



DEPARTMENT OF ECONOMICS
UNIVERSITY OF MILAN - BICOCCA

WORKING PAPER SERIES

**Price Cap e recupero di produttività:
suggerimenti dalla regolazione del settore Gas**

Massimo Beccarello

No. 11 - March 1998

Dipartimento di Economia Politica
Università degli Studi di Milano - Bicocca
<http://dipeco.economia.unimib.it>

Price Cap e recupero di produttività suggerimenti dalla regolazione del settore Gas

Massimo Beccarello

1. Introduzione

Con l'istituzione dell'Authority per l'energia la regolazione dei servizi pubblici italiani trova un esplicito modello di riferimento nel meccanismo di Price Cap¹. Diversamente dalle altre *public utilities* italiane, il settore del Gas Naturale ed in particolare la distribuzione del metano all'utenza civile, può essere considerato a tutti gli effetti un precursore nell'applicazione delle regole di *capping* nella regolazione dinamica delle tariffe. A partire dal 1993 il settore della distribuzione gas ha applicato un meccanismo di regolazione *ibrido* adeguando le tariffe praticate all'utenza civile con un procedimento misto: alcune componenti della tariffa Gas hanno mantenuto la loro natura di componenti di costo recuperate a "piè di lista"; altre componenti, invece, sono state adeguate mediante un meccanismo di indicizzazione simile al Price Cap.

In Italia, la distribuzione finale del gas metano presenta delle peculiarità istituzionali dovuta al gran numero di operatori (quasi 750) che operano in modo capillare su tutto il territorio nazionale. Alcuni operatori hanno natura pubblica (aziende municipalizzate e gestioni comunali), altri sono privati: in entrambi i casi, si tratta di monopolisti che nell'area geografica di competenza operano in regime di concessione.

Le tariffe che questi operatori praticano all'utente finale sono state regolate mediante la sedimentazione di tutta una serie di provvedimenti

legislativi del CIP e più recentemente² del Ministero dell'Industria. Dal 1993 l'adeguamento delle tariffe è determinato mediante un meccanismo che divide il ricavo consentito all'azienda (costo standard) in tre componenti, rispettivamente:

1. quota materia prima: l'impresa può rivalersi sull'utente finale per il recupero dei costi del metano acquistato dalla SNAM sulla base di un contratto periodicamente regolato per tutte le aziende di distribuzione³;
2. una quota investimenti mediante la quale l'impresa può in parte recuperare l'ammortamento finanziario per la propria dotazione;

¹ Art. 3 della legge n. 481 del 9 novembre 1995.

² A partire dal 1994.

³ Il contratto tra la SNAM e le aziende di distribuzione è standardizzato al fine di evitare discriminazioni tra i distributori finali.

3. la terza componente è la quota gestione, cioè la componente tariffaria che dovrebbe consentire all'impresa la copertura dei costi di esercizio e degli oneri generali. Gli elementi di costo che rappresentano la quota gestione sono in prevalenza i costi del lavoro.

La dinamica di queste componenti viene diversamente regolata. La prima può considerarsi a tutti gli effetti una componente tariffaria *pass through*, nel senso che il costo della materia prima viene trasferito totalmente al consumatore. La seconda componente viene adeguata sulla base del grado di sviluppo, ovvero sulla base del rapporto tra quantità venduta e numero di utenti dell'impresa. L'ultima componente, che costituirà l'oggetto della nostra analisi, viene regolata mediante il Price Cap.

Nonostante l'ampia letteratura che si è recentemente sviluppata sulle *public utilities*, solo in rari casi sono stati affrontati problemi metodologici relativi alla stima del recupero di produttività nel Price Cap⁴. Anche in paesi quali il Regno Unito, le modalità analitiche con le quali le Autorità di regolazione determinano il fattore X rimangono sconosciute.

L'obiettivo del *paper* è quello di valutare il livello di efficienza nell'impiego del fattore lavoro nel tentativo di costruire un indicatore di "recupero produttività" per l'applicazione del meccanismo di Price Cap. Lo strumento di analisi utilizzato è stato la stima econometrica di frontiere di efficienza stocastiche mediante Panel Data: i risultati ottenuti sono stati impiegati per stimare un indicatore di efficienza del tipo Battese-Coelli (1988). I dati sono stati resi disponibili dal Ministero dell'Industria su un campione rappresentativo di 115 esercizi di distribuzione relativo a un insieme di 80 aziende. Il campione è suddiviso in 46 imprese Municipalizzate e 32 imprese private (che con un totale di 69 esercizi di distribuzione).

Il lavoro è suddiviso come segue: nel secondo paragrafo sarà analizzato il modello teorico di riferimento; nel terzo parte, analizzando il meccanismo di Price Cap nella tariffa Gas, viene proposta la costruzione di un indicatore di efficienza relativo; il quarto paragrafo presenta il modello di stima, i dati utilizzati e i risultati. Il quinto paragrafo conclude.

