

UNIVERSITA' DELLA CALABRIA  
Scuola di Dottorato in Scienze Economiche e Aziendali

Dipartimento di Economia e Statistica

TESI DI DOTTORATO

**L'IMPATTO DEL CONSOLIDAMENTO DEL SISTEMA BANCARIO  
ITALIANO SULL'EFFICIENZA DI COSTO: FRONTIERE  
STOCASTICHE, DIPENDENZA TRA GLI ERRORI E COPULE**

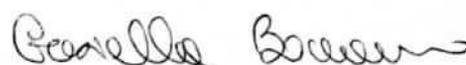
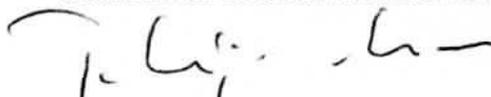
Supervisore: Prof. Giovanni D'ORIO

Tutor: Prof. Filippo DOMMA



Direttore: Prof.ssa Patrizia ORDINE

Dottoranda: Dott.ssa Graziella BONANNO



Settore disciplinare: SECS-P/06, Economia Applicata

XXV Ciclo / 2009-2012

## ***Ringraziamenti***

*Durante i tre anni della Scuola di Dottorato ho avuto il piacere di collaborare con diverse persone ognuna delle quali ha contribuito alla mia crescita umana e professionale.*

*Desidero ringraziare il mio Supervisore, il Professor Giovanni D'Orio, che mi ha regalato preziosi suggerimenti e direttive da seguire per il completamento della tesi.*

*Un grazie particolare va al mio Tutor, il Professor Filippo Domma, con il quale ho cominciato a studiare la Statistica fin dai primi corsi universitari e ad affrontare lo studio delle funzioni copule come strumento per la derivazione di nuove funzioni di distribuzioni.*

*Un altro ringraziamento è riservato al Professor Francesco Aiello, il quale ha contribuito alla nascita del tema di ricerca e ha seguito l'evoluzione della tesi dandomi importanti spunti di riflessione.*

*Grazie al Professor Domenico De Giovanni, la cui collaborazione è stata fondamentale per la stesura di una parte della tesi, e alla Professoressa Sabrina Giordano, con la quale ho avuto alcune informali e stimolanti discussioni.*

*Un grazie alla mia famiglia per il sostegno morale, economico ed affettivo.*

*Inoltre, la mia collega ed amica Alessia merita un ringraziamento particolare per essermi stata accanto nei momenti difficili e anche in quelli più belli.*

## Sintesi

A partire dagli anni '90, il sistema bancario italiano è stato interessato da un intenso processo di concentrazione spinto dall'idea di armonizzare il sistema seguendo le direttive europee. Uno degli obiettivi di questa ristrutturazione è quello di aumentare il livello di efficienza di un settore produttivo che contribuisce ad orientare lo sviluppo del paese. Se e come questo obiettivo sia stato conseguito è una questione ancora aperta. Infatti, dalla letteratura di riferimento, non si ottengono indicazioni chiare sul recupero di efficienza delle banche italiane.

L'obiettivo di questa tesi è di fornire un contributo al dibattito considerando un campione di circa 700 banche italiane osservate annualmente nel periodo 2006-2010. L'efficienza analizzata è quella di costo ottenuta dalla stima di modelli di frontiere stocastiche (SFA) nella specificazione econometrica proposta da Battese e Coelli (1995). Ciò consente di analizzare l'effetto sull'efficienza delle banche che è determinato da fattori di contesto e da variabili specifiche dei singoli istituti, quali, per esempio, la qualità del credito concesso ai mutuatari. Per quanto riguarda la funzione di costo, si fa riferimento all'approccio di intermediazione, poiché è quello che più si avvicina alla funzione di intermediazione finanziaria svolta dalle banche.

L'approccio SFA consente di suddividere la distanza dalla frontiera in una componente di inefficienza e in una componente casuale, presupponendo la formulazione delle ipotesi relative alle distribuzioni e alla dipendenza tra le due componenti. La prassi seguita finora in letteratura è quella di assumere che l'errore casuale si distribuisca secondo una distribuzione Normale e la componente di inefficienza secondo una distribuzione semi-Normale. Si assume, inoltre, che queste due distribuzioni siano indipendenti. Tenendo conto di queste ipotesi, l'analisi condotta nella prima parte della tesi indica una riduzione dell'efficienza del sistema bancario italiano fino al 2008 e una ripresa nel 2010. Il risultato medio nazionale è sempre trainato dalla performance delle Banche di Credito Cooperativo (BCC) e delle banche con sede legale nelle regioni del Nord-Est del paese. Riguardo i fattori che contribuiscono a spiegare l'efficienza bancaria, un risultato da sottolineare è quello che riguarda le sofferenze bancarie. L'evidenza empirica supporta l'ipotesi di *bad management*, in base alla quale esiste una relazione di simultaneità inversa tra efficienza di costo e qualità del credito.

Nella seconda parte della tesi si presentano alcuni aspetti metodologici ed empirici legati all'utilizzo della funzione copula nell'approccio delle frontiere stocastiche. Ciò al fine di creare nuove distribuzioni del termine di errore composto considerando marginali più flessibili rispetto a quelle Normali e introducendo dipendenza fra le due componenti d'errore. La funzione copula considerata è quella di Farlie-Gumbel-Morgenstern (FGM). In tale ambito, si dimostra, innanzitutto, che l'ipotesi di indipendenza non è sempre empiricamente supportata dai dati. Dall'altro lato, si conseguono due importanti risultati di natura metodologica. In primo luogo si dimostra che la dipendenza, misurata mediante l'approccio delle funzioni copule dal parametro  $\theta$ , influenza la varianza dell'errore composto, la quale viene utilizzata per la scomposizione dell'errore composto nella componente di inefficienza e nella componente casuale. In secondo luogo, l'introduzione nel modello delle frontiere stocastiche di funzioni marginali per le due componenti di errore più generali rispetto a quelle Normali rende possibile lo studio di casi in cui le caratteristiche distributive non sono riconducibili al caso di normalità. Si dimostra, infatti, che anche la misura di asimmetria e curtosi della funzione di distribuzione assegnata all'errore casuale ha un impatto sulla varianza dell'errore composto e, quindi, sulle stime dell'efficienza.

## INDICE

<b>Introduzione</b> .....	<b>pg 8</b>
---------------------------	-------------

### CAPITOLO I

<b>La ristrutturazione e l'efficienza del sistema bancario italiano</b> .....	<b>pg 11</b>
---	--------------

1.1 Introduzione.....	pg 11
-----------------------	-------

1.2 Cambiamenti istituzionali ed operativi del sistema bancario italiano .....	pg 12
--	-------

1.3 La ristrutturazione e la dualità del sistema bancario italiano .....	pg 17
--	-------

1.4 La misurazione dell'efficienza .....	pg 21
--	-------

1.5 Rassegna della letteratura empirica sullo studio dell'efficienza dei sistemi bancari ....	pg 24
---	-------

1.5.1 Gli studi basati su metodologie parametriche: analisi dei sistemi bancari internazionali .....	pg 25
--	-------

1.5.2 Gli studi basati su metodologie parametriche: analisi del sistema bancario italiano .....	pg 29
---	-------

1.5.3 Gli studi basati su metodologie non parametriche.....	pg 34
---	-------

1.5.4 Gli studi basati sul confronto fra metodi parametrici e non parametrici .....	pg 37
---	-------

1.6 Conclusioni.....	pg 41
----------------------	-------

### CAPITOLO II

<b>L'efficienza del sistema bancario italiano dal 2006 al 2010: un'applicazione delle frontiere stocastiche</b> .....	<b>pg 42</b>
---	--------------

2.1 Introduzione.....	pg 42
-----------------------	-------

2.2 Efficienza e funzioni di costo .....	pg 43
--	-------

2.3 Descrizione del campione.....	pg 49
-----------------------------------	-------

2.4 La metodologia e la scelta delle variabili .....	pg 53
--	-------

2.5 Le stime su dati <i>panel</i> dei livelli di efficienza di costo delle banche italiane dal 2006 al 2010 .....	pg 65
---	-------

2.5.1 Livelli medi di efficienza di costo delle banche italiane .....	pg 68
---	-------

2.6 Prove di robustezza.....	pg 72
2.7 Conclusioni.....	pg 74
Appendice A1. Il test di Granger.....	pg 76
Appendice A2. Le stime su dati <i>cross section</i> dei livelli di efficienza di costo delle banche italiane dal 2006 al 2010.....	pg 78

### **CAPITOLO III**

<b>La modellizzazione della dipendenza tra le componenti d'errore delle frontiere stocastiche mediante le funzioni copule e specificazioni alternative delle componenti marginali .....</b>	<b>pg 84</b>
---	--------------

3.1 Introduzione.....	pg 84
3.2 La copula bivariata .....	pg 87
3.2.1 Le copule archimedee .....	pg 90
3.3 Frontiere stocastiche e copule.....	pg 91
3.4 La funzione di distribuzione dell'errore composto mediante la copula FGM e marginali Normali.....	pg 93
3.4.1 Stime econometriche della frontiera di costo e dell'equazione dell'efficienza in caso di normalità e dipendenza fra le due componenti d'errore .....	pg 94
3.5 Un approccio metodologico per la costruzione della funzione di distribuzione dell'errore composto mediante la copula FGM ed una legge Esponenziale per la componente di inefficienza .....	pg 98
3.6 Un approccio metodologico per la costruzione della funzione di distribuzione dell'errore composto mediante la copula FGM e marginali Log-Dagum .....	pg 112
3.7 Conclusioni.....	pg 119

<b>Conclusioni .....</b>	<b>pg 120</b>
<b>Riferimenti bibliografici .....</b>	<b>pg 124</b>

## Indice delle tabelle

Tabella 1.1 L'evoluzione della struttura del sistema bancario italiano .....	pg 13
Tabella 1.2 Le operazioni di fusione, incorporazione e trasferimenti di controllo .....	pg 14
Tabella 1.3 Banche e sportelli operanti in Italia nel 1993, nel 2002 e nel 2008 .....	pg 18
Tabella 1.4 Evoluzione del rapporto Impieghi/Depositi nel Sud e nel Centro-Nord .....	pg 19
Tabella 1.5 Numero di banche classificate per area territoriale .....	pg 21
Tabella 1.6 Caratteristiche, vantaggi e limiti degli approcci parametrici e non parametrici utilizzati per misurare l'efficienza .....	pg 24
Tabella 1.7 Gli studi sull'efficienza dei sistemi bancari: un quadro sinottico .....	pg 39
Tabella 2.1 Composizione delle banche del campione per forma giuridica, classe dimensionale e area territoriale .....	pg 50
Tabella 2.2 Totale attività, impieghi, raccolta e costi delle banche italiane incluse nel campione (2006-2010) .....	pg 51
Tabella 2.3 Dipendenti e Sportelli delle banche incluse nel campione (2006-2010) .....	pg 52
Tabella 2.4 Definizione delle variabili incluse nella funzione di costo .....	pg 58
Tabella 2.5 Valori medi delle variabili di input ed output dal 2006 al 2010 .....	pg 60
Tabella 2.6 Definizione delle esplicative della componente di inefficienza .....	pg 62
Tabella 2.7 Evoluzione dei crediti in sofferenza dal 2006 al 2010 .....	pg 63
Tabella 2.8 L'indice di Herfindahl dal 2006 al 2010 .....	pg 65
Tabella 2.9. Stime della funzione di costo delle banche italiane dal 2006 al 2010 .....	pg 66
Tabella 2.10. Stime dell'equazione di inefficienza delle banche italiane dal 2006 al 2010 .....	pg 68
Tabella 2.11. Alcune statistiche descrittive relative ai livelli di efficienza di costo nel 2010 .....	pg 69
Tabella 2.12. Livelli medi di efficienza di costo delle banche italiane per forma giuridica e dimensione .....	pg 70
Tabella 2.13. Livelli di efficienza di costo per area geografica (Solo BCC) .....	pg 71
Tabella 2.14. Livelli medi di efficienza di costo delle banche italiane .....	pg 72
Tabella A1.1 Test di Granger: efficienza di costo <i>versus</i> sofferenze .....	pg 76
Tabella A1.2 Test di Granger: sofferenze <i>versus</i> efficienza di costo .....	pg 76
Tabella A2.1 Stime della funzione di costo delle banche italiane (Modello I) .....	pg 79
Tabella A2.2 Stime dell'equazione di inefficienza delle banche (Modello I) .....	pg 80
Tabella A2.3 Stime della funzione di costo delle banche italiane (Modello II) .....	pg 81

Tabella A2.4 Stime dell'equazione di inefficienza delle banche italiane (Modello II) ....pg	82
Tabella 3.1 Stime della frontiera di costo, dell'equazione dell'inefficienza e del parametro di associazione della copula FGM.....pg	96

## Indice delle figure

Figura 2.1 Efficienza tecnica ed efficienza allocativa .....	pg 44
Figura 2.2 (a) Una frontiera di costo .....	pg 46
Figura 2.2 (b) Una frontiera di costo .....	pg 46
Figura 2.3 Concavità della funzione dei costi .....	pg 48
Figura 2.4 La densità dei livelli medi di efficienza di costo nel 2010.....	pg 69
Figura 2.5. Livelli medi di efficienza di costo per l'intero campione e per categoria giuridica .....	pg 72
Figura 2.6 Densità dei livelli di efficienza nel 2006 .....	pg 73
Figura 2.7 Densità dei livelli di efficienza nel 2007 .....	pg 73
Figura 2.8 Densità dei livelli di efficienza nel 2008 .....	pg 73
Figura 2.9 Densità dei livelli di efficienza nel 2009 .....	pg 73
Figura 2.10 Densità dei livelli di efficienza nel 2010 .....	pg 73
Figura 3.1 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\alpha_v = 1$ .....	pg 105
Figura 3.2 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\alpha_v = 2$ .....	pg 106
Figura 3.3 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\alpha_v = 0.25$ .....	pg 106
Figura 3.4 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\theta = 0$ .....	pg 107
Figura 3.5 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\theta = -1$ .....	pg 108
Figura 3.6 Funzione di densità del termine di errore composto $f_{\theta}^{Ex}(\varepsilon)$ per $\theta = 1$ .....	pg 108

## **Introduzione**

Il sistema bancario italiano è stato interessato, fin dai primi anni Novanta, da una serie di interventi volti a promuovere la concorrenza tra gli intermediari attraverso una profonda revisione della legislazione bancaria in ottemperanza alle direttive comunitarie. Per questa ragione, si è assistito ad un forte processo di riorganizzazione del settore bancario sotto molteplici aspetti, come la dimensione e la forma giuridica.

La deregolamentazione dell'attività creditizia e l'avvento delle tecnologie di comunicazione sono due fenomeni che hanno caratterizzato i cambiamenti del sistema bancario insieme all'avvio dell'Unione Monetaria che ha garantito l'eliminazione del rischio di cambio nell'area dell'euro favorendo l'espansione delle attività delle banche.

Le analisi empiriche sull'efficienza bancaria sono state condotte, nella maggior parte dei casi, mediante analisi della frontiera di efficienza. Si sono osservate differenze nei risultati dovute sia alle tecniche utilizzate sia alla diversa specificazione della forma funzionale della frontiera.

L'obiettivo finale della ricerca è di misurare l'efficienza di costo del sistema bancario italiano dal 2006 al 2010 utilizzando il metodo delle frontiere stocastiche. Tale approccio, assegnando una funzione di distribuzione al termine di errore, garantisce la possibilità di fare inferenza sui parametri stimati. Inoltre, scomponendo l'errore in due componenti, un termine casuale e uno attribuibile all'inefficienza, assicura che le stime dell'efficienza siano depurate da componenti casuali. Tuttavia, per definizione, il metodo richiede assunzioni riguardo alle funzioni di distribuzione delle componenti di errore. In letteratura, il modello più utilizzato prevede una distribuzione Normale per l'errore casuale e una Normale troncata per la componente di inefficienza, che si assume essere non negativa. Pertanto, il punto di avvio dell'analisi empirica è di utilizzare le tradizionali specificazioni delle frontiere di costo in cui le componenti dell'errore sono assunte essere Normali ed indipendenti tra loro. I risultati che si ottengono da queste specificazioni rappresentano il termine di confronto con quelli che si hanno introducendo modifiche alla specificazione della funzione di frontiera.

Alcuni risultati sembrano poco coerenti con gli obiettivi prefissati contenuti negli interventi legislativi degli anni Novanta in poi. Ad esempio, le banche piccole ottengono le migliori performance in termini di efficienza di costo. Inoltre, le BCC rappresentano la categoria di istituti di credito maggiormente efficienti rispetto alle altre. In una prima valutazione, questi due risultati potrebbero essere controintuitivi, ma sono comuni a molti studi presenti in letteratura che ne danno una sostanziale giustificazione. Tuttavia, uno dei principali contributi del lavoro consiste nella stima dell'equazione dell'inefficienza in maniera

simultanea a quella della frontiera di costo secondo l'approccio delle frontiere stocastiche proposto da Battese e Coelli (1995). A tal proposito, interessanti risultati si ottengono relativamente alle determinanti dell'efficienza. Innanzitutto, indagando sulla relazione tra la qualità dei crediti bancari e l'efficienza, si accetta l'ipotesi denominata di *bad management*, secondo la quale c'è una simultaneità inversa fra le due variabili (Berger e De Young, 1997). Un ulteriore risultato riguarda i miglioramenti di efficienza di costo per le banche che vantano un migliore adeguamento patrimoniale secondo le direttive dell'Unione Europea.

Un primo approfondimento dell'approccio applicato in letteratura è quello di rimuovere l'ipotesi di indipendenza tra le distribuzioni della componente di errore casuale e di inefficienza e di considerare, a tal fine, le funzioni copule. Oltre alla dipendenza, un successivo ed importante sviluppo della ricerca è di prevedere che l'errore e l'inefficienza si distribuiscano secondo funzioni diverse e più flessibili della Normale. L'idea è quella di proporre una distribuzione più generale costruita attraverso funzioni copule che consentono di modellare la dipendenza (anche non lineare) fra i due termini di errore e di considerare distribuzioni marginali anche diverse tra di loro. Da un punto di vista empirico, l'introduzione della dipendenza modellata con la copula di Farlie-Gumbel-Morgenstern (FGM) costruita utilizzando le funzioni tradizionalmente proposte in letteratura (Normale per l'errore casuale e Normale troncata per la componente di inefficienza) fornisce stime significativamente diverse da zero del parametro di associazione. Tale risultato è utile per giustificare l'approccio metodologico di costruzione di un modello generale che considera la dipendenza fra i due termini di errore e funzioni di distribuzioni marginali più flessibili rispetto a quelle Normali. Un risultato da evidenziare è legato al fatto che caratteristiche distributive non catturate dalle funzioni Normali influenzano la varianza dell'errore composto utilizzata per la sua scomposizione, influenzando, quindi, anche la stima dei livelli di efficienza.

Il lavoro è organizzato nel modo seguente. Nel primo capitolo si descrivono i maggiori cambiamenti del settore bancario italiano, e, dopo aver discusso le metodologie più comuni in letteratura per stimare l'inefficienza dei sistemi bancari, si approfondisce la discussione sugli studi finalizzati alla valutazione dei livelli di efficienza e allo studio dei fattori che la determinano. A riguardo, si propone una sintesi di alcuni fra i più citati lavori che si sono occupati di stimare gli *score* di efficienza del sistema bancario italiano e dei maggiori paesi industrializzati. Il secondo capitolo presenta le stime dei livelli di efficienza di costo delle banche italiane dal 2006 al 2010 mediante l'approccio delle frontiere stocastiche, considerando indipendenza fra i due termini di errore. Inoltre, sono presenti le stime relative all'equazione dell'inefficienza ottenute seguendo il metodo proposto da Battese e Coelli

(1995). Nell'ultimo capitolo si modella la dipendenza fra l'errore casuale e l'inefficienza mediante la funzione copula FGM da un punto vista metodologico e, inoltre, si riportano le stime della frontiera di costo e degli effetti delle determinanti sull'efficienza nel caso specifico di normalità degli errori e dipendenza.

## CAPITOLO I

### La ristrutturazione e l'efficienza del sistema bancario italiano

#### 1.1 Introduzione

Gli anni Novanta rappresentano un periodo caratterizzato da intensi cambiamenti nei sistemi bancari europei finalizzati a promuovere la concorrenza tra gli intermediari attraverso una profonda revisione della legislazione bancaria. Per questa ragione, si è assistito ad un forte processo di riorganizzazione degli assetti proprietari e delle forme istituzionali delle aziende di credito.

In Italia, questi orientamenti alla ristrutturazione, se da un lato, sono stati necessari per uniformare il sistema italiano a quello europeo al fine di renderlo più competitivo, dall'altro lato, però, proprio per la dualità che caratterizza l'economia italiana, si è assistito all'inasprimento di caratteristiche tradizionali di problematicità legate ai rapporti banche-imprese nelle aree più deboli del Mezzogiorno (Alessandrini, Papi e Zazzaro, 2003).

Il caso italiano è, pertanto, interessante sotto molti punti di vista. Una particolarità riguarda, come sostiene Giannola (2009), l'esistenza di due diversi sistemi bancari per le due aree territoriali del paese (Nord e Sud), le quali presentano distinte caratteristiche strutturali, economiche ed istituzionali. Data la dualità del sistema, la ristrutturazione del settore bancario italiano non è stata affatto uniforme ed omogenea. Da un punto di vista empirico, emerge che, in generale, i comportamenti degli istituti di credito nei confronti della clientela del Sud dipendono in misura rilevante dalle relazioni personali che si instaurano tra banca e privati, mentre al Nord le procedure risultano maggiormente standardizzate (Girardone, Molyneux e Gardener 2004; Lucchetti, Papi e Zazzaro, 1999).

Uno dei fattori di innovazione a livello istituzionale è la profonda deregolamentazione dell'attività creditizia. Un forte stimolo al cambiamento è dovuto ad una serie di interventi legislativi relativi alla liberalizzazione valutaria<sup>1</sup>, alla rimozione dei vincoli all'apertura degli sportelli (Legge Amato-Carli del 1990), al recepimento della II Direttiva comunitaria, all'entrata in vigore del Testo Unico del 1993 (*cf.* § 1.2 e 1.3). Si tratta di interventi necessari, in quanto, a partire dagli anni Novanta, gli intermediari creditizi si confrontano su mercati aperti alla concorrenza, interna ed estera, a differenza di ciò che avveniva in passato, in cui le banche operavano in ambiti protetti.

---

<sup>1</sup> Il processo di liberalizzazione valutaria è stata completata nel 1990 con il Decreto Ministeriale del 27 aprile, in recepimento della Direttiva CEE del 1988.

Altro fattore di stimolo al cambiamento è intervenuto con l'avvento delle tecnologie di comunicazione (Panetta, 2003) che, in qualche modo, hanno agevolato la concentrazione presso una o poche filiali della gestione dei servizi offerti alla clientela, che, al contrario, è localizzata in diverse aree del paese. A tal proposito, occorre evidenziare che tale fenomeno ha garantito maggiori guadagni non derivanti dalle operazioni tradizionali di raccolta e impiego, bensì dovuti a servizi accessori offerti dalle banche.

Un ulteriore fondamentale aspetto è costituito dall'avvio all'Unione Monetaria, poiché l'eliminazione del rischio di cambio nell'area dell'euro ha favorito l'espansione delle attività delle banche in mercati anche diversi da quelli di origine. In definitiva, si è assistito all'abbattimento di barriere geografiche alla concorrenza all'interno del paese e di barriere alla concorrenza tra i sistemi bancari nazionali.

Da un punto di vista degli studi originati dai cambiamenti del sistema bancario italiano, un importante filone di ricerca è quello legato alla misurazione delle performance delle aziende di credito. La ricerca sull'efficienza bancaria è stata condotta, nella maggior parte dei casi, mediante analisi della frontiera di efficienza. L'evidenza empirica sull'efficienza delle banche considera risultati molto eterogenei e tali differenze derivano sia dalle tecniche utilizzate (per esempio, metodi di programmazione lineare oppure metodi parametrici) sia dalla diversa specificazione della forma funzionale della frontiera.

In questo capitolo, dopo aver brevemente descritto i maggiori cambiamenti del settore (cfr. § 1.2 e 1.3), e dopo aver discusso le metodologie più comuni in letteratura per stimare l'inefficienza dei sistemi bancari (cfr. § 1.4), si approfondisce la discussione sugli studi finalizzati alla misurazione della performance e sulle determinanti dei livelli di efficienza (cfr. § 1.5). A riguardo, si propone una sintesi di alcuni fra i più importanti lavori che si sono occupati di stimare gli *score* di efficienza del sistema bancario italiano e dei maggiori paesi industrializzati.

## **1.2 Cambiamenti istituzionali ed operativi del sistema bancario italiano**

Sul finire di degli anni Ottanta, il sistema bancario italiano si presentava segmentato e a controllo prevalentemente pubblico e, quindi, relativamente impermeabile alla concorrenza degli intermediari esteri.

Il primo intervento normativo con l'obiettivo di svincolare il sistema dalle rigidità istituzionali che lo caratterizzavano è costituito dalla Seconda Direttiva CEE di coordinamento bancario (n°646 del 1989), ovvero la seconda direttiva comunitaria

sull'accesso e l'esercizio dell'attività creditizia da parte degli intermediari bancari con la quale si avviavano le procedure per la creazione del Mercato Bancario Unico Europeo.

In ottemperanza alla normativa comunitaria, nel 1990 è stata promulgata la Legge Amato-Carli (Legge n°218 del 1990) per ufficializzare benefici fiscali a favore delle banche di diritto pubblico trasformate in società per azioni. Con questa legge si riconosce agli enti creditizi di tipo pubblico la possibilità di modificare la propria forma giuridica secondo il modello della società per azioni, al fine di accrescere il capitale di rischio attraverso l'emissione di titoli azionari. Grazie a questo intervento legislativo, inoltre, si riconosce la neutralità fiscale delle operazioni di riorganizzazione e ristrutturazione delle banche che necessitano di tali processi (Messori, Tamburini e Zazzaro, 2003).

**Tabella 1.1 L'evoluzione della struttura del sistema bancario italiano**

	1993		2000		2008		2010	
Gruppi bancari	<b>94</b>	-	<b>74</b>	-	<b>75</b>	-	<b>76</b>	-
Banche	<b>1037</b>	100%	<b>841</b>	100%	<b>788</b>	100%	<b>760</b>	100%
di cui: <i>Banche Spa</i>	225	22%	240	29%	247	31%	233	31%
<i>Banche Popolari</i>	100	10%	44	5%	38	5%	37	5%
<i>BCC</i>	671	64%	499	59%	421	54%	415	54%
<i>Succursali di banche estere</i>	41	4%	58	7%	82	10%	75	10%

Fonte: elaborazione su dati Banca d'Italia (Relazione Annuale, 1998-2010).

Dalle tabelle 1.1 e 1.2, emerge l'evoluzione della struttura del sistema bancario italiano caratterizzata da un grande cambiamento verificatosi negli anni Novanta fino ai primi anni del decennio successivo, mentre risulta abbastanza costante nell'ultimo periodo. I gruppi bancari erano 94 nel 1993, mentre tale numero scendeva a 76 nel 2010. Analogamente, agli inizi degli anni Novanta, il numero degli istituti di credito operanti sul territorio italiano superava i mille, mentre nel 2010, era molto minore di 800. La composizione interna è cambiata nella direzione secondo la quale sono aumentate le banche Società per Azioni (Spa; dal 22% nel 1993 al 31% nel 2010) e le succursali di banche estere (dal 4% nel 1993 al 10% nel 2010), mentre, sono diminuite le banche popolari (dal 10% nel 1993 al 5% nel 2010) e le BCC (dal 64% nel 1993 al 54% nel 2010).

Una specificità della legge Amato-Carli è rappresentata dalla regolamentazione delle cessioni delle segmentazioni legate alla durata temporale del credito, poiché, l'operatività delle banche si allargava a tutti i comparti del credito, a medio e a lungo termine. In tal modo, un ente creditizio poteva esercitare tutte le attività finanziarie, ad esclusione di quelle

assicurative. Questo contribuì ad innescare numerosi processi di concentrazione avvenuti a partire dagli anni Novanta (tabella 1.2). Dal 1991 al 2006, si sono registrate 635 operazioni di concentrazione, di cui 426 fusioni e incorporazioni, mentre per la restante parte si tratta di operazioni di acquisizione della maggioranza del capitale.

Con l'entrata in vigore del Testo Unico in materia bancaria e creditizia del 1993 si avvia, inoltre, un processo di sistematizzazione del nuovo quadro normativo. Si procede nella direzione di uniformare le banche sul piano giuridico in un contesto in cui non ci sono limitazioni alle operazioni e ai servizi offerti, nonché, alle scadenze nella raccolta e nell'impiego.

**Tabella 1.2 Le operazioni di fusione, incorporazione e trasferimenti di controllo**

	Fusioni e incorporazioni		Acquisizioni della maggioranza del capitale	Totale operazioni di concentrazione
	Totale	di cui BCC		
1991	33	22	5	38
1992	20	9	1	21
1993	38	25	6	44
1994	42	25	10	52
1995	47	28	19	66
1996	37	25	19	56
1997	24	12	18	42
1998	27	18	23	50
1999	36	23	28	64
2000	33	22	24	57
2001	31	21	9	40
2002	18	16	11	29
2003	20	14	9	29
2004	10	7	7	17
2005	4	3	7	11
2006	6	3	4	10
<i>Totale</i>	<b>426</b>	<b>287</b>	<b>209</b>	<b>635</b>

Fonte: elaborazione su dati Banca d'Italia (Relazione Annuale, 1998-2010).

Riaffermando il concetto della despecializzazione temporale ed operativa, il Testo Unico tende verso il principio dell'universalità del sistema creditizio, ribadito nella Seconda Direttiva comunitaria. Si registrano, pertanto, economie di diversificazione che hanno comportato l'ampliamento della gamma dei servizi offerti alla clientela. In tale contesto, aumentano gli obiettivi della vigilanza esercitata dalla Banca d'Italia e dal Comitato interministeriale per il credito e il risparmio, poiché, ai classici obiettivi di stabilità, di osservanza delle norme e di sana e prudente gestione, si aggiungono quelli di controllo sull'efficienza e sulla competitività del sistema finanziario.

Altra fonte di rinnovamento legislativo in materia bancaria è la Direttiva Dini (Provvedimento emanato dal Ministero del Tesoro “Dini” – n°474 del 1994). Essa facilita le operazioni di concentrazione perché tende a favorire il trasferimento del controllo delle banche dalle Fondazioni a soggetti privati. Su questa strada, si è diffusa la convinzione che vi è concorrenza effettiva solo nel caso in cui gli operatori raggiungano adeguate capacità tecniche ed operative e dimensioni sufficienti per competere sul mercato (Panetta, 2004). Pertanto, le concentrazioni rappresentano un risultato della concorrenza originate dall’esigenza di affrontare un mercato molto competitivo.

Tuttavia, bisogna considerare che la diffusione delle operazioni di concentrazione ha, in qualche modo, provocato effetti distorsivi nella concessione dei prestiti. Infatti, l’incorporazione di piccole banche in organismi maggiori comporta la riduzione dei finanziamenti alle imprese caratterizzate da operatività locale, in grado di contribuire alla crescita della produzione e dell’occupazione. Tutto ciò accade perché il finanziamento di tali attività è fondato sulle cosiddette *soft information* “impalpabili” nei bilanci di queste imprese. In tale contesto, le banche locali (in prevalenza le BCC e le banche popolari) possono affrontare e trattare meglio il finanziamento e l’assistenza creditizia a favore delle piccole imprese. Questo, quindi, in alcuni casi, evita gli scompensi generati da un grave razionamento a seguito della contrazione del volume del credito erogato dai grandi intermediari nati dalle fusioni (Angelini e Cetorelli, 2004).

Dal punto di vista dell’analisi economica, a partire dagli anni Novanta, si è assistito ad una proliferazione di studi sulla misurazione dei livelli di efficienza del sistema bancario italiano evidenziando i cambiamenti dovuti alle operazioni di *merger and acquisitions*. All’interno di questa letteratura, emergono i lavori che focalizzano la loro attenzione sulla relazione tra il grado di efficienza e quello di concorrenza del sistema bancario, quelli che misurano le economie di scala potenzialmente conseguenti alle operazioni di concentrazione, quelli che esaminano la relazione tra l’efficienza e l’allocazione del credito e quelli che studiano la convergenza nei livelli di efficienza tra le diverse aree territoriali di un paese (Lucchetti, Papi e Zazzaro, 1997).

Secondo Ciocca (2004), gli Accordi di Basilea considerati congiuntamente ai principi contabili internazionali<sup>2</sup>, hanno l’obiettivo di garantire guadagni di efficienza e, attraverso questo, una maggiore stabilità. Pertanto, secondo Ciocca, questo si può avere soltanto se si riesce a garantire e mantenere una concorrenza intensificata. In particolare, gli Accordi di

---

<sup>2</sup> I principi internazionali IAS sono stati definiti con l’obiettivo di standardizzare le regole e le procedure contabili che devono essere osservate e rispettate dalle imprese.

Basilea II sollecitano l'efficienza delle banche, in quanto sono premiate con minori obblighi di capitale le banche che usano metodi più precisi nella quantificazione e più efficaci nella gestione dei rischi di credito. In questo senso, il nuovo Accordo rappresenta un fattore di selezione "darwiniana" delle banche migliori, comportando un'intensificazione della concorrenza ed una redistribuzione delle quote di mercato. Il Nuovo Accordo si basa su tre pilastri fondamentali che costituiscono un insieme unitario, poiché si rafforzano reciprocamente e concorrono ad una maggiore stabilità e solidità del sistema finanziario dei paesi che aderiscono all'accordo (D'Agostino, 2003). Il primo pilastro si riferisce ai requisiti patrimoniali minimi e riguarda analiticamente i livelli minimi di capitale, l'aderenza tra rischi e capitale e le nuove categorie di rischio. Il secondo pilastro si riferisce al controllo prudenziale sull'adeguatezza patrimoniale delle banche. Infine, il terzo pilastro riguarda la disciplina dei mercati e mira a favorire una stabilità diffusa supportando l'attività delle autorità di vigilanza.

Un ulteriore fenomeno che caratterizza l'evoluzione del sistema bancario italiano, già accennato peraltro, risulta essere l'avvento delle tecnologie di comunicazione. Negli anni Novanta il progresso tecnologico nel campo dell'informazione e una forte integrazione delle maggiori economie mondiali hanno favorito la formazione di un mercato unico dei capitali a livello mondiale. Si consideri, inoltre, che la deregolamentazione dell'attività finanziaria e la crescente integrazione economica europea hanno creato i presupposti per un confronto competitivo non più circoscritto entro i confini nazionali. Le strategie attuate dalle banche hanno mirato ad ampliare la gamma dei servizi offerti, ad aumentare l'efficienza aziendale e raggiungere dimensioni tali da garantire un certo livello di competizione su scala internazionale. Ciò giustifica la crescente tendenza ad una forte espansione di operazioni di concentrazione tra banche.

Lo studio dell'evoluzione dei livelli di efficienza dovuta a questi cambiamenti nel quadro normativo e negli assetti proprietari è stato molto approfondito in letteratura (Girardone, Molyneux e Gardener, 2004)<sup>3</sup>. Tra gli studi sulle modifiche normative sulla concorrenza nel mercato creditizio, Angelini e Cetorelli (2004)<sup>4</sup> analizzano l'effetto del recepimento della Seconda Direttiva di coordinamento bancario, che mira a rimuovere importanti barriere all'ingresso in nuovi mercati. La cosiddetta deregolamentazione del settore bancario italiano con una conseguente armonizzazione del sistema appariva strumentale alla

---

<sup>3</sup> Si rimanda al paragrafo 1.5 per una breve rassegna di alcuni lavori sulla misurazione dell'efficienza del sistema bancario.

<sup>4</sup> L'analisi empirica ha riguardato il periodo compreso tra il 1983 e il 1997, su dati di bilancio ricavati dalle segnalazioni di Vigilanza. La misura di concorrenza utilizzata è l'indice di Lerner.

creazione di un mercato unico e concorrenziale dei servizi finanziari. Dall'altra parte, però, sembrano legittimi i dubbi degli autori secondo i quali le operazioni di concentrazione strumentali al raggiungimento di livelli di efficienza più alti dovuti, per esempio, all'ottenimento di economie di scala e di scopo, potrebbero avere l'effetto contrario. Ossia una maggiore concentrazione del sistema bancario potrebbe ridurre la concorrenza con l'eventualità di comportamenti monopolistici da parte dei grandi gruppi bancari. Tuttavia, aumenti di efficienza effettivi comportano riduzioni di prezzi, con effetti netti positivi sulla concorrenza. Dalle analisi empiriche (cfr. § 1.5) non emergono, però, conclusioni univoche in merito al conseguimento di più alti livelli di efficienza dovuti alle concentrazioni avvenute dagli anni Novanta in poi. I risultati ottenuti evidenziano un considerevole rafforzamento del grado di concorrenza nel settore bancario italiano e notevoli miglioramenti di efficienza a vantaggio della clientela ottenuti grazie alle numerose operazioni di fusione e acquisizione.

Quindi, da un lato, l'unione degli intermediari accresce il loro potere di mercato e comporta un aumento dei prezzi, danneggiando la clientela. Dall'altro lato, i potenziali guadagni di efficienza consentono alle banche di ridurre il prezzo dei prodotti offerti, garantendo condizioni più vantaggiose ai clienti. Alcuni studi su altri paesi hanno rilevato la prevalenza del potere di mercato sui guadagni di efficienza (Casu e Molyneux, 2003). Tuttavia, è utile considerare che, se da un lato, l'aumento del potere di mercato in seguito ad operazioni di concentrazione potrebbe manifestare i suoi effetti nel breve periodo, i risultati sui livelli di efficienza potrebbero essere visibili solo nel medio-lungo termine.

### **1.3 La ristrutturazione e la dualità del sistema bancario italiano**

Come già accennato, il processo di ristrutturazione del sistema bancario italiano è stato tutt'altro che omogeneo da un punto di vista dimensionale e territoriale. Questo è dovuto alla dualità tipica dei sistemi produttivi, istituzionali, economici e sociali presenti in Italia (Ferri e Messori 2000; Giannola, 2009).

L'aumento della concorrenza e l'allentamento dei vincoli alla "bancarizzazione" del territorio hanno comportato effetti rilevanti per il sistema bancario del Mezzogiorno, ben distinto da quello settentrionale (Giannola, 2009). La difficoltà di contesto in cui le banche del Sud sono costrette a operare e i più contenuti coefficienti patrimoniali hanno contribuito al dissolvimento del sistema creditizio locale che, a partire dalla fine degli anni Novanta, è stato assorbito dalle banche del Centro-Nord. Dall'altro lato, la necessità di coprire il territorio ha comportato la proliferazione di sportelli bancari in tutte le regioni (Rapporto Svimez 2010).

Nel corso del tempo, se da un lato, si è assistito alla corsa allo sportello da parte delle banche registrando una proliferazione delle filiali nelle aree periferiche, dall'altro, si è registrata una diminuzione degli istituti operanti nel territorio nazionale (tabella 1.3). Nello stesso periodo, l'operatività delle banche è aumentata, visto l'aumento considerevole del numero degli sportelli radicati sul territorio (Panetta, 2003). Dalla tabella 1.3 emerge, infatti, che dal 1993 al 2008 gli sportelli aumentano considerevolmente (da 22133 a 34139).

**Tabella 1.3 Banche e sportelli operanti in Italia nel 1993, nel 2002 e nel 2008**

	1993				2002				2008			
	Numero banche		Numero sportelli		Numero banche		Numero sportelli		Numero banche		Numero sportelli	
Banche Spa	225	22%	15961	72%	253	31%	22924	76%	247	31%	26808	78%
Banche Popolari	100	10%	3896	17%	40	5%	3704	12%	38	5%	2997	9%
BCC	671	64%	2226	10%	461	57%	3192	11%	421	54%	4109	12%
Succursali di banche estere	41	4%	50	0.23%	60	7%	106	0.35%	82	10%	225	0.66%
<b>Totale</b>	<b>1037</b>	<b>100%</b>	<b>22133</b>	<b>100%</b>	<b>814</b>	<b>100%</b>	<b>29926</b>	<b>100%</b>	<b>788</b>	<b>100%</b>	<b>34139</b>	<b>100%</b>

Fonte: Baglioni e Petrella (2003), Relazione Annuale Banca d'Italia (1998-2010).

Nel 1990, le banche con sede legale nel Mezzogiorno presentavano gravi condizioni strutturali ed operative, in un contesto in cui i ricavi derivavano soprattutto dalle tradizionali attività di prestito. Per esempio, i costi totali assorbivano oltre il 70% dei ricavi, mentre questo rapporto era pari a circa il 50% per le banche del Centro-Nord. Inoltre, dalle analisi empiriche, emerge che i guadagni di efficienza derivanti dalle operazioni di concentrazione ottenuti al Nord sono di molto superiori a quelli ottenuti nel Mezzogiorno (Baglioni e Petrella, 2003).

Spostando l'attenzione sul costo del credito, la concentrazione della proprietà delle banche del Mezzogiorno nelle mani delle banche del Nord avrebbe potuto creare effetti distorsivi provocando un innalzamento di tali costi. Anche effettuando opportune correzioni relative alla composizione dimensionale e settoriale della clientela, Panetta (2003) evidenzia che il costo del credito nel Mezzogiorno rimane più elevato rispetto al Centro-Nord. Si consideri, inoltre, che la maggiore rischiosità da cui deriva, tra l'altro, anche una parte di differenziale di costo, riflette le difficoltà del sistema produttivo e la mancanza di

infrastrutture le quali frenano la produttività del lavoro. L'autore sostiene che una maggiore rischiosità sia sintomo anche della cattiva qualità della Pubblica Amministrazione e della bassa efficienza delle procedure giudiziarie che determinano, in caso di contenziosi sull'esigibilità del credito, margini di incertezza elevati e tempi di attesa interminabili. Pertanto, il differenziale sui costi non è solo il risultato di un cattivo funzionamento del sistema creditizio quanto piuttosto delle difficoltà strutturali che caratterizzano il Mezzogiorno.

**Tabella 1.4 Evoluzione del rapporto Impieghi/Depositi nel Sud e nel Centro-Nord**

	(Impieghi + Sofferenze)/Depositi Mezzogiorno (a)	(Impieghi + Sofferenze)/Depositi Centro-Nord (b)	a/b
1995	115.97	136.70	0.85
1996	114.98	135.89	0.85
1998	131.12	167.36	0.78
2000	108.06	167.02	0.65
2001	105.70	168.80	0.63
2002	106.43	163.66	0.65
2003	112.34	171.74	0.65
2004	118.31	171.47	0.69
2005	122.30	171.88	0.71
2006	131.59	176.87	0.74

Fonte: Giannola (2009).

