

BANCA D'ITALIA

**L'efficienza
nei servizi pubblici**

Si ringraziano Daniele Franco, Francesco Zollino e gli altri partecipanti al convegno organizzato dalla Banca d'Italia sull'Efficienza dei servizi pubblici, tenutosi presso l'Amministrazione Centrale dell'Istituto l'11 ottobre 2002, nel corso del quale sono state presentate versioni preliminari dei saggi raccolti in questo volume.

Gli autori dei saggi contenuti in questo volume sono i soli responsabili delle idee e delle opinioni ivi espresse. Queste non riflettono la posizione della Banca d'Italia.

INDICE

Prefazione

Antonio Fazio p. 7

Introduzione

Fabrizio Balassone, Maura Francese e Raffaella Giordano

(Banca d'Italia)..... p. 11

1. Efficienza nei servizi pubblici: una rassegna della letteratura

Fabrizio Balassone, Maura Francese e Raffaella Giordano

(Banca d'Italia)..... p. 19

2. L'efficienza degli ospedali pubblici in Italia

Daniele Fabbri

(Università di Bologna) p. 83

Commento

Rosella Levaggi

(Università di Brescia)..... p. 155

3. Valutazione dell'efficienza delle compagnie di bus italiane e svizzere

Roberto Fazioli*, Massimo Filippini°, Michael Künzle+

(* Università di Ferrara)

(° Università di Lugano e Swiss Federal Institute of Technology)

(+ Swiss Federal Institute of Technology).....p. 175

Commento: Caratteristiche del network, meccanismi di sussidio ed efficienza nel trasporto pubblico locale

Giovanni Fraquelli*, Massimiliano Piacenza°

(* Università del Piemonte Orientale, Ceris-CNR e HERMES)

(° Ceris-CNR e HERMES).....p. 215

4. Qualità delle formazione scolastica, scelte formative ed esiti nel mercato del lavoro

Giorgio Brunello*, Daniele Checchi°, Simona Comi+

(* Università degli Studi di Padova, CESifo e IZA)

(° Università degli Studi di Milano)

(+ Università degli Studi di Milano e Essex University).....p. 233

Commento

Dino Rizzi

(Università di Venezia)p. 305

Hanno contribuito:

Fabrizio Balassone (*Banca d'Italia*)

Giorgio Brunello (*Università degli Studi di Padova, CESifo e IZA*)

Daniele Checchi (*Università degli Studi di Milano*)

Simona Comi (*Università degli Studi di Milano e Essex University*)

Daniele Fabbri (*Università di Bologna*)

Roberto Fazioli (*Università di Ferrara*)

Massimo Filippini (*Università di Lugano e Swiss Federal Institute of
Technology*)

Maura Francese (*Banca d'Italia*)

Giovanni Fraquelli (*Università del Piemonte Orientale, Ceris-CNR e
HERMES*)

Raffaella Giordano (*Banca d'Italia*)

Michael Künzle (*Swiss Federal Institute of Technology*)

Rosella Levaggi (*Università di Brescia*)

Massimiliano Piacenza (*Ceris-CNR e HERMES*)

Dino Rizzi (*Università di Venezia*)

*Stampato dal
Centro Stampa della Banca d'Italia
Roma, Luglio 2003*

PREFAZIONE

Antonio Fazio

Negli anni Novanta, il processo di riequilibrio dei conti pubblici si è basato sull'aumento della pressione fiscale; nella media del decennio questa ha raggiunto il 42,4 per cento, contro il 35,6 per cento degli anni Ottanta.

L'azione di controllo della spesa pubblica ha arrestato, ma non invertito, la tendenza crescente del peso delle erogazioni sul prodotto prevalsa fino alla metà dello scorso decennio. Nel 2002 le uscite primarie delle Amministrazioni pubbliche sono state pari al 41,6 per cento del prodotto, lo stesso livello del 1995.

Negli ultimi anni l'attenzione per l'utilizzo delle risorse assorbite dal settore pubblico si è accresciuta in connessione con la necessità di consolidare la riduzione del disavanzo fin qui ottenuta, nel rispetto delle regole di bilancio dell'Unione Monetaria Europea, e di conciliare la disciplina di bilancio con politiche favorevoli alla crescita economica.

La riduzione dell'incidenza della spesa sul prodotto consentirà di dare certezza alla riduzione della pressione fiscale programmata dal Governo. In tal modo si inciderà positivamente sulle aspettative degli operatori economici, stimolando la propensione ad investire delle imprese e la spesa delle famiglie.

La necessità di un'ulteriore opera di razionalizzazione e di contenimento della spesa trae origine anche da altri fattori: le tendenze demografiche, l'esaurimento dei benefici connessi con la riduzione dei tassi di interesse sul debito pubblico, la difficoltà di sostenere gli attuali livelli di prelievo fiscale in un contesto di crescente integrazione economica internazionale.

L'erogazione di servizi che assicurino livelli di assistenza adeguati e prestazioni di qualità soddisfacente richiede recuperi di efficienza nell'utilizzo delle risorse disponibili. Occorre mirare a una configurazione organizzativa nella quale la qualità e la quantità dei servizi vengano innalzate riducendo gli sprechi, aumentando l'efficienza del settore pubblico e l'efficacia degli interventi da esso posti in essere.

Nonostante l'estrema rilevanza del tema, non esiste una consolidata letteratura empirica dedicata all'analisi della produzione dei servizi pubblici. La molteplicità degli obiettivi perseguiti con l'azione pubblica, i tempi necessari perché essa espliciti i suoi effetti, la varietà dei bisogni che essa è chiamata a soddisfare rendono difficile la misurazione della produzione pubblica e dei suoi risultati. In Italia a queste difficoltà si aggiunge l'inadeguatezza dei dati disponibili.

Si tratta di una carenza grave. Analisi di qualità possono fornire sostegno all'attività gestionale degli amministratori e consentire alla collettività di controllarne adeguatamente l'operato.

L'esigenza di ampliare e sviluppare lo studio della produzione pubblica in Italia ha motivato l'organizzazione e la realizzazione di questo progetto di ricerca. Si tratta di un contributo dedicato a tre settori particolarmente rilevanti per la quantità di risorse pubbliche impiegate e per l'impatto sul sistema produttivo: la sanità, i trasporti pubblici locali e l'istruzione.

La spesa pubblica per la sanità in Italia non si discosta significativamente dai livelli medi riscontrabili nei paesi europei e in quelli aderenti all'OCSE. Tra la fine degli anni Ottanta e i primi anni Novanta la spesa era in forte espansione; gli interventi introdotti a partire dal 1992 ne hanno significativamente contenuto la dinamica. Vi hanno contribuito soprattutto le misure amministrative di contenimento della spesa farmaceutica e i più generali interventi attuati per il contenimento della spesa relativa al pubblico impiego. Negli anni più recenti la spesa farmaceutica ha ripreso a crescere.

Nonostante le riforme introdotte negli anni Novanta il grado di efficienza del servizio sanitario resta insoddisfacente. Per il comparto ospedaliero, le stime riportate in questo volume indicano l'esistenza di margini di recupero particolarmente ampi.

Due questioni rilevano per una riflessione sulle caratteristiche della rete ospedaliera: il sottodimensionamento di molti presidi, che incide significativamente sulla performance complessiva del settore; l'ampiezza dell'effettivo bacino di utenza, con cui appare strettamente correlato il grado di efficienza dei singoli presidi.

Come per la sanità anche nel trasporto locale le spese hanno a lungo ecceduto le risorse assegnate. La separazione tra la responsabilità di

finanziamento e quella di produzione del servizio ha prodotto effetti negativi in termini finanziari e di efficienza tecnica.

Seguendo una tendenza comune ad altri paesi europei, l'Italia ha avviato un processo di ristrutturazione volto a estendere al trasporto pubblico i benefici della concorrenza, assicurando la contendibilità del mercato di riferimento attraverso gare d'asta per l'aggiudicazione della concessione dei servizi di trasporto in un determinato bacino.

L'analisi condotta in questo volume mostra che le aziende di trasporto locale di dimensioni medio piccole non sfruttano pienamente i margini a loro disposizione per economie di scala e di scopo. La valutazione di questi margini può incidere sulla definizione dei bacini di riferimento per le gare d'asta, sia sotto il profilo dell'estensione territoriale, sia sotto il profilo delle tipologie di servizi offerti.

L'Italia spende per l'istruzione meno degli altri paesi industrializzati; nel confronto con la media dei paesi dell'OCSE appare bassa la quota di spesa in conto capitale. Viceversa è alta la spesa per il personale docente, in conseguenza dell'alto numero di docenti in rapporto agli studenti, parzialmente compensato da livelli salariali individuali più bassi della media OCSE.

Il maggiore rapporto tra il numero dei docenti e quello degli studenti non sembra influire positivamente sui risultati del sistema scolastico italiano. L'Italia si caratterizza per livelli di scolarità significativamente inferiori alla media dei paesi dell'OCSE. A parità di livello di istruzione, inoltre, gli studenti italiani presentano livelli di apprendimento inferiori alla media. Su questi risultati possono avere influito il sistema di reclutamento e quello di motivazione degli insegnanti.

I risultati dell'analisi offrono una conferma indiretta della bassa correlazione tra qualità dell'istruzione e rapporto tra docenti e studenti: l'incremento di tale rapporto registrato negli ultimi 30-40 anni sembra avere avuto un effetto positivo, ma di entità limitata, sul rendimento marginale dell'istruzione misurato in termini di reddito individuale.

Per ogni settore esaminato in questo volume si sottolineano aspetti critici e si suggeriscono spunti di riflessione. L'augurio è quello di avviare quell'analisi esaustiva, metodologicamente ed empiricamente solida, senza la quale interventi sulle modalità di fornitura dei servizi pubblici difficilmente possono fornire i benefici attesi.

Circa la metà della spesa primaria delle Amministrazioni pubbliche italiane riguarda la produzione di servizi. L'efficienza nell'utilizzo di un tale ammontare di risorse, dell'ordine del 20 per cento del prodotto, è determinante per la competitività del sistema economico e il benessere della collettività.

La notevole differenziazione nei costi e nella qualità delle prestazioni osservabile fra le varie aree geografiche del paese e fra strutture diverse evidenzia l'esistenza di ampi margini di intervento.

Il vero risanamento delle finanze pubbliche in vista di un rafforzamento della competitività e delle prospettive di crescita del nostro paese richiede un aumento della produttività nel settore pubblico che consenta di fornire alla collettività servizi migliori a costi più bassi.

INTRODUZIONE

Fabrizio Balassone, Maura Francese* e Raffaella Giordano**

Questo volume raccoglie le ricerche elaborate nell'ambito di un progetto coordinato dal Servizio Studi della Banca d'Italia. I lavori pubblicati sono stati presentati e discussi durante un convegno tenutosi l'11 ottobre del 2002 presso la sede centrale dell'Istituto.

L'intento del progetto è quello di fornire una panoramica degli strumenti utilizzabili per la valutazione dell'efficienza dei servizi pubblici, di aggiornare – per alcuni settori – i risultati disponibili, di fornire una base per impostare correttamente la discussione di un argomento di così evidente rilievo.

Il volume si apre con una rassegna della letteratura. Seguono tre studi che analizzano settori particolarmente rilevanti per la quantità di risorse pubbliche impiegate: la sanità, i trasporti pubblici locali e l'istruzione. A ciascuno studio fa seguito una nota di commento.

Fabrizio Balassone, Maura Francese e Raffaella Giordano illustrano il concetto economico di efficienza e descrivono le principali tecniche utilizzate in letteratura per la sua stima. Le difficoltà concettuali, di disponibilità dei dati e di interpretazione dei risultati che si incontrano nella valutazione dei servizi pubblici vengono discusse prendendo in esame diversi contributi di ricerca, relativi sia all'Italia sia ad altri paesi. Un'attenzione particolare viene rivolta ai settori analizzati negli altri saggi inseriti in questo volume.

Il lavoro rileva, in primo luogo, la difficoltà che si incontra nell'individuare le variabili più idonee a misurare il prodotto pubblico, in considerazione della sua natura "multiscopo", della variabilità dell'arco temporale durante il quale potranno essere osservati i risultati dei servizi offerti e della disomogeneità della distribuzione spaziale e temporale delle caratteristiche degli utenti.

* Servizio Studi, Banca d'Italia.

Inoltre, gli autori mettono in rilievo la forte dipendenza dei risultati delle analisi empiriche dalla specificazione dei modelli impiegati e dalla presenza di osservazioni estreme. La scarsa robustezza dei risultati empirici e la mancanza di consenso sulle tecniche più appropriate nei diversi ambiti suggeriscono cautela nell'interpretazione e nell'utilizzo delle stime, specie in quei settori, come la sanità e l'istruzione, dove le misure di prodotto presentano maggiori margini di incertezza.

L'analisi delle tecniche disponibili evidenzia limiti e aspetti critici in ciascuna di esse. Le metodologie più semplici, meno impegnative in termini di tecniche di stima e di dati necessari, non consentono di distinguere, nell'analisi del risultato conseguito da una unità produttiva, l'impatto esercitato dal grado di efficienza produttiva da quello determinato da disturbi di tipo stocastico o da elementi comunque esogeni alla sfera decisionale dell'unità produttiva (ad esempio, le condizioni ambientali). Tecniche più sofisticate possono ridurre l'entità del problema, ma non forniscono soluzioni pienamente soddisfacenti. La difficoltà principale è quella di isolare gli elementi che incidono sulla componente sistematica delle deviazioni rispetto al comportamento efficiente. Se i comportamenti di ciascuna unità produttiva sono caratterizzati da regolarità, la diffusione di dati di tipo *panel* può essere un utile strumento per affrontare questo problema.

Le diverse difficoltà sono esemplificate dagli autori con riferimento a uno studio condotto dall'OCSE per conto dell'Organizzazione mondiale della Sanità. Si tratta di uno studio tecnicamente raffinato e che utilizza una definizione di prodotto che cerca di misurare sinteticamente sia l'impatto economico, sia l'efficacia medica delle prestazioni fornite e che tuttavia conclude che "*... the available evidence is neither plentiful enough nor reliable enough, yet, to allow us to draw firm conclusions about variations in performance in the term suggested or about their causes. ...it would be unwise to base policy on such a finding without further investigation*".

Per quanto riguarda l'Italia, la rassegna degli studi sui settori della sanità, dell'istruzione e del trasporto pubblico mostra che nel nostro paese l'analisi applicata incontra, oltre ai problemi metodologici già rilevati, anche difficoltà connesse con la carenza di dati dettagliati e affidabili relativi alle singole unità produttive.

Alla luce dell'analisi condotta gli autori concludono sottolineando l'opportunità di corredare le analisi di efficienza produttiva con indicazioni relative ad altre dimensioni valutative, quali la qualità del servizio, come

percepita dagli utenti, e l'efficacia rispetto agli obiettivi che motivano la produzione pubblica.

Il lavoro di Daniele Fabbri è dedicato all'analisi dei servizi prodotti dagli ospedali pubblici in Italia. Fabbri effettua una stima dell'efficienza tecnica e di scala di ciascuna unità produttiva utilizzando la *Data Envelopment Analysis* (DEA) su dati relativi al 1999. Inoltre, l'autore affronta il problema della misurazione della qualità dei servizi attraverso l'analisi della mobilità dei pazienti.

Dall'analisi emerge che, per quanto riguarda l'efficienza tecnica, esistono margini di miglioramento piuttosto ampi. Le stime suggeriscono che l'attuale produzione dei presidi pubblici italiani potrebbe essere realizzata con il 66 per cento dei letti e del personale impiegati. Inoltre, solo il 5 per cento dei presidi esaminati opera alla scala ottimale. Sebbene non siano poche le strutture sovradimensionate, il problema principale della rete di offerta italiana appare il sottodimensionamento di troppi presidi.

L'analisi evidenzia problemi per tutte le categorie di ospedali esaminate. Il maggior grado di efficienza tecnica viene conseguito dai policlinici, i quali potrebbero continuare a realizzare la loro produzione impiegando circa l'85 per cento dei fattori attualmente utilizzati. Peraltro, i policlinici, in genere sovradimensionati, sono anche le strutture mediamente più lontane dalla dimensione ottimale. Gli ospedali di dimensioni medie sono quelli che meno si discostano dalla dimensione ottimale, ma presentano anche le peggiori *performance* produttive. Gli ospedali più piccoli hanno un grado di efficienza tecnica pari a quello medio del campione, ma risultano sistematicamente sottodimensionati.

Per tenere conto della dimensione qualitativa della produzione ospedaliera Fabbri adotta un approccio fondato sul "valore sociale" di ciascun presidio, misurato dalla perdita di benessere che la chiusura del presidio causerebbe ai suoi utilizzatori. Il valore sociale così definito viene stimato attraverso modelli di scelta individuale che sfruttano informazioni sulla mobilità dei pazienti. Fabbri conduce una analisi dei fattori che spiegano la variabilità del livello di efficienza dei vari ospedali e trova che il valore sociale è la variabile cruciale. Una volta che si controlli per tale fattore, né il tipo né le dimensioni del presidio sembrano essere collegati in

maniera sistematica con il grado di efficienza. L'aspetto più rilevante nel determinare l'efficienza di un presidio è il ruolo funzionale svolto sul territorio, piuttosto che l'insieme delle sue caratteristiche organizzative.

Fabbri conclude che il perseguimento dell'efficienza richiederebbe non tanto la trasformazione dell'assetto organizzativo delle unità produttive, quanto la razionalizzazione della rete di offerta.

Nel suo commento al lavoro di Fabbri, Rosella Levaggi sottolinea il carattere innovativo della stima del valore sociale e l'elevata qualità (e numerosità) del campione di unità produttive esaminato.

Secondo Levaggi, i risultati più interessanti del lavoro riguardano la centralità del valore sociale nel determinare l'efficienza dei presidi, la discontinuità delle economie di scala e il fatto che le Aziende ospedaliere mostrino congiuntamente una buona efficienza tecnica e una tendenza al sovradimensionamento. Per Levaggi quest'ultimo elemento suggerisce che i principi della logica concorrenziale, pur favorendo una maggiore efficienza tecnica, spingono le unità produttive a dotarsi di sovrastrutture, con il rischio di determinare nel tempo un aumento ingiustificato delle prestazioni a causa di fenomeni di domanda indotta.

Levaggi suggerisce qualche cautela con riferimento all'utilizzo dei DRG (*Diagnosis Related Groups*) per la misura del prodotto in quanto ne potrebbe derivare una distorsione da aggregazione nelle stime dell'efficienza tecnica e di scala. Levaggi sottolinea inoltre che la mobilità potrebbe essere legata più alla qualità percepita che a quella effettiva.

Possibili estensioni del lavoro vengono indicate sia nella dimensione temporale (per valutare, ad esempio, come le strutture sanitarie stiano reagendo alle riforme introdotte nel corso degli anni novanta), sia in quella territoriale (per valutare, ad esempio, se i risultati di diversi modelli organizzativi siano invariati rispetto alla loro collocazione regionale).

Roberto Fazioli, Massimo Filippini e Michael Künzle studiano l'efficienza produttiva nel settore del trasporto pubblico locale su gomma.

Seguendo una tendenza comune ad altri paesi europei, il settore del trasporto pubblico locale italiano è al centro di un processo di ristrutturazione volto ad assicurare la contendibilità del mercato. Lo strumento privilegiato è quello della gara d'asta per l'aggiudicazione della concessione di erogare i servizi di trasporto su un determinato bacino. Gli autori sottolineano che l'architettura industriale e istituzionale che si profila alla luce dei recenti interventi normativi è fondata sull'ipotesi, peraltro non sufficientemente sottoposta a verifica, che le economie di rete e di scala eventualmente presenti nel settore siano limitate al perimetro cittadino o alle singole linee di servizio.

Fazioli, Filippini e Künzle effettuano l'analisi attraverso la stima di una frontiera di costo. A scopo comparativo, lo studio è esteso alle aziende di trasporto svizzere. La banca dati utilizzata non copre la totalità delle imprese, ma contiene informazioni economico-finanziarie relative a più esercizi.

I risultati presentati evidenziano come la struttura produttiva delle aziende di trasporto locale italiane di piccola e media dimensione sia caratterizzata da economie di scala e di densità. La maggior parte di queste imprese opera ad una scala di produzione inferiore a quella ottimale. Economie di scala sono riscontrate anche per le aziende svizzere.

Gli autori sostengono che per favorire un migliore sfruttamento delle economie di scala e una migliore efficienza produttiva delle aziende di trasporto locale sarebbe necessario promuovere una politica di fusioni tra compagnie con rete di trasporto adiacente. Un migliore sfruttamento delle economie di scala spaziali richiederebbe la definizione di gare per l'affidamento dei servizi di trasporto pubblico locale che considerino reti di trasporto relativamente estese. I risultati sembrerebbero sconsigliare l'apertura di gare per l'affidamento di singole linee di trasporto.

L'evidenza empirica prodotta dagli autori suggerisce inoltre la presenza di inefficienza tecnica sia tra le imprese italiane, sia tra quelle svizzere. Gli autori rilevano che, nel caso italiano, il grado di efficienza non sembra essere collegato sistematicamente al tipo di servizio (urbano, extraurbano o misto).

Giovanni Fraquelli e Massimiliano Piacenza sottolineano come le indicazioni derivanti dal lavoro di Fazioli, Filippini e Künzle appaiano in controtendenza rispetto ad alcune operazioni di riorganizzazione del servizio di trasporto locale attualmente in corso in Italia.