2. Il modello teorico di analisi

Come anticipato nell'introduzione ogni distributore locale di gas metano opera in condizioni di monopolio. In tali casi la teoria economica ci suggerisce che, in equilibrio, l'output dell'impresa deve eguagliare la funzione di domanda del servizio, ovvero:

[1]
$$Y = f(p)$$

⁴ Si veda ad esempio P. Burns e T. Weyman-Jones (1994).

dove con Y indichiamo l'output dell'impresa e con $f(p)$ la domanda di mercato come funzione del prezzo p . L'impresa svolgerà la propria attività produttiva utilizzando principalmente due *input*, capitale e lavoro:

$$[2] \quad Y = F(K, L)$$

Se invertiamo la funzione di produzione possiamo facilmente derivare le funzioni di domanda dei fattori produttivi, rispettivamente:

$$[3] \quad \frac{\partial L}{\partial p} = \left[\frac{\partial L}{\partial Y} + \frac{\partial L}{\partial K} \frac{\partial K}{\partial Y} \right] \frac{\partial Y}{\partial p}$$

e:

$$[4] \quad \frac{\partial K}{\partial p} = \left[\frac{\partial K}{\partial Y} + \frac{\partial L}{\partial K} \frac{\partial L}{\partial Y} \right] \frac{\partial Y}{\partial p}$$

In entrambe le funzioni il secondo termine tra parentesi non risulta *a priori* determinato nel segno: tuttavia, ipotizzando coefficienti di produzione fissi assumerà valore nullo. Il primo termine delle funzioni di domanda, ovvero le variazioni dei fattori impiegati rispetto alla variazione dell'output, rimane comunque positivo. Date queste ipotesi possiamo riscrivere le equazioni [3] e [4] nel modo seguente:

$$[5] \quad \frac{\partial L}{\partial p} = L'(Y)f'(p) \leq 0$$

$$[6] \quad \frac{\partial K}{\partial p} = L'(K)f'(p) \leq 0$$

Oggetto della nostra analisi è la domanda del fattore lavoro [5]. Tale funzione costituirà il punto di partenza della nostra indagine nel tentativo di inferire in merito all'efficienza delle aziende di distribuzione nell'impiego del fattore lavoro.

Derivando le condizioni di ottimo in un contesto stocastico possiamo riscrivere la [5] per la i -esima impresa al tempo t come segue:

$$[7] \quad L_{it}^* = L'_{it}(Y_{it})f'(p_{it})e^{e_{it}}$$

dove L_{it} è la quantità di fattore lavoro utilizzata per produrre Y_{it} ovvero l'output domandato e prodotto dall'impresa al prezzo p_{it} . Il termine di errore e_{it} può essere ulteriormente decomposto in $e_{it} = v_{it} - i_i$ al fine di considerare l'(in)efficienza specifica dell' i -esima impresa. In particolare, assumeremo che v_{it} sia normalmente distribuito con $N(0, \sigma^2)$, mentre per la seconda componente assumeremo semplicemente⁵ $u_i \geq 0$ per \forall_i . Date queste ipotesi possiamo riscrivere la [7] come segue:

$$[8] \quad L_{it}^* = L'_{it}(Y_{it})f'(p_{it})e^{v_{it}} e^{u_i}$$

Nella [8], l'impresa si trova di fronte a due possibili fonti di errore nella scelta del fattore lavoro, ognuna delle quali è associata a una specifica causa di inefficienza:

1. una causa di inefficienza esogena, v_{it} , che è al di fuori del controllo dell'impresa (ad esempio uno shock dell'offerta che colpisce l'intero settore);
2. una causa di inefficienza endogena, u_i , che rileva l'inefficienza della i -esima impresa nell'adeguare l'input di lavoro alle sue specifiche esigenze di produzione.

3. Il Price Cap nella tariffa Gas

Il meccanismo di Price Cap viene applicato alla tariffa Gas con riferimento alla componente "Quota Gestione" (d'ora in poi Qg) la quale comprende "... i costi di gestione (personale, di esercizio, generali etc)"⁶. La formula di indicizzazione applicata è la seguente:

$$[9] \quad Qg_t^i = \overline{Qg}_{93} (1 + (I - cp_i))$$

dove I è il tasso di inflazione e cp_i è il coefficiente di "recupero produttività" dell' i -esima impresa, meglio noto nella letteratura come "X -Factor". Qg_{93}^i , è la quota gestione predeterminata sulla base del provvedimento CIP 24/88: costituisce la base di applicazione del meccanismo di Price Cap.

⁵ Si veda ad esempio Aigner *et al.* (1977).

⁶ Come da delibera CIPE del 23.12.93 pubblicata nella G.U. del 28.12.93.

L'oggetto della nostra analisi sono le modalità con le quali il coefficiente cp_i viene fissato. Secondo l'ultimo adeguamento tariffario, riportato nel D.m. 19 novembre 96 del MICA "G.U" 31.12.96, il recupero di produttività dipende dal valore della Qg calcolata "per utente" nel periodo precedente:

$$[10] \quad cp_i = f\left(\frac{Qg_{t-1}}{Utenti_{t-1}}\right)$$

I valori previsti sono quattro secondo la tabella⁷ 1:

Tabella 1 - Recupero di produttività Qg

Recupero di produttività	Qg/Qt in lire
0,25	fino a 150.000
0.50	200.000
1	250.000
3	oltre 300.000

L'indicizzazione dinamica della quota gestione è una funzione essenzialmente di tre variabili:

$$Qg_t^i = f\left(\begin{matrix} Qg_{93}^i & I & \frac{Qg_{t-1}}{Utenti_{t-1}} \\ + & + & - \end{matrix}\right)$$

ovvero, è una funzione positiva rispettivamente della quota gestione precedentemente determinata e del tasso di inflazione. Risulta, invece, una funzione negativa del rapporto tra quota di gestione e numero di utenti.

