenza dell'intervento e tale peggioramento. Considerando le difficoltà nello stabilire l'efficacia in positivo per i beneficiari, stabilire tale "efficacia in negativo" e quantificarne l'importanza pare un obiettivo irraggiungibile con i metodi propri dell'analisi di impatto. Un approccio alternativo è quello di usare metodi di tipo qualitativo, come interviste con testimoni privilegiati e *focus groups*, che possono almeno fornire qualche indizio sull'esistenza del fenomeno.

Inoltre, va messa in evidenza una sorta di scala di priorità logica, che spesso viene dimenticata nelle discussioni sugli effetti di spiazzamento. Prima di preoccuparsi operativamente di tali effetti, andrebbe stabilita l'esistenza dell'impatto dell'intervento sui beneficiari diretti. Se l'intervento non è efficace per quelli che serve direttamente, è improbabile che provochi un effetto di segno opposto su altri soggetti.

## **4. LA VALUTAZIONE DELL'EFFICACIA DELLE POLITICHE PASSIVE CONTRO LA POVERTA**

In questa sezione discutiamo come la decomposizione dell'efficacia complessiva proposta nella Figura 6 possa essere applicata alle politiche passive, con particolare attenzione alle politiche di trasferimento di reddito. La valutazione dell'impatto diretto gioca qui un ruolo meno rilevante rispetto al caso delle politiche attive. Maggiore importanza invece assumono la valutazione delle altre componenti, in particolare la scelta dei beneficiari e la scala dell'intervento.

### **4.1 L'impatto diretto sui beneficiari**

Nel caso delle politiche passive, l'impatto diretto sui beneficiari può apparire scontato, non necessitando quindi di alcuna valutazione per essere confermato: se l'intervento ha fornito una prestazione per definizione "utile" ai suoi diretti beneficiari, in quanto diretta ad alleviare una situazione di disagio, l'impatto non può essere che positivo, nonché proporzionale all'intensità della prestazione stessa. Il caso più ovvio è quello in cui la prestazione consista in un trasferimento in denaro, e l'indicatore con cui si misura il successo dell'intervento sia il reddito disponibile dei beneficiari (o la distanza tra il reddito disponibile e una predeterminata soglia di povertà). In questo caso appare quasi scontato che l'impatto diretto sia positivo, e pari all'ammontare del trasferimento stesso.

Questa presunzione di efficacia ignora tuttavia possibili reazioni da parte degli stessi beneficiari o di altri soggetti, tali da ridurre o eliminare del tutto l'aumento di benessere (di reddito) generato dall'intervento. Ad esempio, un aumento della pensione sociale agli anziani può sostituire, almeno in parte, eventuali aiuti in denaro forniti agli stessi anziani da familiari non conviventi. Quando questa eventualità si verifica, l'impatto netto sul reddito disponibile dell'anziano risulta minore rispetto all'aumento della pensione sociale. Lo stesso vale ad esempio per la fornitura pubblica di aiuto domestico a favore degli anziani: in tutto o in parte tale aiuto può sostituire l'aiuto prestato dai parenti o da membri del volontariato, lasciando la situazione effettiva dell'anziano sostanzialmente invariata. Quanto questi effetti di sostituzione siano rilevanti è ovviamente una questione empirica che va esaminata caso per caso. Quello che qui intendiamo sottolineare è che, anche

nel caso delle politiche passive destinate ai poveri, l'impatto diretto dell'intervento sui beneficiari non può essere dato sempre e interamente per scontato. Dal punto di vista della metodologia di analisi, l'approccio resta quello utilizzato per l'analisi di impatto diretto delle politiche attive: si tratta di stabilire la situazione controfattuale, mediante un confronto con gruppi di non-beneficiari.

Nel caso delle politiche passive va notato un limite importante all'utilizzo della valutazione di interventi-pilota: soprattutto le politiche di trasferimento sono difficili da attuare a favore di gruppi limitati e selezionati di beneficiari, come richiedono tali interventi. Anche quando l'intervento "aggiunge qualcosa," possono sorgere problemi legali e soprattutto di "immagine" da parte dell'ente che eroga la prestazione temporanea. Il pericolo è che il trattamento di favore venga visto come un'ingiustizia da coloro che ne vengono a conoscenza ma ne sono esclusi: o, d'altro canto, che tenda a generare da parte dei beneficiari stessi la presunzione di avere acquisito un diritto. Tutti questi pericoli sono presenti in teoria anche per le politiche attive, ma in quel caso la natura incerta dell'impatto dell'intervento o addirittura il costo che esso spesso impone sui beneficiari minimizzano le resistenze e le recriminazioni. Quanto più ci si sposta verso prestazioni di mero trasferimento, tanto più diventa difficile gestirne un'applicazione selettiva e limitata.

