

LA CAPACITÀ PREDITTIVA DEGLI INDICATORI DI BILANCIO: UN METODO PER LE PMI

di *Francesca Pierri, Alberto Burchi, Elena Stanghellini*¹

1. Introduzione

L'attività di risk management da sempre si confronta con il rischio di default. Esso si materializza quando il debitore non è in grado di rispettare le scadenze previste nei termini contrattuali, riguardanti sia il rimborso dei capitali che il pagamento di interessi. A partire dalla seconda metà degli anni '60, grazie ai lavori pionieristici di Beaver (1966) e Altman (1968), il metodo di misurazione del rischio di credito è variato, cominciando a comparire strumenti automatizzati basati su tecniche statistiche (Sironi e Resti, 2008, cap. 11). Da allora, sofisticate metodologie quantitative, originate in contesti diversi, sono state raffinate e adattate alla misurazione del rischio di credito, una rassegna è in Stanghellini (2009, cap. 1). Più in generale, come testimonia la recente monografia di De Laurentis e Maino (2009), strumenti automatizzati sono ormai correntemente utilizzati nel processo di assunzione, gestione e controllo del rischio di credito.

La centralità delle piccole e medie imprese nel tessuto industriale di alcuni paesi sviluppati, fra cui l'Italia (Lusignani e Sironi, 1999), rende necessaria una indagine sulla capacità degli indicatori di bilancio di misurare lo stato di salute delle aziende e di prevederne, con anticipo adeguato, l'evento fallimento. La ricerca accademica è unanime nell'evidenziare l'esigenza di costruire, a partire dalle poste di bilancio, un insieme di indicatori della performance di breve o di lungo periodo (Montrone, 2005) ed ha pertanto prodotto nel corso degli anni una moltitudine di indici con obiettivi cono-

¹ Gli autori intendono ringraziare il Prof. Alessandro Montrone per la costruttiva collaborazione e gli interessanti commenti utili alla interpretazione delle analisi svolte e la Prof.ssa Paola Musile Tanzi per aver letto il manoscritto e per aver fornito preziosi suggerimenti. Gli autori desiderano esprimere la propria gratitudine ai partecipanti al convegno annuale dell'Accademia Italiana di Economia Aziendale (Aidea - Perugia, Ottobre 2011) per i commenti offerti. Si ringraziano i due referee anonimi per gli approfonditi commenti e le indicazioni che hanno permesso di migliorare questo lavoro.

scitivi diversi e particolareggiati. Tuttavia, la pluralità di indicatori a disposizione di chi deve analizzare il bilancio e, in particolare, stimare il rischio di default, non corrisponde ad un altrettanto vasto insieme di fenomeni osservabili. In altri termini, esiste un effetto di sostituzione tra le variabili economiche, indotto dalla loro correlazione, di cui occorre tenere conto (Moody's, 2000).

Diversi studi hanno esplorato la sopravvivenza e il fallimento delle piccole medie imprese in varie nazioni, si veda, negli anni più recenti, Huyhebaert e altri (2000); Watson (2003); Pompe e Bilderbeek (2005). Una estesa rassegna della letteratura è rinvenibile in Altman e Hotchkiss (2006); si veda anche Altman e Sabato (2007).

Obiettivo di questo lavoro è proporre un processo di sviluppo di un modello statistico rigoroso per la valutazione del rischio di credito. Il lavoro è volto ad individuare la capacità predittiva degli indici di bilancio e indicare quale sottoinsieme sia in grado di misurare più accuratamente lo stato di salute di una azienda.

Nella letteratura accademica riguardante i processi di valutazione del merito creditizio, la ricerca di accuratezza attraverso la competenza geografica o la specializzazione settoriale dei modelli è ancora dibattuta. Alcuni studiosi si concentrano verso la specializzazione settoriale (Altman e altri, 1994; Friedman e altri, 2003; Standard e Poor's, 2002) mentre altri sviluppano modelli la cui validità è garantita nell'area geografica di provenienza dei dati. Dal punto di vista statistico, un'analisi su un campione omogeneo, sia con riferimento al settore che al territorio, potrebbe non essere affidabile a causa della bassa incidenza delle aziende fallite rispetto alla numerosità del campione stesso. Una via d'uscita, proposta da alcuni autori come Roggi e Giannozzi (2008), è la creazione di macro settori economici, una seconda possibilità è inserire nel modello variabili dummy (Luppi e altri, 2007). Una ulteriore strategia è quella proposta in questo lavoro dove si è voluto tenere conto sia della specializzazione geografica, riferendosi a società di capitali di una sola regione italiana, l'Umbria, che della specializzazione settoriale, utilizzata nel corso dell'analisi come criterio fondamentale per la composizione di strati omogenei di unità all'interno del campione di studio