Giannola (2009) evidenzia, inoltre, che quando una banca del Nord si fonde con un istituto avente sede legale nel Sud, la clientela di quest'ultima viene penalizzata in quanto le condizioni di accesso al credito peggiorano rispetto alla situazione di partenza. Secondo l'analisi proposta da Giannola (2009), il rapporto impieghi/depositi al Sud rispetto a quello del Nord è passato dallo 0.85 del 1995 allo 0.74 del 2006. Pertanto, nel periodo caratterizzato da operazioni di fusione, incorporazione e acquisizione della maggioranza di capitale con maggiore intensità, si assiste ad un aumento delle distanze tra le due aree dovuto anche all'incapacità delle banche del Mezzogiorno di utilizzare nel proprio territorio i fondi raccolti. Nel 2009, il Mezzogiorno, che vanta il 35% della popolazione e il 25% del PIL nazionale, raggiunge solo il 15% degli impieghi sul totale nazionale. In questo contesto di maggiore difficoltà al Sud, è plausibile che la probabilità di successo di un finanziamento concesso sia

inferiore rispetto a quanto accade nell'altra area del Paese. Pertanto, nel meridione d'Italia, le banche si trovano ad operare in un clima caratterizzato da una notevole rischiosità e da una forte debolezza strutturale. Dalla tabella 1.4 emerge chiaramente che nel Sud i fondi raccolti vengono impiegati con maggiore difficoltà rispetto al Nord, contribuendo ad alimentare il processo di divergenza tra i sistemi bancari delle due regioni. Nel 1995, il rapporto tra impieghi e depositi nel Mezzogiorno raggiunge un valore di circa 116, mentre nel Centro-Nord lo stesso rapporto supera di poco 136. Nel corso del tempo, il distacco aumenta e nel 2006 esso raddoppia rispetto al 1995. Questo è dovuto soprattutto ad un aumento degli impieghi nel Meridione a tassi più contenuti rispetto al Centro-Nord (Mattesini e Messori, 2004).

Un ulteriore aspetto riguarda il ruolo delle banche nella selezione, nel finanziamento e nel controllo di progetti socialmente efficienti. Tuttavia, mentre le banche del Centro-Nord hanno avviato processi di ristrutturazione già a partire dai primi anni Novanta con l'obiettivo di raggiungere livelli più elevati di efficienza, quelle del Sud mostrano pesanti inefficienze gestionali ed allocative che, spesso, si sono tradotte in fallimenti<sup>5</sup>. In generale, il salvataggio degli istituti in crisi si è tradotto nell'incorporazione (fusione o controllo) da parte delle banche settentrionali.

La portata quantitativa di questi cambiamenti è mostrata nella tabella 1.5 dalla quale si nota che i processi di consolidamento, che hanno avuto maggiore intensità nel periodo 1990-2002, hanno determinato una diminuzione del numero di banche che è stata molto più accentuata nel Mezzogiorno, passando dal 29% nel 1990 a 21% nel 2002.

Si noti che nel Centro-Nord il numero della BCC si è ridotto notevolmente tra il 1990 e il 2008. Infatti, il 67% delle banche che hanno sede legale al Centro-Nord è costituito da BCC nel 1990. Tale percentuale è pari al 50% nel 2008. Al Sud, invece, tra il 1990 e il 2008, il numero delle BCC aumenta e la percentuale degli istituti con sede legale al Sud è pari al 68% nel 1990, 75% nel 2008. Questi dati indicano come il processo di riorganizzazione del sistema bancario si sia caratterizzato da una diversa intensità geografica, spostando il controllo del sistema a favore degli istituti del Centro-Nord. Infatti, nel Mezzogiorno, il consolidamento si è avuto ad opera di banche extraregionali, cioè di banche aventi sede legale nell'area settentrionale del paese e tutto ciò ha generato come esito la scomparsa di banche/gruppi bancari meridionali indipendenti rispetto a quelli del resto del paese.

---

<sup>5</sup> Alcune casi di fallimenti di particolare rilevanza a livello nazionale hanno riguardato il Banco di Napoli, il Banco di Sicilia, quello di Sardegna e le Casse di Risparmio dell'area meridionale.

**Tabella 1.5 Numero di banche classificate per area territoriale**

	1990		1995		2002		2008	
Totale Italia	1061	100%	918	100%	814	100%	788	100%
<b>Centro-Nord</b>	<b>748</b>	71%	<b>641</b>	70%	<b>646</b>	79%	<b>640</b>	81%
<i>di cui BCC</i>	500	67%	425	66%	351	54%	321	50%
<b>Mezzogiorno</b>	<b>313</b>	29%	<b>277</b>	30%	<b>168</b>	21%	<b>148</b>	19%
<i>di cui BCC</i>	213	68%	194	70%	121	72%	111	75%

Fonte: Relazione Annuale Banca d'Italia (1998-2008).

#### 1.4 La misurazione dell'efficienza

I cambiamenti strutturali, organizzativi e degli assetti proprietari avvenuti nei settori bancari di molti paesi a partire dai primi anni Novanta hanno comportato la proliferazione degli studi che si sono occupati di studiarne l'efficienza sotto molteplici punti di vista (si veda la tabella 1.7 alla fine del capitolo). Da un punto di vista metodologico, la misurazione dell'efficienza implica la costruzione del luogo geometrico dei punti che rappresenta l'insieme dei piani di produzione più efficienti e il calcolo della distanza tra la combinazione output-input osservata per la singola banca e la frontiera stimata<sup>6</sup>. Esistono in letteratura vari metodi per misurare l'efficienza di un sistema produttivo. La tabella 1.6 riassume i limiti e i vantaggi delle metodologie comunemente più utilizzate, le quali sono la Data Envelopment Analysis (DEA), il Free Disposal Hull (FDH), le frontiere stocastiche (SFA), il Distribution Free Approach (DFA) e il Thick Frontier Approach (TFA). In generale, una prima distinzione è tra metodi parametrici e quelli non parametrici<sup>7</sup>. I primi, tra cui SFA, impongono una forma

<sup>6</sup> Questo lavoro utilizza il concetto di frontiera di costo, la quale rappresenta l'insieme delle combinazioni di input ed output che consentono di produrre una determinata quantità di output, minimizzando i costi. Nel capitolo 2 è presente una trattazione estesa del concetto.

<sup>7</sup> La distinzione tra i due tipi di metodi è relativa al fatto che i metodi parametrici assegnano funzioni di densità alla componente stocastica del modello, mentre i metodi non parametrici definiscono solo la parte deterministica. Si noti che in un modello del tipo  $y = x\beta + \varepsilon$ , la componente deterministica è rappresentata da  $x\beta$  e la componente stocastica è costituita da  $\varepsilon$ .

funzionale alla funzione di produzione o di costo e, in alcuni casi<sup>8</sup>, si fanno delle assunzioni relativamente alle variabili che rappresentano la componente di inefficienza e l'errore casuale. L'approccio non parametrico, qual è la DEA, non impone alcuna funzione di produzione né alcuna distribuzione della componente di inefficienza, ma non prevede una parte casuale di errore (D'Orio, 2003).

Molti dei lavori (Weill, 2004) che presentano il confronto tra i due diversi tipi di approcci hanno evidenziato differenze sostanziali nei risultati. Altri lavori, invece, hanno mostrato soltanto differenze nei livelli di efficienza stimati mediante diversi metodi. In particolare, l'approccio non parametrico tende a sottostimare gli *score* di inefficienza (Beccalli, Casu e Girardone, 2006).

L'approccio delle frontiere stocastiche è un metodo di tipo parametrico che gode del vantaggio di distinguere la distanza dalla frontiera di massima efficienza in due componenti, una casuale e l'altra tipicamente ascrivibile all'inefficienza. Questa metodologia presuppone la scelta a priori sia di una funzione di costo o profitto (che può essere di tipo Cobb-Douglas, Translog, Fourier-flexible) sia della funzione di distribuzione della componente erratica del modello (Battese e Coelli, 1995). La stima può riguardare sia funzioni di costo sia funzioni di profitto. Quando l'obiettivo dell'analisi è quello di verificare l'esistenza di economie o diseconomie di scala e/o di scopo, si stima la frontiera stocastica di costo. Al contrario, si stima una frontiera di profitto quando la questione centrale è quella di misurare il gap di efficienza tecnica (Battese *et al.*, 2005; Kumbhakar e Lovell, 2000). La maggior parte dei lavori esaminati<sup>9</sup> attribuisce una distribuzione Normale e semi-Normale alla componente erratica<sup>10</sup>, ipotizzando indipendenza tra le due. Oltre al vantaggio di distinguere l'efficienza dall'errore casuale, il metodo SFA consente anche di isolare l'eterogeneità individuando fattori che spiegano e influenzano la componente di inefficienza. Tali variabili sono identificate da un generico vettore  $z$  che viene introdotto nelle stime per specificare l'equazione di inefficienza (Battese e Coelli, 1995).

La DEA è, invece, un metodo non parametrico. Si caratterizza per l'impiego di tecniche di programmazione lineare per mezzo delle quali si perviene alla costruzione di indici di inefficienza senza richiedere alcuna specificazione del modello (Ferrier e Lovell,

---

<sup>8</sup> Nel caso dell'approccio delle Frontiere Stocastiche e nel Thick Frontier Approach, si impone la distribuzione degli errori, mentre nel Distribution Free Approach non viene fatta alcuna assunzione sulle distribuzioni della componente erratica.

<sup>9</sup> Dongili, Rossi, Zago (2008), Fiordelisi, Marques-Ibanez e Molyneux (2010), Fontani e Vitali (2008), Girardone, Molyneux e Gardener (2006), Giordano e Lopes (2008), Humprey e Pulley (1997), Molyneux (2003), Turati (2008).

<sup>10</sup> In particolare, si assume che l'errore casuale si distribuisca come una Normale e la componente erratica di inefficienza come una Normale troncata sullo zero.

1990). Tale metodo non consente di distinguere la componente propriamente stocastica da quella ascrivibile all'inefficienza ed essendo una metodologia di tipo non-parametrico, non consente di effettuare test di significatività sui parametri (Farrell, 1957). Il metodo FDH è un caso particolare della DEA (Berger e Humphrey, 1997; Tulkens, 1993). Infatti, a differenza di quello precedente, tale metodo include nella frontiera e, quindi, nel set delle possibilità, non solo i vertici che congiungono la frontiera costruita mediante la DEA, ma anche i punti interni a questi vertici (Destefanis, 2001).

Il Thick Frontier Approach consiste in un metodo parametrico che permette di ottenere i valori previsti di performance a partire da una determinata specificazione funzionale della frontiera (di produzione o di costo). Secondo tale metodo, i valori della frontiera vengono raggruppati in intervalli quartilici, assumendo che le differenze fra i valori previsti all'interno di ciascun intervallo siano espressione di fluttuazioni casuali e che le differenze rilevabili tra intervalli siano dovute all'inefficienza (Berger e Humphrey, 1997). Le frontiere vengono spesso stimate attraverso il metodo OLS, per ogni quartile (Dongili, Rossi e Zago, 2008), infatti, si assume che all'interno di ogni quartile la distanza dalla frontiera sia dovuta solo all'errore casuale. Per questi motivi, è necessario che il grado di inefficienza vari al variare della classe quartilica e che sia presente una componente casuale di errore all'interno di ciascun quartile. Tale approccio non consente di stimare in maniera puntuale il livello di inefficienza, bensì permette di ottenere una stima del livello generale di inefficienza. Altro aspetto critico di tale metodologia risulta essere l'inconsistenza di una divisione arbitraria del campione di banche in quartili (o altra forma di quantili).

Infine, il Distribution Free Approach rappresenta un altro metodo parametrico mediante il quale si intende stimare una frontiera di efficienza, sia essa di costo o di profitto. Sebbene venga classificato come metodo parametrico, non formula ipotesi in merito alla distribuzione delle inefficienze o degli errori (Berger e Humphrey, 1997). Il DFA ipotizza unicamente che, nel tempo, l'efficienza per ogni impresa sia stabile, mentre gli errori tendono verso una media pari allo zero.

**Tabella 1.6 Caratteristiche, vantaggi e limiti degli approcci parametrici e non parametrici utilizzati per misurare l'efficienza delle imprese**

	Approcci non parametrici		Approcci parametrici		
	DEA	FDH	SFA	DFA	TFA
Forma funzionale della frontiera	Non specificata	Non specificata	Specificata	Specificata	Specificata
Termini di errore	Non previsti	Non previsti	Errore composto: - Componente di inefficienza - Errore casuale	Errore composto: - Componente di inefficienza - Errore casuale	Errore composto: - Componente di inefficienza - Errore casuale
Efficienza	- Variabile nel tempo - Stima puntuale	- Variabile nel tempo - Stima puntuale	- Variabile nel tempo - Stima puntuale	- Stabile nel tempo - Stima puntuale	- Variabile nel tempo - Stima del livello generale
Vantaggi	- Nessun vincolo legato alla forma funzionale della frontiera  - Nessun vincolo sulla distribuzione della componente erratica  - Fornisce una stima puntuale del grado di efficienza aziendale	- Nessun vincolo legato alla forma funzionale della frontiera  - Nessun vincolo sulla distribuzione della componente erratica  - Fornisce una stima puntuale del grado di efficienza aziendale  - Non si fanno ipotesi sulla convessità del set di produzione	- Presenza della componente erratica composta da una parte casuale e dalla componente di inefficienza  - Fornisce una stima puntuale del grado di inefficienza aziendale	- Presenza della componente erratica composta da una parte casuale e dalla componente di inefficienza senza vincoli distributivi  - Fornisce una stima puntuale del grado di inefficienza aziendale	- Presenza della componente erratica composta da una parte casuale e dalla componente di inefficienza senza vincoli distributivi
Limiti	- Mancanza della componente erratica  - Impossibilità di fare inferenza sui parametri stimati	- Mancanza della componente erratica  - Impossibilità di fare inferenza sui parametri stimati	- Vincoli sulla scelta della distribuzione della componente erratica  - Vincoli sulla scelta della funzione di produzione o di costo	- Vincoli sulla scelta della funzione di produzione o di costo  - Assunzione della stabilità della inefficienza nel tempo	- Vincoli sulla scelta della funzione di produzione o di costo  - Vincoli sulla scelta della distribuzione della componente erratica  - Non fornisce una stima puntuale della inefficienza  - Divisione arbitraria in quartili

### 1.5 Rassegna della letteratura empirica sullo studio dell'efficienza dei sistemi bancari

I mutamenti istituzionali e normativi del sistema bancario hanno spinto numerosi autori a studiare la relazione esistente tra tali cambiamenti e i livelli di efficienza delle banche.

Di seguito sono riportati gli obiettivi, le metodologie e i risultati ottenuti da una parte degli studi effettuati sull'efficienza dei sistemi bancari. In questa breve rassegna particolare attenzione è attribuita al "caso italiano", e tra questi studi, si è concentrata l'attenzione su

quelli che hanno applicato metodi di tipo parametrico (cfr. § 1.5.2). La presentazione dei diversi saggi segue una classificazione che tiene conto della metodologia, dell'approccio relativo alle variabili e, per quel che riguarda le metodologie parametriche, alla forma funzionale della frontiera. In base a tali criteri, si distinguono, in prima battuta, i lavori che implementano l'approccio delle frontiere stocastiche e quelli che utilizzano la DEA, entrambi a loro volta classificati in studi che utilizzano l'approccio di intermediazione e quelli che utilizzano varianti dello stesso oppure combinazioni con l'approccio di produzione. Infine, nell'ambito degli studi "parametrici", si distinguono quelli che considerano una forma funzionale translogaritmica da quelli che ne considerano altre<sup>11</sup>.

### **1.5.1 Gli studi basati su metodologie parametriche: analisi dei sistemi bancari internazionali**

Tra gli autori che utilizzano le frontiere stocastiche, l'approccio di intermediazione e una frontiera translogartimica, Lang e Welzel (1996) stimano una frontiera di costo per le BCC tedesche nel periodo compreso tra il 1989 e il 1992, al fine di analizzare i processi di concentrazione che hanno caratterizzato le maggiori economie europee fin dai primi anni Novanta. In particolare, il saggio focalizza l'attenzione su alcuni aspetti, il primo dei quali risulta essere la convenienza a creare grandi gruppi bancari allo scopo di sfruttare eventuali economie di scala. Inoltre, gli autori testano l'ipotesi di esistenza di risparmi di costo dovuti alla produzione congiunta di più servizi, verificando, quindi, la presenza di economie di scopo. Gli autori propongono la stima di una funzione di costo di tipo translogaritmica multi-output visto che l'ammontare totale dei crediti viene suddiviso in prestiti verso le banche e verso altri clienti, a loro volta suddivisi tra prestiti e breve e a lungo termine. Per quel che riguarda gli input, vengono introdotti nel modello i prezzi del lavoro, del capitale e della raccolta. Da un punto di vista delle assunzioni delle componenti d'errore, l'approccio seguito è quello "tradizionale", secondo il quale le due componenti si distribuiscono come una variabile casuale Normale e una semi-Normale, considerate indipendenti. I risultati mostrano moderate economie di scala per tutte le classi del campione e, pertanto, si registrano progressi tecnici dovuti a risparmi di costo derivanti dalle operazioni di concentrazione. Risulta, inoltre,

---

<sup>11</sup> Una funzione di tipo translog rappresenta una generalizzazione di una Cobb-Douglas, rispetto alla quale contiene le stesse variabili al quadrato e le interazioni tra di esse. Esiste un'altra forma denominata Fourier-flexible form che rappresenta un'ulteriore estensione della translog. Essa contiene termini di seno e coseno delle variabili di base.

evidente l'esistenza di economie di scopo che supportano l'orientamento verso la banca universale. Mester (1996) propone uno studio di efficienza delle banche operanti nel Terzo Distretto della *Federal Reserve*, evidenziando l'inefficienza di queste banche nell'utilizzo dei loro input. L'autore ha condotto l'analisi per gli anni 1991-1992 su dati provenienti dal *Consolidated Reports of Condition and Income Banks*. Un carattere distintivo di questo studio è che, facendo esplicito riferimento all'approccio di intermediazione, gli output considerati sono rappresentati dai mutui immobiliari, dai prestiti commerciali ed industriali e dai prestiti ai privati. Nell'ambito della stessa categoria di lavori, Molyneux (2003) propone l'analisi dell'efficienza di costo relativamente al sistema bancario dei paesi membri dell'UE-15 per il periodo 1992-2000. I dati sono estratti dal database BankScope. Per la selezione delle variabili da inserire nelle stime, l'autore segue l'approccio di intermediazione introducendo, però, nel modello anche il capitale proprio e un *trend*. A giustificazione di tale scelta, si evidenzia che la derivata parziale rispetto al tempo può essere vista come il cambiamento tecnico nella frontiera di costo. I risultati suggeriscono che le banche più grandi e le banche commerciali mostrano le minori riduzioni di costo, ma, allo stesso tempo, i maggiori guadagni in termini di cambiamento tecnico.

Behr e Tente (2008) propongono lo studio dell'efficienza di costo per un campione di 56 banche tedesche nel 2005, dati Bankscope. Gli autori introducono una novità relativamente alla scelta delle distribuzioni marginali della componente erratica del modello, ipotizzando due diversi scenari. In primo luogo, si considera il caso tradizionale, che prevede una distribuzione Normale per l'errore casuale e una distribuzione Normale troncata per la componente di inefficienza. Inoltre, viene proposto il caso in cui si assume che l'errore casuale si distribuisca come una variabile casuale Normale e che l'inefficienza osservi una regola di tipo Esponenziale. Per entrambi gli scenari, vengono confrontati due metodi di stima, quello della massima verosimiglianza e il metodo dei momenti. A tal proposito, i risultati mostrano la superiorità del primo per i grandi campioni, mentre, se l'analisi è condotta su campioni di piccole dimensioni, è preferibile il metodo dei momenti. Dongili, Rossi e Zago (2008) valutano l'efficienza di costo e di profitto dei maggiori sistemi bancari europei. Gli autori si concentrano sul periodo 1995-2005, un arco temporale caratterizzato dal recepimento delle principali direttive sull'armonizzazione dei sistemi bancari all'interno dell'Unione Europea. La variante riguarda l'utilizzo di due diverse forme funzionali per la frontiera (translog e Fourier-flexible) e il confronto tra le due. L'assunzione sulle distribuzioni della componente erratica composta risulta essere identica a quella presente in letteratura. Altro aspetto su cui gli autori si soffermano riguarda la scelta tra quale tipo di frontiera

studiare e a quale obiettivo di massimizzazione rispondere. Essi sostengono che la scarsa correlazione tra le due misure di efficienza (di costo e di profitto) testimonia quanto sarebbe parziale e incompleta l'analisi basata solo sull'efficienza di costo o solo sull'efficienza di profitto. I risultati mostrano che, in media, i sistemi bancari europei sono caratterizzati da un'elevata efficienza di costo. Per quel che riguarda i profitti, viene evidenziato che il livello di efficienza è minore e decrescente nel tempo. Inoltre, mentre l'efficienza di costo risulta meno variabile nel tempo, quella di profitto mostra una crescente variabilità. Le banche italiane si collocano in posizione intermedia, in quanto mostrano buoni livelli di efficienza di costo, ma minore capacità di ottenere guadagni in termini di profitto. Per spiegare tale risultato, si può considerare la posizione di Turati (2008), il quale mette in evidenza che, da un lato, consumatori e imprese ritengono che i mercati bancari siano gravati da deficit di concorrenzialità che si riflette in prezzi dei servizi elevati per una discreta qualità. Dall'altro lato, le banche sottolineano che le innovazioni normative in merito alla deregolamentazione del mercato bancario comportano un aumento della concorrenza. L'autore propone, quindi, una valutazione del grado di concorrenza del mercato bancario considerandola in un'accezione dinamica, per cui in mercati concorrenziali le imprese meno efficienti dovrebbero essere espulse dall'industria, mentre quelle più efficienti dovrebbero essere quelle più redditizie con quote di mercato più alte. Il test implementato consiste nel verificare se esiste una relazione positiva tra redditività ed efficienza, concludendo, in tal caso, che il settore è caratterizzato da un alto grado di concorrenzialità. Se i dati mostrano, invece, assenza di qualsiasi tipo di relazione, allora si potrebbe concludere che non vi è concorrenza. Lo studio riguarda i sistemi bancari di Italia, Francia, Germania, Spagna e Regno Unito nel periodo 1992-1999 su dati estratti dal database Bankscope, confrontando l'ipotesi di una struttura efficiente con un'ipotesi alternativa in base alla quale esiste un potere di mercato. Successivamente, l'autore propone un modello per spiegare la relazione tra redditività ed efficienza. I risultati sostengono l'ipotesi di mercati non concorrenziali e, pertanto si accetta la tesi secondo cui le operazioni di consolidamento comportino un aumento del potere contrattuale delle banche, sfavorevole per la clientela.

Si differenziano per la forma funzionale due lavori, i quali scelgono di implementare un modello in cui si utilizza una forma Fourier-flexible. Si tratta degli studi proposti da Altunbas, Goddard e Molyneux (1999) e da Altunbaş, Gardener, Molyneux e Moore (2001). I primi stimano la frontiera di efficienza di costo per i paesi membri dell'EU-15 nel periodo 1989-1996. L'analisi è condotta su un campione di 3779 banche estratto dal database Bankscope. I risultati mostrano che il tasso di riduzione dei costi dovuti a variazioni dei livelli

di efficienza è aumentato nel periodo considerato. Inoltre, è emerso che le banche più grandi hanno ottenuto maggiori benefici rispetto a quelli ottenuti dalle banche minori. Altunbaş, Gardener, Molyneux e Moore (2001) estendono l'analisi fatta nel lavoro di Altunbaş, Goddard e Molyneux (1999) per stimare gli *score* di efficienza di costo relativi al sistema bancario di alcuni paesi europei nel periodo 1989-1997<sup>12</sup>. Come in Lucchetti, Papi e Zazzaro (1997), viene utilizzata la tecnica di Jondrow *et al.* (1982) per separare la componente erratica casuale da quella che misura l'inefficienza, calcolando l'aspettativa condizionata di quest'ultima dato l'errore casuale. I risultati mostrano che alcuni paesi come Finlandia, Lussemburgo, Paesi Bassi e Svezia hanno registrato incrementi nei livelli di efficienza del sistema bancario nel corso del tempo. Inoltre, si mostra che i settori bancari più efficienti sono quelli austriaco, danese, tedesco e italiano. Infine, l'impatto del cambiamento tecnologico nel ridurre i costi aumenta all'aumentare della dimensione delle banche.

Al fine di stimare frontiere di costo e di profitto, Huizinga, Nelissen e Vennet (2001) analizzano gli effetti di 52 acquisizioni sull'efficienza delle maggiori banche europee coinvolte in tali operazioni, nel periodo compreso tra il 1994 e il 1998, ovvero il periodo immediatamente precedente all'introduzione dell'euro. Il campione dei dati è stato estratto dal dataset Bankscope. Una delle particolarità presenti nello studio risulta essere la trasformazione della variabile "Risultato d'esercizio" per la stima della frontiera di profitto<sup>13</sup>. Tra le ipotesi che si assumono nella formulazione di tale funzione emerge quella secondo la quale il profitto debba essere una variabile non-negativa. E' inverosimile, però, che un campione di banche osservate in un periodo di medio o lungo termine possa registrare solo risultati di esercizio positivi. Se si utilizzano funzioni Cobb-Douglas o translog che linearizzate includono le trasformazioni logaritmiche delle variabili, è necessario introdurre delle trasformazioni affinché si possa lavorare su valori positivi della variabile dipendente (Kumbhakar e Lovell, 2000). Gli autori trovano evidenza di sostanziali economie di scala non sfruttate e di elevate *X-inefficiencies* nelle banche europee (Leibenstein, 1966). L'analisi ha, inoltre, evidenziato che l'efficienza di costo delle banche appartenenti al campione esaminato è positivamente influenzata dalle operazioni di consolidamento. Dall'altro lato, si osservano miglioramenti solo marginali dei livelli di efficienza di profitto.

---

<sup>12</sup> Austria, Belgio, Danimarca, Finlandia, Francia, Germania, Grecia, Irlanda, Italia, Lussemburgo, Paesi Bassi, Portogallo, Regno Unito, Spagna, Svezia.

<sup>13</sup> Gli autori propongono la seguente trasformazione:  $\ln\left[\left(\frac{\pi}{\text{capitale}}\right) + \left|\frac{\pi}{\text{capitale}}\right|^{\min} + 1\right]$ , in cui il secondo addendo all'interno dell'argomento del logaritmo è il valore assoluto del valore minimo assunto dal rapporto utile (o perdita) / capitale. Si garantisce, in questo modo, la non-negatività della variabile dipendente, in quanto, nel momento in cui il rapporto utile/capitale assume il suo valore minimo, la variabile diventa  $\ln(1) = 0$ .

Fiordelisi, Marques-Ibanez e Molyneux (2010) si propongono di analizzare la relazione tra efficienza, capitale e rischio delle banche europee. Essi stimano frontiere di costo, di profitto e di ricavo per il periodo compreso tra il 1995 e il 2007 e, successivamente, utilizzano il test di Granger per verificare la relazione al centro del loro obiettivo di ricerca. Il lavoro presenta l'analisi di un campione estratto dalla banca dati Bankscope e costituito da un gran numero di banche commerciali dei paesi membri dell'UE-27 escludendo, però, il Lussemburgo. I risultati mostrano che riduzioni dei costi e incrementi nell'efficienza dei ricavi comportano un aumento dei futuri rischi per le banche, mentre miglioramenti nell'efficienza contribuiscono a mantenere e consolidare i livelli di capitale.

Utilizzando un altro metodo parametrico di cui si è parlato in precedenza, il Thick Frontier Approach, Humphrey e Pulley (1997) stimano il grado di efficienza generale del sistema bancario americano nel periodo 1977-1988. Dai risultati emerge che circa l'80% di differenza nella performance esibito dalle banche del campione è riferibile a differenze in termini di efficienza tecnologica. Viene stimata un'equazione di profitto utilizzando una funzione di tipo translog. L'effetto della deregolamentazione avvenuta nel settore bancario americano nel periodo esaminato si riflette, secondo gli autori, su una misura standard della profittabilità delle banche utilizzata come variabile dipendente. Si tratta del ROA (*Return on Assets*), definito come il rapporto tra l'utile netto (ovvero il risultato economico prima degli accantonamenti per rischi di perdite su crediti, delle tasse e dei valori straordinari) e il totale attivo. Nella frontiera di produzione vengono inseriti, come variabili di output, i depositi e i ricavi da servizi di prestiti ed intermediazione, una misura di "altre attività", come residuo rispetto alle variabili già elencate, e il prezzo degli input (costo del lavoro, costo dei depositi, costo del capitale fisico). Infine, si considera un vettore di variabili caratteristiche dell'ambiente di riferimento (la crescita del Pil pro-capite, lo spread tra prestiti e depositi, una misura della produttività degli impiegati, ovvero il rapporto tra il numero di dipendenti e il numero di sportelli).

### **1.5.2 Gli studi basati su metodologie parametriche: analisi del sistema bancario italiano**

Per quanto riguarda gli studi che hanno analizzato il caso del sistema bancario italiano, è utile ricordare innanzitutto quello di Girardone, Molyneux e Gardener (2004) in cui si trova l'evoluzione dell'efficienza del sistema bancario italiano per il periodo 1993-1996, utilizzando dati Bilbank. Dopo aver stimato l'inefficienza, gli autori calcolano la somma delle elasticità di costo di ciascuna banca, al fine di verificare l'esistenza di economie o

diseconomie di scala. Si consideri che nella funzione di costo, oltre agli output (prestiti ai clienti e altri ricavi) e ai prezzi degli input (costo del personale, interessi passivi, altre spese), vengono inseriti anche il rischio e la qualità dell'output. In questo articolo, l'analisi viene condotta differenziando per area geografica e per tipo di banca (Spa, BCC e banche popolari), motivando tale scelta con la considerazione dei diversi obiettivi manageriali (ad esempio, minimizzazione dei costi oppure scopi mutualistici). Dai risultati emerge che l'inefficienza tende a ridursi nel corso del tempo. Tuttavia, i livelli più bassi di efficienza si hanno per quelle banche la cui sede centrale si trova nel Mezzogiorno. Inoltre, i risultati più soddisfacenti provengono dalle BCC e dalle popolari. Dall'analisi emerge un ulteriore aspetto legato al fatto che le banche più grandi non riescono a realizzare economie di scala, esibendo rendimenti costanti. Ovviamente, questo risulta essere in contrasto con i risultati attesi in merito alle operazioni di concentrazione, sebbene gli anni considerati nello studio comprendano il periodo in cui erano appena avviati i processi di ristrutturazione istituzionale del sistema bancario. Infine, viene proposta una regressione logistica per controllare quali variabili possano spiegare l'inefficienza. Dai risultati emerge che essa è correlata negativamente con i livelli di capitale e positivamente con i *non-performing loans*, il che suggerisce una maggiore attenzione per le attività di *monitoring* e di *screening* dei progetti di investimento. Emerge, inoltre, che le banche inefficienti tendono ad avere maggiore orientamento alle famiglie (sia in termini di prestiti che di depositi), maggiori margini di interesse e un numero più elevato di filiali rispetto alle banche più efficienti. Le banche quotate risultano essere più efficienti, quelle del Sud più inefficienti. Al fine di studiare come i cambiamenti strutturali del sistema bancario italiano abbiano esercitato un impatto sull'efficienza di costo delle banche, Fontani e Vitali (2007) concentrano la loro analisi sui principali gruppi bancari italiani dal 1993 al 2004. Secondo gli autori, l'obiettivo della deregolamentazione, ossia quello di consolidare il sistema e favorire i processi di privatizzazione, è stato solo in parte raggiunto dal momento che fra gli azionisti di molti istituti di credito sono presenti ancora le fondazioni bancarie e soggetti ibridi come le banche popolari (ciò fa pensare ad un processo di privatizzazione solo formale). L'analisi si sofferma in particolare sulle numerose operazioni di concentrazione che, secondo gli autori, non necessariamente comportano guadagni di efficienza e riduzione della concorrenza. I risultati cui giungono fanno emergere la presenza di progresso tecnico e di economie di scala.

Nell'ambito dei lavori che utilizzano le frontiere stocastiche come metodologia per stimare l'inefficienza, funzioni translog e approcci ibridi nella scelta di output e input, Giannola, Lopes, Ricci e Scarfiglieri (1997) analizzano l'impatto del consolidamento e della

liberalizzazione del sistema bancario italiano sui livelli di efficienza di costo. L'analisi è stata condotta su dati ABI per il 1994. In tal caso, è stato fatto ricorso ad un approccio ibrido derivante dall'integrazione tra quello di produzione e quello di intermediazione. Sono stati considerati tre tipi di output, ovvero, prestiti ai clienti, depositi, prestiti a banche e titoli. Per quel che riguarda i prezzi dei fattori produttivi, sono stati introdotti il prezzo del lavoro, del capitale fisico e dei fondi raccolti. Il costo totale utilizzato come variabile dipendente è derivante dalla somma tra il costo totale del lavoro, del capitale fisico e quello derivante dalla raccolta dei fondi. Inoltre, seguendo le definizioni della Banca d'Italia si è tenuto conto di una differenziazione di tipo dimensionale individuando banche maggiori, grandi, medie, piccole e minori. Dai risultati emerge che le banche minori e quelle piccole presentano un valore dell'inefficienza prossimo a quello nazionale, con un contributo determinante delle banche del Nord-Est. Le banche grandi presentano un indice di inefficienza superiore alla media nazionale. Quindi, almeno sotto il profilo dimensionale, le evidenze empiriche mostrano risultati diversi da quelli attesi.

Nello studio proposto da Lucchetti, Papi e Zazzaro (1997) viene stimata l'efficienza bancaria delle regioni italiane dal 1982 al 1994. Secondo gli autori, l'efficienza delle banche è una misura che non dovrebbe soffrire del problema di simultaneità in quanto il livello di sviluppo di un'area influenza i costi dell'attività bancaria, ma ha un'influenza minore sulla funzione di produzione dell'attività bancaria. Inoltre, sembra che l'efficienza microeconomica sia adatta a cogliere la funzione allocativa delle banche, considerando che la capacità di utilizzare bene la tecnologia e di combinare gli input in maniera ottimale possano essere considerate come due aspetti fondamentali per l'allocazione delle risorse. Tuttavia, secondo gli autori, se si vuole studiare la relazione tra sistema bancario e crescita economica di un'area, ci si dovrebbe concentrare sull'efficienza allocativa macroeconomica, intesa come la capacità di selezionare i progetti con un più elevato impatto sullo sviluppo. Viene utilizzata la tecnica di Jondrow *et al.* (1982) per separare la componente erratica (calcolando l'aspettativa condizionata della componente di inefficienza dato l'errore casuale). Per quel che riguarda le variabili di output, gli autori utilizzano il totale degli impieghi (da cui sono state sottratte le sofferenze, per cogliere una misura di qualità dell'output) e dei depositi e i ricavi da servizi (quindi sia variabili di stock che di flusso)<sup>14</sup>. Allo stesso tempo, i due input considerati risultano il capitale umano e quello fisico. Successivamente, si è passati ad una misura regionale di efficienza facendo riferimento alla quota territoriale di sportelli per individuare il

---

<sup>14</sup> Considerare i depositi negli output è tipico dell'approccio a valore aggiunto (Dongili, Rossi e Zago 2008).

livello di efficienza raggiunto da ciascuna regione. In tal modo, introducono due forti assunzioni, la prima è che la funzione di produzione della banca sia la stessa in ogni regione; la seconda consiste nel fatto che l'efficienza degli sportelli nella stessa regione sia indipendente dalla localizzazione dello sportello stesso, giustificando, peraltro, tale scelta con una sorta di principio di invarianza del livello di inefficienza di diversi sportelli all'interno di una stessa regione. Pertanto, l'inefficienza del sistema bancario di ciascuna regione è data dalla somma del prodotto tra la stima dell'inefficienza di ogni singola banca e la quota di sportelli di quella banca in quella determinata regione. I risultati suggeriscono che i sistemi bancari più efficienti sono soprattutto quello emiliano e quello veneto. Nel Nord-Ovest, il sistema più efficiente risulta essere quello lombardo. Inoltre, negli anni Ottanta, il Lazio è stato la regione con il sistema bancario più inefficiente, primato che, invece, negli anni '90 è passato alla Campania. Per le regioni del Mezzogiorno, si osserva, altresì, una riduzione del livello di efficienza nel tempo. Infine, gli autori presentano un modello con dati panel in cui si regredisce il valore aggiunto pro capite (per regione) su una serie di quantità di tipo finanziario tra cui la stessa variabile ritardata, il livello di inefficienza regionale, il rapporto tra gli impieghi erogati alla clientela della regione e il Pil regionale, la quota degli impieghi delle aziende di credito erogati al settore privato sul totale impieghi, la quota degli impieghi di BCC e POP sul totale impieghi. Sono state inserite anche variabili di controllo non finanziarie relative al capitale umano (il rapporto tra iscritti alle scuole medie superiori e iscritti alle scuole medie inferiori), ai costi di trasporto (il rapporto tra chilometri di autostrade, strade statali e strade ferrate sui chilometri quadrati della regione), all'efficienza del sistema legale<sup>15</sup>. I risultati dimostrano l'esistenza di un effetto dell'efficienza bancaria sulla crescita reale<sup>16</sup>. Un interessante lavoro è quello proposto da Giordano e Lopes (2006) in cui emerge l'attenzione a come la qualità degli impieghi influenzi l'efficienza operativa. In presenza di asimmetrie informative, si rende necessaria un'attenta valutazione sui rendimenti futuri e sul rischio di un dato progetto di investimento, anche in base ad alcune attività finalizzate a recepire parte delle informazioni rilevanti. Si tratta di attività di *screening*, di *monitoring* e di *auditing*. Banche che riducono il livello di screening sugli impieghi, per esempio, possono essere valutate come banche efficienti perché sosterebbero costi minori rispetto a quelle banche che invece investono molto in queste attività di tipo informativo. Tuttavia, il tipo di output delle due

---

<sup>15</sup> L'indicatore utilizzato è dato dal rapporto tra fallimenti chiusi e fallimenti dichiarati ponderato per la loro dimensione media. Ovvero, il peso è calcolato come la radice quadrata del rapporto tra l'attivo medio di chiusura e i fallimenti chiusi.

<sup>16</sup> I risultati confermano le attese degli autori sui segni degli stimatori, positivi per tutte le variabili considerate, ad eccezione dell'indicatore di inefficienza regionale.

banche potrebbe non essere omogeneo. La questione su cui Giordano e Lopes (2006) hanno indagato riguarda l'endogeneità dei *non-performing loans*. Si sono chiesti se la qualità degli impieghi sia il frutto dei processi gestionali interni, oppure se sia un dato esogeno rispetto alla banca e, quindi in grado di condizionarne l'efficienza. Gli autori effettuano il test di causalità di Granger per scegliere se inserire o meno una proxy della qualità degli impieghi solo come output<sup>17</sup>. I dati utilizzati provengono da informazioni di bilancio contenute nel dataset Bilbank per il periodo 1993-2003. In aggiunta alle variabili tipiche dell'approccio di intermediazione, gli autori considerano una misura dell'intensità dei crediti (crediti verso clienti/totale attivo), una misura dell'avversione al rischio (partecipazioni/totale attivo) e una misura della qualità degli impieghi (sofferenze/totale crediti). Le analisi empiriche hanno condotto a risultati differenziati per tipologia di banche. Ad esempio, le BCC presentano minori (maggiori) livelli di inefficienza (efficienza) di costo, mentre, al contrario, registrano più elevati (più bassi) livelli di inefficienza (efficienza) di profitto. Inoltre, le banche popolari registrano riduzioni dell'inefficienza sia di costo che di profitto. Pertanto, emerge che il processo di consolidamento del sistema bancario italiano motivato dalla volontà di raggiungere più alti livelli di efficienza e caratterizzato dall'adozione diffusa della forma di società per azioni, non sembra trovare fondamento nelle evidenze empiriche. Per quanto riguarda i risultati del test di Granger, emerge che la qualità degli impieghi risulta essere endogena, in favore dell'approccio cosiddetto di *bad management*, secondo cui una cattiva gestione interna determinerebbe la selezione di *non-performing loans* (Berger e De Young, 1997)<sup>18</sup>. Nel loro successivo studio, Giordano e Lopes (2008) si concentrano sui processi di divergenza tra Nord e Sud in riferimento ai livelli di efficienza di costo e di profitto dal 1993 al 2003, utilizzando un campione di 550 banche. I risultati mostrano che vi è un generale declino di efficienza il quale investe soprattutto le società per azioni, protagoniste dei processi di consolidamento. In questo contesto, le banche Spa del Sud conseguono risultati meno soddisfacenti rispetto a quelle del Nord. L'approccio utilizzato per la scelta delle variabili è un approccio ibrido, perché da un lato deriva dall'integrazione dell'approccio di produzione e di

---

<sup>17</sup> Se il test andasse in questa direzione, l'esito sarebbe in favore dell'endogeneità dei *non-performing loans*. Pertanto, la qualità degli impieghi non potrebbe essere utilizzata come variabile esplicativa dell'efficienza delle banche.

<sup>18</sup> I filoni di ricerca che partono da Berger e De Young (1997) riguardano la relazione tra l'efficienza delle banche e la qualità dei crediti. Gli autori evidenziano diverse relazioni tra le due variabili. La prima è detta "*bad management*", secondo la quale vi è endogeneità fra qualità ed efficienza, ovvero una sorta di simultaneità inversa, in quanto i manager potrebbero scegliere di risparmiare anche sui costi relativi alle attività di screening selezionando impieghi di bassa qualità. La seconda ipotesi, detta "*bad luck*", afferma esogeneità fra qualità ed efficienza, in particolare, la cattiva qualità degli impieghi influenzerebbe in maniera esogena i livelli di efficienza delle banche.

quello di intermediazione, dall'altro perché vengono utilizzate variabili sia di flusso che di stock. In particolare, i prodotti della banca sono stati individuati nel totale di impieghi e depositi e nei ricavi da servizi. L'utilizzo di variabili stock (impieghi e depositi) come proxy del valore della produzione in aggiunta a variabili flusso (ricavi da servizi) è motivato dal fatto che le componenti di stato patrimoniale comportano una continua produzione di servizi così da poter essere considerate come buone approssimazioni della produzione bancaria. Il modello prevede altre variabili come quella che misura l'intensità di credito, data dal rapporto tra impieghi alla clientela e attività fruttifere, e quella relativa al grado di avversione al rischio calcolata come il rapporto tra capitale proprio e totale attivo. Secondo gli autori, il rapporto tra sofferenze e totale dei prestiti costituisce la migliore proxy per valutare quanta parte di risorse della banca viene effettivamente impiegata nell'attività di monitoraggio.

