



ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

*Scuola di Dottorato in Scienze Economiche e Statistiche*

Dottorato di Ricerca in  
Economia e Statistica Agroalimentare  
XX ciclo

## L'impatto degli incentivi sull'efficienza delle imprese manifatturiere

Marco Peroni

Dipartimento di Scienze Statistiche "P. Fortunati"

Marzo 2008



**Alma Mater Studiorum**  
**Università degli Studi di Bologna**  
Dipartimento di Scienze Statistiche

---

Dottorato di Ricerca in  
*Economia e statistica agroalimentare – XX ciclo*

**L'impatto degli incentivi sull'efficienza delle  
imprese manifatturiere**

Tesi di Dottorato di:  
MARCO PERONI

Coordinatore:  
Chiar.mo Prof. ROBERTO FANFANI

Relatore:  
Chiar.mo Prof. GUIDO PELLEGRINI

SECS-P/02 Politica economica

*ESAME FINALE*

---

*Anno 2008*



# Indice

INTRODUZIONE.....	I
<b>1 ANALISI LETTERATURA SULL'IMPATTO DEGLI INCENTIVI SULL'EFFICIENZA DELLE IMPRESE</b>	<b>5</b>
Introduzione.....	5
1.1 Una misura di produttività e di efficienza .....	6
1.1.1 L'efficienza produttiva.....	8
1.1.2 La produttività.....	9
1.2 L'efficienza nella teoria microeconomica .....	11
1.2.1 Approccio parametrico .....	12
1.2.1.1 Metodi per dati cross – section .....	13
1.2.1.2 Metodi per dati panel .....	14
1.2.2 Approccio non parametrico .....	15
1.2.3 Data Envelopment Analysis (DEA).....	15
1.2.4 Nozioni teoriche e indagini empiriche sugli effetti delle sovvenzioni .....	16
1.2.5 Lo sviluppo recente e crescita degli incentivi in R&S.....	20

1.3	Approccio Macroeconomico: la Total Factor Productivity ( TFP) nell'analisi della produttività .....	22
1.4	Limiti nella 'misurabilità' del prodotto e stime della TFP: un esercizio basato sui dati per l'industria manifatturiera.....	25
1.5	Differenze nella qualità dell'istruzione e nel contributo del capitale umano .....	25
1.6	Gli effetti regionali degli incentivi nella performance delle imprese: studi empirici .....	26
1.7	Critiche alla TFP .....	28
<b>2</b>	<b>MISURE DI EFFICIENZA DELL'IMPRESA A LIVELLO MICRO</b>	<b>29</b>
	Introduzione.....	29
2.1	Sviluppo dell'analisi tecnica.....	30
2.2	Metodi Parametrici .....	33
2.2.1	Metodi per dati Cross – Section .....	34
2.2.2	Metodi per dati Panel .....	44
2.2.3	Modelli con variabili che determinano l'inefficienza tecnica.....	52
2.3	Metodi non parametrici .....	60
2.3.1	Data Envelopment Analysis.....	60
2.3.2	Free Disposal Hull (Fdh) .....	65
<b>3</b>	<b>MODELLO DELLA STIMA DEGLI EFFETTI PER GLI INCENTIVI ALLE IMPRESE</b>	<b>67</b>
	Introduzione.....	67
3.	Definizione di un modello di stima degli effetti.....	68
3.1	Aspetti metodologici della valutazione di impatto diretto.....	68
3.1.1	Le valutazioni sperimentali.....	75
3.1.2	Le valutazioni non-sperimentali.....	76

3.2	Principali aspetti della valutazione .....	79
3.2.1	Definizione degli obiettivi .....	80
3.2.2	Politics o policy?.....	81
3.3	Problemi e strategie di analisi della valutazione d’impatto .....	82
3.3.1	La stima del controfattuale.....	82
3.3.2	Le “minacce alla validità” della valutazione: l’effetto delle variabili omesse e quello di selezione .....	83
3.4	Modello di analisi .....	85
3.5	Le variabili oggetto di studio.....	86
3.5.1	Principali macro aggregati utilizzati nella riclassificazione del conto economico .....	86
3.6	Struttura dei dati .....	87
3.7	Il metodo Diff-In-Diffs.....	89
3.7.1	La modellizzazione del processo di selezione.....	93
3.8	La selezione del gruppo di controllo per mezzo dello “Statistical Matching” .....	95
3.9	Propensity score .....	97
3.10	Stimatori di Matching della ATT basati sul propensity score .....	99
3.10.1	Nearest Neighbor .....	100
3.10.2	Matching Kernel Matching .....	101
3.10.3	Stratification Matching .....	102
3.11	Modello Econometrico nell’analisi della TFP.....	103
3.12	Modelli Effetti fissi e Modelli Effetti Random .....	105
3.12.1	Il modello con effetti fissi.....	106
3.12.2	Il modello con effetti casuali .....	107
3.12.3	Meglio i modelli panel con effetti fissi o casuali? .....	107
<b>4</b>	<b>L’IMPATTO DEI SUSSIDI AL CAPITALE: UN’ ANALISI EMPIRICA DELLA L.488/92</b> .....	<b>111</b>
	Introduzione .....	111
4.1.	Analisi dei risultati: dati .....	112

## INDICE

4.2. Analisi dei risultati: stima del modello .....	117
4.2.1. Le elasticità e i rendimenti di scala .....	122
4.3 Metodologia non parametrica per il calcolo dell'effetto sui sussidi sulla TFP: tecniche di matching .....	123
4.4 Metodologia parametrica per il calcolo dell'effetto sui sussidi sulla TFP: Diff-In-Diffs.....	125
<b>CONCLUSIONI</b>	<b>127</b>
<b>BIBLIOGRAFIA</b>	<b>131</b>
<b>APPENDICE</b>	
<b>APPENDICE A</b>	<b>147</b>
<b>APPENDICE B</b>	<b>155</b>

# Introduzione

Produttività ed efficienza sono termini comunemente utilizzati per caratterizzare l'abilità di un'impresa nell'utilizzazione delle risorse, sia in ambito privato che pubblico.

Entrambi i concetti sono legati da una teoria della produzione che diventa essenziale per la determinazione dei criteri base con i quali confrontare i risultati dell'attività produttiva e i fattori impiegati per ottenerli. D'altronde, le imprese scelgono di produrre e di investire sulla base delle proprie prospettive di mercato e di costi dei fattori. Quest'ultimi possono essere influenzati dalle politiche dello Stato che fornisce incentivi e sussidi allo scopo di modificare le decisioni riguardanti l'allocazione e la crescita delle imprese. In questo caso le stesse imprese possono preferire di non collocarsi nell'equilibrio produttivo ottimo, massimizzando produttività ed efficienza, per poter invece utilizzare tali incentivi. In questo caso gli stessi incentivi potrebbero distorcere quindi l'allocazione delle risorse delle imprese che sono agevolate.

L'obiettivo di questo lavoro è quello di valutare attraverso metodologie parametriche e non parametriche se incentivi erogati dalla L. 488/92, la principale

politica regionale in Italia nelle regioni meridionali del paese nel periodo 1995-2004, hanno avuto o meno effetti sulla produttività totale dei fattori delle imprese agevolate.

La tesi si compone di quattro capitoli. Nel primo si è condotta una ricognizione rispetto ai principali lavori proposti in letteratura riguardanti la TFP e l'aiuto alle imprese attraverso incentivi al capitale e (in parte) dell'efficienza.

La stima della produttività totale dei fattori richiede di specificare una funzione di produzione, per questo il secondo capitolo contiene una breve rassegna della letteratura esistente circa le funzioni di produzione. L'attenzione è stata posta soprattutto sui modelli di tipo parametrico, che prevedono quindi la specificazione di una determinata forma funzionale relativa a variabili concernenti i fattori di produzione. Da questa si è ricavata la Total Factor Productivity utilizzata nell'analisi empirica che è la misura su cui viene valutata l'efficienza produttiva delle imprese.

Il campione di aziende è dato dal *merge* tra i dati della L.488 e i dati di bilancio della banca dati AIDA. La stima del modello viene sviluppato nell'ultima parte di questo terzo capitolo dove si approfondiscono diversi modelli per la stima della TFP; infine vengono descritti metodi non parametrici (tecniche di matching basate sul propensity score) e metodi parametrici (Diff-In-Diffs) per la valutazione dell'impatto dei sussidi al capitale.

Nel quarto, e ultimo, capitolo si descrive l'analisi empirica condotta. Nella prima parte del capitolo verranno illustrati i passaggi cruciali e i risultati ottenuti a partire dalla elaborazione del dataset. Nella seconda parte, invece, è descritto la stima del modello per la TFP e il confronto tra le metodologie parametriche e non parametriche per valutare se la politica ha influenzato o meno il livello di TFP delle imprese agevolate.

# Capitolo 1

## **ANALISI LETTERATURA SULL'IMPATTO DEGLI INCENTIVI SULL'EFFICIENZA DELLE IMPRESE**

### ***Introduzione***

Recentemente l'interesse fra la Comunità scientifica ed *i policy makers* sull'esplorazione dell'efficienza degli incentivi alle imprese si sta sviluppando velocemente. C'è una credenza diffusa che gli incentivi possono modificare il *dynamics* industriale al livello locale sia nel breve che nel lungo periodo. Questa credenza è stata usata come supporto per le politiche che puntano allo stimolo dello sviluppo delle sovvenzioni assegnanti al settore privato, entrambe ai livelli nazionali e regionali.

## ***1.1. Una misura di Produttività e di Efficienza***

Nella letteratura economica si è sempre cercato di sottolineare come l'efficienza produttiva rappresenti un'importante caratteristica di *performance* del produttore. Nel contesto della teoria neoclassica della produzione, il raggiungimento dell'efficienza produttiva è garantito dal concetto di massimizzazione del profitto, sulla base del quale, dati e prezzi esistenti sul mercato, il produttore è in grado di determinare quale delle tecniche possibili può essere adottata e questa scelta determina, simultaneamente, la quantità di output da produrre e le quantità dei fattori da impiegare.

Quando parliamo di misura dell'efficienza e della produttività ci riferiamo a due concetti che spesso sono usati come sinonimi, ma "produttività" ed "efficienza" afferiscono a concetti diversi, ma che vengono utilizzati per descrivere l'abilità di un'impresa nell'utilizzazione delle risorse. La misurazione, in realtà, può portare a trarre indicatori cui sono associate diverse interpretazioni. L'analisi dell'efficienza produttiva ha avuto un forte sviluppo a partire dagli anni ottanta, ma trae origine dalle definizioni di Koopmans (1951), Debreu (1951) e Farrell (1957).

Un indicatore di produttività può essere definito mediante il rapporto tra il risultato dell'attività produttiva e i fattori impiegati per ottenerlo (Nisticò e Prospetti, 1991). Si può così osservare che il risultato dell'attività produttiva o prodotto viene definito *output*, mentre i fattori utilizzati nel corso per il suo ottenimento sono denominati *input*.

Ad esempio Kuznets (1990) rilevò che il concetto di produttività si compone di tre elementi: gli output, gli input e il progresso tecnico attraverso cui i primi due sono connessi tra di loro.

Il richiamo quindi ad una teoria della produzione diventa necessaria per riuscire a determinare dei criteri in base ai quali confrontare i singoli output e i singoli input.

## 1.1 UNA MISURA DI PRODUTTIVITA' E DI EFFICIENZA

Generalizzando si può dire che la *misura dell'efficienza* di un'unità produttiva può essere definita per confronto tra il processo di produzione effettivamente realizzato e un altro processo, opportunamente scelto, corrispondente quindi al livello ottimo sia nello spazio che nel tempo

Molti autori, come ad esempio Gazzei, Lemmi e Viviani (1997), definiscono la produttività, in ambito soprattutto microeconomico, strettamente dipendente da quella dell'efficienza.

A tutto questo fa seguito un insieme di pensieri e di interpretazioni che ad esempio Schmookler (1952) definisce la produttività come “ un indice di efficienza”, mentre Kendrick (1961) definisce la produttività come il “cambiamento nell'efficienza produttiva”.

Zappa (1957) sottolinea che la “produttività è spesso nota come efficienza economica e che l'efficienza è conosciuta come produttività tecnica”, mentre Hatry (1982) sostiene che “l'efficienza riguarda il rapporto tra risorse utilizzate e quantità di beni o servizi prodotta. Le misure di efficienza si definiscono come la relazione tra la quantità di risorse impiegate ed il prodotto che ne deriva. Il rapporto tra prodotto e risorse impiegate viene chiamato produttività.

L'analisi del concetto di efficienza prendono avvio da Farrell (1957) che fu il primo a misurare l'efficienza produttiva empiricamente. Egli definì l'efficienza di costo e la scompose in due componenti: efficienza tecnica che corrisponde alla capacità di un'impresa di ottenere il massimo output dato un insieme di input; efficienza allocativa che riflette l'abilità ad utilizzare gli input in proporzioni ottimali, dati i rispettivi prezzi. Queste due misure sono poi combinate per ottenere la misura dell'efficienza totale.

Secondo Koopmans “*un produttore è tecnicamente efficiente se, e solo se, è impossibile produrre quantità superiori di un output senza ridurre la produzione di almeno un altro output o senza incrementare gli input*”. Dedreu e Shephard introdussero le funzioni di distanza come metodo per misurare le tecnologie multi-output. Ai fini dell'analisi dell'efficienza tecnica l'introduzione delle funzioni di distanza è importante, in quanto sono un metodo per valutare la distanza radiale tra un produttore e la frontiera, sia dal punto di vista output

orientato (massima espansione dell'output tenendo costanti gli input), che input orientato (utilizzo minimo degli input per un dato output).

### ***1.1.1.L'efficienza produttiva***

Definendo l'efficienza produttiva come l'abilità di una organizzazione di produrre un ben specificato livello di output al minimo costo, emerge la presenza di una componente tecnica, riflettente l'efficienza nella trasformazione fisica di input in output, e di una componente allocativa, che esprime l'efficienza di tipo economico nell'allocazione ottima dei fattori.

Nella formulazione neoclassica della teoria della produzione, si utilizza prevalentemente una rappresentazione analitica del processo produttivo, definito come processo di trasformazione regolato da uomini e descritto mediante una *n-pla* di numeri reali. La costruzione di un modello di scelta del processo "ottimo" è basata sulla determinazione di quale delle tecniche possibili, data la funzione di produzione, riesca ad ottenere il massimo profitto come obiettivo da cui l'imprenditore possa scegliere il giusto mix tra output da produrre e quantità di fattori da impiegare.

Un'altra fondamentale distinzione è tra metodi per dati *cross-section* (più unità produttive osservate in un unico istante temporale) e per dati panel (più unità produttive osservate in diversi istanti temporali).

Al di là delle critiche mosse da diverse interpretazioni del concetto di "media", è necessario sottolineare che la funzione di produzione stimata da dati *cross-section* mediante gli OLS. Per riprodurre i concetti teorici della funzione di produzione, e quindi, per ottenere un'adeguata determinazione della misura di efficienza, il termine di errore del modello econometrico dovrebbe essere caratterizzato da una forma distributiva unilaterale identificando così *funzioni frontiera di produzione*.

In tal modo si è individuato il processo produttivo ottimo per ciascuna impresa, che consente ad ogni unità di operare in una situazione di *efficienza produttiva*.

La frontiera di produzione può essere stimata mediante una funzione parametrica oppure può essere costruita mediante metodologie non parametriche che utilizzano la programmazione lineare.

In letteratura sono stati proposti diversi modelli di frontiere di produzione, una distinzione fondamentale è tra modelli deterministici e stocastici. Nella prima classe di modelli ogni scostamento dalla situazione di ottimo è dovuta all'inefficienza, quindi, nel caso output-orientato, nessun punto può giacere al di sopra della frontiera<sup>1</sup>. Nei modelli stocastici l'allontanamento dalla frontiera può essere causato anche da fattori aleatori, quindi alcuni punti possono giacere sopra la frontiera (output-orientato).

### ***1.1.2 La produttività***

*“The two main sources of economic growth in output are increases in the factors of production (the labour and capital devoted to production) and efficiency or productivity gains that enable an economy to produce more for the same amount of inputs.”*

(Baldwin, Harchaoui, Hosein and Maynard, 2000  
“Productivity: Concepts and Trends”  
Statistics Canada)

*“Productivity is commonly defined as a ratio of a volume measure of output to a volume measure of input use. While there is no disagreement on this general notion, a look at the productivity literature and its various applications reveals very quickly that there is neither a unique purpose for nor a single measure of productivity.”*

(Paul Schreyer OECD Statistics Directorate  
*OECD PRODUCTIVITY MANUAL*, 2001)

Partendo da questa definizione :

---

<sup>1</sup> Nel caso input-orientato nessun punto può giacere al di sotto della frontiera.

***“Productivity:  
A ratio of output to input.”***

(Atkinson, Banker, Kaplan and Young 1995, Management Accounting, p. 514)

La produttività in economia può essere definita in via di prima approssimazione come il rapporto tra la quantità di output e le quantità di uno o più input utilizzati per la sua produzione. Viene calcolata con riferimento alla singola impresa, all'industria o più in generale alla nazione.

Gli indici di produttività possono essere visti sotto due aspetti: la produzione e il consumo.

Dal lato dell'offerta essi assolvono le funzioni di indici di cambiamento tecnologico, efficienza produttiva e competitività, interna ed internazionale. Da tale punto di vista un aumento di produttività può essere visto genericamente come un risparmio di input in termini fisici, sia esso dovuto a progresso tecnico, miglioramento dell'efficienza produttiva ricollegabile ad economie di scala, riduzione della capacità produttiva inutilizzata o ad altro.

Dal lato del consumo sono indici di capacità, effettiva o potenziale, e quindi anche di benessere sociale.

Escluso il caso in cui si abbia un solo input di produzione ed un solo output, caso in cui è facilmente calcolabile la produttività in termini fisici, il calcolo della produttività porta necessariamente con sé il calcolo di indici delle quantità e dei particolari pesi da adottare nella ponderazione dei singoli output o input da aggregare.

A livello settoriale, si è soliti usare come misura della produttività dell'industria il valore aggiunto per lavoratore o ora lavorata. Il valore aggiunto viene di solito preferito alla produzione lorda settoriale perché la produttività calcolata sulla base della produzione lorda risulta influenzata dal livello di integrazione verticale. Questo viene fatto nonostante il calcolo del valore aggiunto reale, ottenuto con il metodo della doppia deflazione, comporti diversi problemi, sia a livello teorico che pratico.

Reder (1943) mette in risalto il fatto che solamente un punto sulla funzione di produzione si riferisce ad una situazione reale, ossia alla combinazione di fattori

utilizzati e al livello di output prodotto corrispondente alle situazione in cui l'unità produttiva massimizza il profitto. Da qui la diversità concettuale con la funzione di produzione Cobb-Douglas stimata a partire con dati osservati, che rappresenta combinazioni di fattori e livelli di output scelti da diverse unità produttive, come conseguenza del comportamento massimizzante.

Si distingue la funzione di produzione teorica per ciascuna unità produttiva, che definisce “*intrafirm*”, dalla funzione stimata mediante l'utilizzazione di dati provenienti da processi posti in essere da diverse unità di produzione denominata “*interfirm*” che ci permette di osservare che tale funzione è l'unione dei punti di equilibrio per diverse unità produttive e rappresenta uno strumento teorico legittimo che può essere utilizzato per la verifica empirica della teoria della distribuzione.

Nonostante sia molto diffuso, soprattutto per via della facilità con cui è possibile reperire i dati e calcolarlo, il rapporto valore aggiunto-lavoro soffre di diversi limiti come indice di produttività settoriale. In particolare, da un lato tiene conto di un solo fattore di produzione, il lavoro; dall'altro non riesce a rendere conto della generale interdipendenza delle industrie, ma considera i singoli settori economici implicitamente come modelli di produzione verticalmente integrati. Per questo, laddove possibile, gli vengono preferiti altri indici.

In particolare, soprattutto in ambito neoclassico, viene spesso utilizzata la *total factor productivity* (TFP), o *produttività totale dei fattori*. Si considerano cioè input le quantità di servizi forniti dai fattori produttivi per unità di tempo e, in aggiunta alle produttività dei singoli fattori, si calcola una produttività totale definita come il rapporto tra un indice di output e un indice di input, media ponderata degli indici di lavoro e capitale.

### ***1.2. L'efficienza nella teoria microeconomica***

I primi studi empirici sulla funzione di produzione ha visto uno sviluppo pressoché basato su serie storiche di dati relativi agli input impiegati e tali output prodotti da ciascuna unità produttiva.

La funzione Cobb-Douglas (1928) è applicata per la prima volta a dati *time series* per osservare la teoria della produttività marginale per la distribuzione del prodotto totale. Critiche furono fatte a questa poiché con il trascorrere del tempo si assiste ad uno sviluppo del livello tecnologico determinando e quindi affermando che le eventuali relazioni stimate tra il prodotto e i fattori produttivi, capitale e lavoro, potrebbero essere esclusivamente il risultato dell'agire del tempo.

Molti lavori poi seguirono utilizzando dati cross-section e in particolare Bronfenbrenner e Douglass (1939) per stimare la funzione di produzione utilizzando il metodo dei minimi quadrati ordinari (OLS) ipotizzando che tutte le deviazioni dalla funzione stimata siano dovute ad errori casuali di misurazione della variabile dipendente o determinate da operazioni aleatorie non incluse nel modello.

Marschack e Andrews (1944) spiegano e trovano che è possibile spiegare la funzione di produzione tramite modelli probabilistici.

### ***1.2.1 Approccio parametrico***

La breve sintesi dei metodi parametrici che verrà presentata si basa sull'assunzione che ogni produttore produce un solo output. Tale assunzione è realizzabile o perché effettivamente viene prodotto un solo output, o perché i diversi output sono aggregabili in un unico indicatore di sintesi. Verranno considerate valutazioni delle performance produttive mediante misure output orientate dell'efficienza tecnica. Non verranno imposte assunzioni relative al comportamento dei produttori e non verranno utilizzate informazioni relative ai prezzi.

### **1.2.1.1 Metodi per dati cross – section**

#### Frontiere di produzione deterministiche

Mediante queste procedure i parametri  $\beta$  vengono calcolati piuttosto che stimati, quindi non è possibile ottenere gli standard error e di conseguenza non si possono applicare procedure inferenziali.

La frontiera di produzione deterministica può essere stimata in due passi (Winsten, 1957). Nel primo passo con i minimi quadrati ordinari (OLS) si ottengono le stime corrette e consistenti dei coefficienti  $\beta_1 \dots \beta_n$ , mentre si hanno stime distorte dell'intercetta  $\beta_0$ . Nel secondo passo la stima di  $\beta_0$  viene corretta per assicurare che la frontiera stimata limiti i dati superiormente. Due diverse tecniche vengono in genere applicate:

- **Minimi quadrati ordinari corretti (COLS)**

Queste stime giacciono al di sotto della frontiera o sulla frontiera (almeno un produttore), e sono semplici da implementare, ma presentano diversi inconvenienti. La frontiera così stimata è parallela alla retta di regressione OLS, quindi la struttura della tecnologia “*best practice*” è la stessa della struttura di tecnologia produttiva “centrale”, mentre dovrebbe essere possibile ottenere una struttura della frontiera che differisce dalla struttura della tecnologia produttiva media.

- **Minimi quadrati modificati (MOLS)**

Un secondo metodo considera la stima mediante i minimi quadrati modificati (MOLS). Quest'ultima parte dall'assunzione che il disturbo segua una distribuzione specifica, per esempio una distribuzione di probabilità esponenziale o una semi normale. Anche in questo caso, il primo passo prevede la stima OLS

dei coefficienti  $\beta$ , seguita dalla stima dell'intercetta con il valore atteso della distribuzione specificata.

### Frontiere di produzione stocastiche

Modelli che prevedono l'esistenza dell'inefficienza tecnica, ma anche di shocks casuali non sotto il controllo dell'azienda, furono proposti simultaneamente da Aigner, Lovell e Schmidt (Aigner, Lovell e Schmidt, 1977) e da Meeusen e Van Den Broeck (Meeusen e Van Den Broeck, 1977).

#### **1.2.1.2 Metodi per dati panel**

La disponibilità di dati panel (osservazioni ripetute nel tempo per ogni produttore), consente di ovviare, almeno in parte, ad alcuni problemi che si presentano nel caso di dati cross-section. In particolare, osservazioni ripetute per un campione di produttori possono essere utilizzate in sostituzione di forti assunzioni circa la forma distributiva. Inoltre, non tutte le tecniche di stima per dati panel richiedono l'indipendenza tra la componente dell'errore attribuibile all'efficienza tecnica e i regressori. L'utilizzo di più osservazioni per lo stesso campione di produttori fornisce un maggiore numero di informazioni sull'efficienza tecnica che l'utilizzo di più produttori nell'analisi cross-section. La stima dell'efficienza tecnica è consistente per  $T \rightarrow \infty$  (dove T è il numero di osservazioni per ogni produttore). Nella pratica questo beneficio è solo parzialmente utilizzabile, in quanto la maggior parte dei data set è composta da un numero relativamente basso di osservazioni temporali.

Nell'analisi dell'efficienza tecnica con dati panel si può supporre che l'efficienza vari tra i produttori, ma rimanga costante nel tempo per ogni produttore, o che vari sia tra produttori che rispetto al tempo.

### ***1.2.2 Approccio Non parametrico***

L'approccio non parametrico richiede un numero limitato di ipotesi relative al processo di produzione, in quanto l'efficienza tecnica di una unità produttiva viene determinata mediante l'applicazione di tecniche di programmazione lineare, quindi senza dover specificare una particolare forma per la funzione di produzione. I metodi non parametrici più comunemente utilizzati sono il FDH (Free Disposal Hull) introdotto da Deprins, Simar e Tulkens (1984) e la DEA (Data Envelopment Analysis). Quest'ultima tecnica fu proposta da Charnes, Cooper e Rhodes (1978) e permette di misurare la massima contrazione (espansione) radiale degli input (output) compatibile con l'insieme delle possibilità produttive esistenti, o lo spazio di inviluppo. La principale distinzione rispetto al metodo FDH è che impone l'ipotesi di convessità delle combinazioni input-output, risulta quindi essere più restrittiva.

### ***1.2.3 Data Envelopment Analysis (DEA)***

La Data Envelopment Analysis (DEA), attribuita a Charnes, Cooper e Rhodes (1978), è una tecnica di programmazione lineare utilizzata per misurare l'efficienza relativa di una unità produttiva rispetto alle prestazioni di altre unità produttive che producono il medesimo insieme di output utilizzando lo stesso insieme di input. Nelle sua formulazione originaria la procedura massimizza un indice di produttività media per ciascuna unità produttiva costruito come rapporto fra una combinazione pesata di output e una combinazione pesata di input.

La DEA tiene conto del fatto che ogni impresa può valutare in modo diverso gli input e output adottando un sistema di pesi diverso per ogni unità produttiva. Il sistema di pesi di ogni impresa sarà quello ad essa più favorevole.

La DEA può essere input orientata o output orientata. Nel primo caso si definisce la frontiera cercando di ridurre gli input e tenendo fisso il livello dell'output, nel secondo invece si ricerca il massimo output dato un certo livello di input.

I due metodi danno gli stessi risultati se si considerano rendimenti di scala costanti, ma in caso di rendimenti variabili i risultati non coincidono più.

#### ***1.2.4 Nozioni teoriche e indagini empiriche sugli effetti delle sovvenzioni***

Accanto alla discussione tradizionale per le sovvenzioni, vale a dire *market failures*, ci sono almeno altre due punti per questa credenza. La prima sviluppa l'idea che in presenza di diminuzione dei ritorni di capitale, gli investimenti dovrebbero convergere, in linea di principio, verso le zone meno sviluppate con più alti ritorni sugli investimenti (Barro e Sala-i Martin, 1998).

Tuttavia, il processo di convergenza è stato troppo lento affinché possa essere socialmente accettabile e quindi si ha avuto la necessità di sovvenzionare le zone sottosviluppate. D'altra parte, la letteratura sulla localizzazione Fujita e Krugman (1995) ha sollecitato l'importanza delle politiche puntando sulle risorse mobili, quali i capitali, gli imprenditori ed ha specializzato il lavoro, che, idiosincronicamente unito con le risorse immobili locali ha permesso effetti positivi esterni da questa aggregazione. Per contro esistono argomenti evidenti, Bergstrom (1998), che a loro volta identificano nello spostamento degli investimenti riservati e nella generazione sia delle inefficienze tecniche che allocative le ragioni fondamentali del fallimento degli sforzi riguardanti gli incentivi alle imprese.

Accanto a queste nozioni teoriche, le indagini empiriche sugli effetti delle sovvenzioni sullo sviluppo ed il rendimento sembrano essere problematici nel migliore dei casi. Studi sull'effetto delle sovvenzioni in paesi asiatici sviluppati, per esempio, indica un effetto negativo delle sovvenzioni sul rendimento (Lee,

1996; Beason e Weinstein, 1996; Harris e Trainor, 2005). Bergstrom (1998), in uno studio sulle sovvenzioni regionali per la Svezia, trova gli effetti contraddittori delle sovvenzioni sul rendimento. Nel breve periodo il rendimento sembra aumentare per le ditte sovvenzionate, mentre dopo un periodo di tempo la tendenza si inverte e più un'impresa ha ricevuto incentivi, meno risulta l'aumento in termini di TFP. Similmente, su dati italiani, Bagella e Becchetti (1998); Pellegrini e Centra (2006); Bronzini e De Blasio (2006) trovano un insieme di risultati parzialmente contraddittori. Nel breve periodo, le sovvenzioni sembrano causare un più alto livello di indebitamento per le imprese che le ricevono con costi non-decrescenti di indebitamento, apparentemente non c'è una allocazione sbagliata di risorse finanziarie. Nel lungo periodo, le imprese sovvenzionate dimostrano un più basso livello di produttività confrontati con quelli non-sovvenzionati, confermando l'idea che questo "Shock" esogeno è riassorbito dal mercato.

Attraverso un esercizio di development accounting, metodologia ispirata al tradizionale approccio di contabilità della crescita, riusciamo ad analizzare l'importanza relativa dell'accumulazione di fattori produttivi e della componente 'residuale' (valutata come *proxy* dell'efficienza o Total Factor Productivity) per l'interpretazione di tale divario.

Inoltre, come viene sottolineato dalla maggior parte degli studi recenti, la quota della variabilità cross country del prodotto per addetto che può essere attribuita alla TFP risulta essere elevata e crescente. Una dettagliata analisi di sensitività sembra confermare la sostanziale solidità di questi risultati rispetto ai potenziali problemi legati alla scelta delle fonti statistiche, alla misurazione dei fenomeni in oggetto e alla specificazione delle variabili.

Nella letteratura economica spesso si considera utile, come passo preliminare, un esercizio di contabilità della crescita' o growth accounting, attraverso il quale si può decomporre l'incremento del prodotto reale pro capite nelle sue componenti associate alla crescita della disponibilità di fattori (capitale e lavoro) e al 'residuo di Solow'. Tale 'residuo' viene comunemente definito 'produttività congiunta dei fattori' o Total Factor Productivity (più brevemente TFP). Barro (1999) sostiene che tale esercizio di growth accounting è particolarmente utile se

c'è la convinzione che le determinanti fondamentali alla base della crescita economica sono sostanzialmente differenti da quelle che regolano la dinamica della tecnologia. Nello stesso articolo Barro ricorda che le basi teoriche della contabilità della crescita sono da ricercare nei contributi di Solow (1957), Kendrick (1961) e Denison (1962), e mostra come sia possibile sviluppare alcune versioni della 'decomposizione' in modo tale da essere collegate a modelli di crescita endogena. Boggio e Seravalli (2002) ricordano che già nel lavoro di Abramovitz (1956) si trova un'applicazione importante della contabilità della crescita per gli Stati Uniti. Gli stessi autori osservano poi che la letteratura si è concentrata sulle caratteristiche e sulla interpretazione del 'residuo', dividendosi fra sostenitori del progresso tecnologico 'incorporato' nelle dotazioni di fattori e sostenitori di modelli non tradizionali (ruolo delle economie di scala, di esternalità, cambiamenti strutturali, ecc.). Più recentemente, all'interno di alcuni studi empirici è stato proposto un approccio simile alla contabilità della crescita per valutare il contributo specifico dell'accumulazione di capitale e del 'residuo', il cosiddetto development accounting .

La scelta è quella di calcolare un valore stimato per il livello delle dotazioni di capitale fisico e di capitale umano per un dato periodo (anno o serie di anni), e valutare quanto tale valore possa 'spiegare' (in base ad una teorica funzione di produzione aggregata) il livello del prodotto reale per addetto e la sua varianza cross country, anziché il suo tasso di crescita. Secondo Caselli (2003), il contributo della TFP calcolato con l'approccio di development accounting può essere interpretato come la differente 'efficienza' nell'utilizzo degli stessi fattori nei vari paesi. Il risultato più rilevante di Klenow e Rodriguez-Clare (1997) è che circa il 60% della variabilità cross country nel livello del prodotto per lavoratore è da attribuire alla componente relativa alla TFP.

Un risultato analogo è ottenuto da Hall e Jones (1999), i quali osservano che, in primo luogo, la correlazione fra le differenze cross country nella TFP calcolata come residuo e il livello del prodotto reale per addetto è significativa ed elevata. Secondariamente, essi mostrano che il contributo della componente residuale (cioè della TFP) è sempre il più elevato per la spiegazione dell'elevato differenziale esistente fra i paesi più ricchi e quelli più poveri . Anche in Easterly

e Levine (2001) viene esaminato, con un approccio molto simile, il rapporto fra accumulazione di fattori e divario nel livello del prodotto reale pro capite.

Non mancano lavori di carattere empirico che utilizzano varie formulazioni di questo approccio. Ad esempio nel lavoro di Bosworth e Collins (2003) si può trovare una combinazione di risultati empirici ottenuti mediante un esercizio di contabilità della crescita e un'analisi di regressione cross country. Questi autori ritengono che l'interpretazione del risultato residuo del growth accounting, cioè la TFP, come indicatore di cambiamento tecnologico è criticabile, in quanto ciò che risulta 'residuale' può essere determinato da molti altri fattori.

Un contributo simile è dato Senhadji (1999), che introduce un passaggio preliminare all'esercizio di growth accounting: una stima delle quote fattoriali sul prodotto nazionali mediante l'analisi di regressione (con tecniche di cointegrazione).

### ***1.2.5 Lo sviluppo recente e crescita degli incentivi in R&S***

Gli anni recenti, tuttavia, inoltre hanno testimoniato uno sviluppo e una crescita di incentivi per l'investimento e la R & S. Nonostante questi cambiamenti istituzionali ed importanza di politica è sufficiente dire che 70% del Programma Operativo Nazionale 2000-2006 è stato distribuito per incentivare le piccole e medie, PMI, la valutazione dell'effetto di queste ha avuto campo nel mercato delle politiche. Come è messo in risalto nelle surveys di Heckman et al.(1999) e di Blundell e Costa-Dias (2000), l'elemento cruciale nell' *impact evaluation* è la specificazione dell'ipotesi controfattuale che potrebbe essere considerata prendendo in esame le imprese sovvenzionate in assenza di intervento. I problemi fondamentali, a tale riguardo, riguardano la omissione di variabile *bias* (collegato alla difficoltà di misurazione degli effetti di intervento esclusivamente da altri fattori) ed effetti dovuti a *selezione bias* (collegata al fatto che le imprese sovvenzionate sono selezionate non a caso, ma in base ad alcuni test di verifica e

possono avere alcune caratteristiche che alterano il loro confronto con le imprese non-sovvenzionate). La soluzione a questi problemi richiede l'approvazione di procedure empiriche adatte (utilizzate non sempre negli studi esistenti), l'utilizzazione dei gruppi di dati che contengono le informazioni sufficienti e ricche sulle caratteristiche delle stesse imprese. A tale riguardo le indagini promosse da Capitalia (precedentemente Mediocredito Centrale) sono di interesse particolare perché contengono, particolarmente per le piccole e ditte medie, le informazioni che al contrario non sono disponibili nelle basi di dati italiane. Per capire meglio come questi dati possono essere utilizzati per migliorare la base empirica e i lavori che già esistono in Italia, è utile brevemente rivalutare le giustificazioni principali di questi interventi di politica, così come i loro problemi potenziali (Carlucci e Pellegrini, 2001). La motivazione di intervento pubblico conta su due tipi differenti di market failure.