Fraquelli e Piacenza confrontano l'evidenza presentata da Fazioli, Filippini e Künzle con la propria, ottenuta dallo studio di un diverso campione di imprese italiane e adottando un approccio metodologico per alcuni aspetti differente.

I loro risultati confermano l'opportunità di promuovere una politica di fusioni tra imprese di trasporto pubblico locale operanti su reti adiacenti al fine di consentire un migliore sfruttamento delle economie di scala. Tuttavia, Fraquelli e Piacenza ritengono che tale strategia debba essere attuata soprattutto tra operatori urbani ed extraurbani, in modo da creare imprese che, fornendo un servizio misto, riescano a sfruttare le economie da produzione congiunta.

L'analisi di Fraquelli e Piacenza suggerisce anche altri spunti di riflessione relativi alla regolamentazione delle imprese del settore e alle condizioni ambientali in cui esse operano. Gli autori sostengono che il ricorso a contratti di servizio a prezzo fisso potrebbe incentivare un maggiore impegno manageriale e che la regolamentazione della viabilità (corsie preferenziali, limiti alla circolazione e alla sosta di taxi e autovetture, incentivi all'uso dei mezzi pubblici), migliorando le condizioni operative delle reti, consentirebbe di conseguire riduzioni significative del livello dei costi.

L'ultimo settore oggetto di analisi è quello dell'istruzione. Giorgio Brunello, Daniele Checchi e Simona Comi analizzano la relazione fra alcuni fattori che influenzano la qualità dell'offerta di formazione e gli esiti osservati sul mercato del lavoro.

Brunello, Checchi e Comi partono dalla considerazione che il numero di studenti per docente si è significativamente ridotto in Italia nel corso della seconda metà del secolo scorso. Tale riduzione non si è distribuita in maniera omogenea né temporalmente, né sotto il profilo territoriale. Gli autori sfruttano tale duplice variabilità per verificare se, a

parità di altre condizioni, l'entità delle risorse investite per studente abbia inciso sulle retribuzioni percepite al termine del percorso formativo.

I dati utilizzati nello studio provengono, per quanto riguarda i redditi individuali, dall'indagine della Banca d'Italia sui bilanci delle famiglie italiane e, per quanto riguarda le caratteristiche del sistema scolastico, dagli annuari del Ministero della Pubblica Istruzione.

I risultati ottenuti suggeriscono che l'incremento del rapporto tra docenti e studenti abbia avuto un effetto positivo, anche se di entità ridotta, sul rendimento marginale dell'istruzione. Viceversa, altri indicatori della qualità del sistema scolastico non sembrano esercitare effetti sistematici se si controlla per gli effetti di coorte e per le caratteristiche del mercato del lavoro locale. Risultati analoghi si ottengono qualora si voglia valutare la relazione tra i fattori che influenzano la qualità dell'istruzione e la probabilità di ingresso nell'occupazione.

Gli autori riconoscono che le stime non consentono di individuare con precisione gli effetti sulle retribuzioni individuali poiché l'impatto positivo di un minor rapporto tra studenti e docenti sul livello di istruzione e sul suo rendimento marginale è in parte o del tutto compensato da un impatto negativo sulle retribuzioni per ogni dato livello di istruzione. Secondo gli autori, quest'ultimo effetto potrebbe essere dovuto al fatto che con l'aumentare della qualità scolastica individui relativamente più abili scelgono livelli di istruzione più elevati, riducendo l'abilità e la produttività medie di coloro che rimangono con livelli di istruzione inferiore.

Dino Rizzi nella sua discussione del contributo di Brunello, Checchi e Comi ricorda che la disponibilità di informazioni rilevanti ai fini dell'analisi dell'efficienza nel settore dell'istruzione è, in generale, molto limitata e sottolinea lo sforzo compiuto dagli autori nel preparare la base dati utilizzata per l'analisi empirica.

I risultati di Brunello, Checchi e Comi suggeriscono che l'aumento di risorse per studente registrato negli ultimi 30-40 anni non abbia dato luogo ad un apprezzabile aumento del livello delle retribuzioni. Rizzi si chiede se questo debba portare a mettere in discussione l'esistenza di una relazione tra qualità del processo formativo ed esiti sul mercato del lavoro

o, piuttosto, l'ipotesi per cui il rapporto tra studenti e docenti sarebbe una *proxy* adeguata della qualità della formazione scolastica.

Rizzi individua alcuni motivi per cui le risorse investite nell'istruzione possono avere avuto un impatto contenuto sulla qualità della formazione e quindi sulle retribuzioni. In particolare, egli suggerisce la possibilità che, oltrepassato un certo livello, un aumento del numero di docenti non migliori la qualità dell'istruzione. Inoltre, secondo Rizzi, nel periodo preso in considerazione da Brunello, Checchi e Comi, il sistema di reclutamento degli insegnanti, il livello delle loro retribuzioni, la mancanza di incentivi per un insegnamento "di qualità" e la riduzione del grado di selettività del sistema scolastico potrebbero avere inciso negativamente sulla qualità del processo formativo, controbilanciando gli effetti positivi di una riduzione del rapporto tra studenti e docenti.

EFFICIENZA NEI SERVIZI PUBBLICI: UNA RASSEGNA DELLA LETTERATURA

Fabrizio Balassone, Maura Francese* e Raffaella Giordano**

1. Introduzione

I risultati di un'attività economica volta alla fornitura di un servizio ad un pubblico di utenti, sia essa pubblica o privata, possono essere valutati sotto due profili: quello dell'efficacia e quello dell'efficienza. Un'attività utilizza risorse per fornire un prodotto attraverso il quale raggiungere uno scopo. L'*efficacia* si misura con riferimento al raggiungimento dello scopo. L'*efficienza* si misura con riferimento al rapporto tra risorse impiegate e prodotto ottenuto.

Nel settore privato la valutazione dell'efficacia di un'attività viene effettuata dal mercato: è il fatto che un prodotto sia acquistato dai consumatori, che un'azienda consegua un profitto, che conquisti quote di mercato, a fornire una misura osservabile dell'efficacia dell'attività svolta. Nel settore pubblico, invece, non esistono in generale indicatori specifici e osservabili dell'efficacia dell'azione svolta, diversi dalla generale sanzione elettorale.

La costruzione di indicatori di efficacia incontra diverse difficoltà, legate soprattutto alla definizione degli obiettivi e alla verifica del loro conseguimento. Fra i problemi più rilevanti si possono elencare i seguenti:

- a) i servizi pubblici spesso perseguono più obiettivi simultaneamente, il che rende necessario un qualche criterio di ponderazione (l'istruzione pubblica, ad esempio, mira a fornire nozioni, creare professionalità, formare cittadini, facilitare l'inserimento nel mondo del lavoro, ecc.);
- b) non sempre a ogni obiettivo corrisponde un risultato misurabile (l'obiettivo di un sistema sanitario può essere l'allungamento della vita media e il miglioramento della sua qualità, ma come si misura quest'ultima?);
- c) i risultati possono prodursi su un arco di tempo prolungato (l'attività di ricerca svolta oggi può produrre risultati tra molti anni);

* Servizio Studi, Banca d'Italia.

- d) gli utenti dei servizi non sono omogeneamente distribuiti (gli studenti di una scuola possono essere mediamente più “dotati” di quelli di un’altra; i pazienti trattati da un ospedale possono essere mediamente più “gravi” di quelli trattati da un altro).

In alcuni casi si è tentato di ovviare a questi problemi utilizzando indagini sul grado di soddisfazione dei cittadini come succedaneo del giudizio del mercato. Tuttavia, tali indagini presentano diverse difficoltà. In particolare, i giudizi possono non essere formulati sulla base di esperienze dirette. Inoltre, essi si fondano sul confronto tra risultati e attese; la disomogeneità di queste ultime può rendere problematico l’utilizzo delle indagini sul grado di soddisfazione per il confronto internazionale fra sistemi e modelli organizzativi dei servizi pubblici.

La valutazione dell’efficienza, seppure non priva di difficoltà specifiche, appare meno complessa. Per valutare se e quanto un’impresa, pubblica o privata, sia efficiente occorre verificare che non sia possibile produrre di più con le stesse risorse, o utilizzare meno risorse per ottenere la stessa quantità di prodotto. Ciò richiede la specificazione di funzioni di produzione o di costo, o almeno l’individuazione delle variabili che figurano in tali funzioni, nonché la misurazione dello scarto fra ciò che si osserva e i valori ottimi teorici.

Anche in questo caso esiste una difficoltà specifica per il settore pubblico: quella di doversi confrontare con prodotti che possono essere di qualità diversa, senza che questa sia approssimata da un prezzo. L’unità di misura del prodotto pubblico deve essere definita precisando anche i requisiti qualitativi standard che devono caratterizzare il prodotto da quantificare.

Questa rassegna costituisce il primo passo di un più ampio progetto di ricerca sui servizi pubblici. L’intento principale è quello di fornire una base di discussione e una panoramica, seppur non esaustiva, degli strumenti e dei risultati disponibili per la valutazione dell’efficienza dei servizi pubblici.

L’esposizione che segue è organizzata secondo la seguente struttura:

- a) definizione economica del concetto di efficienza;
- b) rassegna delle tecniche per la stima dell’efficienza;
- c) esemplificazione, con riferimento al settore sanitario, delle difficoltà concettuali, di disponibilità dei dati e di interpretazione dei risultati

che si incontrano nel tentativo di valutare i risultati dell'azione pubblica;

- d) rassegna degli studi condotti in Italia su alcuni dei principali servizi pubblici (sanità, istruzione, trasporti).

2. Il concetto di efficienza

Nella letteratura economica per efficienza produttiva di un'impresa si intende, nello spazio degli *output*, la corrispondenza tra *output* effettivamente prodotto e *output* massimo potenzialmente producibile, data la tecnologia e gli *input*; o, nello spazio degli *input*, la corrispondenza tra la quantità di *input* effettivamente utilizzata e quantità minima di *input* potenzialmente utilizzabile, data la tecnologia e il livello di *output*.

Da un punto di vista teorico la difficoltà che si incontra una volta che ci si confronta con il problema della non corrispondenza tra comportamenti effettivi e comportamenti potenzialmente ottimi sta nella spiegazione delle cause della suddetta mancata coincidenza (se cioè essa sia imputabile a inefficienza tecnica o allocativa).

Da un punto di vista pratico la difficoltà è invece rappresentata dalla definizione empirica dei livelli potenziali massimi di *output* e potenziali minimi di *input* con cui confrontare i rispettivi livelli effettivamente osservati. Entrambe le difficoltà sono acuite dalla natura multiprodotto delle imprese che implica la necessità di utilizzare indici di *output* e di *input*.

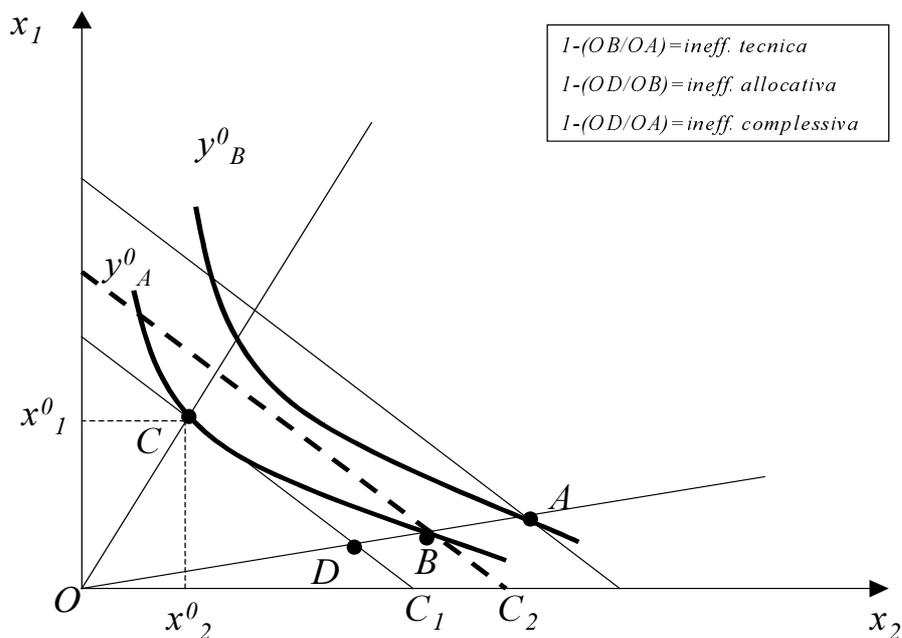
2.1 Inefficienza tecnica e inefficienza allocativa

Un'impresa si trova in una condizione di efficienza produttiva se non sopporta costi in eccesso rispetto a quelli minimi necessari a produrre un determinato *output*. Questa condizione è caratterizzata da assenza di costi indotti, da un lato, da *inefficienza tecnica*, ossia da un uso eccessivo di fattori – dati l'*output* prodotto e la tecnica produttiva – e, dall'altro, da *inefficienza allocativa*, ossia da una proporzione di impiego dei fattori che non minimizza la spesa – dati i prezzi e l'*output* prodotto.

Queste definizioni trovano una semplice illustrazione grafica nella Figura 1, che descrive il classico esempio di una funzione di produzione

Figura 1

Inefficienza tecnica e allocativa



in cui due *input*, x_1 e x_2 , sono utilizzati per produrre un *output*, y . L'isoquanto, denominato con y_A^0 , indica tutte le combinazioni di x_1 e x_2 che danno luogo allo stesso livello di *output*, y_A^0 . La pendenza dell'isoquanto misura il tasso marginale di sostituzione tra i fattori della produzione, ossia di quanto l'impiego di x_2 debba essere ridotto per mantenere costante il livello di *output* se aumenta di un'unità l'impiego di x_1 . La retta dell'isocosto, denominata con C_1 , indica le combinazioni di x_1 e x_2 che determinano lo stesso livello di spesa. La pendenza dell'isocosto è pari al rapporto tra i prezzi dei fattori, $-p_1/p_2$, dove p_i indica il prezzo del fattore x_i .

L'efficienza allocativa si ottiene quando la combinazione degli *input* è tale da minimizzare il costo di produzione, per un dato livello di *output* o, analogamente, da massimizzare il livello di *output*, per un dato livello di

spesa. Nella Figura 1 ciò si verifica nel punto C, dove le pendenze dell'isoquante e dell'isocosto coincidono; l'efficienza allocativa implica pertanto che:

$$\frac{MP_1}{MP_2} = \frac{p_1}{p_2} \quad (1)$$

dove MP_i indica la produttività marginale dell'*input* i .

L'efficienza tecnica si ottiene quando non vi è un utilizzo eccessivo di entrambi i fattori. Supponendo che gli isoquanti y_A^0 e y_B^0 in Figura 1 siano entrambi associati allo stesso livello di *output*, i punti dell'isoquante y_B^0 , che è più lontano dall'origine rispetto a y_A^0 , individuano situazioni di inefficienza tecnica.

Se la situazione di un'impresa è rappresentata dal punto A, essa opera in una situazione di inefficienza, sia tecnica, sia allocativa.

Il grado di efficienza tecnica è misurato dal rapporto OB/OA; esso indica la proporzione degli *input* utilizzati che è effettivamente necessaria per la produzione del livello di *output* osservato. L'inefficienza tecnica è data da $1-(OB/OA)$, che indica la misura in cui potrebbe essere ridotto il costo di produzione del livello di *output* osservato, mantenendo invariato il rapporto tra i fattori.

Il livello di efficienza allocativa è misurato da OD/OB (o, analogamente, da C_1 / C_2) e rappresenta il costo che si sarebbe sostenuto in presenza di efficienza allocativa in rapporto a quello effettivo. L'inefficienza allocativa è data da $1-(OD/OB)$.

L'efficienza complessiva è definita come il rapporto OD/OA ed è il prodotto tra efficienza allocativa ed efficienza tecnica. Infine, l'inefficienza complessiva è data da $1-(OD/OA)$ e misura di quanto i costi eccedono il loro minimo potenziale; essa può essere scomposta, con una certa approssimazione, nella somma dell'inefficienza allocativa e quella tecnica.

I livelli di inefficienza tecnica possono essere utilmente sintetizzati attraverso il concetto di funzione di distanza, duale rispetto a quella di costo. Dato l'insieme $P(x)$ definito come:

$$P(x) = \{y \in R_+^M : x \text{ può produrre } y\} \quad (2)$$

Definiamo la funzione di distanza come:

$$D(x, y) = \min \left\{ \theta : \left(\frac{y}{\theta} \right) \in P(x) \right\} \quad (3)$$

La funzione di distanza così definita assume valori inferiori o uguali a 1, ed è pari a 1 per combinazioni efficienti dei fattori nella produzione dell'output. Si consideri la Figura 2; un'impresa caratterizzata da un vettore degli *input* osservati pari a x_A e un *output* prodotto uguale a y^A è un'impresa tecnicamente efficiente al 100 per cento ed è caratterizzata da un valore della funzione di distanza pari a 1, mentre nel caso in cui insieme ad un pari mix e livello di *input* si osservasse solo un livello di produzione uguale ad y^B il livello di efficienza dell'impresa sarebbe inferiore al 100 per cento e la funzione di distanza avrebbe un valore inferiore a 1.

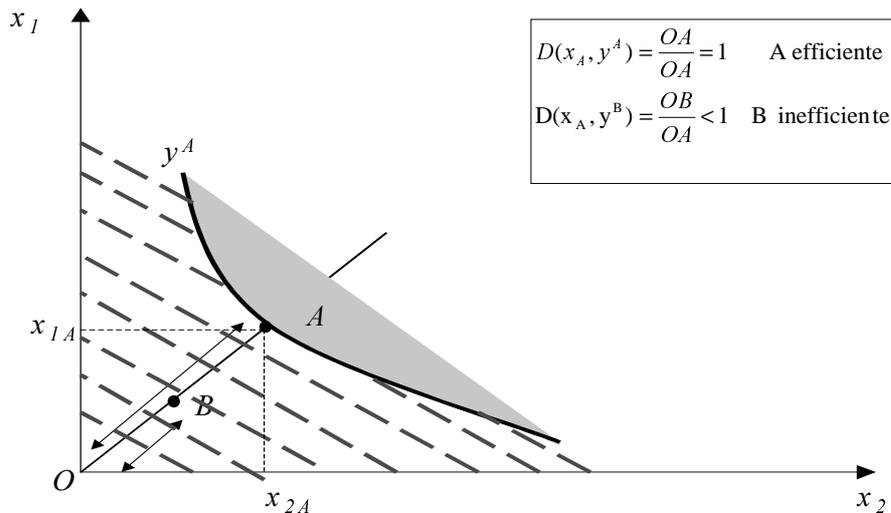
2.2 Inefficienza tecnica pura e inefficienza di scala

Nel contesto della valutazione dell'inefficienza tecnica di un'unità produttiva è possibile approfondire la valutazione considerando anche il

Figura 2

Funzione di distanza

Funzione di Distanza



tipo di rendimenti di scala caratterizzanti la tecnologia impiegata dall'impresa (Färe, Grosskopf e Lovell, 1985).

Il concetto di efficienza di scala fa riferimento al confronto fra il vettore *input-output* effettivamente osservato e quello che si determinerebbe nella situazione teorica di equilibrio concorrenziale Pareto-efficiente di lungo periodo (e dunque con imprese caratterizzate da una condizione di profitti pari a zero).

Per poter effettuare questo confronto è necessario costruire una frontiera di produzione estesa rispetto a quella osservata e caratterizzata da rendimenti di scala costanti. Tale estensione si ottiene considerando la tecnologia con il più piccolo insieme di produzione a rendimenti costanti di scala che sia in grado di contenere l'insieme di produzione a rendimenti di scala variabili effettivamente osservato (ad esempio nella Figura 3 RSC e RSV rispettivamente). La differenza fra la situazione osservata e quella teorica permette di definire l'inefficienza tecnica pura e quella di scala.

Considerando ad esempio una tecnologia a un solo *input* e un solo *output* come nella Figura 3, il punto *A* descrive un'impresa efficiente dal punto di vista tecnico, sia per quanto riguarda l'aspetto tecnico puro che per quanto riguarda la scala di produzione. Al contrario unità produttive descritte dai punti *B* e *B'* sono imprese inefficienti. Si consideri l'impresa *B'*: non solo il livello di *output* $y_{B'}$ potrebbe essere prodotto con un livello inferiore di *input* ($x_{D'} < x_{B'}$) data la tecnologia disponibile, ma rispetto al punto di riferimento teorico di una tecnologia a rendimenti costanti di scala è possibile individuare un ulteriore eccesso di utilizzo dell'*input* ($x_{D'} > x_{C'}$) e quindi un altro fattore di inefficienza. Analogamente per quanto riguarda l'impresa *B* è possibile rintracciare non solo una componente di inefficienza tecnica pura ($x_D < x_B$), ma anche una componente di inefficienza di scala associata alla distanza fra la frontiera di produzione data la tecnologia a rendimenti di scala variabili e quella teorica a rendimenti costanti di scala ($x_D > x_C$).