Battaglia, Farina, Fiordelisi e Ricci (2010) si interrogano su quali sono i fattori individuali e di contesto che influenzano l'efficienza delle BCC italiane dal 2000 al 2005. Utilizzano la specificazione di frontiera stocastica proposta da Battese e Coelli (1995) su dati forniti dalla Federazione Italiana delle Banche di Credito Cooperativo – Casse Rurali ed Artigiane (Federcasse). Gli autori evidenziano che le banche del nord raggiungono livelli di efficienza di costo più alti rispetto a quelle del sud. Tuttavia, la situazione si inverte quando si parla di efficienza di profitto. Infatti, gli istituti di credito cooperativo con sede legale al Sud sembrano, sotto questo punto di vista, maggiormente efficienti. Risultati interessanti emergono quando si tratta dei fattori che influenzano l'inefficienza. Innanzitutto, l'incidenza dei *non performing loans* e del PIL pro-capite sembra non avere effetti significativi. Inoltre, la densità della popolazione e l'indice di concentrazione territoriale (calcolato come rapporto tra la popolazione residente nelle maggiori città della regione in cui la banca ha sede legale e quella residente nel comune) hanno un impatto significativamente negativo sui livelli di efficienza di costo. Effetti positivi, invece, sono determinati dall'indice di imprenditorialità registrato nel comune in cui la banca ha sede legale e dal numero di sportelli con operatività locale dei vari istituti di credito.

### **1.5.3 Gli studi basati su metodologie non parametriche**

Come già detto in precedenza, la metodologia non parametrica più utilizzata per stimare l'inefficienza di vari settori è rappresentata dalla Data Envelopment Analysis (DEA). Anche per il settore bancario e creditizio esiste una vasta letteratura che utilizza tale metodologia al fine di studiarne l'efficienza. Di seguito sono riportati alcuni studi che

utilizzano l'approccio di intermediazione nella scelta delle variabili di frontiera e, successivamente, alcuni lavori che si avvalgono di varianti dello stesso approccio di intermediazione.

Casu e Molyneux (2003), ad esempio, confrontano l'efficienza dei sistemi bancari dei cinque paesi europei con maggiore peso nell'UE<sup>19</sup>, ovvero Francia, Germania, Italia, Spagna e Regno Unito. Si pone l'enfasi sui cambiamenti dei sistemi bancari avvenuti a livello internazionale relativamente alla deregolamentazione e alle operazioni di consolidamento che hanno incentivato le banche a riorganizzarsi al fine di raggiungere livelli più alti di efficienza in un contesto caratterizzato da una crescente concorrenza proveniente anche dall'estero. L'analisi è condotta per il periodo 1993-1997 su un campione di 750 banche (ovvero, le 150 banche più grandi per ogni paese considerato) estratto da Bankscope. Dai risultati emerge che, per tutti i paesi considerati, si registra una tendenza al miglioramento di efficienza media, ad eccezione di quanto accade in Italia, la quale registra livelli di efficienza più bassi rispetto agli altri e con tendenza al peggioramento. Gli autori concludono sostenendo che le differenze negli *score* di efficienza tra i vari paesi sono da addurre a fattori specifici propri di ciascun sistema economico considerato. L'analisi dell'efficienza del sistema bancario di alcuni tra i maggiori paesi europei<sup>20</sup> diventa l'obiettivo di ricerca anche per Casu e Girardone (2006) che stimano frontiere di ricavo per il periodo 1997-2003. Gli autori si concentrano sugli effetti della deregolamentazione dei servizi finanziari e sulla concentrazione del sistema bancario avvenute nell'Unione Europea negli ultimi tempi. Lo scopo è di costruire la statistica  $H$  di Panzar e Rosse (1987) la quale, a seconda dei valori che assume, consente di individuare il grado di concorrenza di un dato settore (concorrenza perfetta, monopolio, oligopolio)<sup>21</sup>. I dati sono estratti dal dataset Bankscope. La variabile output è rappresentata dal rapporto totale ricavi/totale attivo. In aggiunta alle variabili tipiche per un'analisi di frontiera, gli autori introducono altre variabili esplicative che caratterizzano i sistemi bancari, ovvero, il rapporto Capitale proprio/Totale Attivo, il rapporto tra prestiti totali e totale attivo, il rapporto liquidità su totale depositi. Pertanto, si è scelto di utilizzare una variante dell'approccio di intermediazione. I risultati suggeriscono che un più alto grado di concentrazione non è necessariamente sintomo di una più bassa competitività del settore. Inoltre, emerge che i

---

<sup>19</sup> La somma degli *assets* dei cinque paesi rappresenta l'80% del Totale Attivo dell'UE-15 (Banca d'Italia, Report Annuale 1998).

<sup>20</sup> Sono 14 i paesi inseriti in quest'analisi, ovvero, Austria, Belgio, Danimarca, Finlandia, Francia, Germania, Grecia, Irlanda, Italia, Lussemburgo, Paesi Bassi, Portogallo, Spagna e Svezia.

<sup>21</sup> La statistica  $H$  è definita come variazione incrementale dei profitti, a fronte di una variazione unitaria del prezzo degli input. Essa è uguale a zero o negativa quando la struttura è monopolistica, è uguale a 1 se c'è concorrenza perfetta, se  $0 < H < 1$ , vi è concorrenza monopolistica.

sistemi bancari più efficienti, sono anche quelli più competitivi. Tuttavia, questa relazione potrebbe essere poco chiara, ovvero, un aumento di competizione forza le banche a essere più efficienti e, allo stesso tempo, aumenti di efficienza non sembrano promuovere una maggiore competizione. Infine, l'Italia, insieme a Francia, Germania e Lussemburgo, mostra i livelli più bassi di efficienza, mentre, Finlandia, Grecia e Svezia si caratterizzano per sistemi bancari molto efficienti. Tra i numerosi studi che effettuano analisi sulla relazione tra l'efficienza delle banche e il potere di mercato, emerge quello proposto da Chortareas, Garza-Garcia e Girardone (2009) che pongono la loro attenzione sulle operazioni di concentrazione avvenute negli anni Novanta in alcuni paesi dell'America Latina<sup>22</sup>. Differentemente da quanto si è verificato in Italia, le operazioni di concentrazione sono state guidate, in questi paesi, dall'acquisizione delle banche locali da parte di istituzioni straniere. Sorge, pertanto, il rischio che le banche coinvolte attuino comportamenti collusivi i quali, ledendo le regolari pratiche competitive, comportano condizioni svantaggiose per la clientela. Viene utilizzato un campione di 2500 banche estratto dal database Bankscope, nel periodo 1997-2005. Successivamente, gli autori presentano un modello con il quale tentano di esaminare l'effetto dell'efficienza e della concentrazione sul grado di profittabilità delle banche. A tal proposito, la variabile dipendente del modello è stata individuata nel ROA (*Return on Assets*), mentre il grado di concentrazione è espresso da una misura della struttura di mercato calcolata come la somma dei quadrati delle quote di mercato delle banche, ovvero dall'indice di Herfindahl. Dai risultati emerge, per la maggior parte dei paesi considerati, una relazione negativa tra concentrazione e profittabilità e, allo stesso tempo, si deduce che migliori livelli di efficienza comportano migliore profittabilità. Utilizzando la stessa metodologia, Casu e Girardone (2010) focalizzano l'attenzione sulla relazione tra la tendenza all'integrazione dei sistemi bancari e la convergenza in termini di efficienza per le banche aventi sede legale nei paesi membri dell'UE-15 nel periodo 1997-2003. L'analisi è condotta su un campione di 11000 banche estratto dal database Bankscope, ottenendo un panel non bilanciato e giustificando tale scelta mediante la considerazione delle fusioni e acquisizioni avvenute nel periodo in esame. Dai risultati emerge che i membri dell'EU-15 tendono verso una data media del livello di efficienza di costo e, allo stesso tempo, sembra che non ci siano né miglioramenti di efficienza globali né la tendenza verso la *best practice*. Infine, Barra, Destefanis e Lubrano Lavadera (2011) effettuano un'analisi di efficienza focalizzando l'attenzione sulle BCC italiane. Gli autori enfatizzano la relazione tra l'efficienza e alcuni fattori di contesto, la

---

<sup>22</sup> Argentina, Brasile, Cile, Colombia, Costa Rica, Paraguay, Perù, Uruguay e Venezuela.

performance delle banche e la qualità dei crediti rispetto alla quale forniscono evidenze in favore dell'ipotesi di *bad luck*, secondo la quale la qualità dei crediti determina i livelli di efficienza delle banche e non il contrario (Berger e De Young, 1997).

#### **1.5.4 Gli studi basati sul confronto fra metodi parametrici e non parametrici**

In questo tipo di letteratura è frequente trovare studi in cui si effettuano confronti fra le varie metodologie. Per esempio, Weill (2004) propone un confronto tra diverse metodologie per l'analisi degli *score* di efficienza di costo di alcuni paesi europei (Francia, Germania, Italia, Spagna e Svizzera) nel periodo 1992-1998 su dati Bankscope, utilizzando un approccio di intermediazione relativamente alla scelta degli output e dei prezzi dei fattori produttivi. I risultati mostrano che il valore dei livelli di efficienza derivanti dall'applicazione della tecnica delle frontiere stocastiche sono maggiori rispetto ai valori stimati con le altre due metodologie. Successivamente, l'analisi procede con la verifica dei livelli di efficienza distinguendo le banche per categoria di servizi offerti. Sono stati, infatti, individuate tre tipologie di banche, ovvero le banche commerciali, quelle cooperative e le casse di risparmio. I risultati evidenziano livelli di efficienza più alti per le banche cooperative e le casse di risparmio. Questo risultato, che per l'Italia è simile a quello ottenuto in Giordano e Lopes (2008), potrebbe essere contro intuitivo, in quanto ci si aspetta che le banche private siano più efficienti rispetto a quelle pubbliche o a quelle con scopi mutualistici. A tal proposito, si consideri che l'obiettivo dei mutamenti normativi ed organizzativi degli anni Novanta era quello di agevolare e mantenere una situazione in cui emergesse un processo di privatizzazione in maniera da garantire maggiore efficienza. Da quest'analisi emerge che nei lavori basati sulle misure di efficienza, è necessario che vengano utilizzate più tecniche per validare i risultati. Tuttavia, in questo lavoro, i valori stimati con le diverse metodologie, di cui due parametriche e una non parametrica, sono tali da non poterne fare il confronto e, soprattutto, non garantiscono, per uno stesso paese, le stesse conclusioni sui livelli di efficienza raggiunti dal sistema bancario nazionale. Anche Beccalli, Casu e Girardone (2006), attraverso il confronto tra SFA e DEA, hanno analizzato la relazione tra l'efficienza del sistema bancario di Francia, Germania, Italia, Spagna e Regno Unito e il rendimento dei titoli azionari quotati nelle Borse dei paesi considerati, per il biennio 1999-2000. In riferimento all'utilizzo delle frontiere stocastiche, viene stimata una frontiera di costo di tipo translog utilizzando l'approccio di intermediazione per la scelta delle variabili. Anche in questo studio emerge che le stime effettuate con la metodologia delle frontiere stocastiche sono diverse

rispetto a quelle ottenute con la DEA, nella direzione secondo la quale il metodo non parametrico sottostima i livelli di efficienza. Un aspetto interessante del lavoro è il tentativo di studiare la relazione tra i livelli di efficienza e la *bank stock performance* rappresentata dai rendimenti dei titoli negoziati sui mercati borsistici dei paesi in esame. Gli autori propongono tre diversi modelli partendo da quello più ristretto che si propone di evidenziare la relazione tra la variazione di efficienza e la *stock performance* per poi considerare il modello in cui si aggiunge una proxy della dimensione (totale attivo) e una proxy del grado di rischio e profitto delle banche (ROE). Infine, il modello più esteso considera variabili dicotomiche per ciascuno dei paesi presenti nell'analisi. In particolare, la variabile dipendente è rappresentata dalla media dei prezzi dei titoli bancari quotati nel periodo compreso tra gennaio 1999 e giugno 2001. L'attenzione sembra essere focalizzata sulla variabile che coglie il cambiamento nei livelli di efficienza delle banche dei vari paesi tra il 1999 e il 2000. I dati di bilancio delle banche sono stati estratti dal dataset Bankscope, mentre quelli relativi ai prezzi giornalieri dei titoli da Datastream. I risultati mostrano una relazione positiva e significativa tra le due variabili in esame, e pertanto, si può concludere che i cambiamenti dei prezzi dei titoli bancari riflettono i cambiamenti nei livelli di efficienza di costo delle banche. Infine, Casu e Girardone (2009) propongono lo studio della relazione tra il grado di concorrenza e l'efficienza del settore bancario di Francia, Germania, Italia, Spagna e Regno Unito nel periodo 2000-2005. Gli autori stimano gli *score* di efficienza dell'intero campione composto da 2701 banche commerciali sia con la DEA sia con lo SFA. Successivamente, vengono proposti alcuni modelli con lo scopo di studiare il rapporto di causalità tra l'efficienza e una proxy di competitività individuata dall'indice di Lerner. Gli autori propongono il test di causalità di Granger trattando dati longitudinali mediante la stima di un pooled OLS, del modello a effetti fissi e di quello a effetti casuali. I risultati mostrano che un aumento dell'efficienza delle banche comporta una diminuzione dell'indice di Lerner e, quindi, una riduzione del potere monopolistico. Allo stesso tempo, il potere di monopolio può avere un effetto positivo sull'efficienza, se questo permette alle banche di operare a bassi costi.

**Tabella 1.7 Gli studi sull'efficienza dei sistemi bancari: un quadro sinottico**

<b>Autore/i</b>	<b>Tempo</b>	<b>Paese/i</b>	<b>Tecnica</b>	<b>Forma funzionale della frontiera</b>	<b>Funzione di distribuzione degli errori</b>	<b>Variabili utilizzate (Input, Output)</b>
Lang, Welzel (1996)	1989-1992	Germania (Bcc)	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Mester (1996)	1991-1992	Third Federal Reserve District (Usa)	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Giannola, Lopes, Ricci, Scarfiglieri (1997)	1994	Italia	SFA	Translog	Normale e semi-Normale	Approccio ibrido
Humphrey, Pulley (1997)	1977-1988	Usa	TFA	Translog	Normale e semi-Normale	Approccio ibrido
Lucchetti, Papi, Zazzaro (1997)	1982-1994	Regioni italiane	SFA	Cobb-Douglas	Normale e semi-Normale	Approccio ibrido
Altunbas, Goddard, Molyneux (1999)	1989-1996	EU-15	SFA	Fourier-flexible form	Normale e semi-Normale	Approccio di intermediazione
Altunbaş, Gardener, Molyneux, Moore (2001)	1989-1997	Si veda Nota 11	SFA	Fourier-flexible form	Normale e semi-Normale	Approccio di intermediazione
Huizinga, Nelissen, Vennet (2001)	1994-1998	Paesi Europei	SFA	Translog	Normale e semi-Normale	Approccio ibrido
Casu, Molyneux (2003)	1993-1997	Francia, Germania, Italia, Spagna, UK	DEA	-	-	Approccio di intermediazione
Molyneux (2003)	1992-2000	Paesi membri EU-15	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Girardone, Molyneux, Gardener (2004)	1993-1996	Italia	SFA	Translog	Normale e semi-Normale	Variante approccio di intermediazione
Weill (2004)	1992-1998	Francia, Germania, Italia, Spagna, Svizzera	DEA DFA SFA	Translog per l'approccio SF	Normale e Gamma per l'approccio SF	Approccio di intermediazione
Beccalli, Casu, Girardone (2006)	1999-2000	Francia, Germania, Italia, Spagna, UK	DEA SFA	Translog per l'approccio SF	Normale e semi-Normale per l'approccio SF	Approccio di intermediazione
Casu, Girardone (2006)	1997-2003	Paesi membri EU-15 (senza UK)	DEA	-	-	Variante approccio di intermediazione

Giordano, Lopes (2006)	1993-2003	Italia	SFA	Translog	Normale e semi-Normale	Approccio ibrido
Fontani, Vitali (2007)	1993-2004	Italia	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Behr, Tente (2008)	2005	Germania	SFA	Translog	a) Normale- semi-Normale b) Normale-Esponenziale	Approccio di intermediazione
Dongili, Rossi, Zago (2008)	1995-2005	Francia, Germania, Italia, Spagna, UK	SFA	Trnaslog e Fourier-flexible form	Normale e semi-Normale	Approccio di intermediazione
Giordano, Lopes (2008)	1993-2003	Italia (Nord/Sud)	SFA	Translog	Normale e semi-Normale	Approccio ibrido
Turati (2008)	1992-1999	Francia, Germania, Italia, Spagna, UK	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Casu, Girardone (2009)	2000-2005	Francia, Germania, Italia, Spagna, UK	DEA SFA	-	-	...
Chortareas, Garza-Garcia, Girandone (2009)	1997-2005	America Latina	DEA	-	-	Variante approccio di intermediazione
Casu, Girardone (2010)	1997-2003	EU-15	DEA	-	-	Variante approccio di intermediazione
Fiordelisi, Marques-Ibanez, Molyneux (2010)	1995-2007	EU-26	SFA	Translog	Normale e semi-Normale	Approccio ibrido
Battaglia, Farina, Fiordelisi, Ricci (2010)	2000-2005	Italia	SFA	Translog	Normale e semi-Normale	Approccio di intermediazione
Barra, Destefanis, Lavadera (2011)	2006-2008	Italia	DEA	-	-	Approccio ibrido

## 1.6 Conclusioni

La parte centrale di questo capitolo mostra come lo studio dell'efficienza delle banche sia un argomento di ricerca di elevato interesse per numerosi studiosi ed analisti, anche perché i processi di concentrazione che si sono verificati a partire dai primi anni Novanta avevano lo scopo di aumentare la stabilità e l'efficienza di un settore determinante per orientare lo sviluppo di un paese. Dalla rassegna della letteratura si evince che non ci sono chiare ed univoche indicazioni in merito agli effetti dei processi di consolidamento.

Dallo studio dei lavori esaminati emerge, inoltre, che i risultati sono sensibili alla scelta dei metodi di stima dei livelli di efficienza e che esiste una certa arbitrarietà nella scelta delle variabili di input ed output da inserire nei modelli.

Nel capitolo successivo si propone la stima dei livelli di efficienza di costo delle banche italiane nel periodo compreso tra il 2006 e il 2010 attraverso il metodo parametrico delle frontiere stocastiche e, sotto il profilo delle variabili incluse nel modello, la scelta, in questo lavoro, ricade sull'approccio di intermediazione, anziché su quello di produzione (Sealey e Lindley, 1977). A tal proposito, infatti, si tiene conto della particolarità dei servizi di intermediazione offerti dagli istituti di credito e, in considerazione di questo aspetto, tra gli output si inseriscono i prestiti verso i clienti, le commissioni attive e i titoli, mentre tra gli input si includono il lavoro, il capitale e i depositi.

La scelta di utilizzare un metodo parametrico è giustificata dal fatto che, attribuendo una funzione di distribuzione dell'errore, tale metodo consente di fare inferenza sui parametri stimati. In secondo luogo, la specificazione proposta da Battese e Coelli (1995) dell'approccio delle frontiere stocastiche consente di stimare simultaneamente l'equazione dell'inefficienza. Questo comporta una serie di vantaggi tra cui, per esempio, l'inserimento nella funzione dei costi di una parte d'errore esogena. Inoltre, in un contesto in cui è sempre più crescente la competitività, anche a livello internazionale, sembra utile capire quali sono le determinanti dell'efficienza delle banche, una classe di attori che contribuiscono ad innescare meccanismi di rilancio di un'economia.

## **CAPITOLO II: L'efficienza del sistema bancario italiano dal 2006 al 2010: un'applicazione delle frontiere stocastiche**

### **2.1 Introduzione**

Da un punto di vista degli studi originati dai cambiamenti del sistema bancario italiano, un importante filone di ricerca è quello legato alla misurazione delle performance delle aziende di credito. La ricerca sull'efficienza bancaria è stata condotta, nella maggior parte dei casi, mediante analisi della frontiera di efficienza. E' utile ricordare fin d'ora che non esiste in letteratura un concetto univoco della nozione di efficienza. In generale, per efficienza si intende il grado di aderenza del processo di produzione ad uno standard di ottimalità ed indica la capacità dell'impresa di minimizzare le quantità di input per la produzione di un certo output, oppure di massimizzare la quantità di output dato un certo ammontare di fattori produttivi. In questo studio, l'analisi si concentra sull'efficienza di costo delle banche italiane. A tal proposito, la funzione di costo permette di determinare una corrispondenza tra le combinazioni di output e prezzi dei fattori e la minima spesa necessaria per produrre quell'output a quei determinati prezzi (Fabbri, 1996). La distanza tra ciascuna osservazione e la frontiera può essere intesa come misura dell'inefficienza. A tal proposito, Leibenstein (1966) suggerisce che il complesso di scelte e di aggiustamenti determina una situazione di equilibrio diversa da quella corrispondente alla perfetta efficienza delle imprese. Da questo deriva il concetto di *X-efficiency*, che misura la distanza tra i costi sostenuti dall'azienda e la frontiera definita non in base al minimo costo conseguibile in condizioni di efficienza assoluta, ma come distanza dall'impresa che consegue i migliori risultati, ovvero quella che, in qualche modo, fissa la *best-practice frontier* di un particolare campione.

Una delle metodologie comunemente più utilizzate per misurare l'efficienza è l'approccio delle frontiere stocastiche (SFA), che è quello proposto anche in questo studio. Com'è stato evidenziato nel capitolo precedente, lo SFA è un metodo di tipo parametrico che gode del vantaggio di distinguere la distanza dalla frontiera di massima efficienza in due componenti, una casuale e l'altra ascrivibile all'inefficienza. Un vantaggio di questo approccio risiede nel fatto che, rientrando nella classe dei metodi parametrici, assegna una funzione di distribuzione al termine d'errore, permettendo, quindi, di fare inferenza sui parametri stimati (Aiello, Mastromarco e Zago, 2010). Inoltre, tale metodo, se è utilizzato seguendo la specificazione proposta da Battese e Coelli (1995), consente anche di isolare l'eterogeneità nel campione individuando fattori che spiegano e influenzano la componente di

inefficienza. Infatti, secondo tale specificazione, si stima simultaneamente la frontiera di costo e l'equazione dell'inefficienza (Battaglia *et al.*, 2010; Battese e Coelli, 1995).

I dati di bilancio delle banche analizzate in questo capitolo sono estratti dall'archivio dell'Associazione Bancaria Italiana (ABI) e sono relativi al periodo compreso tra il 2006 e il 2010.

Relativamente alla scelta delle variabili da inserire nella funzione dei costi che caratterizzano il settore bancario (Sealey e Lindley, 1977), è stato scelto di seguire l'approccio di intermediazione per considerare la particolare natura delle banche, le quali svolgono, per l'appunto, la funzione di intermediari creditizi.

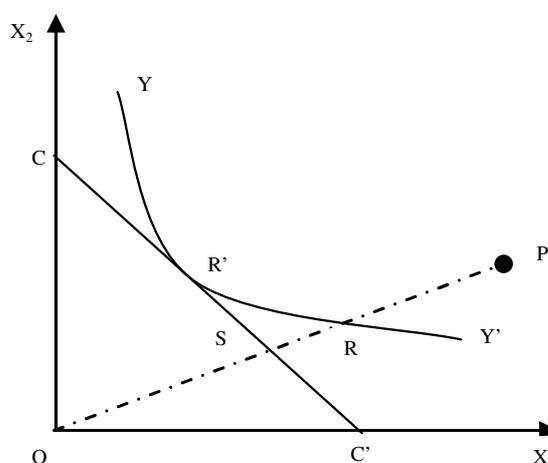
Il capitolo è organizzato nel modo seguente. Il paragrafo 2.2 descrive le caratteristiche e le proprietà delle funzioni di costo e il legame con l'efficienza intesa come la capacità di minimizzare la spesa per la produzione di un certo output. Il paragrafo 2.3 descrive il campione, mentre, quello successivo illustra la metodologia utilizzata e la scelta delle variabili introdotte nella funzione di costo e nell'equazione dell'inefficienza. I risultati ottenuti sono illustrati nel paragrafo 2.5, mentre le prove di robustezza e le conclusioni sono riportate, rispettivamente, nei paragrafi 2.6 e 2.7. Infine, nell'appendice A1, è riportato il test di Granger sulla relazione tra l'efficienza e la qualità dei crediti e, per completezza di analisi, nell'appendice A2 sono riportate le stime su dati cross section.

## **2.2 Efficienza e funzioni di costo**

In letteratura non esiste un concetto univoco della nozione di efficienza. In generale, esprime il grado di aderenza del processo di produzione ad uno standard di ottimalità ed, in senso tecnico, può essere intesa come la capacità dell'impresa di minimizzare le quantità di input per la produzione di un certo output, oppure di massimizzare la quantità di output dato un certo ammontare di fattori produttivi. A riguardo, si fa riferimento alla frontiera di produzione, la quale consente di individuare la produzione massima ottenibile attraverso determinati input, oppure alla frontiera di costo, che consente di individuare i costi minimi da sostenere per ottenere un certo ammontare di output. Pertanto, per funzione (o frontiera) di produzione si intende la mappa, ovvero l'insieme dei punti, in cui vi è corrispondenza tra una combinazione di input e il massimo output ottenibile con quella determinata combinazione. Analogamente, la funzione di costo permette di determinare una corrispondenza tra le combinazioni di output e prezzi dei fattori e la minima spesa necessaria per produrre quell'output a quei determinati prezzi (Fabbri, 1996).

La distanza tra ciascuna osservazione e la frontiera può essere intesa come misura dell'inefficienza e, se si considera soltanto il concetto di efficienza tecnica, l'inefficienza è da intendere come la distanza di ciascuna osservazione dalla frontiera. Tuttavia, come suggeriscono i primi studi sull'efficienza (Farrell, 1957), le analisi effettuate mediante il concetto di frontiera si riferiscono al livello di efficienza totale, che comprende anche la cosiddetta inefficienza allocativa, la quale deriva dall'impiego dei fattori produttivi in proporzioni non ottimali che, in quanto tali, rendono il processo produttivo più costoso.

**Figura 2.1 Efficienza tecnica ed efficienza allocativa**



Fonte: Murillo-Zamorano (2004).

In generale, un'impresa risulta inefficiente in senso tecnico se produce una certa quantità di output utilizzando un ammontare maggiore di input rispetto a quello minimo necessario, oppure se data una certa quantità di input, non riesce a produrre l'output prefissato. Un'impresa risulta inefficiente da un punto di vista allocativo se, dati i prezzi di mercato, la proporzione degli input rispetto agli output è superiore a quella ottimale. Un'impresa è allocativamente efficiente se riesce a selezionare quel mix di input che conduce alla produzione di un certo output minimizzando i costi. Il punto  $R'$  (Figura 2.1) individua una combinazione efficiente sotto entrambi i punti di vista in quanto giace sulla curva di isoquanti nel punto di tangenza alla retta di isocosto  $CC'$  la quale garantisce la minimizzazione dei costi.

Graficamente, è possibile individuare il livello di inefficienza tecnica e di quella allocativa in corrispondenza della combinazione  $P$  dei due input  $X_1$  e  $X_2$ . Tale combinazione risulta essere inefficiente sia dal punto di vista tecnico sia da quello allocativo. L'isoquanti  $YY'$  descrive l'intero set di produzione efficiente in quanto individua le combinazioni del

minimo ammontare di input necessario a produrre lo stesso livello di output. Murillo-Zamorano (2004) suggerisce di misurare l'inefficienza di P attraverso la distanza radiale lungo la retta che lega questo punto all'origine degli assi. L'efficienza tecnica può essere vista come il rapporto tra OR e OP ed intesa come il rapporto tra la combinazione di input tecnicamente efficiente e quella osservata. Una misura di efficienza allocativa, invece, è data dal rapporto tra OS e OR, che identifica il rapporto tra il costo minimo, ovvero quello ottimale, e il costo necessario per produrre l'output. In letteratura (Farrel, 1957; Murillo-Zamorano, 2004), l'efficienza complessiva, ovvero l'efficienza economica (EE), è data dall'interazione tra la componente tecnica (TE) e quella allocativa (AE):

$$EE = TE \times AE = OR/OP \times OS/OR = OS/OP$$

da cui deriva che la distanza dalla combinazione efficiente debba essere intesa in termini di riduzione dei costi. Ad esempio, se si considera la combinazione R, efficiente solo da un punto di vista tecnico, si può affermare che il rapporto SR/OR rappresenta la riduzione del costo che un produttore dovrebbe essere in grado di attuare per ottenere una combinazione produttiva (R') che sia anche allocativamente efficiente (Murillo-Zamorano, 2004).

E' importante considerare che, in merito al criterio utilizzato generalmente per individuare la frontiera, Leibenstein (1966) suggerisce che il complesso di scelte e di aggiustamenti porta ad una situazione di equilibrio diversa da quella corrispondente alla perfetta efficienza delle imprese. Da questo deriva il concetto di *X-efficiency*, che misura la distanza tra la produzione conseguita (o i costi sostenuti) dall'azienda e la frontiera definita non in base al massimo (o al minimo) conseguibile in condizioni di efficienza assoluta, ma con riferimento all'impresa che consegue i migliori risultati, ovvero quella che costituisce la *best-practice frontier*.

In questo studio, l'analisi si concentra sull'efficienza di costo delle banche italiane<sup>23</sup>. In particolare, si stima una frontiera di costo multi-output, poiché l'obiettivo è quello di minimizzare i costi, diversamente da quanto accade quando si stima una frontiera di produzione. In tal caso, infatti, si avrebbe avuto il problema della scelta dell'output da

---

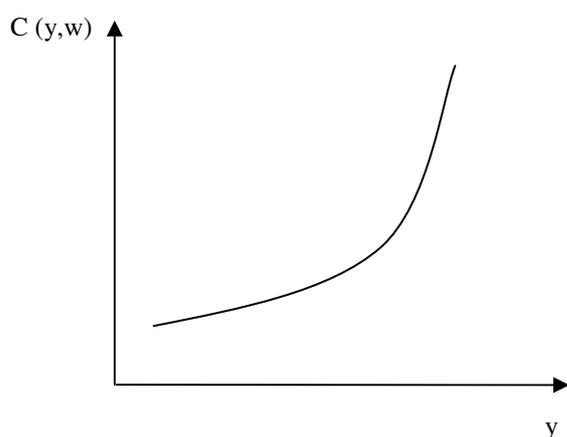
<sup>23</sup> In molti lavori viene affrontato lo studio dell'efficienza di costo insieme a quello dell'efficienza di profitto. In questa analisi si trascurava la misurazione di quest'ultima perché implica la trasformazione logaritmica di una variabile che è essa stessa una trasformazione della proxy del profitto (margine di interesse o margine di intermediazione o risultato netto d'esercizio) in quanto per alcune banche in alcuni anni le osservazioni di questa variabile potrebbero essere negative. Pertanto, in letteratura vengono proposte trasformazioni che rendono tali osservazioni positive (Huizinga *et al.*, 2001).

massimizzare e, comunque, la scelta sarebbe ricaduta preferibilmente solo su uno (Lensink e Meesters, 2012).

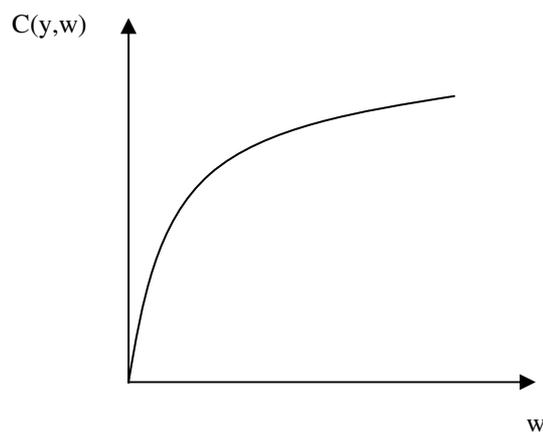
A tal proposito, la frontiera dei costi  $C(y,w)$  rappresenta l'insieme delle combinazioni di input (con prezzi  $w$ ) ed output che consentono alla banca di produrre una determinata quantità di output  $y$ , minimizzando i costi.

La frontiera di costo è una funzione crescente rispetto al livello di output prodotto e al prezzo degli input (figura 2.2 (a) e (b))<sup>24</sup>. La distanza di una banca da questa frontiera fornisce una misura della differenza tra il costo sostenuto e quello che sarebbe stato il costo minimo se la banca fosse pienamente efficiente.

**Figura 2.2 (a) Una frontiera di costo**



**Figura 2.2 (b) Una frontiera di costo**



La funzione di costo presenta, tra le altre, le seguenti proprietà (Chambers, 1988; Kumbhakar e Lovell, 2000):

1.  $C(0,w) = 0$  (Assenza di costi fissi).
2.  $C(y,w) > 0$  per  $y > 0$  (1. e 2. Garantiscono la non negatività della funzione).
3.  $C(y,w') \geq C(y,w)$  per  $w' \geq w$  (non decrescente in  $w$ ).
4.  $C(y,w)$  è concava e continua in  $w$ .
5.  $C(y',w) \leq C(y,w)$  per  $0 \leq y' \leq y$  (non decrescente in  $y$ ).
6.  $C(y,tw) = t C(y,w)$ , è una funzione omogenea di grado 1 rispetto ai prezzi degli input ( $t > 0$ ).

<sup>24</sup> La figura 2.2 (a) indica la presenza di rendimenti di scala della funzione dei costi che si ipotizzano decrescenti dal momento che la curva di costo totale aumenta in maniera più che proporzionale rispetto agli output.

La proprietà della non-negatività legata a quella secondo la quale non sono previsti costi fissi, implica che non è possibile produrre una quantità positiva di output senza sostenere costi. Tale proprietà implica una essenzialità debole della funzione di produzione in quanto è necessaria una quantità strettamente positiva (e almeno pari a uno) di input per avere una quantità positiva di output.

La proprietà 3 indica che la funzione dei costi non è decrescente rispetto al prezzo degli input e, quindi, implica che un aumento dei prezzi dei fattori produttivi non comporta un decremento dei costi. Si supponga che  $w_1 \geq w_2$  e che  $x_1$  e  $x_2$  siano i vettori degli input che minimizzano i costi in corrispondenza dei prezzi  $w_1$  e  $w_2$ . Dal momento che una variazione del prezzo degli input non influenza le scelte relative alle quantità utilizzate dalle imprese,  $x_1$  potrebbe potenzialmente minimizzare i costi quando il prezzo è pari a  $w_2$ . Ma in tali condizioni, la quantità  $x_1$  non verrà realizzata. Pertanto, risulta valida la seguente disuguaglianza:

$$w_2 x_2 \leq w_2 x_1 \quad (2.1)$$

Infatti, solo le combinazioni  $(x_1, w_1)$  e  $(x_2, w_2)$  minimizzano la spesa e, quindi, le imprese non sceglieranno di produrre combinazioni diverse. E', dunque, verificata anche la seguente espressione:

$$w_1 x_1 \leq w_1 x_2 \quad (2.2)$$

Inoltre, dato che  $w_1 \geq w_2$ , è vera anche la seguente disuguaglianza:

$$w_1 x_1 \geq w_2 x_1 \quad (2.3)$$

Introducendo tali disuguaglianze nella funzione di costo, a parità di  $y$ , si ha:

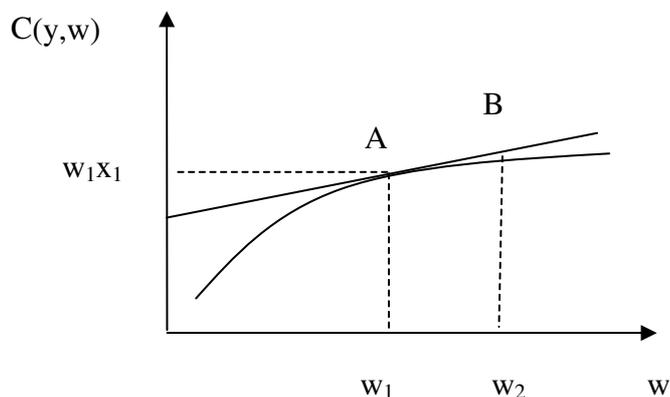
$$C(\bar{y}, w_1) = w_1 x_1 \geq w_2 x_2 = C(\bar{y}, w_2)$$

Dalla considerazione delle espressioni (2.1), (2.2) e (2.3) si giunge all'ineguaglianza fondamentale della minimizzazione dei costi:

$$(w_1 - w_2)(x_1 - x_2) \leq 0 \quad (2.4)$$

Dall'espressione 2.4 deriva anche la proprietà della concavità di cui gode la funzione dei costi.

**Figura 2.3** Concavità della funzione dei costi



La retta che passa attraverso i punti A e B individua l'effetto della variazione del prezzo dell'input  $x_1$  sulla funzione dei costi, tenendo fissa la quantità. Se è valida la disuguaglianza fondamentale della minimizzazione dei costi, la funzione non può essere che concava (figura 2.3). In altri termini, B non potrà appartenere alla frontiera dei costi in quanto la combinazione  $(w_2, x_1)$  non minimizza tale funzione.

Inoltre, considerate due combinazioni  $(w_1, x_1)$  e  $(w_{11}, x_{11})$ , le quali garantiscono entrambe la minimizzazione dei costi e data la combinazione lineare  $w_0 = \theta w_1 + (1 - \theta)w_{11}$ <sup>25</sup>, la concavità si ha se è soddisfatta la seguente disuguaglianza:

$$C(w_0, \bar{y}) \geq \theta C(w_1, \bar{y}) + (1 - \theta)C(w_{11}, \bar{y})$$

Se  $(w_1, x_1)$  e  $(w_{11}, x_{11})$  minimizzano la funzione dei costi, allora è vero che  $w_1 x_0 \geq w_{11} x_0$  e  $w_{11} x_0 \geq w_1 x_0$ . Pertanto, è possibile scrivere che

$$\begin{aligned} C(w_0, \bar{y}) &= w_0 x_0 = [\theta w_1 + (1 - \theta)w_{11}] x_0 = \\ &= \theta w_1 x_0 + (1 - \theta)w_{11} x_0 \geq \theta C(w_1, \bar{y}) + (1 - \theta) C(w_{11}, \bar{y}) \end{aligned}$$

In tal modo, si dimostra che la funzione dei costi gode della proprietà di concavità.

<sup>25</sup> Con  $0 \leq \theta \leq 1$ .

La proprietà 6 dell'omogeneità implica che se il prezzo degli input e il livello di output variano proporzionalmente, allora la scelta della combinazione che minimizza la funzione dei costi non varia, mentre la proprietà secondo la quale non ci sono costi fissi implica che tutti gli input sono perfettamente variabili, ossia tutti dipendenti dalla quantità di output prodotta.

Nel caso specifico, lo studio di frontiera viene condotto su un campione di banche e, a tal proposito, i concetti e le proprietà delle funzioni di costo discussi finora possono essere applicati anche all'analisi dell'efficienza di sistemi bancari. Da un punto di vista tecnico, una banca si dice inefficiente se utilizza troppi fattori produttivi per ottenere un certo output, oppure se, data una certa quantità di input, non riesce ad ottenere la quantità massima di prodotto possibile. In questo caso, la banca si troverà al di sotto della *best-practice frontier*. Risulta difficile determinare con univocità quando una banca possa essere definita più o meno efficiente tecnicamente e quali sono le discriminanti che consentano di poterla definire tale. Una banca, infatti, può essere più efficiente se eroga un ammontare maggiore di crediti raggiungendo livelli più elevati di impieghi, ma allo stesso tempo potrebbe essere più efficiente una banca che, pur avendo un ammontare minore di impieghi, essi siano di maggiore qualità in termini di rischio e solvibilità. Da un punto di vista allocativo, una banca è inefficiente se utilizza un mix sbagliato di input per produrre un certo output, ovvero quando, dato il vettore dei prezzi di mercato, la proporzione degli input rispetto agli output è superiore a quella ottimale. Pertanto, la banca potrebbe trovarsi sulla frontiera senza minimizzare, però, i costi (De Siano, 2001; Mester, 1996). In altre parole, la banca potrebbe essere efficiente da un punto di vista tecnico (raggiungendo un dato livello di output con un impiego minimo di input oppure massimizzando gli output con un dato livello di fattori produttivi), ma non esserlo da un punto di vista allocativo (in quanto potrebbe utilizzare input relativamente troppo costosi).

### **2.3 Descrizione del campione**

I dati utilizzati sono estratti dall'archivio *ABI Banking Data* dell'Associazione Bancaria Italiana (ABI), in cui sono disponibili i bilanci delle banche italiane dal 1993 al 2010. Il periodo scelto ai fini di questa analisi è quello compreso tra il 2006 e il 2010 e tale scelta dipende dalle seguenti considerazioni. In primo luogo, l'efficienza del sistema bancario italiano negli anni Novanta e nei primi anni 2000 è stata già studiata in molti lavori (cfr. § 1.5). Inoltre, il 2005 è da considerarsi come spartiacque tra la legislazione derivante dal decreto legislativo n°87 del 1992 e quella derivante dalla normativa comunitaria che prevede

l'introduzione dei principi internazionali di contabilità, i cosiddetti IAS (*International Accounting Standards*). Nel 2005, l'adozione dei principi IAS era obbligatoria solo a livello consolidato, mentre, era facoltativo per le singole banche che potevano decidere se aderire o meno. Tuttavia, l'obbligatorietà dell'adesione decorre dal 1° gennaio 2006 anche per le singole banche. Poiché, le voci di bilancio "pre-IAS" e quelle imputate secondo i principi internazionali non sono confrontabili, si è preferito concentrare l'analisi nel periodo più recente.

Nella tabella 2.1 si presentano alcuni dati che si riferiscono al campione di banche in esame. Si fa riferimento al numero di banche classificate in base alla forma giuridica, alla dimensione e all'area territoriale in cui è localizzata la sede legale.