E quindi raro, anche nel contesto americano, che pure politiche passive contro la povertà vengano realizzate sotto forma di interventi-pilota. Quando *demonstrations* vengono utilizzate nel contesto delle politiche passive, si tratta solitamente di studiare gli effetti di modifiche nelle modalità di erogazione dell'intervento, non della sua istituzione o abolizione. Un esempio è offerto dalla *food stamp cash-out demonstrations*, condotte in quattro Stati americani alla fine degli anni ottanta, con lo scopo di determinare se l'erogazione dei sussidi alimentari a famiglie a basso reddito sotto forma di denaro, anziché nella tradizionale forma di *vouchers*, avrebbe causato una riduzione nella spesa alimentare da parte di queste famiglie e un deterioramento nel valore nutritivo della loro dieta (Fraker, Martini e Ohls, 1995). L'oggetto di queste *demonstrations* è quindi una politica passiva (erogazione di sussidi), ma l'intervento concretamente valutato è una modifica nelle modalità di erogazione del sussidio: invece che il normale sussidio vincolato nel suo utilizzo all'acquisto di generi alimentari, le famiglie selezionate come "beneficiarie" dell'intervento ricevono un equivalente in denaro (da qui il termine *cash-out*). Le famiglie selezionate come gruppo di controllo continuano invece a ricevere il normale sussidio, e neppure si rendono conto di essere parte di un esperimento fino a quando non vengono incluse nell'indagine finale tesa a misurare il livello e la composizione dei consumi alimentari di entrambi beneficiari e gruppo di controllo. L'impatto dell'intervento (trasformazione del *voucher* in denaro) è misurato come la differenza tra la media dei consumi alimentari dei due gruppi di famiglie, assegnate all'uno o all'altro gruppo mediante casualizzazione.

## 4.2 La scelta dei beneficiari

Questo è l'aspetto forse più rilevante della valutazione di efficacia delle politiche passive. La domanda fondamentale è: "in che misura il trasferimento (o l'aiuto prestato) raggiunge quelli in condizione di (maggior) bisogno o disagio?" Poiché la condizione di bisogno è il fenomeno che l'intervento vuole modificare, l'efficacia complessiva dell'intervento è minore quanto maggiore è la frazione dei beneficiari che "non contano" quando il fenomeno viene misurato, in quanto non si trovano in quella particolare condizione di bisogno. Ad esempio, se il fenomeno è caratterizzato in termini di diffusione della povertà, la politica è tanto meno efficace quanto più alta è la percentuale dei beneficiari dell'intervento che già vivono al di sopra della soglia della povertà.

La strumentazione analitica necessaria consiste sostanzialmente nell'elaborazione di indicatori che mettono a confronto lo stato di bisogno dei beneficiari con la dimensione complessiva del bisogno misurato nella popolazione di riferimento. Dal punto di vista concettuale, la definizione di questi indicatori è relativamente semplice. Le difficoltà sorgono invece nella loro elaborazione empirica, a causa della difficoltà di ottenere i dati necessari.

Prendiamo come punto di riferimento il caso più semplice (e più comune) in cui la politica passiva consiste in un trasferimento in denaro, e lo stato di bisogno è definito in termini di povertà economica (distanza del reddito disponibile, o di un'altra misura di benessere economico quali i consumi, dalla soglia della povertà). Una grande varietà di indicatori di efficacia (intesa qui nel senso ristretto di adeguatezza nella scelta dei beneficiari) può essere elaborata su questa base. (Nella discussione seguente consideriamo i beneficiari come unità generiche, senza addentrarci nel complesso problema di come il benessere di famiglie di diversa dimensione e composizione possa essere confrontato.)