Nel primo caso, alcuni tipi di mercati imperfetti (particolarmente nel credito e nei mercati dei capitali) impediscono una valutazione corretta dei progetti economici ed implicano un livello eccessivamente basso dell'investimento.

Nel secondo caso, l'esistenza di non adeguate externalities nell'investimento, R&D, o progetti simili richiedono l'intervento pubblico per queste attività. In entrambi i casi, l'aiuto potrebbe produrre un'effetto netto basso (si possono sovvenzionare dei progetti finanziati anche se le imprese non ricevono degli incentivi) e di efficienza (se ci sono mercati imperfetti, si potrebbe sovvenzionare progetti erroneamente che il mercato rifiuterebbe).

Il risultato delle evidenze empiriche (corretto) sull'effetto netto di una politica segue naturalmente dalla soluzione dei problemi sopraccennati di specificare l'ipotesi controfattuale. D'altra parte, la valutazione corretta degli effetti di efficienza probabilmente richiede un più originale sforzo nella ricerca, concentrandosi sul calcolo e sull'utilizzazione di misure di performance basate sulla cosiddetta analisi di frontiera di efficienza. Effettivamente, una tipica caratteristica dell'evidenza italiana attuale è che effetti positivi di incentivo sugli investimenti (e sull'occupazione) sono accompagnati da insignificanti (o persino negativi)

effetti su profitto delle imprese, come misurato dagli indicatori di bilancio (tipicamente il ROI).

La distinzione fra i concetti di efficienza tecnica e allocativa non solo permette alcune conclusioni sul risultato dell'aiuto pubblico, ma anche di capire le conseguenze per imprese in una analisi dettagliata.

### ***1.3. Approccio Macroeconomico: la Total Factor Productivity (TFP) nell'analisi della produttività***

Una valutazione particolare meritano i risultati di Baier, Dwyer e Tamura (2002), in quanto il set di dati utilizzato comprende più di cento paesi per un periodo di tempo molto lungo. Anche se si tratta di un dataset non completamente comparabile con le statistiche tradizionalmente utilizzate, la proposta di Baier, Dwyer e Tamura (2002) è interessante per due motivi. Primo, viene evidenziata l'importanza di analizzare l'importanza dell'accumulazione fattoriale e della TFP per differenti (e più lunghi) periodi storici. In particolare, è opinione degli autori che lo studio degli ultimi 25 anni potrebbe condurre a risultati molto diversi dall'analisi degli ultimi 40 anni. Secondo, si sviluppa un approccio per analizzare la variabilità cross country dei tassi di crescita del prodotto per addetto.

Tra le conclusioni di Baier, Dwyer e Tamura è degno di nota che il la TFP sia in grado di 'spiegare' una quota della varianza dei suddetti tassi compresa fra un minimo del 68% ed un massimo pari all'84%.

A partire dal contributo di Robert Solow (1957), il calcolo della TFP venne messo in relazione alla funzione di produzione e alla teoria neoclassica della crescita. In particolare, Solow dimostrò come il tasso di crescita della TFP calcolato come la differenza fra l'indice di Divisia dell'output e l'indice di Divisia degli input risulta uguale al progresso tecnico Hicks-neutral, scorporato dai fattori di produzione e che lascia invariati i rapporti fra le produttività marginali dei singoli fattori.

Dopo diversi studi applicati alla fine degli anni 60 e nella prima metà degli anni 70, negli anni 80 iniziò negli Stati Uniti una misurazione sistematica a livello settoriale della TFP, sotto la denominazione di MFP (Multifactor productivity), da parte del National Bureau of Economic Research (NBER) (cfr. ad es. Gullickson & Harper, 1987). Negli anni 90 gli studi sulla TFP si sono moltiplicati. A questi si sono aggiunti gli studi con approccio econometrico all'analisi della produttività, come ad esempio la Stochastic Frontier Analysis (SFA) (Battese & Coelli, 1992, 1995; Coelli et al., 2005), e quelli che applicano la programmazione lineare per la stima della funzione di produzione, come la Data Envelopment Analysis (DEA) (Cooper et al., 2000).

I suddetti approcci sono comunque da considerare in larga parte complementari e non sostitutivi alle analisi non-parametriche della TFP. Nel 2001 l'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) ha pubblicato un manuale sulle misure di produttività che indirizza gli uffici statistici nazionali e consiglia l'utilizzo della MFP basata sulla produzione lorda, chiamata anche KLEMS (la sigla sta per Kapital, Labour, Energy, Materials e Services) per la stima non parametrica dei tassi di variazione della produttività aggregata.

Recentemente l'Unione Europea ha finanziato un ambizioso progetto, l'EU KLEMS Project, finalizzato alla creazione di un database di serie storiche di misure di produttività settoriale basate sulla TFP. Essendo ormai largamente condiviso e accettato l'uso di misure di produttività totale dei fattori, gli sforzi degli ultimi anni sembrano essere nel senso di:

1. elaborare metodi condivisi di misurazione dello stock di capitale, dei servizi da capitale e del loro costo;
2. migliorare gli indici di quantità per tenere conto dei miglioramenti qualitativi dei beni attraverso la creazione di indici edonistici di prezzo (Triplet, 2004).

Il residuo di Solow è stato utilizzato nella letteratura macroeconomica come misura degli shock tecnologici. Sin dai lavori di Kydland e Prescott (1982) ed in particolare di Prescott (1986) si è sviluppata una controversia sul ciclo economico riguardante la dimensione degli shock tecnologici come causa delle fluttuazioni dell'economia statunitense.

Una crescente letteratura ha criticato questa impostazione sottolineando come il residuo di Solow in realtà rifletta una serie di elementi non tutti correlati con gli shock tecnologici. Mercati non perfettamente concorrenziali, rendimenti di scala crescenti, andamento dei prezzi energetici e varie esternalità sono alcuni fenomeni spesso richiamati dalla letteratura empirica e teorica sul residuo di Solow (tra gli altri, Summers (1986); Eichenbaum (1991); Hall (1988); Burnside (1996); Burnside e Eichenbaum (1994); Finn (1995); Evans (1992); Basu (1995); Caballero e Lyons (1992); Bernanke e Parkinson (1991)).

In molti lavori, la struttura della *growth accounting* indica alcune correlazioni tra una variabile di *policy* e un *output growth* che potrebbe dare due effetti: il primo che la *government policy* può influenzare l'accumulazione di capitale e perciò l'*output growth*. Secondo la *government policy* potrebbe avere effetti sull'*output growth* da influenzare la crescita della TFP.

Nel lavoro di Jong-Wha Lee (1995) si arriva ad un importante risultato, infatti i coefficienti stimati nel *tax incentives* diventano insignificanti fino a che i coefficienti nelle barriere non tariffarie sono ancora fortemente significativi. Il fatto che i *tax incentives* hanno un non significativo effetto nell'*output growth* controllando l'accumulazione di capitale, questo ci suggerisce che essi agiscono sull'accumulazione di capitale fisico, ma non sulla crescita della TFP. In maniera opposta, la forte correlazione tra *import restrictions* e *output growth*, sempre dopo aver controllato l'accumulazione di capitale, mostra significativi effetti di *import restrictions* nella crescita di TFP. In altre parole, dalle stime risulta che *tax incentives* hanno un effetto positivo sul tasso *output growth* stimolando accumulazione di capitale, ma non interessando la crescita della TFP; dall'altra parte si osserva protezionismo specialmente con barriere non tariffarie che fa decrescere accumulazione di capitale fisico e la TFP.

#### ***1.4. Limiti nella 'misurabilità' del prodotto e stime della TFP: un esercizio basato sui dati per l'industria manifatturiera***

All'interno del dibattito "TFP vs factor accumulation" non viene sempre affrontata una delle problematiche più rilevanti: i dati di contabilità nazionale (sui quali è basato il dato per il prodotto reale) per settori quali le costruzioni, la maggior parte dei servizi e il settore pubblico non sono basati su misurazioni effettive. Brevemente, si nota che per gli anni Settanta il contributo della TFP per l'industria è ancor più elevato che nell'aggregato (oltre 63%), e nel passaggio agli anni Novanta tale importanza non si riduce. Ciò deve comunque essere letto anche alla luce del fatto che per molti paesi non OCSE il ventennio è stato caratterizzato da un notevole incremento della produttività grezza, in seguito all'industrializzazione. Tuttavia, il controllo sembra confermare la forte difficoltà che si registra nella spiegazione di tale crescita semplicemente sulla base dell'accumulazione di fattori produttivi.

#### ***1.5. Differenze nella qualità dell'istruzione e nel contributo del capitale umano***

Nonostante i progressi raggiunti dalle metodologie statistiche per fornire un quadro sempre più accurato e preciso delle caratteristiche del capitale umano, il dibattito relativo alla sua corretta misurazione e specificazione non sembra aver proposto soluzioni generalmente accettate. In particolare, si può notare come la letteratura empirica, pur trovando un vasto consenso circa i molti aspetti legati all'investimento per accrescere il livello di preparazione dei lavoratori (spesa per istruzione, formazione continua, ecc.), non abbia ancora proposto misure

abbastanza convincenti per poter valutare il contributo derivante dalla qualità del fattore lavoro.

Dal lavoro di Woessman (2003) sembra emergere che la correzione della variabile di capitale umano in un'applicazione di contabilità della crescita produce risultati notevoli. Il ruolo chiave della TFP nella spiegazione della variabilità crosscountry si riduce notevolmente, mentre si osserva un incremento rilevante nel ruolo della variabile di capitale umano 'corretta'. Per tener conto di questi risultati si è provato ad adattare i dati (disponibili solo per gli anni '90 e 'attribuiti' a molti paesi in base alle medie regionali per l'indice di qualità) per la variabile proposta da Woessman (2003). Sostanzialmente si registra un contributo più elevato per il capitale umano, ma la rilevanza della componente TFP resta comunque non marginale.

Sembrano emergere almeno due aspetti importanti per le future ricerche orientate allo studio delle differenze nei livelli di sviluppo. In primo luogo, il ruolo della TFP è risultato ancor più elevato quando si è scelto di utilizzare i dati relativi al settore manifatturiero. Ciò potrebbe segnalare l'esigenza di abbandonare l'ipotesi legata al sistema produttivo aggregato e analizzare le differenze di efficienza a livello settoriale. In secondo luogo, attraverso un miglioramento nella specificazione e nella misurazione della qualità del capitale umano è sicuramente possibile ridurre la porzione di variabilità cross country che attualmente viene ad essere imputata al 'residuo'.

Nei passi successivi della ricerca si cercherà di determinare quali sono le caratteristiche dei fattori strutturali per i sistemi economici considerati che possono chiarire il ruolo e la dinamica della componente TFP.

### ***1.6. Gli effetti regionali degli incentivi nella performance delle imprese: studi empirici***

Gli effetti economici di *capital subsidies* sono usualmente esaminati da un punto di vista di politiche pubbliche. Si può osservare che tra diversi studi condotti si è

cercato di esaminare gli effetti di *regional capital subsidies* nelle performance delle imprese e specialmente quattro tipi di performance:

- efficiency;
- profitability;
- leverage;
- growth.

Malgrado il fatto che la microeconomia delle sovvenzioni di capitali sia ben nota in maniera automatica o in modo discrezionale ci sono pochissimi lavori empirici che hanno studiato o studiano gli effetti delle sovvenzioni nelle performance delle imprese e/o dei settori.

A livello di settore, Beason and Weinstein (1996) hanno investigato sull'uso di diversi strumenti di *industrial policy*, includendo anche i *capital subsidies*, in Giappone, e contrariamente a quanto convenzionalmente atteso, essi hanno trovato che una sproporzione tra l'ammontare di bassa crescita dei settori e settori con decrescenti ritorni di capitali.

A livello di imprese, Van Tongeren (1998) ha usato modelli di simulazione micro per esaminare la risposta della società industriale negli investimenti nell'Olanda, e anche ha ricercato gli effetti a livello macro-economici.

La *micro simulation model*, come viene definita, rendere diretta la stima degli effetti dei cambiamenti di *profitability*. Uno studio successivo, è stato condotto da Bergstrom (2000) che ha esaminato gli effetti di *capital subsidies* nella crescita della TFP delle imprese della Svezia tra il 1987 e il 1993. L'autore trova che la sovvenzione era positivamente correlata con la crescita del valore aggiunto nel primo anno, dopo che le *subsidies* furono erogate, ma successivamente la previsione di *capital subsidies* appare essere negativamente correlato con la TFP *growth*. Il risultato implica che, sempre se il mercato giustifica *subsidies*, non c'è certezza che le risorse siano efficientemente allocate.

Il modello di Bergstrom (2000) era anche applicato ad un campione di imprese in Danimarca rilevando che, in generale, non c'è correlazione tra la direzione degli incentivi e la crescita della produttività.

### **1.7. Critiche alla TFP**

Nonostante l'utilizzo della TFP sia ampiamente diffuso e accettato, le critiche al suo utilizzo sono state numerose ed in molti hanno evidenziato, nel corso del tempo, limiti ed errori concettuali insiti nell'indicatore.

Già Abramovitz (1956) notava come in realtà il residuo così calcolato era alla fine il risultato non solo del cambiamento tecnologico e del miglioramento nell'efficienza produttiva, ma anche di una serie di possibili errori, come quelli di misura, quelli derivanti da aggregazione e quelli di errata specificazione del modello. Il residuo di Solow risultava così, a conti fatti, essere solo la "misura della nostra ignoranza" ("the measure of our ignorance").

Lo stesso Solow (1987) notava con meraviglia come la TFP non registrasse in alcun modo la rivoluzione digitale: la TFP non aveva registrato tassi di crescita significativi in corrispondenza di nessuna delle rivoluzioni tecnologiche che si erano succedute nel corso degli anni, compresa quella della scoperta e della diffusione dell'energia elettrica.

Negli anni 60, dato il collegamento esplicito posto da Solow (1957) con la funzione di produzione aggregata e con l'ipotesi di progresso tecnico neutrale à la Hicks, la TFP venne investita, in quella che è successivamente divenuta famosa come la Cambridge capital controversy, dalle critiche che colpirono queste ultime. In particolare, da un lato, si negava la possibilità di utilizzare misure aggregate del capitale e la tendenza all'uguaglianza tra tasso di rendimento del capitale e produttività marginale dello stesso, e tutto ciò minava alle basi la funzione di produzione aggregata neoclassica formulata in termini di lavoro e capitale; dall'altro, si criticava la concezione del progresso tecnico, propria del primo Hicks (1964) e dei neoclassici, che distingueva spostamenti lungo la funzione di produzione da spostamenti della funzione stessa.

Di diversa natura sono state le critiche di Read (1968), Rymes (1971, 1972, 1983), Cas & Rymes (1991) e Durand (1996). In particolare Thomas K. Rymes, nei suoi lavori pionieristici, mise in evidenza come l'errore di trattare il capitale come un fattore produttivo scarso, al pari di lavoro e terra, assunzione implicita nella concezione di progresso tecnico Hicks-Meade-Solow, invece di un bene riproducibile nella riproduzione del quale si trasferiscono pertanto gli incrementi di produttività conseguiti dal sistema, finisca per condurre a risultati a volte paradossali. Tra questi, il fatto che la distinzione tra progresso tecnico incorporato e scorporato, l'unico catturato dalla TFP, riposi in ultima istanza sulla possibilità di "incorporarlo" nel capitale a costo zero. La conseguenza è che ridefinizioni "statiche" di cosa è e cosa non è capitale inevitabilmente modificano il tasso stimato della produttività.

Si è osservato fin qui quanto è stato proposto in letteratura negli studi e quanto si è cercato di studiare con l'applicazione di metodi sempre più raffinati lo studio della TFP. Si è parlato di metodi parametrici e non parametrici, si è discusso sui diversi approcci macro e micro. La nostra attenzione è quella di andare a cercare la misura di TFP tale per cui la misura dell'efficienza possa essere misurata e confrontata. Sarà possibile determinare questa misura creando un modello che ci consentirà di fare delle ipotesi e di andare a misura l'efficienza delle imprese manifatturiere cercando di proporre, con tecniche affinate, un valore di un modello di stima degli effetti e considerando come tecnica di confronto tecniche di *matching*. Questi metodi rispondono molto bene al problema di andare a dare una stima degli effetti di programma mediante la stratificazione delle unità beneficiarie e delle unità escluse sulla base delle caratteristiche che condizionano il processo di selezione. L'obiettivo, cioè, sarà quello di minimizzare le differenze tra i due gruppi esistenti, riducendo in tal modo l'effetto della *selection bias*.

## Capitolo 2

### **MISURE DI EFFICIENZA DELL'IMPRESA A LIVELLO MICRO**

#### *Introduzione*

La stima della produttività totale dei fattori richiede di specificare una funzione di produzione. L'attenzione è stata posta soprattutto sui modelli di tipo parametrico, che prevedono quindi la specificazione di una determinata forma funzionale relativa a variabili concernenti i fattori di produzione. Da questa si è ricavata la Total Factor Productivity utilizzata nell'analisi empirica che è la misura su cui viene valutata l'efficienza produttiva delle imprese.

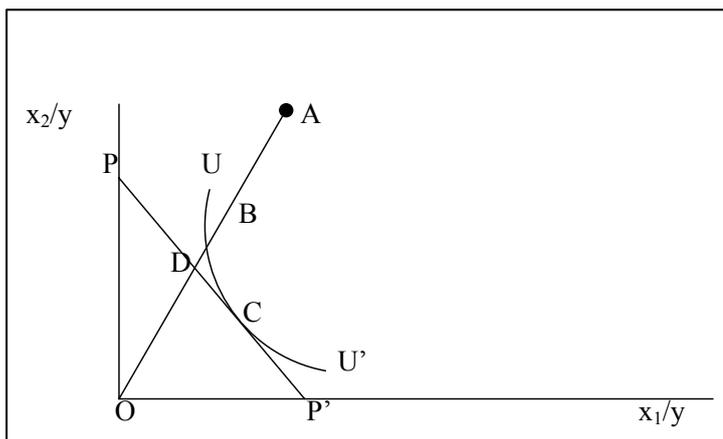
## 2.1 Sviluppo dell'analisi tecnica

Come già detto nel precedente capitolo, Farrell (1957) fu il primo a misurare l'efficienza produttiva empiricamente. Egli definì l'efficienza di costo e la scompose in due componenti: efficienza tecnica che corrisponde alla capacità di un'impresa di ottenere il massimo output dato un insieme di input; efficienza allocativa che riflette l'abilità ad utilizzare gli input in proporzioni ottimali, dati i rispettivi prezzi. Queste due misure sono poi combinate per ottenere la misura dell'efficienza totale.

L'analisi di Farrell sulla misura di efficienza input-orientata può essere illustrata considerando imprese che utilizzano due input ( $x_1, x_2$ ) per produrre un output ( $y$ ). Si assumono rendimenti di scala costanti.

La conoscenza dell'isoquanto unitario ( $UU'$ ) consente la misurazione dell'efficienza tecnica.

Figura 2.1.1 - Misura input orientata dell'efficienza



Si consideri la Figura 2.1.1. Se un'impresa utilizza una quantità di input definita dal punto A per produrre una unità di output, l'inefficienza tecnica può essere rappresentata da BA che è l'ammontare di cui vengono ridotti gli input senza

## 2.1 SVILUPPO DELL'ANALISI TECNICA

ridurre l'output. Il rapporto BA/OA rappresenta la percentuale di cui tutti gli input possono essere ridotti. L'efficienza tecnica (TE) di un'impresa è misurata dal rapporto tra gli input teoricamente sufficienti per produrre l'output e quelli effettivamente utilizzati:

$$TE_1 = OB/OA$$

ed è uguale a 1-BA/OA.  $TE_1$  è uguale a 1 se l'impresa è efficiente e a 0 se è inefficiente.

La conoscenza del rapporto tra i prezzi degli input (PP') consente il calcolo dell'efficienza allocativa, che per il punto A è data da:

$$AE_1 = OD/OB$$

La distanza DB rappresenta la riduzione dei costi di produzione che si verificherebbe se l'impresa occupasse il punto allocativamente (e tecnicamente) efficiente C invece che il punto tecnicamente ma non allocativamente efficiente B.

L'efficienza totale è data da :

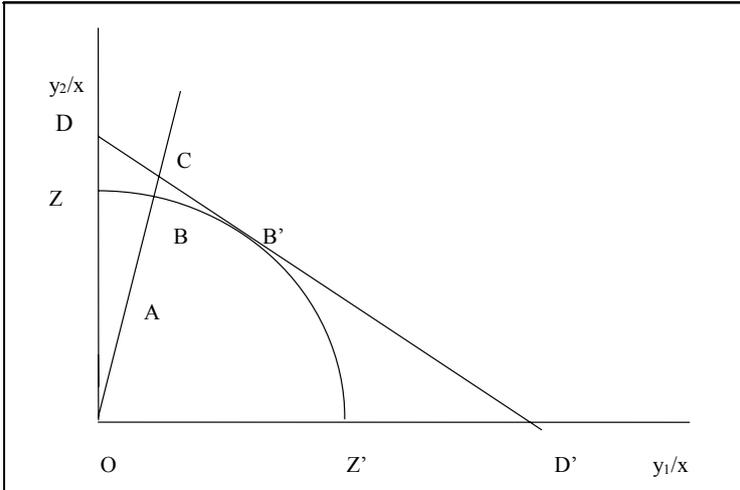
$$EE_1 = OD/OA$$

Il prodotto tra efficienza allocativa e tecnica dà l'efficienza totale:

$$TE_1 * AE_1 = (OB/OA) * (OD/OB) = OD/OA = EE_1$$

Una misura output-orientata è rappresentata in Figura 2.1.2

Figura 2.1.2 - Misura output orientata dell'efficienza



Si considerano due output ( $y_1$  e  $y_2$ ) ed un solo input ( $x$ ). Ipotizzando rendimenti di scala costanti  $ZZ'$  rappresenta la curva delle possibili unità produttive.

L'impresa caratterizzata dal punto  $A$  è inefficiente, essa giace sotto la curva perché  $ZZ'$  rappresenta il limite superiore della possibile produzione.  $AB$  corrisponde all'ammontare di output che può essere incrementato senza aggiunta di input, è quindi l'inefficienza tecnica.

La misura output-orientata di efficienza tecnica è data da:

$$TE_0 = OA/OB$$

Se si dispone di informazioni sui prezzi si può tracciare  $DD'$  e definire l'efficienza allocativa:

$$AE_0 = OB/OC$$

L'efficienza totale è data dal prodotto:

$$EE_0 = (OA/OC) = (OA/OB) * (OB/OC) = TE_0 * AE_0$$

Le misure appena viste sono misure radiali, quindi le relative proporzioni di input ed output sono costanti. Un vantaggio notevole è dato dalla loro invarianza rispetto all'unità di misura.

L'assunzione di rendimenti di scala costanti implica che la misura input-orientata e la misura output-orientata coincidano, ciò non avviene nel caso di rendimenti variabili.

Il luogo geometrico dei punti corrispondenti ad imprese efficienti può essere rappresentato da una funzione di produzione o frontiera di produzione. L'inefficienza delle imprese è definita dalla distanza che le separa dalla frontiera.

Queste misure di efficienza presuppongono che la funzione di produzione delle imprese efficienti sia nota. Nella realtà questo non avviene, quindi viene determinata a partire dal campione.

La frontiera di produzione può essere stimata mediante una funzione parametrica oppure può essere costruita mediante metodologie non parametriche che utilizzano la programmazione lineare.

In letteratura sono stati proposti diversi modelli di frontiere di produzione, una distinzione fondamentale è tra modelli deterministici e stocastici. Nella prima classe di modelli ogni scostamento dalla situazione di ottimo è dovuta all'inefficienza, quindi, nel caso output-orientato, nessun punto può giacere al di sopra della frontiera<sup>1</sup>. Nei modelli stocastici l'allontanamento dalla frontiera può essere causato anche da fattori aleatori, quindi alcuni punti possono giacere sopra la frontiera (output-orientato).

Un'altra fondamentale distinzione è tra metodi per dati cross-section (più unità produttive osservate in un unico istante temporale) e per dati panel (più unità produttive osservate in diversi istanti temporali).

### ***2.2 Metodi Parametrici***

La breve sintesi dei metodi parametrici che verrà presentata si basa sull'assunzione che ogni produttore produce un solo output. Tale assunzione è realizzabile o perché effettivamente viene prodotto un solo output, o perché i diversi output sono aggregabili in un unico indicatore di sintesi. Verranno

---

<sup>1</sup> Nel caso input-orientato nessun punto può giacere al di sotto della frontiera.

considerate valutazioni delle performance produttive mediante misure output orientate dell'efficienza tecnica. Non verranno imposte assunzioni relative al comportamento dei produttori e non verranno utilizzate informazioni relative ai prezzi.

### ***2.2.1 Metodi per dati cross – section***

#### *Frontiere di produzione deterministiche*

La frontiera di produzione deterministica per ogni produttore  $i=1..I$  disponendo di dati cross section relativi a  $N$  input utilizzati per la produzione di un solo output  $Y$  si può scrivere come

$$y_i = f(x_i; \beta) \times TE_i, \quad (2.2.1.1)$$

dove  $y_i$  è uno scalare e corrisponde all'output dell'  $i$ -esimo produttore ( $i=1,..,I$ ),  $x_i$  è il vettore di  $N$  input utilizzati dal produttore  $i$ -esimo,  $f(x_i; \beta)$  è la frontiera di produzione e  $\beta$  è il vettore di parametri tecnologici da stimare. L'efficienza tecnica (output orientata) del produttore  $i$ -esimo è definita come rapporto tra l'output osservato ed il massimo output realizzabile:

$$TE_i = \frac{y_i}{f(x_i; \beta)}, \quad (2.2.1.2)$$

Se  $y_i$  è il massimo output realizzabile il valore di  $TE_i$  sarà uguale a 1, in caso contrario sarà  $TE_i < 1$ , quest'ultimo misura la differenza tra output realizzato e l'output possibile.

In questa specificazione  $f(x_i; \beta)$  è deterministica, quindi ogni spostamento dalla stessa è attribuibile a inefficienza, ignorando che l'output può subire modifiche a causa di shocks casuali non controllabili dall'azienda e che possono essere presenti errori di misura. La ( 2.2.1.1) si può riscrivere come:

$$y_i = f(x_i; \beta) \times \exp\{-u_i\}, \quad (2.2.1.3)$$

dove  $TE_i = \exp\{-u_i\}$ . Essendo richiesto che  $TE_i \leq 1$  avremo che  $u_i \geq 0$ . Se si assume che la funzione sia una Cobb-Douglas log-lineare si otterrà la specificazione:

$$\ln y_i = \beta_0 + \sum_n \beta_n \ln x_{ni} - u_i \quad (2.2.1.4)$$

dove  $u_i \geq 0$  garantisce che  $y_i \leq f(x_i; \beta)$ . L'obiettivo è stimare i parametri  $\beta$ , che descrivono la struttura della frontiera di produzione, e gli  $u_i$ , che vengono utilizzati per ottenere stime dell'efficienza tecnica di ogni produttore mediante la relazione  $TE_i = \exp\{-u_i\}$ .

Aigner and Chu (1968), proposero due metodi di programmazione lineare per la determinazione dei coefficienti della (2.2.1.4).

### 1) *Modello Lineare*

$$\min_{\beta} \sum_i \left| \ln(y_i) - \beta_0 - \sum_n \beta_n \ln(x_{ni}) \right|, \quad (2.2.1.5)$$

dove i  $\beta$  vengono calcolati sotto il vincolo

$$\beta_0 - \sum_n \beta_n \ln(x_{ni}) \geq \ln(y_i) \quad (2.2.1.6)$$

### 2) *Modello quadratico*

$$\min_{\beta} \sum_n \left( \ln(y_i) - \beta_0 - \sum_{ni} \beta_n \ln(x_{ni}) \right)^2, \quad (2.2.1.7)$$

dove i parametri vengono calcolati sotto il vincolo

$$\beta_0 - \sum_n \beta_n \ln(x_{ni}) \geq \ln(y_i) \quad (2.2.1.8)$$

Mediante queste procedure i parametri  $\beta$  vengono calcolati piuttosto che stimati, quindi non è possibile ottenere gli standard error e di conseguenza non si possono applicare procedure inferenziali.

La frontiera di produzione deterministica può essere stimata in due passi (Winsten, 1957). Nel primo passo con i minimi quadrati ordinari (OLS) si ottengono le stime corrette e consistenti dei coefficienti  $\beta_1 \dots \beta_n$ , mentre si hanno stime distorte dell'intercetta  $\beta_0$ . Nel secondo passo la stima di  $\beta_0$  viene corretta per assicurare che la frontiera stimata limiti i dati superiormente. Due diverse tecniche vengono in genere applicate:

### ***Minimi quadrati ordinari corretti (COLS)***

Secondo la procedura dei minimi quadrati ordinari corretti l'intercetta è stimata da

$$\hat{\beta}_0^* = \hat{\beta}_0 + \max_i \{\hat{u}_i\}, \quad (2.2.1.9)$$

dove gli  $\hat{u}_i$  sono i residui OLS. I residui OLS sono corretti nella direzione opposta mediante:

$$-\hat{u}_i^* = \hat{u}_i - \max_i \{\hat{u}_i\}. \quad (2.2.1.10)$$

Queste stime giacciono al di sotto della frontiera o sulla frontiera (almeno un produttore), e sono semplici da implementare, ma presentano diversi inconvenienti. La frontiera così stimata è parallela alla retta di regressione OLS, quindi la struttura della tecnologia “*best practice*” è la stessa della struttura di tecnologia produttiva “centrale”, mentre dovrebbe essere possibile ottenere una struttura della frontiera che differisce dalla struttura della tecnologia produttiva media.

### ***Minimi quadrati modificati (MOLS)***

Un secondo metodo considera la stima mediante i minimi quadrati modificati (MOLS). Quest'ultima parte dall'assunzione che il disturbo segua una distribuzione specifica, per esempio una distribuzione di probabilità esponenziale

o una semi normale. Anche in questo caso, il primo passo prevede la stima OLS dei coefficienti  $\beta$ , seguita dalla stima dell'intercetta con il valore atteso della distribuzione specificata:

$$\hat{\beta}_0^{**} = \hat{\beta}_0 + E(\hat{u}_i), \quad (2.2.1.11)$$

$$\text{dove i residui sono } -\hat{u}_i^{**} = \hat{u}_i - E(\hat{u}_i). \quad (2.2.1.12)$$

Gli inconvenienti di questo secondo tipo di stima sono rappresentati dalla possibilità di avere punteggi di efficienza individuale del produttore maggiori di uno (nel caso in cui il residuo OLS sia sufficientemente grande da garantire  $\hat{u}_i - E(\hat{u}_i) > 0$ ). Inoltre, è possibile che l'intercetta sia spostata troppo in alto e di conseguenza nessun produttore è tecnicamente efficiente e come nel caso dei COLS la frontiera risulta parallela alla retta di regressione.

Le tecniche sopra descritte, pur essendo molto semplici, presentano lo svantaggio di non considerare il termine di errore, quindi tutte le variazioni degli output, non associate a corrispondenti variazioni degli input, sono attribuite all'inefficienza tecnica.

### Frontiere di produzione stocastiche

Modelli che prevedono l'esistenza dell'inefficienza tecnica, ma anche di shocks casuali non sotto il controllo dell'azienda, furono proposti simultaneamente da Aigner, Lovell e Schmidt (Aigner, Lovell e Schmidt, 1977) e da Meeusen e Van Den Broeck (Meeusen e Van Den Broeck, 1977).

Assumendo che la funzione  $f(x_i; \beta)$  sia di tipo Cobb-Douglas nella sua forma lineare, il modello di frontiera stocastica può essere scritto come:

$$\ln(y_i) = \beta_0 + \sum_n \beta_n \ln(x_{ni}) + v_i - u_i, \quad (2.2.1.13)$$

$$\varepsilon_i = v_i - u_i, \quad (2.2.1.14)$$

dove il termine di errore  $\varepsilon_i$  è diviso in due parti:  $v_i$  è la componente di errore casuale, iid e simmetrica, indipendentemente distribuita da  $u_i$  che rappresenta la componente non negativa di inefficienza tecnica.

L'obiettivo dell'analisi è la stima dei parametri  $\beta$  in  $f(x; \beta)$  e le stime dell'efficienza tecnica per ogni produttore. Quest'ultimo obiettivo richiede stime separate per le due componenti dell'errore per ciascun produttore, e quindi assunzioni sulla distribuzione di  $u_i$  e  $v_i$ . Se  $u_i$  è distribuito indipendentemente dagli input, la metodologia OLS fornisce stime consistenti dei parametri tecnologici eccetto che l'intercetta per ogni produttore. Assunzioni ulteriori, e una diversa tecnica di stima sono necessarie per ottenere una stima consistente dell'intercetta e dell'efficienza tecnica di ciascun produttore.

Gli stimatori di massima verosimiglianza vengono comunemente utilizzati per la stima dei  $\beta$  e degli  $u_i$ .

Tre assunzioni principali vengono fatte sulla natura distributiva delle componenti  $v_i$  e  $u_i$ ; cioè

- $v_i \sim \text{iid } N(0, \sigma_v^2)$ ;
- $u_i \sim \text{iid}$  come una semi normale, una normale troncata, un'esponenziale o una funzione gamma
- $v_i$  e  $u_i$  sono indipendenti tra loro e dai regressori.

L'assunzione di indipendenza tra  $u_i$  ed i regressori è la più problematica, in quanto la scelta degli input può essere condizionata da conoscenze degli imprenditori relative alla propria inefficienza.

Considerando il caso in cui  $u_i \sim N^+(0, \sigma_u^2)$ . La funzione di log verosimiglianza per un campione  $X = \{i; i = 1..I\}$  di produttori sarà

$$\ln L = \text{costante} - I \ln \sigma + \sum_i \ln \Phi\left(-\frac{\varepsilon_i \lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2, \quad (2.2.1.15)$$

dove:  $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$ ,  $\lambda = \sigma_u / \sigma_v$  e  $\Phi(\cdot)$  è funzione di distribuzione standard della Normale. La riparametrizzazione in termini di  $\sigma$  e  $\lambda$  è conveniente in quanto

$\lambda$  fornisce indicazioni circa il contributo relativo delle componenti  $v$  e  $u$  rispetto a  $\varepsilon$ .

La massimizzazione della funzione di log verosimiglianza consente di ottenere stime consistenti per  $I \rightarrow +\infty$  dei parametri. Le stime dell'errore  $\varepsilon$  contengono informazioni sulla componente  $u$  e sono utilizzate per ottenere valutazioni dell'efficienza tecnica di ciascun produttore. Si ricorre in questo caso alla distribuzione condizionata di  $u_i$  dato  $\varepsilon_i$ . Jondrow e altri (1982) mostrarono che se  $u_i \sim N^+(0, \sigma_u^2)$  allora:

$$f(u | \varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp\left\{-\frac{(u - \mu_*)^2}{2\sigma_*^2}\right\} \bigg/ \left[1 - \Phi\left(-\frac{\mu_*}{\sigma_*}\right)\right], \quad (2.2.1.16)$$

dove  $\mu_* = -\varepsilon\sigma_u^2/\sigma^2$  e  $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$ . Dal momento che la funzione di densità condizionata  $f(u|\varepsilon)$  segue una  $N^+(\mu_*, \sigma_*^2)$ , sia la media che la moda condizionate sono utilizzate come stimatori puntuali di  $u_i$

$$E(u_i | \varepsilon_i) = \sigma_* \left[ \frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(\varepsilon_i \lambda / \sigma)} - \left(\frac{\varepsilon_i \lambda}{\sigma}\right) \right], \quad (2.2.1.17)$$

$$M(u_i | \varepsilon_i) = \begin{cases} -\varepsilon_i \left(\frac{\sigma_u^2}{\sigma}\right) & \text{se } \varepsilon_i \leq 0 \\ \mathbf{0} & \text{altrimenti} \end{cases}. \quad (2.2.1.18)$$

Nella pratica, l'utilizzo della media condizionata ( $E(u|\varepsilon)$ ) è molto più frequente rispetto alla moda ( $M(u|\varepsilon)$ ), nonostante quest'ultima abbia un' interessante interpretazione come stimatore di massima verosimiglianza.

Le stime dell'efficienza tecnica di ogni produttore sono quindi ottenibili come

$$TE_i = \exp\{-\hat{u}_i\}, \quad (2.2.1.19)$$

dove la stima di  $u_i$  è sostituita dalla media o dalla moda condizionata .