La scelta geografica è stata motivata dal fatto che la popolazione aziendale umbra offre un piano di indagine peculiare, in considerazione della elevata presenza di piccole e medie imprese e di un'economia rivolta alla domanda interna, la cui capacità di crescita non dipende dalla competizione nel mercato aperto, ma dai flussi di spesa privata e pubblica, caratteristica che rende le aziende più facilmente esposte alla crisi (Bracalente, 2010). In un panorama così delineato, il sostegno da parte del mondo del credito diventa di importanza cruciale ed è pertanto rilevante chiarire le criticità e definire i criteri di intervento.

L'analisi dei dati a disposizione ha evidenziato la presenza di un campione sbilanciato a causa di una frequenza percentuale di aziende in fallimento bassa da un punto vista statistico rispetto alla numerosità totale. Ciò ha portato all'implementazione di una estensione del modello logistico tradizionale: il modello logistico condizionato per studi caso-controllo già utilizzato nelle scienze epidemiologiche ed implementato in altre discipline per lo studio di fenomeni che si verificano con bassa frequenza nella popolazione. Il metodo implica in primo luogo la costruzione di un sottocampione che rispetto a quello originario risulta essere bilanciato ovvero composto da circa lo stesso numero di aziende sane e di aziende in default. Questo primo passo permette di ottenere stime più stabili, dal momento che vengono colte meglio le caratteristiche delle unità rare. Inoltre, la selezione delle unità che entrano a fare parte del campione viene effettuata sulla base dell'appartenenza a classi omogenee, o strati, e ciò tende ad accrescere la capacità predittiva del modello. Le difficoltà computazionali derivanti dall'utilizzo di tale modello, che ne limitavano l'impiego, sono ad oggi superate, grazie alla implementazione di appositi algoritmi in due software statistici di largo uso: SAS e STATA.

L'approccio proposto evidenzia la specificità di competenze necessarie nell'applicazione del modello logistico per la valutazione del merito creditizio. I modelli, infatti, abbisognano di adeguate personalizzazioni e perfezionamenti in grado di cogliere le specificità del settore, del tessuto industriale e dei fenomeni indagati. In assenza di tale rigore, l'applicazione di metodi statistici per la misurazione del rischio di credito costituirebbe una leggerezza conducendo ad errori di valutazione (Giampaoli, 2006).

Il lavoro presenta diverse caratteristiche innovative: il metodo di analisi adottato rappresenta una interessante e opportuna evoluzione dei tradizionali modelli logistici; l'attenzione alle numerose fasi di ottimizzazione ha portato alla riduzione del paniere di variabili considerate nei modelli senza inficiare le performance previsionali; i risultati dei coefficienti stimati consentono la quantificazione degli effetti delle variazioni degli indici sulla misura della probabilità di default; la coerenza con precedenti studi delle discipline aziendali conferma la bontà degli approcci automatizzati nella valutazione del merito creditizio.

Il lavoro è così suddiviso: nel paragrafo due viene esposta la metodologia adottata, nel paragrafo tre viene riportata l'analisi del campione e le applicazioni del modello, nel paragrafo quattro sono descritti ulteriori approfondimenti, il quinto paragrafo comprende l'interpretazione economica delle variabili considerate più significative nel processo di valutazione dello stato di salute aziendale, mentre nel sesto ed ultimo paragrafo vengono espone le conclusioni del lavoro.

2. Scelte metodologiche

I modelli di previsione delle insolvenze basati su metodologie di carattere statistico sono oggi ampiamente diffusi, perché consentono una immediata valutazione dell'azienda, tramite l'assegnazione di uno *score*, un punteggio, che sintetizza il suo grado di solvibilità.

Fra questi modelli, noti come modelli di *credit scoring*, quello più diffuso è sicuramente il modello logistico. Per la costruzione di un modello di questo genere è necessario disporre di un campione di soggetti solvibili e insolventi nei quali si è osservato un insieme di variabili con un contenuto informativo tale da permettere la discriminazione fra i due gruppi.