Il più semplice di questo tipo di indicatori di efficacia è il rapporto tra il numero dei beneficiari che vivono al di sopra della soglia della povertà (o non corrispondono a qualche altro profilo di disagio o bisogno) e il numero totale dei beneficiari. Tanto più questo rapporto è vicino all'unità, tanto peggiore sarà il grado di *targeting*, e tanto minore l'efficacia dell'intervento. Spesso viene usato il termine di *leakage*, cioè perdita o spreco, per definire la situazione rappresentata da questo indicatore. Una versione più elaborata mette in rapporto non il numero di beneficiari, bensì i trasferimenti loro erogati: l'indicatore diventa il rapporto tra la spesa che va ai beneficiari non-poveri e la spesa per tutti i beneficiari. Questo secondo indicatore tiene conto della possibile correlazione tra ammontare del trasferimento e stato di bisogno: se ad esempio i beneficiari poveri ricevono in media un sussidio maggiore dei non-poveri, il rapporto sarà più lontano dall'unità che nel caso del semplice rapporto numerico, dando quindi una rappresentazione più pertinente dell'efficacia dell'intervento.

Questi indicatori possono essere costruiti utilizzando fonti diverse per stimare numeratore e denominatore del rapporto. Il numero totale dei beneficiari o la spesa relativa (il denominatore) può essere ottenuto da fonti di tipo amministrativo, mentre per il numeratore è necessaria una fonte che sia anche in grado di fornire, per ogni beneficiario, l'informazione aggiuntiva sullo stato di povertà o bisogno. A questo fine possono essere utilizzate indagini campionarie che contengano sia informazione utili a determinare lo stato di povertà di ciascuna unità (ad esempio il reddito o il consumo della famiglia) sia l'inclusione del sussidio tra le fonti di reddito della famiglia, in modo da determinare la presenza di un beneficiario nella famiglia.

Valutazioni più sofisticate dell'efficacia di un trasferimento possono essere condotte mediante *metodi di microsimulazione*. Invece che costruire semplici rapporti, in questo caso si parte da un campione rappresentativo della popolazione, e per ogni unità (ad esempio, famiglia) si simula cosa succederebbe se il trasferimento fosse rimosso dal suo reddito. La differenza tra il tasso di povertà complessivo ottenuto quando il trasferimento in esame è incluso nel reddito e quando esso è escluso dà una misura dell'efficacia anti-povertà dell'intervento. Quanto più il tasso di povertà è sensibile alla rimozione di un particolare tipo di trasferimento, tanto più mirato ai poveri risulterà l'intervento. Questo ovviamente sotto l'assunto che l'obiettivo dell'intervento sia quello di ridurre la povertà economica nella popolazione.

Una variante sullo stesso tema è ottenuta confrontando il *poverty gap* (cioè la distanza tra il reddito e la linea della povertà, sommata per tutte le famiglie povere) con e senza il trasferimento in questione. Questa seconda variante è più informativa della prima, in quanto riflette l'effetto del trasferimento sull'intensità della povertà invece che sulla sua diffusione. Ad esempio, se tutto il trasferimento va alle famiglie in condizioni di povertà estrema senza riuscire a rimuoverne nessuna dalla povertà, l'efficacia sarà nulla se misurata come riduzione del tasso di povertà, ma sarà alta se misurata come riduzione del *poverty gap*. Definizioni diverse di efficacia dell'intervento sono possibili in questo contesto, con implicazioni potenzialmente molto diverse tra loro. Questo tipo di simulazioni sull'efficacia anti-povertà dei vari trasferimenti pubblici sono prodotte e pubblicate regolarmente negli Stati Uniti da parte della Commissione Finanza del Congresso (*Ways and Means Committee*, vari anni).

Un vincolo all'uso di questo tipo di simulazioni è che lo stato di povertà di ciascuna unità deve essere determinato sulla base del reddito, non dei consumi. Gli esercizi di simulazione appena descritti, basati sulla scomposizione del reddito e sulla rimozione di alcune sue componenti, non avrebbero senso se poi lo stato di povertà venisse determinato sulla base di un'altra grandezza, quali i consumi. Questa limitazione non è molto rilevante negli Stati Uniti, dove di *routine* la povertà è misurata con riferimento al reddito, ma rappresenta un problema in altri contesti, come quello italiano, dove le misure di reddito ottenute nelle indagini campionarie sono ritenute insufficienti e i consumi vengono utilizzati per misurare la povertà.

### 4.3 La scala dell'intervento

Una ovvia fonte di inefficacia di una politica passiva contro la povertà è il fatto che solo una frazione della platea dei potenziali beneficiari effettivamente fruisca della prestazione. Qui ci poniamo in un'ottica diversa rispetto a quella trattata al punto precedente. Qui la fonte di inefficacia non è il *leakage*, lo spreco di risorse a favore di soggetti non in stato di bisogno, piuttosto la scarsa penetrazione dell'intervento.