L'unica variabile che l'azienda può effettivamente controllare è quest'ultima: quanto minore è il costo di gestione per utente tanto più elevato sarà il recupero in termini tariffari associato quota gestione. In generale, la soluzione adottata significa parametrare il recupero di produttività alle sole economie di densità, cioè alla relazione che lega i costi di gestione al numero degli utenti serviti dall'azienda. Rimangono escluse eventuali economie di scala associate alle variazioni dei volumi di vendita e/o ad altre componenti associate ad elementi di "variazione tecnologica" dell'azienda quali, ad esempio, gli investimenti effettuati dall'azienda.

La componente principale di Qg sono sicuramente gli oneri del personale. Nella nostra analisi, riteniamo che sia plausibile ipotizzare una forte correlazione tra il valore di Qg ed il numero dei dipendenti (che chiameremo L), ovvero:

$$[11] \quad Qg \approx L$$

⁷ Inoltre, il recupero di produttività viene modulato con interpolazione lineare per i valori intermedi della quota gestione per utente.

Questa relazione ci consente di riscrivere la [10] come segue:

$$[12] \quad cp_i = \frac{L_i}{Utenti_i}$$

Sulla base della [12], le condizioni di efficienza vengono essenzialmente a dipendere dalle capacità di ottimizzare l'impiego del fattore lavoro da parte dell'impresa. E' evidente che questa specificazione ci impedisce di considerare tutte le variazioni di costo di gestione non direttamente imputabili al personale. Tuttavia, il nostro obiettivo è quello di pervenire ad un modello analitico di stima dell'efficienza relativa dell'*i*-esima impresa sulla base del quale inferire un recupero di produttività che non tenga conto delle sole economie di densità.

E' possibile quindi costruire un coefficiente di recupero produttività sulla base della domanda di lavoro media del settore, cp_m^i sostituendo la [8] nella [12]:

$$[13] \quad cp_m^i = \frac{L_{it}^i(Y_{it})f'(p_{it})e^{e_{it}}}{Utenti_i}$$

dove a numeratore troviamo la domanda di fattore lavoro stimata. A questo punto possiamo facilmente ottenere un indicatore di efficienza "relativa" confrontando il recupero di produttività dell'*i*-esima impresa con il valore medio del settore. Ovvero, dal rapporto tra la [12] e la [13] otteniamo:

$$[14] \quad cp_m^i = \frac{L_{it}^i(Y_{it})f'(p_{it})e^{e_{it}}}{L_i} = I_{e.r.}^i$$

In questo modo il recupero di produttività non viene più ad essere determinato in modo assoluto. Infatti, $I_{e.r.}^i$ consente di valutare l'efficienza dell'*i*-esima azienda rispetto alla situazione "media" eliminando qualsiasi *a priori* che non tenga conto dell'evoluzione tecnologica specifica del settore.

3.1 L'efficienza tecnica nell'impiego del fattore lavoro

Le conclusioni proposte dalla [14] risultano di facile interpretazione sul piano empirico in termini di efficienza tecnologica. Possiamo infatti calcolare l'efficienza relativa dell' i -esimo esercizio di distribuzione stimando il coefficiente di Battese-Coelli (1988). Definiamo l'efficienza tecnologica relativa all'impiego del fattore lavoro, il rapporto tra la quantità media di fattore lavoro che l'impresa avrebbe utilizzato in condizioni di scelta efficienti, ovvero $u_i = 0$, e la quantità di fattore lavoro effettivamente impiegato dall'impresa i -esima. In questo modo, l'efficienza tecnologica della i -esima impresa, $I_{e.r.}^i$, risulta:

$$[15] \quad I_{e.r.}^i = \frac{E(L_{it}^* | u_i = 0, Y_{it})}{E(L_{it}^* | u_i, Y_{it})}$$

dove L_{it}^* è la domanda di fattore lavoro per i -esima impresa al tempo t , Y è l'output venduto e u_i considera l'inefficienza dell'impresa secondo quanto abbiamo visto per la [8].