Nel caso oggetto di studio si è indagato quali indici di bilancio potevano avere una maggiore capacità discriminante all'interno di un campione per il quale lo stato giuridico (attiva o in fallimento) è stato rilevato in un periodo temporale di quattro anni. L'analisi del campione, i cui dettagli sono riportati nel paragrafo successivo, ha mostrato che solo l'1% delle aziende risulta in fallimento. Un evento di tali proporzioni è considerato raro da un punto di vista statistico e non permette l'applicazione del modello logistico tradizionale, in quanto fornirebbe stime instabili ed una elevata probabilità di errata classificazione delle aziende in default; si veda King e Zeng (2001). Si possono invece implementare tecniche di bilanciamento del campione, consistenti nella estrazione casuale dal campione di studio di un pari numero di aziende sane e ed in default, per una dettagliata trattazione si rimanda a Stanghellini (2009, cap. 3). In questo lavoro si è proceduto ad applicare una tecnica che raffina quella precedente, ampiamente utilizzata negli studi epidemiologici per il suo grado di affidabilità, consistente nel suddividere il campione in gruppi omogenei di unità, detti strati, individuati attraverso variabili rilevanti² (in questo caso il codice ATECO e la forma giuridica). All'interno di ogni strato, poi, si è proceduto a formare un campione bilanciato, ovvero ad ogni azienda fallita è stato associato un numero specifico di aziende sane, mediante estrazione casuale dal gruppo di appartenenza. Questo metodo è detto di tipo caso-controllo.

Così facendo è stato raggiunto sia l'obiettivo di effettuare un'analisi statistica rigorosa che quello di tenere nella dovuta considerazione la specificità settoriale e giuridica accanto a quella territoriale.

² Negli studi epidemiologici le variabili più frequentemente usate a questo scopo sono l'età ed il sesso, che hanno una potenziale influenza sulla risposta, ma sulle quali non si può pensare nessun tipo di intervento e sono pertanto considerate variabili di disturbo.

2.1 Il modello logistico condizionato per studi caso-controllo

Si riporta di seguito una sintetica formalizzazione statistica del modello in oggetto utile per comprendere meglio cosa si intende per uno studio caso-controllo, i risultati che si ottengono e la relativa interpretazione.

A ogni unità statistica si associa una variabile casuale binaria Y , che misura l'evento fallimento, o default. All'interno di ciascuno strato si individuano i soggetti/aziende che rappresenteranno i casi ($Y = 1$), ovvero le aziende fallite, ed i controlli ($Y = 0$), cioè le aziende attive. In genere si sceglie un numero fisso di controlli che va da 1 a 5, anche se il numero dei controlli può variare all'interno di ciascuno strato.

Si indichi con $\mathbf{x}' = (x_1, x_2, \dots, x_p)$ il vettore di p variabili esplicative, si definisca con K il numero di strati, con n_{1k} i casi e con n_{0k} i controlli nello strato k , per $k=1,2,\dots,K$. Il modello logistico specifico per ciascuno strato sarà dato da:

$$P(Y = 1 \mid \mathbf{x}) = \pi_k(\mathbf{x}) = \frac{e^{\alpha_k + \beta' \mathbf{x}}}{1 + e^{\alpha_k + \beta' \mathbf{x}}} \quad (2.1.1)$$

dove α_k rappresenta il contributo di tutti i termini costanti (ovvero le variabili utilizzate per la stratificazione) all'interno del k -esimo strato e β il vettore dei p coefficienti, $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$. Si verifica agevolmente che il modello (2.1.1) impone un effetto lineare delle variabili sul logaritmo dell'*odds*³ che un'impresa sia in fallimento. L'interesse conoscitivo risiede nei coefficienti β_j che esprimono l'effetto della covariata x_j sulla probabilità di default. I parametri α_k specifici dello strato sono termini di disturbo e su di essi non vi è interesse conoscitivo. Si può pertanto pensare di eliminarli dall'insieme dei parametri sui quali si vuole fare inferenza per ottenere una migliore stima dei β_j . Questo si ottiene attraverso metodi di inferenza condizionata, che permettono di giungere a stime dei coefficienti di inclinazione β_j che sono consistenti e distribuiti in modo asintoticamente normale. I dettagli matematici si trovano in Cox e Hinkley (1974, cap. 9), mentre quelli applicativi sono in Hosmer e Lemeshow (2000, cap. 7).