Battese e Coelli (1988) proposero uno stimatore puntuale alternativo per  $TE_i$ :

$$TE_i = E(\exp\{-u_i\} | \varepsilon_i) = \left[ \frac{1 - \Phi\left(\frac{\sigma_u - \mu_{*i}}{\sigma_u}\right)}{1 - \Phi\left(\frac{-\mu_{*i}}{\sigma_u}\right)} \right] \cdot \exp\left\{-\mu_{*i} + \frac{1}{2}\sigma_u^2\right\}. \quad (2.2.1.20)$$

Gli stimatori puntuali presentati nella (2.2.1.19) e (2.2.1.20) danno risultati diversi. Lo stimatore proposto da Battese e Coelli in genere è preferito quando  $u_i$  non è prossimo a zero. Le stime dell'efficienza tecnica risultano comunque inconsistenti, in quanto la variazione associata alla distribuzione di  $(u|\varepsilon)$  è indipendente da  $i$ .

Nonostante l'utilizzo di una distribuzione esponenziale o di una semi Normale per la componente  $u$  porti a stime non distorte, ma inconsistenti dell'efficienza tecnica, molti autori pretendono per queste ultime. Infatti, la distribuzione Gamma, anche se più flessibile,

Il termine di errore fin qui considerato omoschedastico, può variare al variare della dimensione del produttore, così come le fonti di inefficienza possono variare a seconda della dimensione dei produttori. Di conseguenza, una delle due componenti dell'errore, o entrambe, potrebbero essere eteroschedastiche.

Se l'eteroschedasticità è presente nella componente simmetrica dell'errore  $v_i$ , ma viene ignorata, si otterranno stime non distorte dei parametri della frontiera  $\beta_n$ , ma distorte per l'intercetta  $\beta_0$ , (essendo  $E(u)$  diverso da zero), che viene sottostimata (Kumbhakar, Lovell, 2000). Infatti, se consideriamo la (3.1.4) e assumiamo che  $v_i \sim N(0, \sigma_{v_i}^2)$  e  $u_i \sim N(0, \sigma_u^2)$  avremo che:

$$E(\ln y_i) = \left[ \beta_0 - \sqrt{2/\pi\sigma_u} \right] + \sum_n \beta_n \ln x_{ni}. \quad (2.2.1.21)$$

Tale stima può essere corretta una volta stimato  $\sigma_u^2$ .

L'effetto dell'eteroschedasticità sulle stime dell'efficienza tecnica sono più rilevanti. Se si considera la moda condizionata  $M(u_i|\varepsilon_i)$ , si avrà<sup>2</sup>

$$M(u_i | \varepsilon_i) = \begin{cases} -\varepsilon_i \left[ \frac{1}{1 + \sigma_{vi}^2 / \sigma_u^2} \right] & \text{se } \varepsilon_i \leq 0 \\ 0 & \text{altrimenti} \end{cases}, \quad (2.2.1.21)$$

dove le fonti di variabilità sono due: i residui e i pesi associati ad essi.

Nell'ipotesi probabile in cui  $\sigma_{vi}^2$  vari direttamente al variare della dimensione del produttore, si avranno sovrastime dell'efficienza tecnica dei piccoli produttori e sottostime di quella dei produttori di dimensione relativamente grande. Disponendo solo di dati cross-section non è possibile stimare un set completo di parametri  $\sigma_{vi}^2$  specifici per ogni produttore, oltre agli altri parametri  $\sigma_u^2$  e  $\beta$  comuni a tutti i produttori. Per questo motivo è necessario esprimere l'eteroschedasticità in funzione di specifiche variabili  $z_i$  relative alla dimensione aziendale:

$$\sigma_{vi}^2 = g_1(z_i; \delta_1). \quad (2.2.1.22)$$

Il modello può essere stimato sia con il metodo della massima verosimiglianza che con il metodo dei momenti.

La funzione di log verosimiglianza, nel caso in cui  $v$  è eteroschedastico può essere così generalizzata

$$\ln L = \text{costante} - \frac{1}{2} \sum_i \ln [g_1(z_i; \delta_1) + \sigma_u^2] + \sum_i \ln \Phi \left( -\frac{\varepsilon_i \lambda_i}{\sigma_i} \right) - \frac{1}{2} \sum_i \frac{\varepsilon_i^2}{\sigma_i^2}, \quad (2.2.1.23)$$

dove

$$\sigma_i^2 = \sigma_u^2 + \sigma_{vi}^2 = \sigma_u^2 + g_1(z_i; \delta_1), \quad (2.2.1.24)$$

$$\lambda_i = \frac{\sigma_u}{\sigma_v} = \frac{\sigma_u}{\sqrt{g_1(z_i; \delta_1)}}. \quad (2.2.1.25)$$

---

<sup>2</sup> L'analisi può essere fatta anche in termini di media condizionata.

La funzione di log verosimiglianza è massimizzata per ottenere stime di  $\beta$ ,  $\sigma_u^2$  e  $\delta_1$ . Allo stesso modo, la massimizzazione della funzione di log-verosimiglianza viene applicata per stimare  $\sigma_{vi}^2$  (ricordando che  $\hat{\sigma} = g_1(z_i; \delta_1)$ ) che è utilizzata, insieme a  $\sigma_u^2$ , per ottenere la stima dell'efficienza tecnica di ciascun produttore.

La presenza di eteroschedasticità nella componente  $u_i$  dell'errore è dovuta a fattori specifici delle aziende che influenzano la distribuzione di  $u_i$ . In questo caso, sia le stime dei parametri della struttura produttiva che l'efficienza tecnica risultano non corrette. I parametri  $\beta$  sono distorti in quanto l'eteroschedasticità di  $u$  genera intercette specifiche per ciascun produttore le quali provocano distorsione anche negli altri parametri della tecnologia. Per quanto riguarda l'efficienza tecnica, l'eteroschedasticità della componente  $u_i$  implica che nella moda condizionata si abbia

$$M(u_i | \varepsilon_i) = \begin{cases} -\varepsilon_i \left[ \frac{1}{1 + \sigma_v^2 / \sigma_{ui}^2} \right] & \text{se } \varepsilon_i \leq 0 \\ 0 & \text{altrimenti} \end{cases} \quad (2.2.1.26)$$

Come evidenziato per l'eteroschedasticità nella componente  $v_i$ , anche in questo caso le fonti di variabilità sono due: i residui e i pesi ad essi associati. L'effetto dell'eteroschedasticità nella componente  $u_i$  è opposto rispetto a quello riscontrato per la componente  $v_i$ . Infatti, se l'eteroschedasticità in  $u$  varia direttamente con la dimensione dei produttori, questo causa una sovrastima della moda condizionata per i piccoli produttori e una sottostima per i grandi produttori. L'efficienza tecnica dei piccoli produttori verrà quindi sottostimata e quella dei grandi produttori sovrastimata.

La componente  $\sigma_{ui}^2$  è espressa in funzione di variabili specifiche per il singolo produttore

$$\sigma_{ui}^2 = g_2(z_i; \delta_2). \quad (2.2.1.27)$$

La funzione di log-verosimiglianza, sotto l'ipotesi che  $v_i \sim \text{iid } N(0, \sigma_v^2)$  e  $u_i \sim N^+(0, \sigma_{ui}^2)$ , sarà

$$\ln L = \text{costante} - \frac{1}{2} \sum_i \ln [g_2(z_i; \delta_2) + \sigma_v^2] + \sum_i \ln \Phi \left( -\frac{\varepsilon_i \lambda_i}{\sigma_i} \right) - \frac{1}{2} \sum_i \frac{\varepsilon_i^2}{\sigma_i^2}, \quad (2.2.1.28)$$

con

$$\sigma_i^2 = \sigma_v^2 + \sigma_{ui}^2 = \sigma_v^2 + g_2(z_i; \delta_2), \quad (2.2.1.29)$$

$$\lambda_i = \frac{\sigma_{ui}}{\sigma_v} = \frac{\sqrt{g_2(z_i; \delta_2)}}{\sigma_v}. \quad (2.2.1.30)$$

La massimizzazione della funzione di log-verosimiglianza consente di ottenere stime di  $\beta$ ,  $\sigma_v^2$  e  $\delta_2$ . Tali stime, insieme a quella di  $\sigma_{ui}^2$ , verranno utilizzate per ottenere stime non distorte dell'efficienza tecnica di ogni produttore. In questo caso, però, il metodo dei momenti non può essere utilizzato.

La presenza di eteroschedasticità nelle due componenti causa distorsioni in entrambi i sensi, l'errore totale  $\varepsilon$  è complessivamente ridotto. L'effetto dell'eteroschedasticità in entrambe le componenti dell'errore è evidente considerando una distribuzione semi Normale e utilizzando la moda condizionata:

$$M(u_i | \varepsilon_i) = \begin{cases} -\varepsilon_i \left[ \frac{1}{1 + \sigma_{vi}^2 / \sigma_{ui}^2} \right] & \text{se } \varepsilon_i \leq 0 \\ 0 & \text{altrimenti} \end{cases}. \quad (2.2.1.31)$$

L'errore nella stima della moda condizionata  $M(u|\varepsilon)$  dipende dal rapporto  $\sigma_{vi}^2 / \sigma_{ui}^2$ . Solo se tale rapporto è costante tra i produttori la stima di  $M(u|\varepsilon)$  non è distorta in caso di eteroschedasticità se si assume omoschedasticità. In caso contrario, la stima di  $\sigma_{ui}^2$  e  $\sigma_{vi}^2$  verrà fatta con il metodo della massima verosimiglianza. Se si assume che  $v_i \sim N(0, \sigma_{vi}^2)$  con  $\sigma_{vi}^2 = g_1(z_i; \delta_1)$  e  $u_i \sim N^+(0, \sigma_{ui}^2)$  con  $\sigma_{ui}^2 = g_2(z_i; \delta_2)$ , la funzione di log verosimiglianza sarà:

$$\ln L = \text{costante} - \frac{1}{2} \sum_i \ln [g_1(z_i; \delta_1) + g_2(z_i; \delta_2)] + \sum_i \ln \Phi \left( -\frac{\varepsilon_i \lambda_i}{\sigma_i} \right) - \frac{1}{2} \sum_i \frac{\varepsilon_i^2}{\sigma_i^2}, \quad (2.2.1.32)$$

dove

$$\sigma_i^2 = \sigma_{vi}^2 + \sigma_{ui}^2 = g_1(z_i; \delta_1) + g_2(z_i; \delta_2), \quad (2.2.1.33)$$

$$\lambda_i = \frac{\sigma_{ui}}{\sigma_{vi}} = \sqrt{\frac{g_2(z_i; \delta_2)}{g_1(z_i; \delta_1)}}. \quad (2.2.1.34)$$

Anche in questo caso, come visto precedentemente, si possono ottenere stime di  $\beta$ ,  $\delta_1$  e  $\delta_2$ , che verranno sostituite nelle funzioni  $g_1(z_i; \delta_1)$  e  $g_2(z_i; \delta_2)$ , per ottenere stime della moda condizionata e successivamente dell'efficienza tecnica di ciascun produttore (ricordando che  $TE_i = \exp\{-\hat{u}_i\}$ ).

In questo contesto, il principale problema posto dai dati cross-section è rappresentato dai risultati, i quali sono fortemente influenzati sia dalle variabili considerate che dalla forma funzionale scelta per la componente eteroschedastica.

### ***2.2.2 Metodi per dati panel***

La disponibilità di dati panel (osservazioni ripetute nel tempo per ogni produttore), consente di ovviare, almeno in parte, ad alcuni problemi che si presentano nel caso di dati cross-section. In particolare, osservazioni ripetute per un campione di produttori possono essere utilizzate in sostituzione di forti assunzioni circa la forma distributiva. Inoltre, non tutte le tecniche di stima per dati panel richiedono l'indipendenza tra la componente dell'errore attribuibile all'efficienza tecnica e i regressori. L'utilizzo di più osservazioni per lo stesso campione di produttori fornisce un maggiore numero di informazioni sull'efficienza tecnica che l'utilizzo di più produttori nell'analisi cross-section. La stima dell'efficienza tecnica è consistente per  $T \rightarrow \infty$  (dove T è il numero di osservazioni per ogni produttore). Nella pratica questo beneficio è solo parzialmente utilizzabile, in quanto la maggior parte dei data set è composta da un numero relativamente basso di osservazioni temporali.

Nell'analisi dell'efficienza tecnica con dati panel si può supporre che l'efficienza vari tra i produttori, ma rimanga costante nel tempo per ogni produttore, o che vari sia tra produttori che rispetto al tempo.

Di seguito si considereranno entrambe le alternative, ipotizzando di disporre di panel bilanciati (ogni produttore  $i=1..I$  è osservato per  $t=1..T$  periodi).

**Efficienza tecnica invariante nel tempo**

Se si considera la forma funzionale Cobb-Douglas, e disponendo di  $T$  osservazioni relative a  $I$  produttori, la frontiera di produzione con efficienza tecnica invariante nel tempo sarà:

$$\ln(y_{it}) = \beta_0 + \sum_n \beta_n \ln(x_{nit}) + v_{it} - u_i, \quad (2.2.2.1)$$

dove  $v_{it}$  rappresenta l'errore casuale e  $u_i$  l'inefficienza tecnica (con  $u_i$  maggiore o uguale a zero). La tecnologia è assunta costante nel tempo, in quanto non viene preso in considerazione il cambiamento tecnologico.

I parametri di questo modello possono essere stimati in diversi modi. Di seguito verranno presentati brevemente tre diversi approcci: il modello ad effetti fissi basato sullo stimatore LSDV (least squares with dummy variables), il modello ad effetti variabili basato sullo stimatore GLS e l'approccio basato sulla funzione di massima verosimiglianza. A seconda delle circostanze un metodo può essere preferibile agli altri.

***Least squares with dummy variables (LSDV)***

Il metodo di stima più semplice è quello ad effetti fissi. L'unica modifica rispetto al classico modello per dati panel ad effetti fissi è l'imposizione  $u_i \geq 0$ . Inoltre, si assume che la componente dell'errore  $v_{it}$  sia iid  $(0, \sigma_v^2)$  e incorrelata con i regressori, mentre non si fanno assunzioni circa la distribuzione di  $u_i$  che può essere correlata con i regressori o con  $v_{it}$ . Le componenti  $u_i$  sono fisse, quindi considerabili come intercette specifiche per ogni produttore e di conseguenza

possono essere stimate insieme ai coefficienti  $\beta_n$ . Il modello può essere stimato applicando il metodo OLS a

$$\ln(y_{it}) = \beta_{0i} + \sum_n \beta_n \ln(x_{nit}) + v_{it}, \quad (2.2.2.2)$$

dove  $\beta_{0i} = (\beta_0 - u_i)$  sono intercette specifiche dei produttori. La stima può essere effettuata eliminando  $\beta_0$  e stimando I intercette specifiche, ciascuna per ogni produttore, oppure mantenendo  $\beta_0$  e stimando (I-1) intercette specifiche per produttore, o applicando la trasformazione *within*, in cui tutti i dati sono espressi in termini di deviazioni dal produttore medio e le I intercette sono ottenute come medie dei residui dei produttori. Dopo la stima si utilizza la normalizzazione:

$$\hat{\beta}_0 = \max_i \{ \hat{\beta}_{0i} \}, \quad (2.2.2.3)$$

e gli  $u_i$  sono stimati da

$$\hat{u}_i = \hat{\beta}_0 - \hat{\beta}_{0i}, \quad (2.2.2.4)$$

tale da assicurare  $u_i \geq 0$ . Stime dell'efficienza tecnica specifica per ogni singolo produttore sono date da :

$$TE_i = \exp\{-\hat{u}_i\}. \quad (2.2.2.5)$$

Nel modello ad effetti fissi almeno un produttore è efficiente al 100% e l'efficienza tecnica degli altri è misurata in relazione al produttore più efficiente.

Questo modello oltre ad essere semplice presenta importanti proprietà di consistenza. Infatti, contrariamente a quanto visto per lo stimatore di massima verosimiglianza per dati cross-section, la stima dell'efficienza tecnica specifica per ogni produttore, nel caso di modello a effetti fissi per panel data, è consistente.

Tale modello ha, però, l'inconveniente di inglobare nella componente  $u_i$  tutti i fenomeni che variano tra i produttori ma sono costanti nel tempo per ogni produttore. L'impossibilità di distinguere tra variazione nell'efficienza tecnica e

altri effetti avviene sia che gli altri effetti siano inclusi o no nel modello come regressori.

**Modello a effetti variabili**

Nel modello a effetti variabili si ipotizza che  $u_i$  siano casualmente distribuiti con media zero e varianza costante, ma incorrelati con i regressori e con  $v_i$ . Non si fanno assunzioni circa la distribuzione di  $u_i$ , si ipotizza soltanto che questi siano non negativi, mentre la componente  $v_i$  ha media zero e varianza costante. Tali assunzioni consentono di includere regressori *time-invariant* nel modello. Il modello può quindi essere riscritto nel seguente modo

$$\begin{aligned} \ln y_{it} &= [\beta_0 - E(u_{it})] + \sum_n \beta_n \ln x_{nit} + v_{it} - [u_i - E(u_i)] = \\ &= \beta_0^* + \sum_n \beta_n \ln x_{nit} + v_{it} - u_i^*, \end{aligned} \tag{2.2.2.6}$$

dove l'assunzione che gli  $u_i$  siano casuali invece che fissi consente ad alcuni input  $x_{nit}$  di essere invarianti nel tempo.

Questo modello può essere stimato con il metodo two-step GLS. Le stime di  $u_i$  vengono ottenute dai residui per mezzo di

$$\hat{u}_i^* = \frac{1}{T} \sum_t \left( \ln y_{it} - \hat{\beta}_0^* - \sum_n \hat{\beta}_n \ln x_{nit} \right), \tag{2.2.2.7}$$

e della normalizzazione :  $\hat{u}_i = \max_i \{ \hat{u}_i^* \} - \hat{u}_i^*$ . Queste stime sono consistenti sia per  $I \rightarrow +\infty$  che per  $T \rightarrow +\infty$ . L'efficienza produttiva di ogni singolo produttore si ottiene sostituendo  $\hat{u}_i$  nella (2.2.2.5).

Uno stimatore alternativo per  $\hat{u}_i$  è lo stimatore BLUP (best linear unbiased predictor), che per T ampio equivale allo stimatore GLS. Entrambi gli stimatori sono consistenti per  $I \rightarrow +\infty$  e per  $T \rightarrow +\infty$ .

Come nel caso dello stimatore ad effetti fissi almeno un produttore sarà tecnicamente efficiente, e l'efficienza tecnica degli altri produttori sarà misurata in funzione del produttore efficiente.

**Metodo della massima verosimiglianza**

La stima di massima verosimiglianza è applicabile anche ai dati panel, facendo le dovute assunzioni circa la distribuzione dei termini di errore. Il modello è strutturalmente molto simile a quello visto per dati cross-section, ma in questo caso la componente di errore  $v_{it}$  varia nel tempo. Se si assume che  $v_{it} \sim iid N(0, \sigma_v^2)$ ,  $u_i \sim iid N^+(0, \sigma_u^2)$ ,  $u_i$  e  $v_i$  sono distribuiti indipendentemente tra loro e dai repressori, allora la funzione di log-verosimiglianza per un campione di I produttori, ciascuno osservato per T periodi, sarà

$$\ln L = costante - \frac{I(T-1)}{2} \ln \sigma_v^2 - \frac{1}{2} \ln(\sigma_v^2 + T\sigma_u^2) + \sum_i \ln \left[ 1 - \Phi \left( \frac{\mu_{*i}}{\sigma_*} \right) \right] - \frac{\sum_i \varepsilon_i \varepsilon_i}{2\sigma_v^2} + \frac{1}{2} \sum_i \left( \frac{\mu_{*i}}{\sigma_*} \right)^2 \quad (2.2.2.8)$$

dove

$$\mu_* = \frac{T\sigma_u^2 \bar{\varepsilon}}{\sigma_v^2 + T\sigma_u^2}, \quad (2.2.2.9)$$

$$\sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma_v^2 + T\sigma_u^2}, \quad (2.2.2.10)$$

$$\bar{\varepsilon} = \frac{1}{T} \sum_i \varepsilon_{it}. \quad (2.2.2.11)$$

La funzione di log verosimiglianza viene massimizzata rispetto ai parametri per ottenere stime di  $\beta$ ,  $\sigma_u^2$  e  $\sigma_v^2$ . Dalla distribuzione di probabilità condizionata di  $(u|\varepsilon)$ :

$$f(u|\varepsilon) = \frac{f(u, \varepsilon)}{f(\varepsilon)} = \frac{1}{(2\pi)^{1/2} \sigma_* [1 - \Phi(-\mu_* / \sigma_*)]} \exp \left\{ -\frac{(u - \mu_*)^2}{2\sigma_*^2} \right\}, \quad (2.2.2.12)$$

si possono ricavare la media o la moda per avere stime puntuali di  $u_i$  (consistenti per  $T \rightarrow +\infty$ ), che vengono poi sostituite nella (2.2.2.5) allo scopo di ottenere stime dell'efficienza tecnica specifiche e *time invariant* per ciascun produttore

$$E(u_i | \varepsilon_i) = \mu_{*i} + \sigma_* \left[ \frac{\phi(-\mu_{*i} / \sigma_*)}{1 - \Phi(-\mu_{*i} / \sigma_*)} \right], \quad (2.2.2.13)$$

$$M(u_i | \varepsilon_i) = \begin{cases} -\mu_{*i} & \text{se } \varepsilon_i \leq 0 \\ 0 & \text{altrimenti} \end{cases}. \quad (2.2.2.13')$$

La componente  $u_i$  può essere modellizzata anche mediante una Normale troncata (Battese e Coelli, 1988).

L'utilizzo di dati panel riduce i problemi legati all'eteroschedasticità delle componenti di errore, anche se non li elimina del tutto. In particolare, se si considera l'eteroschedasticità della componente casuale dell'errore ( $v_i$ ), in modelli per dati panel invarianti nel tempo, l'utilizzo dell'approccio ad effetti fissi o ad effetti variabili, che tengano conto dell'eteroschedasticità, determina stime consistenti sia per i parametri della funzione di produzione che dell'efficienza tecnica. L'utilizzo del metodo della massima verosimiglianza, invece, richiede che il numero di osservazioni  $T$  sia grande rispetto alla numerosità  $I$ , altrimenti non è applicabile in caso di eteroschedasticità.

L'eteroschedasticità della componente di errore attribuibile all'inefficienza tecnica, al contrario, presenta più inconvenienti rispetto al caso precedente. Infatti, il modello ad effetti fissi non è utilizzabile, in quanto gli  $u_i$  non possono essere sia fissi che eteroschedastici contemporaneamente. Nemmeno l'approccio ad effetti casuali è praticabile, e l'utilizzo della massima verosimiglianza anche se possibile a livello teorico, in pratica è inutilizzabile per l'elevato numero di parametri da stimare lo rende inutilizzabile. Solamente il metodo dei momenti può essere

applicato, ed è l'unico possibile anche nel caso in cui entrambe le componenti siano eteroschedastiche il metodo dei momenti è l'unico che è possibile applicare.

**Efficienza tecnica variabile nel tempo**

L'assunzione di efficienza tecnica costante nel tempo è piuttosto forte, soprattutto se si sta analizzando un panel abbastanza lungo. L'ipotesi di efficienza tecnica variabile nel tempo appare in molti casi più realistico, ma impone un maggior numero di parametri da stimare. Anche in questo caso gli approcci all'analisi sono fondamentalmente due: uno in cui la variazione temporale dell'efficienza tecnica è modellata attraverso effetti fissi o variabili e il metodo della massima verosimiglianza.

Il modello per un panel di I produttori, osservati per T periodi, nel caso si faccia riferimento a una funzione di produzione di tipo Cobb-Douglas sarà

$$\begin{aligned} \ln(y_{it}) &= \beta_{0t} + \sum_n \beta_n \ln(x_{nit}) + v_{it} - u_{it} \\ &= \beta_{it} + \sum_n \beta_n \ln(x_{nit}) + v_{it}, \end{aligned} \tag{2.2.2.14}$$

dove  $\beta_{0t}$  è l'intercetta comune a tutti i produttori al tempo t,  $\beta_{it} = (\beta_{0t} - u_{it})$  è l'intercetta per il produttore i-esimo nel periodo t. Per un panel di I×T osservazioni non è possibile stimare I×T intercette ( $\beta_{it}$ ), N pendenze ( $\beta_n$ ) e  $\sigma^2_v$ .

Cornwell, Schmidt e Sickles (1990) proposero di specificare

$$\beta_{it} = \Omega_{i1} + \Omega_{i2}t + \Omega_{i3}t^2, \tag{2.2.2.15}$$

che riduce il numero di intercette a I×3. In questo modo l'efficienza tecnica può variare nel tempo e tra i produttori. Il numero di parametri da stimare rimane comunque elevato, soprattutto se il rapporto I/T è elevato.

La stima può essere effettuata sia mediante l'approccio ad effetti fissi che variabili<sup>3</sup>

La possibilità di fare assunzioni circa la distribuzione e l'indipendenza delle componenti  $u_i$  e  $v_i$  rende possibile utilizzare il metodo della massima verosimiglianza nella stima dell'efficienza tecnica variabile nel tempo.

Se si considera il modello (2.2.2.14), si assume che  $u_{it} = \beta_t u_i$ , dove  $v_{it} \sim \text{iid } N(0, \sigma_v^2)$ ,  $u_i \sim \text{iid } N^+(0, \sigma_u^2)$ , e  $\varepsilon_{it} = v_{it} - u_{it} = v_{it} - \beta_t u_i$  dove  $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$ , allora la funzione di log-verosimiglianza sarà:

$$\ln L = \text{costante} - \frac{I}{2} \ln \sigma_*^2 - \frac{1}{2} \sum_i a_{*i} - \frac{I \cdot T}{2} \ln \sigma_v^2 - \frac{1}{2} \ln \sigma_u^2 + \sum_i \ln \left[ 1 - \Phi \left( -\frac{\mu_{*i}}{\sigma_*} \right) \right], \quad (2.2.2.16)$$

dove

$$\mu_{*i} = \frac{(\sum_t \beta_t \varepsilon_{it}) \sigma_v^2}{(\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2)}, \quad (2.2.2.17)$$

$$\sigma_*^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2}, \quad (2.2.2.18)$$

$$a_{*i} = \frac{1}{\sigma_v^2} \left[ \sum_t \varepsilon_{it}^2 - \frac{\sigma_u^2 (\sum_t \beta_t \varepsilon_{it})^2}{\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2} \right]. \quad (2.2.2.19)$$

Dopo aver ottenuto stime di  $u_i$  utilizzando la media o la moda di  $u_i | \varepsilon_i$ :

$$E(u_i | \varepsilon_i) = \mu_{*i} + \sigma_* \left[ \frac{\phi(-\mu_{*i} / \sigma_*)}{1 - \Phi(-\mu_{*i} / \sigma_*)} \right], \quad (2.2.2.20)$$

$$M(u_i | \varepsilon_i) = \begin{cases} -u_{*i} & \text{se } \sum_t \beta_t \varepsilon_{it} \geq 0 \\ 0 & \text{altrimenti} \end{cases}, \quad (2.2.2.21)$$

<sup>3</sup> Per una descrizione delle metodologie che si possono utilizzare si veda Kumbhakar e Knox Lovell (2000)

le stime di  $u_{it}$  saranno  $\hat{u}_{it} = \hat{u}_i \hat{\beta}_t$ , dove  $\hat{\beta}_t$  sono le stime di massima verosimiglianza di  $\beta_t$ , soggetti alla normalizzazione  $\beta_1=0$  o  $\beta_t=0$ .

La presenza di eteroschedasticità nella componente  $v_i$ , anche nel caso di dati panel con inefficienza tecnica variabile nel tempo, consente di utilizzare la metodologia ad effetti fissi e variabili<sup>4</sup>, mentre per la massima verosimiglianza sorgono i consueti problemi relativi al numero di parametri da stimare. Se è la componente  $u_i$  ad essere eteroschedastica si adottano tecniche che utilizzano il metodo dei momenti o la funzione di massima verosimiglianza, mentre se entrambe le componenti sono eteroschedastiche si considera il metodo dei momenti.

### ***2.2.3 Modelli con variabili che determinano l'inefficienza tecnica***

Gli obiettivi dell'analisi dell'efficienza produttiva sono la stima della frontiera di produzione e dell'efficienza tecnica di ogni produttore. Quest'ultima componente dell'analisi è realizzabile associando le variazioni delle *performances* produttive a variabili esogene che caratterizzano il contesto produttivo in cui i produttori operano. Esempi di variabili che possono determinare l'efficienza tecnica sono il grado di competitività del settore, indicatori di qualità dell'input o dell'output, la forma societaria, la localizzazione geografica ecc.

Sia  $x = (x_1, \dots, x_N) \geq 0$  un vettore di input utilizzati per produrre un output  $y \geq 0$ , e sia  $z = (z_1, \dots, z_Q)$  un vettore di variabili esogene che influenzano l'efficienza tecnica e non la struttura della frontiera di produzione determinata dal legame tra  $x_i$  e  $y_i$ .

Uno tra i primi modelli che è stato proposto per l'analisi delle variabili che influenzano l'efficienza tecnica dei produttori è caratterizzato dalla presenza di due stadi di analisi. Nel primo stadio viene stimata una frontiera stocastica del tipo

---

<sup>4</sup> La metodologia da adottare deve ovviamente tener conto della presenza di eteroschedasticità.

## 2.2 METODI PARAMETRICI

$$\ln y_i = \ln f(x_i, \beta) + v_i - u_i, \quad \text{con } i=1, \dots, I \quad (2.2.3.1)$$

di solito mediante il metodo della massima verosimiglianza, sotto le convenzionali assunzioni di indipendenza, ottenendo una stima dell'efficienza tecnica come visto precedentemente. Nel secondo passo le stime dell'efficienza tecnica vengono regredite sulle variabili  $z_i$  che si ritiene determinino l'efficienza tecnica. Quest'ultima regressione si può formalizzare come:

$$E(u_i | v_i - u_i) = g(z_i; \gamma) + \varepsilon_i. \quad (2.2.3.2)$$

Questo approccio presenta dei problemi al punto di vista econometrico. Si deve assumere che  $z_i$  e  $x_i$  siano incorrelate nella prima parte dell'analisi, altrimenti le stime di massima verosimiglianza di  $\beta$ ,  $\sigma_v^2$ ,  $\sigma_u^2$  risulterebbero distorte per l'omissione di variabili rilevanti  $z_i$  e di conseguenza anche la stima dell'efficienza tecnica sarebbe distorta poiché basata su una rappresentazione non corretta della frontiera di produzione. Nel primo stadio, inoltre, si assume che la componente di inefficienza sia identicamente distribuita, ma questo è contraddetto nel secondo stadio, in cui i termini di inefficienza sono posti in relazione con le  $z_i$ .

Deprins e Simar (1989) espressero la frontiera di produzione come

$$\ln y_i = \ln f(x_i, \beta) + v_i - u_i, \quad (2.2.3.3)$$

$$E(u_i | z_i) = \exp\{\gamma' z_i\}, \quad (2.2.3.4)$$

dove  $\beta$  e  $\gamma$  sono rispettivamente i parametri relativi alla tecnologia e all'efficienza che devono essere stimati,  $u_i \geq 0$  rappresenta l'inefficienza tecnica e  $\exp\{\gamma' z_i\}$  esprime la relazione tra inefficienza tecnica e variabili esogene. Dalle equazioni (2.2.3.3) e (2.2.3.4) e aggiungendo un termine di errore  $\varepsilon_i$  a media zero e varianza costante, si ottiene:

$$\ln y_i = \ln f(x_i, \beta) - \exp\{\gamma' z_i\} + \varepsilon_i, \quad (2.2.3.5)$$

Dal momento che  $u_i \geq 0$ , allora  $\varepsilon_i \leq \exp\{\gamma' z_i\}$ <sup>5</sup>

Il modello(2.2.3.5) è non lineare nei parametri e può essere stimato con il metodo dei minimi quadrati non lineari o quello della massima verosimiglianza, dopo aver specificato una distribuzione per  $u_i = \exp\{\gamma' z_i\} - \varepsilon_i$ . Il limite principale di questo modello è l'assenza di una componente di errore casuale nella forma funzionale della frontiera di produzione, vincolando il modello ad una natura puramente deterministica.

Un' ulteriore difficoltà insita in questi metodi è la definizione di quali variabili siano caratteristiche della tecnologia di produzione e quali dell'efficienza tecnica.

Kumbhakar, Ghosh e McGuckin (1991) proposero un modello che ovvia a questi problemi, inserendo una componente di errore nella funzione di produzione:

$$\ln y_i = \ln f(x_i, \beta) + v_i - (\gamma' z_i + \varepsilon_i), \quad (2.2.3.6)$$

dove  $v_i \sim \text{iid} N(0, \sigma_v^2)$  e  $u_i = \gamma' z_i + \varepsilon_i$ . Poiché è necessario che la componente  $u_i$  sia positiva si impone che  $\varepsilon_i \geq -\gamma' z_i$ . Per ottenere la funzione di massima verosimiglianza, inoltre, sono necessarie assunzioni circa la forma funzionale di  $u_i$  e  $v_i$ . Se si assume che  $v_i \sim \text{iid} N(0, \sigma_v^2)$ , e  $u_i \sim \text{iid} N^+(\gamma' z_i, \sigma_u^2)$  e che  $u_i$  e  $v_i$  sono distribuiti indipendentemente i parametri della (2.2.3.6) possono essere stimati con la massima verosimiglianza. La funzione di log verosimiglianza da massimizzare per ottenere stime di  $\beta$ ,  $\gamma$ ,  $\sigma_v^2$  e  $\sigma_u^2$  è

$$\ln L = \text{costante} - \frac{1}{2} \ln(\sigma_v^2 + \sigma_u^2) + \sum_i \ln \Phi\left(-\frac{\gamma' z_i}{\sigma_u}\right) + \sum_i \ln \Phi\left(\frac{\mu_i^*}{\sigma^*}\right) - \frac{1}{2} \sum_i \frac{(e_i + \gamma' z_i)^2}{\sigma_v^2 + \sigma_u^2}, \quad (2.2.3.7)$$

$$\mu_i^* = \frac{\sigma_v^2 \gamma' z_i - \sigma_u^2 e_i}{\sigma_v^2 + \sigma_u^2}, \quad (2.2.3.8)$$

<sup>5</sup> Il termine  $\varepsilon_i$  non è identicamente distribuito in quanto il suo supporto dipende da  $z_i$ .

$$\sigma^{*2} = \frac{\sigma_u^2 \sigma_v^2}{\sigma_v^2 + \sigma_u^2}, \quad (2.2.3.9)$$

$$e_i = \ln y_i - \ln f(x_i; \beta). \quad (2.2.3.10)$$

Le stime dei parametri sono utilizzate per ottenere la media o la moda condizionate  $E(u_i|e_i)$  e  $M(u_i|e_i)$ , attraverso le quali si stima l'efficienza tecnica di ciascun produttore come visto precedentemente.

Huang e Liu (1994) introdussero un modello in cui è prevista l'interazione tra elementi di  $z_i$  e  $x_i$ . Formalmente il modello è:

$$\ln y_i = \ln f(x_i, \beta) + v_i - [g(z_i; \gamma) + \varepsilon_i], \quad (2.2.3.11)$$

dove la  $g(z_i; \gamma)$  può essere esplicitata come

$$g(z_i, x_i; \gamma) = \sum_q \gamma_q z_{qi} + \sum_q \sum_n \gamma_{qn} \ln x_{ni}. \quad (2.2.3.12)$$

La particolarità di questo modello è proprio l'assunzione che l'efficienza tecnica non sia neutrale rispetto all'utilizzo degli input.<sup>6</sup>

Battese e Coelli (1995) proposero una formulazione delle frontiere di produzione con effetti d'inefficienza per dati panel. Il modello è:

$$\ln y_{it} = \ln f(x_{it}, \beta) + v_{it} - u_{it}, \quad (2.2.3.13)$$

con  $i=1,2,\dots,N$  e  $t=1,2,\dots,T$ ,  $v_i \sim \text{iid } N(0, \sigma_v^2)$  e indipendentemente distribuito da  $u_i \sim \text{iid } N^+(m_{it}, \sigma_u^2)$ , dove  $m_{it} = z_{it} \delta$ . L'inefficienza tecnica viene quindi esplicitata in funzione di un set di variabili ( $z_i$ ) e di un vettore di coefficienti  $\delta$  da stimare. Il vettore  $\mathbf{z}$  può aver il primo valore pari a 1 e includere delle variabili già considerate come input della funzione di produzione, e/o interazioni tra esse,

---

<sup>6</sup> Per maggiori dettagli circa il modello di Huang e Liu, si rimanda all'articolo pubblicato sul Journal of Productivity Analysis (1994).

assieme a variabili specifiche per le aziende. Se le variabili  $z_i$  sono tutte nulle a parte l'intercetta si ritorna al modello specificato da Stevenson (1980) e Battese e Coelli (1988,1992), mentre se tutti gli elementi del vettore  $\delta$  sono nulli l'inefficienza tecnica non è legata alle variabili  $z_i$ , quindi si ha la distribuzione specificata da Aigner, Lovell e Schmidt (1977). La presenza di interazione tra input e variabili specifiche dell'azienda  $z_i$  riporta il modello alla formulazione di frontiera non neutrale proposta da Huang e Liu (1994).