3. La stima dell'efficienza

La valutazione dell'efficienza di un'attività produttiva richiede la specificazione delle variabili che figurano nelle funzioni di produzione o di costo che caratterizzano tale attività. I valori assunti dalle variabili nelle

Tabella 1**Le tecniche di stima dell'efficienza**

	"Di frontiera"		"Non di frontiera"
Non parametriche	Deterministiche	<i>Data Envelopment Analysis</i> (Farrell, 1957)	
	Stocastiche		
Parametriche	Deterministiche	<i>Deterministic Frontier Model</i> (Aigner e Chu, 1968)	<i>Analisi di regressione semplice</i> (Feldstein, 1967)
	Stocastiche	<i>Stochastic Frontier Model</i> (Aigner e altri, 1977; Meeusen e van den Broeck, 1977)	

La tecnica utilizzata può inoltre essere "parametrica", nel caso in cui sia prevista la stima di una funzione di produzione o di costo, oppure "non parametrica", nel caso in cui tale stima non venga effettuata e la frontiera sia individuata specificando alcune proprietà che i punti ad essa appartenenti devono soddisfare.

Infine, si distingue tra misure dell'efficienza stocastiche e deterministiche a seconda del fatto che la tecnica utilizzata consenta o meno, almeno in linea di principio, di tenere conto dei fattori di natura stocastica che possono influire sui risultati delle unità produttive a prescindere dal loro grado di efficienza.

3.1 *Approcci non di frontiera*

Uno dei primi tentativi di misurare l'efficienza sulla base di tecniche standard di regressione è contenuto nel lavoro di Feldstein (1967). Si consideri una funzione di produzione di tipo Cobb-Douglas:

$$\ln y_i = \beta_0 + \sum_j \beta_j \ln x_{ij} + u_i \quad (4)$$

dove y_i denota l'*output* del produttore i , x_{ij} sono gli *input* e u_i è il termine di errore; i coefficienti β_j possono essere interpretati come elasticità del prodotto a variazioni nell'*input* j . La stima di tali coefficienti, applicando ad esempio il metodo dei minimi quadrati ordinari, consente di calcolare i residui \hat{u}_i . Feldstein suggerisce di utilizzare questi ultimi come una misura dell'inefficienza tecnica. Poiché la funzione di produzione è stimata su dati effettivi e non è ottenuta sulla base di valori potenziali, di fatto $\hat{u}_i = 0$ implica un livello di inefficienza pari a quello medio; $\hat{u}_i > 0$ ($\hat{u}_i < 0$) implica un'inefficienza inferiore (superiore) alla media.

Sulla base dell'equazione (4) si può valutare anche la presenza di inefficienza allocativa. Per tale funzione di produzione, la condizione (1) diventa:

$$\frac{\beta_k}{\beta_j} = \frac{p_k x_k}{p_j x_j} \quad (5)$$

ossia, l'efficienza allocativa implica che, per ciascuna coppia di *input*, il rapporto tra le spese per i due fattori sia pari al rapporto tra le elasticità del prodotto a tali fattori. Dal momento che il valore dei coefficienti β non varia al variare dell'ammontare degli *input* utilizzati, il membro di sinistra dell'equazione (5) può essere considerato come dato. Pertanto, qualsiasi differenza tra i due termini dell'eguaglianza può essere attribuita ad un uso non corretto di uno o più *input*. Una volta ottenuta una stima dei β , è possibile determinare se un *input* è sopra/sotto utilizzato rispetto ad un altro. Calcolando le differenze tra i due termini dell'equazione (5) per ciascuna coppia di *input*, si ottiene una stima del costo complessivo dell'inefficienza allocativa¹.

L'efficienza complessiva di un'unità produttiva può essere valutata utilizzando la funzione di costo; si assuma che essa abbia la forma:

$$c_i = \beta_0 + \sum_j \beta_j x_{ij} + u_i \quad (6)$$

dove c_i è il costo totale o medio dell'unità i , x_j include variabili come il livello dell'*output* e i prezzi degli *input*, e u_i è il termine di errore. Feldstein (1967) suggerisce che l'efficienza può essere valutata

¹ Si veda anche Schmidt e Lovell (1979).

confrontando il costo effettivo al costo ‘atteso’ sulla base della (6). Come in precedenza, il residuo può essere utilizzato come misura dell’efficienza: le unità con residui positivi sono meno efficienti rispetto alla media, quelle con residui negativi sono relativamente più efficienti. Tuttavia, va rilevato che le deviazioni dalla media possono essere determinate anche da fattori diversi dall’inefficienza (ad esempio, variabili omesse, cattiva specificazione, errori di misurazione).

3.2 *Approcci di frontiera*

La funzione di produzione fornisce il massimo *output* producibile dati gli *input*, mentre la funzione di costo indica il minor costo sostenibile per la produzione di un certo livello di *output*, dati i prezzi degli *input*. Entrambe le funzioni condividono la caratteristica di frontiera, massima e minima rispettivamente, sulla quale le imprese pienamente efficienti dovrebbero collocarsi, dato lo stato della tecnologia.

Sono stati proposti numerosi metodi di misurazione dell’efficienza basati sulla nozione di frontiera (si vedano Førsund, Lovell e Schmidt (1980) e Greene (1997) per una rassegna). Essi differiscono per vari aspetti. Uno di questi riguarda l’interpretazione della nozione di frontiera: alcuni metodi si propongono di stimare la frontiera “assoluta”, altri quella di *best practice*, che riflette il migliore risultato conseguito tra le imprese incluse nel campione considerato. In pratica, come è stato notato da Førsund e altri (1980), la differenza non dovrebbe essere rilevante: le due frontiere tendono a coincidere in campioni molto grandi.

Una distinzione importante tra le tecniche proposte in letteratura è quella tra tecniche parametriche e tecniche non parametriche. Le prime specificano a priori una forma funzionale con parametri costanti, che vengono stimati con metodi statistico-econometrici in modo tale che il grafico della funzione rappresenti nel modo migliore i dati. Quindi per ciascuna osservazione, la misura dell’efficienza in termini sia di *input*, sia di *output*, è calcolata con riferimento alla funzione stimata, che rappresenta la frontiera della tecnica produttiva in esame. Ovviamente, la misura così ottenuta dipende dalla forma funzionale specificata a priori.

Nelle tecniche non parametriche ciò che viene specificato a priori non è un’esplicita forma funzionale, ma alcune proprietà che i punti appartenenti alla frontiera devono soddisfare. Ipotesi in questo caso implicita è che una tale frontiera esista. I dati vengono quindi valutati

determinando, con tecniche di programmazione matematica, se ciascuna osservazione può essere considerata un elemento della frontiera oppure no, date le ipotesi effettuate. La misura di efficienza ottenuta per ciascuna osservazione dipenderà, in questo caso, dalle proprietà specificate a priori per la tecnologia produttiva.

3.2.1 Il metodo non parametrico

Questo metodo, conosciuto anche come “data envelopment analysis” (DEA), è stato ideato da Farrell (1957). Si consideri un’impresa che usa due *input*, x_1 e x_2 , e produce un *output* y ; si assuma che la funzione di produzione, $y = f(x_1, x_2)$, sia caratterizzata da rendimenti di scala costanti. Essa può pertanto essere scritta come:

$$1 = f\left(\frac{x_1}{y}, \frac{x_2}{y}\right) \quad (7)$$

La frontiera tecnologica può essere rappresentata dall’isoquanto unitario. L’approccio di Farrell è non parametrico nel senso che prevede di costruire, mediante tecniche di programmazione matematica, l’insieme dei rapporti *input/output* osservati nei dati che soddisfa le condizioni di convessità e eliminazione senza costo (*free disposal*)²; ovviamente solo un sottoinsieme del campione soddisferà tali condizioni, mentre la parte rimanente delle osservazioni si collocherà al di sopra dell’isoquanto.

L’approccio può essere illustrato con l’aiuto della Figura 4, dove sono illustrate sei imprese (da A a F) che producono un’unità di *output* y . In assenza di informazioni sulla posizione esatta dell’isoquanto unitario, non è possibile inferire che B, C e D operino in maniera inefficiente: rispetto a C, B utilizza una quantità maggiore di x_1 , ma una quantità inferiore di x_2 ; analogamente, sempre rispetto a C, D utilizza una quantità maggiore di x_2 , ma una inferiore di x_1 . È possibile invece definire la

² La condizione di eliminazione senza costo per essere soddisfatta richiede che dato un vettore *input-output* $z = (x_1, \dots, x_n, y_1, \dots, y_m)$ appartenente all’insieme di produzione Y , anche $z' \leq z$ appartenga all’insieme di produzione Y , ossia che per l’unità produttiva sia possibile eliminare quantità di *input* e/o *output* che non servono senza dover incorrere in un costo. Nel caso di tecnologia monoprodotta la condizione di eliminazione senza costo implica che la funzione di produzione sia non decrescente, e che quindi sia sempre possibile produrre almeno la stessa quantità aumentando la disponibilità di *input*.

tecnologia adottata dall'impresa E inefficiente: essa infatti utilizza rispetto a C una quantità superiore di entrambi gli *input*. L'ipotesi di convessità implica che, se due diverse combinazioni di *input* sono in grado di produrre un'unità di *output*, allora anche una loro combinazione lineare può produrre un'unità di *output*.

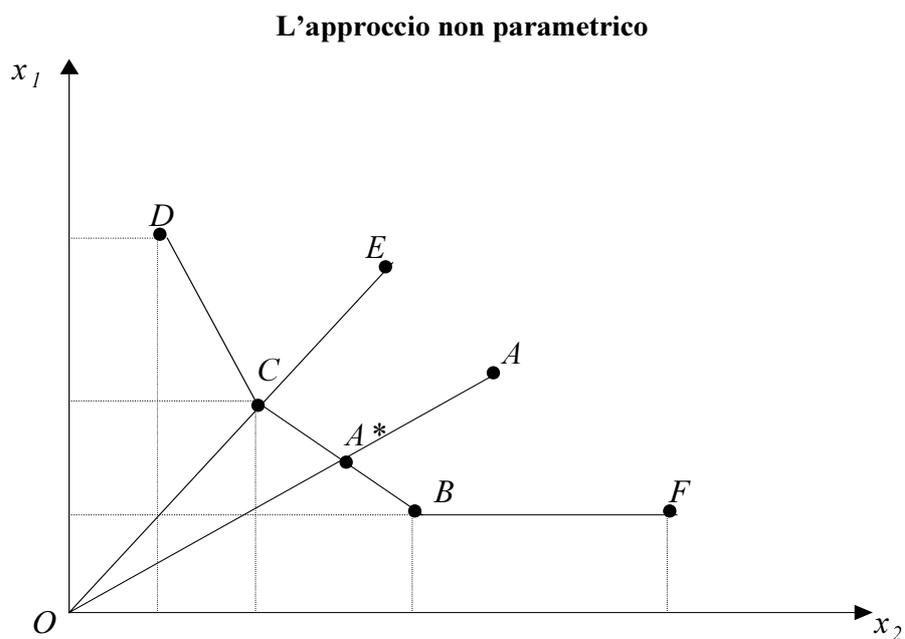
Nella Figura 4, ciò significa che un'impresa può operare in corrispondenza dei punti C o D, o di un qualsiasi punto sul segmento CD. In tal modo, è possibile distinguere combinazioni di *input* efficienti da quelle inefficienti: se il segmento che congiunge due combinazioni adiacenti di *input* ha una pendenza non positiva e nessun'altra combinazione appartenente all'isoquanto giace tra il segmento e l'origine, allora le due combinazioni di *input* selezionate sono efficienti; altrimenti, esse sono definite inefficienti. I segmenti che congiungono tutte le combinazioni di *input* efficienti danno luogo all'isoquanto efficiente. L'efficienza tecnica di un'impresa "dominata" da altre due imprese (ad esempio, A) è misurata confrontando il suo utilizzo di *input* con quello dell'impresa teorica, collocata sull'isoquanto efficiente, che usa gli *input* nella stessa proporzione (l'impresa A*, la cui efficienza tecnica è pari a OA^*/OA). L'efficienza tecnica di un'impresa "dominata" soltanto da un'altra impresa (ad esempio, F) è misurata con riferimento all'impresa che si colloca all'inizio della porzione piatta dell'isoquanto (B). Sebbene sia B sia F vengano definite tecnicamente efficienti, F utilizza una quantità di x_2 superiore a quella necessaria (e si dice essere "debole" nell'uso del fattore x_2).

La tecnica non parametrica DEA può essere utilizzata anche in caso di produzioni multiple e per misurare l'efficienza allocativa: tra i risultati forniti dai metodi di programmazione lineare utilizzati per calcolare l'efficienza tecnica vi è anche il saggio marginale di sostituzione tra i due *input* nella porzione rilevante dell'isoquanto efficiente. Confrontando tale valore con il rapporto tra i prezzi dei fattori è possibile determinare se l'impresa operi in condizioni di efficienza allocativa³.

L'utilizzo di questo approccio fornisce dunque una valutazione di efficienza relativamente alla miglior *performance* osservabile per le unità produttive comprese nel campione.

³ Si veda Charnes, Cooper e Rhodes (1981).

Figura 4



3.2.2 L'analisi parametrica delle frontiere

Questa analisi può essere condotta seguendo due approcci: la stima di frontiere deterministiche e quella di frontiere stocastiche. Il primo attribuisce le deviazioni rispetto alla frontiera unicamente alle scelte del produttore. Il secondo riconosce che tali deviazioni possono essere giustificate anche da circostanze al di fuori del controllo del produttore. Nel primo caso tutte le deviazioni dalla frontiera vengono imputate a inefficienza; nel secondo caso un errore simmetrico viene aggiunto alla funzione da stimare al fine di incorporare fattori attribuibili sia agli stati di natura, sia ad altre circostanze (come omissione di variabili, errori di misurazione, ecc.)⁴.

⁴ Nell'analisi di frontiere deterministiche il problema delle deviazioni di segno "sbagliato" rispetto alla teorizzazione economica – che prevede che dalla stima di funzioni di produzione o di costo emergano solo residui negativi nel primo caso e positivi nel secondo, data l'idea di massimo e minimo associata alle due funzioni – viene risolto correggendo l'intercetta stimata della funzione, che viene spostata fino a che vengono eliminate le deviazioni positive (negative) dalla funzione di produzione (costo). Se si assume invece che la frontiera sia di tipo stocastico il problema non si (continua)

3.2.2.1 Stima di frontiere deterministiche

Il primo modello di stima della frontiera basato su tecniche parametriche è stato il “deterministic frontier model” (DFM) di Aigner e Chu (1968). Essi hanno specificato una frontiera di produzione omogenea, di tipo Cobb-Douglas, con il vincolo che tutte le osservazioni si collochino in corrispondenza della frontiera, o sotto di essa:

$$\ln y_i = \underbrace{\ln f(x_i)}_{\substack{\text{frontiera di} \\ \text{prod. deterministica}}} + u_i = \alpha_0 + \sum_{j=1}^n \alpha_j \ln x_{ij} + u_i \quad u_i \leq 0 \quad (8)$$

dove l'indice i identifica l'impresa e l'indice j il fattore di produzione. Il requisito imposto sull'errore fa sì che $y_i \leq f(x_i)$ ⁵.

Nel caso della frontiera di costo, il residuo deve essere non negativo e implica che l'impresa possa operare in corrispondenza della frontiera, o sopra di essa.

In entrambi i casi la misura dell'inefficienza è data dal residuo, \hat{u}_i : nel caso della funzione di produzione, \hat{u}_i rappresenta l'inefficienza tecnica; nel caso della funzione di costo, \hat{u}_i rappresenta l'inefficienza complessiva.

Il DFM può essere stimato usando diversi metodi; il più semplice tra questi è il metodo dei minimi quadrati ordinari corretti (COLS)⁶. Si supponga che la funzione di produzione sia lineare, del tipo della (8), e si denoti con μ la media degli errori u_i ; si può quindi scrivere:

$$\ln y_i = (\alpha_0 + \mu) + \sum_{j=1}^n \alpha_j \ln x_{ij} + (u_i - \mu) \quad (9)$$

dove il nuovo termine di errore ha media pari a zero; esso soddisfa tutte le solite condizioni ideali, con l'eccezione della normalità. L'equazione (9)

pone: variabili al di fuori del controllo dell'impresa possono in effetti modificare la frontiera, consentendo risultati migliori o peggiori.

⁵ Si noti che, non essendo imposta la condizione $\sum_{j=1}^n \alpha_j = 1$ nella (8), l'analisi si applica anche a tecnologie che presentano rendimenti di scala non costanti.

⁶ Si veda Richmond (1974).

può pertanto essere stimata con i minimi quadrati ordinari per ottenere le migliori stime lineari non distorte della costante ($\alpha_0 + \mu$) e dei parametri α_j . Se si assume che gli u_i siano distribuiti secondo una funzione i cui parametri possono essere derivati dai suoi momenti centrali di ordine superiore, allora è possibile stimare tali parametri in maniera *consistent* dai momenti dei residui della regressione. Poiché μ è funzione di questi parametri, anche esso può essere stimato in maniera *consistent*; tale stima può essere utilizzata per correggere l'intercetta ottenuta con i minimi quadrati ordinari, che è una stima *consistent* di ($\alpha_0 + \mu$). Il metodo COLS fornisce pertanto stime *consistent* di tutti i parametri della frontiera.

Un problema di questo metodo è che, anche dopo aver corretto il termine costante, alcuni dei residui possono ancora presentare il segno "sbagliato", cosicché le osservazioni corrispondenti possono situarsi al di sopra della frontiera. Un modo di risolvere questo problema, oltre a ricorrere alla stima di frontiere stocastiche (si veda il paragrafo seguente), è quello di stimare la (9) con i minimi quadrati ordinari e poi correggere l'intercetta, non con il metodo appena descritto, ma semplicemente spostandola in alto, finché nessun residuo risulta positivo e uno è pari a zero. Gabrielson (1975) e Greene (1980) hanno dimostrato che questa correzione fornisce una stima *consistent* di α_0 . Un altro problema della tecnica COLS è che la correzione del termine costante non è indipendente dalla distribuzione ipotizzata per μ .

3.2.2.2 Stima di frontiere stocastiche

La stima di frontiere stocastiche è caratterizzata dalla possibilità che deviazioni rispetto alla frontiera efficiente siano dovute non solo a scelte errate da parte dell'impresa, ma anche a fattori casuali non controllabili. Nello "stochastic frontier model" (SFM) di Aigner, Lovell e Schmidt (1977) e Meeusen e van den Broeck (1977) il termine di errore è scomposto in due parti. Una componente simmetrica consente variazioni *random* della frontiera tra imprese diverse e cattura gli effetti di errori di misurazione, altri disturbi di natura statistica e *shock* esogeni; una componente, con segno vincolato, cattura gli effetti dell'inefficienza.

Un modello di frontiera della produzione stocastica può essere scritto come:

$$\ln y_i = \underbrace{\ln f(x_i)}_{\substack{\text{parte deterministica} \\ \text{frontiera di prod. stocastica}}} + v_i + u_i \quad u_i \leq 0 \quad (10)$$

La frontiera di produzione stocastica è data da $\ln y_i = \ln f(x_i) + v_i$, dove v_i ha una qualche distribuzione simmetrica e cattura gli effetti *random* di errori di misurazione e di *shock* esogeni, che fanno sì che la posizione delle imprese vari rispetto alla parte deterministica della frontiera di produzione, $f(x_i)$.

L'inefficienza tecnica rispetto alla frontiera stocastica è misurata da u_i ; la condizione $u_i \leq 0$ assicura che tutte le osservazioni si collochino in corrispondenza della frontiera stocastica o sotto di essa (nel caso della frontiera di costo, la condizione sul residuo diventa $u_i \geq 0$).

Purtroppo, non c'è un modo diretto e immediato di determinare se un'eventuale divergenza della posizione di un'impresa dalla componente deterministica della frontiera sia dovuta a inefficienza o a una variazione *random* della frontiera stessa. Ciò rappresenta il principale problema dello SFM: esso non consente di scomporre gli errori individuali nelle due determinanti e quindi di stimare l'inefficienza tecnica di ciascuna osservazione. Il risultato migliore che il metodo può fornire è una stima dell'inefficienza media nel campione.

L'approccio più utilizzato per separare le due componenti è l'analisi dei residui per evidenziare eventuali asimmetrie nella loro distribuzione. Se infatti si assume che la componente v dell'errore sia normalmente distribuita, il fatto che l'inefficienza riduca il prodotto (nel caso della funzione di produzione) o aumenti i costi (nel caso della funzione di costo) implica che l'errore complessivo abbia una distribuzione asimmetrica. Ipotizzando una distribuzione per u ⁷, la componente dell'errore che cattura l'inefficienza, è possibile stimare il modello con il metodo COLS, aggiungendo alle informazioni normalmente utilizzate nella stima di un modello di regressione quella riguardante il grado di asimmetria dei residui (Schmidt e Lovell, 1979). Stime dello SFM possono essere ottenute altrimenti con il metodo della massima verosimiglianza (Greene, 1980 e 1982). Qualunque sia il metodo utilizzato, il risultato fornisce una stima di

⁷ L'ipotesi di una distribuzione *half-normal*, adottata da Aigner, Lovell e Schmidt (1977) e Meeusen e van den Broeck (1977), è quella più comune negli studi applicati.

un residuo per ciascuna impresa, una stima della media degli u_i , ma non quella di ciascun u_i ; è possibile tuttavia stimare $E(u_i | u_i + v_i)$, ossia il valore atteso di u_i , dato il valore dell'errore complessivo (si veda Jondrow, Lovell, Materov e Schmidt, 1982). L'utilizzo di dati *panel* può rappresentare un utile strumento per stimare separatamente le due componenti; ciò è possibile quando la componente specifica di ciascuna unità produttiva è costante nel tempo e cattura unicamente il grado di inefficienza.