Per costruzione $I_{e.r.}^i$ assume dei valori compresi tra zero ed uno⁸. Se $I_{e.r.}^i$ assume ad esempio un valore pari a 0.90, in media, significa che l'impresa i -esima utilizza un quantità di fattore lavoro superiore a quella efficiente, ovvero quando $u_i = 0$. Questo risultato ha delle importanti conseguenze in termini di *welfare*: calcolare il coefficiente di recupero produttività nel modo seguente significa rompere la dipendenza diretta tra le variazioni della tariffa e la specifica struttura dei costi dell'impresa. Il prezzo finale dipenderà (almeno per quanto riguarda la quota gestione) contestualmente anche dalle performance delle altre imprese di settore che concorrono a determinare la frontiera di produzione, rispetto alla quale, abbiamo stimato l'inefficienza dell' i -esima impresa. In altri termini, ciò consentirà di creare un meccanismo *cost reducing* di tipo *Yardstick Competition*⁹. Poiché ogni impresa ha un forte incentivo ad ottimizzare la domanda di fattore lavoro (ovvero $u_i \rightarrow 0$) per ottenere un incremento di tariffa più elevato, nel lungo periodo a livello di sistema conseguirà un miglioramento della frontiera di produzione di tutto il settore. E' evidente che le conclusioni possono essere generalizzate anche nel caso di una valutazione di efficienza che considera una gamma più ampia di elementi di costo sulla base delle informazioni disponibili per il regolatore.

4. Dati, metodologia di analisi, risultati.

⁸ Infatti, a numeratore troviamo la quantità di fattore lavoro che sarebbe stata impiegata in condizioni di produzione medie efficienti, mentre, denominatore, troviamo la quantità di fattore lavoro effettivamente impiegata dall'impresa.

⁹ Si veda A. Shleifer (1985).

4.1 Dati

I dati utilizzati nella nostra analisi sono stati forniti dal Ministero dell'Industria. Il periodo di riferimento è stato 1985-1993. Queste variabili riguardano un campione di 114 esercizi di distribuzione raggruppati in due categorie:

- 1 46 esercizi di distribuzione municipalizzati;
- 2 69 esercizi di distribuzione privati.

Per ogni esercizio di distribuzione sono state utilizzate le seguenti variabili:

1. VEN, vendite in Mcal di Gas Naturale per ogni esercizio;
2. DIP, numero dei dipendenti;
3. UT, utenti totali;
4. INV, investimenti effettuati.

Le fonti di questi dati sono le schede tariffarie compilate dalle aziende di distribuzione e presentate periodicamente al MICA.

4.2 L'analisi empirica

Il punto di partenza dell'analisi statistica è stata la stima del modello teorico suggerito nel paragrafo 2. Abbiamo innanzitutto stimato la funzione di domanda di lavoro sulla base dell'equazione [7] :

$$[16] \quad LDIP_{it} = b_0 + b_1 LVEN_{it} + b_2 LUT_{it} + b_3 LINVR_{it} + e_{it} \quad \begin{cases} i=1, \dots, p \\ t=1, \dots, T \end{cases}$$

Come suggerito nell'equazione [8] il termine di errore può essere a sua volta decomposto come segue:

$$[17] \quad e_{it} = -u_i + v_{it}$$

dove u_i è una variabile casuale identicamente e indipendentemente distribuita con $u_i \approx D(a, s_a^2)$, e indipendente da v_{it} . Il parametro a è la media di u_i : il suo significato economico è quello di *proxy* del livello medio di inefficienza tecnologica. La misura dell'efficienza di una particolare unità si basa

sulla stima della variabile $-u_i$ relativamente alle altre osservazioni del campione. Per quanto riguarda le proprietà statistiche delle frontiere di efficienza si rinvia all'Appendice A.

Tutte le variabili sono state considerate in logaritmi naturali. La variabile dipendente è il numero di dipendenti $LDIP$ per l' i -esima azienda al tempo t . Tra le variabili esplicative abbiamo rispettivamente considerato il volume complessivo delle vendite $LVEN$, il numero di utenti LUT e, quale variabile tecnologico-dimensionale, il valore degli investimenti effettuati nell'anno $LINV^{10}$.

Nella nostra ipotesi, abbiamo tenuto separati nella funzione di produzione media stimata gli effetti sulla produttività legati al volume complessivo di output da quelli che sono gli effetti legati al numero degli utenti. Utilizzare come variabile dipendente il livello del personale è coerente con un'impresa di pubblici servizi. Essa infatti, in quanto obbligata a prestare il servizio agli utenti che ne fanno richiesta, non sceglie di fatto un livello produttivo ma dato l'output richiesto dovrà adeguare i fattori produttivi impiegati. La decisione di valutare la produttività del fattore lavoro in senso tecnico, permette di eliminare tutte le distorsioni associate al costo monetario difficilmente comparabili quali ad esempio, l'anzianità di servizio dei dipendenti, il costo degli straordinari etc.

Date le osservazioni disponibili, abbiamo stimato le frontiere di efficienza con Panel Data. L'utilizzo della tecnica Panel consente di superare uno degli elementi di critica principali dell'analisi in termini di frontiere: l'utilizzo della dimensione temporale nell'analisi *cross section* permette di cogliere le variazioni tecnologiche del campione osservato.

Poiché non si disponeva dello stesso numero di osservazioni su tutti gli esercizi di distribuzione è stato utilizzato un modello Panel *unbalanced*¹¹. I risultati ottenuti dalla frontiera di efficienza sono stati successivamente impiegati per stimare l'equazione [15].