La componente  $u_{it}$  relativa all'inefficienza tecnica può essere specificata dall'equazione

$$u_{it} = z_{it}\delta + w_{it}, \quad (2.2.3.14)$$

dove la variabile casuale  $w_{it}$  è definita dalla distribuzione Normale troncata con media zero e varianza  $\sigma^2$ , in cui il punto di troncamento è  $-z_{it}\delta$ , di conseguenza  $w_{it} = -z_{it}\delta$ . Queste assunzioni sono consistenti con la forma distributiva di  $u_{it}$  secondo una Normale  $N^+(z_{it}\delta, \sigma^2)$ , anche se l'imposizione che  $u_{it}$  e  $v_{it}$  siano indipendentemente distribuite per  $t=1, \dots, T$  e  $i=1, \dots, N$  è una restrizione.

I parametri della frontiera stocastica e del modello relativo all'inefficienza tecnica possono essere stimati simultaneamente con il metodo della massima verosimiglianza (Battese e Coelli, 1993). In questo caso la funzione di log verosimiglianza, espressa in funzione di  $\sigma_s^2 = \sigma_v^2 + \sigma^2$  e  $\gamma \equiv \sigma^2 / \sigma_s^2$ , è

$$\ln L = -\frac{1}{2} \left( \sum_{i=1}^N T_i \right) \left( \ln 2\pi + \ln \sigma_s^2 \right) - \frac{1}{2} \sum_{i=1}^N \sum_{t=1}^{T_i} \left[ \frac{(y_{it} - x_{it}\beta + z_{it}\delta)^2}{\sigma_s^2} \right] - \sum_{i=1}^N \sum_{t=1}^{T_i} \left[ \ln \Phi(d_{it}) - \ln \Phi(d_{it}^*) \right], \quad (2.2.3.15)$$

dove:

$$d_{it} = \frac{z_{it}\delta}{(r\sigma_s^2)^{1/2}} \quad (2.2.3.16)$$

$$d_{it}^* = \frac{\mu_{it}^*}{[\gamma(1-\gamma)\sigma_s^2]^{1/2}} \quad (2.2.3.17)$$

$$\mu_{it}^* = (1 - \gamma)z_{it}\delta - \gamma(y_{it} - x_{it}\beta) \quad (2.2.3.18)$$

$$\sigma^* = [\gamma(1 - \gamma)\sigma_s^2]^{1/2} \quad (2.2.3.19)$$

Questo modello considera la possibilità che il panel sia non bilanciato, infatti si considera  $T_i$  invece di  $T$ .

L'efficienza tecnica per l' $i$ -esima azienda in corrispondenza della  $t$ -esima osservazione sarà

$$TE_{it} = \exp(-u_{it}) = \exp(-z_{it}\delta - w_{it}). \quad (2.2.3.20)$$

Il modello di Battese e Coelli è stato ampiamente utilizzato nell'analisi empirica, ma anche criticato perché non tiene conto della possibile presenza di eteroschedasticità. L'utilizzo di dati aziendali, inoltre, aumenta la possibilità di eteroschedasticità nella componente  $u_i$  che riflette l'inefficienza dovuta a fattori sotto il controllo dell'azienda. E' probabile, infatti, che le aziende di grandi dimensioni abbiano più fattori sotto il loro controllo rispetto alle piccole (Kalirajan e Wu, 1999).

L'obiettivo sottostante il modello di Battese e Coelli è quello di mitigare il vincolo di costanza della moda condizionale della distribuzione Normale troncata, specificandola come funzione di variabili esogene. In questo modo l'inefficienza, che dipende dalla moda della Normale troncata, risulta spiegata dalle variabili esogene.

Un'alternativa è quella di limitare la proprietà di varianza costante della distribuzione assumendola come funzione di variabili esogene; in questo modo si corregge la specificazione del modello per l'eteroschedasticità facendo variare allo stesso tempo l'efficienza tecnica in funzione di variabili specifiche dell'azienda.

Hadri e Whittaker (2003) estesero la correzione per eteroschedasticità proposta da Hadri (1999) al modello di frontiere di produzione stocastiche per panel data. Il modello generale per dati panel è

$$y_{it} = X_{it}\beta + v_{it} - u_{it}, \quad (2.2.3.21)$$

dove  $v_{it}$  è la componente casuale dell'errore e

$$\begin{aligned} E[v_{it}] &= 0 && \forall i \text{ e } t \\ E[v_{it}, v_{jt}] &= 0 && \forall i \text{ e } j \text{ con } i \neq j \text{ e } \forall t \text{ ed } l, \end{aligned}$$

mentre  $u_{it}$  rappresenta la componente non negativa di inefficienza che dipende da fattori sotto il controllo dell'azienda e

$$\begin{aligned} E[u_{it}] &> 0 && \forall i \text{ e } t \\ E[u_{it}, u_{jt}] &= 0 && \forall i \text{ e } j \text{ con } i \neq j \text{ e } \forall t \text{ ed } l. \end{aligned}$$

La funzione di densità di  $\varepsilon_{it} = v_{it} - u_{it}$ , sotto le assunzioni che  $v_{it}$  si distribuisca come una normale e  $u_{it}$  come una semi normale ha la forma

$$f(\varepsilon_{it}) = \frac{2}{\sigma} \phi\left(\frac{\varepsilon_{it}}{\sigma}\right) \left(1 - \Phi\left(\frac{\lambda \varepsilon_{it}}{\sigma}\right)\right), \quad (2.2.3.22)$$

dove  $\sigma^2 = \sigma_v^2 + \sigma_u^2$ ,  $\lambda = \sigma_u^2 / \sigma_v^2$ ,  $\phi$  e  $\Phi$  sono rispettivamente la distribuzione di probabilità e la distribuzione standard della normale.

La specificità di questo modello è la possibilità di avere eteroschedasticità in una delle componenti di errore o in entrambe.

### Componente u eteroschedastica

La deviazione standard della componente  $u_{it}$  è espressa in funzione di variabili  $z$  (è inclusa un'intercetta) legate a caratteristiche specifiche delle aziende

$$\sigma_{uit} = \exp(z_{it}, \alpha), \quad (2.2.3.23)$$

dove  $\alpha$  è un vettore di parametri da stimare. La funzione di densità relativa a questo modello con eteroschedasticità nella componente  $u$  e  $\sigma_v = \exp(\gamma_o)$  è

$$f_{it}(\varepsilon_{it}) = \frac{2}{\sigma_{it}} \phi\left(\frac{\varepsilon_{it}}{\sigma_{it}}\right) \left(1 - \Phi\left(\frac{\lambda_{it}\varepsilon_{it}}{\sigma_{it}}\right)\right), \quad (2.2.3.24)$$

dove  $\sigma_{it}^2 = \sigma_v^2 + \sigma_{uit}^2$ ,  $\lambda_{it} = \sigma_{uit}^2 / \sigma_v^2$ ,  $\phi$  e  $\Phi$  sono state definite precedentemente, mentre la funzione di log-verosimiglianza<sup>7</sup> è data da

$$\ln L = \sum_{i=1}^N \sum_{t=1}^{T_i} \ln(f_{it}(\varepsilon_{it})). \quad (2.2.3.25)$$

### Componente v eteroschedastica

In questo caso è la componente casuale  $v_{it}$  dell'errore ad essere eteroschedastica

$$\sigma_{vit} = \exp(\mathbf{a}_{it}\boldsymbol{\gamma}), \quad (2.2.3.26)$$

dove  $\mathbf{a}_{it}$  è un vettore di variabili esplicative non stocastiche e generalmente legate alla dimensione aziendale ed include un'intercetta. La deviazione standard della componente  $u_{it}$  in questo caso è omoschedastica [ $\sigma_u = \exp(\alpha_0)$ ]. La funzione di densità è definita come la (2.2.3.24), ma in questo caso  $\sigma_{it}^2 = \sigma_{vit}^2 + \sigma_u^2$ ,  $\lambda_{it} = \sigma_u^2 / \sigma_{vit}^2$ .

### Componenti v ed u eteroschedastiche

La specificazione di entrambe le componenti di errore eteroschedastiche è data da

$$\sigma_{uit} = \exp(\mathbf{Z}_{it}\boldsymbol{\alpha}), \quad (2.2.3.27)$$

$$\sigma_{vit} = \exp(\mathbf{A}_{it}\boldsymbol{\gamma}). \quad (2.2.3.28)$$

In questa formulazione si considerano  $\sigma_{it}^2 = \sigma_{vit}^2 + \sigma_{uit}^2$ ,  $\lambda_{it} = \sigma_{uit}^2 / \sigma_{vit}^2$  nella funzione di densità.

<sup>7</sup> Si rimanda all'articolo di Hadri e Whittaker (2003) per le derivate prime della funzione di log verosimiglianza nei tre casi esaminati.

## ***2.3 Metodi non Parametrici***

L'approccio non parametrico richiede un numero limitato di ipotesi relative al processo di produzione, in quanto l'efficienza tecnica di una unità produttiva viene determinata mediante l'applicazione di tecniche di programmazione lineare, quindi senza dover specificare una particolare forma per la funzione di produzione. I metodi non parametrici più comunemente utilizzati sono il FDH (Free Disposal Hull) introdotto da Deprins, Simar e Tulkens (1984) e la DEA (Data Envelopment Analysis). Quest'ultima tecnica fu proposta da Charnes, Cooper e Rhodes (1978) e permette di misurare la massima contrazione (espansione) radiale degli input (output) compatibile con l'insieme delle possibilità produttive esistenti, o lo spazio di inviluppo. La principale distinzione rispetto al metodo FDH è che impone l'ipotesi di convessità delle combinazioni input-output, risulta quindi essere più restrittiva.

### ***2.3.1 Data Envelopment Analysis***

La Data Envelopment Analysis (DEA), attribuita a Charnes, Cooper e Rhodes (1978), è una tecnica di programmazione lineare utilizzata per misurare l'efficienza relativa di una unità produttiva rispetto alle prestazioni di altre unità produttive che producono il medesimo insieme di output utilizzando lo stesso insieme di input. Nelle sua formulazione originaria la procedura massimizza un indice di produttività media per ciascuna unità produttiva costruito come rapporto fra una combinazione pesata di output e una combinazione pesata di input.

La DEA tiene conto del fatto che ogni impresa può valutare in modo diverso gli input e output adottando un sistema di pesi diverso per ogni unità produttiva. Il sistema di pesi di ogni impresa sarà quello ad essa più favorevole.

La DEA può essere input orientata o output orientata. Nel primo caso si definisce la frontiera cercando di ridurre gli input e tenendo fisso il livello dell'output, nel secondo invece si ricerca il massimo output dato un certo livello

di input. I due metodi danno gli stessi risultati se si considerano rendimenti di scala costanti, ma in caso di rendimenti variabili i risultati non coincidono più.

### Rendimenti di scala costanti

L'obiettivo dell'analisi è misurare la performance di ciascuna impresa rispetto alle altre del campione.

Ogni unità produttiva si assume impieghi un vettore di K input  $x$  con  $x \in R^+$  per ottenere un vettore di M output  $y$  con  $y \in R^+$ , il problema viene impostato come

$$\min_{u,v} [v'x_i / u'y_i], \quad (2.3.1.1)$$

con i vincoli

$$v'x_i / u'y_i \geq 1, \quad (2.3.1.2)$$

$$u, v \geq 0, \quad (2.3.1.3)$$

in cui  $u$  e  $v$  sono pesi non negativi che applicati ai singoli vettori di input e output minimizzano il rapporto input/output di ciascuna impresa sotto la condizione che per nessuna unità produttiva tale rapporto sia inferiore a 1 ( $u$  e  $v$  sono variabili del problema). Questo modello si può scrivere in forma lineare:

$$\min_{u,v} [v'x_i], \quad (2.3.1.4)$$

sotto i vincoli

$$u'y_i = 1 \quad (2.3.1.5)$$

$$v'x_i \geq u'y_i \quad i = 1, \dots, N \quad (2.3.1.6)$$

$$u, v \geq 0. \quad (2.3.1.7)$$

Questo schema è duale al problema di programmazione lineare

$$\max_{\Psi, \lambda} \Psi, \quad (2.3.1.8)$$

con i vincoli

$$X\lambda \leq x_i \quad (2.3.1.9)$$

$$\Psi y_i \leq Y\lambda \quad (2.3.1.10)$$

$$\lambda \geq 0, \quad (2.3.1.11)$$

$y_i$  è un vettore  $M \times 1$  di output per l' $i$ -esima impresa;

$x_i$  è un vettore  $K \times 1$  di input per l' $i$ -esima impresa;

$Y$  è una matrice  $N \times M$  di output per le  $N$  imprese;

$X$  è una matrice  $N \times k$  di input per le  $N$  imprese;

$\lambda$  è un vettore  $N \times 1$  di pesi;

$\Psi$  è uno scalare.

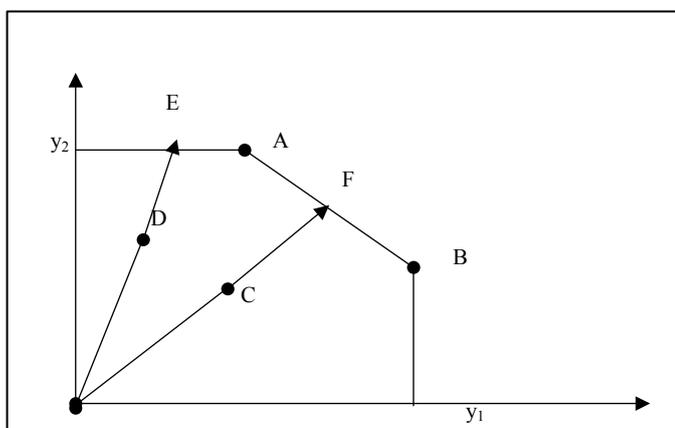
$\Psi$  sarà maggiore o uguale a uno;  $\Psi - 1$  è l'incremento proporzionale nell'output che può ottenere l' $i$ -esima impresa con input costanti. L'efficienza tecnica è data da  $1/\Psi$  e varia tra 0 e 1.

Il precedente problema viene risolto  $N$  volte, una per ogni impresa del campione. Ottenendo la stima del parametro  $\Psi$ , che fornisce informazioni sull'efficienza tecnica dell' $i$ -esima impresa, e del vettore  $\lambda$  sui pesi dei *peers* dell'impresa. I *peers* dell' $i$ -esima impresa sono quelle imprese efficienti, che giacciono sulla frontiera, e sono quindi vicine al punto (sulla frontiera) che occuperebbe la  $i$ -esima unità se fosse efficiente. Nella Figura 2.3.1.1 i *peers* dell'impresa corrispondente al punto C sono A e B, la loro combinazione lineare, considerando come pesi i  $\lambda$ , determina le coordinate del punto efficiente F.

Nella Figura 2.3.1.1 è possibile identificare gli *slacks* o carenze di output (eccessi di input nel caso input-orientato) dell'unità produttiva considerata rispetto a quella di riferimento. Questo problema è legato alla presenza di parti di frontiera parallela agli assi.

L'impresa inefficiente C per diventare efficiente si dovrebbe spostare in F, proiezione radiale di C sulla frontiera. Nel caso dell'impresa D la sua proiezione radiale sulla frontiera è E, ma questo punto non è efficiente in quanto l'output  $y_1$  potrebbe essere incrementato ancora (senza una corrispondente diminuzione di  $y_2$ ) come testimonia il punto A. La differenza tra l'output di A e di E è un *output-slack*. Gli *slacks* possono essere interpretati come inefficienza allocativa.

Figura 2.3.1.1 - Data Envelopment Analysis



### Rendimenti di scala variabili

L'assunzione di rendimenti di scala costanti è spesso troppo rigida<sup>8</sup>. Il modello con rendimenti variabili (considerando sempre il caso output-orientato), si ottiene facilmente da quello con rendimenti di scala costanti aggiungendo il vincolo di convessità:  $\mathbf{1}'_N N\lambda = 1$ . Sarà quindi

$$\max_{\Psi, \lambda} \Psi, \quad (2.3.1.12)$$

con i vincoli

$$\mathbf{X}\lambda \leq \mathbf{x}_i \quad (2.3.1.13)$$

$$\Psi \mathbf{y}_i \leq \mathbf{Y}\lambda \quad (2.3.1.14)$$

$$\mathbf{1}'_N N\lambda = 1 \quad (2.3.1.15)$$

$$\lambda \geq 0, \quad (2.3.1.16)$$

dove  $\mathbf{1}'_N N$  è un vettore  $N \times 1$  di uno. Questo approccio forma una copertura convessa di piani intersecati che avvolge i dati in modo più restrittivo della copertura conica dei rendimenti di scala costanti. I punteggi di efficienza nel caso

<sup>8</sup> Soprattutto non è in linea con la teoria dello sviluppo endogeno a cui si rifanno i modelli distrettuali del "learning by doing" (Arrow, 1962), in cui si afferma che un livello di conoscenza maggiore aumenta il profitto.

di rendimenti variabili saranno quindi maggiori o uguali a quelli ottenuti con rendimenti costanti

Il calcolo dei rendimenti di scala variabili consente inoltre di scindere l'efficienza dei rendimenti costanti in efficienza di scala e "pura" efficienza.

### 2.3.2 Free Disposal Hull (FDH)

Il metodo Free Disposal Hull (Deprins, Simar e Tulkens 1984) si basa sull'assunzione di eliminazione senza costo di input e output.

Dato insieme di produttori  $Y_0$ , l'insieme di riferimento  $Y(Y_0)$  è specificato nel modo seguente

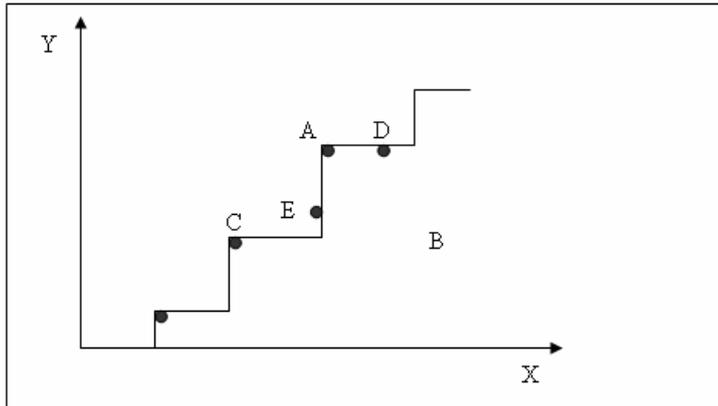
$$Y_{FDH}(Y_0) = \bigcup_{k|(x^k, y^k) \in Y_0} \{(x^k + \alpha, y^k - \beta), \alpha, \beta \geq 0\}, \quad (2.3.2.1)$$

dove  $\alpha$  e  $\beta$  sono vettori di eliminazione senza costo, rispettivamente di input e output. Data l'ipotesi di eliminazione senza costo, l'insieme di riferimento include tutti i produttori che utilizzano input uguali o maggiori e che producono output uguali o minore dell'osservazione  $i$ . Questo insieme di riferimento può indifferentemente essere un insieme di produzione, un *input requirement set* (per dati output) o un insieme delle possibilità produttive (per dati input).

Si consideri una tecnologia basata su 1 input e su 1 output (Figura 2.3.2.1) dove a ogni osservazione corrisponde un'unità produttiva. Partendo da una certa osservazione A, si definisce ogni osservazione posta alla destra e al di sotto di essa (cioè con maggiore input e stesso output come D, o con minore output e stesso input E, o ancora con maggiore input e minore output B) come *dominata* da A. Al fine di misurare l'inefficienza tecnica delle osservazioni dominate occorre stabilire se la misurazione viene effettuata nel senso dell'output o nel senso dell'input. Nel primo caso, l'inefficienza tecnica è data dal rapporto tra l'output dell'unità dominata e l'output dell'unità dominante ad essa relativa. Nel caso particolare in cui un'unità sia contemporaneamente dominata da due o più unità

sulla frontiera di riferimento, viene attribuito all'unità dominata il punteggio di efficienza relativo all'osservazione efficiente da cui viene maggiormente dominata. Se il punteggio di efficienza è pari a 1 allora l'osservazione è tecnicamente efficiente.

Figura 2.3.2.1 - Frontiera di produzione FDH



Nella pratica si è soliti analizzare insiemi di produzione multi-input e multi-output e pertanto le misure di efficienza utilizzate saranno misure radiali. In altre parole, per caratterizzare le unità produttive che si trovano sulla frontiera di produzione si considerano espansioni equiporzionali di tutti gli output o, analogamente, contrazioni di tutti gli input.

Una delle principali differenze tra le tecniche FDH e DEA è che nel primo caso un'osservazione inefficiente è necessariamente dominata da almeno un'altra osservazione realmente esistente, mentre nel secondo caso la frontiera è principalmente costituita da osservazioni virtuali costruite come combinazioni lineari di alcuni prodotti efficienti.

Uno dei maggiori vantaggi del FDH è di evidenziare alcune unità produttive realmente esistenti e di permettere confronti diretti fra queste ultime e le unità che esse dominano. Infine, l'assenza di qualunque ipotesi sulla convessità della tecnologia di produzione fa sì che le frontiere ottenute con il metodo dell'FDH abbiano maggiori probabilità di "accostarsi" ai dati rispetto a quelle ottenute con il metodo della DEA.



## Capitolo 3

### **UN MODELLO DELLA STIMA DEGLI EFFETTI PER GLI INCENTIVI ALLE IMPRESE**

#### *Introduzione*

In questa sezione si intende proporre la valutazione dell'impatto sulla TFP (Total Factor Productivity) di imprese agevolate a seguito di un intervento di sussidio all'accumulazione di capitale privato. Tale valutazione utilizza tecniche quasi-sperimentali e integra modelli econometrici per fornire un modello di stima degli effetti.

### ***3. Definizione di un modello di stima degli effetti***

In un quadro sperimentale, viene calcolato come differenza tra il valore medio di outcome tra agevolati e non agevolati, sotto l'ipotesi che il trattamento siano assegnate in modo casuale tra le imprese. Questa ipotesi è chiaramente violata quando lo strumento di politica utilizzato (come la L. 488/92) determini un processo di selezione. In presenza di selezione, la performance delle aziende non sovvenzionate può differire sistematicamente da quella delle aziende sovvenzionate anche in assenza di incentivi, con una significativa distorsione nella stima dell'impatto dei parametri. Il "bias" di selezione è un problema che può essere affrontato utilizzando un metodo di stima degli effetti opportuno.

#### ***3.1. Aspetti metodologici della valutazione dell'impatto***

L'obiettivo conoscitivo fondamentale della valutazione di impatto è stabilire se il «miglioramento» osservato tra i beneficiari sia da attribuirsi alla loro esposizione all'intervento, e non sia invece da attribuirsi ad altre determinanti che influiscono su di loro indipendentemente dall'intervento. Per «miglioramento» intendiamo il cambiamento nella direzione desiderata in quei comportamenti o in quelle condizioni dei beneficiari su cui l'intervento intende incidere per raggiungere i propri obiettivi.

Poniamo ad esempio che l'intervento in questione, incentivi alle imprese, abbia come obiettivo quello di aumentare la produttività delle imprese presenti in zone svantaggiate. Poniamo che un aumento di produzione (che può essere misurato con una molteplicità di indicatori) sia effettivamente osservato tra i beneficiari (coloro che hanno avuto la politica di incentivo) *dopo* la loro esposizione all'intervento (cioè dopo l'incentivo). L'obiettivo dell'analisi di

impatto in questo caso è di stabilire se questo cambiamento osservato (l'aumento di produttività) sia effettivamente da attribuirsi all'intervento, cioè esista tra i due un legame di causa-effetto.

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

$$\frac{\textit{beneficiari}}{\textit{potenziali} \cdot \textit{beneficiari}} \quad (3.1.1)$$

non fornisce alcuna informazione rispetto all'impatto diretto dell'intervento. Un valore più alto del rapporto (3.1.1) non indica infatti una maggiore efficacia dell'intervento. Solo nel caso in cui l'analisi di impatto abbia dimostrato che l'intervento ha un impatto positivo tra i beneficiari diretti (in questo caso imprese con intervento), allora un alto grado di diffusione suggerisce una maggior efficacia complessiva dell'intervento. Riteniamo quindi fonte di potenziale confusione definire il rapporto (3.1.1) come «indicatore di efficacia» dell'intervento: termini quali «grado di diffusione» o «grado di penetrazione» sono più appropriati a definirne la funzione.

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

$$\frac{\text{Beneficiari per cui si è osservato un miglioramento}}{\text{Totale dei beneficiari dell'intervento}} \quad (3.1.2)$$

non fornisce di per sé una misura dell'impatto diretto dell'intervento. Piuttosto, può essere definito come «tasso lordo di successo». Nell'ambito della letteratura sul controllo di gestione, rapporti simili a quelli rappresentati nella (3.1.2) sono spesso definiti come «indicatori di efficacia». Ad esempio, la «percentuale di

imprese che ottengono sovvenzioni trascorso un certo tempo dalla fine dell'intervento» verrebbe definito come indicatore di efficacia. Tuttavia, questo uso del termine non è legittimo se la nozione di efficacia che sta alla base della valutazione è quella di «capacità dell'intervento di ottenere un risultato migliore di quello che si sarebbe ottenuto in assenza di intervento». Definire il rapporto (3.1.2) come indicatore di efficacia significa presumere che, in assenza di intervento, non si sarebbe potuto verificare un cambiamento favorevole per nessuno dei beneficiari dell'intervento. Il che significa presumere che l'intervento può essere l'unico fattore in grado di influenzare la condizione o il comportamento in questione. Nell'esempio fatto, significa presumere che *nessuna* delle imprese avrebbero potuto svincolarsi dalla situazione di area depressa senza avere gli incentivi.

Quello appena discusso è, nella sua essenza, lo stesso problema metodologico che si deve affrontare per stabilire l'efficacia complessiva dell'intervento. Per questo adottiamo un concetto di causalità basato sulla costruzione di uno scenario controfattuale.

Nel caso dell'analisi di impatto diretto è necessario ricostruire la «situazione controfattuale», cioè *ricostruire cosa sarebbe successo ai beneficiari se non fossero stati esposti all'intervento*. Come si è detto, l'analisi di impatto diretto tende ad essere «miope», nel senso che spesso non è in grado di dire molto sulla capacità dell'intervento di incidere sul fenomeno preso nel suo complesso. Lo spostamento alla dimensione micro ha però anche un indubbio vantaggio: la possibilità di ricostruire la «situazione controfattuale» in questo contesto è resa possibile (almeno in linea di principio) dalla presenza di soggetti che, pur appartenendo alla platea dei potenziali beneficiari dell'intervento, non sono stati esposti all'intervento stesso. I non-beneficiari dell'intervento possono cioè fornire il «gruppo di controllo» per stabilire l'impatto dell'intervento.

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

### 3.1 ASPETTI METODOLOGICI DELLA VALUTAZIONE DELL'IMPATTO

Per spiegare in modo intuitivo come ciò possa avvenire, è utile introdurre il seguente ragionamento. Definiamo per brevità con termine di «migliorati» quei soggetti per cui si è osservato un «risultato favorevole», cioè un miglioramento nella condizione o comportamento che l'intervento intende modificare, rispetto ad un periodo precedente all'intervento. Possiamo quindi riscrivere l'espressione (3.1.2) nel modo seguente:

$$\frac{\text{migliorati}}{\text{beneficiari}} = \frac{\text{migliorati grazie all'intervento}}{\text{beneficiari}} + \frac{\text{migliorati in assenza dell'intervento}}{\text{beneficiari}} \quad (3.1.3)$$

L'impatto dell'intervento è rappresentato dal primo termine sul lato destro della (3.1.3). Il secondo termine sul lato destro rappresenta invece l'effetto di altri fattori che influenzano comportamenti o condizioni dei beneficiari indipendentemente dall'intervento. Una percentuale di coloro per cui si è osservato un cambiamento favorevole avrebbero potuto avere lo stesso cambiamento anche in assenza di intervento. Un modo alternativo per comunicare il contenuto della (3.1.3) è il seguente:

$$\text{miglioramento osservato} = \text{impatto dell'intervento} + \text{situazione controfattuale} \quad (3.1.4)$$

Il fatto che un miglioramento sia osservato tra i beneficiari (cioè che il lato sinistro della (3.1.4) sia positivo) non implica di per sé che l'impatto sia positivo. Nel caso estremo in cui tutto il miglioramento sia dovuto all'effetto di altri fattori, l'impatto dell'intervento è zero. Stabilire quanta parte del miglioramento osservato tra i beneficiari sia imputabile direttamente all'intervento è quindi il compito fondamentale dell'analisi di impatto. La semplice misurazione del miglioramento tra i beneficiari di un intervento, il lato sinistro della (3.1.4), non è sufficiente per condurre un'analisi di impatto, per quanto ne rappresenti una precondizione.

I due termini sul lato sinistro della (3.1.4) non sono osservabili direttamente, sono entrambi quantità ipotetiche. La strategia analitica usata dall'analisi di impatto, nella sua essenza, è quella di trovare un'approssimazione plausibile per la situazione controfattuale, di modo da ottenere, per differenza, una stima dell'impatto. Questo si può vedere mutando l'ordine dei termini della (3.1.4):

$$\text{impatto dell'intervento} = \text{miglioramento osservato} - \text{situazione controfattuale} \quad (3.1.5)$$

Nel caso dell'analisi di impatto diretto, la strategia per ottenere una stima della situazione controfattuale è quella di utilizzare comportamenti o condizioni osservati tra i *non-beneficiari* dell'intervento, utilizzando cioè un gruppo di controllo. L'idea è che, se beneficiari e non-beneficiari sono simili nelle condizioni di partenza (cioè in assenza di intervento), ciò che si osserva tra i non beneficiari approssima ciò che si sarebbe osservato tra i beneficiari se l'intervento non fosse stato attuato. Questa strategia è possibile cioè solo se vale la seguente eguaglianza:

$$\frac{\text{migliorati}}{\text{non - beneficiari}} = \frac{\text{migliorati} \cdot \text{in} \cdot \text{assenza} \cdot \text{di} \cdot \text{int ervento}}{\text{beneficiari}} \quad (3.1.6)$$

Se la (3.1.6) è valida (e questo rappresenta un grosso «se») si può riscrivere la (3.1.3) come segue:

$$\frac{\text{migliorati}}{\text{beneficiari}} = \frac{\text{migliorati} \cdot \text{grazie} \cdot \text{all}' \cdot \text{int ervento}}{\text{beneficiari}} + \frac{\text{migliorati}}{\text{non - beneficiari}} \quad (3.1.7)$$

A questo punto due dei tre termini contenuti nella (3.1.7) sono osservati, e il terzo termine, cioè l'impatto dell'intervento, è «identificato», cioè può essere ottenuto per differenza.

$$\text{impatto} \cdot \text{intervento} = \frac{\text{migliorati}}{\text{beneficiari}} - \frac{\text{migliorati}}{\text{non - beneficiari}} \quad (3.1.8)$$

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

- (i) osservazioni sul cambiamento siano disponibili per i beneficiari dell'intervento;
- (ii) osservazioni sul cambiamento siano disponibili per i non-beneficiari (gruppo di controllo);
- (iii) beneficiari e non-beneficiari siano «simili» in assenza di intervento, cioè valga l'eguaglianza (3.1.6).

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

Se i due gruppi sono diversi nelle condizioni di partenza (sarebbero cioè diversi anche in assenza di intervento), il confronto tra i due gruppi produce una stima *distorta* dell'impatto dell'intervento. Tale distorsione può andare in

entrambe le direzioni, cioè può portare sia a sottostimare l'impatto dell'intervento sui beneficiari sia a sovrastimarli. Se i non-beneficiari tendono ad essere in una situazione peggiore di come sarebbero stati i beneficiari in assenza di intervento, l'impatto sarà sovrastimato, attribuendogli meriti che non ha. Viceversa, se i non-beneficiari tendono ad essere in una situazione di partenza migliore, all'intervento verranno negati meriti che ha.

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

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

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

### ***3.1.1. Le valutazioni sperimentali***

La strategia di casualizzare l'esposizione all'intervento è disponibile in relativamente poche situazioni. Per essere più precisi, tale approccio è raramente applicabile agli interventi a regime, in quanto il tentativo di escludere alcuni individui dalla fruizione di un intervento esistente si scontra contro una lunga serie di ostacoli di natura etica, legale nonché organizzativa. Questo limita in gran parte l'applicazione della strategia sperimentale alla valutazione di interventi-pilota, che sono creati con l'intento esplicito di valutarne effetti e fattibilità. Tuttavia, anche in questo ambito ristretto, l'approccio sperimentale può incontrare difficoltà, in quanto le condizioni ideali in cui un esperimento dovrebbe essere condotto per fornire risultati utili non sono sempre soddisfatte (Hausman and Wise, 1985; Manski e Garfinkel, 1992). La correttezza delle stime dell'impatto può essere minacciata da complicazioni che possono sorgere durante la realizzazione dell'esperimento, quali la possibilità che il gruppo di controllo sia contaminato da prestazioni simili offerte da altri enti pubblici. Inoltre, l'integrità del processo di casualizzazione è spesso difficile da mantenere sul campo, e può essere compromessa dal comportamento degli operatori del servizio, che raramente condividono con i valutatori la stessa fede nella casualizzazione. Ad esempio, gli operatori possono tentare di alterare il risultato dell'assegnazione casuale sulla base delle loro preferenze e opinioni su chi debba usufruire dell'intervento e chi no. In questo modo si ricrea una forma di *selection bias*. In queste situazioni, quando l'equivalenza tra beneficiari e gruppo di controllo viene compromessa, si parla di mancanza di *validità interna* dell'esperimento. Quando ad essere compromessa è invece la rappresentatività e generalizzabilità dei risultati della valutazione sperimentale, si parla di mancanza di *validità esterna*. In questo caso le stime sono corrette, ma possono egualmente risultare di utilità limitata per il successivo disegno delle politiche. Per una serie di ragioni, i risultati dell'esperimento possono essere poco rappresentativi della situazione in cui l'intervento opererebbe una volta a regime: possono contribuire fattori quali,

per citarne solo alcuni, la scala ridotta degli interventi-pilota, la loro durata necessariamente limitata nel tempo, gli effetti indotti sul comportamento degli operatori e dei beneficiari dal fatto di sentirsi sotto osservazione; la dipendenza dei risultati ottenuti dal particolare contesto socio-economico in cui l'esperimento è stato realizzato. Quindi lo strumento della valutazione sperimentale va applicato solo dopo un'attenta verifica delle sue probabilità di successo. Si tratta di uno strumento assai potente, ma anche estremamente delicato nel suo uso.

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

### ***3.1.2. Le valutazioni non-sperimentali***

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

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

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

Un caso particolare è quello in cui siano state utilizzate graduatorie, e i beneficiari siano stati scelti come coloro che si situano al di sopra (o al di sotto) di una certa soglia o punteggio nella graduatoria. In questo caso una semplice stima dell'impatto può essere ottenuta confrontando la situazione post-intervento di coloro immediatamente al di sotto e immediatamente al di sopra della soglia (quindi molto simili tra loro, in quanto inclusi o esclusi dall'intervento in un modo *quasi* casuale: inoltre, quanto più la graduatoria è imperfetta, tanto gli ultimi inclusi e i primi esclusi saranno simili). Se la dimensione campionaria di tali casi marginali non è sufficiente, tecniche di regressione possono essere usate.