Il metodo COLS è analiticamente più semplice di quello della massima verosimiglianza ma produce stime asintoticamente meno efficienti. Olson, Schmidt e Waldman (1980) presentano un'analisi Monte Carlo che dimostra che con il metodo COLS si ottengono risultati qualitativamente analoghi a quelli forniti dal metodo della massima verosimiglianza, anche in campioni piuttosto grandi.

3.3 *Una valutazione comparativa*

In letteratura non è emersa una preferenza chiara per una delle tecniche esaminate.

Il difetto principale dell'approccio deterministico sta nel fatto che esso ignora la possibilità che il risultato di una unità produttiva sia influenzato, oltre che da variabili da essa controllate, da *shock* esogeni e da disturbi di tipo statistico (dovuti a errori di misurazione o cattiva specificazione del modello). Peraltro le tecniche stocastiche non offrono una soluzione soddisfacente a tale problema.

Il vantaggio principale dell'approccio non parametrico deriva dal fatto che nessuna forma funzionale è imposta sui dati, mentre per l'approccio parametrico è cruciale la corretta specificazione della funzione di produzione.

Le tecniche non di frontiera non consentono di misurare il livello di inefficienza, ma solo di valutare la posizione relativa delle unità produttive comprese nel campione. Peraltro, nelle tecniche di frontiera la stima di quest'ultima è basata su un sottoinsieme di osservazioni e può risentire in misura significativa della presenza di *outliers*.

Una breve rassegna degli studi settoriali effettuati sulla base di alcuni degli approcci presentati – senza alcuna pretesa di esaustività – può essere utile per individuare sia le difficoltà specifiche delle varie tecniche, sia quelle che si incontrano indipendentemente dall’approccio metodologico prescelto.

3.4 *Alcune applicazioni*

Le tecniche descritte sopra sono state impiegate per valutare l’efficienza tecnica, quella allocativa e quella complessiva in numerose applicazioni su una pluralità di settori. Una breve panoramica su alcuni degli studi effettuati consente di mettere in rilievo le differenze fra i risultati ottenuti utilizzando i diversi approcci e i problemi di valutazione delle conclusioni che possono essere tratte dall’analisi empirica.

3.4.1 *Le tecniche di regressione standard*

Per quanto riguarda l’efficienza tecnica, Feldstein (1967) ha analizzato 177 ospedali britannici utilizzando come misura dell’*output* la somma ponderata dei casi trattati, classificati per categorie, con pesi pari al costo medio di ciascuna categoria stimato sulla base di una funzione di costo. Come è stato sottolineato più volte in precedenza, la definizione stessa di *output* da utilizzare nell’analisi empirica costituisce un passaggio delicato nel caso in cui le tecniche di valutazione dell’efficienza vengano applicate a servizi pubblici complessi come quello sanitario.

Levitt e Joyce (1987) hanno utilizzato lo stesso approccio per stimare l’efficienza (tecnica) delle forze di polizia. Essi usano una misura del prodotto non ponderata (la percentuale di crimini risolti) e stimano un modello di equazioni simultanee che mette in relazione tale misura con una serie di variabili di *input*, quali il numero di agenti di polizia per abitante e le risorse per agente. Per tenere conto della natura disomogenea dei reati trattati e delle “condizioni ambientali” Levitt e Joyce utilizzano un indice della casistica dei crimini commessi in ciascuna area e alcune variabili socio-economiche (ad esempio, la densità di popolazione).

Tecniche di regressione standard sono state usate anche per il settore dell’istruzione, oltre che dagli stessi Levitt e Joyce (1987), dal Department of Education and Science, DES (1983 e 1984). In tutti e tre gli studi l’*output* è misurato in maniera non ponderata. Le misure sono connesse con

i risultati conseguiti negli esami. Tra gli *input* sono incluse sia le spese per studente direttamente connesse con l'attività di insegnamento, sia quelle non legate all'insegnamento; sono inoltre considerate altre variabili, quali il tasso di ricambio degli insegnanti e la fornitura di servizi scolastici aggiuntivi. Per tenere conto della disomogeneità dei casi trattati, tra le variabili esplicative è incluso il ceto sociale di provenienza degli studenti.

Da questi lavori emergono con chiarezza alcuni aspetti particolarmente problematici. Fra questi, la capacità delle variabili utilizzate di catturare l'inevitabile disomogeneità degli utenti del servizio: i malati, gli studenti, le vittime dei crimini. Levitt e Joyce (1987), ad esempio, confrontano la graduatoria delle unità di polizia ottenuta sulla base del semplice tasso di risoluzione dei casi con quella basata sui residui del modello di regressione da essi stimato ed evidenziano la difformità dei risultati ottenuti con i due metodi, specialmente alle estremità della graduatoria: delle unità classificate tra le prime otto sulla base del tasso di risoluzione dei casi, solo due risultano nella stessa posizione con il modello di regressione.

Si tratta di un problema di non facile soluzione. La scelta operata da Feldstein (1967) di ponderare i trattamenti in base al loro costo determina una contaminazione tra misura di prodotto e misura di costo che potrebbe distorcere le stime dell'efficienza. Per gli studi sulle forze di polizia e sull'istruzione di Levitt e Joyce (1987) e del DES (1983 e 1984), l'utilizzo di una misura di prodotto non ponderata sposta il problema sulle variabili esplicative: in entrambi i casi le variabili indipendenti utilizzate, l'indice di casistica per le forze di polizia e il ceto sociale di provenienza per gli studenti, non sono soddisfacenti; nel primo caso, con l'utilizzo di un indice sintetico, che pure consente in linea di principio di tenere conto di molte cause di variabilità, si corre il rischio di "nascondere" variabili significative dentro una media che non mostra alcuna correlazione con costi e prodotto; nel secondo, si utilizza uno solo degli elementi potenzialmente rilevanti (il ceto sociale; mentre, ad esempio, occorrerebbe tenere conto anche del curriculum scolastico precedente).

Feldstein (1967) presenta anche valutazioni relative all'efficienza allocativa e conclude che negli ospedali britannici da lui esaminati "si spende eccessivamente per infermieri, vitto e per altre forniture e non sufficientemente per medici, farmaci e abbigliamento". Lavers e Whynes (1978) considerano gli ospedali per maternità inglesi. Essi utilizzano sia funzioni di tipo Cobb-Douglas, sia forme funzionali meno restrittive (translogaritmiche) e concludono invece che la spesa per personale medico

è eccessiva rispetto a quella per il personale infermieristico, per i farmaci e per l'abbigliamento. McGuire (1987), stimando una funzione translogaritmica su un campione di ospedali scozzesi, conclude che "si spende troppo per il personale paramedico e per il personale addetto alle pulizie e poco per quello medico e infermieristico".

I risultati di questi studi evidenziano la sensibilità dei risultati alla specificazione delle funzioni di produzione e alla definizione del settore di studio.

Invece della stima della funzione di produzione, la stima di funzioni di costo per la valutazione dell'efficienza complessiva delle unità di produzione è stata impiegata, oltre che da Feldstein (1967) per la sanità, tra gli altri anche da Domberger, Meadowcroft e Thompson (1986), per il settore della raccolta dei rifiuti.

3.4.2 *La Data Envelopment Analysis*

Oltre alla regressione standard Levitt e Joyce (1987) hanno utilizzato anche la tecnica DEA per analizzare l'efficienza dei servizi di polizia; Jesson, Mayston e Smith (1987) l'hanno applicata all'istruzione secondaria; Cubbin, Domberger e Meadowcroft (1987) al servizio di raccolta dei rifiuti.

L'analisi di Levitt e Joyce (1987) è particolarmente interessante perché consente di confrontare i risultati ottenuti con questo metodo con quelli derivati con l'applicazione di tecniche di regressione utilizzando lo stesso *data set*: le graduatorie delle unità ottenute con le due tecniche presentano un notevole grado di sovrapposizione, ma con alcune importanti eccezioni. In particolare, delle quattro unità giudicate più inefficienti sulla base della tecnica della regressione, tre hanno ricevuto un punteggio per il grado di efficienza pari al 100 per cento utilizzando la DEA.

L'analisi di Jesson e altri (1987) per l'istruzione secondaria utilizza due definizioni di *output*: 1) la percentuale di studenti che ottengono cinque o più sufficienze al livello "O"; 2) la percentuale di studenti che ottengono tre o più sufficienze al livello "CSE" od "O"⁸. L'*input* è dato

⁸ Nel sistema scolastico britannico gli studenti intorno ai 15-16 anni di età sostengono esami su un certo numero di materie al livello *ordinary* ("O"); a tale livello ne segue uno *advanced* che si ottiene su base volontaria e, generalmente, su un numero più limitato di materie. Il *Certificate of* (continua)

dalla spesa per l'istruzione secondaria per studente. Inoltre sono considerate variabili ambientali, quali la percentuale di studenti il cui capo famiglia è un lavoratore specializzato, la percentuale di studenti appartenenti a famiglie mono-parentali, la percentuale di studenti nati nel Regno Unito, in Irlanda, negli Stati Uniti o nel Commonwealth.

L'analisi mette in luce come i risultati varino al variare della misura di *output* utilizzata e siano molto sensibili alla presenza di osservazioni estreme.

Inoltre il confronto tra i risultati di Jesson e altri (1987) e quelli ottenuti in DES (1983 e 1984) con tecniche di regressione confermano la variabilità dei risultati nelle code delle graduatorie: delle otto scuole più efficienti nello studio del DES, sei ricevono un punteggio per l'efficienza pari al 100 per cento secondo il metodo DEA; le altre due ottengono un punteggio pari al 97 e 95 per cento e si collocano al 48esimo e 66esimo posto della graduatoria. Delle cinque scuole più inefficienti nello studio del DES, tre ottengono un punteggio basso secondo il metodo DEA, una ottiene un punteggio pari al 97 per cento (51esima in graduatoria), mentre l'ultima ottiene un punteggio per l'efficienza pari al 100 per cento.

Questi risultati suggeriscono l'opportunità di usare particolare cautela nell'impiego dei risultati di queste analisi dal momento che è proprio agli estremi delle graduatorie di efficienza che in linea di principio dovrebbero concentrarsi eventuali interventi correttivi.

3.4.3 *La stima delle frontiere*

Questa tecnica è particolarmente complessa e richiede basi informative ricche. Perciò il suo utilizzo è relativamente poco frequente.

Rubinstein, Schwartz e Stiefel (1998) considerano il settore dell'istruzione utilizzando dati a livello di singola scuola e mettendo a confronto tecniche diverse. La stima dell'efficienza viene effettuata utilizzando un campione panel di tre anni riguardante dati relativi a scuole della Georgia. Essi specificano una funzione di produzione in cui l'*output* è la percentuale di studenti che nei test standard ottiene un risultato superiore alla media. Dall'analisi dei risultati empirici è possibile concludere che la

Secondary Education ("CSE"), conseguito al termine del secondo ciclo di studi, è analogo al nostro diploma di maturità.

correlazione tra le graduatorie ottenute con i diversi metodi dipende significativamente dalla scelta dell'anno considerato, ma sembra essere meno sensibile alla specificazione adottata.

Grosskopf e altri (1998) analizzano il problema dell'inefficienza nella produzione scolastica con un approccio stocastico di frontiera sulla funzione di distanza (duale rispetto alla funzione di costo). Essi utilizzano come misura di *output* i risultati degli studenti in test standard, e come *input* misure relative alla quantità di personale docente e amministrativo, ad altri fattori di produzione, oltre che alle condizioni familiari e demografiche. L'attenzione in questo lavoro è concentrata sulla relazione fra livello di efficienza nella produzione e situazione competitiva nelle aree in cui operano i distretti scolastici. Per misurare il livello di competitività viene utilizzato un indice della concentrazione scolastica rispetto al mercato locale. I risultati evidenziano l'esistenza di una correlazione positiva tra l'inefficienza dei distretti scolastici e la mancanza di pressioni competitive nell'ottenere l'iscrizione degli studenti.

Tra gli studi più recenti e metodologicamente sofisticati che si sono occupati dell'efficienza nel settore dei trasporti, va segnalato quello di Parisio (1999) che analizza l'efficienza di otto compagnie ferroviarie europee nel periodo 1973-89 con un modello di frontiera di costo stocastica⁹. I risultati indicano la presenza di un livello significativo di inefficienza tecnica, mentre l'inefficienza allocativa appare contenuta entro livelli "fisiologici". Va rilevato che per le ferrovie italiane il costo indotto da inefficienza allocativa è più elevato di quello stimato per le altre compagnie e appare attribuibile ad un numero eccessivo di addetti. Al contrario, il costo indotto da inefficienza tecnica è minore in Italia che negli altri paesi. L'analisi tuttavia sembra trascurare aspetti di qualità del servizio quali la puntualità e la frequenza delle corse che sono invece aspetti rilevanti dell'efficienza del servizio percepita dagli utenti.

Un confronto fra le tecniche di valutazione dell'efficienza in campo sanitario si trova in Giuffrida e Gravelle (2001). In questo lavoro gli autori utilizzano dati relativi alle Family Health Service Authority (che forniscono le prestazioni di base nel sistema sanitario inglese) per confrontare i risultati che si ottengono dall'applicazione delle tecniche econometriche di stima delle frontiere deterministiche e stocastiche con

⁹ Si veda anche Bosco e Parisio (1996) per una analisi delle tecniche di stima dell'efficienza nella produzione pubblica, con particolare riferimento al settore dei trasporti su rotaia.

quelli ottenibili con il metodo DEA, non essendoci da un punto di vista teorico forti motivazioni per ritenere una metodologia strettamente superiore ad un'altra. I dati utilizzati si riferiscono agli anni finanziari 1993-94 e 1994-95. La stima è effettuata con tecniche diverse e riguarda sia frontiere di costo che frontiere di produzione. La sensibilità dei risultati viene valutata anche rispetto a specificazioni diverse della tecnologia. Il confronto fra le diverse metodologie e specificazioni considerate porta gli autori a concludere che l'applicazione delle varie tecniche genera graduatorie di efficienza delle unità produttive significativamente diverse.

4. Il caso della sanità

Il settore sanitario fornisce un esempio delle difficoltà che si incontrano nel valutare l'efficienza e l'efficacia nei servizi pubblici e nell'utilizzare tali valutazioni per indicare possibili linee di intervento. Si tratta di un caso particolarmente rilevante: negli ultimi decenni i tassi di crescita della spesa nel settore sanitario sono stati elevati; ciò ha reso più evidente la necessità di assicurare un uso efficiente delle risorse e di adottare assetti istituzionali in grado di garantire un maggiore controllo delle erogazioni.

Un confronto internazionale basato unicamente sulla spesa complessiva per la sanità non è particolarmente informativo. La quota del reddito nazionale spesa per la sanità tende a crescere con il grado di sviluppo: la correlazione tra la spesa sanitaria pro capite e il PIL pro capite appare rilevante (Figura 5). Tuttavia, una spesa più elevata non si traduce immediatamente in una maggiore "produzione" di servizi sanitari, né in un migliore risultato in termini di stato di salute e di attenzione alle esigenze dei pazienti. Secondo OCSE (1999), circa il 25 per cento dell'aumento della spesa pro capite registrato nei paesi dell'OCSE tra il 1980 e il 1996 è attribuibile ad un aumento dei prezzi relativi; occorrerebbe pertanto disporre di misure fisiche e non monetarie delle risorse destinate al servizio sanitario. Inoltre, a parità di servizio reso, la spesa può essere più elevata per un uso inefficiente delle risorse.

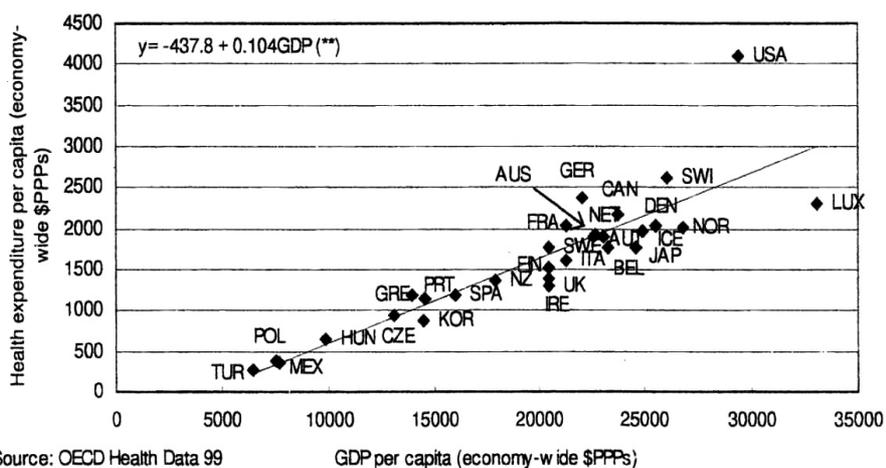
L'OCSE costituisce da tempo un punto di riferimento per l'analisi delle politiche e dei sistemi sanitari: ha condotto studi sulle cause dell'aumento della spesa e sulle riforme adottate nei paesi membri dell'organizzazione, ha raccolto informazioni su indicatori quantitativi di

costo e di produzione presso tali paesi e organizzato una banca dati per i confronti internazionali.

È significativo che in OCSE (1999), un rapporto preparato in risposta ad una richiesta dell'Organizzazione Mondiale per la Sanità (OMS) per una valutazione dei risultati ottenuti dai sistemi sanitari dei paesi membri dell'OCSE, si giunga alla conclusione che *"...the available evidence is neither plentiful enough nor reliable enough, yet, to allow us to draw firm conclusions about variations in performance in the term suggested or about their causes"*.

Figura 5

Chart 1: GDP per capita and health expenditure per capita (economy-wide \$PPPs), 1996



Fonte: OCSE (1999).

Come rilevato, il primo ostacolo da superare è la definizione di risultato. A tale riguardo, nella sua richiesta all'OCSE, l'OMS ha individuato quattro elementi principali:

- a) lo stato di salute della popolazione;
- b) l'attenzione alle esigenze degli assistiti;

- c) l'impatto economico;
- d) il coinvolgimento dei cittadini.

In linea di principio la definizione fornita dall'OMS "elimina" il problema dell'arbitrarietà della definizione degli obiettivi del servizio pubblico e consente di tentare una valutazione dell'efficacia in relazione ai suoi obiettivi.

Resta tuttavia il problema di individuare misure appropriate del grado di raggiungimento degli obiettivi. Da questo punto di vista la definizione dell'OMS è vaga, lontana dall'essere operativa: manca un sistema di ponderazione dei quattro elementi; non è chiaro cosa si intenda per impatto economico e come questo possa essere messo in relazione con le risorse utilizzate; non è immediatamente evidente come possa essere misurato il coinvolgimento dei cittadini.

Per tentare una applicazione quantitativa delle indicazioni fornite dall'OMS, OCSE (1999) ha proceduto come segue:

- a) ha tralasciato il grado di coinvolgimento dei cittadini;
- b) ha utilizzato la spesa sanitaria come misura dell'impatto economico;
- c) ha raccolto dati su alcuni indicatori dello stato di salute e dell'attenzione alle esigenze degli assistiti (misure di efficacia);
- d) ha utilizzato i rapporti tra i dati di cui al punto (c) e la spesa sanitaria come indicatori di efficienza (misurata quindi non in relazione ai servizi prodotti, ma ai risultati raggiunti)¹⁰.

Lo stato di salute della popolazione può essere misurato in diversi modi. Le misure più diffuse sono quelle basate su indici di mortalità (vita attesa a diverse età; mortalità infantile e perinatale; anni di vita potenziali persi prima dei 70 anni). I risultati non sono invarianti rispetto alla misura prescelta. Inoltre, lo stato di salute, comunque definito, dipende in misura significativa da una pluralità di fattori che sono, in buona parte, indipendenti dal servizio sanitario prestatato: il livello di istruzione (Auster e altri, 1969; Hadley, 1982); la distribuzione del reddito (Wilkinson, 1992; Kaplan e altri, 1996; Kennedy e altri, 1996); il grado di sviluppo (Elola e

¹⁰ Secondo l'OMS l'efficienza di un sistema sanitario andrebbe definita in base al rapporto tra la misurazione quantitativa dei risultati ottenuti e le risorse finanziarie destinate alla sanità.

altri, 1995); Frech e Miller, 1999); le abitudini alimentari (Grubalugh e Santerre, 1994; Frech e Miller, 1999)¹¹.

Non è facile individuare una misura quantitativa pienamente soddisfacente del grado di attenzione prestato alle esigenze degli assistiti. Rapidità del trattamento, informazioni fornite, condizioni generali in cui avviene la prestazione medica, continuità del rapporto con i medici sono tutte dimensioni rilevanti per il grado di soddisfazione degli utenti.

Una possibilità è costituita dall'utilizzo di "indagini di mercato". Tuttavia, tali indagini presentano diversi aspetti insoddisfacenti; in particolare: a) il campione di intervistati può includere persone che non usano i servizi sanitari da diversi anni; b) i giudizi possono essere formulati sulla base di notizie di stampa piuttosto che di esperienze dirette; c) nel confronto internazionale occorrerebbe tenere conto della possibile disomogeneità delle aspettative degli utenti di diversi paesi.

Una possibilità alternativa è quella di utilizzare informazioni sui tempi di attesa per l'accesso ai servizi. Anche in questo caso, tuttavia, la misurazione è meno semplice di quanto possa apparire. Con riferimento agli interventi di chirurgia, ad esempio, occorre decidere se misurare il tempo di attesa dal momento in cui la visita specialistica individua la necessità dell'intervento o da quello in cui il paziente viene posto formalmente in lista d'attesa; se includere o escludere i pazienti trattati in emergenza; se utilizzare il numero di pazienti in lista d'attesa o il tempo atteso in media dai pazienti effettivamente trattati in un dato periodo. I risultati possono variare in misura anche significativa in funzione delle scelte effettuate.