4.3 Risultati

Mediante la stima dell'equazione [16] è stato possibile valutare le variazioni di domanda del fattore lavoro nelle diverse aziende a partire dal 1986. L'utilizzo delle tre variabili esplicative ci consente di separare gli effetti sulle economie di scala relative all'impiego del fattore lavoro, considerando separatamente gli effetti associati al volume di vendite, gli effetti di densità associati al numero di utenti serviti ed, infine, gli effetti tecnologici e di sviluppo degli esercizi rilevati dagli investimenti annualmente effettuati. Inoltre, le variabili esplicative utilizzate consentono indirettamente di rilevare tutta una serie di caratteristiche specifiche legate alla collocazione geografica dell'esercizio quali ad esempio densità urbana del territorio in cui l'esercizio opera,

¹⁰ Il valore monetario degli investimenti annui è stato considerato in termini reali utilizzando il deflatore del PIL anno base 1990.

tipologia di utenti dell'area territoriale, ed infine la senescenza ed il tasso di crescita della rete di distribuzione. La tabella 2 riporta i risultati delle stime effettuate:

Tabella 2 - Panel Aziende di distribuzione

Anno 1986/93	Numero di osservazioni 936	
Variabile dipendente	Numero dipendenti (LDIP)	
Variabili esplicative	Coefficienti	T-stat
Costante	-3,7495	-9,726
Vendite Totali (LVEN)	0,024	2,257
Numero Utenti (LUT)	1,035	45,219
Investimenti (LINVVR)	0.012	4.05
s_u	9,39	2,841
s_v		
$\sqrt{s_u + s_v}$	0,0088	36,342
σ_u	0,159	
σ_v	0,019	
R ² OLS	0,975	
Std. Res. OLS	0,96	
Coefficiente di scala	0.98	

La stima ha riguardato 114 esercizi di distribuzione per un numero complessivo di 936 osservazioni.

I risultati riportati nella tavola 2 suggeriscono rispettivamente:

1. esistono delle forti economie di scala nella domanda del fattore lavoro rispetto all'output totale: infatti, l'elasticità parziale suggerisce che a fronte di un aumento percentuale unitario del volume di vendite sufficiente un incremento del fattore lavoro impiegato del 0.024 %;
2. diversamente, l'elasticità parziale rispetto al numero di utenti, suggerisce che a fronte di un aumento percentuale unitario del numero di utenti l'aumento del fattore lavoro è pari al 1.035 %. I risultati stimati sembrano non confermare la presenza di economie di densità che si traducono in una diminuzione del fattore lavoro all'aumentare del numero di utenti;
3. infine, l'ultimo coefficiente suggerisce che a fronte di un aumento percentuale unitario negli investimenti la variazione del fattore lavoro è dello 0.012 %.

Tutti i coefficienti risultano statisticamente significativi (livello di significatività del 5%).

Dall'equazione stimata possiamo trarre un prima conclusione: l'applicazione di un Price Cap scalato con un recupero di produttività basato sulle economie di densità, almeno per la componente di quota gestione associata al costo del personale, appare non confermato sul piano empirico. In media la "scelta tecnica stimata" degli operatori del nostro campione sembra suggerire un comportamento tale per cui, in seguito ad una variazione percentuale del numero di utenti, consegue un adeguamento più che proporzionale del numero di dipendenti. In secondo luogo, la verifica empirica evidenzia forti

¹¹ Si veda, ad esempio, J. Seale (1990).

economie di scala rispetto all'output suggerendo una riformulazione analitica del recupero di produttività con un metodo scalato che privilegi il riferimento ad economie di scala rispetto al volume complessivo di vendite.

Successivamente, sulla base dei risultati ottenuti stimando la frontiera di domanda del fattore lavoro, abbiamo calcolato l'efficienza relativa degli esercizi di distribuzione con il coefficiente di Battese-Coelli come suggerito dalla [15]. La tabella 3 presenta i risultati stimati classificando per classi i valori assunti dal coefficiente. Poiché i valori sono compresi da $0 < I_{e.r.}^i < 1$, abbiamo diviso le frequenze del coefficiente in dieci classi di ampiezza 0,1. Nella tabella 3, le classi di valore assunte dal coefficiente sono rappresentate lungo la prima colonna. La seconda e la terza colonna riportano il numero di frequenze nelle diverse classi rispettivamente degli esercizi pubblici e di quelli privati:

Tabella 3 - Coefficiente di efficienza relativa

Classi	Esercizi pubblici	Esercizi privati	Totale
0-0,1	3	3	6
0,1-0,2	9	13	22
0,2-0,3	8	15	23
0,3-0,4	4	9	13
0,4-0,5	9	10	19
0,5-0,6	7	9	16
0,6-0,7	2	6	8
0,7-0,8	1	3	4
0,8-0,9	2	0	2
0,9-1	1	0	1
Val. min.	0,019	0,015	
Val. max	0,96	0,75	
Val. Medio	0,37	0,36	
Val. mediano	0,35	0,015	

I risultati della tabella 3 sembrano suggerire un livello di inefficienza nell'impiego del fattore lavoro piuttosto consistente. La maggior parte delle aziende complessivamente considerate presentano un indicatore $I_{e.r.}^i < 0.5$. In particolare, sono gli esercizi privati a presentare delle performance leggermente inferiori rispetto agli esercizi di distribuzione municipalizzati. Infatti, il valore medio del coefficiente stimato per le imprese pubbliche è di 0,37 contro lo 0,36 degli esercizi privati, ed il valore mediano ripropone un valore di 0,35 nelle imprese pubbliche contro uno 0,33 in quelle private. I valori più efficienti sono rispettivamente 0,96 per le imprese pubbliche e 0,75 per quelle private; i valori meno efficienti sono 0,019 per le imprese pubbliche e 0,015 per quelle private. Questi ultimi presentano una distribuzione più uniforme nelle diverse classi di inefficienza. Gli esercizi di distribuzione privati mostrano una maggiore concentrazione nelle classi di inefficienza

più elevata. Nell'appendice B sono riportati i risultati con i relativi valori per le diverse aziende del campione.

