

Efficienza

PAOLO BERTA, GIANMARIA MARTINI, GIORGIO VITTADINI

SOMMARIO: 1. L'efficienza tecnica e allocativa: concetti base, 63 – 2. La stima dell'efficienza tecnica, 64 – 3. L'analisi dell'efficienza tecnica degli ospedali lombardi, 66 – 4. Indicatore di efficienza tecnica e assegnazione del punteggio alle strutture, 66.

Un ospedale, come qualsiasi impresa produttiva, consuma risorse (input) per realizzare un prodotto / servizio (output) attraverso il quale raggiungere uno scopo: l'efficienza si misura in riferimento al rapporto tra input impiegati ed output prodotti.

Per valutare se e quanto un ospedale sia efficiente, occorre verificare che non sia possibile produrre di più con le medesime risorse, ovvero utilizzare meno risorse per ottenere la stessa quantità di prodotto, e che gli input siano utilizzati nelle proporzioni ottimali, dati i loro prezzi, per minimizzare il costo di produzione. Concretamente questo richiede la specificazione di una funzione di produzione o di costo, la determinazione del valore ottimo teorico e la misurazione della differenza tra la produzione osservata e l'ottimo teorico.

Ai fini di questo sistema di valutazione della *performance* si è scelto di utilizzare come indicatore di efficienza il numero di ricoveri standardizzato per il peso dei DRG. Tale indicatore è validato dalla letteratura internazionale ed esprime in modo appropriato l'efficienza tecnica di una struttura.

1. L'efficienza tecnica e allocativa: concetti base

Nella letteratura economica, l'efficienza produttiva di un'impresa è definita come la capacità di produrre il massimo quantitativo potenzialmente producibile data la tecnologia e i fattori produttivi disponibili, in un'ottica orientata all'output prodotto oppure, in un'ottica orientata

agli input, la capacità di ottenere un dato livello di produzione utilizzando la quantità minima potenzialmente utilizzabile di fattori produttivi, data la tecnologia disponibile (Aigner D., Chu S., 1967; Battese G., Rao D. e Coelli T., 1997).

Quando si verifichi la mancata corrispondenza tra il comportamento osservato ed il comportamento ottimo potenziale, ovvero si sia in presenza di inefficienza, risulta di fondamentale importanza essere in grado di comprendere e spiegare le cause dello scostamento, ed in particolare se essa sia dovuta alla presenza di inefficienza tecnica o allocativa.

Un'impresa si trova in una condizione di efficienza produttiva se produce al costo minimo necessario per realizzare un determinato livello di output. Due condizioni sono indispensabili affinché questo si verifichi: da un lato, l'assenza di uso eccessivo di risorse, dato l'output prodotto e la tecnica produttiva utilizzata (efficienza tecnica) (Chikos T., Sear A., 1994); dall'altro, l'impiego dei fattori produttivi in proporzioni tali da minimizzare la spesa, dati i prezzi e l'output prodotto (efficienza allocativa).

Per chiarire meglio l'analisi di efficienza nel suo complesso in appendice vengono approfonditi sia il concetto di efficienza tecnica che quello di efficienza allocativa: bisogna sottolineare che ai fini del modello di valutazione lombardo l'indicatore utilizzato è quello di efficienza tecnica, che può essere definita come l'abilità di un'impresa ad ottenere il massimo output dati gli input a disposizione (orientamento all'output), ovvero la sua abilità a minimizzare l'impiego di input necessari per ottenere un dato livello produttivo (orientamento all'input) (Gori E., Vittadini G., 1999).

2. La stima dell'efficienza tecnica

In accordo con la teoria economica della produzione, gli ospedali possono essere considerati come imprese che utilizzano fattori produttivi o input (medici, infermieri, tecnici, amministrativi, apparecchiature, materiale sanitario, materiale non sanitario, ...) e la tecnologia disponibile, per la produzione di servizi ospedalieri (output).

La funzione di produzione $y = f(x)$ identifica il massimo *output* producibile dati gli *input*. Questa funzione identifica delle frontie-

re, massima e minima rispettivamente, sulla quale le imprese pienamente efficienti dovrebbero collocarsi, dato lo stato della tecnologia (Newhouse J. P., 1994). La valutazione dell'efficienza di un ospedale richiede la specificazione delle variabili presenti nella funzione di produzione. I valori assunti dalle variabili nelle diverse unità produttive possono essere utilizzati per formulare valutazioni in merito all'efficienza individuale.