Le difficoltà aumentano se si vuole passare a misure di efficienza basate sulla relazione tra *input* e risultati, come nel caso di OCSE (1999).

Per quanto riguarda i risultati in termini di stato di salute, OCSE (1999) fa riferimento a Or (1997) che utilizza la banca dati dell'OCSE (21 paesi per il periodo 1970-92) e misura l'inefficienza tecnica con un approccio non di frontiera basato sulla stima della funzione di produzione, come in Feldstein (1967). Le variabili utilizzate sono:

- a) gli anni di vita potenziali persi prima dei 70 anni (AVPP) come misura inversa del "prodotto sanitario";

¹¹ Quest'ultima variabile dipende anche dall'efficacia della funzione di informazione e prevenzione svolta dal servizio sanitario.

- b) diverse misure di *input*, sia monetarie (la spesa sanitaria pro capite complessiva e la relativa quota pubblica; la spesa farmaceutica), sia fisiche (numero di medici e di infermieri; numero di ricoveri ospedalieri);
- c) diverse variabili di controllo (reddito pro capite; percentuale della popolazione con mansioni di tipo impiegatizio; inquinamento; consumi di tabacco, alcool, grassi animali e zuccheri).

I principali risultati sono così riassumibili:

- a) la spesa sanitaria è correlata negativamente con gli AVPP; la correlazione è significativa solo per le donne; l'elasticità è bassa (0,2);
- b) il numero di medici e infermieri, il numero di ricoveri ospedalieri, la quota pubblica della spesa sanitaria e la spesa farmaceutica sono correlati negativamente con gli AVPP; la correlazione è significativa per uomini e donne; l'elasticità è bassa (inferiore a 0,15);
- c) la percentuale della popolazione con mansioni impiegatizie ha una correlazione negativa, significativa, con gli AVPP, con un'elasticità elevata (compresa tra 0,5 e 0,9 a seconda delle equazioni stimate); risultati analoghi riguardano il reddito pro capite;
- d) inquinamento, consumo di tabacco, di alcool, di grassi animali e di zuccheri hanno una correlazione positiva (in genere significativa) con gli AVPP, sebbene con una bassa elasticità.

Anche in questo studio, come in quello di Feldstein (1967), la misura di prodotto è cruciale. Si tratta di una misura più sofisticata e più vicina al concetto di risultato (Feldstein misurava i casi trattati, senza riguardo all'esito del trattamento). Tuttavia, se nel caso di Feldstein era la ponderazione in base ai costi a "sporcare" la misura di prodotto, qui la misura è evidentemente teorica, basata su stime dell'efficacia dei trattamenti.

Relativamente all'attenzione per le esigenze degli utenti, OCSE (1999) fa riferimento ai dati sui tempi di attesa per il primo appuntamento con uno specialista riportati in Royal College of General Practitioners (1992)¹² (Figura 7).

OCSE (1999) utilizza le informazioni così raccolte per cercare di fornire una rappresentazione sintetica di misure di efficienza/efficacia

¹² Vengono riportati anche i risultati dell'indagine "Eurobarometro" (Mossialos, 1997) che sembrano indicare una correlazione positiva tra spesa e grado di soddisfazione (Figura 6).

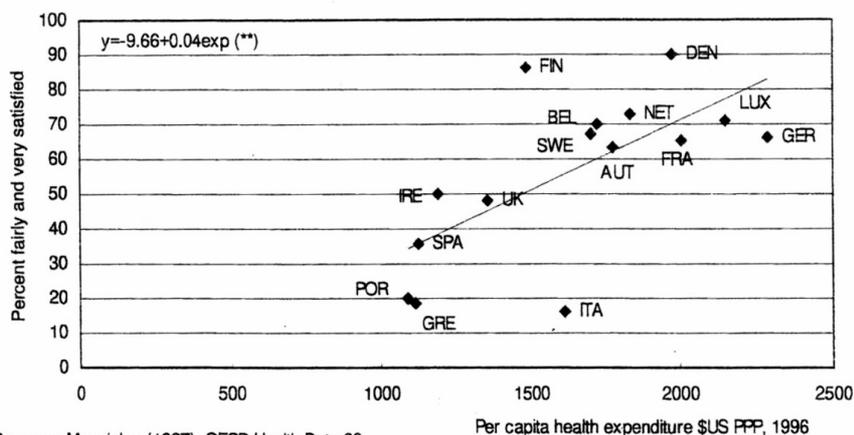
basate su stato di salute e attenzione alle esigenze degli utenti attraverso un “diagramma di possibilità di efficienza” (Figura 8). L’asse delle ordinate misura l’inverso rapporto tra il tempo di attesa per il primo appuntamento con uno specialista e la spesa sanitaria pro capite (valori più elevati indicano una performance migliore). L’asse delle ascisse misura il rapporto tra la riduzione attesa in termini di AVPP (stimata nello studio di Or, 1997) e la spesa pro capite (in definitiva, si ipotizza una eguale produttività dei medici nei vari paesi e la variabile segue l’andamento del rapporto tra numero di medici e spesa sanitaria).

Il diagramma fornisce deboli indizi che, a parità di condizioni:

- a) i paesi dove il numero di medici è più elevato hanno *performance* migliori in termini di tempi d’attesa e di riduzioni di AVPP (il numero di medici pro capite è illustrato in Figura 7; essa mostra peraltro una debole correlazione negativa tra i tempi d’attesa e il numero di medici pro capite)¹³;

Figura 6

Chart 8: Per capita expenditure on health in \$US PPP as compared to satisfaction with health care system, 1996

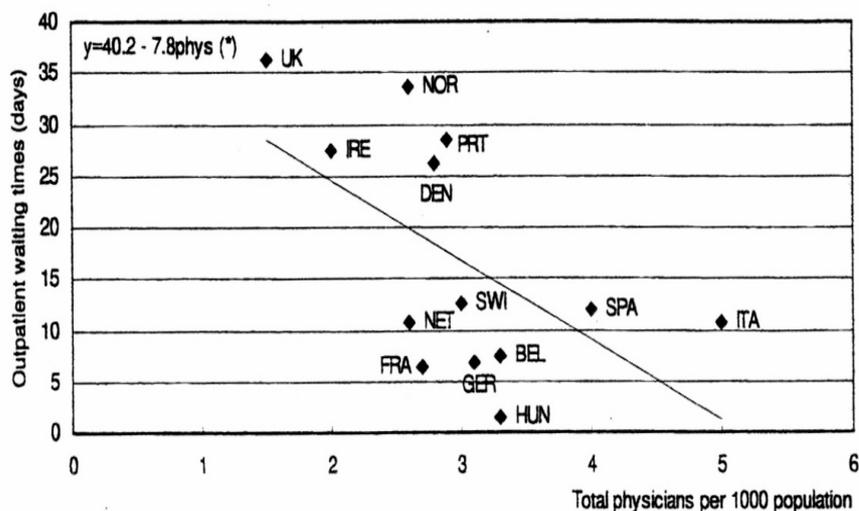


Fonte: OCSE (1999).

¹³ Si veda anche Blendon e altri (1995).

Figura 7

Chart 10: Outpatient waiting times and total physicians per 1000 population, 1992



Source: Royal College of General Practitioners, 1992 and OECD Health Data 99

Fonte: OCSE (1999).

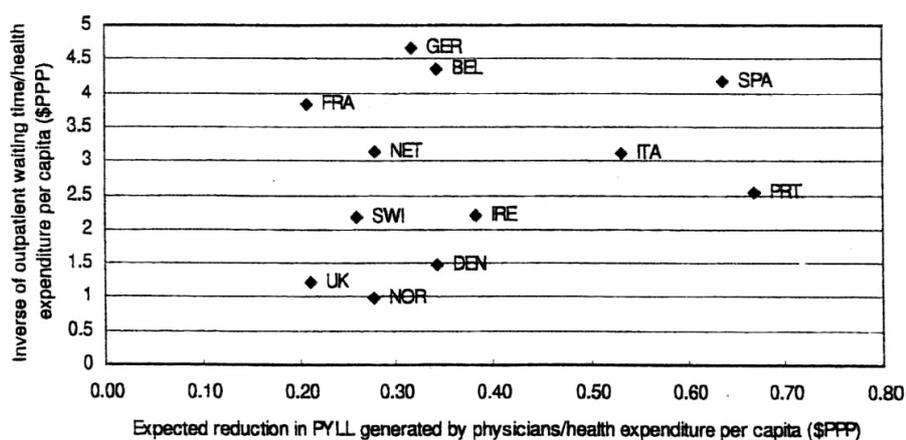
b) la remunerazione dei medici in base alle prestazioni implica una migliore prestazione in termini di tempi d'attesa rispetto ad una remunerazione di tipo salariale.

Infatti, Svizzera, Irlanda, Danimarca, Norvegia e Regno Unito, paesi con un numero relativamente basso di medici e in cui – con l'eccezione della Svizzera – i medici sono salariati, hanno risultati inferiori alla media in entrambe le dimensioni del diagramma. Spagna e Italia, che hanno un numero di medici relativamente elevato, hanno risultati superiori alla media in entrambe le dimensioni. Germania, Belgio, Francia e Olanda, che adottano remunerazioni dei medici basate sulle prestazioni, hanno una performance superiore alla media in termini di tempi di attesa.

Tuttavia, il carattere sperimentale dell'esercizio, le limitazioni in termini di disponibilità e comparabilità dei dati e il carattere restrittivo di

Figura 8

Chart 18: Tentative estimates of outpatient waiting efficiency variations per \$ and expected reduction in PYLL per \$



Source: Royal College of General Practitioners 1992, OECD Health Data 99

Fonte: OCSE (1999).

alcune delle ipotesi adottate¹⁴ suggeriscono di valutare con prudenza questi risultati; come rilevato in OCSE (1999; p. 13): “*it would be unwise to base policy on such a finding without further investigation*”.

Alla prudenza porta anche il confronto tra i risultati ottenuti con l'applicazione delle diverse tecniche di stima su di uno stesso insieme di dati. Si consideri infatti la discussione svolta nella sezione 3.3 (si veda, in particolare, Giuffrida e Gravelle, 2001).

5. L'efficienza dei servizi pubblici in Italia

5.1 Un quadro quantitativo

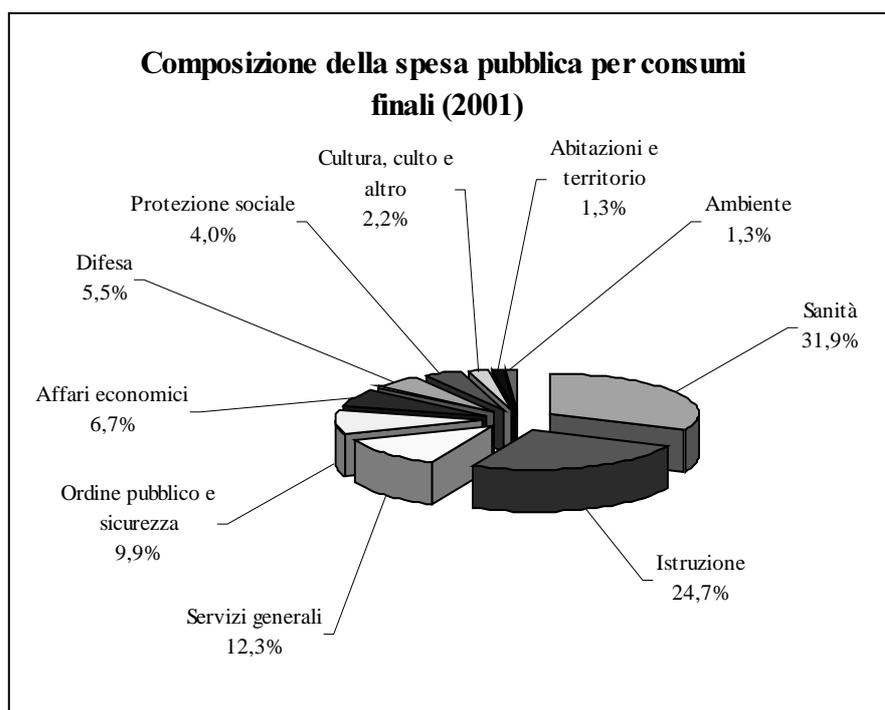
Nel 2001 le spese correnti primarie delle Amministrazioni pubbliche sono risultate pari al 37,6 per cento del prodotto (Tabella 2). Di queste,

¹⁴ Sui risultati può avere influito, in particolare, l'ipotesi di produttività eguale in tutti i paesi: alcuni paesi potrebbero aver raggiunto le zone della produzione di servizi sanitari dove i rendimenti marginali (in termini di riduzione di AVPP) sono decrescenti.

quasi il 51 per cento riguarda trasferimenti (il 44 per cento per prestazioni sociali in favore delle famiglie; il 7 per cento riguarda soprattutto le imprese); la parte rimanente, pari al 18,5 per cento del PIL, finanzia la produzione di servizi pubblici. Composizione e incidenza della spesa per consumi finali sul prodotto sono risultate sostanzialmente stabili nella seconda metà degli anni novanta e in crescita nell'ultimo biennio¹⁵.

Nell'ambito di tale spesa, sanità e istruzione sono i settori che assorbono la quota maggiore di risorse, pari rispettivamente al 5,9 e al 4,6 per cento del PIL nel 2001 (Figura 9 e Tabella 3). Spese dell'ordine del 2 per cento del PIL vengono sostenute per i servizi generali (il funzionamento dell'amministrazione in senso proprio) e per l'ordine pubblico e la

Figura 9



Fonte: ISTAT.

¹⁵ Dati elaborati in base al nuovo sistema europeo dei conti (SEC95).

Tabella 2

Spese correnti delle Amministrazioni Pubbliche

(milioni di euro, incidenza sulla spesa primaria, incidenza sul PIL)

	1995		1996		1997		1998		1999		2000		2001	
	milioni di euro	% spesa primaria												
Consumi finali	164,819	48,3	177,568	48,1	186,426	48,0	192,223	47,9	199,692	47,8	212,187	48,6	224,663	49,2
di cui:														
Redditi da lavoro dipendente	103,560	30,3	112,876	30,6	119,109	30,6	114,450	28,5	117,739	28,2	122,810	28,2	129,028	28,2
Prestazioni sociali in denaro	154,293	45,2	165,810	44,8	177,771	45,7	181,894	45,3	190,000	45,5	195,344	44,8	202,728	44,4
Altro	22,364	6,5	26,198	7,1	24,452	6,3	27,273	6,8	27,763	6,7	28,636	6,6	29,701	6,5
Spese correnti primarie	341,475	37,0	369,335	37,7	388,649	37,7	401,390	37,4	417,455	37,4	436,167	37,7	457,092	37,6
Interessi	106,385	11,5	112,950	11,5	96,105	9,4	86,011	8,0	74,834	6,8	75,265	6,8	77,133	6,3
Totale spese correnti	447,861	48,5	482,285	49,2	494,755	49,2	487,401	45,4	497,289	44,4	511,432	44,4	534,225	43,9

Fonte: Elaborazioni da Relazione generale sulla situazione economica del Paese e ISTAT.

Tabella 3

Composizione della spesa pubblica per consumi finali
(milioni di euro, valori percentuali, incidenza sul PIL)

	1995		1996		1997		1998		1999		2000		2001		
	milioni di euro	%	milioni di euro	%											
Sonito'	47,431	28,778	51,099	28,777	55,210	29,615	57,137	22,913	59,972	30,032	66,231	31,214	71,776	31,9	5,9
Istruzione	41,952	25,5	45,614	25,7	47,695	25,6	49,167	19,7	50,579	25,3	52,904	24,9	55,467	24,7	4,6
Servizi generali	20,946	12,7	22,537	12,7	23,005	12,3	23,618	9,5	24,252	12,1	26,131	12,3	27,709	12,3	2,3
Ordine pubblico e sicurezza	18,279	11,1	20,362	11,5	20,916	11,2	21,523	8,6	21,846	10,9	21,966	10,4	22,298	9,9	1,8
Affari economici	12,435	7,5	13,122	7,4	13,784	7,4	14,026	5,6	14,160	7,1	14,433	6,8	15,048	6,7	1,2
Difesa	10,378	6,3	10,640	6,0	10,476	5,6	10,732	4,3	11,751	5,9	11,995	5,7	12,417	5,5	1,0
Protezione sociale	6,349	3,9	6,745	3,8	6,906	3,7	7,069	2,8	7,432	3,7	8,273	3,9	9,086	4,0	0,7
Cultura, culto e altro	3,791	2,3	4,146	2,3	4,562	2,4	4,548	1,8	4,560	2,3	4,684	2,2	4,980	2,2	0,4
Abitazioni e territorio	2,183	1,3	2,125	1,2	2,371	1,3	2,470	1,0	2,752	1,4	2,837	1,3	3,002	1,3	0,2
Ambiente	1,074	0,7	1,178	0,7	1,501	0,8	1,933	0,8	2,388	1,2	2,733	1,3	2,880	1,3	0,2
Totale	164,819	100,0	177,568	100,0	186,426	100,0	249,300	100,0	199,692	100,0	212,187	100,0	224,663	100,0	16,5

Fonte: ISTAT.

sicurezza (1,8). Agli affari economici e alla difesa viene destinato, rispettivamente, l'1,2 e l'1,0 per cento del prodotto. Gli altri servizi assorbono finanziamenti di entità inferiore al punto percentuale del prodotto.

I servizi di trasporto pubblico sono inclusi tra gli affari economici. Va tuttavia rilevato che il finanziamento dei servizi di trasporto avviene soprattutto attraverso trasferimenti in favore di imprese pubbliche (le Ferrovie, ma anche le aziende municipali o regionali) che non rientrano tra i consumi finali delle Amministrazioni pubbliche e sono invece compresi sia nella voce "altro" della tabella 2, sia tra i trasferimenti in conto capitale. Nel 2001 i trasferimenti da parte dell'amministrazione centrale e delle amministrazioni locali per il finanziamento del trasporto pubblico su strada e rotaia sono stati pari a circa 7 miliardi di euro (0,6 per cento del PIL).

Dal punto di vista della quantità di risorse loro destinate, quindi, i settori più rilevanti della spesa pubblica per servizi sono, nell'ordine: la sanità, l'istruzione, l'amministrazione generale, l'ordine pubblico e i trasporti. In quanto segue si fornisce una breve rassegna di alcuni studi effettuati sull'efficienza dei servizi pubblici in Italia nei settori della sanità, dell'istruzione e dei trasporti.

5.2 *La sanità*

Una prima misura di efficienza dei servizi sanitari italiani è offerta in Formez (1987), dove si valutano i carichi di lavoro per dipendente degli ospedali pubblici come rapporto tra il numero di giornate di degenza e il numero dei dipendenti (inclusi gli amministrativi). L'analisi evidenzia una perdita di efficienza così misurata pari al 40 per cento tra il 1970 e il 1978. I limiti dell'indicatore sono evidenti: si trattano le giornate di degenza allo stesso modo, a prescindere dalla patologia che le determina; si identifica una riduzione delle giornate di degenza come una perdita di efficienza quando in realtà essa potrebbe indicare una maggiore efficacia delle cure prestate¹⁶.

La maggior parte degli studi sull'efficienza dei servizi sanitari italiani utilizza indicatori simili, più o meno complessi, ma non si basa sulla nozione di distanza da una frontiera efficiente (Barbetta e Turati,

¹⁶ Si vedano, ad esempio, Franco (1993) e i riferimenti ivi indicati.

2000). Solo negli anni più recenti hanno trovato applicazione la DEA e la stima di funzioni di produzione o di costo.

Cellini e altri (2000) misurano l'efficienza della produzione ospedaliera italiana utilizzando la DEA e ne analizzano le determinanti attraverso l'analisi di regressione. I dati si riferiscono al 1996 e riguardano circa i due terzi degli ospedali pubblici e privati accreditati ad operare nell'ambito del servizio sanitario nazionale (1.183 ospedali su 1.789).

Gli autori usano diverse misure composte di *output*: tutte le misure includono il numero di casi di day-hospital affiancato, per i ricoveri ordinari, dal numero di pazienti dimessi e dai giorni di degenza. L'aggregazione di questi elementi in una misura composta viene effettuata utilizzando come pesi i costi standard di trattamento delle varie patologie (classificate secondo il sistema dei *diagnostic related groups*, DRG) stabiliti nell'ambito del Servizio Sanitario Nazionale per il finanziamento dell'attività ospedaliera. Gli *input* considerati sono: la numerosità del personale medico, paramedico e amministrativo e il numero di letti disponibili nell'ospedale.