Lo scarto esistente tra i valori minimi e massimi dei coefficienti di efficienza relativa, nonché la distribuzione dei vari esercizi nelle diverse classi, suggerisce che un meccanismo di "recupero produttività" come quello proposto dalla normativa (tabella 1), risulta difficilmente in grado di cogliere il differenziale di produttività esistente negli esercizi del campione. Il metodo di stima proposto consente di valutare le diverse classi di efficienza con un diretto riferimento alla situazione media di settore. Il regolatore potrebbe quindi determinare un livello massimo di "recupero produttività" consentito e, successivamente, attribuire a ogni singolo operatore un valore scalato sulla base del livello di efficienza stimato.

L'adozione di un meccanismo di recupero produttività definito in termini relativi consentirebbe indirettamente di recuperare dei meccanismi di incentivo di tipo *Yardstick*: infatti, l'efficienza (e quindi il recupero di produttività di ogni singola impresa verrebbe a essere calcolata confrontando le performance della stessa con le performance delle altre aziende che condividono una tecnologia di produzione simile.

5. Conclusioni

Il vantaggio principale dell'impiego delle frontiere di produzione consiste nel riuscire a tener conto in maniera adeguata del legame funzionale esistente tra i fattori impiegati in una realtà produttiva e la produzione realizzata, separando il fenomeno dell'efficienza produttiva dai fenomeni, che risultano collegati nei consueti indicatori di produttività, del progresso tecnico e dei rendimenti di scala. Sulla base dei risultati ottenuti con le frontiere di efficienza abbiamo proposto l'utilizzo del coefficiente di Battese-Coelli quale possibile stima del "recupero produttività" associato al meccanismo di dinamica tariffaria Price Cap. L'analisi empirica effettuata su un campione di aziende del settore Gas ci ha consentito di verificare sul piano empirico la consistenza dell'attuale meccanismo di calcolo basato sul rapporto "Quota Gestione/Utenti". Possiamo riassumere brevemente le principali conclusioni come segue:

1. la stima della funzione di domanda di lavoro per il campione di imprese pubbliche e private non sembra suggerire la presenza di economie di scala rispetto al numero di utenti. Poiché gli oneri di

gestione (costi variabili) sono in prevalenza costituiti dai costi del personale, non sembra empiricamente fondato parametrare il recupero di produttività sulle sole economie di densità;

2. in alternativa, abbiamo proposto un recupero di produttività basato su un concetto di efficienza tecnica che, accanto alle economie di densità, consideri le economie di scala associate al volume di output venduto nonché ad altre variabili tecnologiche quali ad esempio investimenti nella struttura di rete. Sulla base dei risultati ottenuti abbiamo costruito la frontiera di efficienza del settore e, utilizzando il coefficiente di Battese-Coelli abbiamo misurato le performance delle diverse aziende di distribuzione;
3. dopo aver valutato il campione di imprese pubbliche e private, gli indicatori di produttività suggeriscono un livello di inefficienza piuttosto consistente. Per determinare il recupero di produttività il regolatore potrebbe adottare una procedura a due stadi: in una prima fase fissare un *performance* medie di settore; successivamente, scalare il coefficiente riconosciuto sulla base del grado di efficienza di ogni singola impresa;
4. il risultato più interessante di questa metodologia è che possiamo valutare la produttività di un'impresa relativamente a quella delle altre che operano nel settore. In pratica, l'adeguamento tariffario viene a dipendere contestualmente anche dalle performance delle altre imprese. Eliminando (in parte) la diretta dipendenza tra costi interni ed X-factor, il recupero di produttività così calcolato consente di introdurre dei meccanismi di *cost reducing* simili a quelli che sarebbero ottenuti in un contesto di *Yardstick Competition*.

Il modello può essere facilmente esteso per valutare l'efficienza anche da un punto di vista economico-tecnico. Infatti, a mano a mano che ulteriori informazioni sulle misurazioni monetarie dei costi di produzione saranno disponibili al regolatore, la valutazione dell'efficienza relativa può essere facilmente estesa alle funzioni di costo e/o produzione senza alcuna perdita di generalità.

Riferimenti bibliografici:

Aigner D.J., S.F. Chu (1968). "On Estimating the Industry Production Function". American Economic Review, vol. 54, n. 4.

Aigner D.J., Lovell C.A.K. e Schmidt P. (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models". Journal of Econometrics, vol. 6.

Battese G. E., e T.J. Coelli (1988). "Production of FirmLevel Technical Efficiencies with a Generalized Frontier Production Function". Journal of Econometrics, 38, pag. 387-399.