Le tecniche proposte in letteratura per la formulazione di tali valutazioni appartengono a due grandi famiglie. Da una parte vi sono le tecniche econometriche o parametriche (COLS, Frontiera Stocastica), dall'altra quelle di programmazione lineare o non parametriche (DEA) (Banker R. D., Charnes A. e Cooper W., 1984; Banker R. D., Conrad R. e Strauss R., 1986; Battese G., Coelli T. e Rao D., 1997). Le prime possono essere ulteriormente distinte in deterministiche o stocastiche. Nel primo caso (deterministiche), le deviazioni rispetto alla frontiera sono attribuite unicamente alle scelte del produttore e, pertanto, sono interamente imputate a inefficienza. Nel secondo caso (stocastiche), le deviazioni possono essere giustificate anche da fattori al di fuori del controllo del produttore quali stati di natura e altre circostanze come l'omissione di variabili o la presenza di errori di misurazione.

In appendice vengono brevemente illustrati i due principali metodi di analisi parametrici, COLS e Frontiera Stocastica, che fanno riferimento, rispettivamente, alla stima di frontiere deterministiche e stocastiche (Jacobs R., 2000).

Il metodo utilizzato per misurare l'efficienza di produzione in Regione Lombardia è la Frontiera Stocastica (Hollingsworth B., 2003; Rosko M. D., Mutter R. L., 2008).

Si è scelto di utilizzare tale metodo proprio perché, a differenza delle tecniche non parametriche (come la DEA) e di quelle deterministiche, esso permette di tenere in considerazione eventuali errori di misurazione delle variabili che possono influire sui risultati di *performance* di un ospedale.

3. L'analisi dell'efficienza tecnica degli ospedali lombardi

Nel corso degli ultimi anni, la Direzione Generale Sanità della Regione Lombardia ha iniziato a studiare con sistematicità l'efficienza del proprio sistema sanitario con l'obiettivo, da un lato di analizzare e comprendere le cause che spiegano i diversi livelli di efficienza degli ospedali e, dall'altro, di creare un unico ambiente integrato di analisi che coniugasse le tematiche di efficienza e di efficacia (Gori E., Rossi C. e Grassetti L., 2001).

In appendice viene illustrata la metodologia utilizzata per stimare gli scostamenti delle osservazioni rispetto alla frontiera ottima e studiare le determinanti di tali scostamenti. Per compiere la valutazione viene specificata una funzione di produzione e sono determinati gli scores di efficienza degli ospedali lombardi attraverso la tecnica della Frontiera Stocastica (Kumbhakar S. C., Knox Lovell C. A., 2000).

4. Indicatore di efficienza tecnica e assegnazione del punteggio alle strutture

L'indicatore selezionato concerne la valutazione dell'efficienza di produzione del numero di ricoveri standardizzato per il peso dei DRG.

In altre parole la valutazione di efficienza tecnica, eseguita con il metodo della frontiera stocastica sopra riportato, produce un indicatore che rappresenta per ogni ospedale lo score di efficienza tecnica rispetto ad una teorica frontiera di produzione stimata, che rappresenta il massimo della produzione che l'ospedale potrebbe ottenere dati gli input.

La differenza tra la produzione massima teorica e la stima puntuale della produzione osservata può essere scomposta in due componenti: una componente casuale che rappresenta il residuo del modello e una componente che, invece, misura l'efficienza dell'ospedale.

Pertanto dall'analisi di efficienza tecnica si ottiene un unico indicatore per ogni struttura ospedaliera.

Anche in questo caso i risultati del modello sono poi analizzati secondo i ranghi percentili.

Per fare la valutazione si è scelto di assegnare alle strutture un punteggio da 1 a 3 in base al rango percentile occupato dall'ospedale

secondo questa modalità:

- Rango percentile minore del 25°: 1
- Rango percentile compreso tra il 25° e il 75°: 2
- Rango percentile maggiore del 75°: 3

Anche in questo caso 3 indica una buona *performance* mentre 1 indica una cattiva *performance*.

La valutazione dell'efficienza attraverso la stima delle frontiere di produzione fornisce due importanti indicazioni. Il segno e la significatività statistica dei coefficienti della regressione consentono di identificare le variabili che aumentano e quelle che riducono l'efficienza dell'ospedale, e rappresentano pertanto un importante strumento di governo per il management delle strutture. I residui della regressione, invece, misurano il livello d'inefficienza dei singoli ospedali che costituiscono il sistema sanitario lombardo. Gli ospedali per i quali avviene che $u_i = 0$, a seconda del modello stimato, sono posizionati sulla frontiera. Questo significa che sono pienamente efficienti, cioè che utilizzano al meglio la propria capacità produttiva producendo il massimo output possibile senza sprechi di risorse. Quanto più elevato è il valore del residuo di un ospedale, tanto maggiore è il livello di inefficienza nel quale l'ospedale opera.

L'analisi dell'efficienza tecnica fornisce un'importante indicazione della qualità della "efficienza gestionale" della singola struttura ospedaliera, ovvero la sua capacità di erogare assistenza evitando sprechi. Tali indicazioni costituiscono una declinazione della qualità complessiva delle strutture ospedaliere, che va necessariamente posta in relazione con le altre dimensioni oggetto di valutazione presenti nel modello.

Paolo Berta, Gianmaria Martini, Giorgio Vittadini