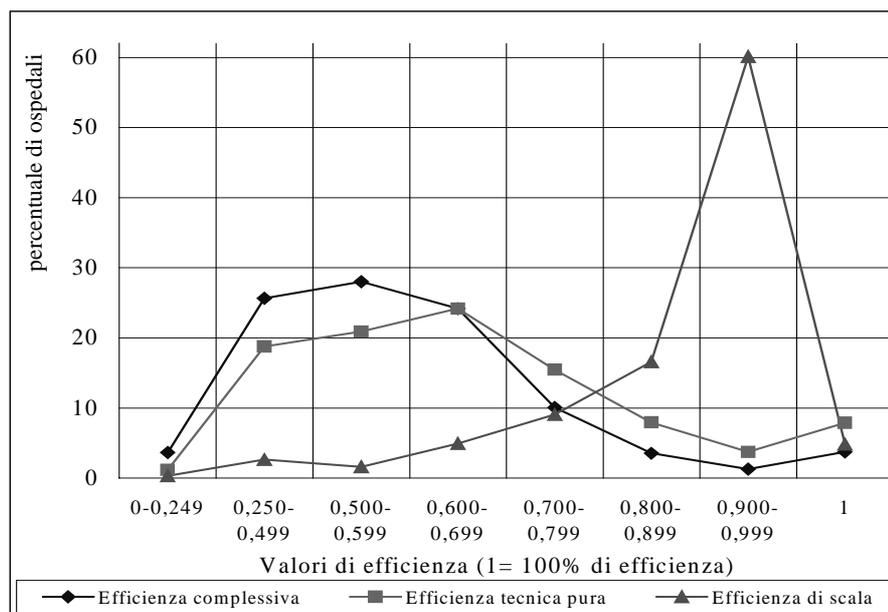
I risultati indicano la presenza di un elevato grado di inefficienza: solo 44 ospedali (il 3,72 per cento del campione) si situano sulla frontiera di efficienza stimata; fatto pari a 100 l'*input* necessario per una data quantità di prodotto, in media gli ospedali inclusi nel campione ne utilizzano una quantità pari a 142,5. La distribuzione dei valori di efficienza riscontrata per i 1.183 ospedali considerati è riportata in Figura 10. Le stime mostrano inoltre che l'inefficienza è legata solo in minima parte a diseconomie di scala (11 punti sui 42,5 totali, dovuti quasi in egual misura a strutture sotto e sovradimensionate).

Gli autori mostrano come il grado di efficienza non sia collegato alla proprietà pubblica o privata della struttura¹⁷. Gli ospedali pubblici con amministrazione autonoma ottengono risultati analoghi a quelli ottenuti dai privati. Sono le cliniche universitarie e gli ospedali gestiti dalle ASL ad avere i risultati peggiori. Gli autori suggeriscono che la causa di questa cattiva *performance* vada rintracciata, per quanto riguarda le prime, nella loro natura di azienda multiprodotto (cure sanitarie, formazione e ricerca) che non viene colta dal modello stimato; per quanto riguarda i secondi, nel fatto che, al contrario degli ospedali autonomi, essi non beneficiano della

¹⁷ Risultati analoghi sono ottenuti da Barbetta e Turati (2000) per gli ospedali della Lombardia.

Figura 10

**Distribuzione dei valori relativi all'efficienza negli ospedali italiani
(Cellini e altri, 2000)**



separazione tra produttore e acquirente del servizio ospedaliero introdotta dalla riforma del 1992¹⁸.

Il grado di efficienza non varia significativamente per area geografica. È interessante notare, comunque, che mentre al Nord gli ospedali più efficienti sono quelli pubblici, al Sud sono quelli privati.

Infine, per quanto riguarda la relazione tra caratteristiche del mercato ed efficienza degli operatori, l'evidenza prodotta da Cellini e altri non è univoca: un'alta densità di ospedali sembra determinare perdite di efficienza, ma così pure un elevato grado di concentrazione. Più problematica appare la valutazione dell'effetto della presenza di ospedali privati sul grado di efficienza di quelli pubblici; in particolare tale presenza

¹⁸ Con la riforma le prestazioni effettuate dagli ospedali pubblici con amministrazione autonoma sono rimborsate in base alla classificazione DRG.

riduce l'efficienza se la quota di mercato privata è misurata in termini di posti letto, la aumenta se è misurata in termini di numero di ospedali.

Un altro studio recente in cui viene utilizzata la tecnica DEA per la valutazione dell'efficienza degli ospedali pubblici italiani è quello di Giuffrida, Lapecorella e Pignataro (2000). In questo lavoro gli autori analizzano l'efficienza tecnica dei presidi ospedalieri e delle aziende ospedaliere della Lombardia, con l'obiettivo di mettere in evidenza eventuali differenze tra i due tipi di strutture. L'analisi viene condotta considerando come *input* il personale (medico, infermieristico e amministrativo) e i posti letto. Gli autori considerano due *output* intermedi rispetto al miglioramento delle condizioni di salute dei pazienti: il numero dei ricoveri effettuati e le giornate di degenza. Infine la dimensione della qualità delle cure effettuate viene colta utilizzando l'indice di mortalità per alcuni tipi di ricoveri particolarmente gravi. Per tutte le specificazioni del modello considerate l'efficienza tecnica delle aziende ospedaliere è superiore a quella dei presidi, e questa differenza è dovuta maggiormente alla componente di efficienza tecnica pura. Per quanto riguarda la scala di produzione, i risultati suggeriscono che le aziende ospedaliere tendono ad assumere dimensioni troppo grandi, mentre i presidi hanno la tendenza opposta.

Anche Fabbri (2001) analizza l'efficienza degli ospedali in Italia. Il campione considerato in questo caso è quello degli ospedali pubblici nella regione Emilia Romagna per gli anni 1994-95. La tecnica utilizzata è anche questa volta la DEA, e l'oggetto dell'analisi è limitato alla pura efficienza tecnica nella produzione. I risultati suggeriscono un recupero di efficienza degli ospedali pubblici nei due anni considerati; in particolare i risultati migliori in termini di aumento dell'efficienza sono ottenuti dagli ospedali in gestione diretta delle ASL, mentre le aziende ospedaliere sembrano reagire meno agli interventi normativi intervenuti negli anni considerati. Si noti comunque che la situazione di partenza è caratterizzata per le aziende ospedaliere da un livello di efficienza tecnica piuttosto elevato. L'analisi di sensibilità dei risultati rispetto alla definizione di *output* suggerisce una certa robustezza delle conclusioni rispetto alle specificazioni adottate.

Si può osservare che gli studi di efficienza del sistema ospedaliero italiano fin qui esaminati non tengono conto di alcuni aspetti considerati di particolare rilevanza dagli utenti e in generale legati alle caratteristiche qualitative del servizio. Elementi per una discussione di questi aspetti e per una valutazione dell'efficienza percepita dagli utenti dei servizi sanitari sono forniti da un'indagine campionaria dell'Istat che riporta dati relativi al

grado di soddisfazione nella fruizione dei servizi ospedalieri. In questa indagine vengono considerati aspetti come la soddisfazione rispetto al vitto, ai servizi igienici, all'assistenza infermieristica oltre che medica.

I dati Istat disponibili sono raccolti nell'ambito delle cosiddette Indagini Multiscopo sulle famiglie. Nella definizione ufficiale, "...Alla fine di ogni anno vengono rilevati gli aspetti fondamentali della vita quotidiana e i comportamenti relativi all'anno in corso. Le aree tematiche presenti nei questionari permettono di cogliere la vita quotidiana della popolazione e se è soddisfatta dei servizi di pubblica utilità." (Istat, 2001, p. 7).

Nell'indagine relativa al 2000 viene presentata una sezione riguardante l'utilizzo dei servizi ospedalieri. Il 3,5 per cento della popolazione ha dichiarato di essere stato ricoverato in ospedale almeno una volta nei tre mesi precedenti l'intervista. Gli aspetti del ricovero di cui i pazienti si dichiarano maggiormente soddisfatti sono quelli concernenti l'assistenza, sia essa medica (85,6 per cento) o infermieristica (83,2 per cento), mentre riguardo agli aspetti legati al *comfort* la quota di soddisfatti risulta notevolmente inferiore (66,9 per cento per il vitto e 73,9 per i servizi igienici), soprattutto nelle strutture dei grandi centri urbani.

A livello territoriale, chi risiede nel centro-nord manifesta una maggiore soddisfazione per tutti gli aspetti del servizio rispetto ai residenti nelle altre aree geografiche. Le valutazioni meno positive provengono dai cittadini del sud e delle isole (Tabella 4).

La valutazione dell'efficienza nel settore della sanità in Italia è stata oggetto, soprattutto negli ultimi anni, di un numero crescente di analisi. Negli studi disponibili sono state applicate anche tecniche sofisticate. Gli sviluppi, promettenti dal punto di vista delle metodologie, incontrano alcuni ostacoli. Da un lato la povertà di dati sulle singole unità produttive, dall'altro l'affidabilità di alcuni risultati interessanti, quali quelli relativi al confronto fra pubblico e privato e agli effetti di forme di concorrenza, risente dei problemi di robustezza delle tecniche disponibili. Inoltre l'analisi di efficienza del settore ospedaliero condotta secondo le metodologie descritte in questa rassegna, trascura aspetti considerati particolarmente rilevanti dagli utenti finali e legati in particolare alle caratteristiche qualitative del servizio offerto. Infine, oltre all'assistenza ospedaliera, esistono altri comparti della sanità pubblica a cui sarebbe necessario estendere l'analisi di efficienza nell'impiego delle risorse disponibili, quali ad esempio la medicina di base e i servizi diagnostici.

Tabella 4**Grado di soddisfazione per i vari aspetti del ricovero**

Anno 2000	Persone molto o abbastanza soddisfatte*			
	Assistenza		Vitto	Servizi Igienici
	Medica	Infermieristica		
Nord-Ovest	85,5	84,9	70,6	79,3
Nord-Est	89,9	89,1	74,2	86,0
Centro	88,3	87,2	65,9	75,8
Sud	81,5	77,8	63,0	65,0
Isole	83,3	74,7	56,2	57,5
Italia	85,6	83,2	66,9	73,9
	Persone molto o abbastanza soddisfatte*			
	Assistenza		Vitto	Servizi Igienici
	Medica	Infermieristica		
Italia				
1998	88,4	86,5	70,5	74,0
1999	89,3	85,9	72,9	76,4
2000	85,6	83,2	66,9	73,9

* per cento persone della stessa zona con almeno un ricovero.

Fonte: Istat (2001).

5.3 L'istruzione

L'analisi dei carichi di lavoro svolta in Formez (1987) riguarda, oltre al settore sanitario, anche quello dell'istruzione. Ne emerge un calo costante del carico di lavoro (misurato in termini di popolazione scolastica totale) per dipendente tra il 1970 e il 1978: fatto cento l'indice nel primo anno, si scende a 80,2 nel 1978. Risultati analoghi sono ottenuti da Del Santo (1990) per il periodo 1980-87 attraverso l'analisi delle ore di lezione per insegnante, corrette con un indice di congestione delle classi. Del Santo rapporta il suo indice ai consumi collettivi a prezzi costanti del settore dell'istruzione in modo da ottenere un indice del prodotto per unità di costo: una misura di efficienza che scende da 100 nel 1980 a 87,8 nel 1987.

All'inizio degli anni '90 il numero di alunni per insegnante in Italia risulta tra i più bassi nel confronto internazionale (Istat, 1995).

Franco (1993) mostra come il rapporto tra numero degli addetti e numero di studenti sia passato dal 6,9 per cento del 1960 all'8,4 del 1965 e, dopo una fase di stabilità nel quinquennio successivo, abbia ripreso a crescere in maniera sostanzialmente continua (solo tra il 1978 e il 1980 si registra una lieve flessione) sino a raggiungere il 13,5 per cento nel 1990. Nello stesso periodo l'incidenza della spesa sul PIL passa dal 2,1 al 3,7 per cento. Alla fine degli anni novanta l'incidenza della spesa per istruzione sul PIL raggiunge il 4,6 per cento.

Altre linee di indagine hanno analizzato il tasso di riuscita nei vari livelli scolastici (la percentuale di studenti che terminano con successo un dato livello) e la fase di transizione scuola-lavoro. Nel primo caso è stato segnalato un tasso di riuscita particolarmente basso nelle fasi estreme del corso di studi: scuola elementare ed università (si tratta di valori bassi anche nel confronto internazionale; cfr. Istat, 1995). Nel secondo sono stati sottolineati, da un lato, i punti di mancata corrispondenza quali-quantitativa tra offerta di istruzione da parte del sistema scolastico e domanda di lavoro qualificato da parte del sistema produttivo; dall'altro, il tasso di occupazione particolarmente basso di alcuni gruppi di laurea (Bernardi e altri, 1989; Alessi e Bruni, 1990; FGB-MPI, 1990; Istat, 1990; Tronti e Mariani, 1994; Cuffaro e Andolina, 1996).

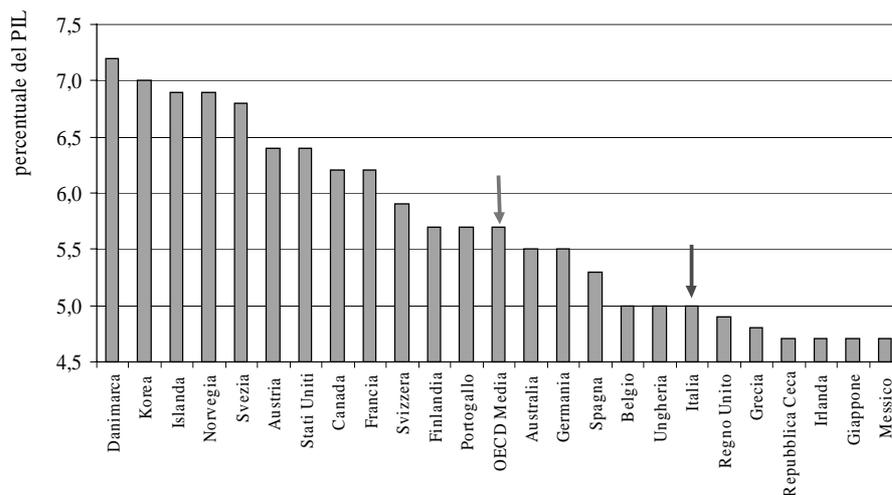
Nel complesso, l'analisi dell'efficacia dell'utilizzo delle risorse nel settore dell'istruzione non appare particolarmente sviluppata nel nostro Paese (cfr. Checchi, 1999).

Alcune indicazioni possono essere tratte dai confronti internazionali fra sistemi di istruzione (OCSE, 2001a e 2001b).

La spesa totale per l'istruzione in rapporto al PIL non appare in Italia particolarmente elevata, collocandosi al di sotto della media OCSE (5,7 per cento) (Figura 11). La spesa corrente costituisce il 96 per cento della spesa totale, con solo il 4 per cento destinato ad investimenti in conto capitale (Figura 12), mentre in media nei paesi OCSE l'8 per cento delle risorse è destinato agli investimenti. Le spese correnti sono costituite in parte preponderante dai compensi per il personale (83 per cento, di cui il 69 per il personale docente – contro una media OCSE pari, rispettivamente, all'80 e al 67 per cento).

Figura 11

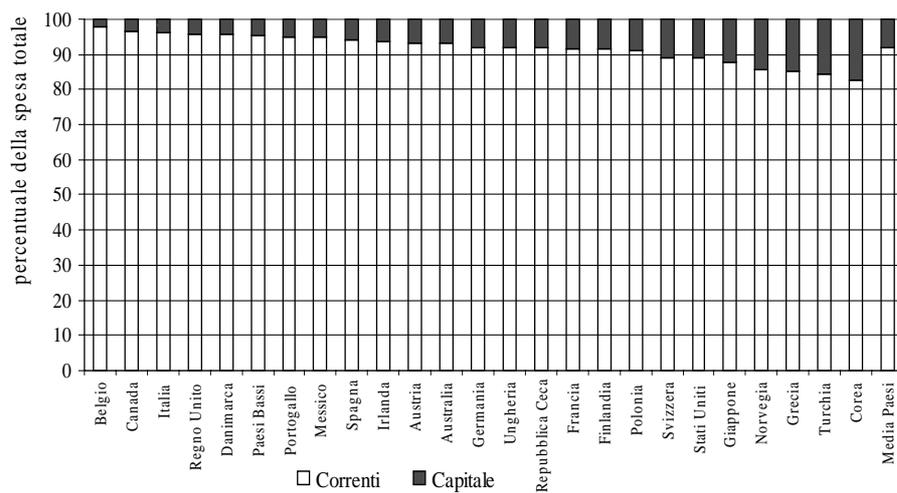
Spese totali per l'istruzione in percentuale del PIL (1998)



Fonte: OCSE (2001a).

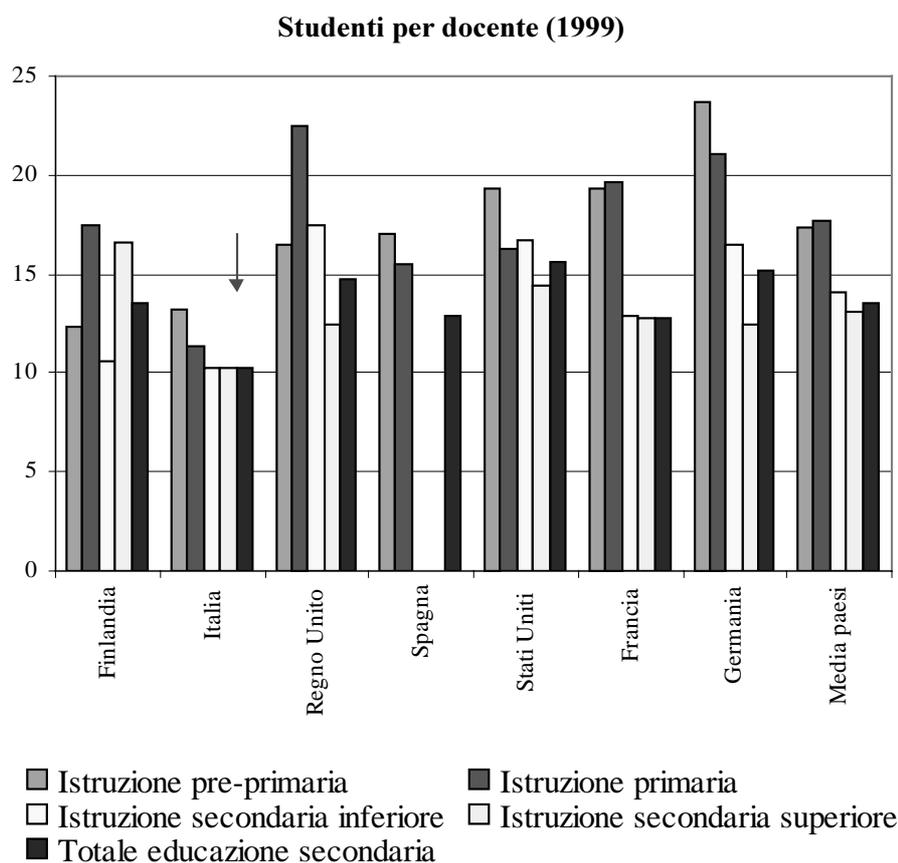
Figura 12

Spese in istruzione per categoria (1998)



Fonte: OCSE (2001a).

Figura 13



Fonte: OCSE (2001a).

Il rapporto studenti/personale docente continua a essere relativamente basso se confrontato con la media dei paesi OCSE (Figura 13).

Per quanto riguarda la misurazione dei livelli di istruzione raggiunti dai sistemi scolastici dei diversi paesi, sono disponibili i risultati per l'anno 2000 del *Programme for International Student Assessment* (PISA) dell'OCSE. In ciascuno dei 32 paesi partecipanti al programma un campione rappresentativo di studenti di 15 anni è stato sottoposto a test per verificare il livello di apprendimento in tre aree:

- 1) alfabetizzazione matematica;
- 2) alfabetizzazione scientifica;
- 3) capacità di lettura e comprensione dei testi.

Lo studio dell'OCSE considera dunque come misura di efficacia dei sistemi scolastici i risultati conseguiti in questi test standard. Tali risultati mettono in evidenza una posizione relativamente debole degli studenti italiani rispetto alla media dei paesi considerati (Tabella 5 e Figura 14).

La posizione relativa degli studenti italiani è sistematicamente al di sotto della media dei paesi considerati, con una significativa differenza rispetto ai livelli più elevati.

I dati dello studio dell'OCSE mettono in evidenza aspetti rilevanti dell'efficacia dei sistemi di istruzione nei diversi paesi. In Italia la disuguaglianza fra i risultati ottenuti dagli studenti è bassa se confrontata con quella degli altri paesi europei (soltanto in Spagna essa risulta inferiore). Tuttavia, livelli elevati di performance e bassi di disuguaglianza non si escludono¹⁹.

Anche se in generale livelli superiori di *performance* sono accompagnati da livelli elevati di PIL pro capite e spese in istruzione per studente, l'Italia ottiene risultati inferiori rispetto a quanto si potrebbe prevedere sulla base di questi indicatori, essendo caratterizzata da una spesa per studente elevata rispetto alla media dei paesi appartenenti al campione.

Non sono tuttavia disponibili analisi empiriche relative all'efficienza tecnica e/o allocativa nell'uso delle risorse investite nel sistema di istruzione italiano.

I problemi più rilevanti appaiono quelli della definizione e misurazione di una variabile di *output*, legata all'individuazione operativa degli obiettivi del sistema scolastico²⁰, e la disponibilità di dati per unità produttiva (istituti), o almeno per unità produttive omogenee, che permettano di mettere in relazione le *performance* osservate con gli *input* misurati.

¹⁹ Ad esempio, si vedano i casi di Giappone, Corea e Finlandia in OCSE (2001b).

²⁰ La problematicità della relazione fra misure empiriche di *performance* e obiettivi impliciti del sistema di istruzione è rilevata anche in Mayston (2002) con riferimento in particolare al sistema scolastico inglese.