Dal punto di vista metodologico, non resta molto da aggiungere rispetto a quanto discusso nel caso delle politiche attive. Anche in questo caso l'approccio è quello di stimare il numero dei beneficiari effettivi e di quelli potenziali, e calcolare il relativo rapporto, il *take-up rate*. La definizione di "beneficiari potenziali" qui tende a coincidere con gli "aventi diritto" al sussidio. La definizione di questi ultimi è concettualmente più semplice che nel caso delle politiche attive, in quanto è probabile che esistano specifiche condizioni che stabiliscono chi ha diritto a ricevere il sussidio e chi no, mentre nel caso delle politiche attive la nozione di "avente diritto" è più sfocata. D'altro canto, nel caso delle politiche passive può essere più difficile implementare il denominatore del rapporto dal punto di vista empirico. Le condizioni che danno diritto a ricevere il sussidio sono solitamente legate alla mancanza o insufficienza di altre forme di reddito. Se nell'indagine campionaria usata per determinare il numero degli aventi diritto il reddito viene sistematicamente sottodichiarato, il numero dei beneficiari potenziali viene sovrastimato, portando a concludere che il *take-up rate* sia più basso di quello che è realmente.

Negli Stati Uniti si osservano *take-up rates* sistematicamente inferiori al 100 per cento, soprattutto per alcuni interventi anti-povertà quali *i food stamps* (Doyle and Beebout, 1988), dove il tasso scende a circa il 60 per cento. La causa non sembra essere un artefatto statistico (la sovrastima degli aventi diritto), bensì in buona parte un fenomeno reale, dovuto a mancanza di informazione, disagio psicologico derivante al ricorrere all'aiuto pubblico, e difficoltà nell'accesso alle strutture amministrative che erogano i sussidi (*General Accounting Office*, 1988).

Un tipo diverso di valutazione della scala dell'intervento può essere condotto su un piano normativo invece che descrittivo. Piuttosto che riprodurre il numero degli aventi diritto, sulla base della legislazione corrente e dei dati disponibili, la preoccupazione in questo caso è di valutare in che misura la legislazione stessa è adeguata, cioè in che misura gli aventi diritto "coprono" la popolazione in stato di bisogno economico. E possibile infatti che le condizioni che danno diritto ad un sussidio siano troppo restrittive rispetto a qualche nozione di bisogno che si ritenga normativamente più adeguata. Questo avviene ad esempio quando esiste una molteplicità di sussidi, ciascuno dei quali copre un settore limitato della popolazione mediante regole di ammissibilità al sussidio che dipendono da una complessa combinazione di condizioni (età, reddito, storia lavorativa, composizione fami-

liare, disabilità). In questa situazione (che caratterizza il caso statunitense ma ancor più quello italiano, come argomentato da Rizzi e Rossi, 1996) è possibile che settori della popolazione in stato di bisogno restino "scoperti", in quanto non rientrano in nessuna delle complesse combinazioni di ammissibilità. La valutazione dell'efficacia si sposta in questo caso dal singolo sussidio al complesso dell'assistenza pubblica. Il rapporto interessante è il grado di "copertura" (*coverage*), cioè tra il totale dei beneficiari e il totale della popolazione in stato di bisogno o disagio economico.

#### 4.4 La presenza di effetti negativi

Nel caso delle politiche passive gli effetti di spiazzamento sono poco rilevanti, a parte il caso limite in cui l'accesso all'intervento sia razionato, e la corresponsione del sussidio ad un individuo implica la negazione dello stesso sussidio ad un altro. Tuttavia, tale situazione è più utilmente analizzata sotto il profilo del *targeting* delle risorse, piuttosto che di quello dello spiazzamento.