**Tabella 2.1 Composizione delle banche del campione per forma giuridica, classe dimensionale e area territoriale<sup>26</sup>**

	2006	2007	2008	2009	2010
<b>Banche</b>	<b>686</b>	<b>692</b>	<b>689</b>	<b>686</b>	<b>648</b>
di cui:					
<i>Spa</i>	218	218	222	233	207
<i>Popolari</i>	37	38	39	39	35
<i>BCC</i>	431	436	428	414	406
<b>Banche</b>	<b>686</b>	<b>692</b>	<b>689</b>	<b>686</b>	<b>648</b>
di cui:					
<i>Banche Minori</i>	514	520	521	517	496
<i>Banche Piccole</i>	118	119	118	121	109
<i>Banche Medie</i>	32	35	33	29	29
<i>Banche Grandi</i>	13	11	10	12	10
<i>Banche Maggiori</i>	9	7	7	7	4
<b>Banche</b>	<b>686</b>	<b>692</b>	<b>689</b>	<b>686</b>	<b>648</b>
con sede:					
<i>Area Nord Ovest</i>	151	149	144	152	138
<i>Area Nord Est</i>	241	242	242	239	231
<i>Area Centro</i>	151	150	154	150	144
<i>Area Sud e Isole</i>	143	151	149	145	135

Fonte: elaborazione su dati ABI.

<sup>26</sup> La classificazione dimensionale e quella territoriale presenti nella banca dati ABI sono quelle proposte dalla Banca d'Italia. In particolare, le dimensioni sono definite in base agli impieghi e ai depositi, al numero di dipendenti e al numero di sportelli. La classificazione territoriale consiste nei seguenti raggruppamenti: Nord-Ovest (Liguria, Lombardia, Piemonte, Valle d'Aosta); Nord-Est (Emilia Romagna, Friuli Venezia Giulia, Trentino, Veneto); Centro (Abruzzo, Lazio, Marche, Toscana, Umbria); Sud e Isole (Basilicata, Calabria, Campania, Molise, Puglia, Sicilia, Sardegna).

Nel campione sono presenti 686 banche nel 2006, 692 nel 2007, 689 nel 2008, 686 nel 2009 e 648 nell'ultimo anno e risulta costituito prevalentemente da BCC che rappresentano il 63% del totale, seguite da Spa (32%) e dalle banche popolari (6%). Come si può notare, la maggior parte delle banche sono piccole e minori (92% del campione nel 2006 e 88% nel 2010).

**Tabella 2.2 Totale attività, impieghi, raccolta e costi delle banche italiane incluse nel campione (2006-2010)**

(Valori costanti in milioni di euro - Indice NIC Istat, anno base = 1995)

	2006	2007	2008	2009	2010
<b>Totale Attivo</b>	<b>1,896,071</b>	<b>2,064,221</b>	<b>2,243,777</b>	<b>2,239,651</b>	<b>2,022,890</b>
di cui:					
<i>Spa</i>	1,597,072	1,710,287	1,910,688	1,880,901	1,659,514
<i>Popolari</i>	195,223	241,982	214,495	233,842	233,906
<i>BCC</i>	103,775	111,952	118,594	124,909	129,471
<b>Totale Impieghi</b> (Crediti verso clienti)	<b>1,028,274</b>	<b>1,122,542</b>	<b>1,180,088</b>	<b>1,161,220</b>	<b>1,204,578</b>
di cui:					
<i>Spa</i>	847,458	915,280	966,325	929,557	972,818
<i>Popolari</i>	107,858	126,324	127,379	142,338	138,054
<i>BCC</i>	72,958	80,938	86,384	89,325	93,705
<b>Totale Raccolta</b> (Debiti verso clienti)	<b>668,798</b>	<b>697,579</b>	<b>716,911</b>	<b>733,914</b>	<b>741,774</b>
di cui:					
<i>Spa</i>	540,062	559,002	576,867	578,068	586,552
<i>Popolari</i>	75,193	83,207	83,455	95,134	92,990
<i>BCC</i>	53,544	55,371	56,589	60,712	62,232
<b>Totale Costi<sup>27</sup></b>	<b>110,398</b>	<b>123,147</b>	<b>137,701</b>	<b>104,575</b>	<b>89,500</b>
di cui:					
<i>Spa</i>	92,450	102,118	114,934	85,694	72,246
<i>Popolari</i>	11,402	13,389	14,055	11,402	10,157
<i>BCC</i>	6,546	7,640	8,712	7,480	7,097

Fonte: elaborazione su dati ABI.

<sup>27</sup> Il totale dei costi include le spese amministrative, i costi operativi, gli ammortamenti delle immobilizzazioni, le commissioni passive e gli interessi passivi.

Inoltre, il numero di banche che hanno sede legale nel Nord è pari al 60% del campione. Si tratta di un valore molto più elevato di quello associato alle banche che hanno sede legale nel Sud (20%).

La dimensione espressa in termini di Totale Attivo oscilla tra 1896 milioni di euro nel 2006 e 2023 milioni di euro nel 2010. Occorre ricordare, inoltre, che circa l'85% del totale attivo dell'intero settore proviene dalle Spa, il 10% dalle banche popolari e il 5% dalle BCC (tabella 2.2).

Si può notare che per tutte le variabili considerate emerge una tendenziale crescita dal 2006 al 2008. Nel 2009 e nel 2010, invece, sono evidenti delle riduzioni. Si può osservare, inoltre, che il maggiore contributo alla composizione delle variabili è dato dalle banche Spa ed in misura minore dalle BCC.

La tabella 2.3 riporta i dati relativi al numero dei dipendenti e al numero degli sportelli per i quattro anni inclusi nell'analisi. Se da un lato gli sportelli aumentano fino al 2009 e subiscono una riduzione nell'ultimo anno, il numero dei dipendenti registra una contrazione di circa 6 mila unità dal 2007 al 2008, di circa 12 mila unità nel 2009, per continuare a diminuire nel 2010 fino a circa 309 mila unità.

**Tabella 2.3 Dipendenti e Sportelli delle banche incluse nel campione (2006-2010)**

	2006	2007	2008	2009	2010
<b>Dipendenti</b>	<b>338,520</b>	<b>340,914</b>	<b>335,164</b>	<b>321,903</b>	<b>309,307</b>
di cui:					
<i>Banche Spa</i>	270,744	269,557	265,508	251,603	240,715
<i>Banche Popolari</i>	40,146	42,578	39,614	39,980	37,561
<i>BCC</i>	27,630	28,779	30,042	30,320	31,031
<b>Sportelli</b>	<b>32,866</b>	<b>33,098</b>	<b>34,065</b>	<b>34,211</b>	<b>33,184</b>
di cui:					
<i>Banche Spa</i>	25,215	25,140	25,524	25,603	24,667
<i>Banche Popolari</i>	3,858	4,074	4,380	4,427	4,200
<i>BCC</i>	3,793	3,884	4,161	4,181	4,317

Fonte: elaborazioni su dati ABI.

## 2.4 La metodologia e la scelta delle variabili

In questo paragrafo si descrive l'approccio delle frontiere stocastiche, che è la metodologia utilizzata per misurare l'efficienza delle banche italiane dal 2006 al 2010.

Come già accennato (Cfr. § 1.4), esistono metodi parametrici e non parametrici in questa classe di studi<sup>28</sup>. I metodi parametrici, assegnando una distribuzione all'errore, consentono di fare inferenza sui parametri stimati, attraverso i test parametrici di significatività. Altra classificazione è quella che distingue i metodi stocastici da quelli deterministici. I primi tengono conto del fatto che una banca possa allontanarsi dalla frontiera efficiente anche per motivi di natura casuale e non solo per l'inefficienza.

Il metodo parametrico più utilizzato è quello delle frontiere stocastiche, il quale permette di scomporre l'errore in due parti, una attribuibile all'inefficienza e una ad errori *random* (da cui l'aggettivo stocastiche assegnato al nome del metodo). Tale approccio appare preferibile rispetto ad uno dei metodi non parametrici più utilizzati, la DEA (*Data Envelopment Analysis*), la quale suppone che la distanza dalla frontiera sia spiegata interamente dall'inefficienza. La DEA non considera, quindi, errori casuali come possono essere, ad esempio, gli errori di misurazione delle variabili oppure quelli legati ad eventi non previsti. Il metodo delle frontiere stocastiche appare preferibile anche rispetto ad altri metodi parametrici, come il Thick Frontier Approach (TFA), perché garantisce la stima dei livelli di inefficienza per ogni singola *Decision Making Unit* (DMU) e, quindi, anche di un livello di efficienza del settore che si considera<sup>29</sup>.

Un ulteriore vantaggio del metodo SFA è la possibilità di inserire contestualmente nel modello di frontiera un insieme di variabili che spiegano la componente di inefficienza. In particolare, si fa riferimento all'approccio proposto da Battese e Coelli (1995), in cui si stimano in modo simultaneo il modello di frontiera e l'equazione dell'inefficienza. Questo metodo offre, quindi, la garanzia di considerare nelle stime della frontiera una componente di inefficienza resa esogena. Inoltre, Battaglia *et al.* (2010) sottolinea un ulteriore vantaggio derivante dall'applicazione di tale specificazione, che è quello relativo al fatto di tener conto dell'eterogeneità delle banche incluse nel campione attraverso l'inserimento nell'equazione dell'inefficienza di variabili individuali e di contesto delle singole DMU. In particolare, dal

---

<sup>28</sup> La distinzione tra i due tipi di metodi è relativa al fatto che i metodi parametrici assegnano funzioni di densità alla componente stocastica del modello, mentre i metodi non parametrici definiscono solo la parte deterministica.

<sup>29</sup> Il TFA è un metodo parametrico che permette di ottenere i valori previsti di performance a partire da una determinata specificazione funzionale della frontiera (di produzione o di costo). Secondo tale metodo, i valori della frontiera vengono raggruppati in intervalli quartili, assumendo che le differenze fra i valori previsti all'interno di ciascun intervallo siano espressione di fluttuazioni casuali, mentre le differenze rilevabili tra intervalli siano dovute all'inefficienza (Berger e Humphrey, 1997).

momento che si tratta di una stima simultanea, si tiene contemporaneamente conto di variabili che influenzano i costi delle banche in maniera indiretta perché ne influenzano l'efficienza (Lensink e Meesters, 2012).

Una delle questioni dirimenti di questo approccio è che, essendo di tipo parametrico, richiede l'attribuzione *ex ante* di una specifica forma funzionale della frontiera di costo e di una specifica distribuzione degli errori. Per quel che riguarda la forma funzionale della frontiera, in letteratura, si è consolidato l'uso di funzioni Cobb-Douglas (o di sue generalizzazioni). Mentre, per quanto riguarda le ipotesi relative alla distribuzione delle componenti erratiche, si ipotizza una funzione Normale per la parte casuale e una distribuzione semi-Normale per la componente di inefficienza (Dongili, Rossi e Zago, 2008; Turati, 2008). Inoltre, gli errori casuali sono considerati indipendenti dal termine associato all'inefficienza (Fiordelisi, Marques-Ibanez e Molyneux, 2010; Giordano e Lopes, 2006).

L'approccio delle frontiere stocastiche assume che tutte le banche siano confrontate con quella più efficiente che rappresenta, quindi, il *benchmark* (Liebenstein, 1966). Si assume, inoltre, che esse condividano la stessa funzione di produzione e che la tecnologia e il contesto istituzionale siano comuni a tutte le banche.

Per poter giungere alla stima dell'efficienza attraverso il metodo SFA, come già detto, occorre assegnare alla frontiera di costo una forma funzionale, che, in questo lavoro viste le assunzioni del paragrafo 2.2, si assume sia di tipo Cobb-Douglas:

$$Costi_{it} = C(y, w) \exp(v) \exp(u) = \prod_j y_{ij}^{\beta_j} \cdot \prod_n w_{in}^{\gamma_n} \cdot e^{v_{it} + u_{it}} \quad (2.5)$$

La funzione di costo indica il costo minimo per produrre un vettore di output  $y$  dato un vettore di prezzi  $w$ . La frontiera stocastica in forma log-lineare assume la seguente espressione:

$$\log Costi_{it} = \sum_{j=1}^J \beta_j \log y_{ij} + \sum_{n=1}^N \gamma_n \log w_{in} + u_{it} + v_{it} \quad (2.6)$$

in cui  $Costi_{it}$  è il costo totale sostenuto dalla  $i$ -esima banca al tempo  $t$ ;  $y_{ij}$  rappresenta il valore del  $j$ -esimo output ottenuto dalla banca  $i$  nell'anno  $t$ ;  $w_{in}$  è il costo dell' $n$ -esimo input nell'anno  $t$ ;  $\beta_j$  e  $\gamma_n$  sono i rispettivi parametri da stimare;  $u_{it}$  rappresenta la componente erratica

che misura l'inefficienza. Si tratta di una variabile non negativa;  $v_{it}$  è, invece, l'errore casuale; con  $j = 1, \dots, J$  e  $n = 1, \dots, N$ .

La funzione di costo stocastica è, quindi, composta da due parti, la componente deterministica e una componente stocastica, la quale comprende l'errore casuale e l'inefficienza. Da un punto di vista economico, questo aspetto assume particolare importanza in quanto l'interesse è di capire il ruolo di fattori che influenzano i costi di una banca non in modo diretto, ma attraverso il loro legame con l'efficienza.

$$\log Costi_{it} = \underbrace{\sum_j \beta_j \log y_{itj}}_{\text{COMPONENTE DETERMINISTICA}} + \underbrace{\sum_n \gamma_n \log w_{itn}}_{\text{COMPONENTE STOCASTICA}} + \underbrace{u_{it} + v_{it}}_{\text{COMPONENTE STOCASTICA}} \quad (2.6\text{-bis})$$

Il metodo di stima è quello della massima verosimiglianza. In letteratura (Kumbhakar e Lovell, 2000; Mastromarco, 2007), sono presenti altri metodi di stima sia per dati *cross section* che per dati *panel*, ma quello della massima verosimiglianza appare preferibile perché consente di ottenere stimatori consistenti ed efficienti<sup>30</sup>.

Tra le altre proprietà (cfr. § 2.2), si assume che la funzione di costo rispetti il vincolo di omogeneità lineare relativamente al prezzo degli input, ovvero:

$$\sum_n \gamma_n = 1 \quad (2.7)$$

Al fine di tenerne conto, il modello stimato può essere espresso nel seguente modo:

$$\log \frac{Costi_{it}}{w_{ir}} = \sum_{j=1}^J \beta_j \log y_{itj} + \sum_{n=1}^{N-1} \gamma_n \log \frac{w_{itn}}{w_{itr}} + u_{it} + v_{it} \quad (2.8)$$

Inoltre, la misura di efficienza di costo può essere espressa come il rapporto tra il livello minimo di costo sostenuto dalla banca se fosse efficiente e il livello di costo effettivamente osservato, ovvero:

---

<sup>30</sup> Altri metodi di stime dei livelli di efficienza nell'approccio delle frontiere stocastiche sono i metodi dei minimi quadrati corretti e modificati (COLS e MOLS) ed il metodo dei momenti.

$$CE_{it} = \frac{C(y_{it}, w_{it}) \exp(v_{it})}{C(y_{it}, w_{it}) \exp(v_{it}) \exp(u_{it})} = \exp(-u_{it}) \quad (2.9)$$

Il numeratore di questo rapporto rappresenta il costo minimo, in quanto la spesa osservata non può che essere maggiore o uguale a questa espressione, dal momento che l'eccesso è attribuibile alla componente di inefficienza (Battese *et al.*, 2005). L'espressione (2.9) indica, quindi, il rapporto fra il costo sostenibile, nell'ipotesi in cui l'unità statistica risulti efficiente, e il costo effettivamente sostenuto.

In modo analogo alle applicazioni effettuate in altri lavori (cfr. § 1.5), in questo capitolo si assume che l'errore casuale si distribuisca come una Normale con media zero e l'inefficienza come una Normale troncata secondo la specificazione di Battese e Coelli (1995). Si assume, inoltre, che l'errore casuale sia indipendentemente e identicamente distribuito. Formalmente,

$$v_{it} \sim iidN(0, \sigma_v^2) \quad (2.10)$$

$$u_{it} \sim N^+(z' \eta, \sigma_u^2) \quad (2.11)$$

Battese e Coelli (1995) suggeriscono di inserire nel modello di frontiera per dati panel variabili che influenzano direttamente i livelli di inefficienza e indirettamente anche i costi delle banche<sup>31</sup>. La componente di inefficienza può, quindi, essere specificata nel modo seguente:

$$u_{it} = \sum_{k=1}^K \eta_k z_{itk} + e_{it} \quad (2.12)$$

dove  $z_{ik}$  rappresenta la k-esima variabile che, al tempo  $t$ , influenza la banca i-esima; con  $k = 1, \dots, K$ .

L'idea è di stimare in maniera simultanea le equazioni (2.8) e (2.12), in modo da inserire nella funzione di costo la componente di inefficienza spiegata da variabili  $z_i$  che la influenzano (Battese e Coelli, 1995). In particolare, il termine di errore  $u$  è costituito da una componente deterministica, che dipende da un vettore di fattori osservabili,  $z$ , e da una

---

<sup>31</sup> Battese e Coelli (1995) estendono ai dati panel un approccio già utilizzato per i dati cross section. A riguardo, si vedano gli studi di Huang e Liu (1994), Kumbhakar, Ghosh e McGuckin (1991), Reifschneider e Stevenson (1991).

componente puramente casuale non osservabile,  $e$ . Infine, per garantire la non-negatività di  $u$  deve valere la disuguaglianza (espressa in forma matriciale)  $e \geq -z'\eta$ .

Un ulteriore vantaggio di tale specificazione è legato al fatto di considerare l'eterogeneità delle banche del campione che consente di ottenere stime non distorte (Battaglia *et al.*, 2010; Bos *et al.*, 2005). Per tenerne conto ed evitare eventuali problemi di distorsione, si stima simultaneamente l'equazione (2.12) eliminando i limiti dell'approccio cosiddetto "two steps" (Greene, 1993), secondo il quale l'inefficienza viene prima stimata nell'ambito di un modello base (equazione 2.8) e poi utilizzata come variabile dipendente in un successivo modello, comportando distorsione nelle stime. In particolare, come evidenziano Lensink e Meesters (2012) e Wang e Schimdt (2002), nell'approccio *two steps* potrebbero incorrere problemi di distorsione derivanti dal fatto che, nella stima della frontiera di costo, si ipotizza che la componente di inefficienza si distribuisca in maniera identica e indipendente, mentre, nella stima dell'equazione dell'inefficienza, si assume che essa dipende da una serie di variabili esplicative.

Per quel che riguarda le variabili da utilizzare nell'analisi econometrica, occorre ricordare che vi è un ampio dibattito tra l'approccio del valore aggiunto (o di produzione) e quello di intermediazione (Sealey e Lindley, 1977). Questi due approcci si distinguono essenzialmente per la definizione degli input e degli output da inserire nella stima di una frontiera di efficienza. L'approccio del valore aggiunto considera prestiti e depositi come output del processo produttivo di una banca, mentre il lavoro e il capitale fisico sono gli input. Pertanto, sembra che l'approccio di produzione tenda a considerare le banche al pari di altre imprese produttrici di servizi e prodotti. Al contrario, l'approccio di intermediazione individua gli impieghi come output di una banca e il lavoro, il capitale e i depositi come fattori produttivi. In tal caso, quindi, la banca, pur essendo considerata come un'impresa, viene intesa come un particolare soggetto che raccoglie e gestisce i depositi bancari per offrire impieghi alla clientela. Tra queste due posizioni estreme, esistono frequenti combinazioni dei due approcci, in cui, per esempio, nelle stime delle funzioni di costo, si considerano i prestiti e i depositi come output e il costo del lavoro, del capitale e della raccolta come prezzi degli input. In alcuni lavori (Giordano e Lopes, 2008), si tende a considerare nell'ambito degli output anche i ricavi provenienti dai servizi, ossia le commissioni, introducendo, quindi, variabili di flusso e non solo variabili di stock, quali sono il totale degli impieghi e il totale dei depositi. Nella rassegna sugli studi che misurano l'efficienza dei sistemi bancari proposta da Berger e Humphrey (1997), si giunge alla conclusione che l'approccio di intermediazione risulta più appropriato per valutare le istituzioni finanziarie. In base a tali considerazioni, le

variabili da inserire nel modello in questo studio sono scelte seguendo tale approccio, tenuto conto della particolarità dei servizi offerti dalle banche. Poiché si fariferimento ad una estensione dell'approccio di intermediazione, si introduce nel modello una variabile di flusso come altro output, ossia le commissioni bancarie (Dongili e Zago, 2005). Tale scelta è giustificata dalla considerazione che oggi le banche offrono una serie di servizi “collaterali” per i quali ottengono elevate commissioni attive<sup>32</sup>. La tabella 2.4 riassume le variabili introdotte nella stima della funzione di costo.

**Tabella 2.4 Definizione delle variabili incluse nella funzione di costo**

Variabili	Nome	Definizione
$y_1$	Prestiti	Crediti verso la clientela (Variabile di stock)
$y_2$	Commissioni Attive	Guadagni non derivanti dalle attività tradizionali di impiego e di raccolta delle banche (Variabile di flusso)
$y_3$	Titoli	Somma di crediti verso altre banche, titoli di capitale e titoli obbligazionari (Variabile di stock)
$x_1$	Lavoro	Numero di dipendenti
$x_2$	Capitale	Prodotto Bancario Lordo
$x_3$	Depositi	Totale debiti verso la clientela
$w_1$	Costo del lavoro	Rapporto tra spese del personale e numero dei dipendenti
$w_2$	Costo del capitale	Rapporto tra altre spese <sup>33</sup> e Prodotto Bancario Lordo
$w_3$	Costo della raccolta	Rapporto tra interessi passivi su debiti verso i clienti e debiti verso la clientela
$Costi(y, w)$	Costi totali	$w_1x_1 + w_2x_2 + w_3x_3 =$ Spese amministrative + Ammortamento immobilizzazioni + Interessi passivi + Costi operativi + Commissioni passive

<sup>32</sup> Nello specifico, si tratta di ricavi provenienti da operazioni di negoziazione di strumenti finanziari e valute, di custodia e amministrazione titoli, di attività di consulenza, di gestione di prodotti assicurativi, di incasso e pagamento, di esattorie e ricevitorie (ABI).

<sup>33</sup> Le altre spese comprendono i costi che non sono stati impiegati altrove nella funzione di costi. Tale voce comprende le spese relative alle commissioni passive, ai costi operativi, agli ammortamenti delle immobilizzazioni, alle spese amministrative che non fanno riferimento alle spese del personale e agli interessi passivi che non si riferiscono a quelli calcolati sui depositi. Sostanzialmente, al fine di calcolare il costo del capitale inteso in senso ampio, si fa riferimento a costi come canoni di locazione, spese per la pulizia dei locali, manutenzione, energia elettrica, riscaldamento e acqua, spese telefoniche, cancelleria varia.

Sebbene esista molto dibattito sulle specificazioni di input e output da scegliere negli studi di efficienza delle banche, vi è un certo consenso nel considerare i crediti verso clienti ( $y_1$ ) come l'output bancario per eccellenza<sup>34</sup>. Secondo quanto proposto da vari autori<sup>35</sup>, si considerano, inoltre, tra gli output, i guadagni derivanti da attività non tradizionali, utilizzando come proxy le cosiddette commissioni attive ( $y_2$ ), le quali comprendono proventi derivanti da servizi di varia natura e costituiscono un'importante variabile che aiuta ad identificare la performance delle banche. Il terzo output utilizzato in questo lavoro è rappresentato dai titoli ( $y_3$ ), ossia dai prestiti verso altre banche, titoli di capitale e obbligazioni (Barra, Destefanis e Lubrano Lavadera, 2011).

Dal lato degli input, la scelta è ricaduta sul lavoro, sul capitale e sulla raccolta, così come suggerisce l'approccio di intermediazione. Il lavoro ( $x_1$ ) è misurato come il numero dei dipendenti delle singole banche. Il costo del lavoro ( $w_1$ ) è calcolato come il rapporto tra le spese del personale e il numero dei dipendenti.

Nel caso del capitale, i problemi di misurazione sono più complessi e, non a caso, si tratta della variabile in cui la soggettività degli studiosi emerge con maggiore frequenza. Alcuni autori considerano il costo del capitale fisico, calcolato come il rapporto tra il deprezzamento e l'ammontare delle immobilizzazioni materiali (Fiordelisi *et al.*, 2010). Tuttavia, appare restrittivo includere nel capitale delle banche solo le immobilizzazioni materiali anche perché non si terrebbe conto del capitale, inteso in senso ampio come ciò di cui le banche si servono per svolgere la loro attività. Si considerino, a titolo di esempio, le banche telematiche che svolgono le loro attività con un capitale fisico molto ridotto. Barra *et al.* (2011) misurano il prezzo del capitale come rapporto fra la somma di spese amministrative, ammortamenti e costi operativi e il numero degli sportelli. Giordano e Lopes (2008) considerano lo stesso numeratore utilizzato in Barra *et al.* (2011) e lo rapportano al totale dei fondi raccolti.

---

<sup>34</sup> Tale voce comprende conti correnti, pronti contro termine attivi, mutui, carte di credito, prestiti personali e cessioni del quinto, operazioni relative a locazione finanziaria e factoring, attività deteriorate, titoli di debito strutturati e altri titoli.

<sup>35</sup> Barra, Destefanis e Lubrano-Lavadera (2011), Fiordelisi, Marques-Ibanez e Molyneux (2010), Giordano e Lopes (2008), Hunter e Timme (1995).

Un'accezione più ampia del concetto di capitale è proposta da Fontani e Vitali (2007), i quali calcolano il costo del capitale come rapporto tra patrimonio di vigilanza e prodotto bancario loro<sup>36</sup>.

Tenendo conto delle precedenti considerazioni, il costo del capitale ( $w_2$ ) è misurato come rapporto tra le spese che non sono considerate nelle altre variabili di input del modello di frontiera e il prodotto bancario lordo ( $x_2$ ). Il numeratore comprende, quindi, le spese amministrative al netto delle spese del personale, i costi operativi, gli interessi passivi al netto di quelli su debiti verso clienti, gli ammortamenti delle immobilizzazioni e le commissioni passive.

**Tabella 2.5 Valori medi delle variabili di input ed output dal 2006 al 2010**

(valori costanti in migliaia di euro - Indice NIC Istat, anno base = 1995)

Variabili	2006	2007	2008	2009	2010
<i>Costi</i>	130,897	182,003	202,113	152,768	142,597
$y_1 = \text{crediti verso clienti}$	1,302,929	1,654,131	1,735,899	1,711,587	1,947,854
$y_2 = \text{commissioni attive}$	31,037	28,700	25,731	25,755	23,216
$y_3 = \text{titoli}$	456,295	793,723	918,017	904,780	564,279
$w_1 = \text{costo del lavoro}$	52.32	52.88	53.97	53.26	52.97
$w_2 = \text{costo del capitale}$	0.0286	0.0891	0.0443	0.0991	0.0667
$w_3 = \text{costo della raccolta}$	0.0112	0.0149	0.0172	0.0119	0.0073
<i>Osservazioni</i>	597	634	680	639	596

Fonte: elaborazioni su dati ABI.

<sup>36</sup> Il patrimonio di vigilanza, il cui calcolo trova fondamento normativo negli Accordi di Basilea, è il principale parametro di riferimento per le valutazioni in merito alla solidità bancaria. Esso viene calcolato partendo dal patrimonio di base aumentato del cosiddetto patrimonio supplementare calcolati mediante l'applicazione dei filtri prudenziali. Rappresenta così una grandezza che esprime un presidio contro i rischi di mercato. Il prodotto bancario lordo, invece, è dato dalla somma tra gli impieghi, la raccolta diretta e la raccolta indiretta. Gli impieghi sono sostanzialmente costituiti dai crediti verso clienti, la raccolta diretta dai debiti verso clienti, mentre la raccolta indiretta è data dalla somma tra risparmio gestito e risparmio amministrato. Il risparmio gestito rappresenta l'ammontare dei fondi che la banca gestisce per conto dei risparmiatori ed è rappresentato dai fondi comuni di investimento, fondi pensione ed altro. Il risparmio amministrato è dato da investimenti in titoli azionari ed obbligazionari che la banca amministra appunto per conto dei suoi clienti.

Tra le spese amministrative sono comprese voci di costo, quali quelle relative ad energia elettrica, canoni di locazione e manutenzioni di vario tipo<sup>37</sup>.

Infine, il terzo input considerato è costituito dalla raccolta presso la clientela ( $x_3$ ) il cui costo ( $w_3$ ) è dato dal rapporto tra gli interessi passivi su debiti verso i clienti e l'ammontare dei depositi.

La variabile dipendente della funzione di costo,  $Costi(y, w)$ , indica il totale dei costi sostenuti dalle singole banche e sono calcolati come la somma tra spese amministrative, interessi passivi, costi operativi, commissioni passive e ammortamento delle immobilizzazioni.

Nella tabella 2.5 sono riportati i valori medi annuali delle variabili incluse nelle stime della frontiera dei costi. Il livello medio dei costi, in migliaia di euro, aumenta tra il 2006 e il 2008, per poi ridursi in misura notevole nei due anni successivi. I crediti verso i clienti ( $y_1$ ), invece, registrano una tendenziale crescita nel periodo esaminato, ad eccezione del 2009 in cui mostrano un leggero calo. Questi dati sembrano indicare la presenza di un buon andamento delle performance degli istituti di credito, poiché da un lato si assiste ad una tendenziale riduzione dei costi e dall'altro si osserva una crescita dei crediti verso la clientela. Per quel che riguarda gli altri due output inseriti nella funzione di costo, si noti che le commissioni attive tendono a diminuire nel corso del tempo, mentre i titoli aumentano dal 2006 al 2008 per ridursi negli ultimi due anni.

Per quanto riguarda il costo del lavoro, si osservi che nel 2008 ha avuto un considerevole aumento legato a due variazioni dello stesso segno delle variabili che lo determinano. Infatti, il numero dei dipendenti si è ridotto (tabella 2.3) e tale variazione è legata, verosimilmente, alle operazioni di ristrutturazione dovute alla crisi finanziaria. Inoltre, contemporaneamente, le spese del personale sono rimaste pressoché costanti tra il 2007 e il 2008 (18641 milioni nel 2007 e 18584 milioni nel 2008 – dati ABI, valori costanti).

E' interessante analizzare il costo del capitale ( $w_2$ ) che segue un andamento molto altalenante e che oscilla tra il 3% del 2006, il 9% nel 2007, il 4% nel 2008, il 10% nel 2009 e il 7% nell'ultimo anno.

---

<sup>37</sup> Esistono altre tecniche di misurazione del patrimonio delle banche quali, ad esempio, l'approccio dell'*Asset and Liabilities Management* (ALM) e quello del *Weighted Average Cost of Capital* (WACC). Al fine di definire il concetto di capitale, l'ALM consiste nella gestione integrata dell'attivo e del passivo finalizzata ad allocare le risorse in un'ottica di ottimizzazione del rapporto rischio-rendimento. Si tratta di una tecnica che spesso viene affidata a soggetti specializzati esterni alle banche. Il WACC, invece, consiste nella media ponderata tra il costo del capitale proprio e il costo del capitale di debito. In tal caso, la scelta di non utilizzare queste tecniche è stata dettata dal fatto che si tratta di effettuare calcoli relativamente laboriosi che, in questa sede, non costituiscono una questione centrale.

Infine, la tabella 2.5 evidenzia che il costo della raccolta è in media basso, infatti, negli anni compresi nel periodo esaminato, gli interessi passivi sui debiti verso clienti incidono in maniera marginale sui fondi raccolti, nell'ordine dell'1-2%.

In precedenza, si è già discusso dell'utilità di stimare una frontiera stocastica di costo utilizzando la specificazione di Battese e Coelli (1995), la quale consiste nella stima simultanea delle equazioni (2.8) e (2.12). Ciò implica la necessità di definire anche le determinanti dell'inefficienza, descritte nella tabella 2.6. Una prima variabile che spiega l'efficienza delle banche è rappresentata dalla qualità degli impieghi ( $z_1$ ), data dal rapporto tra i crediti in sofferenza e il totale dei crediti verso i clienti (Berger e De Young, 1997; Giordano e Lopes, 2008).

**Tabella 2.6 Definizione delle esplicative della componente di inefficienza**

Variabili	Nome	Definizione
$z_1$	Qualità dei crediti	Rapporto tra sofferenze e crediti verso clienti
$z_2$	Indice di solvibilità	Rapporto tra patrimonio di vigilanza e attività ponderate per il rischio; si tratta di una proxy del rischio cui sono soggette le banche. Tale misura tiene conto delle direttive presenti nella normativa di Basilea
$z_3$	<i>Cost/Income</i>	Rapporto tra costi operativi e margine di intermediazione <sup>38</sup> ; si tratta di una proxy della performance delle banche
$z_4$	Indice di Herfindahl	Calcolato come la somma del quadrato delle quote espresse come rapporto tra l'attivo di ciascuna banca e il totale attivo del settore; è una proxy della concentrazione del settore di attività
$z_5$	Tempo	<i>Time trend</i>
$d_{spa}, d_{pop}$	Forma giuridica	Dummies per forma giuridica (Gruppo di controllo: BCC)
$d_{minori}, d_{medie}, d_{grandi}, d_{maggiori}$	Dimensione	Dummies per classe dimensionale (Gruppo di controllo: banche piccole)
$d_{nordovest}, d_{centro}, d_{sud}$	Localizzazione geografica	Dummies per area territoriale (Gruppo di controllo: banche con sede legale nel Nord-Est)

<sup>38</sup> I costi operativi comprendono le spese del personale e le altre spese amministrative, mentre il margine di intermediazione rappresenta i ricavi netti di un intermediario finanziario. Esso è dato dalla somma tra il margine di interesse loro e del margine finanziario e di servizi.

La tabella 2.7 illustra l'evoluzione delle Sofferenze dal 2006 al 2010 in livelli ed in percentuale. Emerge che nel 2008 tale variabile si riduce dell'8% circa e che negli anni successivi subisce una crescita sostenuta registrando variazioni percentuali di oltre il 30%. Il rapporto Sofferenze/Totale Crediti segue lo stesso andamento. Tuttavia, tra il 2006 e il 2007 la variazione di tale indice è negativa, ma contenuta (-0,75%), mentre le Sofferenze aumentano del 7% circa, evidenziando una più bassa incidenza delle crediti anomali rispetto al totale.

**Tabella 2.7 Evoluzione dei crediti in sofferenza dal 2006 al 2010**

(valori costanti in migliaia di euro - Indice NIC Istat, anno base = 1995)

Variabili	2006	2007	2008	2009	2010
Sofferenze	13,200,000	14,200,000	13,000,000	17,700,000	24,500,000
(in % rispetto all'anno precedente)	-	+7.58%	-8.45%	+36.15%	+38.42%
Sofferenze/Totale Crediti	0.0128	0.0127	0.0110	0.0152	0.0204
(in % rispetto all'anno precedente)	-	-0.75%	-13.41%	+38.53%	+34.16%

Fonte: elaborazioni su dati ABI.

A tal proposito, è interessante capire quale sia la natura del nesso di causalità che lega i *non performing loans* e l'efficienza degli intermediari. Berger e De Young (1997) propongono alcune ipotesi interpretative riguardo alla relazione tra l'efficienza e il comportamento dei manager. La prima, quella del *bad management*, assume che siano le inefficienze gestionali della banca a determinare un deterioramento della qualità dei crediti. Secondo Giordano e Lopes (2008), un aumento dei livelli di efficienza di costo potrebbe essere determinato da un risparmio di costi relativi alle attività di *screening* e di *monitoring* degli impieghi con conseguenze negative sulla loro qualità in termini di rischiosità. Tale ipotesi interpretativa sarebbe a favore dell'ipotesi di endogeneità della variabile  $z_I$  (tabella 2.6). La seconda ipotesi, denominata *bad luck*, assume che le sofferenze siano determinate in maniera esogena e che dipendano da shock esterni. Pertanto, lo shock fa aumentare le sofferenze e ciò si ripercuote sull'efficienza. Si assume, quindi, che sia il contesto ambientale a determinare peggioramenti della qualità degli impieghi. Un'ulteriore tesi è quella denominata *skimping behaviour*, secondo la quale ad un aumento di efficienza segue un

aumento delle sofferenze. L'ipotesi di fondo di questo approccio è che i manager scelgano di minimizzare i costi di breve periodo tralasciando la qualità dei crediti di lungo periodo.

Al fine di tenere conto della posizione di rischiosità in cui si trova l'istituto di credito e dell'effetto che ciò potrebbe avere sugli *score* di efficienza, si introduce nell'equazione (2.12) un indicatore di solvibilità delle banche ( $z_2$ ), calcolato come il rapporto tra il patrimonio di vigilanza e le attività ponderate per il rischio<sup>39</sup>. Esso rappresenta una misura di adeguatezza patrimoniale delle banche in relazione al rischio di credito da esse sostenuto.

La variabile "*Cost/Income*" ( $z_3$ ) è stata introdotta come variabile di controllo. Essa viene spesso utilizzata in letteratura (Weill, 2004) come indicatore di performance degli istituti di credito. Valori elevati di tale rapporto indicano più basse performance dell'istituto di credito dovute ad un'incidenza maggiore dei costi sul risultato della gestione caratteristica. Il segno atteso del coefficiente legato a tale variabile va nella direzione di misurare un effetto positivo sull'inefficienza, ovvero per valori più alti dell'indice ci si aspetta un più elevato livello di inefficienza.

Ulteriore variabile inserita è rappresentata da una proxy della concentrazione del settore, ovvero l'indice di Herfindahl ( $z_4$ ), calcolato come la somma del quadrato delle quote di mercato di ciascuna azienda di credito. Si tratta di una questione già affrontata in molti lavori (Casu e Girardone, 2009; Dongili, Rossi e Zago, 2008; Fontani e Vitali, 2007) interessati a verificare se una maggiore concentrazione del settore, come quella avvenuta per il settore bancario italiano a partire dagli anni '90, possa aver influenzato i livelli di efficienza delle banche. In generale, l'esito è incerto, poiché, se da un lato, le operazioni di consolidamento hanno comportato un aumento delle dimensioni con probabili e attesi aumenti dei livelli di efficienza, dall'altro, infatti, potrebbero provocare una diminuzione della competitività delle banche. Turati (2008) propone un modello che cattura la relazione tra redditività ed efficienza. I risultati sostengono l'idea di un settore bancario non concorrenziale e, secondo l'autore, le operazioni di consolidamento comportano un aumento del potere contrattuale delle banche, sfavorevole per la clientela. Nella tabella 2.8 è riportata l'evoluzione dell'indice di Herfindahl ( $z_4$ ) calcolato per il campione in esame dal 2006 al 2010. Come si nota, in questi anni il settore bancario italiano è stato interessato da un tendenziale aumento della concentrazione (l'indice raddoppia dal 2006 al 2010, passando da 0.0277 a 0.0559).

---

<sup>39</sup> I rischi di credito considerati dalle direttive di Basilea sono il rischio operativo (il rischio di perdita dovuto ad errori nella gestione delle ordinarie attività bancarie), il rischio di mercato (il rischio di perdita per la variazione del valore per gli strumenti finanziari) e il rischio di credito (rischio di perdita per controparte insolvente).

Infine, si sono inserite alcune variabili dicotomiche allo scopo di tener conto del fatto che eventuali differenze nei livelli di efficienza di costo possano essere determinate dalla categoria giuridica, dalla localizzazione geografica e dalla dimensione degli intermediari<sup>40</sup>, mentre la variabile  $z_5$  rappresenta il *time trend* ed è stato considerato per depurare le stime dall'effetto temporale. Si è scelto di inserire un trend crescente di tipo lineare<sup>41</sup>. Esso cattura gli effetti di natura esogena sull'efficienza quali, ad esempio, quelli relativi ai cambiamenti tecnologici e alle dinamiche strutturali del settore.

**Tabella 2.8 L'indice di Herfindahl dal 2006 al 2010**

	<b>2006</b>	<b>2007</b>	<b>2008</b>	<b>2009</b>	<b>2010</b>
	0.0277	0.0420	0.0449	0.0455	0.0559
<i>Oss.</i>	686	682	689	686	648

Fonte: elaborazioni su dati ABI.

## 2.5 Le stime su dati *panel* dei livelli di efficienza di costo delle banche italiane dal 2006 al 2010

Tenendo conto delle variabili inserite nella funzione di costo, il modello da stimare diventa (cfr. § 2.4)<sup>42</sup>:

$$\log \frac{Costi_{it}}{w_{it}^r} = \beta_0 + \beta_1 \log y_{1it} + \beta_2 \log y_{2it} + \beta_3 \log y_{3it} + \gamma_1 \log \frac{w_{1it}}{w_{it}^r} + \gamma_2 \log \frac{w_{2it}}{w_{it}^r} + u_{it} + v_{it} \quad (2.13)$$

in maniera simultanea alla stima dell'equazione di inefficienza.

La tabella 2.9 mostra i risultati relativi alla funzione dei costi definita nell'equazione (2.13). I coefficienti stimati rappresentano le elasticità dei costi rispetto alle variabili inserite nel modello. Tali elasticità sono tutte positive e significative. In particolare, i costi sono più sensibili a variazioni degli impieghi tradizionali, ossia i crediti verso clienti (0.61). In tal caso, infatti, l'elasticità assume un valore più elevato rispetto ai parametri stimati per gli altri due

<sup>40</sup> Per tale motivo, si considerano variabili binarie che identificano i vari istituti a seconda della categoria giuridica, della localizzazione territoriale e della dimensione.

<sup>41</sup> Il trend è così costruito: 2006 = 1, 2007 = 2, 2008 = 3, 2009 = 4, 2010 = 5.

<sup>42</sup> Le stime sono state realizzate mediante l'uso del software *R-project*.

output, le commissioni attive e i titoli (rispettivamente, 0.19 e 0.11). Per quanto riguarda le elasticità dei costi rispetto al prezzo degli input, la funzione di costo risponde con maggiore intensità a variazione del costo del lavoro (0.51), a seguire le entità della sensibilità dei costi rispetto al prezzo del capitale (0.43) e a quello dei fondi raccolti (0.06)<sup>43</sup>.