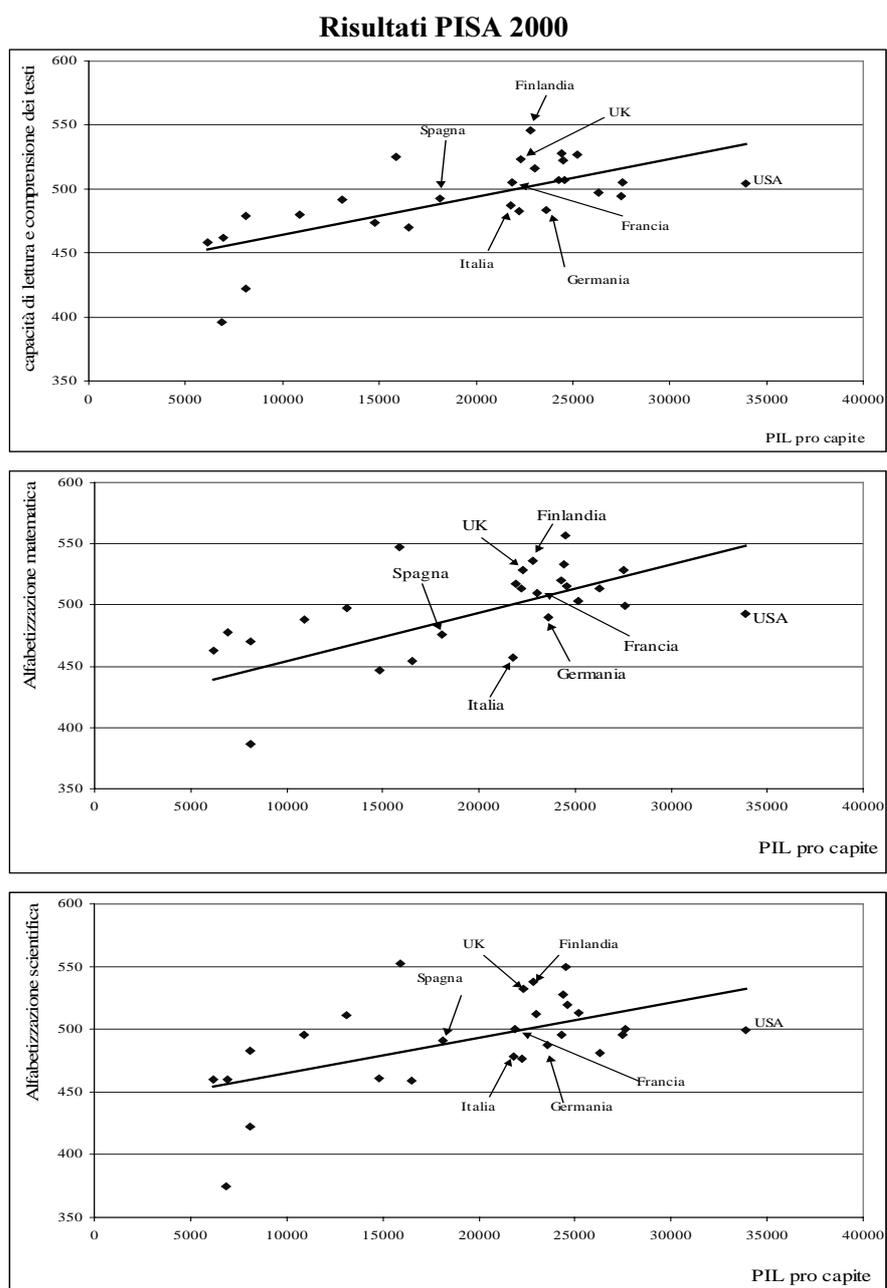
Tabella 5

Paese	Scala relativa all'alfabetizzazione per la comprensione e capacità di lettura		Scala relativa all'alfabetizzazione matematica		Scala relativa all'alfabetizzazione scientifica		PIL pro capite (dollari USA) (1999)	Spese per istituzioni educative e d'istruzione per studente (dollari USA) (1998)		
	Punteggio medio	Posizione	Punteggio medio	Posizione	Punteggio medio	Posizione		Posizione	Posizione	
Australia	528	4	533	5	528	6	24400	8	44623	12
Austria	507	12	515	11	519	8	24600	6	71387	1
Belgio	507	11	520	9	496	15	24300	9	46338	10
Canada	534	2	533	6	529	5				
Repubblica Ceca	492	19	498	18	511	11	13100	21	21384	20
Danimarca	497	16	514	12	481	22	26300	4	65794	3
Finlandia	546	1	536	4	538	3	22800	12	45363	11
Francia	505	14	517	10	500	12	21900	15	50481	9
Germania	484	21	490	20	487	20	23600	10	41978	14
Grecia	474	25	447	28	461	25	14800	20	27356	19
Ungheria	480	23	488	21	496	16	10900	22	20277	21
Islanda	507	10	514	13	496	17				
Irlanda	527	5	503	16	513	9	25200	5	31015	17
Italia	487	20	457	26	478	23	21800	16	60824	6
Giappone	522	8	557	1	550	2	24500	7	53255	8
Corea	525	6	547	2	552	1	15900	19	30844	18
Lussemburgo	441	29	446	29	443	29				
Messico	422	30	387	30	422	30	8100	23	11239	23
Nuova Zelanda	529	3	537	3	528	7				
Norvegia	505	13	499	17	500	13	27600	2	61677	5
Polonia	479	24	470	24	483	21	8100	24	16154	22
Portogallo	470	26	454	27	459	28	16500	18	36521	16
Spagna	493	18	476	23	491	19	18100	17	36699	15
Svezia	516	9	510	15	512	10	23000	11	53386	7
Svizzera	494	17	529	7	496	18	27500	3	64266	4
Regno Unito	523	7	529	8	532	4	22300	13	42793	13
Stati Uniti	504	15	493	19	499	14	33900	1	67313	2
Brasile	396	31	334	31	375	31	6840	26	9231	24
Lettonia	458	28	463	25	460	26	6164	27	m	
Liechtenstein	483	22	514	14	476	24	22235	14	m	
Federazione Russa	462	27	478	22	460	27	6930	25	m	
OCSE Totale	499		498		502					
OCSE Media	500		500		500					
Indici di Correlazione (1)			lett/mat		lett/scien		lett/Pil		lett/spese	
			0,88		0,92		0,66		0,47	
					mat/scien		mat/Pil		mat/spese	
					0,86		0,58		0,48	
							scien/PIL		scien/spese	
							0,50		0,33	
									Pil/spese	
									0,85	

(1) Coefficienti di correlazione tra i ranking derivanti dai punteggi medi delle relative variabili.

Fonte: OCSE (2001b).

Figura 14



Fonte: OCSE (2001b).

Per quanto riguarda in particolare il caso italiano, inoltre, è necessario tener conto che non esistono margini ampi per una politica “autonoma” nella gestione delle risorse da parte degli istituti scolastici. Per quanto riguarda i fattori produttivi occorrerebbe essere in grado di valutare la qualità del capitale umano impiegato (docenti), al di là di indicatori quali il titolo di studio (requisito omogeneo per classi di concorso ai fini dell’accesso all’insegnamento nel sistema di istruzione italiano).

Una strategia di analisi alternativa si basa sull’idea che il miglioramento delle capacità di reddito individuale (e quindi la futura *performance* degli studenti sul mercato del lavoro) o nazionale sia l’obiettivo del sistema educativo e su tali ipotesi costruisce una misura dell’efficacia/efficienza dei sistemi educativi osservati.

Di Liberto e Symons (2001), ad esempio, considerano il contributo dei livelli di istruzione raggiunti dalla popolazione ai tassi di crescita regionali italiani. I loro risultati sottolineano l’importanza dei livelli di istruzione primaria e il grado di alfabetizzazione nel promuovere la crescita nelle regioni meridionali durante gli anni ‘60, mentre la crescita dei livelli di istruzione superiore non sembrerebbe aver avuto un impatto significativo.

Analisi che utilizzano il reddito individuale come misura di *output* sono state effettuate per altri paesi. Un contributo significativo è quello di Card e Krueger (1992); essi adottano un approccio micro per stimare gli effetti della qualità dell’istruzione scolastica sui tassi di rendimento raggiunti dagli studenti maschi nati tra il 1920 e il 1949 nei diversi stati degli Stati Uniti. La qualità dell’istruzione scolastica viene misurata utilizzando tre indicatori: il rapporto fra studenti iscritti rispetto allo staff accademico per ogni stato, la lunghezza media dei trimestri e il salario medio annuo degli insegnanti. I risultati suggeriscono un effetto rilevante della qualità dell’educazione scolastica sui rendimenti dell’istruzione.

5.4 I trasporti

Per quanto riguarda il settore dei trasporti si riscontra una maggiore disponibilità di analisi applicate tese a mettere in evidenza il livello di efficienza del servizio.

Nell’ultimo ventennio le aziende di trasporto locale hanno accumulato notevoli disavanzi. Ciononostante, l’incidenza della spesa sul

PIL è rimasta costante intorno allo 0,7 per cento dal 1981 al 1987, riducendosi successivamente fino a circa lo 0,5 per cento nel 1997, ultimo anno per il quale sono disponibili dati di consuntivo. I disavanzi sarebbero pertanto riconducibili all'inadeguatezza delle risorse messe a disposizione delle aziende²¹. Esistevano tuttavia margini per un recupero di produttività. Le aziende pubbliche presentavano un numero di addetti per vettura e un costo medio per addetto superiori a quelli di analoghe aziende private. Vincoli di bilancio "morbidi" e rimborsi a piè di lista non fornivano incentivi al perseguimento dell'efficienza.

Una rassegna delle prime ricerche condotte in Italia su questi argomenti è contenuta in Cispel (1986); indicatori descrittivi sintetici e una loro applicazione al caso dell'azienda di trasporto Atac di Roma sono analizzati in Cerisola (1989).

Levaggi (1994) applica a un campione di 55 aziende di trasporto locali sia l'analisi basata sulle frontiere stocastiche, sia la DEA. Le due tecniche forniscono risultati analoghi: l'autrice rileva un notevole eccesso di capacità e un uso inefficiente del capitale; l'inefficienza incide per circa il 50 per cento sui costi delle aziende. Fabbri (1994) perviene a conclusioni analoghe esaminando le aziende di trasporto dell'Emilia Romagna.

Del trasporto locale si è occupato ancora Fabbri (1998) con una rassegna sugli studi di efficienza delle imprese pubbliche in questo settore. I risultati riportati sottolineano la presenza di rilevanti economie di scala di breve periodo (mentre la loro presenza è più incerta nel lungo) e una dimensione produttiva non ottimale; gli indici di inefficienza indicano l'esistenza di significativi spazi di miglioramento nella produzione del servizio.

L'analisi dei servizi di trasporto pubblico locale viene effettuata anche da Fraquelli, Piacenza e Abrate (2001) attraverso la stima di una funzione di costo variabile translogaritmica. Il campione considerato comprende 45 imprese pubbliche che operano nel trasporto urbano e in quello extraurbano negli anni 1996-98. I risultati delle stime mettono in evidenza rilevanti economie di scala non solo di breve, ma anche di lungo periodo, suggerendo che il settore del trasporto locale possa essere considerato un monopolio naturale a livello locale. In continuità rispetto a studi precedenti gli autori trovano anche che la tecnologia utilizzata in

²¹ Tra il 1982 e il 1991 il Fondo nazionale trasporti si è ridotto del 21 per cento in termini reali; le tariffe del trasporto pubblico sono rimaste a lungo bloccate da apposite direttive governative.

questo settore consente scarse possibilità di sostituzione fra i fattori, messe in evidenza da ridotte elasticità.

Piacenza (2001) analizza il settore del trasporto locale italiano studiandone i livelli di efficienza tecnica e il modo in cui questi sono influenzati dai meccanismi di sussidio alle imprese pubbliche di trasporto locale (in particolare gli effetti della regolamentazione attraverso contratti di tipo *cost-plus* o *fixed price*)²² per un campione di 45 imprese su un arco temporale di 7 anni (1993-99 – gruppi di imprese sono osservate per sottoperiodi fra gli anni 1993 e 1999). Lo scopo dell'analisi è quello di mettere in evidenza come l'evoluzione della regolamentazione pubblica nel settore del trasporto urbano possa avere inciso sui livelli di inefficienza nella fornitura del servizio negli anni '90, anni in cui il settore ha subito profonde riforme a livello normativo. I risultati empirici conducono a rigettare l'ipotesi di assenza di inefficienza tecnica delle imprese comprese nel campione considerato. C'è evidenza di capacità produttiva inutilizzata e che il monopolio naturale locale sia il riferimento organizzativo rilevante nel settore del trasporto locale. Per quanto riguarda gli effetti delle condizioni esterne nell'ambito delle quali le imprese di trasporto si trovano ad operare, si ha che un aumento della velocità di percorrenza delle tratte coperte porterebbe ad una diminuzione significativa del costo medio, il che condurrebbe a considerare politiche rivolte al contenimento del traffico privato fra le misure utili a favorire il conseguimento dell'efficienza produttiva del trasporto pubblico. In media l'inefficienza delle imprese del trasporto locale è peggiorata durante il periodo considerato. Tuttavia, l'inefficienza di costo si è ridotta per la maggior parte delle imprese che passano da una regolamentazione di tipo *cost-plus* ad una di tipo *fixed-price*. Quando l'inefficienza intrinseca della rete di trasporto²³ è molto elevata l'impatto dei sistemi di regolamentazione non è più rintracciabile nei dati. Analogamente è possibile ricondurre il peggioramento dei risultati di efficienza di alcune imprese al deterioramento delle condizioni ambientali in cui esse operano.

²² Dal punto di vista teorico nel caso di contratto *cost-plus* vi è un basso potere incentivante per l'impresa che viene rimborsata interamente dei costi sostenuti e riceve inoltre una maggiorazione fissata a priori e indipendente dalle scelte produttive dell'impresa stessa. Con un contratto di tipo *fixed-price* si prevede invece anticipatamente e in forma fissa l'ammontare complessivo del pagamento. La componente di incentivazione alla riduzione dei costi in questo tipo di contratto è quindi superiore poiché qualsiasi riduzione del costo determina un aumento del profitto dell'impresa.

²³ Ad esempio, una velocità molto ridotta dovuta a cattive condizioni ambientali.

I principali problemi nella valutazione e nell'utilizzo dei risultati degli studi empirici come quelli qui esaminati risiede nel fatto che queste analisi dell'efficienza del trasporto pubblico locale trascurano alcuni aspetti importanti e percepiti come rilevanti dagli utenti: puntualità, frequenza delle corse, qualità e capillarità del servizio, interconnessione dei diversi sistemi di trasporto (su gomma, tram, ferrovie locali, connessione ai trasporti su media e lunga distanza – ferrovie nazionali, autolinee, aeroporti).

Dall'indagine Istat (2001) emerge infatti anche un problema di qualità del servizio: solo poco più della metà degli utenti nel 2000 si è dichiarata soddisfatta dei diversi aspetti del trasporto urbano; dopo gli aumenti registrati tra il 1993 e il 1996, la percentuale di utenti soddisfatti diminuisce per tutti gli aspetti ad eccezione del costo del biglietto ed evidenza di nuovo un lieve aumento solo nell'ultimo anno considerato (Tabella 6). Le riduzioni maggiori tra il 1996 e il 2000 riguardano la puntualità (-5,8 punti), la comodità degli orari (-3,1) e i collegamenti (-3,8). I livelli più bassi di soddisfazione riguardano la comodità di attesa (36,6 per cento) e il costo (45,7 per cento).

Nella letteratura che ha analizzato l'efficienza dei sistemi di trasporto una parte rilevante dell'attenzione è stata rivolta, oltre che al trasporto urbano, anche a quello ferroviario. Il trasporto ferroviario come quello locale ha avuto notevoli problemi finanziari. Larga parte dei trasferimenti statali in favore di imprese infatti hanno riguardato e riguardano tuttora le ferrovie. Anche in questo caso è stata a lungo sostenuta l'esistenza di margini per un recupero di produttività; le ferrovie hanno avviato programmi di ristrutturazione che hanno riguardato sia il personale, sia il ridimensionamento del servizio offerto su alcune linee.

Coelli e Perelman (2000) hanno studiato l'efficienza dei sistemi di trasporto ferroviario nei paesi europei, considerando dati relativi a 17 paesi per il periodo 1988-93; il loro approccio non mira a stimare una frontiera di costo, ma utilizza il concetto di funzione di distanza nello spazio degli *inputs*. A questa scelta gli autori pervengono dopo avere adottato e confrontato vari metodi di stima. L'approccio basato sulla funzione di distanza consente di fare a meno di alcune ipotesi, come quelle relative all'obiettivo di minimizzazione dei costi in un contesto, come quello del trasporto ferroviario in cui prevalgono situazioni di monopolio, alto livello di regolamentazione pubblica del servizio e sussidi economici da parte dei governi. Il concetto di efficienza analizzato è ristretto all'ambito dell'efficienza tecnica. I risultati ottenuti mettono in evidenza un livello di

Tabella 6

Grado di soddisfazione per i vari aspetti del servizio di trasporto urbano*

Anno	Persone molto o abbastanza soddisfatte (percentuale)									
	frequenza corse	puntualità	posto a sedere	velocità corse	pulizia vettura	comodità attesa	collegamenti	comodità orari	costo	
2000										
Nord-Ovest	62,8	59,4	52,5	64,7	55,6	43,9	65,0	63,1	47,2	
Nord-Est	71,1	72,2	58,9	76,4	66,8	52,6	64,1	67,2	45,8	
Centro	48,7	44,4	42,6	58,0	47,7	28,3	51,0	50,8	52,1	
Sud	46,8	43,9	45,0	53,5	41,4	27,9	46,6	41,6	40,3	
Isole	33,5	30,8	30,8	43,7	29,8	21,6	36,0	32,4	32,8	
Italia	55,2	52,6	47,8	61,2	50,8	36,6	55,4	54,1	45,7	
Italia	Persone molto o abbastanza soddisfatte (percentuale)									
	frequenza corse	puntualità	posto a sedere	velocità corse	pulizia vettura	comodità attesa	collegamenti	comodità orari	costo	
1993	49,7	49,1	41,4	56,7	47,3	30,9	50,4	49,5	35,1	
1994	55,6	55,6	45,8	61,8	51,0	34,2	55,0	54,7	37,2	
1995	56,0	56,0	47,2	61,4	50,9	35,2	55,1	55,3	35,5	
1996	58,2	58,4	49,7	62,2	52,5	38,3	59,2	57,2	40,1	
1997	58,2	55,6	48,4	62,5	51,2	35,1	57,4	56,7	40,0	
1998	53,8	52,0	46,7	60,2	50,1	33,0	54,7	53,4	42,6	
1999	53,5	51,2	46,9	59,8	49,3	34,3	54,5	53,6	43,2	
2000	55,2	52,6	47,8	61,2	50,8	36,6	55,4	54,1	45,7	

*Campione: persone di età superiore ai 14 anni che utilizzano i mezzi di trasporto urbani.

Fonte: Istat (2001).

Tabella 7

Grado di soddisfazione per i vari aspetti del servizio ferroviario*

Anno	Persone molto o abbastanza soddisfatte (percentuale)							informazioni
	frequenza corse	puntualità	posto a sedere	pulizia	comodità orari	costo		
2000								
Nord-Ovest	69,1	50,5	65,9	28,8	63,4	41,9	53,2	
Nord-Est	72,0	61,2	65,9	36,3	65,9	43,6	59,4	
Centro	68,1	56,6	67,7	39,2	62,1	43,2	55,5	
Sud	63,8	55,2	60,6	33,3	55,7	35,0	50,1	
Isole	51,4	44,5	68,5	32,0	47,7	34,3	44,7	
Italia	67,1	54,5	65,2	33,6	60,8	40,5	53,6	
	Persone molto o abbastanza soddisfatte (percentuale)							
Italia	frequenza corse	puntualità	posto a sedere	pulizia	comodità orari	costo	informazioni	
1993	68,4	60,0	65,7	41,2	62,2	48,0	60,6	
1994	71,8	65,7	67,6	43,2	64,6	47,2	63,0	
1995	69,9	64,1	67,5	42,5	63,5	43,3	59,6	
1996	70,7	66,8	67,4	42,1	64,7	45,0	59,3	
1997	65,8	55,5	65,2	36,4	60,1	40,2	55,8	
1998	63,3	50,3	63,9	31,7	57,6	40,7	51,3	
1999	64,4	49,7	64,0	30,6	57,9	36,9	51,8	
2000	67,1	54,5	65,2	33,6	60,8	40,5	53,6	

*Campione: persone di età superiore ai 14 anni che utilizzano il treno.

Fonte: Istat (2001).

efficienza media delle ferrovie europee pari a 0,863 (dove il valore può variare tra 0 e 1, con 1 corrispondente a livelli di efficienza tecnica del 100 per cento). I livelli di efficienza stimati per 17 paesi europei mostrano come i risultati migliori siano raggiunti dalle ferrovie olandesi (0,980) mentre il valore più basso sia quello ottenuto per l'Italia (0,784) (Figura 15).

Lo studio confronta i risultati in termini di efficienza ottenuti dal trasporto ferroviario per gli anni 1978-83 con quelli relativi al periodo 1988-93 (Figura 16). Tale confronto consente di evidenziare come ci siano stati guadagni di efficienza in tutti i paesi (tranne la Grecia) e come questo sembri essere imputabile principalmente alla riduzione nell'uso degli *inputs*, in particolare del fattore lavoro (in gran parte dovuto alle pressioni per un riequilibrio dei bilanci da parte dei governi imposte anche dalle regolamentazioni introdotte dall'Unione Europea in materia di sussidi per le imprese di trasporto ferroviario).