Altri effetti negativi sono invece potenzialmente rilevanti nel caso delle politiche passive, soprattutto gli effetti di attrazione, o *entry effects*. Ad esempio, la disponibilità di un sussidio o prestazione condizionata alla presenza di certi requisiti può creare l'incentivo da parte di alcuni individui ad alterare il proprio comportamento in modo da soddisfare tali condizioni. Senza la disponibilità del sussidio, l'individuo si sarebbe comportato diversamente. La concessione di un'abitazione a condizioni favorevoli ai senza-dimora può indurre alcuni individui che vivono in abitazioni molto degradate a presentarsi come (o addirittura a diventare) senza-dimora allo scopo di usufruire di tale concessione, in particolare se le abitazioni concesse siano di qualità migliore di quelle in cui costoro vivono correntemente. Si può certo argomentare che la platea dei potenziali beneficiari dell'intervento è in questo caso definita in modo troppo restrittivo, e dovrebbe includere sia i senza-dimora sia coloro che vivono in condizioni al di sotto di uno standard minimo. Tuttavia, se l'obiettivo dell'intervento è quello di ridurre la numerosità assoluta dei senza-dimora, la presenza dell'effetto di attrazione può portare a risultati apparentemente contraddittori, e l'intervento può essere giudicato del tutto inefficace, se non addirittura dannoso.

Il caso più comune di *entry effects* è quello che si manifesta quando le condizioni di ammissibilità al trasferimento includono l'insufficienza del reddito. Alcuni individui possono avere un incentivo a ridurre il proprio reddito, ad esempio rinunciando a lavorare, o ad alterare il modo con cui il reddito è ricevuto in modo da poterlo più agevolmente mascherare, ad esempio scegliendo forme di lavoro sommerso. In entrambi i casi la dimensione (reale o misurata) della platea degli aventi diritto al sussidio aumenta per effetto dell'intervento stesso. Un altro esempio è quello di un sussidio che sia condizionato al fatto che il beneficiario sia un capofamiglia non sposato. Questo

può creare l'effetto perverso di aumentare i divorzi (quantomeno di rendere il divorzio un'opzione più praticabile in caso di crisi matrimoniale) e di scoraggiare i matrimoni (ad esempio scoraggiando la regolarizzazione di unioni di fatto.)

Quello degli *entry effects* è un caso particolare tra gli incentivi perversi che molti affermano essere connaturati agli interventi di sostegno al reddito, in particolare quelli rivolti a persone povere in età lavorativa. Una vastissima letteratura esiste negli Stati Uniti su questo tema, in particolare sugli effetti che forme di sostegno hanno sull'offerta di lavoro e sulla durata della disoccupazione. Moffitt (1992) fornisce un'ampia rassegna della ricerca sugli *incentive effects* del sistema di *welfare* americano. Con questo tipo di ricerca, la valutazione di efficacia si fonde con il più generale programma di ricerca delle scienze sociali, che indaga sulle determinanti del comportamento umano e sulle sue interazioni con le politiche pubbliche. I metodi analitici usati in questo programma di ricerca si differenziano da quelli discussi in questo lavoro, e fanno largo uso di tecniche econometriche applicate a dati non-sperimentali.



## RIFERIMENTI BIBLIOGRAFICI

- Blalock, Ann (a cura di), *Evaluating Social Programs at the State and Local Level*, Kalamazoo: The Upjohn Institute, 1990.
- Bondonio, Piervincenzo e Francesco Scacciati, *Efficienza e produttività negli Enti Locali*, Roma: Nuova Italia Scientifica, 1991.
- Burkhauser, Richard e Aldrich Finegan, "Poverty and Minimum Wage: The End of a Relationship"; in *Journal of Policy Analysis and Management*, vol. 7, no. 4, 1988.
- Decker, Paul e Katherine Dickinson, *Evaluation of Workers Profiling and Reemployment Services Systems*, Unemployment Insurance Occasional Papers 96-1, Washington: US Department of Labor, 1996.
- Dickinson, Katherine e altri, *Evaluation of the Effects of JTPA Performance Standards on Clients, Services and Costs*, Washington: National Commission on Employment Policy, 1988.
- Doyle, Patricia e Harold Beebout, *Food Stamp Program Participation Rates*, Alexandria: Food and Nutrition Service, 1988.
- Duran, Patrice, Eric Monnier e Andy Smith, "Evaluation à la française: Towards a New Relationship between Social Science and Public Action", in *Evaluation*, Vol. 1(1), 1995.
- Fraker, Torn, Alberto Martini e Jim Ohls, "The Effects of Food Stamp Cash-out on Food Expenditures: An Assessment of the Findings from Four Demonstrations", *Journal of Human Resources*, vol. 30 no. 4, Fall 1995.
- General Accounting Office, *Food Stamps: Reasons for Nonparticipation*, Washington: US Government Printing Office, 1988.
- Hausman, Jerry e David Wise (a cura di), *Social Experimentation*, Chicago: University of Chicago Press, 1985.
- Heckman, James e Joe Hotz, "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs", in *Journal of the American Statistical Association*, vol. 84, 1989.
- Heckman, James e Richard Robb, *Alternative Methods for Evaluating the Impact of Interventions*, in Heckman e Singer (a cura di), *Longitudinal Analysis of Labor Market Data*, Cambridge: Cambridge University Press, 1985.
- Kemper, Peter e altri, *The Evaluation of the National Long Term Care Demonstration*, numero speciale di *Health Services Research*, Vol. 23, No. 1, 1988.
- Manski, Charles, *Identification Problems in the Social Sciences*, Cambridge: Harvard University Press, 1995.