**Tabella 2.9. Stime della funzione di costo delle banche italiane dal 2006 al 2010**

	Coefficienti	SE	z-value	p-value
Intercetta	-0.50***	0.090	-5.58	0
log $y_1$	0.61***	0.007	92.15	0
log $y_2$	0.19***	0.007	28.35	0
log $y_3$	0.11***	0.004	29.10	0
log ( $w_1/w_3$ )	0.51***	0.012	41.70	0
log ( $w_2/w_3$ )	0.43***	0.013	33.61	0
sigma2 <sup>44</sup>	71.06***	17.45	4.07	0
gamma <sup>45</sup>	0.9996***	0.0001	8984	0
Log-likelihood	165.7756			

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Una maggiore attenzione è rivolta alla tabella successiva (tabella 2.10), nella quale sono riportate le stime relative all'equazione dell'inefficienza (equazione 2.12) e dalla quale emerge l'effetto di variabili individuali e di contesto sui livelli di  $u$ .

La prima evidenza è la relazione negativa tra l'incidenza delle sofferenze ( $z_1$ ) e i livelli di inefficienza. Tale risultato, al contrario del contributo di Battaglia *et al.* (2010), in cui la relazione tra efficienza e *non-performing loans* risulta non significativa, indica che un valore più elevato di tale rapporto, ossia una minore qualità dei crediti, comporta una minore (maggiore) inefficienza (efficienza). Una possibile spiegazione di questo risultato è quella

<sup>43</sup> Il valore dell'elasticità della funzione dei costi rispetto al prezzo dei fondi raccolti è calcolato come complemento a 1 della somma tra le altre due elasticità visto che nel modello stimato (espressione 2.8) si tiene conto dell'ipotesi di omogeneità lineare dei costi rispetto al prezzo degli input (cfr. § 2.4).

<sup>44</sup>  $\sigma^2_u = \sigma^2_u + \sigma^2_v$ ; si tratta della varianza dell'errore composto, data dalla somma delle varianze delle due componenti.

<sup>45</sup>  $\gamma = \frac{\sigma^2_u}{\sigma^2}$ ; il valore zero di tale parametro indica che deviazioni dalla frontiera sono dovute solo all'errore casuale; mentre valori prossimi a uno di gamma comportano che la distanza dalla frontiera sia dovuta all'inefficienza. Esso rappresenta il parametro che, nella tecnica di Jondrow *et al.* (1982) viene usato per incorporare la componente di inefficienza (tecnica JLMS).

secondo la quale la qualità dei crediti è dipendente dalle operazioni di controllo e monitoraggio degli impieghi (Giordano e Lopes, 2006). Negli anni considerati, si è, evidentemente, assistito a scelte manageriali che hanno ridotto gli sforzi nella selezione del credito e questo si è riflesso in un maggiore tasso di insolvenza dei mutuatari. Tuttavia, il risparmio sui costi di selezione degli output potrebbe comportare aumenti dell'efficienza di costo delle banche italiane. Questi risultati sembrano essere in linea con l'ipotesi di *bad management* (Berger e De Young, 1997; Giordano e Lopes, 2008; Williams, 2004)<sup>46</sup>. A tal proposito, secondo Giordano e Lopes (2006), un aumento dei livelli di efficienza di costo potrebbe essere determinato da risparmi di costi relativi alle attività di *screening* e di *monitoring* degli impieghi con conseguenze negative sulla loro qualità in termini di rischiosità.

Risulta anche interessante evidenziare i risultati relativi alla relazione tra l'indice di solvibilità ( $z_2$ ) e l'inefficienza di costo. Valori alti di questo indice, ovvero un minor rischio cui sono soggette le banche (cfr. § 2.4), indicano una maggiore adeguatezza patrimoniale. La stima del coefficiente associato all'indice di solvibilità è pari a -2.07 ed indica che in presenza di un aumento di tale indice, gli *score* di efficienza seguono lo stesso andamento. Si tratta di un importante risultato poiché le banche, seguendo le direttive europee, registrano livelli di patrimonializzazione più alti per far fronte ai rischi operativi, di mercato e di credito, e ciò determina un miglioramento nei livelli di efficienza di costo.

Un effetto negativo deriva, invece, dal rapporto *Cost/Income* ( $z_3$ ), infatti, dalle stime ottenute, risulta che minore è il valore espresso da questo indicatore, maggiore è l'efficienza gestionale della banca. Pertanto, a valori più elevati di tale indice (segnale di scarsa efficienza gestionale), corrispondono più elevati *score* di inefficienza di costo.

Per quanto riguarda la relazione tra la concentrazione del mercato ( $z_4$ ) e l'efficienza, si noti che il relativo coefficiente è negativo, ed indica, pertanto, che, a parità di altre condizioni, una maggiore concentrazione consente di raggiungere più elevati livelli di efficienza di costo<sup>47</sup>. Facendo anche riferimento all'indice di Herfindahl (tabella 2.8), si può sostenere che nel corso del periodo considerato in questo studio, i vantaggi derivanti dalla creazione di grandi banche per sfruttare economie di scala hanno prevalso sui potenziali effetti negativi per

---

<sup>46</sup> In Appendice A1 (tabelle A1.1 e A1.2) vengono riportati i risultati relativi al test di Granger implementato per indagare sulla relazione tra l'efficienza e la qualità dei crediti delle banche italiane. Il test ha prodotto risultati a favore dell'ipotesi di *bad management*. Infatti, le stime dei modelli ausiliari necessarie per il test di Granger mostrano che il valore di significatività effettiva della statistica utilizzata per condurre il test evidenzia una relazione di simultaneità inversa tra i livelli di efficienza e le sofferenze.

<sup>47</sup> Simili risultati sono stati raggiunti in Fontani e Vitali (2007) e, tuttavia, sembrano in contrasto con quelli raggiunti in Turati (2008), secondo il quale una maggiore concentrazione del settore comporta livelli di efficienza di costo più bassi.

la concorrenza derivanti da una maggiore concentrazione del settore (Fontani e Vitali, 2007)<sup>48</sup>.

I coefficienti stimati per le dummy relative alla forma giuridica ( $d_{spa}$  e  $d_{pop}$ ) indicano che i migliori livelli di efficienza sono registrati dalle BCC. Inoltre, seguendo la classificazione dimensionale della Banca d'Italia, si ottiene che, rispetto agli istituti di credito classificati come banche piccole, quelle minori, medie e grandi raggiungono risultati più soddisfacenti in termini di efficienza di costo, mentre il contrario si ha nel caso delle banche maggiori. Infine, dalla tabella 2.10 si può osservare che, rispetto alle banche del Nord-Est, quelle localizzate in altre aree tendono a registrare risultati meno soddisfacenti in termini di efficienza di costo.

**Tabella 2.10. Stime dell'equazione di inefficienza delle banche italiane dal 2006 al 2010**

	Coefficienti	SE	z-value	p-value
$z_1 = bad\ loans$	-13.94***	3.18	-4.38	0
$z_2 = indice\ di\ solvibilità$	-2.07***	0.50	-4.13	0
$z_3 = Cost/Income$	3.19***	0.81	3.92	0
$z_4 = indice\ di\ Herfindahl$	-26471***	6417.3	-4.12	0
$z_5 = time\ trend$	154.43***	37.41	4.13	0
$d_{spa}$	271.28***	65.69	4.13	0
$d_{pop}$	290.53***	71.50	4.06	0
$d_{minori}$	-133.78***	32.69	-4.09	0
$d_{medie}$	-105.33***	25.19	-4.18	0
$d_{grandi}$	-125***	28.59	-4.37	0
$d_{maggiori}$	52.23***	14.41	3.63	0
$d_{nordovest}$	114.48***	27.65	4.14	0
$d_{centro}$	34.89***	9.30	3.75	0
$d_{sud}$	43.48***	11.72	3.71	0

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

### 2.5.1 Livelli medi di efficienza di costo delle banche italiane

Dopo aver presentato le stime del modello di efficienza di costo, si riportano i risultati relativi ai livelli medi di efficienza stimati dal 2006 al 2010, per l'intero campione e per alcuni

<sup>48</sup> Si consideri che i valori dell'indice di Herfindhal sono cinque, ovvero uno ogni anno, pertanto, non è sorprendente che la stima dello *standard error* del coefficiente associato a tale variabile possa essere distorta (Lensink e Meesters, 2012).

gruppi di banche. La tabella 2.11 riporta alcune statistiche descrittive relative alla stima dei livelli di efficienza di costo per il 2010<sup>49</sup>.

Dalla tabella emerge che il 50% delle banche, nel periodo considerato, raggiunge livelli di efficienza di costo superiori o uguali a 0.9125 e che solo l'1% del campione ottiene *score* al più pari a 0.4038. Inoltre, si registra un'asimmetria negativa e un indice di curtosi molto elevato: la distribuzione dell'efficienza è, quindi, leptocurtica.

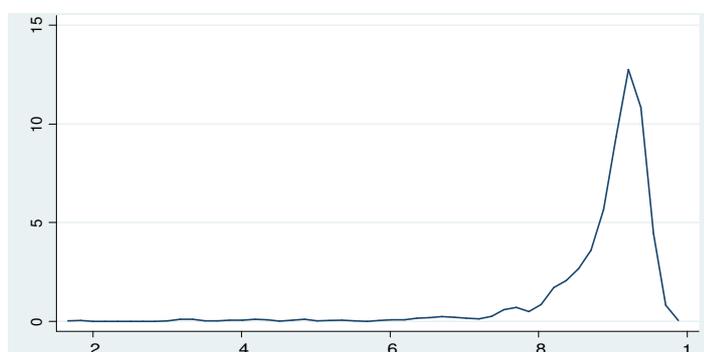
**Tabella 2.11. Alcune statistiche descrittive relative ai livelli di efficienza di costo nel 2010**

	<b>Percentili</b>	<b>Valori minori</b>		
1%	0.4038	0.1769		
5%	0.7494	0.3144	<b>Osservazioni</b>	572
10%	0.8180	0.3267		
25%	0.8766	0.3408	<b>Media</b>	0.8855
50%	0.9125			
		<b>Valori maggiori</b>	<b>DS</b>	0.0917
75%	0.9305	0.9655		
90%	0.9435	0.9667	<b>Varianza</b>	0.0084
95%	0.9500	0.9686	<b>Simmetria</b>	-3.8793
99%	0.9631	0.9779	<b>Curtosi</b>	21.9023

Fonte: elaborazioni su dati ABI.

Gli stessi risultati si ottengono dalla figura 2.4, in cui è riportata la distribuzione dei livelli di efficienza stimati per il 2010. La densità mostra una distribuzione leptocurtica intorno alla media (0.8855).

**Figura 2.4. La densità dei livelli di efficienza di costo nel 2010**



Fonte: elaborazioni su dati ABI (Software: Stata 10).

<sup>49</sup> Il valore delle statistiche e la distribuzione delle stime dell'efficienza negli altri anni analizzati non si discostano sostanzialmente da quanto emerge per il 2010 (i dati sono disponibili su richiesta).

Nella tabella 2.12 sono riportati i valori medi di efficienza di costo classificati secondo la categoria giuridica e le dimensioni delle banche. Il livello medio dell'efficienza di costo delle banche incluse nel campione è pari a 0.8838 nel 2006, aumenta leggermente nel 2007 (0.8884), mentre nel 2008 e nel 2009 registra una diminuzione. In questi anni, il livello medio di efficienza è pari, rispettivamente, a 0.8717 e 0.8313. Nel 2010, le banche italiane registrano un livello di efficienza pari a 0.8855, più elevato rispetto ai due anni precedenti.

Differenziando per categoria giuridica, si rileva che le BCC ottengono le migliori performance per tutto il periodo considerato. Si tratta di un risultato apparentemente inatteso visto che i processi di concentrazione sono stati affiancati da trasformazioni delle banche in società per azioni, al fine di aumentare la dimensione, migliorare la gestione con una spinta radicalizzazione agli obiettivi di profitto e, quindi, raggiungere livelli di efficienza più elevati. Tuttavia, questo è un risultato ottenuto da vari studi (Giordano e Lopes, 2006; Girardone, Molyneux e Gardener, 2004), per esempio, quello proposto da Giordano e Lopes (2008) in cui si mostra che, anche nel caso dell'efficienza di profitto, le Spa presentano stime dei livelli medi inferiori a quelli ottenuti dalle BCC.

**Tabella 2.12. Livelli medi di efficienza di costo delle banche italiane per forma giuridica e dimensione**

	2006	2007	2008	2009	2010
Intero campione	0.8838	0.8884	0.8717	0.8713	0.8855
Spa	0.8201	0.8289	0.8017	0.8103	0.8249
Pop	0.8213	0.8589	0.8060	0.7966	0.8612
Bcc	0.9151	0.9169	0.9098	0.9042	0.9120
Minori	0.9078	0.9101	0.9028	0.8958	0.8997
Piccole	0.8117	0.8207	0.7794	0.7963	0.8423
Medie	0.8436	0.8645	0.8099	0.8287	0.8455
Grandi	0.8007	0.8527	0.7752	0.8464	0.8577
Maggiori	0.7304	0.7572	0.6638	0.6931	0.7441

Fonte: elaborazioni su dati ABI.

In merito alla classificazione dimensionale, le banche minori, che rappresentano il 70% circa del campione, registrano i più elevati livelli di efficienza. Anche questo è un risultato già noto (Giannola *et al.*, 1997; Molyneux, 2003). All'aumentare delle dimensioni, i livelli di efficienza di costo si riducono fino ad arrivare ai livelli più bassi registrati dalle banche maggiori nel 2008 e nel 2009 (0.6638 e 0.6931, rispettivamente). La questione della relazione tra efficienza di costo e dimensioni è molto dibattuta in letteratura e non è da escludere che a questi risultati possa concorrere anche un problema di misurazione degli *score* di efficienza delle banche più grandi (Inanoglu, Jacobs e Sickles, 2010).

Infine, la tabella 2.13 riporta i livelli medi di efficienza di costo solo per le BCC. L'attenzione è rivolta unicamente alle BCC a causa dello specifico aspetto territoriale che caratterizza la loro operatività (Battaglia *et al.*, 2010). Come già accennato, questo tipo di banche raggiunge i livelli di efficienza più alti rispetto alle altre categorie giuridiche, seguendo la stessa dinamica osservata per l'intero campione di banche (tabella 2.12). Inoltre, le BCC del Centro e del Nord-Est sono quelle che mostrano i più alti livelli di efficienza, con una netta superiorità delle seconde nell'ultimo anno (tabella 2.13). Questo risultato sulla specificità delle BCC è analogo a quello ottenuto da Giannola *et al.* (1997).

**Tabella 2.13. Livelli di efficienza di costo per area geografica (Solo BCC)**

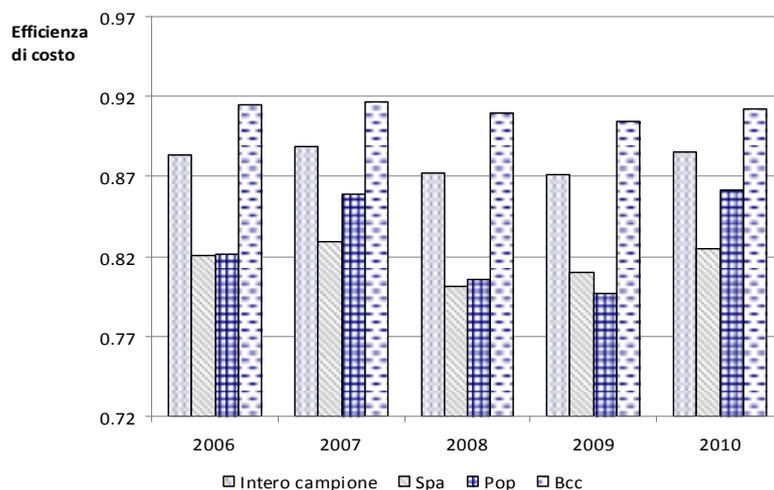
	2006	2007	2008	2009	2010
BCC	0.9151	0.9169	0.9098	0.9042	0.9120
Bcc NO	0.8902	0.8924	0.8892	0.8806	0.8914
Bcc NE	0.9229	0.9213	0.9186	0.9187	0.9277
Bcc Centro	0.9262	0.9270	0.9098	0.9129	0.9175
Bcc Sud	0.9075	0.9147	0.9065	0.8861	0.8913

Fonte: elaborazioni su dati ABI.

Un risultato inatteso, invece, è relativo al fatto che le BCC del Sud, pur raggiungendo livelli nettamente inferiori ai livelli medi, sembrano più efficienti di quelle del Nord-ovest, le quali presentano i più bassi livelli di efficienza (tabella 2.13 e figura 2.5). Si tratta di un

risultato diverso da quello mostrato da Barra *et al.* (2011), i quali, però, stimano l'efficienza mediante la DEA<sup>50</sup>.

**Figura 2.5. Livelli medi di efficienza di costo per l'intero campione e per categoria giuridica**



Fonte: elaborazioni su dati ABI.

## 2.6 Prove di robustezza

In questo paragrafo, si propone una verifica di robustezza dei risultati ottenuti sui dati panel confrontando le distribuzioni dei livelli di efficienza di costo stimati e il livello medio di efficienza, quando dal campione si escludono le banche che presentano i valori inferiori al primo percentile e quelli maggiori del 99-esimo percentile della distribuzione di  $\log \frac{Cost_{it}}{w_{it}^r}$ , che rappresenta la variabile dipendente della funzione di costo.

**Tabella 2.14. Livelli medi di efficienza di costo delle banche italiane**

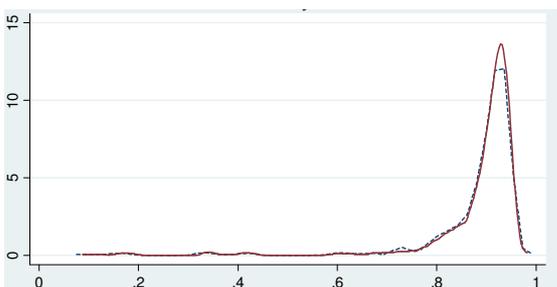
	2006	2007	2008	2009	2010
Intero campione	0.8838	0.8884	0.8717	0.8713	0.8855
Sottocampione	0.8876	0.8927	0.8756	0.8739	0.8880

Fonte: elaborazioni su dati ABI.

<sup>50</sup> Occorre osservare che Barra *et al.* (2011) stimano una frontiera di costo per le BCC e una per tutte le altre banche. Pertanto, i risultati sembrano essere sensibili non solo al tipo di scelta riguardo alle metodologie utilizzate, ma anche perché è diversa la definizione del *benchmark*. In un tale scenario, è plausibile che esistano differenze nei risultati ottenuti.

Nella tabella 2.14 sono riportati i livelli medi di efficienza ottenuti per l'intero campione e per il sottoinsieme di banche estratto nel modo appena descritto. Come si può notare, le differenze in media sono marginali. Tale risultato indica che le stime dell'efficienza sono abbastanza robuste in media e tali indicazioni sono confermate dall'analisi delle funzioni di densità dei livelli di efficienza (figure 2.6-2.10). Emerge una sostanziale sovrapposizione delle funzioni di densità dei due campioni di banche considerati.

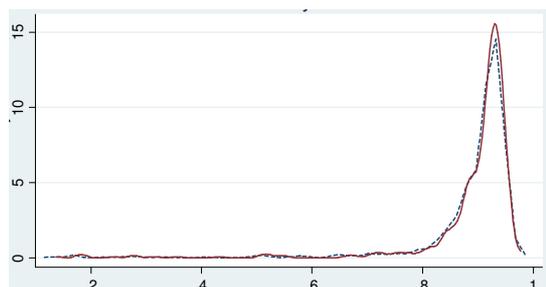
**Figura 2.6. Densità dei livelli di efficienza nel 2006**



Legenda: ---- intero campione; — sottocampione.

Fonte: elaborazioni su dati ABI.

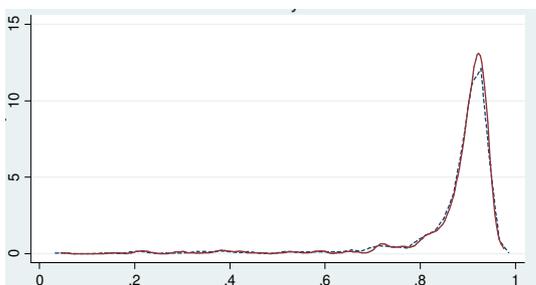
**Figura 2.7. Densità dei livelli di efficienza nel 2007**



Legenda: ---- intero campione; — sottocampione.

Fonte: elaborazioni su dati ABI.

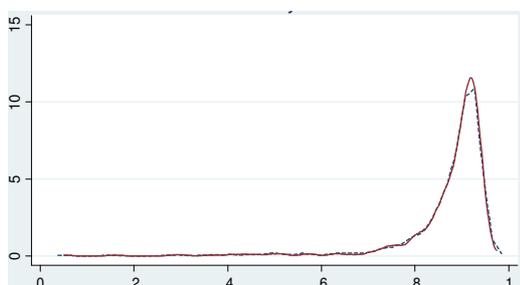
**Figura 2.8. Densità dei livelli di efficienza nel 2008**



Legenda: ---- intero campione; — sottocampione.

Fonte: elaborazioni su dati ABI.

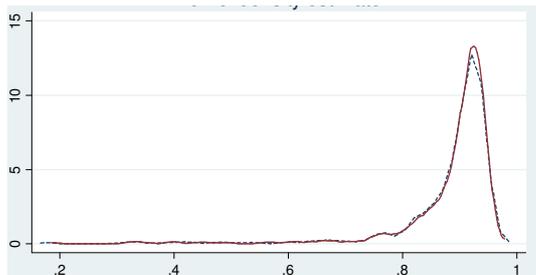
**Figura 2.9. Densità dei livelli di efficienza nel 2009**



Legenda: ---- intero campione; — sottocampione.

Fonte: elaborazioni su dati ABI.

**Figura 2.10. Densità dei livelli di efficienza nel 2010**



Legenda: ---- intero campione; — sottocampione.

Fonte: elaborazioni su dati ABI.

## 2.7 Conclusioni

Questo studio propone un'analisi dell'efficienza delle banche italiane nel periodo 2006-2010 mediante l'approccio delle frontiere stocastiche. Nell'esaminare un periodo molto recente, sono stati utilizzati i dati di bilancio delle banche italiane forniti dall'Associazione Bancaria Italiana (ABI) imputati seguendo i principi contabili internazionali (IAS). L'importanza di tale aspetto risiede nel fatto che gli IAS sono volti a migliorare l'informazione e la trasparenza. Essi, infatti, consentendo una rappresentazione più realistica dei rischi di credito e di mercato, contribuiscono ad accrescere l'efficienza e, quindi, la stabilità bancaria (Ciocca, 2004).

Le stime sono state effettuate seguendo la specificazione proposta da Battese e Coelli (1995). In questo modo, si stimano simultaneamente la frontiera di costo e l'equazione dell'inefficienza. Inoltre, tale specificazione consente di ottenere le stime dei coefficienti relativi ai fattori individuali e di contesto che influenzano l'inefficienza. Inoltre, utilizzare tale specificazione consente, di tener conto dell'eterogeneità nel campione simultaneamente alla stima della funzione di costo in modo da evitare i problemi di distorsione legati all'orientamento “two steps” che rappresenta l'alternativa all'approccio proposto (Battaglia *et al.*, 2010; Greene, 1993).

Il principale risultato è che nel 2008 si è verificata una notevole riduzione dei livelli medi di efficienza delle banche incluse nel campione e che questa dinamica è comune in tutte le singole categorie giuridiche. Inoltre, si è mostrato che, per tutti gli anni compresi nell'analisi, le BCC hanno raggiunto i livelli più alti di efficienza trainando, quindi, l'evoluzione dell'intero settore. Si evidenzia anche che le banche piccole e quelle minori ottengono i risultati più soddisfacenti e le BCC del Centro e del Nord-Est risultano essere più efficienti rispetto alle BCC della altre aree del paese.

Per quanto concerne le determinanti dell'efficienza, emerge che una minore qualità dei crediti comporta maggiore efficienza di costo. E' un risultato in linea con l'ipotesi di *bad management* che è anche empiricamente supportato dal test di Granger. Il risparmio di costo che si ottiene riducendo alcune tipologie di spese comporta, infatti, guadagni di efficienza di costo, ma, allo stesso tempo, le banche selezionano progetti più rischiosi che presentano maggiore probabilità di tramutarsi in sofferenze.

Inoltre, una maggiore solvibilità degli istituti di credito e una maggiore concentrazione del settore determinano il raggiungimento di più elevati livelli di efficienza di costo. Infine, le banche che presentano una scarsa efficienza gestionale, evidenziata da livelli più elevati del rapporto *Cost/Income*, registrano livelli di efficienza di costo più bassi.

Lo studio dei livelli di efficienza secondo la specificazione di Battese e Coelli (1995) contribuisce ad indirizzare le scelte strategiche, di *corporate* e operative delle banche allo scopo di tenere sotto “controllo” i fattori che determinano le performance degli istituti di credito. A tal fine, diventa importante specificare modelli econometrici che consentono di ridurre problemi di distorsione ed errata misurazione del fenomeno. La caratterizzazione dell’applicazione dell’approccio SFA al caso delle aziende di credito presente in questo lavoro può essere considerata il punto di partenza per introdurre alcuni miglioramenti legati alla specificazione delle frontiere stocastiche. In particolare, è ragionevole pensare, in primo luogo, che l’errore casuale e l’inefficienza si distribuiscano secondo variabili casuali più flessibili e, in tal modo, si rilassa l’ipotesi di normalità che potrebbe essere un caso particolare di un modello più generale. Inoltre, appare opportuno introdurre l’ipotesi di dipendenza tra le due componenti di errore di una frontiera stocastica, poiché da un punto di vista strettamente statistico questa ipotesi incorpora come caso specifico l’ipotesi di indipendenza. Un metodo che permette di incorporare nell’approccio delle frontiere stocastiche sia distribuzioni diverse dalla Normale sia ipotesi di dipendenza anche non lineare è rappresentato dalle funzioni copule che rappresenta uno stimolante argomento di ricerca già avviato, peraltro, grazie al contributo di Smtih (2008). Nel prossimo capitolo, infatti, si sottopone a verifica empirica la dipendenza tra errore casuale e componente di inefficienza mantenendo funzioni di distribuzione marginali Normali. Dopo aver verificato che non si può escludere a priori tale ipotesi, la quale incorpora il caso di indipendenza, si propone un approccio di tipo metodologico per la costruzione di un modello più generale che tenga conto dell’ipotesi di dipendenza tra i termini di errore e marginali più flessibili rispetto a quelle Normali. Come si vedrà, tale approccio fornisce alcune importanti evidenze, quali, ad esempio, l’influenza dei parametri di forma delle distribuzioni marginali e del parametro di associazione della copula sulla scomposizione del termine di errore composto e, quindi, sulla stima dei livelli medi di efficienza di costo.

## APPENDICE A1

### Il test di Granger

**Tabella A1.1 Test di Granger: efficienza di costo *versus* sofferenze**

CE	OLS	FE	RE
intercetta	0.1918***	1.0930***	0.2175***
CE <sub>t-1</sub>	0.5885***	-0.1191***	0.4595***
CE <sub>t-2</sub>	0.0935***	-0.1674***	0.1575***
CE <sub>t-3</sub>	0.0964	0.0997*	0.1339***
BL <sub>t-1</sub>	-0.0021	0.0065**	0.0002
BL <sub>t-2</sub>	0.0007	0.0034	0.0003
BL <sub>t-3</sub>	-0.0004	-0.0005	-0.0016
Test di Wald (p-value)	0.5255	0.0154	0.8052

**Tabella A1.2 Test di Granger: sofferenze *versus* efficienza di costo**

BL	OLS	FE	RE
intercetta	-1.6104***	-0.7698	-1.7920***
CE <sub>t-1</sub>	-0.0451	-0.8294*	0.0691
CE <sub>t-2</sub>	0.0774	-1.8413***	-0.0060
CE <sub>t-3</sub>	1.1533***	0.0688	1.2511***
BL <sub>t-1</sub>	0.7543***	0.1560***	0.6238***
BL <sub>t-2</sub>	0.1118***	0.1080*	0.2077***
BL <sub>t-3</sub>	-0.0538 .	-0.7698	-0.0403 .
Test di Wald (p-value)	0	0.0002	0

Livelli di significatività: '\*\*\*' = 0.01; '\*\*' = 0.001; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Il test di Granger è stato realizzato mediante stime pooled OLS, effetti fissi ed effetti casuali. Nella tabella A1.1, sono riportate le regressioni in cui la variabile dipendente è il livello di efficienza di costo stimato con la specificazione di Battesi e Coelli (1995). Il *p-value* relativo al test di Wald si riferisce alla verifica dell'ipotesi secondo la quale i tre coefficienti delle variabili *Bad Loans* ritardate ( $BL_{t-s}$ ) siano uguali a zero (con  $s=1, \dots, 3$ ). La variabile BL è rappresentata dalla variabile  $z_l$  dell'equazione (2.12). Nella tabella A1.2, sono riportate le regressioni in cui la variabile dipendente è la misura della qualità dei crediti. Il *p-value* relativo al test di Wald si riferisce alla verifica dell'ipotesi secondo la quale i tre coefficienti delle variabili *Cost Efficiency* ritardate ( $CE_{t-s}$ ) siano uguali a zero (con  $s=1, \dots, 3$ ).

I risultati ottenuti mediante il metodo OLS e ad effetti casuali hanno condotto all'ipotesi di *skimping behaviour*, secondo la quale emerge una relazione positiva tra le due variabili in esame nella direzione verso cui i livelli di efficienza influenzano il livello delle sofferenze. Tuttavia, il test di Hausman ha indirizzato la scelta verso il modello a effetti fissi. I risultati del test di Granger, in tal caso, evidenziano una simultaneità inversa tra i livelli di efficienza e le sofferenze a favore dell'ipotesi di *bad management*.

## APPENDICE A2

### **Le stime su dati *cross section* dei livelli di efficienza di costo delle banche italiane dal 2006 al 2010**

Per completezza di analisi, si propongono le stime *cross section* che si riferiscono all'efficienza di costo delle banche italiane incluse nel campione dal 2006 al 2010, sebbene tali risultati implicino aspetti critici e problemi che, in parte, sono risolvibili mediante analisi *panel*.

Come più volte detto, il settore bancario italiano è stato caratterizzato da un numero notevole di operazioni di concentrazione, tra cui fusioni e incorporazioni. Lavorando su dati *cross section*, si ha un campione diverso ogni anno, in quanto si considerano le banche esistenti in quel periodo temporale. Effettuando un'analisi *panel*, si potrebbe avere una perdita di informazioni nel momento in cui si verificano operazioni di concentrazione. Si consideri, infatti, una situazione in cui una banca viene incorporata in un'altra. In questo caso, da un certo anno in poi, la banca incorporata non esiste più, ma non è uscita dal mercato perché inefficiente, bensì ha adottato una politica che potenzialmente potrebbe comportare guadagni di efficienza. Pertanto, il caso della banca incorporata e il caso della banca in uscita dal mercato sono assimilati e trattati allo stesso modo. Inoltre, uno dei problemi è che non si riescono a cogliere i vantaggi che l'incorporazione dovrebbe comportare, ovvero maggiori livelli di efficienza. La stessa cosa accade nel caso delle fusioni dalle quali nasce un terzo soggetto. Si precisa che da un punto di vista tecnico, tale questione non comporta problemi perché si tratta di lavorare con un set di dati *panel* non bilanciato, ma da un punto di vista interpretativo bisogna prestare molta attenzione ai risultati.

Tuttavia, oltre ai problemi classici che si presentano quando si effettuano analisi *cross section*, risolvibili mediante analisi *panel*, come collinearità tra i regressori e distorsione da variabili omesse, emerge uno svantaggio di notevole importanza. Esso riguarda l'impossibilità di eseguire confronti fra i risultati ottenuti di anno in anno perché cambia il campione e, poiché gli istituti di credito non sono sempre gli stessi, la frontiera stimata ha come *benchmark* banche diverse secondo l'anno analizzato.

Da tutte queste considerazioni nasce l'idea di proporre in appendice i risultati ottenuti per analisi *cross section*<sup>51</sup>.

---

<sup>51</sup> Si consideri che nei lavori esaminati nel capitolo I di questo lavoro sono presenti alcuni contributi che effettuano analisi di efficienza su dati *cross section* (Behr e Tente, 2008; Giannola *et al.*, 1997).

Nelle tabelle A2.1 e A2.2 sono presenti le stime effettuate su dati cross section per tutti gli anni compresi nel periodo in esame inserendo nell'equazione dell'inefficienza la qualità dei crediti ( $z_1$ ), l'indice di solvibilità ( $z_2$ ), il rapporto *Cost/Income* ( $z_3$ ) e le variabili dicotomiche di categoria giuridica, dimensione e area geografica (Modello I)<sup>52</sup>. Nelle due successive tabelle, invece, sono riportati i risultati relativi alle stime del modello ristretto, ovvero quello in cui non si considera  $z_3$  (Modello II)<sup>53</sup>.

**Tabella A2.1 Stime della funzione di costo delle banche italiane (Modello I)**

	2006	2007	2008	2009	2010
Intercetta	-1.26***	-2.01***	-1.08***	-1.64***	-2.08***
log( $y_1$ )	0.61***	0.59***	0.52***	0.55***	0.58***
log( $y_2$ )	0.19***	0.17***	0.31***	0.21***	0.19***
log( $y_3$ )	0.13***	0.12***	0.07***	0.09***	0.08***
log( $w_1/w_3$ )	0.56***	0.59***	0.68***	0.62***	0.65***
log( $w_2/w_3$ )	0.38***	0.38***	0.23***	0.34***	0.35***
sigma2	144	0.1051***	1.3977	0.3326**	0.1243**
gamma	0.9998***	0.7541***	0.9814***	0.9222***	0.7765***
Log-likelihood	23.1357	99.1415	75.5888	77.8750	28.3549

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Come si vede nella tabella A2.1, le elasticità della funzione di costo sono statisticamente significative ed i valori stimati sono simili di anno in anno. In generale, risulta che i costi sono sempre più sensibili alle variazioni dei crediti ( $y_1$ ), i quali rappresentano l'output bancario più importante. Seguono l'incidenza delle commissioni attive ( $y_2$ ) e quella dei titoli ( $y_3$ ). Per quel che riguarda le elasticità dei costi rispetto al prezzo degli input, si evidenzia che variazioni del prezzo relativo del lavoro rispetto alla raccolta ( $w_1/w_3$ )

<sup>52</sup> Le stime sono state realizzate mediante l'uso del software *R-project*.

<sup>53</sup> Data la mancanza di una teoria di riferimento sulle determinanti dell'efficienza delle banche, sono stati stimati diversi modelli in cui si sono inserite altre variabili rispetto a quelle elencate nella tabella 2.7, come ad esempio la quota di mercato (calcolata come il rapporto tra il totale attivo di ciascuna banca e il totale attivo del settore), il numero degli sportelli ed, infine, una proxy dell'incidenza di informazioni provenienti da rating esterni (tale indice è stato calcolato come lo scarto tra l'ammontare delle due classi di rating migliori e quello delle due classi peggiori in rapporto al totale attivo di ciascuna banca). Tuttavia, le stime dei coefficienti relativi a tali variabili risultano non statisticamente significative per tutti gli anni in esame. I risultati sono disponibili su richiesta.

comportano variazioni dei costi maggiori rispetto a quelle derivanti da variazioni del prezzo relativo del capitale rispetto alla raccolta ( $w_2/w_3$ ). Infatti, i coefficienti stimati delle elasticità della funzione di costo rispetto a  $w_1/w_3$  sono pari a 0.60 circa per tutti gli anni esaminati, mentre quelli relativi a  $w_2/w_3$  registrano valori sempre inferiori a 0.40.

Maggiore attenzione è rivolta ai risultati relativi all'effetto delle variabili individuali e di contesto sui livelli di inefficienza. Dalla tabella A2.2, emerge che la significatività dei parametri stimati cambia di anno in anno e che non c'è un orientamento univoco circa l'entità di tali effetti. Non risulta che il modello stimato per gli anni 2006 e 2008 fornisca informazioni significative. Infatti, emerge che l'effetto della variabile che in letteratura viene maggiormente studiata nella relazione con l'efficienza delle banche, ovvero la qualità dei crediti, risulta significativo al 10% solo nel 2007 (Berger e De Young, 1997; Giordano e Lopes, 2008). Tuttavia, nel 2007 e nel 2009 si può notare che l'indice di solvibilità ha un effetto positivo sull'inefficienza e ciò è controintuitivo, poiché valori più alti di tale indice, che indicano una maggiore solvibilità degli istituti di credito, dovrebbero comportare migliori performance. Quest'ultimo risultato è evidente, invece, nel paragrafo 2.5, in cui sono riportate le stime realizzate sul panel di dati dal 2006 al 2010.

**Tabella A2.2 Stime dell'equazione di inefficienza delle banche italiane (Modello I)**

	2006	2007	2008	2009	2010
$z_1 = bad\ loans$	107.91	0.09 .	0.61	-0.07	-0.01
$z_2 = indice\ di\ solvibilità$	1.37	1.26***	0.81	0.03***	-0.001
$z_3 = Cost/Income$	-677.47	-0.16	1.26	-2.37**	0.01
$d\_spa$	268.64	0.22 .	-1.48	-0.07	0.02
$d\_pop$	370.98	0.29**	-1.06	1.02***	0.13
$d\_minori$	-370.68	-0.24*	-13.82	-1.23*	-0.38 .
$d\_medie$	-689.44	-0.09	-0.25	-0.03	-0.01
$d\_grandi$	-303.80	-0.02	2.10	-0.64	0.13
$d\_maggiori$	-150.78	0.44**	2.90	1.87***	0.86***
$d\_nordovest$	295.70	0.33**	-0.24	0.30	0.12
$d\_centro$	115.11	-0.09	-0.08	0.03	-0.17
$d\_sud$	-59.38	0.20*	-0.17	0.96***	0.37***
<i>Efficienza di costo media</i>	0.8792	0.8634	0.8784	0.8639	0.8251

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Si consideri che in questo modello, il coefficiente associato al rapporto *Cost/Income* risulta significativo solo nel 2009. Anche questo si tratta di un risultato inatteso in quanto valori più elevati di questa variabile indicano performance delle banche più basse, alle quali dovrebbero corrispondere valori di inefficienza più elevati.

Per quel che riguarda la classificazione giuridica, le stime sottolineano differenze fra le banche popolari e le altre categorie nel 2007 e nel 2009, evidenziando risultati migliori per le BCC e le Spa, mentre, dal lato della classificazione dimensionale, emerge che, negli stessi anni e anche nel 2010, le banche minori raggiungono livelli di efficienza più elevati rispetto alle banche piccole (che rappresentano il gruppo di controllo) e le banche maggiori quelli più bassi. Sempre negli stessi anni, le banche del Sud registrano livelli di efficienza più bassi rispetto a quelle del Nord-Est, le quali non sembrano ottenere diversi risultati rispetto alle banche del Centro e del Nord-Ovest.

Le stime dei livelli medi di efficienza variano tra il valore minimo registrato nel 2010 (0.8251) ed il valore massimo registrato nel 2006 (0.8792), che, tuttavia, è circa uguale a quello registrato nel 2008 (0.8784). Nel 2007 e nel 2009, le banche italiane registrano un livello medio di efficienza di costo pari, rispettivamente, a 0.8634 e 0.8639. Tuttavia, come già accennato, tali valori non sono confrontabili anno per anno in quanto le stime si riferiscono a campioni di banche diversi.

**Tabella A2.3 Stime della funzione di costo delle banche italiane (Modello II)**

	2006	2007	2008	2009	2010
Intercetta	-1.32***	-2.03***	-1.06***	-1.57***	-2.05***
log( $y_1$ )	0.62***	0.59***	0.51***	0.54***	0.58***
log( $y_2$ )	0.18***	0.17***	0.31***	0.21***	0.19***
log( $y_3$ )	0.13***	0.12***	0.07***	0.09***	0.08***
log( $w_1/w_3$ )	0.56***	0.59***	0.68***	0.63***	0.65***
log( $w_2/w_3$ )	0.37***	0.38***	0.23***	0.33***	0.35***
sigma2	199.29**	0.0987***	2.2674***	0.3124**	0.1270**
gamma	0.9999***	0.7383***	0.9883***	0.9163***	0.7824***
Log-likelihood	20.6567	98.8018	73.8401	63.9951	28.2389

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Nella tabella A2.3, sono riportate le stime relative al modello II che si differenzia dal modello I per l'assenza della variabile *Cost/Income*, la cui inclusione nel modello I è dipesa dalla considerazione del fatto che essa rappresenta uno degli indici maggiormente utilizzati anche dalle banche per evidenziare le loro performance. Tuttavia, emerge che solo nel 2009 il coefficiente ad essa associato è statisticamente significativo.

**Tabella A2.4 Stime dell'equazione di inefficienza delle banche italiane (Modello II)**

	2006	2007	2008	2009	2010
$z_1 = \text{bad loans}$	233.06**	0.09 .	0.79	0.06	-0.005
$z_2 = \text{indice di solvibilità}$	2.05**	1.23***	1.45	0.03***	-0.001
$d_{spa}$	309.19**	0.18 .	-2.52	-0.61	0.02
$d_{pop}$	425.40**	0.25**	-1.14	0.40	0.13
$d_{minori}$	-657.66**	-0.26**	-23.36	-2.08*	-0.38 .
$d_{medie}$	-912.59**	-0.08	0.18	0.19	-0.01
$d_{grandi}$	-490.22**	-0.02	2.67	-0.17	0.13
$d_{maggiori}$	-195.43**	0.42***	4.58	1.59**	0.87**
$d_{nordovest}$	373.92**	0.30**	-0.57	0.01	0.12
$d_{centro}$	24.15***	-0.09	-0.33	-0.01	-0.18
$d_{sud}$	-243.18**	0.16*	-0.39	0.56*	0.37***
<i>Efficienza di costo media</i>	0.8762	0.8611	0.8817	0.8580	0.8254

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Le stime relative alla frontiera di costo cambiano solo marginalmente (tabella A2.3), mentre variazioni più importanti si ottengono per le stime dell'equazione dell'efficienza (tabella A2.4). Migliorano i risultati in termini di significatività nel 2006, tuttavia la situazione rimane pressoché invariata negli altri anni. Anche sui livelli medi di efficienza di costo si evidenziano variazioni solo marginali ad eccezione del valore ottenuto per il 2009 (0.8580), leggermente diverso da quello raggiunto dal sistema bancario nel 2007 (0.8611).

Seppur si tenga conto della considerazione che la scelta tra più modelli non possa essere realizzata solo attraverso un test, viene implementato, per la sua semplicità, il test del rapporto di verosimiglianza (LR test), il cui risultato supporta la scelta del modello ridotto

(modello II) per tutti gli anni, ad eccezione del 2009, con un livello di significatività dell'1%<sup>54</sup>.