Si parla in questo caso di *regression-discontinuity design*, in quanto la discontinuità creata dalla graduatoria viene usata per identificare l'impatto dell'intervento.

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

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

La disponibilità di dati longitudinali consente anche di verificare la validità dei metodi per l'eliminazione del *selection bias* basati sulla modellazione del processo di selezione. L'idea è quella di applicare questi metodi ai dati pre-

intervento. Prima dell'intervento beneficiari e non-beneficiari differiscono non a causa dell'intervento (che non è ancora avvenuto) ma a causa di tutto ciò che li indurrà a fare una scelta diversa (motivazione, storia passata, ecc.). Se il metodo statistico proposto davvero elimina il *selection bias*, dovrebbe eliminare ogni differenza tra i due gruppi nell'indicatore osservato prima dell'intervento (Heckman e Hotz, 1989).

Due limitazioni all'applicabilità dei metodi basati sui dati longitudinali vanno notate. La prima è alquanto banale: i dati longitudinali sono costosi e difficili da raccogliere, e la decisione di raccogliarli richiede una dose notevole di lungimiranza. La seconda limitazione è che la condizione o il comportamento che l'intervento vuole modificare può non essere «ripetibile» nel tempo, cioè può non produrre osservazioni ripetute e confrontabili. Ad esempio, mentre il fatturato di una impresa consente osservazioni ripetute nel tempo, il peso di una specifica azione di incentivo al sostegno dell'impresa potrebbe non essere un'osservazione ripetibile nel tempo: per ogni impresa (esposta o meno all'intervento) si ha una sola osservazione. L'impatto di un intervento a favore delle imprese può essere valutato solo confrontando una singola osservazione per ciascun beneficiario e una per ciascun non-beneficiario.

### ***3.2. Principali aspetti della valutazione***

La maggior parte della letteratura sulla valutazione dell'impatto dei sussidi pubbliche alle imprese riguarda la stima degli effetti di causalità di un trattamento su uno o più risultati di outcome, come ad esempio l'occupazione e gli investimenti (Bronzini e De Blasio, 2006, e Bernini, Centra e Pellegrini, 2006, per l'Italia; Roper e Hewitt - Dundas, 2001, per l'Irlanda). Tuttavia, diverse politiche consentono diversi livelli di sovvenzioni a seconda dei progetti di investimento, della dimensione dell'impresa, la zona geografica, la regione e anche la "selezione" dell'impresa.

Il tema delle sovvenzioni riguarda i sussidi assegnate dalla legge 488/1992, la più importante politica di intervento per sovvenzionare l'accumulo di

capitale privato nelle regioni più povere italiane nel corso degli ultimi dieci anni. Le imprese sovvenzionate dalla L. 488/92 sono esposte a diversi livelli di trattamento, pertanto, studiarne l'impatto di tale trattamento, potrebbe trascurare la sua eterogeneità e nascondere alcune importanti caratteristiche della politica di intervento. Per esempio, i *policy makers* potrebbero essere interessati a diversi effetti connessi alle differenze nel livello di trattamento.

La letteratura relativa alla valutazione ex-post dell'impatto della L. 488 è sempre più recente e in forte crescita: effetti positivi della L. 488 sugli investimenti si sono trovati in Bronzini e De Blasio (2006) e Bernini, Centra e Pellegrini (2006). Pellegrini e Carlucci (2003) evidenziano un effetto positivo, Bernini, Centra e Pellegrini (2006) anche il fatturato, ma non sulla produttività. Bronzini e De Blasio (2006) indicano la presenza di (moderata) sostituzione intertemporale: aziende finanziate rallentano significativamente i loro investimenti negli anni seguenti il programma.

### ***3.2.1. Definizione degli obiettivi***

La domanda che tipicamente si pone la valutazione riguarda, principalmente, il *successo* di una politica pubblica e più specificamente la sua capacità di *produrre effetti in una direzione desiderata* (o di non causare effetti indesiderati). In entrambi i casi, ciò che è oggetto di indagine è la relazione di causa – effetto che si instaura tra l'attuazione della politica e i cambiamenti osservati nei fenomeni su cui la politica intende (o si trova a ) incidere.

Alla domanda “cosa è la politica pubblica?”, possiamo rispondere dicendo che il termine è utilizzato nella letteratura specialistica e nel parlar comune con una gran varietà di significati. Anche lasciando da parte la “politica” intesa come *politics* e concentrandosi esclusivamente sulla politica intesa come *policy*, restano comunque possibili diverse accezioni e sfumature.

### 3.2.2. *Politics o policy?*

In italiano esiste un solo termine, politica, per definire due concetti assai diversi tra loro che li anglosassoni tengono separati anche sul piano terminologico grazie all'uso dei vocaboli *politics* e *policy*. Per evitare possibili fraintendimenti chiariamo velocemente quali sono le differenze tra le due accezioni di "politica". La politica in senso di *politics* può essere definita come "la competizione esistente tra gruppi e/o tra individui per la conquista e il mantenimento delle risorse necessarie all'esercizio del potere". La politica nel senso di *policy*, spesso tradotta con l'espressione "politica pubblica", può essere definita come "l'insieme delle azioni messe in atto da attori pubblici al fine di affrontare un problema collettivo". È utile, ai fini della nostra trattazione, dare del concetto di politica pubblica un'interpretazione piuttosto restrittiva, pur sapendo che non si tratta della sola interpretazione possibile, né probabilmente della più esaustiva. Definiamo la (singola) politica pubblica come lo *sforzo consapevole, condotto da parte di attori pubblici, di cui "curare" un certo problema collettivo*.

Per problema collettivo intendiamo una situazione socialmente percepita come insoddisfacente e quindi meritevole di cambiamento (ed eventualmente meritevole di un intervento pubblico). Quando ad un problema fa seguito un intervento consapevole per affrontarlo (anche solo marginalmente o simbolicamente), noi parliamo di "politica pubblica". La domanda che spesso ricorre è cosa vogliamo valutare, e cioè "la politica ha avuto successo nel *curare* il problema?". Nel nostro caso potremmo dire: "Quale è stato l'effetto delle politiche sulle imprese?".

Il "successo" di una politica pubblica può senza dubbio essere valutato lungo diverse dimensioni. Più semplicemente diciamo che una politica abbia avuto successo per il fatto stesso di essere attuata, per il fatto cioè di passare dallo stato "dettato legislativo" a quello di "pratica amministrativa", dallo stato in cui vengono enunciate buone intenzioni a quello in cui le risorse disponibili sono effettivamente utilizzate, servizi reali sono stati davvero erogati, obblighi o divieti sono effettivamente imposti.

Le definizioni più accreditate di "valutazione" pongono al loro centro gli effetti dell'azione pubblica e la loro misurazione.

### ***3.3. Problemi e strategie di analisi della valutazione d'impatto***

Le maggiori difficoltà nel produrre uno studio di valutazione d'impatto derivano dal fatto che il valutatore deve essere in grado di confrontare che cosa è realmente accaduto alle unità *target* del programma con quello che sarebbe accaduto alle stesse unità, nelle medesime condizioni, ma in assenza del programma sotto analisi.

#### ***3.3.1. La stima del “controfattuale”***

L'impatto di un programma economico, e di un qualsiasi programma pubblico in generale, sulle unità *target* del programma è infatti comunemente definito come la differenza tra il valore della variabile che rappresenta il risultato di interesse osservato dopo che le unità *target* sono state esposte al programma ed il valore che la stessa variabile avrebbe assunto se le medesime unità non fossero state esposte al programma.

Quest'ultimo valore viene definito come il “controfattuale” (*counterfactual*, Moffit 1991) ed è ovviamente per sua stessa definizione non direttamente osservabile, generando il fondamentale problema metodologico di come ottenerne una stima accettabile. Nel caso, ad esempio, dei lavori di valutazione dei programmi statali statunitensi di incentivi alle imprese per favorirne la localizzazione in aree urbane particolarmente degradate (*Enterprise Zone Programs*), il “controfattuale” è tipicamente definito come l'incremento del numero di occupati registrato nelle aree *target* del programma (*Enterprise Zones*), tra un momento precedente ed uno successivo all'implementazione del programma, in una ipotetica situazione in cui le imprese localizzate in tali aree non abbiano invece ricevuto gli incentivi del programma (ad esempio: Dowall 1996 e Boarnet e Bogart 1995). Questo in quanto, in tale esempio, le unità *target*

del programma sono tipicamente rappresentate dai distretti in cui le imprese beneficiano degli incentivi degli *Enterprise Zone Programs*, mentre la variabile risultato di interesse è generalmente il numero di occupati presenti in tali distretti (ad esempio: Bostic 1996, Glaser e Yeager 1990, Papke 1994, Erickson e Friedman 1990, Talbot 1988).

#### ***3.3.2. Le “minacce alla validità” della valutazione: l’ “effetto delle variabili omesse” e quello di “selezione”***

I due maggiori problemi, generalmente riferiti come “minacce alla validità” della valutazione (*threats to validity*, Bartik e Bingham 1995), che si pongono nello stimare il “controfattuale” sono causati dall’influenza di avvenimenti esogeni al programma da valutare e dal processo di selezione dei partecipanti allo stesso.

Il primo di tali problemi è costituito dal fatto che cambiamenti estranei al programma possono influire sulle unità *target* in modo tale da fare pensare che il programma abbia funzionato, riducendo ad esempio il tasso di disoccupazione di una certa area economicamente depressa, quando invece tale risultato è il frutto di avvenimenti esogeni al programma stesso (per esempio il tasso di disoccupazione di un area si è ridotto per la generale ripresa economica del paese e non per lo specifico effetto del programma). Questo tipo di problema che la valutazione d'impatto deve affrontare è definito come l’ “effetto delle variabili omesse” (*omitted variable bias*), per indicare il rischio che i risultati della valutazione contengano una distorsione sistematica (che può essere positiva o negativa) dovuta ad una variabile esogena al programma di cui il valutatore non è in grado di quantificare l'effetto sulla variabile risultato osservata.

Il secondo problema riscontrabile nello stimare il "controfattuale" è legato al fatto che le unità *target* del programma, quelle cioè che ne ricevono i benefici, possono differire in maniera sistematica da quelle invece escluse dal programma.

Tale differenza sistematica, riferita in letteratura come “effetto di selezione” (*selection bias*, Heckman 1979, Heckman and Robb 1985, Robinson 1989), può causare una distorsione nella stima dell’impatto del programma quando quest’ultima è ottenuta confrontando la differenza tra la *performance* del gruppo di unità *target* e quella del gruppo di unità escluse (ed utilizzando quindi la performance delle imprese escluse dal programma come stima del “controfattuale”).

Nell’esempio degli *Enterprise Zone Programs* implementati negli Stati Uniti, una possibile fonte delle differenze sistematiche tra il gruppo di unità *target* del programma e quello delle unità escluse è rappresentata dal processo di selezione delle aree geografiche nelle quali gli incentivi alle imprese vengono offerti. Tali aree, come nella grande maggioranza dei programmi di incentivo allo sviluppo economico, non vengono infatti selezionate mediante un procedimento di assegnazione casuale che garantisce statisticamente l’omogeneità tra il gruppo di aree in cui gli incentivi del programma vengono offerti e le aree escluse. Esse sono invece selezionate dal legislatore secondo particolari criteri che tendono a favorire le aree maggiormente disagiate dal punto di vista economico (Erickson e Friedman 1991, Brintnall e Green 1988). In questo modo le imprese che ricevono gli incentivi del programma, in quanto localizzate in tali aree, operano in condizioni di svantaggio competitivo rispetto a quelle escluse dal programma. Tale svantaggio differenzerebbe la *performance* (rendendola tendenzialmente peggiore) delle imprese che ricevono gli incentivi del programma, rispetto a quella delle imprese escluse dal programma, anche in assenza di quest’ultimo. L’ “effetto di selezione”, quindi, tenderebbe in questo caso a distorcere la stima dell’impatto del programma sottostimandone gli effetti. In altre situazioni, l’ “effetto di selezione” può invece indurre a sovrastimare l’impatto di un programma economico. Questo può essere il caso, ad esempio, dei programmi in cui le imprese si autoselezionano per ricevere gli incentivi. In questo caso, è probabile che le imprese che richiedono i benefici del programma siano quelle con il *management* migliore e con il maggiore desiderio e capacità ad espandersi. E’ probabile quindi che esse otterrebbero una *performance* migliore delle altre imprese anche in assenza dei benefici del programma.

### **3.4. *Modello di analisi***

La valutazione dell'impatto degli incentivi dei programmi di aiuto alle imprese richiede idealmente di confrontare variazioni di TFP verificatesi nelle imprese beneficiarie, a cavallo del periodo di erogazione dei contributi, con una stima adeguata delle variazioni in termini di efficienza e di Total factor productivity che si sarebbero prodotte nello stesso periodo in assenza di contributi.

Un semplice confronto tra il livello di TFP registrato nelle imprese beneficiarie prima e dopo l'erogazione non fornirebbe invece indicazioni attendibili circa l'impatto netto delle misure di aiuto in esame. Un confronto di questo tipo non permetterebbe infatti di stimare l'effettivo contributo netto prodotto dagli interventi in esame in quanto parte del cambiamento della funzione di produzione in termini di TFP e di efficienza potrebbe essere stata causata dall'andamento congiunturale economico prodottosi in concomitanza al periodo di erogazione dei contributi e non dalla disponibilità di quest'ultimi: in presenza di congiunture economiche positive l'impatto dei contributi erogati tenderebbe così ad essere sovrastimato, mentre, in presenza di trend economici negativi, lo stesso impatto tenderebbe ad essere sottostimato.

Per individuare quanta parte del cambiamento in termini di efficienza e di produttività nelle imprese beneficiarie sia attribuibile all'effetto dei contributi erogati anziché all'effetto di effetti economici locali/regionali e/o di specifici settori di attività, è indispensabile disporre di dati occupazionali non solo delle imprese beneficiarie, ma anche delle imprese non beneficiarie da alcuna misura di intervento. Il confronto, con modelli di analisi, dei dati di bilancio registrati nelle imprese beneficiarie e non beneficiarie segue una strategia valutativa denominata "*comparison group design*" (Bondonio 1998, 2000). Tale strategia trova fondamento, in termini intuitivi, nel seguente ragionamento: se il cambiamento di efficienza e di produttività, nella fatti specie la TFP, rilevato nelle imprese beneficiarie fosse attribuibile al solo effetto di trend economici indipendenti dai contributi erogati, allora il medesimo cambiamento di efficienza e TFP dovrebbe

essere rilevabile anche in quelle imprese non beneficiarie con caratteristiche del tutto simili alle beneficiarie.

Lo specifico modello di analisi sviluppato per questo obiettivo è un modello parametrico di regressione lineare multipla che permette di separare l'effetto netto dei contributi erogati dall'effetto sia delle congiunture economiche generali, di settore e/o di distretto geografico, sia delle eventuali diversità sistematiche tra le caratteristiche iniziali delle imprese beneficiarie e non-beneficiarie.

### ***3.5. Le variabili oggetto di studio***

In questo paragrafo si analizzeranno le principali caratteristiche strutturali, economiche e finanziarie delle aziende. Ovviamente non si intende generalizzare il risultato ottenuto, ma ci si limita a studiare nello specifico il campione oggetto d'indagine. Il valore di sintesi che si è scelto di usare nell'analisi degli indicatori di bilancio è la mediana, in quanto meno sensibile all'influenza di dati anomali rispetto alla media.

#### ***3.5.1. Principali macroaggregati utilizzati nella riclassificazione del conto economico:***

##### **STATO PATRIMONIALE**

- **Totale immobilizzazioni sep.ind. di quelle conc. in loc. finanz.**
  - Totale Immobilizzazioni Immateriali;
  - Totale Immobilizzazioni Materiali;
- **Totale attivo**
- **Totale passivo**
- **Patrimonio netto**
- **Totale debiti**

##### **CONTO ECONOMICO**

- **Valore aggiunto** = differenza tra valore della produzione e costi della produzione;
- **Margine operativo lordo** = valore aggiunto meno costi del personale;
- **Margine operativo netto** = valore aggiunto al netto dei costi del personale, degli ammortamenti immateriali e materiali.

#### **3.6. *Struttura dei dati***

I dati utilizzati per l'analisi sono raccolti da due diverse indagini: i dati amministrativi della L. 488 e la banca dati AIDA, che contiene i bilanci consegnati da un sottoinsieme delle imprese italiane alle Camere di Commercio. L'integrazione tra i diversi insiemi di dati ha richiesto un processo complesso di pulizia e *merge* tra i due dataset. Il dataset ultimo, dove sono state poi effettuate le analisi, consiste di 665 progetti finanziati e 1.228 progetti non finanziati nel periodo 1995-2004. Per la validità del gruppo di controllo è stato fatto un confronto tra le caratteristiche principali dei progetti e degli indicatori nel campione di sovvenzioni e quello di imprese non sovvenzionate.

La valutazione dei dati propone inoltre un'indagine sulle tecniche che soddisfano le esigenze di analisi, ne descrive la possibile applicabilità e fornisce dei parametri di valutazione.

- **Trasformazione dei dati**

Le trasformazioni potrebbero essere dettate da esigenze di migliore comprensione del fenomeno e dalla massimizzazione delle informazioni contenute nei dati: una combinazione di variabili può portare alla creazione di una nuova dimensione maggiormente esplicativa. La trasformazione è dettata anche da esigenze di sintesi. Sostanzialmente le trasformazioni possono essere di due tipi come abbiamo visto prima:

1. *Trasformazioni sulla distribuzione dei dati: modificazioni matematiche alla distribuzione delle variabili.*
2. *Creazioni di dati: creazione di nuove variabili dalla combinazione di variabili esistenti per eseguire tassi, differenze e così via.*

Per le analisi statistiche, la fase della trasformazione dei dati è critica poiché alcune metodologie statistiche richiedono che i dati siano linearmente collegate ad una variabile obiettivo, normalmente distribuiti e liberi dagli outliers.

- **Eliminare gli effetti degli outliers**

Se i valori degli outliers sono estremi possono seriamente alterare l'accuratezza del modello che sarà costruito. È per questa ragione che si è deciso di eliminare il primo centile e l'ultimo centile della distribuzione.

- **Rendere i dati "facilmente" interpretabili.**

Molte variabili di transazioni, come valore aggiunto e il numero di dipendenti, hanno una distribuzione asimmetrica. Utilizzando i dati in questo formato le loro visualizzazioni sono difficili da interpretare. Usando uno schema di discretizzazione o prendendo il logaritmo, si trasformano tali variabili in modo da distribuire normalmente i dati, in questo modo il risultato è facile da interpretare e qualche volta si migliora anche la qualità dei risultati.

- **Discretizzazioni usando range**

La discretizzazione dei dati numerici usando dei range (di tutti i possibili valori) è un buon modo per normalizzare i dati. La creazione di range discreti permette una facile comprensione dei dati. Altre volte sono proprio i tool a dettare la rappresentazione dei dati. Per esempio, essere "agevolato" o "non agevolato" lavorano bene su variabili dicotomiche "1" o "0". Molte classificazioni richiedono valori continui, come la dimensione aziendale, e raggruppati in range come "Grandi", "Medie" e "Piccole". La trasformazione in range dei dati può essere

utile anche a rimuovere gli effetti degli outliers. Da notare che la codifica che si sceglie può influenzare il risultato del modello che si sta generando.

- **Trasformazioni matematiche**

Le funzioni matematiche applicate per trasformare i dati sono utili per standardizzare distribuzioni anormali e quando si tenta di linearizzare una variabile. Alcune funzioni matematiche includono trasformazioni logaritmiche, trasformazioni ad hoc e trasformazioni polinomiali. Le *trasformazioni logaritmiche* sono usate per normalizzare una variabile che ha una distribuzione asimmetrica. Queste tendono anche a ridurre gli effetti degli outliers. Le *trasformazioni polinomiali* sono utili quando si vogliono linearizzare i dati che sono distribuiti in modo continuo.

Operazioni che sono state condotte sui dati hanno riguardato anche “l’aggiustamento” delle variabili analizzate. I principali aggregati che abbiamo considerato sono stati: valore aggiunto, numero di dipendenti e immobilizzazioni, abbiamo considerato per questa variabile la somma delle immobilizzazioni materiali e di quelle immateriali. Il valore aggiunto e le immobilizzazioni sono state destagionalizzate e rese confrontabili per poter elaborare correttamente i nostri dati. Si è usato come indice deflazionante quello degli investimenti lordi, trasformando da prezzi correnti a prezzi costanti. Abbiamo reso il nostro dataset un panel, cioè abbiamo trasformato il nostro dataset originario da un formato *wide* ad un formato *long*. Oltre che ad eliminare i percentili estremi delle nostre distribuzioni, si è scelto di eliminare per la variabile Valore Aggiunto i valori inferiori a zero.

### ***3.7. Il metodo Diff-In-Diffs***

Nell’ambito della generale strategia di analisi del “disegno con gruppo di controllo non sperimentale”, la valutazione d’impatto con il metodo “*difference in difference*” (ad esempio: Papke 1994 e Boarnet e Bogart 1995) presuppone la disponibilità di due o più osservazioni nel tempo della variabile risultato (*Y*) di

ciascun *i.esima* azienda (sia esso appartenente al gruppo delle N imprese *target* del programma, sia essa appartenente invece al gruppo delle M imprese escluse da esso). Il vantaggio di avere a disposizione dati di questo tipo è che le osservazioni  $Y_{i,t-r}, Y_{i,t-r-1}, Y_{i,t-r-2}, Y_{i,t-r-3}, \dots, Y_{i,t-r-T}$ , antecedenti il momento (*t*) dell'implementazione del programma, possono essere utilizzate per fornire una stima più accurata del “controfattuale”, limitando le distorsioni da “effetto di selezione”. Questo in quanto la differenza tra i passati livelli dell'occupazione ( $Y_{i,t-r}, \dots, Y_{i,t-r-T}$ ) registrati nelle N aziende trattate dal programma e quelle registrate nelle rimanenti M aziende della regione può essere indicativa delle differenze sistematiche esistenti tra i due gruppi di aziende (dovute, ad esempio, a diversi livelli di reddito medio, di densità della popolazione e di capacità gestionale degli amministratori locali). Tali differenze, se non adeguatamente controllate, costituirebbero infatti una non trascurabile fonte di distorsione da “effetto di selezione” in quanto esse influirebbero, in maniera esogena rispetto al programma, sul livello di occupazione registrato sia nel gruppo dei distretti trattati che nel gruppo dei distretti esclusi dal programma.

E' importante notare, a questo punto, che la semplice valutazione dell'impatto del programma effettuata con due sole osservazioni per ciascun distretto (definibile come metodo “*difference in difference*” in senso stretto) è, in alcune precise condizioni, già in grado di controllare la distorsione da “effetto di selezione”. Il vantaggio di utilizzare un numero maggiore di osservazioni rilevate in periodi precedenti all'implementazione del programma (definibile come metodo “*difference in difference, in difference*” o “*difference in difference*” in senso lato) è quello che esse permettono di alleggerire considerevolmente le restrizioni poste da tali condizioni (Moffit 1991).

Un semplice modello formale è utile per chiarire questo concetto. Con l'utilizzo di due sole osservazioni per ciascuna azienda, l'impatto del programma (*a*) è idealmente definibile come:

$$a = E(Y_{i,t+p}^1 - Y_{i,t-r} | D_i=1) - E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i=1). \quad (3.7.1)$$

Dove:

$Y_{i,t+p}^1$  = il livello di occupazione registrato al tempo *t+p* nella *i.esima* azienda nell'ipotesi che quest'ultima sia stata “trattata” dal programma al tempo *t*;

$Y_{i,t+p}^0$  = il livello di occupazione registrato al tempo  $t+p$  nella medesima azienda  $i$ .esima nell'ipotesi, invece, che quest'ultima sia stata esclusa dal programma;

$D_i = 1$  se la  $i$ .esima azienda appartiene effettivamente al gruppo delle  $N$  unità *target* del programma;

$D_i = 0$  se la  $i$ .esima azienda appartiene effettivamente al gruppo degli  $M$  aziende escluse dal programma.

L'equazione (3.7.1) definisce l'impatto ( $a$ ) come la differenza tra il valore atteso  $E$  della crescita dell'occupazione registrata nel gruppo di imprese ( $D_i = 1$ ) *target* del programma e quella registrata nello stesso gruppo di imprese nella ipotetica situazione in cui essi siano invece escluse dal programma. Il valore dell'impatto ( $a$ ), ovviamente, non può essere direttamente calcolato in quanto  $E(Y_{i,t+p}^0 | D_i = 1)$ , che rappresenta la misura del "controfattuale", non è osservabile. Ciò che può essere calcolato è invece una stima ( $\hat{a}$ ) dell'impatto del programma, definibile come:

$$\hat{a} = E(Y_{i,t+p}^1 - Y_{i,t-r} | D_i = 1) - E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i = 0). \quad (3.7.2)$$

La stima ( $\hat{a}$ ) coincide con il valore ideale ( $a$ ) dell'impatto soltanto a patto che:

$$E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i = 0) - E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i = 1) = 0. \quad (3.7.3)$$

La condizione posta dall'equazione (3.7.3) richiede che il valore atteso della crescita dell'occupazione registrato nel gruppo delle  $M$  imprese escluse dal programma sia uguale a quello registrato nel gruppo dei rimanenti  $N$  imprese *target* del programma nella ipotetica situazione che queste ultime non ricevano invece i benefici del programma. Questa condizione può essere efficacemente illustrata graficamente (Figure B.1 e B.2). In Figura B.1 (Appendice B), essa è soddisfatta in quanto la differenza ( $A2' - A1$ ), che rappresenta  $E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i = 1)$ , è pari alla differenza ( $B2 - B1$ ) che rappresenta invece  $E(Y_{i,t+p}^0 - Y_{i,t-r} | D_i = 0)$ . La stima ( $\hat{a}$ ) dell'impatto del programma, pari alla differenza  $(B1 - A1) - (B2 - A2)$ , risulta quindi corretta. In Figura B.2 (Appendice B), invece, tale condizione non è soddisfatta, in quanto l'incremento ( $A2' - A1$ ) è minore di quello ( $B2 - B1$ , che è equivalente all'incremento  $A2'' - A1$ ) registrato nel gruppo delle imprese escluse dall'intervento. In questa situazione la stima ( $\hat{a}$ ) dell'impatto sottovaluta l'effettivo

impatto ( $a$ ) del programma che è invece rappresentato in Figura B.2 dalla differenza ( $A2-A2'$ ). La differenza ( $A2''-A2'$ ) costituisce quindi la misura della distorsione da “effetto di selezione” che l'utilizzo di due sole osservazioni per ciascuna impresa non è stato, in questa circostanza, in grado di eliminare. L'utilizzo di una terza osservazione – Figura B.3 (Appendice B) - rilevata al tempo ( $t-r-1$ ) per ciascuna *i*-esima impresa, permette invece di utilizzare la differenza ( $B1-B0$ ) - ( $A1-A0$ ) nei tassi medi di crescita dei due gruppi di imprese osservate nel periodo ( $t-r-1, t-r$ ) per stimare con più accuratezza il “controfattuale” ( $A2'$ ). Tale differenza è infatti utilizzata per correggere il tasso medio di crescita ( $B2-B1$ ), registrato tra il tempo ( $t-r$ ) e ( $t+p$ ) nel gruppo delle imprese escluse dal programma ( $D_i=0$ ), che è il tasso di crescita utilizzato per stimare il “controfattuale” ( $A2'$ ). In Figura B.3, ad esempio, la stima del “controfattuale” che si otterrebbe con l'utilizzo di due sole osservazioni ( $A2''$ ), viene corretta verso il basso ( $A2'''$ ), in quanto nel periodo ( $t-r-1, t-r$ ) il tasso di crescita “medio” ( $B1-B0$ ) del gruppo di imprese ( $D_i=0$ ) escluso dal programma è stato più alto di quello ( $A1-A0$ ) delle imprese ( $D_i=1$ ) *target* del programma. In questa circostanza sarebbe quindi errato assumere che il tasso di crescita “medio” ( $A2'-A1$ ) del gruppo di imprese trattate ( $D_i=1$ ), nell'ipotesi che esse non ricevano invece i benefici del programma, sarebbe stato analogo a quello ( $B2-B1$ ) delle imprese escluse. L'informazione aggiuntiva, ottenuta grazie ad una terza osservazione, permette dunque, nell'esempio di Figura B.3, di correggere la distorsione da “effetto di selezione” ( $A2''-A2'$ ) che si otterrebbe stimando l'impatto con due sole osservazioni e di ottenere quindi una più corretta stima dell'impatto del programma pari ( $\hat{a}=A2-A2'''$ ). Tale stima ( $\hat{a}$ ) è formalmente calcolata come:

$$\hat{a} = E [Y_{i,t+p}^1 - (Y_{i,t-r} - Y_{i,t-r-1}) \mid D_i=1] - E [Y_{i,t+p}^0 - (Y_{i,t-r} - Y_{i,t-r-1}) \mid D_i=0] \quad (3.7.4)$$

La stima ( $\hat{a}$ ) coincide con il valore ideale ( $a$ ) dell'impatto del programma soltanto a patto che la seguente condizione sia soddisfatta:

$$E [Y_{i,t+p}^0 - (Y_{i,t-r} - Y_{i,t-r-1}) \mid D_i=1] - E [Y_{i,t+p}^0 - (Y_{i,t-r} - Y_{i,t-r-1}) \mid D_i=0] = 0 \quad (3.7.5)$$

E' importante notare, tuttavia, come questa condizione sia molto meno stringente di quella dell'equazione (3.7.1): essa non assume più che i due gruppi ( $D_i=0$ ) e ( $D_i=1$ ) abbiano eguali valori attesi di crescita in assenza dell'intervento, ma richiede solamente che tali valori attesi siano eguali una volta corretti per tenere

conto delle differenze rilevate nei *trend* temporali antecedenti all'implementazione del programma.

In maniera analoga, l'utilizzo di un numero ancora maggiore di osservazioni, antecedenti l'attuazione del programma, permette di raffinare ulteriormente la correzione del tasso di crescita ( $B2-B1$ ) utilizzato per stimare il “controfattuale” ( $A2'$ ), alleggerendo in maniera ancora maggiore le restrizioni poste dalle condizioni rappresentate nelle equazioni (3.7.1) e (3.7.5).

### 3.7.1. *La modellizzazione del processo di selezione*

La modellizzazione del processo di selezione (Heckman e Hotz 1989 e Ashenfelter e Card 1985) adotta invece una diversa strategia per controllare la differenza tra il tasso di crescita ( $B2-B1$ ) di Figura B.2 e quella ( $A2'-A1$ ) che definisce idealmente il “controfattuale”. Questa strategia, che necessita di due sole osservazioni per ciascuno delle  $N+M$  imprese della regione, consiste nell'individuare esplicitamente le caratteristiche (che hanno un impatto sul livello  $Y$  dell'occupazione, in questo caso la variabile risultato di interesse della valutazione) in base alle quali il gruppo di imprese *target* è stato selezionato.

L'individuazione di queste caratteristiche permette di controllare tutte le differenze sistematiche osservabili dal valutatore esistenti tra il gruppo di imprese trattate e quello delle imprese escluse dal programma.

Questa operazione è compiuta stimando, per prima cosa, i parametri  $\beta$  del seguente modello:

$$P(D_i=1) = F(\beta_0 + \beta_1 X^1_{i,t-r} + \beta_2 X^2_{i,t-r}). \quad (3.7.1.1)$$

Dove, ad esempio:

$X^1_{i,t-r}$  è il reddito medio degli addetti dell'*i.esima* impresa al tempo  $t-r$ ;

$X^2_{i,t-r}$  è la densità dell'*i.esima* impresa al tempo  $t-r$ .

Il modello descritto nell'equazione (3.7.1.1) rappresenta una regressione *Logit* (o in altri casi *Probit*), dove  $F$  è la “funzione caratteristica” (*cumulative distribution function*) della distribuzione logistica (oppure la “funzione caratteristica” della distribuzione *normale standard*). La finalità di questo modello è di calcolare i

coefficienti  $\beta$  in modo tale che i valori stimati ( $\hat{D}_i$ ) della variabile dipendente risultino sempre compresi tra zero ed uno. Ciò permette di interpretare ciascun valore stimato ( $\hat{D}_i$ ) come la probabilità che la rispettivo impresa *i.esima* sia compreso tra quelle  $N$  trattate dal programma. La stima ( $\hat{D}_i$ ) riassume così l'effetto di tutte le caratteristiche osservabili delle imprese che differenziano sistematicamente il gruppo delle  $N$  trattate da quello delle  $M$  escluse dal programma.

Per controllare l'effetto di tali caratteristiche, che altrimenti sarebbero confuse con l'impatto del programma, generando un problema di distorsione da “effetto di selezione”, i valori stimati ( $\hat{D}_i$ ) sono successivamente inseriti nel seguente modello di regressione lineare:

$$(Y_{i,t+p} - Y_{i,t-r}) / (Y_{i,t-r}) = a_0 + aD_i + a_1\hat{D}_i + \varepsilon \quad (3.7.1.2)$$

Nel modello dell'equazione (3.7.1.2), dove la variabile dipendente rappresenta il la TFP (la variabile risultato  $Y$ ), l'impatto del programma è quindi stimato mediante il calcolo del coefficiente ( $a$ ). Tale coefficiente è separato da ( $a_1$ ) che incorpora invece l'impatto delle caratteristiche iniziali dell'impresa sulla TFP. Il valore della costante ( $a_0$ ), nella medesima equazione, stima invece il tasso di crescita osservabile nel caso in cui un distretto abbia una probabilità tendente a

zero di essere incluso nel gruppo degli  $N$  trattati dal programma e che esso non abbia effettivamente ricevuto i benefici del programma al tempo  $t$ .

La capacità di controllare ogni fonte di distorsione da “effetto di selezione” con questo metodo dipende, comunque, dalla possibilità di osservare e misurare tutte le caratteristiche delle imprese che differiscono sistematicamente tra il gruppo di quelli *target* e quello degli esclusi dal programma e che hanno un impatto non nullo sulla variabile risultato. Se infatti il valutatore non fosse in grado di osservare e misurare correttamente una di tali caratteristiche (ad esempio la capacità gestionale degli amministratori locali), essa verrebbe ad essere rappresentata nell'errore ( $\varepsilon$ ) dell'equazione (3.7.1.2). In questo modo tale errore risulterebbe correlato con ( $D_i$ ):

$$E(\varepsilon | D_i) \neq 0. \quad (3.7.1.3)$$

La stima dell'impatto del programma, in questo caso, sarebbe distorta, anche se tale metodologia garantirebbe comunque una notevole riduzione della portata del problema in quanto permette di controllare l'effetto delle differenze sistematiche osservabili ( $X^1$  e  $X^2$ ).

### **3.8. La selezione del gruppo di controllo per mezzo dello “statistical matching”**

Il metodo dello “*statistical matching*” (ad esempio: Barnow 1987, Rosebaum e Rubin 1984, 1985, Rubin 1973)<sup>1</sup>, per semplicità parliamo di imprese, mira a stimare l'impatto del programma selezionando per ciascun *i.esima* impresa, appartenente agli  $N$  trattati, un'impresa “gemella” ( $i^*$ ) scelta nel gruppo delle  $M$  escluse dal programma. Tale impresa  $i^*.esima$  è individuata in modo tale da minimizzare le differenze sistematiche, con l'omologa impresa *i.esima*, che possono influire sul livello  $Y$  della variabile risultato in maniera esogena dall'effetto del programma. Il problema della distorsione da “effetto di selezione” viene quindi affrontato costruendo un gruppo di controllo, della stessa dimensione del gruppo di unità *target*, il più simile possibile a quest'ultimo. Una restrizione implicita all'utilizzo dello “*statistical matching*” è che il numero delle  $N$  imprese *target* deve essere minore di quello delle  $M$  imprese escluse dal programma. Inoltre, tanto maggiore è la differenza ( $M-N$ ), tanto più grandi sono le potenzialità di tale metodo, in quanto tanto maggiore è il numero di imprese escluse, tanto più grande è il numero delle imprese tra cui scegliere le imprese “gemelle”, aumentando così la probabilità di ottenere un accoppiamento migliore.