Burns P. e T.W. Weyman Jones (1994) "Cost Drivers and Cost Efficiency in Electricity Distribution: a Stochastic Frontier Approach". Technical Paper 2 CRI, London.

Cornwell, C., P. Schmidt e R.C. Sickels (1990). "Production Frontiers with Cross Sectional and Time-Series Variation in Efficiency". Journal of Econometrics, 46, oag. 185-200.

Fare R. e Lovell C.A.K. (1978). "Measuring the Tecnical Efficiency of Production". Journal of Economic Theory, vol. 19.

Farrel M.J. (1957). "The Measurement of Production Efficiency". Journal of Royal Statistical Society, vol. 120.

Greene W. (1980). "Maximum Likelihood Estimation of Econometrics Frontier Function". Journal of Econometrics, 13.

Greene W. (1996). Econometric Analysis. Prentice Hall.

Schmidt P., (1976). "On the Statistical Estimation of Parametric Frontier Production Function". Review of Economic Studies, 58.

\bibitem{se} Seale J.L. (1990). "Estimating Stochastic Frontier System with Unbalanced Panel Data: the Case of Floor Tile Manufactories in Egypt" Journal of Applied Econometrics, vol. 5, pag. 59-74.

Shleifer A. (1985). "A theory of Yardstick Competition". Rand Journal of Economics, Vol. 16, no. 3, Autumn 1985.

Appendice A

L'applicazione dei Panel Data nella stima di frontiere di efficienza é stata proposta in origine da Schmitd (1976) e da Cornwell, Schmitd e Sickels (1990). Utilizzando la procedura di stima OLS per la [16], come suggerito da Greene (1980), possiamo ottenere uno stimatore di \mathbf{b} consistente. In concetto di efficienza relativa emerge nel momento in cui lo stimatore \mathbf{b}_0 è ottenuto con uno spostamento *ad hoc* del termine costante. In tal modo i residui della stima OLS risultano tutti negativi, ovvero:

$$[18] \quad \tilde{\mathbf{b}} = \hat{\mathbf{b}}_0 + \max_{i,t} \mathbf{e}_{it}$$

dove \mathbf{e}_{it} sono i residui OLS dell'equazione [16]. A questo punto siamo in grado di determinare l'efficienza di ogni

$$[19] \quad eff_{it} = \exp \left(\mathbf{e}_{it} - \max_{i,t} \mathbf{e}_{it} \right)$$

Per separare gli effetti legati alla specifica impresa o gli effetti di tipo temporale nelle variazioni del livello di efficienza si può facilmente utilizzare un metodo ANOVA. La stima dell'efficienza della i – esima impresa può essere ottenuta come media temporale. Il limite principale del metodo deterministico consiste nel fatto che tutte le "osservazioni" si trovano da una parte della frontiera: in pratica le stime risentono ancora delle osservazioni "outliers" e nello stesso tempo non sono considerati eventuali *shock* casuali attorno alla funzione di produzione media.

Poiché le osservazioni sull'insieme degli esercizi di distribuzione, sia pubblici che privati, considerava un periodo di nove anni, abbiamo stimato la nostra frontiera di efficienza mediante un Panel Random Effect. Tuttavia, poiché non tutti gli esercizi presentavano lo stesso numero di osservazioni è stato necessario utilizzare un Panel Unbalanced. Data la struttura dei dati, l'equazione [16] può essere facilmente specificata nel modo seguente:

$$[20] \quad y_{it} = \mathbf{b}_0 + x_{it}' \mathbf{b} - u_i + v_{it} \quad \begin{cases} i = 1, \dots, p \\ t = 1, \dots, T \end{cases}$$

dove u_i è una variabile casuale identicamente e indipendentemente distribuita, indipendente da v_{it} con $u_i \approx D(a, \mathbf{s}_u^2)$.

La componente stocastica è in questo caso data da:

$$[21] \quad \mathbf{e}_{it} = -u_i + v_{it}$$

Il parametro a è la media di queste variabili e rappresenta una sorta di approssimazione del livello di inefficienza tecnologica. La misura dell'efficienza di una particolare unità si basa, in questo caso, sulla stima della variabile casuale $-u_i$ relativamente alle altre osservazioni del campione. In termini matriciali, considerando T osservazioni per ogni impresa, otteniamo:

$$(11) \quad \mathbf{y} = \begin{bmatrix} \mathbf{x}_{pT} & \mathbf{X} \end{bmatrix} \begin{pmatrix} \mathbf{b}_0 \\ \mathbf{b} \end{pmatrix} + \mathbf{e}_{it}$$