---

<sup>54</sup> Il test del rapporto di verosimiglianza viene implementato per selezionare il modello migliore tra due modelli *nested*, e quando, come in questo caso, entrambi sono stimati con il metodo di massima verosimiglianza. La statistica LR (*Log-Likelihood Ratio*) è calcolata come  $-2[\text{Log-Lik}(\text{modello II}) - \text{Log-Lik}(\text{modello I})]$  e si distribuisce secondo una variabile casuale Chi-quadro,  $\chi^2_{gdl}$ , con gradi di libertà (gdl) pari al numero di parametri extra del modello esteso. La regola di decisione prevede che se la statistica calcolata supera il valore critico presente nella tavola  $\chi^2_{gdl}$ , si rifiuta l'ipotesi nulla sotto la quale vale la restrizione sul parametro, ovvero si rifiuta il modello ristretto in favore di quello esteso.

## CAPITOLO III

### La modellizzazione della dipendenza tra le componenti d'errore delle frontiere stocastiche mediante le funzioni copule e specificazioni alternative delle componenti marginali

#### 3.1 Introduzione

In questo capitolo, si propone una generalizzazione dei modelli di frontiere stocastiche mediante la costruzione di funzioni di distribuzione più generali delle componenti dell'errore. Di seguito, si cerca di spiegare in che cosa consiste la generalizzazione a cui si fa riferimento.

Nei capitoli precedenti è emerso chiaramente che nella fase di specificazione di una frontiera stocastica *classica*, vi sono almeno due ipotesi restrittive non sempre supportate dalle evidenze empiriche. In particolare, si fa riferimento (1) alle ipotesi sulle forme funzionali delle distribuzioni marginali delle componenti stocastiche, errore casuale ed inefficienza, e (2) all'ipotesi di indipendenza tra dette componenti stocastiche. In relazione alla prima ipotesi, facendo riferimento ad una letteratura che va anche oltre le frontiere stocastiche, la normalità dell'errore casuale risulta essere più una comodità di calcolo che una ipotesi confortata dall'evidenza empirica. Diversi autori, suggeriscono l'utilizzo di distribuzioni più flessibili che possano cogliere, con uno o più parametri, situazioni di simmetria e di asimmetria (positiva o negativa) della distribuzione oppure distribuzioni platicurtiche o leptocurtiche. Tali generalizzazioni, in ogni caso, non escludono la possibilità che dette distribuzioni possano avere, sostanzialmente, andamento gaussiano (ovvero andamento campanulare, simmetrico con due flessi e unimodale con code leggere). Per fare qualche esempio di distribuzioni che rientrano in tale contesto, si considerino le Normali asimmetriche e le Logistiche generalizzate. La seconda ipotesi, relativa all'indipendenza tra le due componenti d'errore, appare fortemente restrittiva in quanto riguarda solo un caso molto particolare del modo in cui possono interagire due variabili casuali. Inoltre, si vedrà, nell'applicazione presentata in questo capitolo, che anche da un punto di vista empirico vi sono situazioni in cui la dipendenza, stimata attraverso un parametro, non può essere rigettata da un punto di vista statistico. In effetti, come si è evidenziato nei capitoli precedenti, l'ipotesi di indipendenza semplifica molto la specificazione della distribuzione congiunta tra le componenti stocastiche della frontiera in quanto, sotto tale ipotesi, detta distribuzione coincide con la fattorizzazione delle distribuzioni marginali. Anche in tale situazione, si evidenzia la necessità di specificare distribuzioni congiunte molto flessibili sia per modellare

forme di dipendenza anche non lineari - che non escludono la possibilità di avere come caso particolare l'indipendenza - sia che presentino distribuzioni marginali con le caratteristiche descritte in precedenza. Per meglio descrivere la problematica che si vuole affrontare, si fa riferimento ad un "controesempio" evidenziando che spesso in letteratura, in un contesto più generale, al fine di tener conto di eventuali dipendenze, si utilizza la Normale multivariata. Come è noto, vi sono due conseguenze immediate dell'utilizzo di tale ipotesi, cioè: a) le distribuzioni marginali sono Normali e b) l'unica forma di dipendenza catturata da tale modello è quella lineare. Evidentemente, sotto tale modello non si ha la possibilità di specificare distribuzioni marginali diverse dalla Normale e dipendenze diverse da quella lineare.

In questo lavoro di tesi si utilizza la cosiddetta funzione copula che appare essere lo strumento metodologico più appropriato all'obiettivo di specificare in modo flessibile sia le distribuzioni marginali delle componenti d'errore sia la struttura di dipendenza sottostante dette componenti. Infatti, le copule consentono di costruire distribuzioni bivariate (multivariate) specificando separatamente le distribuzioni marginali dalla forma della dipendenza (Nelsen, 1999).

L'obiettivo è, quindi, duplice. Da un lato, si considerano funzioni di distribuzione diverse da quelle comunemente utilizzate in letteratura (ossia funzioni Normali e Normali troncate), dall'altro lato, si introduce nel modello l'ipotesi di dipendenza fra le due componenti di errore.

Come già detto, l'obiettivo è di utilizzare le funzioni copule per introdurre nella frontiera stocastica funzioni di distribuzione marginali più flessibili di quelle Normali e, in tale direzione, ottenere un migliore adattamento del modello ai dati, catturando in modo più adeguato le caratteristiche della distribuzione dell'errore. Ad esempio, alcune distribuzioni, come la distribuzione Log-Dagum, presentano curtosi e asimmetria, differenziandosi dalle funzioni Normali che presentano simmetria e curtosi nulla. Nel paragrafo 3.5 si dimostra come le caratteristiche proprie delle marginali influenzino la varianza dell'errore composto  $\varepsilon$ , utilizzata nell'ambito del metodo delle frontiere stocastiche per la scomposizione di  $\varepsilon$  nelle due componenti, una casuale e una attribuibile all'inefficienza. In particolare, si deriva che il parametro che governa la forma della distribuzione Log-Dagum, assegnata al termine di errore casuale, entra nell'espressione della varianza dell'errore composto. Pertanto, l'utilizzo di distribuzioni che non riescono a catturare tali caratteristiche quando, al contrario, sono presenti nei dati è causa di una cattiva specificazione del modello con l'implicazione che,

alterando la scomposizione dell'errore composto, si ottengono stime distorte dei livelli di efficienza.

In letteratura, il lavoro che introduce la dipendenza tra le due componenti di errore di una frontiera stocastica è quello proposto da Smith (2008), il quale considera distribuzioni relativamente semplici, una distribuzione Esponenziale negativa per la componente di inefficienza e una distribuzione Logistica Standard per l'errore casuale. Quest'ultima, in particolare, non cattura elementi di base delle distribuzioni, quali, per esempio, l'asimmetria. Tuttavia, il lavoro di Smith (2008) è fortemente innovativo rispetto alla letteratura di riferimento, poiché mostra la presenza di importanti effetti sulle stime delle frontiere di costo derivanti dalla dipendenza tra le componenti d'errore. Questo risultato consente all'autore di giustificare alcuni limiti noti di tale approccio che, al contrario, altri autori (Kumbhakar e Lovell, 2000) attribuiscono ad errate specificazioni dei modelli.

Da un punto di vista economico, l'importanza di non escludere *a priori* la dipendenza dell'errore casuale dalla componente di inefficienza risiede nel fatto che shock inattesi potrebbero influenzare entrambe le componenti. In particolare, variabili non introdotte né nella funzione di costo né nell'equazione dell'inefficienza (ad esempio, l'abilità dei direttori di filiale perché non si riesce a misurarla) e che hanno influenza sui costi, anche indirettamente perché influenzano l'efficienza, sono incluse nel termine di errore. Il vantaggio della frontiera stocastica è che separa l'errore casuale dalla componente di inefficienza e l'utilità di tale metodo risiede nel fatto che variabili come l'abilità dei direttori di filiale possano confluire nella parte di errore attribuita all'inefficienza. Tuttavia, non si ha la certezza che ciò avvenga, pertanto, non si può escludere a priori che entrambe le componenti di errore siano influenzate contemporaneamente da tali effetti dai quali potrebbe derivarne la presenza di dipendenza. Data l'impossibilità di prevedere dinamiche di questo genere, è teoricamente importante considerare l'ipotesi di dipendenza tra le due parti dell'errore.

Il capitolo è organizzato come segue. Il paragrafo 3.2 illustra alcune nozioni di base della copula bivariata. Il paragrafo successivo introduce l'utilizzo delle funzioni copule nell'ambito delle frontiere stocastiche ed evidenzia le problematiche sottostanti la costruzione di tali frontiere. Nel paragrafo 3.4 si presenta la costruzione della funzione di distribuzione dell'errore composto mediante la copula FGM (Farlie-Gumble-Morgenstern) per modellare la dipendenza fra le componenti erratiche distribuite normalmente. Nel sottoparagrafo 3.4.1 è riportata la trattazione empirica di tale modello. I risultati empirici ottenuti su dati reali, già utilizzati nei capitoli precedenti, confermano la necessità di considerare la dipendenza tra le componenti stocastiche delle frontiere stocastiche, ma anche l'esigenza di indagare

ulteriormente sul ruolo della dipendenza e della forma delle distribuzioni marginali. Nei paragrafi 3.5 e 3.6 si presentano alcuni risultati metodologici relativi all'impatto della dipendenza e della forma delle distribuzioni marginali sulle stime delle frontiere stocastiche. In particolare, nel paragrafo 3.5 si propone una frontiera stocastica nella quale la dipendenza è modellata tramite una copula FGM, la componente di inefficienza è modellata tramite una distribuzione Esponenziale negativa e, infine, l'errore casuale è modellato tramite una distribuzione Logistica Generalizzata di I tipo (riparametrizzazione della distribuzione Log-Dagum). Si evidenzia che quest'ultima distribuzione presenta come caso particolare la distribuzione Logistica Standard e, quindi, il modello proposto in questo paragrafo risulta essere una generalizzazione del modello di Smith (2008). Il risultato che si ritiene essere quello più importante è relativo all'impatto della dipendenza e del parametro di forma della distribuzione dell'errore accidentale sulla varianza della componente dell'errore composto. Nel paragrafo successivo, si riportano i risultati relativi ad una ulteriore generalizzazione in cui nella specificazione utilizzata nel paragrafo 3.5, invece della Esponenziale negativa, si utilizza una distribuzione Log-Dagum troncata nello zero.

### 3.2 La Copula bivariata

Una copula bivariata (o 2-copula o semplicemente copula) è una funzione di due variabili che ha come supporto la cosiddetta *unit square*  $\mathbf{I}^2 = [0,1] \times [0,1]$  e immagine nell'intervallo  $[0,1]$ , ovvero:

$$C: [0,1] \times [0,1] \rightarrow [0,1]$$

Si tratta cioè di una funzione di distribuzione bivariata con margini uniformi (0,1).

In generale, una copula presenta le seguenti proprietà:

1.  $C(u,v)$  è una funzione non decrescente in ogni argomento  $u,v$ ;
2. per ogni  $u,v \in [0,1]$ , vale che

$$C(u,0) = C(0,v) = 0, \quad C(u,1) = u, \quad C(1,v) = v$$

3. per ogni  $u_1, u_2, v_1, v_2 \in [0,1]$  tali che  $u_1 \leq u_2, v_1 \leq v_2$ ,

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$$

Le funzioni copule sono ampiamente utilizzate per costruire funzioni di distribuzione e, a tal proposito, si fa riferimento al Teorema di Sklar (Nelsen, 1999), che può essere così sintetizzato:

sia  $F_{XY}(x,y)$  una funzione di distribuzione congiunta con marginali  $F(x)$  e  $G(y)$ , allora esiste una copula  $C$  tale che per ogni  $x, y \in \mathfrak{X}$ ,

$$F_{XY}(x,y) = C(F(x), G(y)) \quad (3.1)$$

Se  $F(x)$  e  $G(y)$  sono continue allora  $C$  è unica.

Inoltre, risulta utile il seguente corollario:

sia  $F_{XY}(x,y)$  una distribuzione bivariata di due variabili  $X, Y$  con distribuzioni marginali  $F(x)$  e  $G(y)$  continue. Siano  $F^{-1}$  e  $G^{-1}$  le inverse delle distribuzioni marginali. Allora esiste un'unica copula  $C: [0,1] \times [0,1] \rightarrow [0,1]$  tale che  $\forall (u,v) \in [0,1] \times [0,1]$

$$C(u,v) = F_{XY}(F_x^{-1}(u), G_y^{-1}(v)) \quad (3.2)$$

La (3.2) informa che ogni funzioni di ripartizione bivariata possiede una copula, unica nel caso di variabili casuali continue, e può essere determinata conoscendo le funzioni inverse delle distribuzioni marginali. La (3.1) afferma che, nel caso di variabili casuali continue, esiste un'unica copula associata ad una data funzione di ripartizione bivariata. D'altra parte, la (3.1) indica anche che date due funzioni di ripartizione marginali,  $F(x)$  e  $G(y)$  ed una copula  $C(u,v)$ , la posizione  $u = F(x)$  e  $v = G(y)$  definisce una funzione di ripartizione bivariata. Di conseguenza, specificate le distribuzioni marginali  $F(x)$  e  $G(y)$ , la copula  $C(u,v)$  viene utilizzata per definire/stabilire il "legame" tra le funzioni di ripartizione marginali. Da qui il termine "copula" (*link*, legame). A questo punto, risulta evidente la separazione esistente tra la fase di specificazione delle distribuzioni marginali e la fase di specificazione della dipendenza.

In seguito, si utilizzerà la cosiddetta *density copula*, cioè la densità bivariata della copula  $C(u,v)$ , definita nel seguente modo:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} \quad (3.3)$$

Differenziando l'espressione riportata nella (3.1) rispetto a  $x$  ed  $y$ , si ottiene la funzione di densità congiunta delle variabili casuali  $X$  ed  $Y$ , cioè:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = c(F(x), G(y)) f(x) g(y) \quad (3.3\text{-bis})$$

Detto che la copula di indipendenza è data  $C(u,v)=uv$ , dalla (3.3-bis) risulta evidente che nel caso di indipendenza la densità congiunta è uguale alla fattorizzazione della densità marginali e, quindi, ancora una volta la struttura di dipendenza è modellata dalla copula tramite la *density copula*.

Dato che la struttura di dipendenza è modellata esclusivamente dalla funzione copula  $C(u,v)$ , risulta immediato constatare che tutti gli indici di dipendenza possono essere espressi solo ed esclusivamente in funzione della copula. Ad esempio, l'indice  $\tau$  di Kendall ( $\tau_{XY}$ ), definito nel seguente modo:

$$\tau_{XY} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

e indica la probabilità di “discordanza” tra due vettori di variabili casuali,  $(X_1, Y_1)$  e  $(X_2, Y_2)$ .

Utilizzando la funzione copula, può essere riscritto nel modo seguente:

$$\tau_{XY} = 4 \int_{I^2} C(u,v) \frac{\partial^2 C(u,v)}{\partial u \partial v} dudv - 1 \quad (3.4)$$

Si evidenzia che la covarianza non può essere espressa solo ed esclusivamente in funzione della copula dato che per definizione dipende anche dalle medie marginali e, quindi, dalle distribuzioni marginali.

Un aspetto che in economia può assumere una notevole importanza è il concetto di dipendenza nelle code di una distribuzione multivariata. Questo tipo di dipendenza descrive il comportamento delle variabili quando si verificano eventi estremi. La dipendenza in coda a destra (sinistra) è definita come la probabilità che un margine superi (non superi) una certa soglia dato che un altro margine ha già superato (non ha già superato) una soglia. I seguenti indici sono considerati misure di *upper* e *lower tail dependence*, rispettivamente:

$$\lambda_U = \lim_{p \rightarrow 1^-} P[X > F_X^{-1}(p) | Y > F_Y^{-1}(p)] \quad (3.5)$$

$$\lambda_L = \lim_{p \rightarrow 0^+} P[X < F_X^{-1}(p) | Y < F_Y^{-1}(p)] \quad (3.6)$$

Se  $\lambda_U = 0$  e  $\lambda_L = 0$ , le variabili sono indipendenti nelle code.

In questo capitolo, verrà utilizzata una particolare copula, ovvero la copula Farlie-Gumbel-Morgenstern (FGM), la quale rappresenta uno strumento molto utilizzato per la sua semplicità. Inoltre, il parametro di associazione che la caratterizza assume valori nell'intervallo  $[-1, 1]$ , pertanto, modella dipendenza sia negativa sia positiva (anche se debole). Di tale copula ne verrà data una più ampia trattazione nei prossimi paragrafi (cfr. §3.4).

### 3.2.1 Le copule archimedee

Nell'ambito delle funzioni copule, una particolare classe di copule di diffuso utilizzo è quella delle copule archimedee. Ciò dipende, in primo luogo, dalla non particolare difficoltà con cui possono essere costruite. In secondo luogo, per l'ampia varietà di famiglie di funzioni che appartengono a tale classe e, infine, per le proprietà matematico-statistiche che presentano. In generale, esse possono essere espresse nel modo seguente:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) \quad (3.7)$$

dove  $\varphi(t): [0,1] \rightarrow [0, \infty)$  è la funzione generatrice della copula, continua, strettamente decrescente e convessa. Vale, inoltre, che  $\varphi(1) = 0$ .

Questa famiglia di copule risulta molto utile per modellare la dipendenza nelle code della distribuzione. Tra le copule archimedee si trova la copula di Clayton che assume la seguente forma:

$$C^C(u, v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad (3.8)$$

per la quale  $\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$  e il parametro di dipendenza  $\theta$  può assumere valori compresi nell'intervallo  $(0, \infty)$ . Da questo si deduce che la copula di Clayton non modella la dipendenza negativa. Questa famiglia di copule presenta una forte dipendenza nelle code a sinistra, ma debole a destra.

Una funzione copula molto utilizzata che appartiene alla classe delle copule archimedee è la copula di Frank<sup>55</sup>, espressa nel modo seguente:

$$C^F(u, v; \theta) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right) \quad (3.9)$$

La funzione generatrice assume la seguente forma:  $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$ . Si consideri che in questo caso,  $\theta$  appartiene all'intervallo  $(-\infty, +\infty)$ . Tuttavia, la copula di Frank presenta una debole dipendenza nei valori estremi. Secondo Joe (1997), questa copula presenta un ordinamento positivo e una dipendenza positiva in termini di ordinamento stocastico.

La copula di Gumbel<sup>56</sup>, invece, è data dalla seguente espressione:

$$C^G(u, v; \theta) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right) \quad (3.10)$$

<sup>55</sup> Copula B3 nella definizione di Joe (1997).

<sup>56</sup> Copula B6 nella classificazione proposta da Joe (1997).

per la quale  $\varphi(t) = (-\ln t)^\theta$ . Inoltre, il parametro di dipendenza varia nell'intervallo  $[1, \infty)$ . Questa famiglia di copule presenta una forte dipendenza nelle code a destra, debole a sinistra. Anche la copula di Gumbel presenta un ordinamento positivo e una dipendenza positiva in termini di ordinamento stocastico.

### 3.3 Frontiere stocastiche e copule

In questo paragrafo si mostra cosa cambia nella costruzione della frontiera stocastica nel caso si tenga conto della dipendenza tra le componenti stocastiche. Si vedrà che, in tale contesto, la verosimiglianza sarà funzione della densità dell'errore composto e, quindi, della specificazione della funzione di densità congiunta tra le componenti stocastiche. Il modello di frontiera stocastica per dati cross section risulta essere il seguente:

$$Costi_i = C(y, w) \exp(v) \exp(u) = \prod_j y_{ij}^{\beta_j} \prod_n w_{in}^{\gamma_n} e^{v_i + u_i}$$

ed in forma logaritmica:

$$\log Costi_i = \sum_j \beta_j \log y_{ij} + \sum_n \gamma_n \log w_{in} + u_i + v_i$$

In questo capitolo si è scelto di utilizzare l'approccio proposto da Battese e Coelli (1995)<sup>57</sup>, secondo il quale si stima simultaneamente la frontiera dei costi e l'equazione dell'inefficienza che risulta dalla seguente espressione:

$$u_i = \sum_{k=1}^K \eta_k z_{ik} + e_i$$

Posto  $\varepsilon_i = u_i + v_i$  e seguendo il procedimento illustrato in Smith (2008), l'idea è di costruire la funzione di verosimiglianza assegnando all'errore composto la funzione di distribuzione costruita mediante la copula. Come più volte detto, questo metodo consente di introdurre funzioni marginali più flessibili e di tenere conto di strutture di dipendenza non solo lineari. Formalmente, la funzione di verosimiglianza risulta:

$$L = \prod_{i=1}^I f_\theta(\varepsilon) = \prod_{i=1}^I f_\theta \left( \log Costi_i - \sum_j \beta_j \log y_{ij} - \sum_n \gamma_n \log w_{in} \right) \quad (3.11)$$

<sup>57</sup> Si veda anche Kumbhakar, Ghosh e McGuckin (1991).

dove

$$f_{\theta}(\varepsilon) = \int_{\mathfrak{R}^+} f(u, \varepsilon - u) du \quad (3.12)$$

e

$$f(u, \varepsilon - u) = \frac{\partial^2}{\partial u \partial \varepsilon} F(u, \varepsilon - u) = f(u)g(\varepsilon - u)c_{\theta}(F(u), G(\varepsilon - u)) \quad (3.13)$$

Confrontando le espressioni (3.11)-(3.13) con quelle analoghe ottenute nel caso di indipendenza (Kumbhakar e Lovell, 2000), risultano evidenti le difficoltà di natura statistico-matematica derivanti non solo dalla specificazione della funzione di densità congiunta tra le componenti d'errore ma anche, se non soprattutto, dalla necessità di risolvere l'integrale presente nella (3.12).

Infine, per scorporare la componente di inefficienza dall'errore composto, l'efficienza di costo è ottenuta secondo la seguente espressione (Smith, 2008):

$$CE_{\theta} = E[\exp(-u) | \varepsilon = \varepsilon^*] = \frac{1}{f_{\theta}(\varepsilon)} \int_{\mathfrak{R}^+} \exp(-u) f(u, \varepsilon - u) du \quad (3.14)$$

Riepilogando, i passaggi per costruire la funzione di distribuzione dell'errore composto  $\varepsilon = u + v$  di una frontiera stocastica e scorporare l'efficienza sono i seguenti:

1. Scelta e definizione delle funzioni di distribuzione marginali;
2. Costruzione della copula,  $C_{\theta}(F(u), G(v))$ ;
3. Derivazione della *density copula*,  $c_{\theta}(F(u), G(v))$ ;
4. Costruzione della funzione congiunta,  $f(u, v)$  [equazione 3.13];
5. Derivazione della funzione di distribuzione del termine d'errore composto  $f_{\theta}(\varepsilon) = \int_{\mathfrak{R}^+} f(u, \varepsilon - u) du$ , tenendo conto che  $\varepsilon = v + u$  [equazione 3.12];
6. Costruzione della funzione di distribuzione condizionata al fine di scomporre le due componenti di errore,  $f_{\theta}(u | \varepsilon) = \frac{f_{\theta}(u, \varepsilon - u)}{f_{\theta}(\varepsilon)}$ ;
7. Derivazione dell'efficienza di costo  $CE_{\theta} = E[\exp(-u) | \varepsilon = \varepsilon^*]$  [equazione 3.14].

### 3.4 La funzione di distribuzione dell'errore composto mediante la copula FGM e marginali Normali

In questo paragrafo, si propone la costruzione della funzione di distribuzione dell'errore composto mediante la copula Farlie-Gumbel-Morgenstern (FGM) e considerando funzioni marginali Normali. La forma di una copula FGM assume la seguente espressione (Nelsen, 1999):

$$C_{\theta}(F(u), G(v)) = F(u) G(v) + F(u) G(v) \theta [1 - F(u)] [1 - G(v)] \quad (3.15)$$

in cui il parametro di associazione  $\theta$  assume valori all'interno dell'intervallo  $[-1, 1]$ .

La *density copula* FGM assume la seguente forma:

$$c_{\theta}(F(u), G(v)) = 1 + \theta [1 - 2 F(u)] [1 - 2 G(v)] \quad (3.16)$$

Secondo Trivedi e Zimmer (2005), tale copula risulta utile quando la dipendenza fra le due componenti non assume valori elevati.

Di seguito è riportato il modello costruito mediante la copula FGM che considera la dipendenza tra variabili distribuite seguendo funzioni Normali.

La componente di inefficienza, non-negativa, si ipotizza che si distribuisca secondo una Normale troncata, ovvero  $U \sim N^+(z'\eta, \sigma_u^2)$ , ed abbia la seguente funzione di densità:

$$f_0(u; \sigma_u^2) = \frac{f(u; \sigma_u^2)}{1 - F(0)} = \frac{\frac{1}{\sigma_u \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{u - z'\eta}{\sigma_u}\right)^2\right\}}{\int_0^{+\infty} \frac{1}{\sigma_u \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{u - z'\eta}{\sigma_u}\right)^2\right\}} \quad (3.17)$$

con  $u > 0$  e  $\sigma_u^2 > 0$ , e la seguente funzione di ripartizione

$$F_0(u; \sigma_u^2) = \Phi_0(u) = \frac{\Phi(u) - F(0)}{1 - F(0)} \quad (3.18)$$

$f_0(u; \sigma_u^2)$  è la funzione di densità Normale troncata allo zero della componente di efficienza con media  $z'\eta$  e parametro di scala  $\sigma_u^2$ , dove  $z'\eta$  è il predittore lineare dell'equazione dell'inefficienza (cfr. § 3.3).

Inoltre, si ipotizza che l'errore casuale, definito in  $\mathfrak{R}$ , segua una legge Normale, ossia  $V \sim N(0, \sigma_v^2)$ . La funzione di densità e di ripartizione, rispettivamente, sono:

$$g(v; \sigma_v^2) = \frac{1}{\sigma_v \sqrt{2\pi}} \exp\left(-\frac{v^2}{2\sigma_v^2}\right) \quad (3.19)$$

$$G_v(v; \sigma_v^2) = \Phi(v) \quad (3.20)$$

con  $v \in \mathfrak{R}$  e  $\sigma_v > 0$ , dove  $\sigma_v^2$  rappresenta il parametro di scala della distribuzione marginale dell'errore casuale. Sotto tali ipotesi, la copula FGM assume la seguente forma:

$$c_\theta^N(F_0(u), G(v)) = \Phi_0(u) \Phi(v) + \Phi_0(u) \Phi(v) \theta [1 - \Phi_0(u)] [1 - \Phi(v)] \quad (3.21)$$

e la *density copula* FGM è data da:

$$c_\theta^N(F_0(u), G(v)) = 1 + \theta [1 - 2\Phi_0(u)] [1 - 2\Phi(v)] \quad (3.22)$$

Pertanto, la funzione di distribuzione congiunta delle due componenti di errore è uguale a:

$$f_\theta^N(u, v) = f_0(u) g(v) \{1 + \theta [1 - 2\Phi_0(u)] [1 - 2\Phi(v)]\} \quad (3.23)$$

Sapendo che  $\mathcal{E} = u + v$  e sostituendo nella  $f_\theta^N(u, v)$ , si ottiene:

$$f_\theta^N(\mathcal{E}) = \int_{\mathfrak{R}^+} f(u, \mathcal{E} - u) du = \int_0^{+\infty} f_0(u) g(\mathcal{E} - u) \{1 + \theta [1 - 2\Phi_0(u)] [1 - 2\Phi(\mathcal{E} - u)]\} du \quad (3.24)$$

Data la presenza nella (3.24) della funzione di ripartizione Normale di cui non si ha forma esplicita, anche tale espressione rimane in forma implicita. Ciò non preclude la possibilità di massimizzare la funzione di verosimiglianza in cui ogni elemento è dato da questa funzione di distribuzione dell'errore composto. A tal proposito, nel prossimo sottoparagrafo sono riportate le stime della funzione di costo e dell'equazione dell'inefficienza relative al modello appena proposto.

### 3.4.1 Stime econometriche della frontiera di costo e dell'equazione dell'efficienza in caso di normalità e dipendenza fra le due componenti d'errore

In questo sottoparagrafo, si riportano le stime relative al modello appena presentato costruito assegnando alle componenti di errore funzioni di distribuzione Normale e Normale

troncata e modellando la dipendenza mediante una copula FGM<sup>58</sup>. In termini di scelta delle variabili è stato replicato il modello II di cui si è discusso nel capitolo precedente (cfr. Appendice A2).

Di particolare interesse è la stima del parametro di associazione della copula,  $\theta$  (ultima riga della tabella 3.1). Esso è, infatti, significativo nei primi tre anni considerati, mentre, non è significativo negli ultimi due. In altre parole, se sono valide le ipotesi del modello (funzione copula FGM e distribuzioni Normali) emerge che le due componenti di errore sono dipendenti nel triennio 2006-2008, mentre sono indipendenti nel 2009 e nel 2010. Inoltre, è importante evidenziare che l'introduzione dell'ipotesi di dipendenza fra le due componenti di errore ha prodotto rilevanti novità e che la non-significatività del parametro  $\theta$  in alcuni anni di analisi non sembra destare particolare preoccupazione. Quest'ultimo aspetto, infatti, può essere giustificato da varie ragioni ma, per prima cosa, occorre sottolineare che il caso di indipendenza è un caso particolare del modello proposto nel paragrafo 3.4. Il fatto che i parametri stimati della frontiera di costo presenti nelle ultime due colonne della tabella 3.1 e quelli stimati nel caso di indipendenza (cfr. Appendice A2, tabella A2.3) siano solo marginalmente diversi dimostra l'importanza di considerare un modello più generale che contempli anche la dipendenza tra l'errore casuale e la componente di inefficienza.

Differenze più marcate sono ravvisabili nei primi tre anni di analisi per i quali il coefficiente  $\theta$  risulta statisticamente significativo. Le elasticità delle funzioni di costo stimate considerando l'ipotesi di dipendenza, infatti, cambiano in misura maggiore rispetto al caso in cui si rifiuta l'ipotesi di dipendenza.

E' utile sottolineare che nei casi in cui si rifiuta la dipendenza tra le due componenti di errore, l'aspettativa è quella di ottenere risultati uguali o marginalmente diversi da quelli ottenuti stimando il modello che impone l'ipotesi di indipendenza. Tuttavia, eventuali differenze potrebbero trovare giustificazione in tre ordini di motivi. Il primo è relativo alla possibilità di ottenere approssimazioni diverse dovute ai diversi processi di ottimizzazione implementati (per esempio, nel caso delle stime riportate nella tabella 3.1 la routine di ottimizzazione è stata costruita in modo autonomo e differisce da quella dei pacchetti applicativi utilizzati per ottenere i risultati del cap. 2 relativi al caso *classico*). La seconda questione riguarda l'introduzione nel modello di un parametro in più, ovvero il parametro di associazione  $\theta$  della copula FGM. Questo è evidente soprattutto nei differenti risultati in termini di significatività dei parametri, poiché cambia la dimensione della matrice di varianza

---

<sup>58</sup> Le stime sono realizzate mediante il software *Matlab*. Il metodo utilizzato è quello della massima verosimiglianza.

e covarianza utilizzata per la stima degli errori standard. La terza motivazione è relativa alla possibile esistenza di forme diverse di dipendenza che la copula FGM non riesce a catturare. In sintesi, l'introduzione della dipendenza tra le componenti di errore mediante la copula FGM suggerisce che il confronto dei risultati ottenuti nei due casi (dipendenza e indipendenza delle due componenti) debba essere effettuato con estrema cautela.

**Tabella 3.1 Stime della frontiera di costo, dell'equazione dell'inefficienza e del parametro di associazione della copula FGM**

	2006		2007		2008		2009		2010	
	Coeff.	t								
Intercetta	-2.38	-1.56	-1.56***	-9.53	-0.66***	-3.36	-5.11***	-53.05	-5.43***	-33.70
log(y <sub>1</sub> )	0.61***	9.77	0.48***	50.53	0.42***	38.00	0.53***	77.71	0.66***	79.91
log(y <sub>2</sub> )	0.16*	2.11	0.24***	16.88	0.33***	22.74	0.25***	36.10	0.11***	18.75
log(y <sub>3</sub> )	0.23**	3.05	0.10***	9.13	0.09***	8.32	0.10***	14.18	0.14***	17.40
log(w <sub>1</sub> /w <sub>3</sub> )	0.45*	2.42	0.63***	33.43	0.65***	29.33	0.71***	52.14	0.63***	34.83
log(w <sub>2</sub> /w <sub>3</sub> )	0.44*	2.19	0.31***	20.56	0.20***	12.29	0.29***	23.02	0.32***	29.62
z1 = l_bl	-0.22***	-40.13	-0.07***	-29.26	-0.08***	-39.74	-0.39***	-41.28	-0.39***	-37.54
z2 = ind_solv	0.07***	95.57	0.57***	29.64	0.01	0.49	0.01***	3.50	0	0
z3 = dspa	0.17***	5.70	0.10***	8.62	0.09***	7.76	0.44***	8.17	0.50***	9.32
z4 = dpop	0.29***	5.84	0.14***	5.58	0.15***	8.66	0.73***	8.65	0.76***	7.45
z5 = dminori	0.32***	9.96	0.14***	11.29	0.19***	14.64	0.88***	18.18	0.94***	16.66
z6 = dmedie	0.15*	2.02	1.03***	26.28	1.05***	33.17	0.45***	3.23	0.46**	2.97
z7 = dgrandi	-0.11	-1.36	0.90***	35.11	0.84***	32.39	-0.23*	-2.22	-0.30**	-2.97
z8 = dmaggiori	-0.64*	-2.28	1.03***	14.09	0.97***	20.13	0.72	1.71	0.59	0.79
z9 = dnordovest	0.14***	4.38	0.04***	3.31	0.06***	4.92	0.30***	4.60	0.32***	5.57
z10 = dcentro	0.16***	4.83	0.06***	4.46	0.08***	6.04	0.32***	4.84	0.26***	4.15
z11 = dsud	0.29***	8.42	0.09***	6.25	0.12***	8.08	0.50***	6.81	0.51***	5.82
1/sigma_u <sup>59</sup>	4.44***	31.66	9.37***	37.96	8.80***	37.68	1.91***	44.83	1.90***	39.65
1/sigma_v	1.11***	81.77	3.60***	37.09	3.23***	53.36	4.78***	43.29	4.58***	41.86
theta(FGM)	-0.60*	-2.16	0.31*	2.32	0.73***	5.43	-0.09	-0.59	-0.05	-0.36

Livelli di significatività: '\*\*\*' = 0.001; '\*\*' = 0.01; '\*' = 0.05; '.' = 0.1; ' ' = 1.

Fonte: elaborazioni su dati ABI.

Tenendo conto di queste considerazioni, si sottolinea, innanzitutto, che le elasticità della funzione di costo rispetto agli output e al prezzo dei fattori produttivi risultano sempre statisticamente significative. Inoltre, i valori assunti dai parametri stimati sono sempre positivi e coerenti con le aspettative (cfr. § 2.2).

<sup>59</sup> Sigma\_u è la radice quadrata della varianza di u e sigma\_v è la radice quadrata della varianza di v.

Uno dei principali contributi di questa applicazione è quello di proporre un'analisi di efficienza mediante l'approccio delle frontiere stocastiche proposto da Battese e Coelli (1995) tenendo conto della dipendenza tra gli errori. Per quel che riguarda le stime dell'equazione dell'inefficienza, si può notare come molti coefficienti siano statisticamente significativi. Inoltre, per il 2006 i risultati riportati nella tabella 3.1 differiscono da quelli ottenuti sotto l'ipotesi di indipendenza (cfr. Appendice A2, tabella A2.4), ma sono più in linea con quelli degli altri anni.

Un risultato da evidenziare è legato alle stime del coefficiente associato alla qualità dei crediti ( $z_1$ ). Si consideri che tale indice è misurato come il rapporto tra i crediti in sofferenza e il totale dei crediti verso clienti, quindi, valori più elevati di  $z_1$  segnalano una minore qualità dei crediti. Relativamente alle stime realizzate con indipendenza (tabella A2.4), i coefficienti risultano non statisticamente significativi negli ultimi tre anni di analisi, mentre risultano significativi e positivi nel 2006 e nel 2007. Nella tabella 3.1 emerge, invece, che i coefficienti associati alla variabile  $z_1$ , sempre statisticamente significativi, assumono valori negativi in tutti gli anni coerentemente con quanto si ottiene nelle stime realizzate sui dati panel sotto l'ipotesi di indipendenza (cfr. § 2.5, tabella 2.10). Tale risultato indica che una minore qualità dei crediti comporta una minore (maggiore) inefficienza (efficienza). Come osservano Giordano e Lopes (2006), la qualità dei crediti è dipendente dalle operazioni di controllo e monitoraggio degli impieghi, per cui scelte manageriali che ridimensionano gli sforzi nella selezione del credito si riflettono in un maggiore tasso di insolvenza dei debitori. Tuttavia, il risparmio sui costi di monitoraggio e *screening* potrebbe comportare un aumento dell'efficienza di costo delle banche italiane (cfr. § 2.5).

Per quanto concerne le *dummies* di categoria giuridica, si ottengono risultati coerenti per tutti gli anni che vanno nella direzione di mostrare migliori performance per le BCC. Si tratta di un risultato che non emerge nelle stime realizzate sotto l'ipotesi di indipendenza, ma che si riscontra nelle stime sui dati panel (cfr. § 2.5-§2.6). Facendo riferimento al ruolo della dimensione, si nota che le banche minori e medie raggiungono livelli di efficienza più bassi rispetto alle banche piccole, mentre per le altre banche i risultati non sono univoci. Infine, anche relativamente alle *dummies* territoriali, si hanno risultati più convincenti rispetto a quelli ottenuti in caso di indipendenza in base ai quali le banche del Sud ottengono, in alcuni casi, livelli di efficienza più alti rispetto alle banche del Nord-Est, che rappresentano il gruppo di controllo. Dalla tabella 3.1 emerge, invece, che in tutti gli anni di analisi, le banche del Nord-Est registrano i livelli maggiori di efficienza, a parità di altre condizioni.

In base a tali considerazioni e tenendo conto dei risultati relativi al parametro di associazione della funzione copula FGM, nei prossimi paragrafi si propone la derivazione metodologica di due modelli che rappresentano un tentativo di generalizzazione di quello utilizzato finora. In particolare, nel paragrafo 3.5 si generalizza il modello proposto da Smith (2008), mentre, nel paragrafo 3.6, si utilizza la copula FGM e si ipotizza che le componenti di errore si distribuiscano secondo funzioni Log-Dagum.

### 3.5 Un approccio metodologico per la costruzione della funzione di distribuzione dell'errore composto mediante la copula FGM ed una legge Esponenziale per la componente di inefficienza

In questo paragrafo si deriva la funzione di distribuzione dell'errore composto,  $\varepsilon$ , nel caso in cui si assegna alla componente di inefficienza,  $u$ , una funzione Esponenziale e all'errore casuale,  $v$ , una funzione Logistica Generalizzata di I tipo<sup>60</sup>. Si tratta, quindi, di un modello che include, come caso particolare, quello proposto da Smith (2008), il quale utilizza un modello in cui  $u$  si distribuisce secondo una legge Esponenziale e l'errore casuale secondo una funzione Logistica Standard. Come si è detto in precedenza, la generalizzazione proposta ha l'obiettivo di valutare l'impatto della forma (in particolare l'asimmetria) della distribuzione dell'errore accidentale sulla distribuzione dell'errore composto nel caso di dipendenza tra gli errori della frontiera stocastica.

Formalmente, si assume che la variabile casuale (v.c.)  $U$  segua una distribuzione Esponenziale negativa di parametro  $\delta_u$ , in sintesi  $U \sim Exp(\delta_u)$ , con funzioni di densità e ripartizione, rispettivamente, date da:

$$f(u; \delta_u) = \frac{1}{\delta_u} e^{-\frac{u}{\delta_u}} \quad (3.25)$$

$$F(u; \delta_u) = 1 - e^{-\frac{u}{\delta_u}} \quad (3.26)$$

con  $u > 0$  e  $\delta_u > 0$ . Mentre, la v.c.  $V$  segue una distribuzione Logistica Generalizzata di I tipo, con parametri,  $\delta_v$ ,  $\alpha_v$  e  $\lambda_v$ , in sintesi  $V \sim LDa(\alpha_v, \delta_v, \lambda_v)$ , con funzione di densità e di ripartizione date da:

---

<sup>60</sup> Come già accennato, la funzione Logistica Generalizzata di I tipo è una ri-parametrizzazione di una funzione Log-Dagum. Si veda Domma e Condino (2013) per i recenti sviluppi relativi a tale distribuzione.

$$g(v; \alpha_v, \delta_v, \lambda_v) = \frac{\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v-1} \quad (3.27)$$

$$G(v; \alpha_v, \delta_v, \lambda_v) = \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \quad (3.28)$$

con  $v \in \mathfrak{R}$  e  $\alpha_v, \delta_v, \lambda_v > 0$ , dove  $\lambda_v$  è il parametro di locazione,  $\alpha_v$  è un parametro che regola la forma della distribuzione e  $\delta_v$  è il parametro di scala. Si evidenzia che la  $LDa(\alpha_v, \delta_v, \lambda_v)$  è simmetrica se  $\alpha_v = 1$ , asimmetrica negativa se  $\alpha_v \in (0, 1)$  e, infine, asimmetrica positiva se  $\alpha_v > 1$ .

Considerando la (3.12), la (3.13) e la *density copula* FGM (3.16), la seguente espressione rappresenta la funzione di distribuzione congiunta:

$$f_{\theta}^{Ex}(u, v) = \frac{1}{\delta_u} e^{-\frac{u}{\delta_u}} \frac{\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v-1} \left\{ 1 + \theta \left[ 1 - 2 \left( 1 - e^{-\frac{u}{\delta_u}} \right) \right] \left[ 1 - 2 \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \right] \right\} \quad (3.29)$$

Sostituendo  $\mathcal{E} = u + v$  nell'espressione 3.29 e marginalizzando rispetto alla variabile  $u$ , si ottiene la funzione di distribuzione dell'errore composto  $f_{\theta}^{Exp}(\mathcal{E})$ , la quale rappresenta l'argomento della funzione di verosimiglianza da costruire al fine di poter stimare i parametri delle distribuzioni e gli *score* di efficienza. In formula:

$$\begin{aligned} f_{\theta}^{Ex}(\mathcal{E}) &= \int_0^{+\infty} f_{\theta}^{Ex}(u, \mathcal{E}-u) du = \\ &= \int_0^{+\infty} f(u) g(\mathcal{E}-u) \{ 1 + \theta [1 - 2 F(u)] [1 - 2 G(\mathcal{E}-u)] \} du \end{aligned} \quad (3.30)$$

E' possibile giungere ad una forma esplicita della funzione di distribuzione dell'errore composto seguendo il procedimento di seguito illustrato.