Il parametro di partenza in base al quale selezionare le imprese “gemelle” è ottenuto mediante la stima dei parametri ( $\beta$ ) dello stesso modello *Logit* (o *Probit*)

---

<sup>1</sup> Il metodo dello “*statistical matching*”, nella forma qui presentata (denominata “*nearest available matching on the estimated propensity score*”) è stato sviluppato da Rubin (1973) nei primi anni settanta e successivamente perfezionata da Rosebaum e Rubin (1984 e 1985). Alcune delle sue numerose applicazioni nel campo della valutazione dei programmi pubblici sono riviste da Barnow (1987).

dell'equazione (3.7.1.1), dove le variabili  $X$  contengono tutte le caratteristiche osservabili delle  $N+M$  imprese della regione. Una volta calcolati i valori dei coefficienti ( $\beta$ ), il parametro di selezione è individuato, per ciascuna impresa, come il valore della stima ( $\hat{D}_i$ ), denominata in questo caso *propensity score* (Rosembaum e Rubin 1985). La selezione delle imprese  $i^*.esima$  che compongono il gruppo di controllo avviene quindi con il seguente procedimento. In primo luogo le  $N$  imprese *target* vengono separate dai rimanenti  $M$  ed ordinati con una procedura casuale. Il primo delle  $N$  imprese *target* viene quindi appaiato con l'impresa, delle  $M$  escluse, avente il *propensity score* ( $\hat{D}_i$ ) più simile al suo. Questi due imprese sono poi rimossi dalla lista dei due gruppi e lo stesso procedimento viene quindi ripetuto per il secondo ed ognuno delle successive imprese della lista delle  $N$  *target*. Una variante di questo procedimento è quella di condizionare la scelta dell'impresa "gemella" in base ad uno o più parametri vincolanti. Supponiamo, ad esempio, che la regione in cui le  $N+M$  imprese sono localizzati sia caratterizzata da tre distinte aree geografiche. Se l'appartenenza a due diverse aree geografiche fosse giudicata dal valutatore come una importante fonte di differenze sistematiche tra due imprese, l'impresa "gemella"  $i^*$  non verrebbe più selezionata come quella avente un valore del *propensity score* ( $\hat{D}_i$ ) in assoluto più simile a quello della impresa  $i.esima$ . Tale impresa sarebbe selezionata, invece, come quella avente il *propensity score* più simile nell'ambito delle sole imprese localizzate nella medesima regione della impresa  $i.esima$ .

Una volta che ciascuno delle  $N$  imprese *target*, mediante uno di questi procedimenti, sia stata appaiata con una corrispondente  $i^*.esima$  impresa gemella, l'impatto del programma ( $\hat{a}_i$ ), per ciascuna impresa  $i.esima$ , è semplicemente stimato come:

$$\hat{a}_i = (Y_{i,t+p} - Y_{i,t-r}) - (Y_{i^*,t+p} - Y_{i^*,t-r}). \quad (3.8.1)$$

A livello aggregato, una stima generale dell'impatto del programma ( $\hat{a}$ ) sull'insieme delle  $N$   $i.esime$  imprese *target* è invece ottenuta come:

$$\hat{a} = E(Y_{i,t+p} - Y_{i,t-r}) - E(Y_{i^*,t+p} - Y_{i^*,t-r})^2. \quad (3.8.2)$$

---

<sup>2</sup> Nell'equazione (3.8.2), il valore  $E(Y_{i^*,t+p})$  corrisponde al valore atteso dell'occupazione registrato nelle  $i^*.esime$  imprese "gemelle" al tempo  $(t+p)$ , e rappresenta la stima del "controfattuale".

### 3.9. Propensity score

Il *propensity score* è definito da Rosenbaum e Rubin (1983) come la probabilità condizionata di ricevere un trattamento prese le caratteristiche di pre-trattamento:

$$p(X) \equiv \Pr\{D = 1 | X\} = E\{D | X\} \quad (3.9.1)$$

Dove  $D = \{0,1\}$  è l'indicatore di esposizione al trattamento e la  $X$  è il vettore multidimensionale con le caratteristiche di pre-trattamento. Rosenbaum e Rubin (1983) mostrano che, se l'esposizione al trattamento è casuale all'interno delle celle definita da  $X$ , è anche casuale all'interno delle celle definite dai valori della variabile mono-dimensionale  $p(X)$ . Come risultato, dato una popolazione di unità indicate con  $i$ , se il *propensity score*  $p(X_i)$  è nota, la media dell' effetto del trattamento sui trattati (ATT) può essere stimata come segue:

$$\begin{aligned} \tau &\equiv E\{Y_{1i} - Y_{0i} | D_i = 1\} \\ &\equiv E\{E\{Y_{1i} - Y_{0i} | D_i = 1\}, p(X_i)\} \\ &\equiv E\{E\{Y_{1i} | D_i = 1\}, p(X_i)\} - \{E\{Y_{0i} | D_i = 1\}, p(X_i)\} \end{aligned} \quad (3.9.2)$$

dove il valore atteso esterno è una distribuzione finita di  $(p(X_i) | D_i = 1)$ , mentre  $Y_{1i}$  e  $Y_{0i}$  sono i potenziali risultati nelle due situazioni controfattuali (rispettivamente) di trattamento e di non trattamento. Formalmente, le seguenti due ipotesi sono necessari per derivare (3.9.2) data la (3.9.1) .

**Lemma 1.** Bilanciamento delle variabili pre-trattamento dato il *propensity score*.  
Se  $p(X)$  è il *propensity score*, allora:

$$D \perp X | p(X) \quad (3.9.3)$$

**Lemma 2.** Unconfoundedness è data dal propensity score.  
Supponiamo che l'assegnazione dei trattati sia unconfounded, cioè

$$Y_1, Y_0 \perp D | X \quad (3.9.4)$$

Allora l'assegnamento di trattati è unconfounded dato dal propensity score, cioè

$$Y_1, Y_0 \perp D \mid p(X) \quad (3.9.5)$$

Se l'ipotesi di bilanciamento è soddisfatta del Lemma 1, le osservazioni con lo stesso *propensity score* deve avere la stessa distribuzione di caratteristiche osservabili (e non osservabili) indipendentemente dal trattamento. In altre parole, per un dato *propensity score*, l'esposizione al trattamento è casuale, e quindi trattati e unità di controllo dovrebbero essere in media osservazionalmente identici. La scelta di quale termini includere è determinata solamente dalla necessità di ottenere delle stime del propensity score che soddisfi il *Balancing Hypothesis*, cioè Ipotesi di Bilanciamento.

Il programma `pscore.do` stima il propensity score e testa l'ipotesi di bilanciamento (Lemma 1) seguendo questo algoritmo<sup>3</sup>:

1. Stima del modello logit (o probit)

$$\Pr\{D_i = 1 \mid X_i\} = \Phi(h(X_i)) \quad (3.9.6)$$

dove  $\Phi$  denota la normale (logistica) e  $h(X_i)$  è una specificazione di partenza che comprende tutte le covariate come termini lineare, senza interazioni o termini superiori.

2. Divide il campione in  $k$  intervalli di parti uguali del *propensity score*, dove  $k$  è determinato da un utente e il valore predefinito è 5.
3. All'interno di ogni intervallo testa che la media del *propensity score* di trattati e delle unità di controllo non differisca.
4. Se il test fallisce in un intervallo, si suddivide l'intervallo in tempi e prova di nuovo.
5. Si continua fino a quando, in tutti gli intervalli, la media del *propensity score* di trattamento e di controllo non differiscono.

---

<sup>3</sup> L'ipotesi di Unconfoundedness del Lemma 2 non può essere testata.

6. All'interno di ciascun intervallo, si testa che la media di ogni caratteristica non differisca tra trattati e unità di controllo. Questa è una condizione necessaria per il Bilanciamento Hypothesis.
7. Se le medie di una o più caratteristiche differiscono, informa l'utente che la proprietà di bilanciamento non è soddisfatta e che si ha bisogno di una specificazione meno parsimoniosa di  $h(X_i)$ .

I passaggi 2-7 dell'algoritmo può essere ristretta. Questa restrizione implica che il test della proprietà di bilanciamento viene eseguita solo sulle osservazioni che hanno un *propensity score* appartenente alla intersezione che supporta il propensity score dei trattati e del gruppo di controllo. Imporre questa restrizione può migliorare la qualità del matching usando la stima della ATT.

### ***3.10. Stimatori di Matching della ATT basati sul Propensity Score***

Una stima del *propensity score* non è sufficiente per la stima della ATT utilizzando l'equazione (3.9.2). Il motivo è che la probabilità di osservare due unità con esattamente lo stesso valore del *propensity score*, in linea di principio è pari a zero in quanto  $p(X)$  è una variabile continua. Vari metodi sono stati proposti in letteratura per superare questo problema e quelli usati in questa tesi sono stati: *Nearest Neighbor Matching*, *Kernel Matching* e *Stratification Matching*.

Iniziando con quest'ultimo, lo *Stratification method* consiste dividere il range di variazione del *propensity score* in intervalli tale che all'interno di ciascuno di essi sia trattati che unità di controllo hanno, in media, lo stesso *propensity score*. Per motivi pratici si utilizzano gli stessi blocchi individuati dall'algoritmo che stima il *propensity score*. Quindi, all'interno di ciascun intervallo, in cui entrambi i trattati e le unità di controllo sono presenti, la differenza tra la media dei risultati dei trattati e dei controlli è così calcolata. L'ATT di interesse è infine ottenuta come media degli ATT di ogni blocco con peso dato dalla distribuzione dei trattati di unità in tutti blocchi. Uno degli scogli del metodo di stratificazione metodo è che si rigetti se le osservazioni nei blocchi, sia trattati o unità di controllo, sono

assenti. Questa osservazione suggerisce un alternativo metodo per la corrispondenza tra trattati e unità di controllo, che consiste nel prendere ogni unità trattata e cercare l'unità di controllo con il *propensity score* più vicino, vale a dire il più “vicino prossimo” che in letteratura va sotto il nome di *Nearest Neighbor*. Anche se non è necessario, il metodo è applicato di solito con sostituzione, nel senso che una unità di controllo può avere una migliore corrispondenza per più di una unità di trattamento. Una volta che ogni unità di trattamento è affiancata da una unità di controllo, la differenza è calcolata come il risultato tra l'unità di trattamento e il risultato della corrispondente unità di controllo. L'ATT di interesse è poi ottenuta da una media di tali differenze.

Mentre nel caso del metodo di Stratificazione ci possono essere trattati che possono essere unità scartate perché non è disponibile il controllo e il loro blocco, nel caso del metodo *Nearest Neighbor* le unità di tutti i trattati trovano una corrispondenza. Tuttavia, è evidente che alcuni di queste corrispondenze sono abbastanza povere, perché per alcuni trattati di unità il “prossimo più vicino” può avere un differente il *propensity score*. Il *Kernel Matching* offre una soluzione a questo problema. Infatti tutti i trattati sono abbinati con una media ponderata a tutte le unità di controllo, con pesi che sono inversamente proporzionale alla distanza tra il *propensity score* dei trattati e del gruppo di controllo.

### **3.10.1. *Nearest Neighbor Matching***

Sia  $T$  il set delle unità trattate e  $C$  il set delle unità di controllo, e  $Y_i^T$  e  $Y_i^C$  rispettivamente gli *outcomes* dei trattati e delle unità di controllo. Noto che  $C(i)$  è il set delle unità di controllo “abinate”(match) alle  $i$  unità trattate con un valore del propensity score  $p_i$ . Il set Nearest neighbor matching:

$$C(i) = \min_j \|p_i - p_j\| \quad (3.10.1.1)$$

Nel matching *nearest neighbor* si osserva che il *match* tra il numero di controls e le osservazioni  $i \in T$  da  $N_i^C$  è definito i pesi  $w_{ij} = \frac{1}{N_i^C}$  se  $j \in C(i)$  e  $w_{ij} = 0$  altrimenti. Allora la formula dello stimatore Nearest Neighbor può essere scritto come segue (dove M sta per Nearest Neighbor matching e il numero di unità del gruppo dei trattati è definito da  $N^T$ )

$$\begin{aligned} \tau^M &= \frac{1}{N^T} \sum_{i \in T} \left[ Y_i^T - \sum_{j \in C(i)} w_{ij} Y_j^C \right] \\ &= \frac{1}{N^T} \left[ \sum_{i \in T} Y_i^T - \sum_{i \in T} \sum_{j \in C(i)} w_{ij} Y_j^C \right] \\ &= \frac{1}{N^T} \sum_{i \in T} Y_i^T - \frac{1}{N^T} \sum_{j \in C} w_j Y_j^C \end{aligned} \quad (3.10.1.2)$$

dove i pesi  $w_j$  sono definiti da  $w_j = \sum_i w_{ij}$ . La varianza di questo stimatore assume pesi fissati e negli outcomes è assunta indipendenza tra le unità.

$$Var(\tau^M) = \frac{1}{N^T} Var(Y_i^T) + \frac{1}{(N^T)^2} \sum_{j \in C} (w_j)^2 Var(Y_j^C) \quad (3.10.1.3)$$

### 3.10.2. **Kernel Matching**

Lo stimatore Kernel è dato da:

$$\tau^K = \frac{1}{N^T} \sum_{i \in T} \left\{ Y_i^T - \frac{\sum_{j \in C} Y_j^C G\left(\frac{p_j - p_i}{h_n}\right)}{\sum_{k \in C} G\left(\frac{p_k - p_i}{h_n}\right)} \right\} \quad (3.10.2.1)$$

dove  $G(\cdot)$  è la funzione *kernel* e  $h_n$  è il parametro *bandwidth*. Il termine

$$\frac{\sum_{j \in C} Y_j^C G\left(\frac{p_j - p_i}{h_n}\right)}{\sum_{k \in C} G\left(\frac{p_k - p_i}{h_n}\right)}$$

è uno stimatore consistente dell'outcome controfattuale  $Y_{0i}$ .

### 3.10.3. Stratification Matching

Questo metodo è basato sulla stessa procedura di stratificazione usato per la stima del propensity score. In ciascun blocco definito da queste procedure le covariate sono bilanciate e l'assegnamento al trattamento deve essere considerato random. Sia  $q$  l'indice che identifica i blocchi definiti negli intervalli finiti del propensity score, in ciascun blocco si può scrivere che:

$$\tau_q^S = \frac{\sum_{i \in I(q)} Y_i^T}{N_q^T} - \frac{\sum_{j \in I(q)} Y_j^C}{N_q^C} \quad (3.10.3.1)$$

dove  $I(q)$  è il set di unità in  $q$  blocchi;  $N_q^T$  e  $N_q^C$  sono il numero di unità di trattati e di controllo nel  $q$ -esimo blocco. Lo stimatore ATT dell'equazione (3.9.2) nel metodo Stratification può essere così scritta:

$$\tau^S = \sum_{q=1}^Q \tau_q^S \frac{\sum_{i \in I(q)} D_i}{\sum_{\forall i} D_i} \quad (3.10.3.2)$$

assumendo indipendenti gli outcomes, la varianza di  $\tau^S$  è così ricavata:

$$Var(\tau^S) = \frac{1}{N^T} \left[ Var(Y_i^T) + \sum_{q=1}^Q \frac{N_q^T}{N^T} \frac{N_q^T}{N_q^C} Var(Y_j^C) \right] \quad (3.10.3.3)$$

### 3.11. *Modello Econometrico nell'analisi della TFP*

Molti sono stati i lavori in letteratura riguardanti sia l'efficienza sia la determinazione di una funzione di produzione, in cui la Total Factor Productivity sia la determinante del nostro modello. Negli ultimi decenni si è assistito ad un enorme concentrazione verso lo studio macroeconomico, con studi sulla performance a livello di aggregati nazionali (Tonini, Jorgeneel 2004), sia e soprattutto a livello microeconomico (Karagiannis e Tzouvelekas, 2004, Tzouvelekas, Pantrzios e Fotopulos, 2001 Battese e Borca 1997), anche grazie alla crescente disponibilità di dati a livello aziendale.

Il diffuso uso di incentivi ha generato molto interesse nell'effetto che gli stessi incentivi potrebbero avere e hanno avuto alla crescita dell'occupazione e la strategica interazione nelle imprese in negoziati nei quali le imprese annunciano la creazione di posti di lavoro in cambio di incentivi da parte del governo.

Consideriamo, in questa prima analisi, la relazione tra TFP e *inputs* espressa attraverso questa funzione di produzione:

$$\ln Y_i = \sum_{k=1}^K \ln X_{ki}$$

Scrivendo la relazione:

$$TFP = y - f(K,L) \quad (3.11.1)$$

La relazione (3.5.1) può essere scritta evidenziando tali relazioni tra TFP e *inputs*:

$$\begin{aligned} \ln TFP &= y_{it} - \alpha_0 + \alpha_1(\ln lav_{it}) + \alpha_2(\ln cap_{it}) + \alpha_3(\ln lav_{it})(\ln cap_{it}) + \alpha_4 t(\ln cap_{it}) \\ &+ \alpha_5 t(\ln lav_{it}) + \alpha_6(\ln cap_{it})^2 + \alpha_7(\ln lav_{it})^2 + \alpha_8 t + \alpha_9 t^2 - u_{it} \end{aligned} \quad (3.11.2)$$

dove  $y_{it}$  è espresso nel logaritmo e rappresenta l'output nell'*i*-esimo stabilimento al tempo *t* e misurato come il fatturato o valore aggiunto;  $lav$  si riferisce al logaritmo della media dell'occupazione nell'*i*-esimo stabilimento e nel tempo *t*;  $cap$  si riferisce al logaritmo degli input intermedi reali nell'*i*-esimo stabilimento nel tempo *t*, cioè lo *stock di capitale* misurato come il valore degli investimenti fissi dell'*i*-esima azienda al tempo *t*. Nel calcolare la TFP, un comune approccio è quello di ottenere stime dell'elasticità di output con rispettivi

*inputs* e allora trattare la TFP come combinazione del residuo  $u_{it}$  (3.11.2) e tempo  $t$  che rappresenta il cambiamento tecnologico. Un problema che si riscontra nella (3.11.2), usando le elasticità di output trovate è che avremmo bisogno di modellare le determinanti della TFP in modo da poter rispondere alla domanda se le sovvenzioni migliorano o peggiorano la performance delle imprese non incentivate. La stima della TFP ha bisogno di essere regredita attraverso un set di determinanti (come se ad esempio l'impresa ha ricevuto incentivo) e da un chiaro termine di errore  $u_{it}$  che cattura tutto ciò che non è possibile spiegare dal modello, dove  $u_{it} \sim \text{n.i.d}(0, \sigma^2)$  richiesto per stime efficienti e non distorte del modello. Da un altro punto di vista è possibile considerare la seguente relazione:

$$TFP = f(X, SUB_{it}) + e \quad (3.11.3)$$

dove TFP è la variabile dipendente legata da funzione  $f(\cdot)$  a SUB e ad una serie di covariate  $X_{it}$ . Considerando questa equazione:

$$\ln TFP = \beta_0 + \beta_1 SUB_{it} + \gamma_i X_{it} + \varepsilon_{it} \quad (3.11.4)$$

dove  $SUB_i$  è una variabile *dummy* che ha valore 1 se l'impresa ha avuto incentivo in qualsiasi momento durante il periodo di riferimento (cioè, questo permette di intercettare il valore medio della TFP nelle differenze di questi sottogruppi di imprese che sono state assistite durante il periodo)<sup>4</sup>;  $SUB_{it}$  fa riferimento al valore reale dell'incentivo durante il periodo che l'impresa ha partecipato al programma di intervento<sup>5</sup>. Infine  $X$  è un vettore di variabili che include tutti gli altri effetti.

Includendo queste variabili possiamo sostanzialmente rispondere al problema se la produttività differisce tra le imprese e le regioni (Harris, 2001), e che varia con l'età dell'impianto. In altri termini, si vuole conoscere se le imprese hanno avuto una maggiore o minore produttività media durante il periodo di assistenza e se si è avuto effetto sulla produttività. Nella specificazione dell'equazione (3.11.4)

<sup>4</sup> Con la specificazione del modello si fa interagire la  $SUB_i$  con altre variabili chiave (es: *factor inputs*, fattori di inputs) e quindi la nostra specificazione è abbastanza generalizzata (al massimo potrebbe essere equivalente stimando separatamente la funzione di produzione per due gruppi – trattati e non trattati – se tutti i termini del nostro modello sono significativi).

<sup>5</sup> Se si tratta come potenziale variabile endogena, si potrebbe dire che non siamo di fronte ad un problema di selezione del campione (come presentato in Faulk 2002, e Gabe e Kraybill 2002) nel nostro modello.

l'ipotesi che è stata fatta sulle imprese manifatturiere è che le unità eleggibili per il sussidio provengono dalla stessa popolazione. In contrasto con i recenti studi di Faulk (2002) e Gabe e Kraybill (2002) i quali usano entrambi il modello di selezione di Heckman (usando dati cross-sectional) stimando prima la propensione a partecipare al programma di incentivo (usando un probit model) e poi stimano il modello di crescita degli stabilimenti tenendo conto del fatto che la partecipazione è scelta endogena (e ne consegue che la stima dei loro parametri del modello di crescita degli stabilimenti hanno ricevuto assistenza parziale, in quanto non sono rappresentative della popolazione di tutte le imprese). Questi studi, che utilizzano l'approccio di tipo Heckman, tentano di tener conto della potenziale endogeneità.

Lo scopo è di garantire e di confrontare "like-with-like" in termini di SUB facendo la differenza (Klette, Møen, e Griliches, 2000), tra il TFP prima dell'intervento e successivo l'intervento per tutte le imprese oggetto d'esame.

### ***3.12. Modelli Effetti fissi e Modelli Effetti Random***

L'insieme dei dati osservabili, include le realizzazioni delle variabili del modello ( $y$  e  $x$ ), o di proxy (nel caso in cui le variabili teoriche di interesse non siano osservabili). Le osservazioni campionarie possono essere classificate in:

- (i) dati sezionali (*cross-section*); le osservazioni si riferiscono a diversi individui (paesi, famiglie, imprese, ecc.) rilevate nello stesso periodo (ad esempio, nel 1995);
- (ii) *serie storiche*; le osservazioni sono relative allo stesso individuo, o allo stesso aggregato, misurate in diversi periodi (ad esempio: il valore aggiunto dal 1995 al 1998);
- (iii) pooling di dati sezionali e serie storiche; le osservazioni sono bidimensionali, in quanto variano sia per individuo (o aggregato) sia nel periodo temporale di rilevazione.

I *panel data* appartengono a questa categoria di dati. A seconda del tipo di osservazioni campionarie di cui si dispone, la fase di stima incontra problematiche in parte diverse. Con riferimento ai casi (i) e (ii), possono essere sottolineati i seguenti tratti distintivi. L'utilizzo di dati cross-section per la stima è spesso accompagnato da problemi di eteroschedasticità dei residui, a causa della mancata spiegazione della elevata varianza delle osservazioni. Queste difficoltà sono riconducibili al fatto che la variabilità nei dati tende a concentrarsi per gruppi (famiglie ricche e famiglie povere, imprese di piccole e imprese di medio-grandi dimensioni, paesi europei a bassa e alta inflazione).

### 3.12.1. *Il modello con effetti fissi*

Nel modello con effetti fissi (fixed effects model, FE) si ipotizza che l'intercetta sia deterministica e vari da individuo a individuo (Cfr. equazione 3.12.1.1), oppure per individuo e nel tempo (Cfr. equazione 3.12.1.2).

$$y_{it} = a + b x_{it} + \sum_{j=1}^{N-1} \mu_j D_{ji} + \varepsilon_{it} \quad (3.12.1.1)$$

$$y_{it} = a + b x_{it} + \sum_{j=1}^{N-1} \mu_j D_{ji} + \sum_{s=1}^{T-1} \mu_s D_{st} + \varepsilon_{it} \quad (3.12.1.2)$$

Nel primo caso vengono utilizzate  $N-1$  variabili dummy individuali  $D_{ji}$ , nel secondo  $(N-1)+(T-1)$  dummy  $D_{ji}$  e  $D_{st}$  tali che:

$D_{ji}=1$  se  $i=j$ ,  $D_{ji}=0$  se  $i \neq j$ ;  $D_{st}=1$  se  $t=s$ ,  $D_{st}=0$  se  $t \neq s$ .

Gli  $N-1$  parametri  $\mu_i$  misurano gli scostamenti delle intercette di  $N-1$  individui da quella dell'individuo preso come base. Allo stesso modo, i  $T-1$  parametri  $\tau_t$  rappresentano gli scostamenti delle intercette di  $T-1$  periodi dal periodo-base.

Nota: se si inseriscono nei modelli (oltre all'intercetta)  $N$  dummy individuali e/o  $T$  dummy temporali, qualsiasi stima è resaimpossibile dalla perfetta collinearità con l'intercetta, in quanto

$$\sum_{j=1}^N D_{ji} = 1 \quad \left( \sum_{s=1}^T D_{st} = 1 \right)$$

### 3.12.2. *Il modello con effetti casuali*

Una alternativa al modello FE è il modello con effetti casuali (random effects model, RE) in cui  $\mu_i$  (e/o  $\tau_t$ ) vengono considerate stocastiche.

In particolare, si suppone che:  $E(\mu_i) = 0$ ;  $E(\mu_i^2) = \sigma_\mu^2$ ;  $E(\mu_i \mu_j) = 0$  (quando  $i \neq j$ ).

Ipotesi analoghe (media nulla, omoschedasticità e incorrelazione) sono formulate per gli effetti temporali  $\tau_t$ . In breve:

$$\mu_i \sim \text{IID}(0, \sigma_\mu^2); \quad \tau_t \sim \text{IID}(0, \sigma_\tau^2)$$

Inoltre, nel modello con effetti casuali si suppone che i movimenti stocastici di  $\mu$  e  $\tau$  siano incorrelati con tutti gli altri regressori e con l'errore  $\varepsilon$  dell'equazione. I modelli con effetti casuali possono essere riscritti nel seguente modo:

$$y_{it} = a + b x_{it} + v_{2it} \quad v_{2it} = \varepsilon_{it} + \mu_i \quad (3.12.2.1)$$

$$y_{it} = a + b x_{it} + v_{3it} \quad v_{3it} = \varepsilon_{it} + \mu_i + \tau_t \quad (3.12.2.2)$$

dove  $a$  e  $b$  sono parametri costanti (per individuo e nel tempo) e i movimenti stocastici dell'intercetta sono inglobati nei termini stocastici di errore composto  $v_{2it}$  (per il caso modello one way) e  $v_{3it}$  (per il caso modello two ways). Per questa ragione il modello RE viene talvolta definito come "error components model".

### 3.12.3. *Meglio i modelli panel con effetti fissi o casuali?*

Spesso nella scelta del modello ci si chiede quale di questi è il miglior modello per la stima e per i dati che si ha di fronte. Molte sono le argomentazioni

che favoriscono l'impiego del modello FE. I pregi di questa scelta sono dettati da dal fatto che è facile e immediato da stimare (no approssimazioni come in GLS). Inoltre lo stimatore LSDV è robusto all'omissione di variabili esplicative time-invariant. Anche nei casi in cui è valido il modello RE, lo stimatore LSDV è comunque consistente (perde solo l'efficienza). Infine vista la sistematicità delle ipotesi di specificazione, costituisce il primo passo dell'analisi panel.

Altre argomentazioni favoriscono invece l'impiego del modello RE perché permette di risparmiare molti gradi di libertà (specialmente per N grande); tiene conto anche della varianza between e non solo di quella within: efficiente uso (combinazione) di entrambe le fonti di informazione; tratta l'effetto individuale allo stesso modo del termine di errore (entrambi stocastici); permette di studiare l'influenza esercitata da variabili esplicative costanti nel tempo (genere, anni di studio, settore di appartenenza, area geografica); permette di fare inferenza sulla popolazione da cui il campione di dati è stato estratto.

La scelta di trattare gli effetti individuali come fissi o come casuali è delicata e può essere ricondotta ad una serie di fattori.

- Le determinanti degli effetti individuali, infatti se sono motivati da un elevato numero di circostanze casuali e non osservabili, è più indicato il modello random.
- Il numero di individui: per N grande e T piccolo, il modello FE ha pochi gradi di libertà (stime non troppo affidabili). Se l'interesse è soprattutto sulle pendenze, meglio filtrare le differenze individuali (senza stimarle esplicitamente) con il modello RE.
- La natura del campione: quando il campione è chiuso ed esaustivo (come nel caso di paesi o di settori), gli effetti fissi sono i naturali candidati. Quando il campione è aperto (N individui sono estratti da una popolazione), la specificazione random è più interessante.
- Il tipo di inferenza: spetta al ricercatore scegliere se desidera fare inferenza sulle caratteristiche della popolazione (perché interessato anche al comportamento degli individui esclusi dal campione) mediante inferenza non condizionale, o concentrarsi sugli effetti presenti nel campione (perché l'interesse è proprio su quelli) mediante inferenza condizionale ai soli individui nel campione.

### Test di Hausmann

Supponiamo di riprendere il discorso tra modelli ad effetti fissi e modelli ad effetti random, e considerando un modello che utilizza dati in forma panel:

$$y_{it} = \alpha_0 + x_{it}'\beta + z_{it}'\alpha + \varepsilon_{it} \quad (3.12.3.1)$$

con  $i=1, \dots, n$ ,  $t=1, \dots, T$ .  $\alpha_0$  è la costante,  $\beta$  rappresenta il vettore dei coefficienti,  $x_{it}$  contiene  $K$  regressori e la matrice  $z_{it}$  è un insieme di variabili non osservate che catturano gli effetti specifici legati alle caratteristiche degli individui  $\varepsilon_{it}$  è il termine di errore. Le variabili contenute in  $z_{it}$  sono non osservate e possono essere correlate o non correlate con i regressori. Nel primo caso, nel modello (3.12.3.1) il valore dell'intercetta varia da individuo ad individuo, ed è costante nel tempo, e viene definito ad effetti fissi (Fixed Effects), assumendo la seguente relazione funzionale:

$$y_{it} = \alpha_0 + x_{it}'\beta + \alpha_i + \varepsilon_{it} \quad (3.12.3.2)$$

Nel secondo caso, il modello viene definito a effetti casuali (Random Effects) nel senso che le variabili della matrice  $z_{it}$  sono non osservate e non correlate con le  $x_{it}$ . In tal caso la relazione funzionale diventa:

$$y_{it} = \alpha_0 + x_{it}'\beta + u_i + \varepsilon_{it} \quad (3.12.3.3)$$

dove  $u_i$  è la componente casuale specifica al gruppo degli individui. La differenza tra il modello ad effetti fissi ed effetti casuali sta proprio nella differenza tra la componente individuale  $\alpha_i$  nel modello ad effetti fissi e la componente sempre individuale  $u_i$  nel modello ad effetti casuali. Nel modello ad effetti fissi il termine  $\alpha_i$  è deterministico e cattura le caratteristiche del singolo individuo. Per questo motivo esso assume un valore diverso per ciascuno degli individui facenti parte del Panel, è costante nel tempo ed essendo legato alle caratteristiche individuali, esso è correlato con le variabili  $x_i$ . Nel modello definito ad effetti casuali, il termine  $u_i$  ha una distribuzione casuale specifica per ciascun gruppo di individui.

Per la sua natura stocastica il termine  $u_i$  non è correlato con le  $x_i$  essendo queste variabili di natura deterministica. L'ipotesi del modello a effetti fissi è plausibile per stime nelle quali, come il nostro caso, le osservazioni sono unità territoriali perchè è ragionevole supporre che le caratteristiche rappresentate dalle variabili non osservate siano costanti nel tempo (Green, 2003). Tuttavia la scelta relativa a quale dei due modelli rappresenta la migliore specificazione per il caso considerato può essere fatta con il test di Hausmann.

Tale Test si basa sulla statistica:  $W = (\beta_f - \beta_r)'(V_f - V_r)^{-1}(\beta_f - \beta_r)$  dove  $\beta_f$  e  $\beta_r$  sono rispettivamente i parametri stimati con il metodo a effetti fissi ed a effetti casuali e sono le relative matrici varianze-covarianze  $V_f$  e  $V_r$ . Secondo l'ipotesi nulla la statistica  $W$  si distribuisce come una  $\chi^2(k)$  dove  $k$  è il numero dei coefficienti presenti nel vettore  $\beta$  intercetta esclusa. L'ipotesi nulla del test è l'assenza di correlazione tra la componente stocastica e le  $x_{it}$  e pertanto l'assenza di una differenza sistematica tra i valori dei parametri stimati nelle due regressioni. Se tale ipotesi è verificata le stime Random Effects sono preferibili perchè più efficienti. Nel caso contrario il metodo di stima migliore è quello ad effetti fissi in quanto le stime risultano essere consistenti.

## Capitolo 4

### **L'IMPATTO DEI SUSSIDI AL CAPITALE: UN' ANALISI EMPIRICA DELLA L. 488/92**

#### *Introduzione*

In questa sezione si andranno ad illustrare quelli che sono stati i risultati dell'analisi valutativa sull'impatto dei sussidi al capitale erogati dalla L. 488/92 sulla TFP. Si è costruito un dataset incrociando i dati della L.488/92 e i dati di bilancio della banca dati AIDA ottenendo come ultimo un dataset per le analisi. La TFP è la misura di efficienza calcolata attraverso metodi parametrici; la valutazione dell'impatto di sussidi al capitale ha visto l'applicazione dei due metodi di stima: uno con metodo parametrico, sulla base di una stima più tradizionale DID, e uno non parametrico basato su tecniche di matching. Entrambe hanno portato alla stessa conclusione, mettendo in evidenza la mancanza di un effetto degli incentivi sulle imprese agevolate.

#### **4.1      *Analisi dei risultati: i dati***

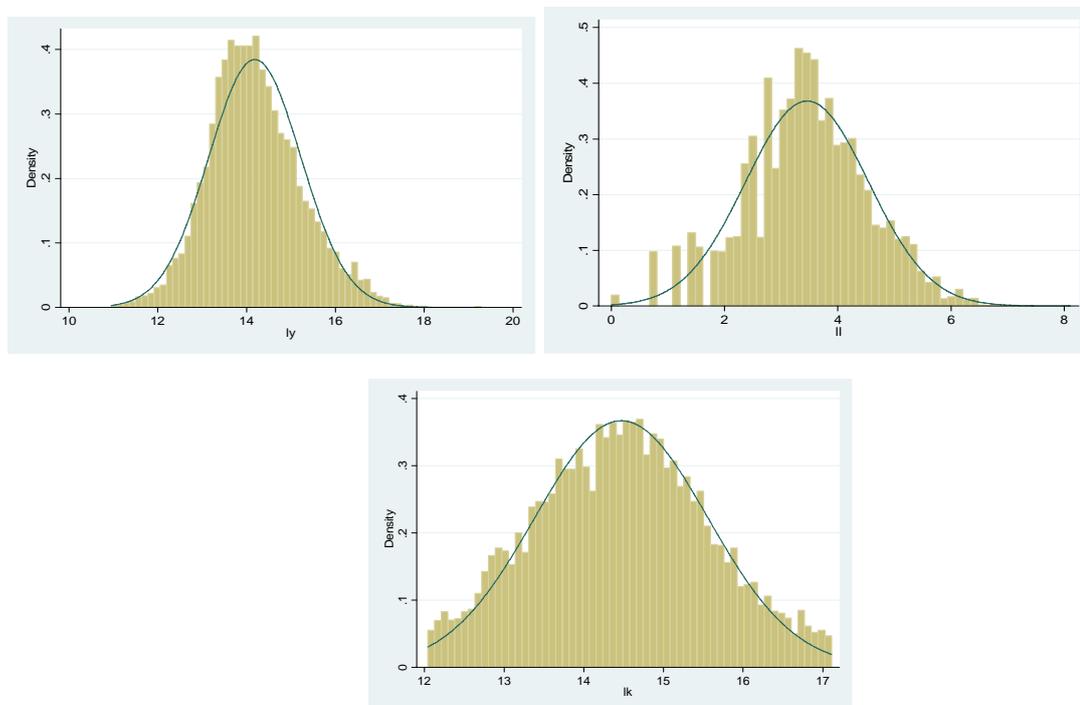
È di fondamentale importanza rilevare quali siano state le problematiche riguardanti la costruzione del dataset utilizzato. Per la validità dei risultati è decisivo disporre di dati che siano “qualitativamente buoni”. Il dataset originario era composto da 1.893 aziende di cui 665 finanziate e 1.228 non finanziate. Si è osservato che dalle prime analisi esplorative i dati presentavano alcuni problemi, come ad esempio la presenza di outliers e di dati mancanti. Si è quindi cercato di capire le ragioni di queste problematiche e giungere ad una soluzione. La principale variabile in cui si sono riscontrati tali problemi è stata il valore aggiunto. Infatti si è notata la presenza di valori negativi e di valori *missing* nelle imprese. La decisione che si è presa è stata appunto quella di eliminare in primo luogo i valori negativi, mentre i *missing* sono stati trattati direttamente in fase di stima. Eliminando i valori negativi del valore aggiunto, l'impresa veniva completamente eliminata dal dataset e allo stesso tempo si avevano 337 osservazioni in meno, cioè pari a quasi il 18%. Attraverso l'ausilio di statistiche riassuntive e grafici esplorativi sulle variabili, si è osservato che il dataset non era ancora abbastanza depurato da valori anomali e quindi si è deciso di eliminare i valori estremi delle nostre distribuzioni.