La verosimiglianza condizionata per il k -esimo strato è ottenuta come la probabilità del verificarsi della configurazione di casi e controlli dello strato, condizionata al numero n_{1k} e n_{0k} di casi e controlli. Ad esempio, per un modello 1-3 (in cui il rapporto fra i casi e i controlli è 1 su 3) si denotino i valori delle covariate nello strato k con \mathbf{x}_{k1} per il caso e con $\mathbf{x}_{k2}, \mathbf{x}_{k3}, \mathbf{x}_{k4}$

³ Si definisce odds di un evento A il rapporto fra la probabilità che A si verifichi ($P(A)$) e la probabilità che A non si verifichi ($1-P(A)$)

per i controlli. Il contributo alla verosimiglianza dello strato k è dato dalla seguente espressione:

$$l_k(\beta) = \frac{e^{\beta' \mathbf{x}_{k1}}}{e^{\beta' \mathbf{x}_{k1}} + e^{\beta' \mathbf{x}_{k2}} + e^{\beta' \mathbf{x}_{k3}} + e^{\beta' \mathbf{x}_{k4}}} \quad (2.1.2)$$

Dati i valori dei coefficienti, la (2.1.2) fornisce la probabilità che condizionatamente a tutte le possibili assegnazioni di 1 caso e 3 controlli all'interno dello strato, il soggetto con i dati osservati \mathbf{x}_{kl} sia il caso. Maggiori dettagli in Hosmer e Lemeshow (2000, pag. 225). Si noti che se le covariate sono identiche per tutti e quattro i soggetti, allora lo strato risulta essere non informativo per la stima dei coefficienti, visto che $l_k(\beta)=0.25$ per ogni valore di β . Per una singola covariata ci deve essere almeno un controllo che abbia un valore diverso dal caso, altrimenti lo strato non risulterà informativo per quel coefficiente. Per questo motivo, alcuni autori suggeriscono di far precedere l'analisi multivariata da modelli univariati e di osservare la stima degli errori standard e degli intervalli di confidenza (Hosmer e Lemeshow, 2000, pag. 231). Inoltre, alcuni software, fra cui SAS 9.2 utilizzato per l'analisi dei dati, mettono a disposizione comandi specifici per massimizzare la verosimiglianza condizionata e forniscono indicazione del numero degli strati non informativi non utilizzati nel processo di stima.

3. Analisi dati

I dati per questa analisi sono stati estratti dalla Banca Dati AIDA. Sono state prese in considerazione le aziende umbre con ricavi delle vendite maggiori o uguali a 0 nell'anno 2006. Per queste stesse aziende a maggio 2010 si è estratto lo stato giuridico: se ne riporta la distribuzione in Tabella 1. Il campione iniziale di riferimento, all'interno di questo studio, è quindi rappresentato da 59 aziende in fallimento e da 6103 attive.

Tab. 1 - Distribuzione delle aziende in base allo stato giuridico

Stato giuridico	Frequenza	Percentuale
Ditta attiva	6103	93.63
Ditta in fallimento	59	0.91
Ditta in liquidazione	324	4.97
Ditta inattiva	32	0.49

E' opportuno chiarire che la Banca Dati AIDA non dà la possibilità di conoscere lo stato giuridico di un'azienda relativamente ad ogni anno. Quando lo stato giuridico di un'azienda cambia, questa voce viene aggiornata.

nata nel database non lasciando traccia della data di modifica. Per esemplificare, per le aziende attive nel 2006 che a maggio 2010 risultano fallite, non si conosce in quale anno queste siano fallite (si conosce solo il fatto che l'evento è accaduto fra il 2006 e il 2010). Questa censura non permette di utilizzare metodi di sopravvivenza, che risulterebbero altrettanto informativi per lo studio in corso. La ricerca accademica (Roggi e Giannozzi, 2008), inoltre, ha evidenziato la presenza di un ritardo tra la disponibilità del bilancio di esercizio e il verificarsi dell'evento fallimento. Ciò è dovuto alla prassi aziendale, in ottemperanza alle norme civilistiche e fiscali, di rendere disponibili i bilanci di esercizio circa sei mesi dopo la chiusura dell'esercizio stesso. Per questa ragione molti modelli considerano il bilancio di esercizio di due anni precedenti al fine di valutare la probabilità di default dell'azienda: ad esempio, si analizza il bilancio 2005 per valutare la probabilità di default dell'azienda nel 2007. La scelta da noi compiuta di analizzare i bilanci del 2006 considerando le aziende fallite nel più ampio arco temporale successivo fino al maggio 2010 è dunque coerente. La lista delle variabili considerate, in via preliminare, è in Tabella 2.