Il termine di errore, in questo caso, risulta strutturato come segue:

$$[22] \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{M} \\ \mathbf{e}_p \end{pmatrix}$$

con ogni:

$$[23] \quad \mathbf{e}_i = u_i \mathbf{i}_T + v_i \quad i = 1, \dots, p$$

Il cui valore atteso:

$$[24] \quad E(\mathbf{e}) = a \mathbf{i}_{pT}$$

Nel caso specifico la matrice varianze-covarianze non è più uno scalare e la procedura di stima OLS non risulta statisticamente efficiente. La struttura della matrice è data dalla seguente:

$$[25] \quad \sum_{v_i} = \mathbf{s}_u^2 \mathbf{i}_T \mathbf{i}_T' + \mathbf{s}_v^2 \mathbf{I}_T = \mathbf{A} = \begin{pmatrix} \mathbf{s}_u^2 + \mathbf{s}_v^2 & \mathbf{s}_u^2 & \mathbf{K} & \mathbf{s}_u^2 \\ \mathbf{s}_u^2 & \mathbf{s}_u^2 + \mathbf{s}_v^2 & \mathbf{K} & \mathbf{s}_u^2 \\ \mathbf{M} & \mathbf{M} & \mathbf{K} & \mathbf{M} \\ \mathbf{s}_u^2 & \mathbf{s}_u^2 & \mathbf{K} & \mathbf{s}_u^2 + \mathbf{s}_v^2 \end{pmatrix}$$

Infine, utilizzando il prodotto di Kronecker otteniamo per tutte le unità osservate:

$$[26] \quad \Sigma_v = I_p \otimes A = \begin{pmatrix} A & O & K & O \\ O & A & K & O \\ M & M & K & M \\ O & O & K & A \end{pmatrix}$$

La matrice inversa di A è data dalla seguente:

$$[15] \quad A^{-1} = \left(\frac{-\mathbf{s}_u^2}{\mathbf{s}_v^2 (\mathbf{s}_v^2 + T\mathbf{s}_u^2)} \right) i_T i_T' + \frac{I_T}{\mathbf{s}_v^2}$$

Per ottenere uno stimatore GLS di β_0 e β dobbiamo utilizzare il metodo dei minimi quadrati generalizzati. Innanzitutto possiamo ottenere degli stimatori consistenti di σ_ε^2 e σ_α^2 dai residui OLS ottenuti dalla stima dell'equazione [20]¹².

In tal modo possiamo ottenere lo stimatore GLS con la seguente espressione: applicati alla seguente equazione:

$$[28] \quad \begin{pmatrix} \beta_0 \\ \beta \end{pmatrix} = \left[\begin{matrix} i_n & X \end{matrix} \right]' \hat{\Sigma}_v^{-1} \left[\begin{matrix} i_n & X \end{matrix} \right] \left[\begin{matrix} i_n & X \end{matrix} \right]' \hat{\Sigma}_v^{-1} y$$

In realtà questo calcolo può essere evitato poiché è noto che gli stimatori GLS di β_0 e β possono essere ottenuti con dei semplici OLS sui dati così trasformati:

$$[29] \quad \begin{aligned} y_{it}^* &= y_{it} - c\bar{y}_i \\ x_{it}^* &= x_{it} - c\bar{x}_i \end{aligned}$$

dove c è il parametro che rappresenta la di *quasi*-devianza ottenuto da:

$$[30] \quad c = 1 - \sqrt{\frac{\mathbf{s}_u^2}{\mathbf{s}_u^2 + T\mathbf{s}_u^2}}$$

Applicando gli OLS sulla *quasi*-devianza dei residui otteniamo delle stime consistenti dei parametri β :

¹² Gli stimatori consistenti di σ_ε^2 e σ_α^2 si ottengono mediante la decomposizione della variabilità dei residui come segue:

$$(a) \quad E \left(\sum_{i,t} (v_{it} - \bar{v}_i)^2 \right) = p(T-1)\sigma_\varepsilon^2$$

$$(b) \quad E \left(\sum_{i,t} (\bar{v}_i - \bar{v}_{..})^2 \right) = (p-1)(\sigma_\varepsilon^2 + T\sigma_\alpha^2)$$

$$[31] \quad y_{it}^* = \beta_0(1-c) + \beta' x_{it}^* + v_{it}^* \quad \begin{cases} i = 1, \dots, p \\ t = 1, \dots, T \end{cases}$$

Lo stimatore di β_0 viene spostato “ad hoc” al fine di ottenere delle stime positive di u_i

Infine per stimare le (in)efficienze é necessario procedere partendo dalle stime di α_i . Il punto di partenza é costituito dai residui v_{it} che sono stati ricalcolati al fine di considerare le stime di β più efficienti ottenute con I GLS. La relazione tra e_{it}^* e e_{it} é data da:

$$[32] \quad e_{it} = e_{it}^* + \frac{c}{1-c} \bar{e}_i^*$$

Poiché $E(e_{it}) = 0$ e $u_i = v_{it} - e_{it}$, gli stimatori naturali di α_i , sono dati semplicemente da:

$$[33] \quad u_i = \max_j \bar{e}_j - \bar{e}_i$$

dove il valore massimo é considerato al fine di ottenere valori positivi per u_i .

Lo stimatore OLS di β_0 si ottiene mediante il solito “spostamento” della frontiera:

$$[34] \quad \tilde{\beta}_0 = \hat{\beta}_0 \max_j \bar{v}_j$$

per avere una misura globale si potrebbe ad esempio considerare una media di tutti gli u_i .

Appendice B

Tabella A - Coefficiente Battese-Coelli stimato per i 46 esercizi di distribuzione pubblici.

Tabella B - Coefficiente di Battese-Coelli stimato per i 69 esercizi di distribuzione privati