Diversi sono stati i tentativi sui percentili che ha riguardato il valore aggiunto e il capitale (immobilizzazioni materiali e immateriali): eliminando il 5% dal primo e dall'ultimo percentile dalla distribuzione di entrambi gli aggregati al netto dei valori negativi per il valore aggiunto, ottenendo un totale di 1.525 imprese (553 agevolate e 972 non agevolate). Il successivo tentativo è stato combinando il 5% sul primo e ultimo percentile del valore aggiunto e l'1% sul primo e ultimo percentile del capitale ottenendo così un totale di 1.587 ( 578 agevolate e 1.009 non agevolate). Ulteriore due prove sono state fatte: la prima togliendo l'1% sul primo e ultimo percentile del valore aggiunto e il 5% dagli estremi della distribuzione delle immobilizzazioni ottenendo un totale di 1.585 (578 agevolate e 1.007 non agevolate); la secondo si è provato a togliere l'1% dal primo e ultimo

## 4.1 ANALISI DEI DATI

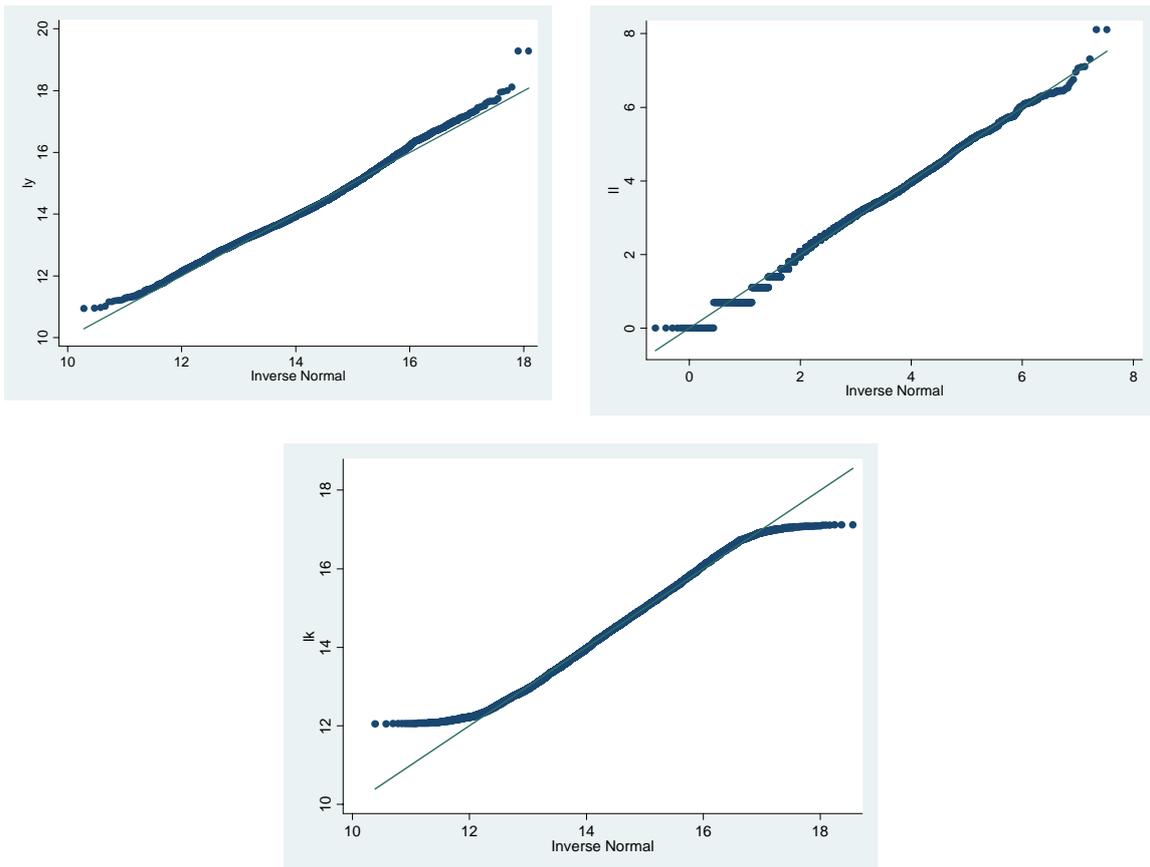
percentile di entrambi le variabili: valore aggiunto (al netto dei valori negativi) e del capitale (immobilizzazioni materiali e immateriali) ottenendo così un totale di 1.661, di cui 603 finanziate e 1.058 non finanziate. Le variabili valore aggiunto e immobilizzazioni (materiali e immateriali) hanno avuto una trasformazione per renderle confrontabili nel tempo: sono state deflazionate per l'indice degli investimenti fissi, e rese a prezzi costanti. Il presupposto teorico per la trasformazione delle variabili non implica che i dati debbano essere distribuiti come una normale. Rispetto alle trasformazioni che sono state rese necessarie, questo comunque fornisce un buon risultato da cui poter partire. Il dataset che si è deciso di adottare per l'analisi è quello riguardante 1.661 imprese (si è eliminato l'1% dei valori estremi da capitale e lavoro). Andando a rappresentare graficamente la distribuzione delle nostre variabili osservate, espresse nei logaritmi e deflazionate, attraverso un istogramma a barre (Figura 4.1.1) e il Q-Q plot (Figura 4.1.2), si osserva la normalità dei dati dopo il trattamento rispetto agli outliers. Le variabili del valore aggiunto, del lavoro e del capitale, espressi nei logaritmi rispettivamente come  $\ln(va)$ ,  $\ln(lav)$ ,  $\ln(k)$ .

FIGURA 4.1.1 – Distribuzione di  $\ln(va)$ ,  $\ln(lav)$ ,  $\ln(k)$



Possiamo dire che dalle distribuzioni di densità si osserva una normalità distributiva, infatti rispetto alla retta, si ha una buona approssimazione dei nostri dati alla retta ideale gaussiana.

FIGURA 4.1.2 – Q-Q plot di  $\ln(va)$ ,  $\ln(lav)$ ,  $\ln(k)$



Dalle prime operazioni eseguite sui dati possiamo ritenerci abbastanza soddisfatti delle distribuzioni ottenute. Molto complesso è stato il lavoro per poter ottenere questi risultati che sicuramente hanno migliorato il *dataset* originario. Queste distribuzioni sono state utilizzate per costruire il modello e per la stima.

Lo scopo della nostra analisi è quella di andare ad osservare attraverso metodologie diverse se l'incentivo alle imprese della L. 488 ha avuto effetto su una misura di TFP (Total Factor Productivity). Partendo e mantenendo rigore logico nell'analisi, si è voluto costruire una tavola che riassume in maniera più dettagliata alcune statistiche riguardanti le nostre variabili (percentili, medie,

#### 4.1 ANALISI DEI DATI

kurtosis e Skewness) per riuscire a valutare la dimensionalità delle variabili base per la stima del modello *trans-log*, descritto nel capitolo 3 (Tabella 4.1.1).

Tabella 4.1.1 - Statistiche descrittive del  $\ln(va)$ ,  $\ln(lav)$ ,  $\ln(k)$

ly				
	Percentiles	Smallest		
1%	<b>11. 79431</b>	<b>10. 94468</b>		
5%	<b>12. 53969</b>	<b>10. 95583</b>		
10%	<b>12. 90171</b>	<b>10. 97718</b>	Obs	<b>12635</b>
25%	<b>13. 46099</b>	<b>11. 01452</b>	Sum of Wgt.	<b>12635</b>
50%	<b>14. 11458</b>		Mean	<b>14. 27497</b>
		Largest	Std. Dev.	<b>1. 255611</b>
75%	<b>14. 90684</b>	<b>19. 86546</b>		
90%	<b>15. 86389</b>	<b>19. 86546</b>	Variance	<b>1. 57656</b>
95%	<b>16. 57248</b>	<b>19. 88715</b>	Skewness	<b>. 9449502</b>
99%	<b>18. 60326</b>	<b>19. 88715</b>	Kurtosis	<b>4. 808265</b>
ll				
	Percentiles	Smallest		
1%	<b>. 6931472</b>	<b>0</b>		
5%	<b>1. 386294</b>	<b>0</b>		
10%	<b>2. 079442</b>	<b>0</b>	Obs	<b>12392</b>
25%	<b>2. 772589</b>	<b>0</b>	Sum of Wgt.	<b>12392</b>
50%	<b>3. 465736</b>		Mean	<b>3. 535104</b>
		Largest	Std. Dev.	<b>1. 238444</b>
75%	<b>4. 234107</b>	<b>8. 641356</b>		
90%	<b>5. 09375</b>	<b>8. 641356</b>	Variance	<b>1. 533743</b>
95%	<b>5. 634789</b>	<b>8. 846066</b>	Skewness	<b>. 4017515</b>
99%	<b>7. 192934</b>	<b>8. 877103</b>	Kurtosis	<b>3. 999603</b>
lk				
	Percentiles	Smallest		
1%	<b>11. 43734</b>	<b>10. 98684</b>		
5%	<b>12. 20486</b>	<b>10. 98785</b>		
10%	<b>12. 79622</b>	<b>10. 98794</b>	Obs	<b>12635</b>
25%	<b>13. 60615</b>	<b>11. 00187</b>	Sum of Wgt.	<b>12635</b>
50%	<b>14. 46949</b>		Mean	<b>14. 49681</b>
		Largest	Std. Dev.	<b>1. 387197</b>
75%	<b>15. 3014</b>	<b>19. 24986</b>		
90%	<b>16. 19814</b>	<b>19. 26377</b>	Variance	<b>1. 924315</b>
95%	<b>16. 92007</b>	<b>19. 27589</b>	Skewness	<b>. 3287599</b>
99%	<b>18. 5395</b>	<b>19. 27589</b>	Kurtosis	<b>3. 546859</b>

Fonte: elaborazioni su dati L.488 e AIDA

Osservando attentamente questa tabella si osserva come la mediana e la media siano quasi coincidenti evidenziando ancora maggiormente la simmetria della distribuzione precedentemente descritta nelle Figure 4.1.1 e 4.1.2. Si osserva anche la distribuzione dei percentili e delle concentrazioni delle osservazioni attorno ai valori mediani. Le osservazioni sono di tipo panel, N imprese individuate ciascuna in T tempi. Il periodo preso in considerazione va dal 1995 al 2004<sup>1</sup>.

<sup>1</sup> I dati sono quindi stati trasformati da tipo *wide* a tipo *long* attraverso il software Stata 10 utilizzato nell'elaborazione degli output dell'intero lavoro.

Il panel è sbilanciato, cioè per ogni unità statistica non sono disponibili osservazioni per uno stesso numero di periodi (Tabella 4.1.2).

Tabella 4.1.2 – Caratteristiche dei Pattern

i d:	<b>1, 2, ..., 1893</b>	n =	<b>1661</b>
t:	<b>1, 2, ..., 10</b>	T =	<b>10</b>
	Del ta(t) = <b>1 uni t</b>		
	Span(t) = <b>10 peri ods</b>		
	(i d*t uni quel y i denti fi es each observati on)		

Freq.	Percent	Cum.	Pattern
269	16.20	16.20	1111111111
200	12.04	28.24	.1111111111
164	9.87	38.11	111111111.
121	7.28	45.39	.11111111.
63	3.79	49.19	11111111..
61	3.67	52.86	.111111..
53	3.19	56.05	.111111..
50	3.01	59.06	111111..
33	1.99	61.05	..11111111
647	38.95	100.00	(other patterns)
1661	100.00		XXXXXXXXXX

Fonte: elaborazioni su dati L.488 e AIDA

Il controllo del bilanciamento del panel è una operazione molto importante e indispensabile, così come l'inizializzazione di T e di N. Dalla tabella 4.1.2 si osserva la distribuzione delle 1.661 aziende nei T periodi (T=1,...10, cioè dal 1995 al 2004). La colonna *Pattern* restituisce la presenza o l'assenza nel *t-esimo* periodo dell'impresa appartenente al nostro campione. In totale solo 269 sono presenti in tutti i periodi, mentre le imprese presenti solamente in un solo anno sono solamente 29. Le imprese, considerando il panel sbilanciato, sono state prese tutte in esame. A partire dalla struttura del panel è anche possibile individuare la distribuzione temporale delle imprese agevolate e non agevolate ordinate per T periodi (Tabella 4.1.3).

Tabella 4.1.3 – Distribuzione temporale delle imprese

agev_	anno									
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
0	440	815	894	858	924	936	924	872	751	481
1	319	493	510	506	531	551	561	522	469	278
Totale	759	1.308	1.404	1.364	1.455	1.487	1.485	1.394	1.220	759

Fonte: elaborazioni su dati L.488 e AIDA

## 4.2 *Analisi dei risultati: Stima del modello*

La prima scelta da fare nell'affrontare il tema dell'analisi di una funzione di produzione è quella dell'utilizzo di metodi di stima parametrici o non parametrici. Come è stato già ampiamente detto, in questo contesto si è preferita l'impostazione parametrica, in quanto permette di ottenere informazioni non solo sul livello TFP delle imprese, ma anche sulla funzione di produzione, tramite la stima delle elasticità dei fattori produttivi. I metodi parametrici, inoltre, per la loro caratteristica di robustezza verso la presenza di possibili *outliers*, sono spesso impiegati nell'analisi di dati di bilancio. La possibilità di fare test inferenziali sui parametri, infine, rappresenta un notevole vantaggio a favore dei metodi parametrici, anche se si è consapevoli che le imposizioni fatte sulla forma funzionale possono essere molto restrittive.

La specificazione del modello è stata effettuata mediante una funzione di produzione di tipo trans-log, scelta per la sua flessibilità. D'altro canto però, l'elevato numero di parametri da stimare e la non immediata interpretazione degli stessi pone la necessità di testare l'effettiva preferibilità di questa forma funzionale rispetto per esempio ad una specificazione più semplice quale la Cobb-Douglas. Il modello *Cobb-Douglas* rispetto al modello trans-log presenta una chiave interpretativa molto più semplice, essendo le elasticità pari ai coefficienti del capitale e del lavoro. Molte prove sono state fatte e hanno mostrato come la funzione trans-log sia da preferire alla Cobb-Douglas.

Sono stati quindi effettuati dei test per verificare se la forma funzionale trans-log è quella che meglio si adatta ai dati. In particolare, è stato applicato il test classico del rapporto di verosimiglianza (Likelihood Ratio) *LR* definito come

$$\lambda = -2[\ln L(H_0) - \ln L(H_1)] \quad (4.2.1)$$

dove  $L(H_0)$  rappresenta il valore della funzione di verosimiglianza quando si impongono restrizioni sui parametri mediante l'ipotesi  $H_0$ , mentre  $L(H_1)$  è la

verosimiglianza del modello generale della funzione di produzione. La statistica test si distribuisce come un  $\chi^2$  con gradi di libertà pari al numero di restrizioni imposte.

La prima ipotesi esaminata (Tabella. 4.2.1) è stata la possibilità di rappresentare i dati con una forma funzionale semplice come la Cobb-Douglas con una componente di trend ( $\beta_3=\beta_6=\beta_7=\beta_9=0$ ). In questo caso si è rifiutata l'ipotesi nulla in quanto il valore di  $\lambda$  era decisamente superiore al valore teorico del  $\chi^2$  tabulato. La seconda ipotesi formulata è stata quella di considerare una funzione trans-log completa visto il valore di  $\lambda$  era decisamente inferiore al valore teorico del  $\chi^2$  tabulato. Di conseguenza, il modello di riferimento sarà questo la funzione trans-log.

Tabella. 4.2.1 Test LR per la valutazione del modello

Funzione	Restrizioni imposte	Log-verosim.	$\lambda$	$\chi^2$	Decisione
<b>Trans-log</b>		<b>-24913</b>	<b>1.56</b>	<b>5.76</b>	<b>Non Rifiuto</b>
Cobb-Douglas (con trend)	$\beta_3=\beta_6=\beta_7=\beta_9=0$	-2611.8	231.96	9.43	Rifiuto

Il modello proposto è:

$$\ln TFP = y_{it} - \alpha_0 + \alpha_1(\ln lav_{it}) + \alpha_2(\ln cap_{it}) + \alpha_3(\ln lav_{it})(\ln cap_{it}) + \alpha_4 t(\ln cap_{it}) + \alpha_5 t(\ln lav_{it}) + \alpha_6(\ln cap_{it})^2 + \alpha_7(\ln lav_{it})^2 + \alpha_8 t + \alpha_9 t^2 - u_{it}$$

La TFP è stimata come una regressione in cui si osserva una componente costante, una componente ad effetto fisso, una componente di trend, una componente di trend quadratica e l'errore. Nella componente ad effetto fisso vengono introdotte delle interazioni tra il capitale e il lavoro e la componente di trend. I parametri della funzione di produzione nel caso della forma trans-log non sono direttamente interpretabili, ma di questo è opportuno calcolare le elasticità di ogni input (tabella 4.2.1.1).

## 4.2 ANALISI DEI RISULTATI: STIMA DEL MODELLO

La stima del modello ad Effetti Fissi (Tabella 4.2.1) ha evidenziato un trend negativo che ha avuto effetto nelle interazioni con il capitale (coefficiente positivo  $t_{lk} = 0.0067$ ), mentre con il lavoro (coefficiente negativo  $t_{ll} = -0.0074$ ). I segni dei coefficienti sono legati principalmente ai coefficienti del lavoro e del capitale rispettivamente  $lk = -0.8952$  e  $ll = 1.0024$ . Si può osservare che i coefficienti, di tutte le componenti del modello, siano tutti significativi<sup>2</sup>. Il modello stimato, per un panel non bilanciato, ha visto un vettore  $X$  di variabili indipendenti, e come dipendente il valore aggiunto. A ciascun coefficiente troviamo l'errore standard associato e il valore del test  $t$  dove l'ipotesi nulla è che il coefficiente non sia significativamente diverso da zero.

Tabella 4.2.1 – Stima del modello Fixed-effects

ly	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
t	-.0466421	.01663	-2.80	0.005	-.07924 - .0140442
lk	-.8952413	.0847658	-10.56	0.000	-1.061398 - .7290846
ll	1.002418	.0601784	16.66	0.000	.8844571 1.120379
lk_ll	-.0742101	.0048261	-15.38	0.000	-.0836701 - .0647502
t_lk	.0067142	.0012922	5.20	0.000	.0041812 .0092473
t_ll	-.0074418	.0014575	-5.11	0.000	-.0102987 - .0045848
lk2	.046003	.0033121	13.89	0.000	.0395106 .0524954
ll2	.0738043	.0039077	18.89	0.000	.0661445 .0814642
t2	-.001391	.0004861	-2.86	0.004	-.0023439 - .0004381
_cons	16.72505	.5554601	30.11	0.000	15.63625 17.81386
sigma_u	.58123344				
sigma_e	.33748842				
rho	.74786212	(fraction of variance due to u_i)			
F test that all u_i=0: F( 1660, 10722) = 9.28 Prob > F = 0.0000					

Fonte: elaborazioni su dati L.488 e AIDA

Per la determinante scelta del modello ad effetti fissi è stata favorita dal fatto di aver compiuto il test di Hausman tra un modello ad effetto fisso e un modello ad effetto random<sup>3</sup> che giustifica la nostra scelta (tabella 4.2.2).

<sup>2</sup> L'output ottenuto è stato ottenuto attraverso la stima di un panel con effetti fissi.

<sup>3</sup> Hausman è un test del confronto fra il risultato di stimatori alternativi.  $H_0: Cov(xit, \mu_i) = 0$  (non correlazione fra gli effetti casuali e i regressori del modello)  $H_1: Cov(xit, \mu_i) \neq 0$ . Sotto  $H_0$  il modello RE è il migliore, mentre, sotto  $H_1$ , le proprietà statistiche dello stimatore GLS del modello RE vengono meno. La stima LSDV del modello FE è consistente sia sotto  $H_0$  che sotto  $H_1$ , ma non

Tabella 4.2.2 – Test di Hausman

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) fe	(B) re		
t	-. 0466421	-. 0519803	. 0053382	. 0007229
lk	-. 8952413	-1. 290704	. 3954632	. 0434051
ll	1. 002418	1. 278399	-. 2759807	. 025535
lk_ll	-. 0742101	-. 1020074	. 0277973	. 0018026
t_lk	. 0067142	. 0046621	. 0020521	.
t_ll	-. 0074418	-. 0049095	-. 0025323	.
lk2	. 046003	. 0654913	-. 0194883	. 0016646
ll2	. 0738043	. 110099	-. 0362947	. 0019875
t2	-. 001391	-. 0001402	-. 0012508	.

b = consistent under H<sub>0</sub> and H<sub>1</sub>; obtained from xtreg  
B = inconsistent under H<sub>1</sub>, efficient under H<sub>0</sub>; obtained from xtreg

Test: H<sub>0</sub>: difference in coefficients not systematic

$$\chi^2(9) = (b-B)' [(V_b-V_B)^{-1}] (b-B) = 897.61$$

Prob>chi2 = 0.0000  
(V\_b-V\_B is not positive definite)

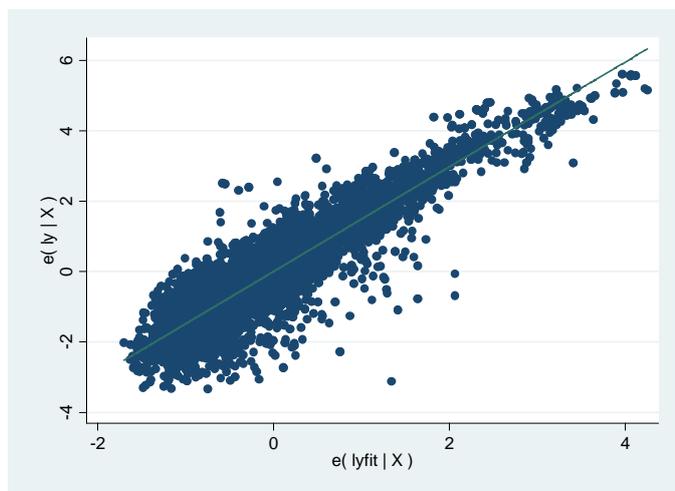
Fonte: elaborazioni su dati L.488 e AIDA

A partire dal modello e dalle stime si è evidenziato e calcolata la misura di Total Factor Productivity per ciascuna azienda avendo ben presente che il modello andava a stimare un effetto fisso per tutte le aziende e ponendo come base teorica l'effetto individuale. La TFP è calcolata rispetto all'anno prima dell'investimento e l'anno successivo la fine dell'investimento, così da ottenere una valutazione dell'effetto dell'incentivo sulla misura di TFP.

Un aspetto importante conseguente al calcolo della TFP è nell'andare ad osservare i residui data dalla differenza dell'outcome  $ly$  e quello stimato dal modello  $ly\_fit$  (Figura 4.2.1).

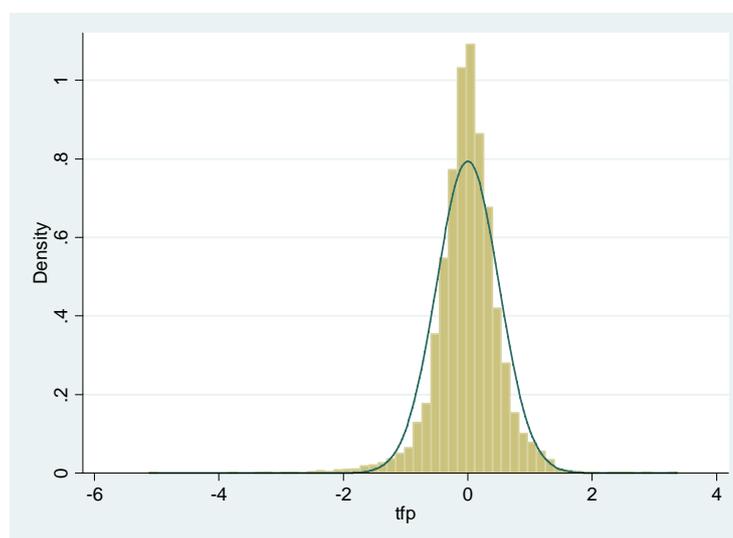
---

è efficiente sotto H<sub>0</sub>. Perciò, sotto H<sub>0</sub>, le stime saranno statisticamente simili e, quindi, la scelta ricade su quelle GLS del modello RE; viceversa sotto l'alternativa. Formalmente, si definisce la differenza fra stime:  $q = b^{LSDV} - b^{GLS}$ ; da cui  $b^{LSDV} = q + b^{GLS}$  e quindi  $Var(b^{LSDV}) = Var(q) + Var(b^{GLS})$  [Hausman dimostra che  $Cov(q, b^{GLS}) = 0$ ]. Da quest'ultimo risultato si ha che:  $Var(q) = Var(b^{LSDV}) - Var(b^{GLS})$ . La statistica test è quindi:  $m = q' [Var(q)]^{-1} q$  che si distribuisce, sotto H<sub>0</sub>:  $q = 0$ , come  $\chi^2_{(K)}$ . Il modello RE "passa" il test anche quando la non significatività dello scarto fra stime dipende da una loro bassa significatività.

FIGURA 4.2.1 – Distribuzione dei Residui tra  $ly$  e  $ly\_fit$ 

La differenza verticale tra ciascuna coppia di  $ly$  e  $ly\_$  è approssimativamente normale, come dovrebbe essere per una corretta applicazione (Figura 4.2.2).

FIGURA 4.2.2 – TFP Density



### 4.2.1. *Le elasticità ed i rendimenti di scala*

I parametri della funzione di produzione trans-log non hanno una immediata interpretazione economica, è quindi necessario calcolare le elasticità del prodotto rispetto agli input. Le espressioni 4.2.1.1 e 4.2.1.2 descrivono la derivata del logaritmo del valore atteso di Y rispetto al logaritmo del capitale e del lavoro.

$$\frac{\partial \ln E(Y)}{\partial \ln Cap} = \alpha_2 + \alpha_3 \ln(Lav) + \alpha_4 t + 2 \cdot \alpha_6 \ln(Cap) \quad (4.2.1.1)$$

$$\frac{\partial \ln E(Y)}{\partial \ln Lav} = \alpha_1 + \alpha_3 \ln(Cap) + \alpha_5 t + 2 \cdot \alpha_7 \ln(Lav) \quad (4.2.1.2)$$

I risultati del calcolo delle elasticità nel punto pari al valore medio degli input sono riportati in Tabella 4.2.1.1. Un incremento del capitale dell'1% determina un aumento dell'output dello 0.41%, mentre per il fattore lavoro l'aumento dell'output è circa la metà. I rendimenti di scala stimati in questo modello, derivanti dalla somma delle elasticità degli input, sono decrescenti (il valore è significativamente inferiore a 1).

Tabella. 4.2.1.1 – Elasticità media del prodotto rispetto agli input e rendimenti di scala

	<b>Elasticità</b>
<b>Capitale</b>	0,4080
<b>Lavoro</b>	0,2126
<b>Rendimenti di scala</b>	0,6206

Fonte: elaborazioni su dati L.488 e AIDA

### 4.3 *Metodologia non parametrica per il calcolo dell'effetto dei sussidi sulla TFP: tecniche di Matching*

Fino ad ora abbiamo argomentato su come si è arrivati ad ottenere una misura di TFP. Il lavoro che si è condotto ha visto l'implementazione di una metodologia non parametrica e una parametrica per il calcolo dell'effetto dei sussidi sulla TFP. Si vuole in una qualche maniera andare a studiare l'effetto degli incentivi sulla TFP attraverso l'implementazione di differenti metodologie, descritte nel precedente capitolo, attraverso un confronto tra le stesse tecniche per individuare se si ha avuto qualche effetto sulla misura di TFP.

Il *matching* è una tecnica per stimare l'ATT in presenza di un elevato numero di osservazioni e variabili relative a unità trattate e non trattate. Può essere applicata quando non è possibile un esperimento (assegnazione casuale) e quando non esiste la possibilità di fare un esperimento naturale (ovvero non esiste una variabile-strumento per cui, data la sua distribuzione, le unità vengono trattate e non trattate in modo casuale). Lo scopo del *matching* è di confrontare (*matching*) unità trattate e non trattate il più simili rispetto alle loro caratteristiche osservabili. In questo modo è possibile comparare direttamente trattati e non trattati per ottenere l'effetto del trattamento, in quanto l'unica differenza che rimane tra i due gruppi dovrebbe essere quella relativa alla partecipazione al programma, come nel caso di assegnazione casuale<sup>4</sup>. Nell'applicazione delle metodologie si è calcolato il *pscore*<sup>5</sup> (Tabella 4.3.1). Il *pscore* ha soddisfatto la proprietà di bilanciamento rispetto alle variabili introdotte. Successivamente lo si è usato per determinare la  $ATT_K$ , la  $ATT_{ND}$  e la  $ATT_S$ <sup>6</sup>.

---

<sup>4</sup> Bisogna però ricordare che tutte queste tecniche, non dicono nulla rispetto alla “scatola nera” della politica. Ovvero non spiegano perché il programma ha o non ha funzionato. Infine non da una risposta “risolutiva” una volta per tutte. Gli effetti dipendono inevitabilmente dalle condizioni di contesto, dalle altre politiche, dal ciclo economico ecc.

<sup>5</sup> Il calcolo del *pscore* ha visto l'individuazione di variabili dimensionali (*dum\_g* e interazioni *dum\_m1 dum\_g1 dum\_g2*), variabili di localizzazione (*area1, area2, area3*), variabili ateco, variabili riguardanti i bandi. Il calcolo del *pscore* è fornito dall'utilizzazione delle procedure scritte da Andrea Ichino (<http://www2.dse.unibo.it/ichino>) con il software Stata 10.

<sup>6</sup>  $ATT_{ND}$ : ATT estimation with the Stratification method;  $ATT_K$ : ATT estimation with the Kernel Matching method;  $ATT_S$ : ATT estimation with Nearest Neighbor Matching method

Tabella 4.3.1. - *Stima del propensity score*

```

*****
Algorithm to estimate the propensity score
*****

```

The treatment is agev

agev	Freq.	Percent	Cum.
0	1,058	63.70	63.70
1	603	36.30	100.00
Total	1,661	100.00	

Estimated propensity score

Percentiles	Smallest	Largest	Obs	Sum of Wgt.
1%	.0172534	.0159264		1455
5%	.02542	.0160811		1455
10%	.0420207	.0161716		
25%	.1096944	.0162663		
50%	.3207718		Mean	.4134829
75%	.7037131		Std. Dev.	.3294501
90%	.940599	.9992626		
95%	.9748759	.9995139	Variance	.1085374
99%	.997379	.9995325	Skewness	.4506497
		.9999707	Kurtosis	1.757627

Fonte: elaborazioni su dati L.488 e AIDA

Riassumendo per ciascuna metodologia le informazioni si è ottenuto il seguente risultato (Tabella 4.3.2). In totale le aziende agevolate sono 403, mentre quelle non agevolate sono pari a 498.

Tabella 4.3.2 – Stima ATT Matching<sup>7</sup>

Metodologia	AGEV	NON AGEV	ATT	STD ERROR	T-TEST <sup>8</sup>
ATT <sub>ND</sub>	403	498	-0.027	0.033	-0.832*
ATT <sub>K</sub>	403	498	-0.019	0.019	-0.956*
ATT <sub>S</sub>	403	498	-0.015	0.020	-0.747*

Fonte: elaborazioni su dati L.488 e AIDA

<sup>7</sup> Bootstrap degli standard errors a 100 reps.

<sup>8</sup> Il t-test è calcolato con un  $\alpha=0.01$

Si osserva come il numero delle agevolate per ciascuna delle metodologie proposte sia uguale e come il valore del t-test è non significativo in tutte. Il coefficiente è in tutti e tre i casi risulta essere negativo rispetto alle agevolate. Si può dire che l'incentivo dato alle aziende non ha avuto l'effetto sperato di un aumento di produttività, ma al contrario l'effetto è statisticamente nullo.

#### ***4.4 Metodologia parametrica per il calcolo dell'effetto dei sussidi sulla TFP: Diff-In-Diffs***

Ponendo a verifica un modello parametrico Diff-In-Diffs nel nostro dataset panel, si osserva che il numero di osservazioni sono 1.478. La strategia empirica per valutare il ruolo avuto dai sussidi nell'attività di investimento dovrebbe confrontare gli investimenti delle imprese agevolate con quelli che le stesse imprese avrebbero effettuato se non avessero ricevuto i sussidi. Poiché non si hanno informazioni sull'attività di investimento delle imprese agevolate nell'ipotesi in cui esse non avessero ricevuto i finanziamenti, si tratta di una valutazione controfattuale. Si è quindi deciso di avvalersi di un meccanismo attraverso il quale si è posto un confronto tra le imprese agevolate (*treated*) con il gruppo di imprese che hanno richiesto i finanziamenti, ma che non hanno ottenuto agevolazioni perché si sono collocate in basso alla graduatoria (*non-treated*). Le differenze sistematiche nei livelli non sono il principale problema perché possono essere controllate usando la metodologia *diff-in-diffs*.

La logica del metodo *diff-in-diffs* è quella di stimare l'effetto di una misura di incentivo attraverso un duplice confronto (differenze nelle differenze). Innanzitutto, per ciascun gruppo di imprese, rispettivamente trattate e non trattate si misurano le differenze della variabile oggetto di studio prima e dopo il trattamento; successivamente si calcola la differenza di queste variazioni tra i due gruppi di imprese localizzate nella stessa area evidenziando, attraverso una regressione, il valore della differenza media dopo il trattamento dei non trattati e per i trattati. Nel caso più semplice con un'unica osservazione per-intervento e

un'unica osservazione post-intervento, l'effetto di trattamento potrà essere misurato con una stima del tipo:  $\alpha_{dd}=(y^*_{j,1}- y^*_{j,0})- (y^*_{k,1}- y^*_{k,0})$ ; dove  $y^*_j$  e  $y^*_k$  sono le medie della variabile oggetto di osservazione per il gruppo delle imprese trattate  $j$  e non trattate  $k$ ; 0 e 1 rappresentano il periodo di tempo prima e dopo il trattamento, rispettivamente.

Il risultato che viene fuori è quello descritto dalla tabella 4.4.1. L'equazione che si è posta a stima assume la forma seguente:

$$TFP_{it} = cons + \alpha_1 dummytrat + \alpha_2 dummydopo + \alpha_3 ddopotrat + \varepsilon_{it}$$

Si tratta quindi di una versione a periodi dello stimatore Diff-In-Diffs. La  $TFP_{it}$  è la variabile oggetto di interesse, pari a una misura produttività per l'impresa  $i$ -esima nell'anno  $t$ ,  $dummytrat$  misura la differenza dalla media prima del trattamento dei non trattati e dei trattati,  $dummydopotrat$  misura la differenza dalla media prima del trattamento e dopo il trattamento per i non trattati ed infine  $dummydopotrat$  indica la differenza dalla media dopo il trattamento dei non trattati e dei trattati. Il coefficiente di nostro interesse è  $\alpha_3$ , cioè misura l'impatto della L.488 sulle imprese trattate.