Da semplici passaggi algebrici, si può scrivere la (3.30) nel seguente modo:

$$f_{\theta}^{Ex}(\varepsilon) = (1+\theta) \underbrace{\int_0^{\infty} f(u) g(\varepsilon-u) du}_{(1)} - 2\theta \left\{ \underbrace{\int_0^{\infty} f(u) g(\varepsilon-u) F(u) du}_{(2)} + \underbrace{\int_0^{\infty} f(u) g(\varepsilon-u) G(\varepsilon-u) du}_{(3)} - 2 \underbrace{\int_0^{\infty} f(u) g(\varepsilon-u) F(u) G(\varepsilon-u) du}_{(4)} \right\} \quad (3.31)$$

Si tratta di una funzione composta da quattro integrali, laddove i primi tre sono riconducibili al caso più generale rappresentato dalla componente (4) della (3.31). Evidentemente, i casi particolari si ottengono quando  $F(u)$  e/o  $G(\varepsilon-u)$  sono pari a 1. Per comodità, la quarta componente verrà indicata come “integrale tipo”; risolto tale integrale, i primi tre si ottengono come casi particolari. Sostituendo nell’integrale tipo le funzioni di densità e di ripartizione della Esponenziale negativa e della Logistica Generalizzata di I tipo, si ottiene:

$$I = \int_0^{\infty} f_0(u) g(\varepsilon-u) F_0(u) G(\varepsilon-u) du = \int_0^{\infty} \frac{1}{\delta_u} e^{-\frac{u}{\delta_u}} \frac{\alpha_v}{\delta_v} e^{-\frac{(\varepsilon-u)-\lambda_v}{\delta_v}} \left( 1 + e^{-\frac{(\varepsilon-u)-\lambda_v}{\delta_v}} \right)^{-\alpha_v-1} \left( 1 - e^{-\frac{u}{\delta_u}} \right) \left( 1 + e^{-\frac{(\varepsilon-u)-\lambda_v}{\delta_v}} \right)^{-\alpha_v} du = \frac{\alpha_v c_2}{\delta_u \delta_v} \left\{ \int_0^{\infty} e^{-\frac{u}{\delta_u} + \frac{u}{\delta_v}} \left[ 1 + c_2 e^{\frac{u}{\delta_v}} \right]^{-2\alpha_v-1} du - \int_0^{\infty} e^{-\frac{2u}{\delta_u} + \frac{u}{\delta_v}} \left[ 1 + c_2 e^{\frac{u}{\delta_v}} \right]^{-2\alpha_v-1} du \right\} \quad (3.32)$$

$$\text{con } c_2 = e^{-\frac{\varepsilon-\lambda_v}{\delta_v}} .$$

Dopo alcuni passaggi algebrici e utilizzando la trasformazione  $y = e^{-u}$  si ottiene la seguente espressione dell'integrale tipo:

$$\begin{aligned}
I &= \frac{\alpha_v c_2}{\delta_u \delta_v} \left\{ \int_0^1 y^{\frac{1}{\delta_u} - \frac{1}{\delta_v} - 1} \left( 1 + c_2 y^{\frac{1}{\delta_v}} \right)^{-2\alpha_v - 1} dy - \int_0^1 y^{\frac{2}{\delta_u} - \frac{1}{\delta_v} - 1} \left( 1 + c_2 y^{\frac{1}{\delta_v}} \right)^{-2\alpha_v - 1} dy \right\} = \\
&= \frac{\alpha_v c_2}{\delta_u \delta_v} \left\{ \int_0^1 y^{\frac{1}{\delta_u} - \frac{1}{\delta_v} - 1} \frac{c_2^{-2\alpha_v - 1}}{y^{\frac{1}{\delta_v}} (-2\alpha_v - 1)} \left( 1 + \frac{y^{\frac{1}{\delta_v}}}{c_2} \right)^{-2\alpha_v - 1} dy - \int_0^1 y^{\frac{2}{\delta_u} - \frac{1}{\delta_v} - 1} \frac{c_2^{-2\alpha_v - 1}}{y^{\frac{1}{\delta_v}} (-2\alpha_v - 1)} \left( 1 + \frac{y^{\frac{1}{\delta_v}}}{c_2} \right)^{-2\alpha_v - 1} dy \right\} = \\
&= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta_u \delta_v} \left\{ \int_0^1 y^{\frac{1}{\delta_u} + \frac{2\alpha_v}{\delta_v} - 1} \left( 1 + \frac{y^{\frac{1}{\delta_v}}}{c_2} \right)^{-2\alpha_v - 1} dy - \int_0^1 y^{\frac{2}{\delta_u} + \frac{2\alpha_v}{\delta_v} - 1} \left( 1 + \frac{y^{\frac{1}{\delta_v}}}{c_2} \right)^{-2\alpha_v - 1} dy \right\}
\end{aligned} \tag{3.33}$$

A questo punto, è utile sfruttare i seguenti risultati noti in letteratura (Abramowitz e Stegun, 1972):

- sviluppo binomiale

$$(1+z)^{-a} = \sum_{p=0}^{\infty} \frac{(-1)^p (a)_p z^p}{p!} \tag{3.34}$$

dove  $(a)_p = a(a+1)\dots(a+p-1) = \frac{\Gamma(a+p)}{\Gamma(a)}$ , noto come simbolo di Pochhammer, e  $\Gamma(\cdot)$  è la

funzione matematica Gamma definita dalla seguente espressione:  $\Gamma(a) = \int_0^{\infty} w^{a-1} e^{-w} dw$ ; è utile

ricordare che per la funzione Gamma vale la seguente relazione di ricorrenza:

$$\Gamma(a+1) = a\Gamma(a).$$

- serie ipergeometrica generalizzata

$${}_p F_q [\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; z] = \sum_{k=0}^{\infty} \frac{(\alpha_1)_k \dots (\alpha_p)_k}{(\beta_1)_k \dots (\beta_q)_k} \frac{z^k}{k!} \tag{3.35}$$

- serie di Kampé de Fériet (si tratta di una generalizzazione della funzione ipergeometrica e comprende anche la funzione ipergeometrica di Appell)

$$F_{l;m;n}^{p;q;k} [a_p : b_q; c_k; \alpha_l : \beta_m; \gamma_n; -x, -y] = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\prod_{j=1}^p (a_j)_{r+s} \prod_{j=1}^q (b_j)_r \prod_{j=1}^k (c_j)_s}{\prod_{j=1}^l (\alpha_j)_{r+s} \prod_{j=1}^m (\beta_j)_r \prod_{j=1}^n (\gamma_j)_s} \frac{x^r}{r!} \frac{y^s}{s!} \tag{3.36}$$

Ritornando alla (3.33), utilizzando lo sviluppo binomiale si può scrivere:

$$\left(1 + \frac{y}{c_2}\right)^{-2\alpha_v - 1} = \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i!} \frac{y^i}{c_2^i} \quad (3.37)$$

Sostituendo la (3.37) nell'ultima espressione dell'integrale tipo si ottiene:

$$\begin{aligned} I &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta_u \delta_v} \left\{ \int_0^1 y^{\frac{1}{\delta_u} + \frac{2\alpha_v}{\delta_v} - 1} \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i!} \frac{y^i}{c_2^i} dy - \int_0^1 y^{\frac{2}{\delta_u} + \frac{2\alpha_v}{\delta_v} - 1} \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i!} \frac{y^i}{c_2^i} dy \right\} = \\ &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta_u \delta_v} \left\{ \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \left[ \int_0^1 y^{\frac{1}{\delta_u} + \frac{2\alpha_v + i}{\delta_v} - 1} dy - \int_0^1 y^{\frac{2}{\delta_u} + \frac{2\alpha_v + i}{\delta_v} - 1} dy \right] \right\} = \\ &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta_u \delta_v} \left\{ \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \left[ \frac{1}{\frac{1}{\delta_u} + \frac{2\alpha_v + i}{\delta_v}} - \frac{1}{\frac{2}{\delta_u} + \frac{2\alpha_v + i}{\delta_v}} \right] \right\} \end{aligned} \quad (3.38)$$

Sotto l'ipotesi  $\delta_u = \delta_v = \delta$ , si ottiene la seguente espressione:

$$\begin{aligned} I &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta^2} \left\{ \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \left[ \frac{\delta}{(2\alpha_v + i + 1)} - \frac{\delta}{(2\alpha_v + i + 2)} \right] \right\} = \\ &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta} \left\{ \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \left[ \frac{1}{(2\alpha_v + i + 1)} - \frac{1}{(2\alpha_v + i + 2)} \right] \right\} \end{aligned} \quad (3.39)$$

Utilizzando, ora, la formula ricorsiva della funzione Gamma, si osserva che

$$\frac{(2\alpha_v + 1)_i}{(2\alpha_v + 2)_i} = \frac{\Gamma(2\alpha_v + i + 1)}{\Gamma(2\alpha_v + 2)} = \frac{\Gamma(2\alpha_v + i + 1)}{\Gamma(2\alpha_v + 1)} \frac{(2\alpha_v + 1)\Gamma(2\alpha_v + 1)}{(2\alpha_v + i + 1)\Gamma(2\alpha_v + i + 1)} = \frac{2\alpha_v + 1}{2\alpha_v + i + 1} \quad (3.40)$$

In tal modo, si ottengono le espressioni:

$$\frac{1}{2\alpha_v + i + 1} = \frac{(2\alpha_v + 1)_i}{(2\alpha_v + 2)_i} \frac{1}{(2\alpha_v + 1)} \quad (3.41)$$

$$\frac{1}{2\alpha_v + i + 2} = \frac{(2\alpha_v + 2)_i}{(2\alpha_v + 3)_i} \frac{1}{(2\alpha_v + 2)} \quad (3.42)$$

Utilizzando le relazioni (3.41) e (3.42), l'integrale tipo si può scrivere nel modo seguente:

$$\begin{aligned}
 I &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta} \left\{ \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \left[ \frac{(2\alpha_v + 1)_i}{(2\alpha_v + 2)_i} \frac{1}{(2\alpha_v + 1)} - \frac{(2\alpha_v + 2)_i}{(2\alpha_v + 3)_i} \frac{1}{(2\alpha_v + 2)} \right] \right\} = \\
 &= \frac{\alpha_v c_2^{-2\alpha_v}}{\delta} \left\{ \frac{1}{(2\alpha_v + 1)} \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \frac{(2\alpha_v + 1)_i}{(2\alpha_v + 2)_i} - \frac{1}{(2\alpha_v + 2)} \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \frac{(2\alpha_v + 2)_i}{(2\alpha_v + 3)_i} \right\} \quad (3.43)
 \end{aligned}$$

A questo punto, considerando la forma generale di una serie ipergeometrica generalizzata (3.35), si giunge ad una soluzione esplicita dell'integrale tipo come somma algebrica di funzioni ipergeometriche:

$$I = \frac{\alpha_v c_2^{-2\alpha_v}}{\delta} \left\{ \frac{1}{(2\alpha_v + 1)} {}_2F_1 \left[ (2\alpha_v + 1), (2\alpha_v + 1); (2\alpha_v + 2); -c_2^{-1} \right] - \frac{1}{(2\alpha_v + 2)} {}_2F_1 \left[ (2\alpha_v + 1), (2\alpha_v + 2); (2\alpha_v + 3); -c_2^{-1} \right] \right\} \quad (3.44)$$

Seguendo lo stesso procedimento per le altre tre componenti della funzione di distribuzione congiunta data dall'espressione 3.31 si ottengono i seguenti integrali come casi particolari dell'integrale tipo:

$$I_{(1)} = \int_0^{\infty} f_0(u) g(\varepsilon - u) du = \frac{\alpha_v c_2^{-\alpha_v}}{\delta_u \delta_v} \int_0^1 y^{\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left( 1 + c_2^{-1} y \frac{1}{\delta_v} \right)^{-\alpha_v - 1} dy \quad (3.45)$$

$$I_{(2)} = \int_0^{\infty} f_0(u) g(\varepsilon - u) F_0(u) du = \frac{\alpha_v c_2^{-\alpha_v}}{\delta_u \delta_v} \left\{ \int_0^1 y^{\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left( 1 + c_2^{-1} y \frac{1}{\delta_v} \right)^{-\alpha_v - 1} dy - \int_0^1 y^{\frac{\alpha_v}{\delta_v} + \frac{2}{\delta_u} - 1} \left( 1 + c_2^{-1} y \frac{1}{\delta_v} \right)^{-\alpha_v - 1} dy \right\} \quad (3.46)$$

$$I_{(3)} = \int_0^{\infty} f_0(u) g(\varepsilon - u) G(\varepsilon - u) du = \frac{\alpha_v c_2^{-2\alpha_v}}{\delta_u \delta_v} \int_0^1 y^{2\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left( 1 + c_2^{-1} y \frac{1}{\delta_v} \right)^{-2\alpha_v - 1} dy \quad (3.47)$$

In analogia a quanto fatto per l'integrale tipo si può dimostrare che le forme esplicite delle altre tre componenti della funzione dell'errore composto sono:

$$I_{(1)} = \frac{\alpha_v c_2^{-\alpha_v}}{(\alpha_v + 1) \delta} {}_2F_1[(\alpha_v + 1), (\alpha_v + 1); (\alpha_v + 2); -c_2^{-1}] \quad (3.48)$$

$$I_{(2)} = \frac{\alpha_v c_2^{-\alpha_v}}{\delta} \left\{ \frac{1}{(\alpha_v + 1)} {}_2F_1[(\alpha_v + 1), (\alpha_v + 1); (\alpha_v + 2); -c_2^{-1}] - \frac{1}{(\alpha_v + 2)} {}_2F_1[(\alpha_v + 1), (\alpha_v + 2); (\alpha_v + 3); -c_2^{-1}] \right\} \quad (3.49)$$

$$I_{(3)} = \frac{\alpha_v c_2^{-2\alpha_v}}{(2\alpha_v + 1) \delta} {}_2F_1[(2\alpha_v + 1), (2\alpha_v + 1); (2\alpha_v + 2); -c_2^{-1}] \quad (3.50)$$

Pertanto, la funzione di distribuzione dell'errore composto,  $f_{\theta}^{Ex}(\varepsilon)$ , assume la seguente forma esplicita:

$$\begin{aligned} f_{\theta}^{Ex}(\varepsilon) = & (1 + \theta) \frac{\alpha_v c_2^{-\alpha_v}}{(\alpha_v + 1) \delta} {}_2F_1[(\alpha_v + 1), (\alpha_v + 1); (\alpha_v + 2); -c_2^{-1}] - \\ & - 2\theta \frac{\alpha_v c_2^{-\alpha_v}}{\delta} \left\{ \frac{1}{(\alpha_v + 1)} {}_2F_1[(\alpha_v + 1), (\alpha_v + 1); (\alpha_v + 2); -c_2^{-1}] - \frac{1}{(\alpha_v + 2)} {}_2F_1[(\alpha_v + 1), (\alpha_v + 2); (\alpha_v + 3); -c_2^{-1}] \right\} - \\ & - 2\theta \frac{\alpha_v c_2^{-2\alpha_v}}{(2\alpha_v + 1) \delta} {}_2F_1[(2\alpha_v + 1), (2\alpha_v + 1); (2\alpha_v + 2); -c_2^{-1}] + \\ & + 4\theta \frac{\alpha_v c_2^{-2\alpha_v}}{\delta} \cdot \\ & \cdot \left\{ \frac{1}{(2\alpha_v + 1)} {}_2F_1[(2\alpha_v + 1), (2\alpha_v + 1); (2\alpha_v + 2); -c_2^{-1}] - \frac{1}{(2\alpha_v + 2)} {}_2F_1[(2\alpha_v + 1), (2\alpha_v + 2); (2\alpha_v + 3); -c_2^{-1}] \right\} \end{aligned} \quad (3.51)$$

Si consideri che le serie ipergeometriche incluse nella (3.51) convergono solo se è verificata la seguente disuguaglianza  $|-c_2^{-1}| < 1$ , ovvero se vale che  $|-e^{\frac{\varepsilon - \lambda_v}{\delta_v}}| < 1$  (Smith, 2008).

In particolare,  $-1 < -e^{\frac{\varepsilon - \lambda_v}{\delta_v}} < 1$  quando  $\varepsilon < \lambda_v$ .

Appare opportuno evidenziare che la soluzione analitica ottenuta risulta essere un risultato rilevante in quanto fornisce la possibilità di valutare direttamente l'effetto della

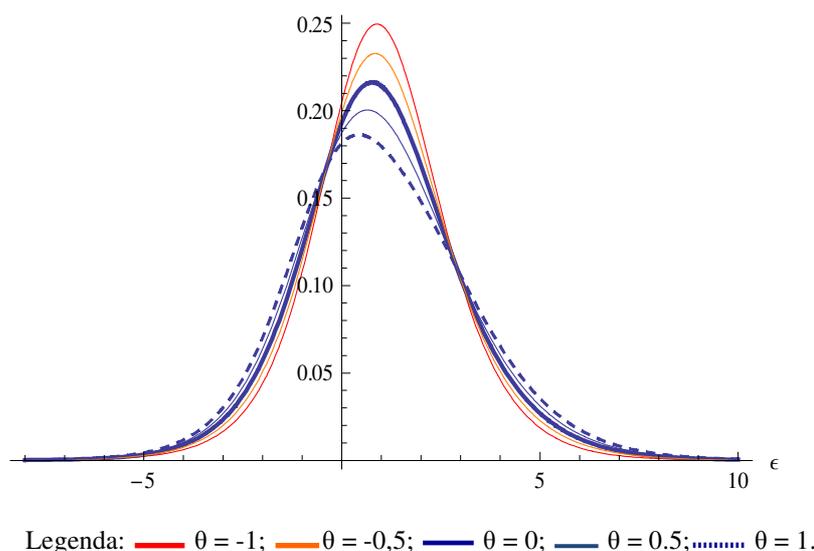
dipendenza (parametro  $\theta$ ) e l'effetto della forma della distribuzione dell'errore accidentale (parametro  $\alpha_v$ ) sulla distribuzione dell'errore composto.

Nelle figure 3.1-3.3 sono riportati i grafici della funzione (3.51), dati alcuni valori per i parametri delle distribuzioni marginali ( $\delta_u = \delta_v = \delta = 1$  e  $\lambda_v = 0$ ) e considerando cinque valori del parametro di associazione (-1, -0.5, 0, 0.5 e 1), al variare di  $\alpha_v$ , che rappresenta la misura di asimmetria e curtosi della funzione di distribuzione Log-Dagum assegnata alla componente casuale. Si considerano tre valori diversi di  $\alpha_v$ , ovvero  $\alpha_v = 1$  che segnala simmetria della distribuzione dell'errore accidentale,  $\alpha_v = 2$  che indica, invece, asimmetria positiva (a destra) e, infine,  $\alpha_v = 0.25$  che segnala la presenza di asimmetria negativa (a sinistra).

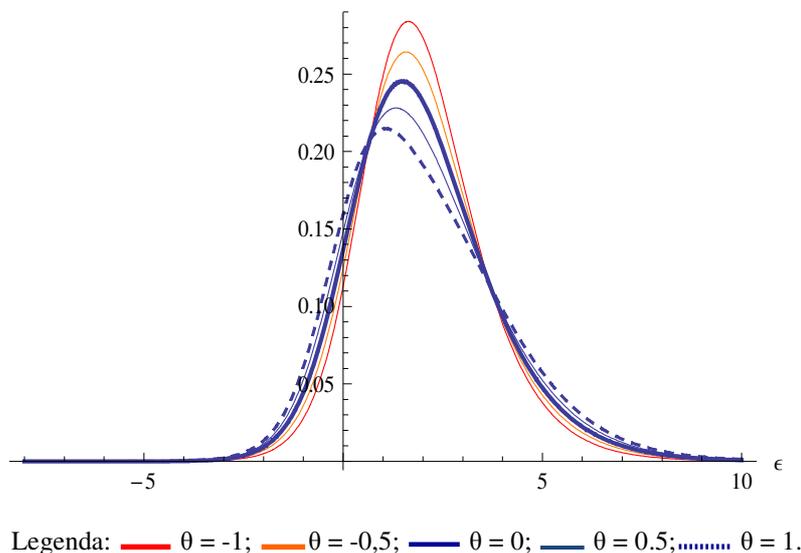
Il parametro di associazione in una copula FGM assume valori compresi nell'intervallo [-1, 1]. Secondo quanto emerge in Smith (2008), tutte le funzioni di densità giacciono all'interno delle due funzioni calcolate negli estremi dell'intervallo di definizione di  $\theta$ . Come si vede nella figura 3.1, le funzioni sono comprese tra la curva rossa ( $\theta = -1$ ) e quella tratteggiata ( $\theta = 1$ ).

Inoltre, come sottolinea Smith (2008), l'asimmetria mostrata dalla funzione di densità dell'errore composto può derivare dalla dipendenza tra l'errore casuale e la componente di inefficienza, in contrasto con quanto sostengono Kumbhakar e Lovell (2000), i quali attribuiscono tale asimmetria ad una scorretta specificazione del modello.

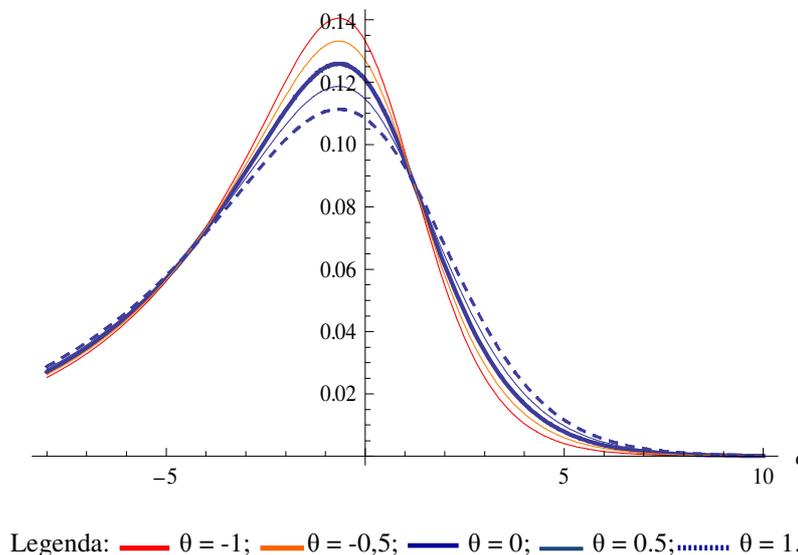
**Figura 3.1 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\alpha_v = 1$**



**Figura 3.2 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\alpha_v = 2$**



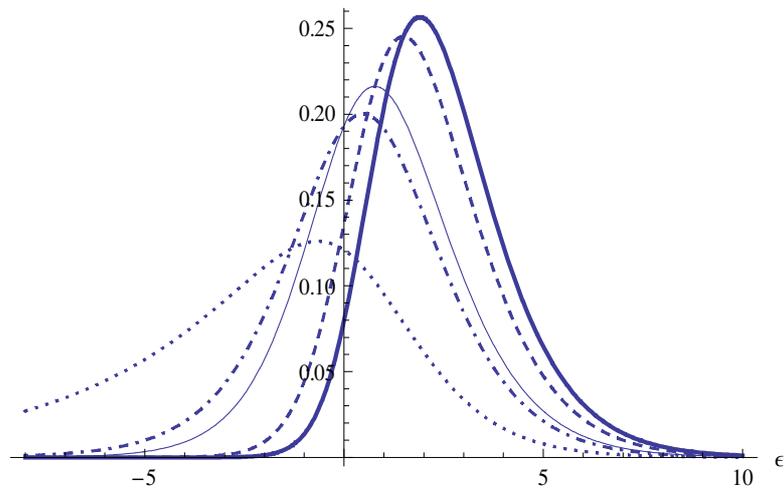
**Figura 3.3 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\alpha_v = 0.25$**



Dai grafici delle funzioni presenti nelle figure 3.1-3.3 emerge che la dipendenza tra le due componenti di errore della frontiera stocastica influenza la forma della distribuzione dell'errore composto in condizioni di simmetria della componente casuale e anche nel caso di asimmetria (sia essa positiva o negativa).

Inoltre, è utile capire quale sia il comportamento della funzione di distribuzione dell'errore composto  $f_{\theta}^{Ex}(\varepsilon)$  in presenza di asimmetria dell'errore casuale, dato un livello di dipendenza. Nelle figure 3.4-3.6 sono riportati i grafici della (3.51) dati alcuni valori per i parametri delle distribuzioni marginali ( $\delta_u = \delta_v = \delta = 1$  e  $\lambda_v = 0$ ) e considerando cinque valori del parametro di forma della distribuzione Log-Dagum assegnata alla componente casuale dell'errore ( $\alpha_v = 1, \alpha_v = 2, \alpha_v = 3, \alpha_v = 0.25$  e  $\alpha_v = 0.75$ )<sup>61</sup>, al variare del valore del parametro di associazione della copula,  $\theta$ . In particolare, si mostrano le funzioni di distribuzione di  $\varepsilon$  in presenza di indipendenza ( $\theta = 0$ ), di dipendenza negativa ( $\theta = -1$ ) e di dipendenza positiva ( $\theta = 1$ ).

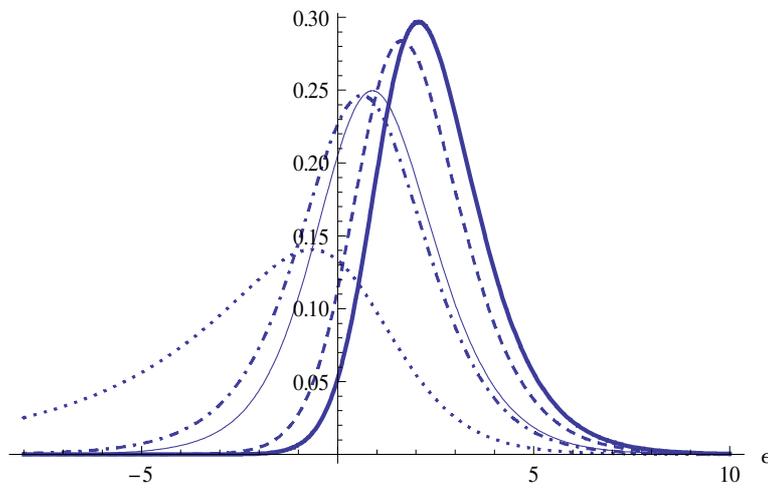
**Figura 3.4 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\theta = 0$**



Legenda: .....  $\alpha_v = 0.25$ ; - - -  $\alpha_v = 0.75$ ; —  $\alpha_v = 1$ ; - · - ·  $\alpha_v = 2$ ; —  $\alpha_v = 3$ .

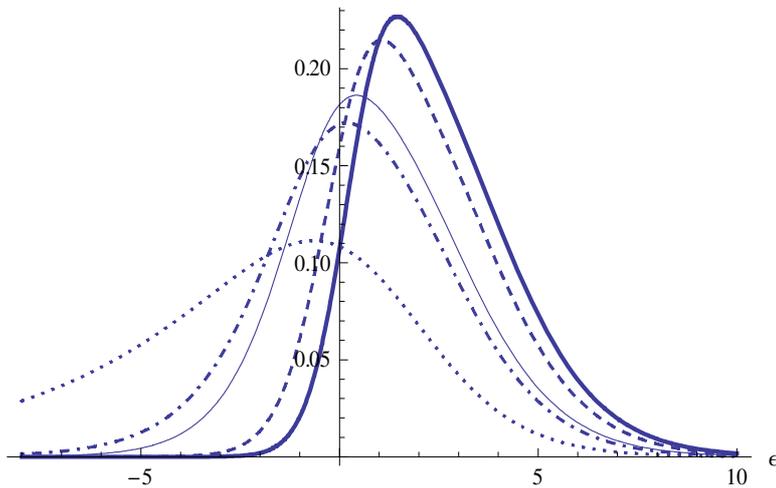
<sup>61</sup> Si consideri che per  $\alpha_v = 1$  si ha simmetria nella funzione di distribuzione dell'errore casuale, per  $0 < \alpha_v < 1$  si ha asimmetria negativa, mentre, per  $\alpha_v > 1$  si ha asimmetria positiva.

**Figura 3.5 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\theta = -1$**



Legenda: .....  $\alpha_v = 0.25$ ; - - -  $\alpha_v = 0.75$ ; —  $\alpha_v = 1$ ; - · - ·  $\alpha_v = 2$ ; —  $\alpha_v = 3$ .

**Figura 3.6 Funzione di densità del termine di errore composto  $f_{\theta}^{Ex}(\varepsilon)$  per  $\theta = 1$**



Legenda: .....  $\alpha_v = 0.25$ ; - - -  $\alpha_v = 0.75$ ; —  $\alpha_v = 1$ ; - · - ·  $\alpha_v = 2$ ; —  $\alpha_v = 3$ .

Dai grafici, si può notare che, per dati livelli di dipendenza, la funzione di distribuzione dell'errore cambia al variare delle condizioni di asimmetria.

Dopo aver derivato la funzione di distribuzione, è importante valutare le caratteristiche delle distribuzioni marginali di  $u$  e di  $v$  e quelle della funzione di distribuzione dell'errore

composto  $\varepsilon$  e come queste possano influenzare anche la scomposizione dell'errore composto nell'ambito delle frontiere stocastiche. A tal fine, si consideri che il valore atteso e la varianza della componente di inefficienza distribuita secondo una legge Esponenziale sono dati dalle seguenti espressioni:

$$E[U] = \delta_u \quad (3.52)$$

$$Var[U] = \delta_u^2 \quad (3.53)$$

Inoltre, il valore atteso e la varianza dell'errore casuale distribuito secondo una Logistica Generalizzata di I tipo assumono i valori (Domma, 2004; Domma e Perri, 2008):

$$E[V] = \lambda_v + \delta_v [\Psi(\alpha_v) - \Psi(1)] \quad (3.54)$$

$$Var[V] = \delta_v^2 [\Psi'(\alpha_v) + \Psi'(1)] \quad (3.55)$$

dove  $\Psi(\cdot)$  e  $\Psi'(\cdot)$  sono, rispettivamente, la funzioni matematiche Digamma e Trigamma, date dalla derivata prima e seconda del logaritmo di una funzione Gamma (Abramowitz e Stegun, 1972). Si evidenzia che Smith (2008) utilizza una copula FGM con marginali  $Exp(\delta_u)$  ed una Logistica Standard di parametro  $\delta_v$ , simmetrica intorno allo zero con varianza pari a  $Var(V) = \delta_v$ . Evidentemente, posti  $a_v = 1$  e  $\lambda_v = 0$  nella Logistica Generalizzata qui considerata, si ottiene il caso particolare studiato da Smith (2008).

Date le seguenti informazioni, si ottengono il valore atteso e la varianza dell'errore composto:

$$E[\varepsilon] = E[U] + E[V] = \delta_u + \lambda_v + \delta_v [\Psi(\alpha_v) - \Psi(1)] \quad (3.56)$$

$$Var[\varepsilon] = Var[U] + Var[V] + 2Cov[U, V] = \delta_u^2 + \delta_v^2 [\Psi'(\alpha_v) + \Psi'(1)] + 2Cov[U, V] \quad (3.57)$$

Evidentemente la determinazione della (3.57) richiede il calcolo della covarianza fra le due componenti di errore  $U$  e  $V$ . A tal fine, ricordando l'espressione della funzione di densità congiunta, si ha che:

$$\begin{aligned}
Cov[U, V] &= E[UV] - E[U]E[V] = \\
&= \int_0^{+\infty} \int_{-\infty}^{+\infty} uv f(u, v) dudv - E[U]E[V] = \\
&= \int_0^{+\infty} \int_{-\infty}^{+\infty} uv f(u) g(v) \{1 + \theta [1 - 2F(u)][1 - 2G(v)]\} dudv - E[U]E[V] = \\
&= \int_0^{+\infty} \int_{-\infty}^{+\infty} uv f(u) g(v) dudv + \theta \int_0^{+\infty} \int_{-\infty}^{+\infty} uv f(u) g(v) [1 - 2F(u)][1 - 2G(v)] dudv - E[U]E[V] = \\
&= E[U]E[V] + \theta \left\{ \left[ \int_0^{+\infty} u f(u) du - 2 \int_0^{+\infty} u f(u) F(u) du \right] \left[ \int_{-\infty}^{+\infty} v g(v) dv - 2 \int_{-\infty}^{+\infty} v g(v) G(v) dv \right] \right\} - E[U]E[V] = \\
&= \theta \{E[U] - 2 E[U \cdot F[U]]\} \{E[V] - 2 E[V \cdot G[V]]\}
\end{aligned} \tag{3.58}$$

Date le espressioni dei valori attesi dell'errore casuale  $u$  e della componente di inefficienza  $v$ , è necessario calcolare  $E[U \cdot F[U]]$  e  $E[V \cdot G[V]]$ . In formula:

$$\begin{aligned}
E[U \cdot F[U]] &= \int_0^{+\infty} u f(u) F(u) du = \int_0^{+\infty} u \left( 1 - e^{-\frac{u}{\delta_u}} \right) \frac{1}{\delta_u} e^{-\frac{u}{\delta_u}} du = \int_0^{+\infty} u \frac{1}{\delta_u} e^{-\frac{u}{\delta_u}} du - \int_0^{+\infty} u \frac{1}{\delta_u} e^{-\frac{2u}{\delta_u}} du = \\
&= E[U] - \frac{1}{2} \int_0^{+\infty} u \frac{2}{\delta_u} e^{-\frac{2u}{\delta_u}} du = \delta_u - \frac{\delta_u}{4} = \frac{3}{4} \delta_u
\end{aligned} \tag{3.59}$$

e, analogamente:

$$\begin{aligned}
E[V \cdot G[V]] &= \int_{-\infty}^{+\infty} v g(v) G(v) dv = \int_{-\infty}^{+\infty} v \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v-1} \frac{\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} dv = \\
&= \frac{1}{2} \int_{-\infty}^{+\infty} v \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-2\alpha_v-1} \frac{2\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} dv = \frac{\lambda_v}{2} + \frac{\delta_v}{2} [\Psi(2\alpha_v) - \Psi(1)]
\end{aligned} \tag{3.60}$$

Riprendendo l'espressione 3.58 e sostituendo la (3.59) e la (3.60), si ottiene:

$$\begin{aligned}
Cov[U, V] &= \theta \{E[U] - 2 E[U \cdot F[U]]\} \{E[V] - 2 E[V \cdot G[V]]\} = \\
&= \theta \left[ \delta_u - 2 \frac{3}{4} \delta_u \right] \left\{ \lambda_v + \delta_v [\Psi(\alpha_v) - \Psi(1)] - 2 \left[ \frac{\lambda_v}{2} + \frac{\delta_v}{2} [\Psi(2\alpha_v) - \Psi(1)] \right] \right\} = \\
&= \theta \left( -\frac{\delta_u}{2} \right) \{ \lambda_v + \delta_v [\Psi(\alpha_v) - \Psi(1)] - \lambda_v - \delta_v [\Psi(2\alpha_v) - \Psi(1)] \} = \\
&= \theta \left( -\frac{\delta_u}{2} \right) \{ \delta_v \Psi(\alpha_v) - \delta_v \Psi(1) - \delta_v \Psi(2\alpha_v) + \delta_v \Psi(1) \} = \\
&= \theta \left( -\frac{\delta_u}{2} \right) \{ \delta_v [\Psi(\alpha_v) - \Psi(2\alpha_v)] \} = \\
&= \theta \frac{\delta_u \delta_v}{2} [\Psi(2\alpha_v) - \Psi(\alpha_v)]
\end{aligned} \tag{3.61}$$

Al fine di valutare il risultato suesposto, si osservi che ponendo  $\alpha_v = 1$  nella (3.61) si ottiene, come caso particolare, l'analogo risultato proposto da Smith (2008) per una frontiera di produzione dato da:

$$Cov[U, V] = \theta \frac{\delta_u \delta_v}{2} [\Psi(1)+1 - \Psi(1)] = \theta \frac{\delta_u \delta_v}{2} \quad (3.62)$$

ottenuto sfruttando la relazione di ricorrenza della funzioni digamma secondo la quale  $\Psi(z+1) = \Psi(z) + \frac{1}{z}$ , per cui si ha che  $\Psi(2\alpha_v) = \Psi(2) = \Psi(1+1) = \Psi(1) + 1$ . Confrontando la (3.61) con la (3.62) risulta immediato constatare che la forma della distribuzione incide direttamente sulla covarianza tra gli errori.

Dopo aver calcolato la covarianza fra i due termini di errore (3.61), si può calcolare la varianza dell'errore composto, che rappresenta un parametro utilizzato per la sua scomposizione e, quindi, per la definizione dell'efficienza.

In formula, riprendendo la (3.57), si ottiene la seguente espressione:

$$\begin{aligned} Var[\varepsilon] &= Var[U] + Var[V] + 2Cov[U, V] = \delta_u^2 + \delta_v^2 [\Psi'(\alpha_v) + \Psi'(1)] + 2Cov[U, V] = \\ &= \delta_u^2 + \delta_v^2 [\Psi'(\alpha_v) + \Psi'(1)] + 2\theta \frac{\delta_u \delta_v}{2} [\Psi(2\alpha_v) - \Psi(\alpha_v)] = \\ &= \delta_u^2 + \delta_v^2 [\Psi'(\alpha_v) + \Psi'(1)] + \theta \delta_u \delta_v [\Psi(2\alpha_v) - \Psi(\alpha_v)] \end{aligned} \quad (3.63)$$

Al fine di valutare l'espressione della varianza dell'errore composto, innanzitutto, si evidenzia che la funzione  $\Psi(\cdot)$  è una funzione sempre crescente (Abramowitz e Stegun, 1972), quindi,  $\forall \alpha_v > 0$  vale la seguente disuguaglianza:

$$\Psi(2\alpha_v) > \Psi(\alpha_v) \quad (3.64)$$

Questa osservazione permette di dire che la quantità  $[\Psi(2\alpha_v) - \Psi(\alpha_v)]$  è sempre maggiore di zero. Detto che la funzione  $\Psi'(\cdot)$  è sempre positiva, è evidente che quando le due componenti di errore sono positivamente dipendenti (cioè quando  $\theta > 0$ ), la varianza dell'errore composto risulta più elevata rispetto al caso di indipendenza (cioè quando  $\theta = 0$ ). Al contrario, quando la dipendenza ha segno negativo (cioè quando  $\theta < 0$ ), la varianza dell'errore composto si riduce.

Inoltre, la forma della distribuzione dell'errore casuale, misurata dal parametro  $\alpha_v$ <sup>62</sup>, influenza la varianza dell'errore composto. Nel lavoro di Smith (2008) non si riesce a catturare un'eventuale asimmetria perché l'errore casuale  $v$  si distribuisce secondo una funzione Logistica Standard che è simmetrica. Se  $v$  fosse asimmetrico in un modello che non riesce a catturare tale caratteristica di forma, allora la varianza di  $\varepsilon$  sarebbe distorta. Nel metodo delle frontiere stocastiche questo comporta forti distorsioni nella scomposizione dell'errore composto e, quindi, nella determinazione del livello di efficienza. Nell'ambito di questa metodologia, una delle tecniche utilizzate per la scomposizione dell'errore composto è quella proposta da Jondrow *et al.* (1982)<sup>63</sup>, la quale considera il caso in cui l'errore casuale si distribuisce secondo una Normale e la componente di inefficienza secondo una semi-Normale. Tale tecnica utilizza il coefficiente  $\frac{\sigma_u^2}{\sigma^2}$  al fine di scorporare la componente di inefficienza  $u$  da  $\varepsilon$ , dove  $\sigma_u^2$  rappresenta la varianza di  $u$  e  $\sigma^2 = \sigma_u^2 + \sigma_v^2$  la varianza dell'errore composto  $\varepsilon$ , con  $\sigma_v^2$  che rappresenta la varianza dell'errore casuale  $v$ . Il valore zero di tale parametro indica che deviazioni dalla frontiera sono dovute solo all'errore casuale, mentre valori prossimi a uno comportano che la distanza dalla frontiera sia dovuta all'inefficienza.

Il contributo principale del lavoro di Jondrow *et al.* (1982) è rappresentato dalla definizione dei parametri di locazione e di scala della funzione di distribuzione condizionata di  $u$  data da una Normale di media  $\mu_*$  e varianza  $\sigma_*^2$ , ovvero  $N(\mu_*, \sigma_*^2)$ , dove  $\mu_* = -\frac{\varepsilon \sigma_u^2}{\sigma^2}$  e  $\sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma^2}$  (Jondrow *et al.*, 1982). Pertanto, si può notare che la definizione della componente di inefficienza e, quindi, del livello di efficienza dipende in modo particolare dalla varianza dell'errore composto.

### 3.6 Un approccio metodologico per la costruzione della funzione di distribuzione dell'errore composto mediante la copula FGM e marginali Log-Dagum

In questo paragrafo, si presenta il metodo che consente di giungere alla forma esplicita della funzione di distribuzione dell'errore composto  $f_\theta(\varepsilon)$  caratterizzata dalle marginali Log-

<sup>62</sup> Tale parametro misura sia l'asimmetria che la curtosi di una distribuzione Log-Dagum (Domma, 2004).

<sup>63</sup> Le stime dei livelli di efficienza presenti nel capitolo 2 sono state realizzate mediante la tecnica di Jondrow *et al.* (1982) (cfr. § 2.5; Appendice A2).

Dagum e Log-Dagum troncata allo zero e dalla dipendenza modellata attraverso la copula FGM<sup>64</sup>.

L'idea di inserire marginali diverse da quelle Normali nasce dalla necessità di costruire modelli più flessibili che riescano a catturare caratteristiche distributive che l'ipotesi di normalità non consente di fare. La distribuzione Log-Dagum, infatti, può assumere forme diverse a seconda dei valori assunti dai parametri che la governano presentando anche asimmetria e curtosi (Domma, 2002; Domma, 2004).

Si ipotizza che la componente di inefficienza si distribuisca secondo una funzione Log-Dagum troncata allo 0,  $U \sim LDa_0(\alpha_u, \delta_u, \lambda_u)$ , con funzione di densità e di ripartizione, rispettivamente:

$$f_0(u; \alpha_u, \delta_u, \lambda_u) = \frac{f(u; \alpha_u, \delta_u, \lambda_u)}{1 - F(0)} = \frac{\frac{\alpha_u}{\delta_u} e^{-\frac{u-\lambda_u}{\delta_u}} \left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u - 1}}{1 - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}} \quad (3.65)$$

$$F_0(u; \alpha_u, \delta_u, \lambda_u) = \frac{F(u; \alpha_u, \delta_u, \lambda_u) - F(0)}{1 - F(0)} = \frac{\left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u} - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}}{1 - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}} \quad (3.66)$$

con  $u > 0$  e  $\alpha_u, \delta_u, \lambda_u > 0$ , dove  $\lambda_u$  è il parametro di scala, mentre  $\alpha_u$  e  $\delta_u$  regolano la forma della distribuzione<sup>65</sup>.