- Manski, Charles, e Irwin Garfinkel (a cura di), *Evaluating Welfare and Training Programs*, Cambridge: Harvard University Press, 1992.
- Manpower Demonstration Research Corporation, *Summary and Findings of the National Supported Work Demonstration*, Cambridge: Ballinger Publishing Company, 1980.
- Mancini, Andrea, "Il controllo nelle strategie per l'efficienza" in Zuliani e altri, *Sistemi di controllo e valutazione di efficienza negli enti locali italiani*, Bologna: Il Mulino, 1993.
- Martini, Alberto, *Metodi statistici per la valutazione degli interventi pubblici*, in Società Italiana di Statistica, Atti della XXXVI Riunione Scientifica, vol. 1, 1992.
- Mohr, Lawrence, *Impact Analysis for Program Evaluation*, Pacific Grove: Brooks/Cole, 1988.
- Moffitt, Robert, *Incentive Effects of the US Welfare System*, in *Journal of Economic Literature*, Vol. XXX, No. 1, 1992
- Moffitt, Robert, "Evaluation Methods for Program Entry Effects", in Manski e Garfinkel (a cura di), *Evaluating Welfare and Training Programs*, Cambridge: Harvard University Press, 1992.
- Negri, Nicola e Chiara Saraceno, *Le politiche contro la povertà in Italia*, Bologna: Il Mulino, 1996.
- Rizzi, Dino e Nicola Rossi, "Minimo vitale e imposta sul reddito proporzionale", dattiloscritto, luglio 1996.
- Rossi, Peter, e Howard Freeman, *Evaluation: A Systematic Approach*, Newbury Park: Sage Publications, 1989.
- Ways and Means Committee, *Overview of Entitlement Programs (Green Book)*, Washington: Government Printing Office, vari anni.
- Wooldridge, Judith, e Jennifer Shore, *Evaluation of the National Long Term Care Demonstration: Channeling Effects on Hospital, Nursing Home, and Other Medical Services*, Princeton: Mathematica Policy Research, 1986.
- Zuliani, Alberto, Andrea Mancini e Giuseppina Filacchione, *Sistemi di controllo e valutazione di efficienza negli enti locali italiani*, Bologna: Il Mulino, 1993.

**COLLANA  
DELLA COMMISSIONE DI INDAGINE  
SULLA POVERTÀ  
E SULL'EMARGINAZIONE**

- *Verso una politica di lotta alla povertà. L'assegno per i figli e il minimo vitale.*  
Luglio 1995.
  
- *La povertà in Italia  
1980-1994.*  
Marzo 1996
  
- *Le politiche locali contro l'esclusione sociale*  
Giugno 1996
  
- *La povertà in Italia 1995*  
Luglio 1996
  
- *Le misure della povertà in Italia: scale di equivalenza e  
aspetti demografici*  
Ottobre 1996
  
- *Povertà abitativa in Italia 1989-1993*  
Gennaio 1997

Dipartimento per gli Affari Sociali  
Via Veneto, 56 - Roma

---

Pubblicazione della

***Presidenza del Consiglio dei Ministri***

***Dipartimento per l'informazione e l'editoria*** Direttore: Mauro Masi

Via Po, 14/16A - 00198 Roma - Tel. 06/85981

---

*Direttore*

Mirella Boncompagni

*Coordinamento editoriale*

Diana Agosti

*Realizzazione grafica*

Ufficio Grafico dell'Istituto Poligrafico  
e Zecca dello Stato presso il  
Dipartimento per l'informazione e l'editoria

*Stampa e diffusione*

Istituto Poligrafico e Zecca dello Stato  
Roma, Giugno 1997

---



*In copertina*

"XXI secolo"

Tecnica mista su tela cm. 35 x 50,  
serie anno 1989/92