Tabella 4.4.1 – *Differenza media dopo il trattamento dei non trattati e per i trattati*

Source	SS	df	MS	Number of obs = 1478		
Model	10.3042993	3	3.43476644	F( 3, 1474) =	9.62	
Residual	526.379873	1474	.357109819	Prob > F =	0.0000	
Total	536.684172	1477	.363360983	R-squared =	0.0192	
				Adj R-squared =	0.0172	
				Root MSE =	.59759	

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
dummytrat	27.93179	66.06264	0.42	0.672	-101.655	157.5186
dummydopo	-.0765296	.0226352	-3.38	0.001	-.1209303	-.0321289
ddopotrat	-.0139458	.0330159	-0.42	0.673	-.0787088	.0508173
_cons	153.1086	45.29602	3.38	0.001	64.25712	241.9602

Fonte: elaborazioni su dati L.488 e AIDA

Il coefficiente  $\alpha_3 = -0.0139$  è non significativo e negativo; l'impatto e quindi l'effetto dell'incentivo sulla TFP è negativo. Anche in questo caso non si hanno avuti effetti sulla produttività totale dei fattori.

# Conclusioni

Lo scopo di questa tesi è stato di valutare se gli incentivi erogati alle imprese tramite la L. 488/92 avessero effetti sulla TFP (Total Factor Productivity) delle imprese. Per questo sono state adeguate metodologie parametriche (Diff-In-Diffs) e non parametriche (tecniche di Matching basate sul propensity score).

La Legge 488/92 ha consentito alle imprese di ricevere un sussidio pubblico che copre una frazione della spesa per investimenti. Gli incentivi sono assegnati attraverso bandi di gara secondo specifici criteri predeterminati, come la proporzione di fondi propri investiti nel progetto; il numero di occupati coinvolti e la proporzione di assistenza richiesta.

L'analisi è stata effettuata utilizzando un campione di aziende costruito a partire dai dati della L. 488/92 e della banca dati AIDA che contiene i bilanci uniformati per le aziende. Si è costituito un dataset contenente un totale di 1.661 aziende per il periodo 1995-2004, di cui 603 imprese finanziate e 1.058 imprese non finanziate.

Il lavoro ha avuto come obiettivo principale il calcolo del valore della TFP come misura dell'efficienza ottenuta attraverso l'integrazione e l'applicazione di modelli econometrici che hanno permesso di ottenere informazioni non solo sul livello di TFP delle imprese, ma anche sulla funzione di produzione attraverso la

stima delle elasticità dei fattori produttivi. La specificazione del modello ha visto l'adozione di una funzione di produzione trans-log, tra diverse specificazioni, per la sua flessibilità, anche se l'elevato numero di parametri da stimare e la non immediata interpretazione degli stessi pone la necessità di testarne l'adozione, per esempio, con una specificazione più semplice quale la Cobb-Douglas che rispetto al modello trans-log presenta una chiave interpretativa molto più semplice essendo le elasticità pari ai coefficienti del capitale e del lavoro.

Per determinare quale funzione di produzione adottare è stato effettuato il test LR che risultando non significativo per una specificazione di tipo Cobb-Douglas, mentre significativo per una specifica di tipo trans-log completa, si è deciso di adottare un modello di tipo trans-log. Conseguente alla scelta della funzione di produzione, la TFP è stata calcolata stimando un modello ad effetti fissi per dati panel.

I risultati del calcolo delle elasticità nel punto pari al valore medio degli input hanno visto un incremento del capitale dell'1% che ha determinato un aumento dell'output dello 0.41%, mentre per il fattore lavoro l'aumento dell'output è stato di circa la metà. I rendimenti di scala stimati in questo modello, derivanti dalla somma delle elasticità degli input, sono risultati essere decrescenti

Successivamente il lavoro che si è condotto ha visto l'implementazione di una metodologia non parametrica e una parametrica per il calcolo dell'effetto dei sussidi sulla TFP. Si è voluto andare a studiare l'effetto degli incentivi sulla TFP attraverso l'implementazione di differenti metodologie attraverso un confronto tra le stesse tecniche per individuare se si ha avuto qualche effetto sulla misura di TFP.

Nell'applicazione delle metodologie si è calcolato il *p*score che ha soddisfatto la proprietà di bilanciamento rispetto alle variabili introdotte. Successivamente lo si è usato per determinare la ATT (Effect of treatment on the treated) rispettivamente per la metodologia ATT<sub>S</sub>: ATT stimata con metodo delle stratificazione; ATT<sub>K</sub>: ATT stimata con metodo del Kernel Matching; ATT<sub>S</sub>: ATT stimata con metodo delle Nearest Neighbor Matching. In totale le aziende agevolate sono risultate 403, mentre quelle non agevolate sono risultate pari a 498.

## CONCLUSIONI

Si è osservato come il numero delle agevolate per ciascuna delle metodologie proposte sia uguale e come il valore del t-test è risultato non significativo in tutte. Il coefficiente è in tutti e tre i casi è risultato essere negativo rispetto alle agevolate. Si può dire che l'incentivo dato alle aziende non ha avuto l'effetto sperato di un aumento di produttività, ma al contrario l'effetto è statisticamente nullo.

Ponendo a verifica un modello parametrico Diff-In-Diffs nel nostro dataset panel, si è osservato che il numero di osservazioni si sono ridotte, cioè pari a 1.478 aziende. La logica del metodo *diff-in-diffs* è quella di stimare l'effetto di una misura di incentivo attraverso un duplice confronto (differenze nelle differenze). Innanzitutto, per ciascun gruppo di imprese, rispettivamente trattate e non trattate si sono misurate le differenze della variabile oggetto di studio prima e dopo il trattamento; successivamente si è calcolata la differenza di queste variazioni tra i due gruppi di imprese localizzate nella stessa area evidenziando, attraverso una regressione, il valore della differenza media dopo il trattamento dei non trattati e per i trattati. Il modello ha considerato la  $TFP_{it}$ , che è la variabile oggetto di interesse, pari a una misura di produttività per l'impresa *i-esima* nell'anno *t*, *dummytrat* la differenza dalla media prima del trattamento dei non trattati e dei trattati, *dummydopotrat* la differenza dalla media prima del trattamento e dopo il trattamento per i non trattati ed infine *dummydopotrat* la differenza dalla media dopo il trattamento dei non trattati e dei trattati. Il coefficiente di interesse era proprio di quest'ultima variabile che misura l'impatto della L.488 sulle imprese trattate. Il coefficiente è risultato non significativo e negativo; l'impatto, e quindi l'effetto dell'incentivo sulla TFP è risultato negativo. Anche in questo caso non si hanno avuti effetti sulla produttività totale dei fattori.

I risultati, sostanzialmente, evidenziano una convergenza delle risposte date alla domanda: l'incentivo ha avuto effetto nella misura di TFP? Utilizzando le diverse metodologie si è risposto a questa domanda: l'impatto della L. 488 sulle aziende sovvenzionante è negativo e statisticamente non significativo.

Questo effetto nullo dell'incentivo non sorprende. Infatti anche in altri lavori si è osservato come gli incentivi alle imprese e la valutazione su una misura di TFP non ha avuto effetto. Ad esempio Lee (1996) nella sua analisi ritiene che il

## CONCLUSIONI

*government targeting* dei settori della Korea del Sud ha avuto bassa produttività; Bergstrom (1998) ha rilevato che l'impatto di crescita dei *capital subsidies* in Svezia senza però aver influenzato la TFP. Un limite del lavoro è che il risultato è contrario a quanto aspettato. Infatti sarebbe ragionevole presumere che le imprese agevolate utilizzassero impianti più nuovi e quindi più produttivi. È evidente che questo effetto è ridotto e si confronta con un recupero di efficienza necessario da parte delle imprese non agevolate.

Il confronto tra il trattati e non trattati sarebbe più omogeneo rispetto al livello di trattamento. Tuttavia, i metodi di matching utilizzati non sono in grado di spiegare l'eterogeneità. In effetti, le imprese con diverso livello di trattamento potrebbero avere caratteristiche dissimili, e discrepanze a livello di trattamento potrebbero essere imputate a questa eterogeneità. La valutazione dell'incentivo è stato però valutato solo sulla misura di TFP, ma ulteriori analisi lasciano il margine di poter osservare cosa con queste metodologie potrebbe accadere se andassimo a valutare come l'impatto di tale incentivo ha avuto effetto anche rispetto all'efficienza delle stesse aziende.

# Bibliografia

- Abadie, A. and Imbens, G.(2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235-267
- Abramovitz, M., 1956. Resource and output trends in United States since 1870. *American Economic Review* **46**, pp. 5–23.
- Aigner, D., Lovell, C.A.K., and Schmidt, P. (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, 6, pp.21-37.
- Athey, S. and Imbens, G.(2006). Identification and inference in nonlinear difference-in-difference models. *Econometrica*, vol. 74 (n.2)
- Ahn, S.C. e P. Schmidt, (1995). Efficient Estimation of Models for Dynamic Panel Data. *Journal of Econometrics*, 68, 5-28.

- Anderson, T.W. e C. Hsiao (1981). Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics*, 18, 570-606.
- Arellano, M. (2003). *Panel Data Econometrics*. Oxford: Oxford University Press
- Arellano, M. e S. Bond, (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277-297.
- Arellano, M. e O. Bover, (1995). Another look at the instrumental variable estimation of error components models. *Journal of Econometrics*, 68, 29-51.
- Arellano, M. e B. Honorè, (2001). Panel Data Models: Some Recent Developments. In *Handbook of Econometrics* vol. 5 a cura di J.J. Heckman e E. Leamer. Amsterdam: Elsevier.
- Bagella, M., Becchetti, L., (1998). The optimal financing strategy of high-tech firm: the role of warrants, *Journal of Economic Behaviour and Organisation*, 35, 1-23.
- Baier, S. L., Dwyer, G. P. e R. Tamura, (2002), "How important are capital and total factor productivity for economic growth?", *Federal Reserve Bank of Atlanta Working Paper*, n. 2a.
- Baltagi, B. (2005). *Econometric Analysis of Panel Data*. New York: Wiley & Sons.
- Barca, F., and G. Pellegrini (2000). Politiche per la competitività territoriale in Europa: note sul programma 2000-2006 per il Mezzogiorno d'Italia. Paper presented at the First Annual Workshop of University Bologna Center Buenos Aires, Buenos Aires, 26/27 April.

## BIBLIOGRAFIA

- Barnow B., Cain G., Goldberger A., (1981); Issues in the analysis of selectivity bias, in *Evaluation Studies*, Annual 5, Sage, Beverly Hills
- Barro, R. J., Sala-i-Martin, X.,(1998) *Economic Growth*, New York, McGraw Hill.
- Barro, R. J., (1999), “Notes on growth accounting”, *Journal of Economic Growth*, n. 4, pp. 119-137.
- Battese G. E. and Coelli T. J. (1988), *Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data*, *Journal of Econometrics*, n. 38.
- Battese .G.E., Coelli T. (1992) *Frontier production function and technical efficiency: A survey of empirical applications in agricultural economics*. *Agricultural Economics* 7:185-208
- Battese .G.E., Coelli T. (1993) *A stochastic frontier production function incorporating a model for technical inefficiency effects* Department of Econometrics , University of New England Working Paper n.69
- Battese G., Coelli T. (1995), *A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data*, *Empirical Economics*, n. 20, p. 325-332.
- Battese G., Broca S. (1996), *Functional Forms of Stochastic Frontier Production Functions and Model for Technical Inefficiency Effects: A comparative Study for Wheat Farmers in Pakistan*, *Journal of Productivity Analysis* 8:395-414.
- Basu S.,(1995). «Procyclical Productivity: Increasing Returns or Cyclical Utilization? », *NBER Working Paper*, n. 5336

- Baumol, W J, (1990), Entrepreneurship: Productive, unproductive, and destructive. *Journal of Political Economy*, Vol. 98, No. 5, pp. 893-921.
- Beason, R & Weinstein, D E, (1996), "Growth, economies of scale, and targeting in Japan (1955-1990)". *The Review of Economics and Statistics*, Vol. 78, No. 2, May, pp. 286-295.
- Benfratello L. (2002) *Beyond profitability: effects of acquisitions on technical efficiency in the Italian pasta industry*, Working Paper CERIS-CNR
- Bergstrom, F.(1998), "Capital subsidies and the performance of firms", *Small Business Economics*, Vol.8, pp.175-80.
- Bernanke B.S. - Parkinson M.L.,(1991). «Procyclical Labor Productivity and Competing Theories of the Business Cycle: Some Evidence from Interwar US Manufacturing Industries», *Journal of Political Economy*, n. 7, pp. 340-59.
- Bernini, C., Centra, M., Pellegrini, G. (2006). Growth and efficiency in subsidized firm. *Mimeo*.
- Blomström, M, Lipsey, R E, & Zejan, M, (1996), "Is fixed investment the key to economic growth?". *The Quarterly Journal of Economics*, Vol. 111, February, pp. 267-276.
- Boggio, L. e G. Seravalli, (2002), *Lo sviluppo economico. Fatti, teorie, politiche*, Bologna: Il Mulino.
- Bondonio D. (2000). "Statistical methods to evaluate geographically-targeted economic development programs", *Statistica Applicata*, 12 (2), pp. 177-204.

## BIBLIOGRAFIA

- Bondonio D. (2002). "Evaluating decentralized policies: A method to compare the performance of economic development programmes across different regions or states", *Evaluation*, 8(1), pp. 101-124.
- Bondonio D. and J. Engberg (2000). "Enterprise Zones and local employment: evidence from the states' programs", *Regional Science and Urban Economics*, 30(5), pp. 519-549.
- Bottasso A, Sembenelli A. (2004) *Does ownership affect firms' efficiency? Panel data evidence on Italy* *Empirical Economics* 29:769-786.
- Burnside C., «Production Function Regressions, Returns to Scale and Externalities », *Journal of Monetary Economics*, apr. 1996, pp. 177-201.
- Burnside C. - Eichenbaum M.,(1994). «Factor Hoarding and the Propagation of Business Cycle Shocks», *NBER Working Papers*, n. 4675.
- Blundell, R. e S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, 87, 115-143.
- Bosworth, B. P. e S. M. Collins, (2003), "The empirics of growth: an update", Washington DC: Brookings Institution.
- Bowsher, C. G. (2002). On testing overidentifying restrictions in dynamic panel data models. *Economics Letters* 77, 211-220.
- Bronfenbrenner, M. and Paul H. Douglas (1939), "Cross-Section Studies in the Cobb-Douglas Function," *Journal of Political Economy*, 47:6 (December), 761-85.
- Bronzini, R., de Blasio, G. (2006), Qual'è l'effetto degli incentivi alle imprese? Una valutazione della 488/92, Temi di discussione Banca d'Italia, n482

- Bruno, G.S.F. (2005a). Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*, 87, 361-366.
- Bruno, G.S.F. (2005b.) Estimation and inference in dynamic unbalanced panel data models with a small number of individuals. CESPRI WP n.165 , Università Bocconi-CEPRI, Milan.
- Bruno, G.S.F, A.M. Falzoni e R. Helg, (2004). Masuring the effect of globalization on labour demand elasticity: An empirical application to OECD countries. CESPRI WP n.153 , Università Bocconi-CEPRI, Milan.
- Bun, M.J.G., Kiviet, J.F., (2003). On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters*, 79, 145-152.
- Caballero R.J. - Lyons R.K.,(1992) «External Effects in the US Procyclical Productivity», *Journal of Monetary Economics*, n. 29, pp. 209-26.
- Cannari, L., D'Aurizio, L., de Blasio, G., “The effectiveness of investment subsidies: evidence from survey data”, *Occasional Papers Banca d'Italia*, 4, 2006.
- Carlucci, C., Pellegrini, G. (2003), Gli effetti della legge 488/92: una valutazione dell'impatto occupazionale sulle imprese agevolate, *Rivista Italiana degli Economisti*, 2003/2, pp. 267-286
- Carlucci, C., Pellegrini, G. (2005), Nonparametric analysis of the effects on employment of public subsidies to capital accumulation: the case of law 488/92 in Italy, mimeo
- Cas, A. & Rymes, T. K. (1991) *On Concepts and Measures of Multifactor Productivity in Canada, 1961-1980*, Cambridge, Cambridge University Press;

## BIBLIOGRAFIA

- Caselli, F., (2003), "The missing input: accounting for cross-country income differences", bozza per il Capitolo 8 di P. Aghion e S. Durlauf (a cura di), *Handbook of Economic Growth*, North Holland, Amsterdam, di prossima pubblicazione.
- Caves, R E, & Barton, D R, (1990), *Efficiency in U.S. manufacturing industries*. The MIT Press, Cambridge, Massachusetts.
- Chanda, A. e C. J. Dalgaard, (2003), "Dual economies and international total factor productivity differences", mimeo.
- Charnes A., Cooper W.W., Rodes E.,(1978) "*Measuring the efficiency of decision making units*", *European Journal of Operational Research* 2, pp. 429-444.
- Chiri, S. and Pellegrini, G. (1995). Gli aiuti alle imprese nelle aree depresse. *Rivista economica del Mezzogiorno*, n.3
- Cobb, C. W. and Douglas, P. H. (1928). "A Theory of Production", *American Economic Review*, 18 (Supplement), 139-165.
- Daly, M., Gorman, I., Lenjosek, G., MacNevin, A., and Phiriyapreunt, W. (1993). The impact of regional investment incentives on employment and productivity. *Regional Science and Urban Economics*, (23), pp. 559-575.
- Debreu, G. (1951). "*The Coefficient of Resource utilization*", *Econometria*, vol. 19, n.3, pp.273-292.
- Dehejia, R. (2005). Practical propensity score matching: a reply to smith and todd. *Journal of Econometrics*, vol. 125 (issues 1-2), pp. 355-364

- Dehejia, R. and Wahba, S. (1998b). Propensity score matching methods for non-experimental causal studies. Technical report.
- De Long, J.B. and Summers, L.H., (1991). "Equipment investment and economic growth", *The Quarterly Journal of Economics*, Vol. 106, pp. 445-502.
- Denison, E., (1962). United States economic growth. *Journal of Business*, pp. 109–121
- Deprins D., Simar L., Tulkens H., Measuring Labor-efficiency in Post Offices, Elsevier Science Pb. Marchant, Amsterdam, North Holland, 1984
- Durand, R. (1996) Canadian input-output-based multi-factor productivity accounts, *Economic Systems Research*, 8(4), pp. 367–389
- Eichenbaum M., «Real Business Cycle Theory: Wisdom or Whimsy», *Journal of Economic Dynamics and Control*, ott. 1991, pp. 607-26.
- Easterly, W. e R. Levine, (2001), "What have we learned from a decade of empirical research on growth? It's not factor accumulation: stylized facts and growth models", *World Bank Economic Review*, 15, pp. 177-219.
- Evans C.L.(1992), «Productivity Shocks and Real Business Cycles», *Journal of Monetary Economics*, n. 29, pp. 191-208.
- Faini, R., Schiantarelli, F., (1987). "Incentives and investment decisions: The effectiveness of regional policy", *Oxford Economic Paper*, 39, pp. 516-533.
- Farrell, M.J. (1957). The Measurement of Productive Efficiency of Production, *Journal of the Royal Statistical society. Series A, General*, vol. 120, pp.253-281.

## BIBLIOGRAFIA

- Finn M.G.,(1995). «Variance Properties of Solow's Productivity Residual and Their Cyclical Implications», *Journal of Economic Dynamics and Control*, n. 19, pp. 1249-81.
- Fujita,M.and Krugman, P. (1995) When is the economy monocentric?: von Thuenen and Chamberlin unified. *Regional Science and Urban Economics* 25, pp. 505-28.
- Fama, E. e J. MacBeth, (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 81, 607-636.
- Gazzei, D. S., Lemmi A., Viviani A.(1997). Misure statistiche di performance produttiva. Un percorso di metodi e di evidenze empiriche. CLEUP
- Gabe, T.M., Kraybill, D.S. (2002). "The effects of state economic development incentives on employment growth of establishment", *Journal of Regional Science*, 42, pp. 703, 730.
- Hahn, J. (1999). "How informative is the initial condition in the dynamic panel data model with fixed effects?" *Journal of Econometrics*, 93, 309-326.
- Hall R.E., «The Relation Between Price and Marginal Cost in US Industry», *Journal of Political Economy*, ott. 1988, pp. 921-47.
- Harris R. and Trainor, M.(2005). "Innovations and R & D in Northern Ireland Manufacturing: A Schumpeterian Approach," *Regional Studies*, Taylor and Francis Journals, vol. 29(7), pages 593-604, April.
- Hatry H.P., (1982). Misurazione della Performance: Principi e Tecniche, in *Problemi di Amministrazione Pubblica*, Quaderni FORMEZ.

- Heckman, J.J., Ichimura, H., Todd, P.(1998) "Matching as an econometric evaluation estimator", *Review of Economic Studies*, 65, pp. 261-294.
- Heckman, J.J., LaLonde, R., Smith, J.(1999). *The economics and econometrics of active labor market programs*, in Ashenfelter, O., Card, D (eds.) *Handbook of Labor Economics*, vol. 3A, Amsterdam, North-Holland, pp. 1865-1097.
- Hirano, K, Imbens, G.W., Ridder, G. (2003). "Efficient estimation of average treatment effects using the estimated propensity score", *Econometrica*, 71, pp. 1161-1189.
- Holland, P.,(1986) "Statistics and causal inference", *Journal of the American Statistical Association*, 81, pp. 945-970.
- Hulten, C. H., (2000), "Total factor productivity: a short biography", *NBER WorkingPaper*, n. 7471.
- Ichino, A. and Becker, S. (2002). "Estimation of Average Treatment Effects Based on Propensity Scores" (with Sascha Becker), *The Stata Journal* , 2 (4), 358-377.
- Ichino, A. (2002). *The Problem of Causality in the Analysis of Educational Choices and Labor Market Outcomes*, Lectures Notes.
- Istat, *Classificazione delle attività economiche*, Rome, 2002.
- Jondrow J., Lovell C., Materov L, Schmidt P. (1982), *On Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model*, *Journal of Econometrics*, nr. 19.
- Jorgenson, Dale W. and Kevin J. Stiroh (2000), "Raising the speed limit: U.S. Economic Growth in the Information Age", *Brookings Papers on Economic Activity*, 1, 125-211.

## BIBLIOGRAFIA

- Kiviet, J.F., (1995). On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models. *Journal of Econometrics*, 68, 53-78.
- Kiviet, J.F., (1999). Expectation of Expansions for Estimators in a Dynamic Panel Data Model; Some Results for Weakly Exogenous Regressors. In: Hsiao, C., Lahiri, K., Lee, L.-F., Pesaran, M.H. (Eds.), *Analysis of Panel Data and Limited Dependent Variables*. Cambridge University Press, Cambridge.
- Kiviet, J.F. e M.J.G. Bun, (2001). The Accuracy of Inference in Small Samples of Dynamic Panel Data Models. Tinbergen Institute Discussion Paper TI 2001-006/4.
- Kendrick, (1961), *Productivity trends in the United States*, New York: Princeton University Press for NBER.
- Kydland F.E. - Prescott E.C.,(1982) «Time to Build and Aggregate Fluctuations», *Econometrica*, vol. 50, pp. 1345-70.
- Klenow, P. J. and Rodriguez-Clare, A., (1997), “The neoclassical revival in growth economics: has it gone too far?”, *NBER Macroeconomics Annual*, 12, pp. 73-102.
- Koopmans, T.C. (1951) *Analysis of Production as an Efficient Combinations of Activities*, trad. italiano in Koopmans, T.C. (1987)
- Koopmans, T.C. (1987) *Econometria, analisi delle attività, crescita ottimale*, il Mulino, Bologna
- Kuznets, S. (1990) *Popolazione, tecnologia, sviluppo*, Il Mulino, Bologna.
- Lee, J.W.,(1995) “Government intervention and productivity growth”, *Journal of Economic Growth*, 1, pp. 391-414.

- Mankiw, Romer e Weil (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, 107, pp. 407-438.
- Marschack, J. and Andrews, W.H. (1944) "Random Simultaneous Equations and the Theory of Production" *Econometrica* **12**(3,4), 143-205.
- McLeod, A.I., K.W. Hipel (1978). Smulation Procedures for Box-Jenkins Models. *Water Resources Research*, 14, 969-975.
- Meeusen, W. and Van den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* **18** 2, pp. 435-444
- Mohr, L.B., (1988) *Impact analysis for program evaluation*, Pacific Grove (CA), Brooks/Cole Publishing
- Newey, W. e K. West, (1987). A simple positive-definite, heteroskedastic and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- Nickell, S.J., (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49, 1417-1426.
- Nisticò, A. e Prosperetti, L. (1991) Produzione e Produttività, in *Statistica Economica*, a cura di Marbach, G., Utet, Torino
- Pellegrini, G., Bernini, C. , Adorno V. (2007), The impact of capital subsidies: new estimations under continuous treatment , in via di pubblicazione presso il *Giornale degli Economisti*
- Petersen, M. A. (2005). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Working Paper NBER n. 11280.

## BIBLIOGRAFIA

- Prescott E.C., (1986) «Theory Ahead of Business Cycle Measurement», in BRUNNER K. - MELTZER A.H. (a cura di), *Real Business Cycles, Real Exchange Rates and Actual Policies, Carnegie-Rochester Conference on Public Policy*, vol. 25, pp. 11-44.
- Reati, A. (2001), “Total factor productivity-a misleading concept”, *BNL Quarterly Review*, 218, pp. 313-332.
- Reder, M. W. (1943), An alternative interpretation of the Cobb-Douglas production function. *Econometrica* 11:259-64
- Read, L. M. (1968) The measure of total factor productivity appropriate to wage-price guidelines, *Canadian Journal of Economics*, 1(2), pp. 349–358;
- Rymes, T. K. (1971) *On Concepts of Capital and Technical Change*, Cambridge, Cambridge University Press
- Rymes, T. K. (1972) The measurement of total factor productivity in the context of the Cambridge theory of capital, *Review of Income and Wealth*, 18(1), pp. 79–108
- Rymes, T. K. (1983) More on the measurement of total factor productivity, *Review of Income and Wealth*, 29, pp. 297–316
- Rettore, E., Trivellato, U., Martini, A., (2003) “La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?”, *Politica Economica*, XIX (3), pp. 301-341.
- Rogers, W. (1993). Regression Standard Errors in Clustered Samples. *Stata Technical Bulletin*, 13, 19-23.

- Rosenbaum, P., Rubin, D., (1983) "The central role of the propensity score in observational studies for causal effects", *Biometrika*, 70, pp. 41-55.
- Rosenbaum, P., Rubin, D., (1984) "Reducing bias in observational studies using subclassification on the propensity score, *Journal of the American Statistical Association*, 79, pp. 516-524.
- Rubin, D. B., (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies", *Journal of Educational Psychology*, 66, pp. 688-701.
- Rubin, D.B, Thomas, N., (1996) "Matching using estimated propensity score, relating theory to practice, *Biometrics*, 52, pp. 249-264.
- Senhadji, A., (1999), "Sources of economic growth: an extensive growth accounting exercise", *International Monetary Fund Working Paper*, n. 99/77.
- Schmookler J.(1952), *The Changing Efficiency of the American Economy, 1869-1938*, *Review of Economics and Statistics*, 34 (3), pp. 214-231
- Scott L. Baier & Gerald P. Dwyer, Jr. & Robert Tamura,(2002). "How important are capital and total factor productivity for economic growth?," Working Paper 2002-2a, Federal Reserve Bank of Atlanta, revised
- Solow, R., (1956), "A contribution to the theory of economic growth", *Quarterly Journal of Economics*, Vol. 70, pp. 65-94.
- Solow, R. (1957): "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics*, **39**:312-320.
- Stiroh, K. J., (2001), "What drives productivity growth?", *Federal Reserve Bank of New York Economic Policy Review*, Marzo.

## BIBLIOGRAFIA

- Summers L., «Some Skeptical Observations on Real Business Cycle Theory», Federal Reserve Bank of Minneapolis, *Quarterly Review*, autunno 1986, pp. 237.
- Van Tongeren, F. W., (1998). "Microsimulation of Corporate Response to Investment Subsidies," *Journal of Policy Modeling*, Elsevier, vol. 20(1), pages 55-75, February.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient GMM twostep estimators. *Journal of Econometrics*, 126, 25-51.
- Winsten, C. B. (1957): "Discussion on Mr. Farrell's Paper." *Journal of the Royal Statistical Society, Series A*, 120 (III), 282-284.
- Woessmann, L. (2003). Specifying human capital. *Journal of Economic Surveys*, **17**, 239-270.
- Zappa, G. (1957), *Le produzioni nell'economia delle imprese*, Giuffrè, Milan, (Production in the Economics of the Firms), .



# Appendice A



Estimated propensity score

Percentiles		Smallest		
1%	.0172534	.0159264		
5%	.02542	.0160811		
10%	.0420207	.0161716	Obs	1455
25%	.1096944	.0162663	Sum of Wgt.	1455
50%	.3207718		Mean	.4134829
75%	.7037131	Largest	Std. Dev.	.3294501
90%	.940599	.9992626		
95%	.9748759	.9995139	Variance	.1085374
99%	.997379	.9995325	Skewness	.4506497
		.9999707	Kurtosis	1.757627

\*\*\*\*\*  
**Step 1: Identification of the optimal number of blocks**  
**Use option detail if you want more detailed output**  
 \*\*\*\*\*

The final number of blocks is 8

This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks

Seguendo l’algoritmo descritto nella sessione 2 di questa Appendice, i blocks per il quale la media del *propensity score* di trattati e del gruppo di controllo sono divisi a metà.

\*\*\*\*\*  
**Step 2: Test of balancing property of the propensity score**  
**Use option detail if you want more detailed output**  
 \*\*\*\*\*

The balancing property is satisfied

Nel caso in questione, la proprietà di bilanciamento è soddisfatta, la distribuzione finale dei trattati e del gruppo di controllo attorno ai blocchi sono così tabulati assieme all’ *Inferior* per ciascun blocco:

This table shows the inferior bound, the number of treated and the number of controls for each block

Inferior of block of pscore	agev		Total
	0	1	
.0159264	312	25	337
.1	171	34	205
.2	201	70	271
.4	103	78	181
.6	51	128	179
.8	11	28	39
.85	3	49	52
.9	0	191	191
Total	852	603	1,455

Note: the common support option has been selected

\*\*\*\*\*  
**End of the algorithm to estimate the pscore**  
 \*\*\*\*\*

Si impone l'uso della condizione di common support usando l'opzione *comsup*. Conseguentemente i blocchi identificano missing per le osservazioni del gruppo di controllo per il common support e il numero di osservazioni nella tabella sono 1.455 contro 1.893. Dopo l'uso del *pscore*, si può procedere alla stima dell'ATE (Average Treatment Effect) usando i metodi per il calcolo della *att\**.



### A.3 Output *attk*

*attk dv agev, pscore(pp1) cmsup epan detail boot reps(100) dots*

#### ATT estimation with the Kernel Matching method

n. treat.	n. contr.	ATT	Std. Err.	t
403	498	-0.019	.	.

Note: Analytical standard errors cannot be computed. Use the bootstrap option to get bootstrapped standard errors.

#### Bootstrapping of standard errors

```
command:      attk dv agev , pscore(pp1) cmsup epan bwidth(.06)
statistic:    attk      = r(attk)
```

```
Bootstrap statistics                                Number of obs = 1066
                                                    Replications = 100
```

Variable	Reps	Observed	Bias	Std. Err.	[95% Conf. Interval]
attk	100	-.0186121	.0026595	.0194691	-.0572431 .0200189 (N)
					-.0585092 .0166899 (P)
					-.0753931 .0077171 (BC)

Note: N = normal  
P = percentile  
BC = bias-corrected

#### ATT estimation with the Kernel Matching method Bootstrapped standard errors

n. treat.	n. contr.	ATT	Std. Err.	t
403	498	-0.019	0.019	-0.956

## A.4 Output *atts*

*atts dv agev, pscore(pp1) comsup blockid(dd1) detail boot reps(100) dots*

### The treatment is agev

agev	Freq.	Percent	Cum.
0	498	55.27	55.27
1	403	44.73	100.00
Total	901	100.00	

### The distribution of the pscore is

Estimated propensity score

Percentiles	Smallest		
1%	.0223378	.020373	
5%	.0303776	.0207126	
10%	.0523279	.020834	
25%	.1300283	.0209915	Obs 901
			Sum of Wgt. 901
50%	.3464759		Mean .4337372
		Largest	Std. Dev. .3275879
75%	.7234814	.9983438	
90%	.9470227	.9983438	Variance .1073138
95%	.976002	.9986768	Skewness .3821644
99%	.9968588	.9999707	Kurtosis 1.705715

### The structure of blocks is

Key
<i>frequency</i>
<i>column percentage</i>

Number of block	agev		Total
	0	1	
1	164 32.93	19 4.71	183 20.31
2	100 20.08	24 5.96	124 13.76
3	126 25.30	47 11.66	173 19.20
4	66 13.25	53 13.15	119 13.21
5	31 6.22	83 20.60	114 12.65
6	8 1.61	16 3.97	24 2.66
7	3 0.60	36 8.93	39 4.33
8	0 0.00	125 31.02	125 13.87
Total	498 100.00	403 100.00	901 100.00



# Appendice B

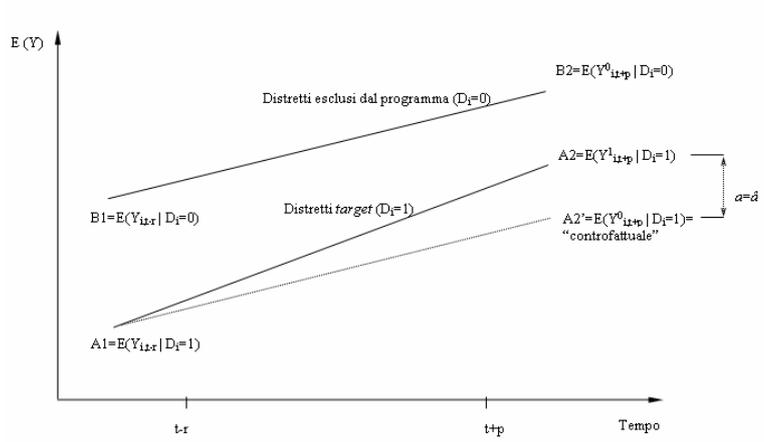


Figura B.1

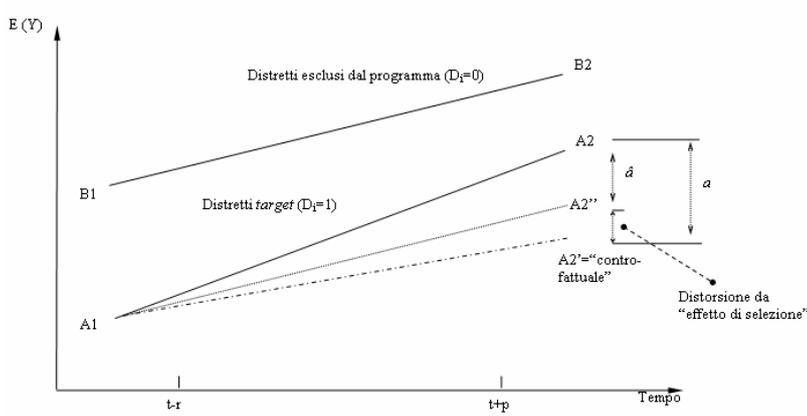


Figura B.2

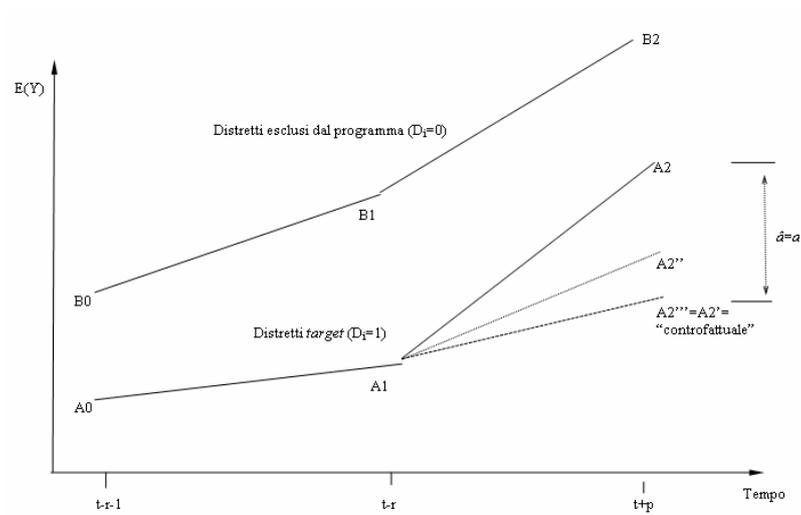


Figura B.3