In questo modello, l'errore casuale si distribuisce secondo una funzione Log-Dagum,  $V \sim LDa(\alpha_v, \delta_v, \lambda_v)$ . La funzione di densità e quella di ripartizione assumono le seguenti espressioni:

$$g(v; \alpha_v, \delta_v, \lambda_v) = \frac{\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} \left(1 + e^{-\frac{v-\lambda_v}{\delta_v}}\right)^{-\alpha_v - 1} \quad (3.67)$$

<sup>64</sup> In particolare, si utilizza la stessa ri-parametrizzazione della Log-Dagum proposta nel paragrafo 3.5.

<sup>65</sup> In Domma (2002), si dimostra che  $\alpha_u$  rappresenta il parametro di simmetria della funzione.

$$G(v; \alpha_v, \delta_v, \lambda_v) = \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \quad (3.68)$$

con  $v \in \mathfrak{R}$  e  $\alpha_v, \delta_v, \lambda_v > 0$ , dove  $\lambda_v$  è il parametro di locazione,  $\alpha_v$  è un parametro che regola la forma della distribuzione, mentre  $\delta_v$  è il parametro di scala.

Sotto tali ipotesi, la copula FGM risulta:

$$C_{\theta}^{LDa}(F(u), G(v)) = \left[ \frac{\left( 1 + e^{-\frac{u-\lambda_u}{\delta_u}} \right)^{-\alpha_u}}{1 - \left( 1 + e^{-\frac{\lambda_u}{\delta_u}} \right)^{-\alpha_u}} \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \right] \quad (3.69)$$

$$\left\{ 1 + \theta \left[ 1 - \frac{\left( 1 + e^{-\frac{u-\lambda_u}{\delta_u}} \right)^{-\alpha_u}}{1 - \left( 1 + e^{-\frac{\lambda_u}{\delta_u}} \right)^{-\alpha_u}} \right] \left[ 1 - \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \right] \right\}$$

Conseguentemente, la *density copula* assume la seguente forma:

$$c_{\theta}^{LDa}(F(u), G(v)) = 1 + \theta \left[ 1 - 2 \left[ 1 - \frac{\left( 1 + e^{-\frac{u-\lambda_u}{\delta_u}} \right)^{-\alpha_u}}{1 - \left( 1 + e^{-\frac{\lambda_u}{\delta_u}} \right)^{-\alpha_u}} \right] \left[ 1 - 2 \left( 1 + e^{-\frac{v-\lambda_v}{\delta_v}} \right)^{-\alpha_v} \right] \right] \quad (3.70)$$

Pertanto, la seguente espressione rappresenta la funzione di distribuzione congiunta:

$$f_{\theta}^{LDa}(u, v) = \frac{\frac{\alpha_u}{\delta_u} e^{-\frac{u-\lambda_u}{\delta_u}} \left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u-1}}{1 - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}} \frac{\frac{\alpha_v}{\delta_v} e^{-\frac{v-\lambda_v}{\delta_v}} \left(1 + e^{-\frac{v-\lambda_v}{\delta_v}}\right)^{-\alpha_v-1}}{\left\{1 + \theta \left[1 - 2 \frac{\left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u}}{1 - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}}\right] \left[1 - 2 \left(1 + e^{-\frac{v-\lambda_v}{\delta_v}}\right)^{-\alpha_v}\right]\right\}}. \quad (3.71)$$

Sostituendo  $\mathcal{E} = u + v$  nell'espressione 3.71 e marginalizzando rispetto a  $u$ , si ottiene la funzione di distribuzione dell'errore composto  $f_{\theta}^{LDa}(\mathcal{E})$ , la quale rappresenta l'argomento della funzione di verosimiglianza. In formula:

$$f_{\theta}^{LDa}(\mathcal{E}) = \int_0^{+\infty} f_{\theta}^{LDa}(u, v) du = \int_0^{+\infty} f(u) g(\mathcal{E}-u) \{1 + \theta [1 - 2 F(u)] [1 - 2 G(\mathcal{E}-u)]\} du \quad (3.72)$$

Analogamente a quanto esposto nel paragrafo precedente, si giunge alla seguente espressione dell'integrale tipo:

$$I = \int_0^{\infty} k^{-2} \frac{\alpha_u}{\delta_u} e^{-\frac{u-\lambda_u}{\delta_u}} \left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u-1} \frac{\alpha_v}{\delta_v} e^{-\frac{(\mathcal{E}-u)-\lambda_v}{\delta_v}} \left(1 + e^{-\frac{(\mathcal{E}-u)-\lambda_v}{\delta_v}}\right)^{-\alpha_v-1} \left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-\alpha_u} \left(1 + e^{-\frac{(\mathcal{E}-u)-\lambda_v}{\delta_v}}\right)^{-\alpha_v} du =$$

$$= k^{-2} \underbrace{\frac{\alpha_u}{\delta_u} \frac{\alpha_v}{\delta_v} e^{\frac{\lambda_u}{\delta_u}} e^{-\frac{\mathcal{E}-\lambda_v}{\delta_v}}}_A \int_0^{\infty} e^{-\frac{u}{\delta_u}} \left(1 + e^{-\frac{u-\lambda_u}{\delta_u}}\right)^{-2\alpha_u-1} e^{\frac{u}{\delta_v}} \left(1 + e^{-\frac{(\mathcal{E}-u)-\lambda_v}{\delta_v}}\right)^{-2\alpha_v-1} du \quad (3.73)$$

$$\text{dove } K = 1 - \left(1 + e^{-\frac{\lambda_u}{\delta_u}}\right)^{-\alpha_u}$$

Dopo alcuni passaggi algebrici non riportati per brevità e usando la trasformazione  $y = e^{-u}$ , l'integrale tipo assume la seguente espressione:

$$I = Ac_2^{-2\alpha_v-1} \int_0^1 y^{\frac{2\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left(1 + c_1 y \frac{1}{\delta_u}\right)^{-2\alpha_u-1} \left(1 + c_2^{-1} y \frac{1}{\delta_v}\right)^{-2\alpha_v-1} dy \quad (3.74)$$

dove  $A = \frac{\alpha_u \alpha_v c_1 c_2}{K^2 \delta_u \delta_v}$ , con  $c_1 = e^{\frac{\lambda_u}{\delta_u}}$  e  $c_2 = e^{\frac{\varepsilon - \lambda_v}{\delta_v}}$ .

Come già sottolineato nel paragrafo precedente, si può utilizzare lo sviluppo binomiale e sfruttare la seguente uguaglianza:

$$\left(1 + c_1 y \frac{1}{\delta_u}\right)^{-2\alpha_u-1} = \sum_{j=0}^{\infty} \frac{(-1)^j (2\alpha_u + 1)_j}{j!} y^{\frac{j}{\delta_u}} c_1^j \quad (3.75)$$

$$\left(1 + c_2^{-1} y \frac{1}{\delta_v}\right)^{-2\alpha_v-1} = \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i!} y^{\frac{i}{\delta_v}} c_2^{-i} \quad (3.76)$$

Sostituendo (3.75) e (3.76) nell'espressione di  $I$  (3.74) e dopo alcuni passaggi algebrici, si ottiene:

$$\begin{aligned} I &= Ac_2^{-2\alpha_v-1} \sum_{j=0}^{\infty} \frac{(-1)^j (2\alpha_u + 1)_j c_1^j}{j!} \sum_{i=0}^{\infty} \frac{(-1)^i (2\alpha_v + 1)_i}{i! c_2^i} \int_0^1 y^{\frac{2\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1 + \frac{j}{\delta_u} + \frac{i}{\delta_v}} dy = \\ &= Ac_2^{-2\alpha_v-1} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(2\alpha_u + 1)_j (2\alpha_v + 1)_i}{i! j!} (-c_1)^j (-c_2)^{-i} \frac{1}{\frac{1}{\delta_v} (2\alpha_v + i) + \frac{1}{\delta_u} (1 + j)} \end{aligned} \quad (3.77)$$

Se si considera l'ipotesi secondo la quale  $\delta_u = \delta_v = \delta$ , allora si può scrivere:

$$\frac{1}{\frac{1}{\delta_v} (2\alpha_v + i) + \frac{1}{\delta_u} (1 + j)} = \frac{\delta}{2\alpha_v + i + j + 1}.$$

Inoltre, utilizzando la relazione ricorsiva delle funzioni Gamma, secondo la quale  $\Gamma(n+1) = n\Gamma(n)$ , si ha che:

$$\frac{(2\alpha_v + 1)_{i+j}}{(2\alpha_v + 2)_{i+j}} = \frac{\frac{\Gamma(2\alpha_v + i + j + 1)}{\Gamma(2\alpha_v + 1)}}{\frac{\Gamma(2\alpha_v + i + j + 2)}{\Gamma(2\alpha_v + 2)}} = \frac{\Gamma(2\alpha_v + i + j + 1)}{\Gamma(2\alpha_v + 1)} \frac{(2\alpha_v + 1)\Gamma(2\alpha_v + 1)}{(2\alpha_v + i + j + 1)\Gamma(2\alpha_v + i + j + 1)} = \frac{2\alpha_v + 1}{2\alpha_v + i + j + 1} \quad (3.78)$$

dalla quale si ottiene:

$$\frac{1}{2\alpha_v + i + j + 1} = \frac{(2\alpha_v + 1)_{i+j}}{(2\alpha_v + 2)_{i+j}} \frac{1}{(2\alpha_v + 1)} \quad (3.79)$$

Sostituendo la (3.79) nella (3.77), riordinando i termini e considerando la (3.36), si ottiene la soluzione dell'integrale tipo, ovvero una funzione ipergeometrica generalizzata data dalla seguente espressione:

$$\begin{aligned} I &= \frac{Ac_2^{-2\alpha_v - 1}}{(2\alpha_v + 1)} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(2\alpha_v + 1)_{i+j} (2\alpha_u + 1)_j (2\alpha_v + 1)_i}{(2\alpha_v + 1)_{i+j}} \frac{(-c_1)^j}{j!} \frac{(-c_2)^{-i}}{i!} = \\ &= \frac{Ac_2^{-2\alpha_v - 1}}{(2\alpha_v + 1)} F_{1:0;0}^{1:1;1} \left[ (2\alpha_v + 1); (2\alpha_v + 1); (2\alpha_u + 1); (2\alpha_v + 2); 0; 0; c_1, c_2^{-1} \right] \end{aligned} \quad (3.80)$$

Tale risultato consente di affermare che la distribuzione dell'errore composto sia una somma ponderata di integrali di tipo  $I$ .

In particolare, seguendo lo stesso procedimento per gli altri tre elementi della funzione di distribuzione congiunta (3.30) e analogamente all'integrale tipo si ha che:

$$I_{(1)} = \int_0^{\infty} f_0(u)g(\varepsilon - u)du = A_2 c_2^{-\alpha_v - 1} \int_0^1 y^{\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left( 1 + c_1 y^{\frac{1}{\delta_u}} \right)^{-\alpha_u - 1} \left( 1 + c_2^{-1} y^{\frac{1}{\delta_v}} \right)^{-\alpha_v - 1} dy \quad (3.81)$$

$$\text{con } A_2 = \frac{\alpha_u \alpha_v c_1 c_2}{K \delta_u \delta_v}$$

$$I_{(2)} = \int_0^{\infty} f_0(u)g(\varepsilon - u)F_0(u)du = Ac_2^{-\alpha_v - 1} \int_0^1 y^{\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left( 1 + c_1 y^{\frac{1}{\delta_u}} \right)^{-2\alpha_u - 1} \left( 1 + c_2^{-1} y^{\frac{1}{\delta_v}} \right)^{-\alpha_v - 1} dy \quad (3.82)$$

$$I_{(3)} = \int_0^{\infty} f_0(u)g(\varepsilon - u)G(\varepsilon - u)du = A_2 c_2^{-2\alpha_v - 1} \int_0^1 y^{2\frac{\alpha_v}{\delta_v} + \frac{1}{\delta_u} - 1} \left(1 + c_1 y \frac{1}{\delta_u}\right)^{-\alpha_u - 1} \left(1 + c_2^{-1} y \frac{1}{\delta_v}\right)^{-2\alpha_v - 1} dy \quad (3.83)$$

Ripetendo i passaggi seguiti per giungere alla forma esplicita dell'integrale tipo (che non sono riportati per brevità), si ottengono le forme esplicite delle altre tre componenti della funzione dell'errore composto:

$$I_{(1)} = \frac{A_2 c_2^{-\alpha_v - 1}}{(\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(\alpha_v + 1); (\alpha_v + 1); (\alpha_u + 1); (\alpha_v + 2); 0; 0; c_1, c_2^{-1}] \quad (3.84)$$

$$I_{(2)} = \frac{A c_2^{-\alpha_v - 1}}{(\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(\alpha_v + 1); (\alpha_v + 1); (2\alpha_u + 1); (\alpha_v + 2); 0; 0; c_1, c_2^{-1}] \quad (3.85)$$

$$I_{(3)} = \frac{A_2 c_2^{-2\alpha_v - 1}}{(2\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(2\alpha_v + 1); (2\alpha_v + 1); (\alpha_u + 1); (2\alpha_v + 2); 0; 0; c_1, c_2^{-1}] \quad (3.86)$$

Pertanto la funzione di distribuzione dell'errore composto assume la seguente forma esplicita:

$$\begin{aligned} f_{\theta}^{LDa}(\varepsilon) = & (1 + \theta) \frac{A_2 c_2^{-\alpha_v - 1}}{(\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(\alpha_v + 1); (\alpha_v + 1); (\alpha_u + 1); (\alpha_v + 2); 0; 0; c_1, c_2^{-1}] - \\ & - 2\theta \frac{A c_2^{-\alpha_v - 1}}{(\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(\alpha_v + 1); (\alpha_v + 1); (2\alpha_u + 1); (\alpha_v + 2); 0; 0; c_1, c_2^{-1}] - \\ & - 2\theta \frac{A_2 c_2^{-2\alpha_v - 1}}{(2\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(2\alpha_v + 1); (2\alpha_v + 1); (\alpha_u + 1); (2\alpha_v + 2); 0; 0; c_1, c_2^{-1}] + \\ & + 4\theta \frac{A c_2^{-2\alpha_v - 1}}{(2\alpha_v + 1)} F_{1;0;0}^{1;1;1} [(2\alpha_v + 1); (2\alpha_v + 1); (2\alpha_u + 1); (2\alpha_v + 2); 0; 0; c_1, c_2^{-1}] \end{aligned} \quad (3.87)$$

Si tratta di un risultato originale rispetto alla letteratura esistente, poiché dà origine ad un modello generale costruito considerando l'ipotesi di dipendenza fra le componenti erratiche di una frontiera stocastica e due funzioni di distribuzione marginali diverse da quelle tradizionalmente utilizzate in questa tipologia di studi.

### 3.7 Conclusioni

L'obiettivo del presente capitolo è rappresentato dal tentativo di proporre modelli più generali di frontiere stocastiche che possano catturare la dipendenza fra l'errore casuale e la componente di inefficienza e caratteristiche distributive diverse da quelle presentate da distribuzioni Normali.

Seguendo il lavoro proposto da Smith (2008) e dopo aver presentato le caratteristiche principali delle funzioni copule bivariate e la loro utilità nell'ambito delle frontiere stocastiche, si costruisce un modello in cui l'unica variante rispetto ai modelli presenti nel capitolo precedente è l'introduzione della dipendenza tra i due termini di errore mediante una copula FGM. Dalle stime ottenute emergono risultati tali da giustificare l'ipotesi di dipendenza tra le due componenti di errore. Inoltre, si ottengono stime più convincenti dei parametri associati alle determinanti dell'inefficienza rispetto a quanto emerge dalle stime sotto l'ipotesi di indipendenza.

I risultati ottenuti a favore della opportunità di considerare modelli che non escludono a priori la dipendenza fra le due componenti di errore sono alla base del tentativo di proporre un approccio di tipo metodologico per la costruzione di un modello utilizzabile in più contesti.

In particolare, si giunge a due importanti conclusioni. La prima è relativa all'importanza di considerare la dipendenza fra l'errore casuale e la componente di inefficienza. A riguardo si è empiricamente mostrato come la dipendenza tra queste due componenti possa essere statisticamente significativa nel caso dell'utilizzo di una funzione copula FGM e marginali Normali. Inoltre, un'ulteriore motivazione alla base dell'ipotesi di dipendenza è la dimostrazione analitica che il parametro di associazione  $\theta$  influenza la varianza dell'errore composto e, quindi, la stima dell'efficienza. La seconda conclusione del capitolo è relativa al vantaggio di introdurre nel modello funzioni marginali per le due componenti di errore più generali rispetto a quelle Normali, in modo da catturare caratteristiche distributive diverse. Infatti, si è dimostrato come anche l'asimmetria e la curtosi della funzione di distribuzione dell'errore casuale della frontiera stocastica influenzino la varianza dell'errore composto.

Si tratta di risultati relativamente originali da un punto di vista metodologico e con enormi potenzialità di sviluppi futuri, anche di tipo empirico.

## Conclusioni

L'efficienza del settore bancario italiano è un tema analizzato da molti autori perché l'industria bancaria è stata oggetto di radicali ristrutturazioni negli ultimi vent'anni. Un ulteriore motivo di interesse al tema è dovuto allo sviluppo di numerose tecniche di stima dell'efficienza che hanno trovato ampio spazio applicativo al caso delle banche. Per quanto riguarda il sistema bancario italiano, dalla rassegna della letteratura presente nel primo capitolo di questa tesi, si evince che non ci sono chiare ed univoche indicazioni in merito agli effetti sull'efficienza determinati dai processi di consolidamento osservati a partire dagli anni novanta. Peraltro, dall'analisi della letteratura emerge un secondo elemento che rappresenta la motivazione della tesi. Ci si riferisce al fatto che i risultati relativi all'impatto delle riforme del settore bancario sono molto sensibili al modello utilizzato per stimare l'efficienza.

In questo lavoro, dopo aver illustrato le caratteristiche del sistema bancario italiano e l'evoluzione dei processi di consolidamento, si propone la stima dei livelli di efficienza di costo delle banche italiane nel periodo compreso tra il 2006 e il 2010. La metodologia di stima è rappresentata dal metodo parametrico delle frontiere stocastiche. Inoltre, in tutte le applicazioni presenti nella tesi, per la scelta delle variabili da includere nella funzione di costo si fa riferimento all'approccio di intermediazione (Sealey e Lindley, 1977), poiché è quello che più da vicino tiene conto della particolarità dei servizi di intermediazione offerti dalle banche. Seguendo tale approccio, la scelta degli output bancari è ricaduta sui prestiti verso i clienti, sulle commissioni attive e sui titoli in portafoglio, mentre i fattori produttivi sono il lavoro, il capitale delle banche e i depositi bancari.

Nell'analizzare un periodo molto recente (2006-2010), i dati utilizzati sono estratti dall'archivio *ABI Banking Data* dell'Associazione Bancaria Italiana (ABI). Rispetto alla letteratura di riferimento sul caso italiano, si stima, pertanto, l'efficienza delle banche utilizzando per la prima volta i dati di bilancio delle banche imputati seguendo i principi contabili internazionali (IAS).

Le stime presentate nel capitolo II sono state effettuate seguendo la specificazione delle frontiere stocastiche proposta da Battese e Coelli (1995), la quale prevede la stima simultanea della frontiera di costo e dell'equazione dell'inefficienza. Tale specificazione consente di ottenere direttamente le stime dei coefficienti relativi ai fattori individuali e di contesto che influenzano l'inefficienza in modo simultaneo alla stima della funzione di costo. E' un metodo che contrasta con l'approccio "two steps" proposto da Greene (1993), il quale prevede la stima sequenziale dei due modelli. In questo modo, seguendo il modello di Battese e Coelli (1995) si tiene conto dell'eterogeneità nel campione simultaneamente alla stima della

funzione di costo in modo da evitare eventuali problemi di distorsione legati alla procedura “two steps” (Battaglia *et al.*, 2010; Lensink e Meesters, 2012). Inoltre, un ulteriore aspetto che caratterizza il lavoro risiede nella specificazione dell’equazione di efficienza, in quanto si sono considerati numerosi fattori individuali e di contesto in grado di influenzare l’efficienza delle banche (qualità del credito, solvibilità, dimensione, forma giuridica, localizzazione).

Dall’analisi dei risultati ottenuti utilizzando dati panel emerge che le BCC raggiungono i livelli più elevati di efficienza tanto da trainare l’evoluzione dell’intero settore. Inoltre, le banche piccole e le banche minori ottengono risultati più soddisfacenti rispetto alle banche medio-grandi. Da un punto di vista territoriale, le BCC del Centro e del Nord-est risultano essere più efficienti rispetto alle BCC della altre aree del paese.

Come già accennato, il vantaggio di utilizzare la specificazione delle frontiere stocastiche proposta da Battese e Coelli (1995) è quello di ottenere le stime degli effetti di variabili individuali e di contesto sull’efficienza degli istituti di credito. A tal proposito, emerge che una minore qualità dei crediti comporta maggiore efficienza di costo. E’ un risultato in linea con l’ipotesi di *bad management* (Berger e De Young, 1997), così come emerge anche dal test di Granger. L’interpretazione è che il risparmio di costo che si ottiene riducendo alcune tipologie di spese di selezione e monitoraggio dei mutuatari comporta, sì, guadagni di efficienza di costo, ma, allo stesso tempo, le banche selezionano progetti più rischiosi che presentano maggiore probabilità di insolvenza.

Per quanto riguarda le altre determinanti inserite nell’equazione dell’inefficienza, si ottiene che una maggiore solvibilità degli istituti di credito e una maggiore concentrazione del settore determinano il raggiungimento di più elevati livelli di efficienza di costo. Inoltre, le banche che presentano una scarsa efficienza gestionale, ovvero le banche che registrano livelli più elevati del rapporto *Cost/Income*, raggiungono livelli di efficienza di costo più bassi.

Lo studio presentato nel capitolo II sulle determinanti dei livelli di efficienza offre un modo per indirizzare le scelte strategiche, di *corporate* e operative delle banche allo scopo di tenere sotto “controllo” i fattori che ne determinano le performance. A tal fine, diventa importante specificare modelli econometrici che consentano di ridurre problemi di distorsione ed errata misurazione del fenomeno. Su questo aspetto è utile ricordare che tutti i modelli che utilizzano l’approccio delle frontiere stocastiche - discussi e presentati nei capitoli I e II della tesi - si basano su specifiche ipotesi sulle componenti dell’inefficienza e dell’errore casuale. In particolare, tutti gli autori ipotizzano che queste due componenti siano indipendenti e che, nella quasi totalità dei casi, si distribuiscano secondo funzioni Normali. La discussione sulla validità di queste due ipotesi rappresenta il punto di partenza del terzo capitolo della tesi,

poiché è ragionevole pensare che l'errore casuale e l'inefficienza si distribuiscano secondo variabili casuali più flessibili rispetto alle funzioni Normali. Inoltre, appare opportuno introdurre l'ipotesi di dipendenza tra le due componenti, poiché da un punto di vista strettamente statistico questa ipotesi incorpora, come caso specifico, quella di indipendenza.

In base a tali considerazioni, nell'ultimo capitolo di questo lavoro, si propone un metodo che permette di incorporare nell'approccio delle frontiere stocastiche sia distribuzioni diverse da quelle Normali sia l'ipotesi di dipendenza. Secondo quanto è stato avviato nel contributo di Smith (2008), si utilizzano le funzioni copule per la costruzione di modelli più generali che riescano a misurare caratteristiche distributive che le leggi Normali non catturano e a modellare la dipendenza tra le due componenti di errore di una frontiera stocastica.

Il primo elemento da evidenziare è che la rimozione dell'ipotesi di indipendenza fra le due componenti erratiche ha prodotto interessanti risultati sia da un punto di vista empirico sia da un punto di vista metodologico.

Da un punto di vista empirico, si è sottoposto a verifica econometrica un modello in cui l'unica variante rispetto ai modelli "tradizionali" è rappresentata dall'introduzione della dipendenza tra i due termini di errore mediante la copula Farlie-Gumbel-Morgenstern (FGM). La procedura di stima è basata sulla massimizzazione della funzione di verosimiglianza che è stata analiticamente derivata al fine di incorporare sia la funzione di costo sia quella dell'efficienza, secondo la specificazione di Battese e Coelli (1995). Il fatto di stimare simultaneamente le due equazioni, rappresenta un ulteriore approfondimento anche del modello di Smith (2008), poiché questo autore utilizza la specificazione precedentemente proposta da Battese e Coelli (1992) che prevede la stima della sola funzione di costo. Le stime sono state effettuate su dati cross-section per ciascun anno del periodo 2006-2010.

Dall'analisi emergono risultati tali da giustificare l'ipotesi di dipendenza tra le due componenti di errore, poiché, per alcuni anni, si ottengono stime statisticamente significative del parametro di associazione della copula. Inoltre, si ottengono stime che sono economicamente più convincenti rispetto ai risultati presentati nel capitolo II per dati cross-section e relativi a modelli che ipotizzano l'indipendenza tra le componenti dell'errore composto. In particolare, la più importante differenza rispetto alle stime cross-section ottenute con il modello "tradizionale" è che in caso di dipendenza cambia il segno del coefficiente associato alla qualità dei crediti. In questo caso, infatti, risulta statisticamente significativo e assume valori negativi in tutti gli anni. E' un dato coerente con la visione secondo la quale ad aumenti della qualità dei crediti corrispondono aumenti (diminuzioni) di inefficienza

(efficienza) di costo. Questa evidenza è in linea con l'ipotesi di *bad management* già discussa in precedenza relativa alle stime panel.

La prima considerazione che la stima di questo modello consente di formulare riguarda l'importanza di considerare metodi che non escludono a priori la dipendenza fra le due componenti di errore. Per questo motivo, nella tesi si propone un approccio metodologico per la costruzione di un modello utilizzabile per misurare l'efficienza in tutti i casi in cui le due componenti dell'errore composto siano dipendenti e si distribuiscano secondo funzioni più flessibili e più generali delle funzioni Normali.

Due sono le implicazioni più importanti di questo approccio. La prima è relativa all'importanza, derivata, come si è visto, anche empiricamente, di considerare dipendenza fra l'errore casuale e la componente di inefficienza. Nella tesi si dimostra che questa dipendenza, misurata mediante l'approccio delle funzioni copule dal parametro  $\theta$ , influenza la varianza dell'errore composto  $\varepsilon$ , la quale viene utilizzata per la scomposizione di  $\varepsilon$  nella componente di inefficienza e nella componente casuale. E' evidente che si ottengono stime distorte dell'efficienza quando esiste dipendenza e si considerano, invece, modelli che la escludono.

Il secondo importante risultato deriva dall'introduzione nel modello delle frontiere stocastiche di funzioni marginali per le due componenti di errore più generali rispetto a quelle Normali. In tal modo, è possibile studiare e modellare i casi in cui le caratteristiche distributive non siano riconducibili al caso di normalità. Infatti, si è anche dimostrato che il parametro che regola la forma della distribuzione assegnata all'errore casuale della frontiera stocastica (ovvero, una Logistica Generalizzata di I tipo) ha un impatto sulla varianza dell'errore composto e, quindi, sulle stime dei livelli di efficienza.

Si tratta di risultati originali che costituiscono il punto di partenza per futuri sviluppi metodologici ed empirici.

## Riferimenti bibliografici

- Abramowitz, M., Stegun, I. A., (1972) *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, New York
- Aiello, F., Mastromarco, C., Zago, A., (2010) “Be Productive or Face Decline. On the Sources and Determinants of Output Growth in Italian Manufacturing Firms”, *Empirical Economics*, Online First, 9 Ottobre
- Alessandrini, P., Papi, L., Zazzaro, A., (2003) “Banche, territorio e sviluppo”, *Moneta e Credito*, Vol. 221
- Altunbaş, Y., Gardener, E. P. M., Molyneux, P., Moore, B., (2001) “Efficiency in European Banking”, *European Economic Review*, Vol. 45, pp. 1931-1955
- Altunbaş, Y., Goddard, J., Molyneux, P., (1999) “Technical change in Banking”, *Economics Letters*, n° 64, pp. 215-221
- Angelini, P., Cetorelli, N., (2004) *Gli effetti delle modifiche normative sulla concorrenza nel mercato creditizio*, in Panetta F. (a cura di) *Il sistema bancario italiano negli anni novanta*, Il Mulino, Bologna
- Baglioni, A., Petrella, G., (2003) “L’evoluzione del sistema bancario italiano nel decennio 1993-2002”, in ASSBB, Osservatorio Monetario, n°3
- Barra, C., Destefanis, C., Lubrano Lavadera, G., (2011) “Risk Regulation: the Efficiency of Italian Cooperative Banks”, *Centre for Studies in Economics and Finance, University of Naples*, WP, n. 290
- Battaglia, F., Farina, V., Fiordelisi, F., Ricci, O., (2010) “The Efficiency of Cooperative Banks: the Impact of Environmental Economic Conditions”, *Applied Financial Economics*, Vol. 20, n° 17, pp. 1363-1376
- Battese, G. E., Coelli, T. J., (1992) “Frontier Production Functions, Technical Efficiency and Panel Data: with Application to Paddy Farmers in India”, *Journal of Productivity Analysis*, Vol. 3, pp. 153-169
- Battese, G. E., Coelli, T. J., (1995) “A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data”, *Empirical Economics*, Vol. 20, n° 2, pp. 325-332
- Battese, G. E., Coelli, T. J., Rao, D. S. P., O’Donnell, C. J., (2005) *An Introduction to Efficiency and Productivity Analysis*, Springer, New York
- Beccalli, E., Casu, B., Girandone, C., (2006) “Efficiency and Stock Performance in European Banking”, *Journal of Business Finance & Accounting*, Vol. 33, n° 1-2, pp. 245-262
- Behr, A., Tente, S., (2008) “Stochastic Frontier Analysis by Means of Maximum Likelihood and Method of Moments” *Deutsche Bundesbank – Banking and Financial Studies*, Discussion Paper n°19

- Berger, A. N., De Young, R., (1997) “Problem Loans and Cost Efficiency in Commercial Banking”, *Journal Banking and Finance*, Vol. 21, pp. 849-870
- Berger, A. N., Humphrey, D. B., (1997) “Efficiency of Financial Institutions: International Survey and Directions for Future Research”, *European Journal of Operation Research*, Vol. 98, pp. 175–212
- Bos, J. W. B., Heid, F., Koetter, M., Kolari, J. W., Kool, C. J. M., (2005) “Inefficient or just Different? Effects of Heterogeneity on Bank Efficiency Scores”, *Deutsche Bundesbank*, Discussion Paper, N. 15
- Casu, B., Girardone, C., (2006) “Bank Competition, Concentration and Efficiency in the Single European Market”, *The Manchester School*, Vol. 74, n° 4, pp. 441-468
- Casu, B., Girardone, C., (2009) “Testing the Relationship between Competition and Efficiency in Banking: a Panel Data Analysis”, *Economics Letters*, n° 105, pp. 134-137
- Casu, B., Girardone, C., (2010) “Integration and Efficiency Convergence in EU Banking Markets”, *Omega*, n° 38, pp. 260-267
- Casu, B., Molyneux, P., (2003) “A Comparative Study of Efficiency in European Banking”, *Applied Economics*, Vol. 35, n° 17, pp. 1865-1876
- Chambers, R. G., (1988) *Applied Production Analysis*, Cambridge University Press
- Chortareas, G. E., Garza-Garcia, J. G., Girardone, C., (2009) “Banking Sector Performance in Latin America: Market Power versus Efficiency”, *University of the West of England, Department of Economics*, Discussion Papers, n° 0905
- Ciocca, P., (2004) “Basilea 2 e IAS: più concorrenza, minori rischi”, VIII Convention ABI, Roma
- D’Agostino, M., (2003) *Gli effetti di Basilea 2 sulle imprese*, Bonanno Editore, Acireale-Roma
- D’Orio, G., (2003) *Un modello per valutare l’efficienza delle imprese nella fase di transizione*, in Infante D. (a cura di) *Il mercato automobilistico russo fra concorrenza e protezionismo*, Franco Angeli, Milano
- De Siano, R., (2001) “La valutazione dell’efficienza nella banca come impresa multi-prodotto”, *Università degli Studi di Napoli Parthenope, Istituto di Studi Economici*, WP n° 8
- Destefanis, S., (2001) “Differenziali territoriali di produttività ed efficienza negli anni ’90: i livelli e l’andamento”, *Università degli Studi di Salerno, Dipartimento di Scienze Economiche, Centro di Economia del Lavoro e di Politica Economica*, Discussion Paper, n° 59

- Domma, F., (2002) “L’andamento della *hazard function* nel modello di Dagum a tre parametri”, *Quaderni di Statistica*, Vol. 4
- Domma, F., (2004) “Kurtosis Diagram for the Log-Dagum Distribution”, *Statistica & Applicazioni*, Vol.2, n.2
- Domma, F., Condino, F., (2013) “The Beta-Dagum Distribution: Definition and Properties”, *Communication in Statistics – Theory and Methods*, di prossima pubblicazione (Marzo 2013)
- Domma, F., Perri, P. F., (2008) “Some Developments on the Log-Dagum Distribution”, *Statistical Methods and Applications*, Vol. 18, n° 2, pp. 205-220
- Dongili, P., Rossi, S. P. S., Zago, A., (2008) *Efficienza e competitività delle banche italiane: un confronto con alcuni sistemi bancari europei*, in Cella G., Zago, A. (a cura di) *Competitività ed efficienza dell’economia italiana: fattori sistemici e valutazioni quantitative*, Il Mulino, Bologna
- Dongili, P., Zago, A., (2005) “Bad Loans and Efficiencies in Italian Bank”, *Dipartimento di Scienze Economiche, Università di Verona*, WP n° 28
- Fabbri, D., (1996) “La stima di frontiere di costo nel trasporto pubblico locale: una rassegna e un’applicazione”, *Dipartimento di Scienze Economiche, Università di Bologna*, WP
- Farrel, M. J., (1957) “The Measurement of Productive Efficiency”, *Journal of Royal Statistical Society*, Vol. 120, n°3, pp. 253-290
- Ferri, G., Messori, M., (2000) “Bank-Firm Relationships and Allocative Efficiency in Northeastern and Central Italy and in the South”, *Journal of Banking and Finance*, n° 24, pp. 1067-1095
- Ferrier, G., Lovell, C. A. K., (1990) “Measuring Cost Efficiency in Banking: Econometric and Linear Programmino Evidence”, *Journal of Econometrics*, Vol. 46, pp. 229-245
- Fiordelisi, F., Marques-Ibanez, D., Molyneux, P., (2010) “Efficiency and Risk in European Banking”, *European Central Bank*, WP n° 1211
- Fontani, A., Vitali, L., (2007) “L’efficienza di costo dei gruppi bancari italiani: un’analisi mediante frontiera stocastica”, *Department of Economics and Business, Luiss, Roma*
- Giannola, A., (2009) *Bank Mergers and Credit Allocation Among Italian Regions*, in Silipo D. B. (a cura di) *The Banks and the Italian Economy*, Springer Phisyca-Verlag, Berlin Heidelberg
- Giannola, A., Lopes, A., Ricci, C., Scarfiglieri, G., (1997) *Divari territoriali ed efficienza del sistema bancario italiano*, in Quintieri B. (a cura di) *Finanza, istituzioni e sviluppo regionale*, Il Mulino, Bologna
- Giordano, L., Lopes, A., (2006) *Preferenza al rischio e qualità degli impieghi come determinanti dell’efficienza del sistema bancario italiano*, in Giannola A. (a cura di)

*Riforme istituzionali e mutamento strutturale. Mercati, imprese e istituzioni in un sistema dualistico*, Carocci Editore, Roma

- Giordano, L., Lopes, A., (2008) “L’efficienza del sistema bancario italiano in un contesto dualistico e la rilevanza della dimensione e degli aspetti giuridici”, *Il Risparmio*, n° 2
- Girardone, C., Molyneux, P., Gardener, E. P. M., (2004) “Analysing the Determinants of Bank Efficiency: the Case of Italian Banks”, *Applied Economics*, Vol. 36, n°3, pp. 215-227
- Greene, W. H., (1993) *The Econometric Approach to Efficiency Analysis*, in *The Measurement of Productivity Efficiency: Techniques and Applications*, Oxford University Press, New York, pp. 92-250
- Huang, C. J., Liu, J. T., (1994) “Estimation of a non-neutral Stochastic Frontier Production Function”, *Journal of Productivity Analysis*, Vol. 5, pp. 171-180
- Huizinga, H. P., Nelissen, J. H. M., Vennet, R. V., (2001) “Efficiency Effect of Bank Mergers and Acquisitions in Europe”, *Tinbergen Institute, Discussion Paper*, Vol. 88, n° 3
- Humphrey, D. B., Pulley, L. B., (1997) “Banks’ Responses to Deregulation: Profits, Technology and Efficiency”, *Journal of Money, Credit and Banking*, Vol. 29, n° 1, pp. 73-93
- Hunter, W., C., Timme, S. G., (1995) “Core Deposits and Physical Capital: a Reexamination of Bank Scale Economies with Quasi-Fixed Inputs”, *Journal of Money, Credit and Banking*, Vol. 27, n°1, pp. 165-185
- Inanoglu, H., Jacobs, Jr. M., Sickles, R., (2010) “Analyzing Bank Efficiency: are ‘Too-Big-to-Fail’ Bank Efficient?”, forthcoming in *the Journal of Efficiency*
- Joe, H., (1997) *Multivariate Models and Dependence Concepts*, Chapman and Hall, London
- Jondrow, J., Lovell, C. A. K., Materov, I. S., Schimdt, P., (1982) “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model”, *Journal of Econometrics*, Vol. 19, pp. 233-238
- Kumbhakar, S. C., Ghosh, S., McGuckin, J. T., (1991) “A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in US Dairy Farms”, *Journal of Business and Economic Statistics*, Vol. 9, pp. 279-286
- Kumbhakar, S. C., Lovell, C. A. K., (2000) *Stochastic Frontier Analysis*, Cambridge University Press
- Lang, G., Welzel, P., (1996) “Efficiency and Technical Progress in Banking Empirical Results for a Panel of German Cooperative Banks”, *Journal of Banking and Finance*, n° 20, pp. 1003-1023
- Leibenstein, H., (1966) “Allocative Efficiency versus “X-Efficiency””, *American Economic Review*, Vol. 56, pp. 392-415

- Lensink, R., Meesters, A., (2012) “Institutions and Bank Performance: a Stochastic Frontier Analysis”, *Oxford Bulletin of Economics and Statistics*
- Lucchetti, R., Papi, L., Zazzaro, A., (1999) “Efficienza del sistema bancario e crescita economica nelle regioni italiane”, *Dipartimento di Economia, Università degli Studi di Ancona, Quaderni di Ricerca n°121*
- Mastromarco, C., (2007) “Stochastic Frontier Models”, *Dipartimento di Economia e Matematica-Statistica, Università di Lecce, CIDE 2007*
- Mattesini, F., Messori, M., (2004) *L’evoluzione del sistema bancario meridionale: problemi aperti e possibili soluzioni*, Il Mulino, Bologna
- Messori M., Tamburini R., Zazzaro A., (2003) *Il sistema bancario italiano*, Carocci Editore, Roma
- Mester, L. J., (1996) “A Study of Bank Efficiency Taking into Account Risk-Preferences”, *Journal of Banking and Finance*, n°20, pp. 1025-1045
- Molyneux, P., (2003) “Technical Change, Costs and Profits in European Banking”, *United Nations University, Institute for New Technologies, Technology and Finance WP*, n° 21
- Murillo-Zamorano, L. R., (2004) “Economic Efficiency and Frontier Techniques”, *Journal of Economic Survey*, Vol. 18, n° 1, pp. 33-45
- Nelsen, R. B. (1999) *An Introduction to Copula*, Lecture Notes in Statistics, Springer, New York
- Panetta, F., (2003) “Evoluzione del sistema bancario e finanziamento dell’economia nel Mezzogiorno”, *Tema di Discussione della Banca d’Italia*, WP n° 467
- Panetta, F., (2004) *La trasformazione del sistema bancario e i suoi effetti sull’economia italiana*, in Panetta F. (a cura di) *Il sistema bancario italiano negli anni novanta*, Il Mulino, Bologna
- Panzar, J. C., Rosse, J. N., (1987) “Testing for Monopoly Equilibrium”, *Journal of Industrial Economics*, Vol. 35, pp. 443-456
- Rapporto Svimez 2010 sull’economia del Mezzogiorno
- Reifschneider, D., Stevenson. R., (1991) “Systematic Departures from the Frontier. A Framework for the Analysis of Firm Inefficiency”, *International Economic Review*, Vol. 32, pp.715-723
- Relazione Annuale (1998-2010), *Banca d’Italia*
- Sealey, C. W., Lindley, J. T., (1977) “Input, output and a theory of production and cost at depository financial institutions”, *The Journal of Finance*, n° 4

- Smith, M. D. (2008) “Stochastic frontier models with dependent error components”, *Econometrics Journal*, Vol. 11, pp. 172-192
- Trivedi, P. K., Zimmer, D. M., (2005) “Copula Modeling: An Introduction for Practitioners”, *Foundations and Trends in Econometrics*
- Tulkens, H. (1993), “On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit”, *The Journal of Productivity Analysis*, Vol. 4
- Turati, G., (2008) *La valutazione del grado di concorrenza nell’industria bancaria negli anni Novanta*, in Cella G., Zago, A. (a cura di) *Competitività ed efficienza dell’economia italiana: fattori sistemici e valutazioni quantitative*, Il Mulino, Bologna
- Wang, H.-J., Schimdt, P., (2002) “One-Step and Two-Step Estimation of the Effect of Exogenous Variables on Technical Efficiency Levels”, *Journal of Productivity Analysis*, Vol. 18, pp. 129-144
- Weill, L., (2004) “Measuring Cost Efficiency in European Banking: a Comparison of Frontier Techniques”, *Journal of Productivity Analysis*, Vol. 21, n° 2, pp. 133-152
- Williams, J., (2004) “Determining Management Behaviour in European Banking”, *Journal of Banking & Finance*, Vol. 28, pp. 2427-2460