Anche in questo caso vengono comunque trascurati aspetti qualitativi importanti nella valutazione dell'efficienza del servizio come il tipo di servizio offerto, la puntualità e la frequenza delle corse; elementi che indubbiamente incidono sulla percezione che gli utenti hanno dell'efficienza del servizio di trasporto.

Istat (2001) rileva una tendenziale riduzione della soddisfazione degli utenti del servizio ferroviario in Italia. Tra il 1996 e il 2000 si riscontra una diminuzione della percentuale di utenti che si dichiara molto o abbastanza soddisfatta per tutti gli aspetti del servizio ferroviario presi in considerazione nell'inchiesta (Tabella 7).

I cali più significativi riguardano la puntualità delle corse (-12,3 punti), la pulizia delle vetture (-8,5) e le informazioni (-5,7); mentre le percentuali di soddisfazione più basse si riferiscono alla pulizia delle vetture (33,6 per cento), al costo del biglietto (40,5 per cento), alle informazioni (53,6 per cento) e alla puntualità delle corse (54,5 per cento). La contrazione dei livelli di soddisfazione è generalizzata in tutto il paese, sebbene più accentuata nel nord e nel centro.

Figura 15

**Efficienza tecnica (media anni 1988-1993)
nel trasporto ferroviario europeo – Coelli e Perelman(2000)**

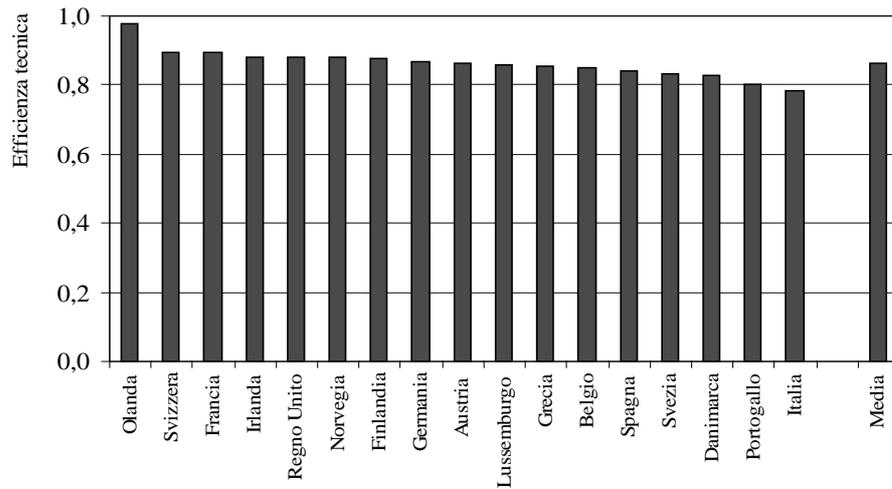
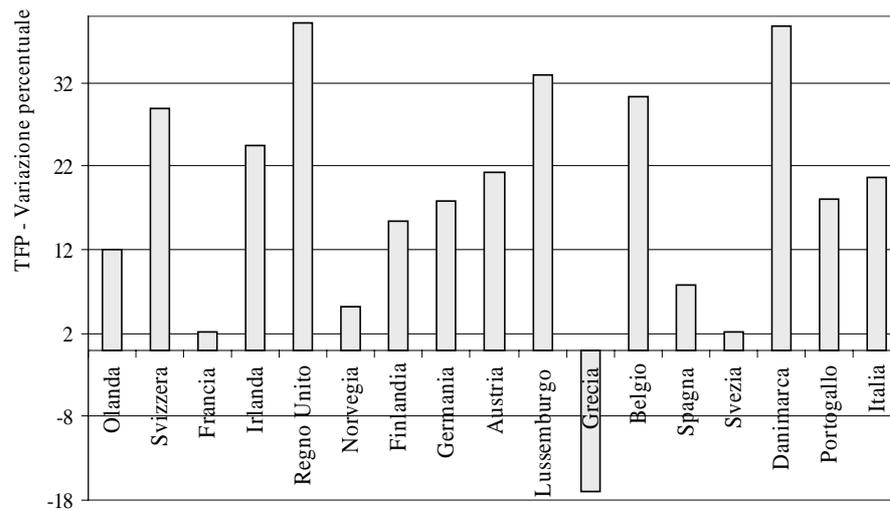


Figura 16

**Variazione percentuale della produttività totale
dei fattori (TFP) dal 1978-83 al 1988-93
nel trasporto ferroviario europeo – Coelli e Perelman(2000)**



5. Conclusioni

Il tema affrontato in questa rassegna è quello dell'efficienza dei servizi pubblici; l'intento quello di fornire una base per la discussione anche attraverso una panoramica (non esaustiva) degli studi disponibili.

L'utilizzo di risorse e la loro organizzazione ai fini della fornitura di un servizio possono essere valutati in termini sia di *efficacia*, sia di *efficienza*. L'efficacia si valuta in relazione al raggiungimento dello scopo per il quale viene fornito il servizio. Le difficoltà che si incontrano nel formulare una definizione operativa degli obiettivi dei servizi pubblici rendono la costruzione di indicatori di efficacia difficile e complessa. La valutazione dell'efficienza nella produzione, almeno da un punto di vista teorico e in prima approssimazione, appare più semplice. L'oggetto dell'analisi può essere identificato infatti nel rapporto fra risorse impiegate e prodotto ottenuto data la tecnologia disponibile.

In questa rassegna dopo aver trattato brevemente i temi della definizione teorica dell'efficienza produttiva e delle tecniche di stima adottate in letteratura, siamo passati all'analisi di alcuni contributi in aree rilevanti nel contesto dei servizi pubblici (come ad esempio la sanità e l'istruzione) ponendo particolare attenzione ai risultati disponibili per i servizi pubblici italiani.

Dalla discussione svolta sono emerse alcune aree critiche di particolare rilevanza.

In primo luogo, non vi è consenso sulle variabili più rilevanti e sulle tecniche più appropriate nei diversi settori. La discussione e comparazione dei risultati disponibili in letteratura mostra come l'adozione di tecniche differenti possa portare quasi sistematicamente a risultati contrastanti. I problemi metodologici si pongono fin dagli stadi iniziali dell'analisi empirica in cui devono essere definite non solo le tecniche di stima, ma anche le variabili di *input* e *output* da utilizzare. La criticità della scelta delle variabili da utilizzare è in parte connessa con la difficoltà già ricordata di definire e di ponderare gli obiettivi perseguiti dai servizi pubblici, in parte con quella di individuare variabili di prodotto che, oltre che misurabili, siano anche significative. Un altro aspetto di cui è necessario tenere conto nell'analisi empirica riguarda l'eterogeneità degli utenti dei servizi pubblici e gli aspetti qualitativi del servizio offerto.

Questi problemi si riflettono frequentemente nella complessiva scarsa robustezza dei risultati che si evidenzia in graduatorie di merito relative all'efficienza delle diverse unità produttive che sono fortemente dipendenti dalle specificazioni e metodologie adottate. È necessario ricordare inoltre che le tecniche di valutazione dell'efficienza sono molto sensibili alla presenza di osservazioni estreme. Si tratta di un problema particolarmente significativo, dato che proprio su tali osservazioni dovrebbero concentrarsi gli interventi per il miglioramento dell'efficienza nella fornitura del servizio. Appare pertanto necessaria estrema cautela nell'utilizzo dei risultati di questi studi; la loro interpretazione rimane infatti in molti casi ambigua e aperta ad obiezioni. La disomogeneità dei risultati che si ottengono utilizzando approcci diversi e la sensibilità dei risultati alla specificazione della variabile dipendente sembrerebbero inoltre consigliare particolare cautela nell'applicazione di tecniche di stima dell'efficienza a quei settori della produzione pubblica dove le misure quantitative presentano margini maggiori di ambiguità, quali ad esempio l'istruzione o la sanità. D'altro canto tali settori, per la loro particolare rilevanza, non solo in termini di risorse pubbliche investite, costituiscono aree di indagine verso cui si indirizza un forte interesse.

Dall'esame degli studi disponibili sui servizi pubblici in Italia emerge l'esistenza di aree in cui l'utilizzo e la valutazione di misure di efficienza non si è ancora diffusa (è il caso dell'istruzione), mentre per quanto riguarda altri settori, come la sanità e i trasporti, esiste un percorso di ricerca avviato e sono disponibili applicazioni che utilizzano tecniche sofisticate. Inoltre, in alcuni casi mancano dati soddisfacenti sulla produzione e sulle risorse utilizzate a livello di singola unità produttiva, indispensabili per approfondire la conoscenza del funzionamento dei settori considerati. Inoltre, agli studi sull'efficienza produttiva dovrebbero aggiungersi analisi riguardanti altri aspetti della fornitura di servizi pubblici, come la qualità percepita dagli utenti (il *comfort* del servizio e il rapporto con i medici nel caso della sanità; la frequenza, puntualità e interconnessione dei servizi nel sistema di trasporto pubblico) e l'efficacia delle prestazioni offerte.

RIFERIMENTI BIBLIOGRAFICI

- Aigner, D.J. e S.F. Chu (1968), *On Estimating the Industry Production Function*, in "American Economic Review", Vol. 58, n. 4, pp. 826-39.
- Aigner, D.J., C.A.K. Lovell e P.J. Schmidt (1977), *Formulation and Estimation of Stochastic Frontier Production Function Models*, in "Journal of Econometrics", Vol. 6, n. 1, pp. 21-37.
- Alessi, T. e M. Bruni (1990), *Sistema formativo e professioni. Dalla disoccupazione intellettuale al deficit educativo*, F. Angeli, Milano.
- Auster, R., I. Leveson e D. Sarachek (1969), *The Production of Health, An Exploratory Study*, in "Journal of Human Resources", Vol. 4, n. 4, pp. 411-36.
- Barbetta, G.P. e G. Turati (2000), *L'analisi dell'efficienza tecnica nel settore della sanità. Un'applicazione al caso della Lombardia*, Quaderni dell'Istituto di Economia e Finanza, Università Cattolica del Sacro Cuore, Milano, n. 28.
- Bernardi, L., G. Lovison e U. Trivellato (a cura di) (1989), *Sistema scolastico provinciale. Indicatori di funzionamento e previsioni*, COSES, Venezia.
- Blendon, R.J. e altri (1995), *Who Has the Best Health Care System? A Second Look*, in "Health Affairs", Vol. 14, n. 4, pp. 220-30.
- Bosco, B. e L. Parisio (1996), *Efficienza nella produzione pubblica di beni e servizi. Modelli teorici ed analisi econometrica*, Roma, La Nuova Italia Scientifica.
- Card, D. e A.B. Krueger (1992), *Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States*, in "Journal of Political Economy", Vol. 100, n. 2, pp. 1-40.
- Cellini, R., G. Pignataro e I. Rizzo (2000), *Competition and Efficiency in Health Care: An Analysis of the Italian Case*, in "International Tax and Public Finance", Vol. 7, n. 4-5, pp. 503-19.
- Cerisola, M. (1989), *Produttività delle aziende municipalizzate di trasporto*, in "Economia Pubblica", Vol. 19, n. 12, pp. 663-71.
- CISPTEL (1986), *Metodologia e parametri di efficienza per le aziende municipalizzate*, Roma, Cispel.

- Cecchi, D. (1999), *Istruzione e mercato*, Bologna, Il Mulino.
- Charnes, A., W.W. Cooper e E. Rhodes (1981), *Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through*, in "Management Science", Vol. 27, n. 6, pp. 668-97.
- Coelli, T. e S. Perelman (2000), *Technical Efficiency of European Railways: a Distance Function Approach*, in "Applied Economics", Vol. 32, n. 15, pp. 1967-76.
- Cubbin, J., S. Domberger e S. Meadowcroft (1987), *Competitive Tendering and Refuse Collection: Identifying the Source of Efficiency Gains*, in "Fiscal Studies", Vol. 8, n. 3, pp. 49-58.
- Cuffaro, M. e P. Andolina (1996), *Efficienza ed efficacia dell'istruzione universitaria: un'indagine campionaria sui laureati dell'ateneo palermitano*, in C. Quintano (a cura di), *Scritti di Statistica Economica 2*, Napoli, Rocco Curto Editore.
- Del Santo, A. (1990), *Il prodotto e la produttività nel settore delle amministrazioni pubbliche*, in "Economia e Lavoro", Vol. 24, n. 1, pp. 45-59.
- Department of Education and Science (1983), *School Standards and Spending: A Statistical Analysis*, in "Statistical Bulletin", n. 16/83, London.
- (1984), *School Standards and Spending: A Further Appreciation*, in "Statistical Bulletin", n. 16/84, London.
- Di Liberto, A. e J. Symons (2001), *Education and Italian Regional Development*, CEPR Discussion Paper, n. 496.
- Domberger, S., S. Meadowcroft e D. Thompson (1986), *Competitive Tendering and Efficiency: the Case of Refuse Collection*, in "Fiscal Studies", Vol. 7, n. 4, pp. 69-87.
- Elola, J., A. Daponte e V. Navarro (1995), *Health Indicators and the Organisation of Health Care Systems in Western Europe*, in "American Journal of Public Health", Vol. 85, n. 10, pp. 1397-401.
- Fabbri, D. (1994), *La struttura produttiva e del costo del trasporto pubblico urbano*, Quaderni Cifrel, Università di Ferrara.

- (1998), *La stima di frontiere di costo nel trasporto pubblico locale: una rassegna e un'applicazione*, in "Economia Pubblica", Vol. 28, n. 3, pp. 55-94.
- (2001), *Efficienza tecnica e produzione ospedaliera*, in "Economia Pubblica", Vol. 31, n. 1, pp. 37-70.
- Färe, R., S. Grosskopf e K.C.A. Lovell (1985), *The Measurement of Efficiency of Production*, Boston, Kluwer.
- (1994), *Production Frontiers*, Cambridge, Cambridge University Press.
- Farrell, M.J. (1957), *The Measurement of Productive Efficiency*, in "Journal of the Royal Statistical Society", Serie A, General, n. 120, parte 3, pp. 253-81.
- Feldstein, M. (1967), *Economic Analysis for Health Service Efficiency: Econometric Studies of the British National Health Service*, Amsterdam, North-Holland.
- FGB-MPI (1990), *Scuola e Sud. Gli esiti scolastici nel mezzogiorno*, Roma, Ministero della Pubblica Istruzione.
- Formez (1987), *La produttività nella Pubblica Amministrazione*, in CNEL (a cura di), *La produttività in Italia*, Milano, Edizioni del Sole-24 Ore.
- Førsund, F., C.A.K. Lovell e P. Schmidt (1980), *A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement*, in "Journal of Econometrics", Vol. 13, n. 1, pp. 5-25.
- Franco, D. (1993), *L'espansione della spesa pubblica in Italia*, Bologna, Il Mulino.
- Fraquelli, G., M. Piacenza, e G. Abrate (2001), *Costs and Technology of Public Transit Systems in Italy: Some Insights to Face Inefficiency*, Hermes Working Paper, n. 1/2001, Moncalieri (Torino).
- Frech, H.E. e R.D. Miller (1999), *The Productivity of Health Care and Pharmaceuticals: An International Comparison*, Washington, D.C., The AEI Press.
- Gabrielson, A. (1975), *On Estimating Efficient Production Functions*, Chr. Michelsen Institute, Department of Humanities and Social Sciences, Bergen, Norway, Working Paper, n. A-85.

- Giuffrida, A. e H. Gravelle (2001), *Measuring Performance in Primary Care: Econometric Analysis and DEA*, in "Applied Economics", Vol. 33, n. 2, pp. 163-75.
- Giuffrida, A., F. Lapecorella e G. Pignataro (2000), *Organizzazione dell'assistenza ospedaliera: analisi dell'efficienza delle aziende ospedaliere e dei presidi ospedalieri*, in "Economia Pubblica", Vol. 30, n. 4, pp. 101-24.
- Greene, W.H. (1980), *Maximum Likelihood Estimation of Econometric Frontier Functions*, in "Journal of Econometrics", Vol. 13, n. 1, pp. 27-56.
- (1982), *Maximum Likelihood Estimation of Stochastic Frontier Production Models*, in "Journal of Econometrics", Vol. 18, n. 2, pp. 285-89.
- (1997), *Frontier Production Function*, in M.H. Pesaran e P. Schmidt, *Handbook of Applied Econometrics. Volume II: Microeconomics*, Blackwell, Handbooks in Economics, pp. 81-166.
- Grosskopf, S., K. Hayes, L.L. Taylor e W.L. Weber (1998), *Allocative Inefficiency and School Competition*, in "National Tax Journal", Suppl., pp. 282-90.
- Grubalugh, S.G. e R.E. Santerre (1994), *Comparing the Performance of Health Care Systems*, in "Southern Economic Journal", Vol. 60, n. 4, pp. 1030-42.
- Hadley, J. (1982), *More Medical Care, Better Health?*, The Urban Institute.
- ISTAT (1990), *Indagine 1989 sugli sbocchi professionali dei laureati*, Roma, ISTAT.
- (1995), *Sistema educativo e mercato del lavoro nel contesto internazionale*, Roma, ISTAT.
- (2001), *I servizi pubblici e di pubblica utilità: utilizzo e soddisfazione*, Roma, ISTAT.
- Jesson, D., D. Mayston e P. Smith (1987), *Performance Assessment in the Education Sector: Educational and Economic Perspective*, in "Oxford Review of Education", Vol. 13, pp. 249-66.
- Jondrow, J., C.A.K. Lovell, I. Materov e P. Schmidt (1982), *On the Estimation of Technical Inefficiency in the Stochastic Frontier*

- Production Model*, in "Journal of Econometrics", Vol. 19, n. 2-3, pp. 233-38.
- Kaplan G. e altri (1996), *Inequality in Income and Mortality in the United States: Analysis of Mortality and Potential Pathways*, in "British Medical Journal", Vol. 312, n. 20, pp. 999-1003.
- Kennedy, B.P., I. Kawachi e D. Prothrow-Stith (1996), *Income Distribution and Mortality: Cross Sectional Study of the Robin Hood Index in the United States*, in "British Medical Journal", Vol. 312, n. 20, pp. 1004-07.
- Lavers, R.J. e D.K. Whynes (1978), *A Production Function Analysis of English Maternity Hospitals*, in "Socioeconomic Planning Science", Vol. 12, pp. 85-93.
- Levaggi, R. (1994), *Parametric and Non-Parametric Approach to Efficiency: the Case of Urban Transport in Italy*, in "Rivista di Politica Economica", Vol. 84, n. 10, pp. 67-88.
- Levitt, M.S. e M.A.S. Joyce (1987), *The Growth and Efficiency of Public Spending*, Cambridge, Cambridge University Press.
- Mayston, D. (2002), *Efficiency and Effectiveness in Public Services: Outstanding Problems of Performance Measurement*, mimeo.
- McGuire, A. (1987), *The Measurement of Hospital Efficiency*, in "Social Science and Medicine", Vol. 24, pp. 719-24.
- Meeusen, W. e J. van den Broeck (1977), *Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error*, in "International Economic Review", Vol. 18, n. 2, pp. 435-44.
- Mossialos, E. (1997), *Citizen's View on Health Care Systems in the 15 Member States of the EU*, in "Health Economics", Vol. 6, pp. 109-16.
- OCSE (1999), *What Causes Variations in the Performance of Health Care Systems?*, Paris, mimeo.
- (2001a), *Education at a Glance: OECD Indicators*, 2001 Edition, Paris.
- (2001b), *Knowledge and Skills for Life, First Results from PISA 2000*, Complete Edition, Paris.

- Olson, J.A., P. Schmidt e D.M. Waldman (1980), *A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions*, in "Journal of Econometrics", Vol. 13.
- Or, Z. (1997), *Out-Performing or Under-Performing: An Analysis of Health Outcomes in France and Other OECD Countries*, in "Health and System Science", Vol. 1, n. 3-4, pp. 321-44.
- Parisio, L. (1999), *A Comparative Analysis of European Railroads Efficiency: a Cost Frontier Approach*, in "Applied Economics", Vol. 31, n. 7, pp. 815-23.
- Piacenza, M. (2001), *Regulatory Constraints and Cost Efficiency of the Italian Public Transit System: an Exploratory Stochastic Frontier Model*, Torino, mimeo.
- Richmond, J. (1974), *Estimating the Efficiency of Production*, in "International Economic Review", Vol. 15, n. 2, pp. 515-21.
- The Royal College of General Practitioners (1992), *The European Study of Referrals from Primary to Secondary Care*, Report to the Concerted Action Committee of Health Services Research for the European Community, Occasional Paper, n. 56, Exeter, Devon.
- Rubenstein, R., A.E. Schwartz e L. Stiefel (1998), *Conceptual and Empirical Issues in the Measurement of School Efficiency*, in "National Tax Journal", Suppl., pp. 267-74.
- Schmidt, P. e C.A.N. Lovell (1979), *Estimating Technical and Allocative Efficiency Relative to Stochastic Production and Cost Frontiers*, in "Journal of Econometrics", Vol. 9, n. 3, pp. 343-66.
- Tronti, L. e P. Mariani (1994), *La transizione università-lavoro in Italia*, in "Economia e Lavoro", Vol. 28, n. 2, pp. 3-26.
- Wilkinson, R.G. (1992), *Income Distribution and Life Expectancy*, in "British Medical Journal", Vol. 304, n. 18, pp. 165-68.