Come già anticipato nel secondo paragrafo, la metodologia statistica utilizzata è rappresentata dal modello logistico condizionato per studi caso-controllo, dove il caso è rappresentato dall'azienda fallita ed il controllo da un numero di aziende sane all'interno dello stesso strato. L'analisi dei dati inizia quindi con la selezione delle variabili per la formazione degli strati (3.1) e, ai fini della definizione di un modello, è seguita dall'individuazione degli indici maggiormente associati con la variabile risposta e dall'accertamento della presenza di correlazione fra questi (3.2); successivamente si perviene alla definizione di un modello logistico condizionato per studi caso-controllo (3.3) e alla verifica delle assunzioni di base di tale modello (3.4).

3.1 Definizione del campione: scelta delle variabili per la formazione degli strati

L'impiego del modello logistico condizionato per studi caso-controllo richiede come presupposto una base dati opportunamente costruita secondo un principio di stratificazione. Per formare gli strati inizialmente sono state prese in considerazione, con riferimento all'anno 2006, il totale delle immobilizzazioni, il costo del personale e i salari e stipendi, le prime due cifre del codice ATECO e la forma giuridica. La variabile costo del personale è stata ritenuta non utilizzabile in quanto la frequenza dell'immissione del dato era molto bassa. Per le variabili che misurano il totale delle immobilizzazioni e i salari e gli stipendi sono state costruite classi di ampiezza opportuna verificando successivamente gli odds delle fallite sulle attive all'interno di ciascuna classe con l'obiettivo di massimizzare la differenza degli odds fra una classe e l'altra.

Tab. 2 - Elenco delle variabili utilizzate per le analisi

N	Nome Variabile	Codice
1	Ragione Sociale	RagSoc
2	Numero CCIAA	NCCIA
3	ATECO 2007 Codice	Ateco07
4	Stato Giuridico	SG
5	Forma Giuridica	FG
6	Provincia	PROV
7	Ricavi delle vendite	RV
8	Utile Netto	UN
Indici finanziari		
1	Costo denaro a prestito (%)	CDP
2	Debiti verso le banche su fatturato (%)	DVB
3	Debt/EBITDA ratio	DEBITDA
4	Debt/Equity ratio	DER
5	Grado di ammortamento	GA
6	Grado di copertura degli interessi passivi	CIP
7	Grado di indipendenza da terzi	GIT
8	Indice di copertura delle immobilizzazioni (finanziario)	ICIF
9	Indice di copertura delle immobilizzazioni (patrimoniale)	ICIP
10	Indice di disponibilità	ID
11	Indice di indebitamento a breve	IIB
12	Indice di indebitamento a lungo	IIL
13	Indice di indipendenza finanziaria (%)	IIF
14	Indice di liquidità	IL09
15	Oneri finanziari su fatturato (%)	OF
16	Posizione finanziaria netta	PFN
17	Rapporto di indebitamento	RI
Indici della gestione corrente		
1	Durata Ciclo Commerciale (gg)	DCC
2	Durata media dei crediti al lordo IVA (gg)	DMC
3	Durata media dei debiti al lordo IVA (gg)	DMD
4	Rotazione del capitale circolante lordo (volte)	RCCL
5	Rotazione del capitale investito (volte)	RCI
6	Incidenza circolante operativo (%)	ICO
Indici di redditività		
1	EBITDA/Vendite (%)	EBITDAV
2	Redditività del capitale proprio (ROE) (%)	ROE

continua ...

...continua

3	Redditività delle vendite (ROS) (\%)	ROS
4	Redditività di tutto il capitale investito (ROI) (%)	ROI
5	Incidenza oneri / Proventi extragestione (%)	IOPE
Immobilizzazioni e costo del personale		
1	Totale immobilizzazioni materiali	TIM
2	Terreni e fabbricati	TF
3	Totale immobilizzazioni immateriali	TII
4	Totale immobilizzazioni finanziarie	TIF
5	Totale costi del personale	TCP
6	Salari e stipendi	SS

Utilizzando queste quattro variabili si riusciva ad ottenere un matching solo per 16 aziende fallite, pertanto, dopo successivi tentativi, si è deciso di utilizzare due sole variabili: le prime due cifre del codice ATECO e la forma giuridica. La distribuzione della seconda è riportata in Tabella 3.

Tab. 3 - Distribuzione delle aziende in base alla forma giuridica

Forma Giuridica	Frequenza	Percentuale
S.C.A.R.L.	4	1.79
S.C.A.R.L.P.A	4	0.91
S.P.A.	16	7.14
S.R.L.	191	85.27
S.R.L. a socio unico	9	4.02

Data la difficoltà di matching per le S.R.L a socio unico, si decide di accorparle alle S.R.L. Analizzando la frequenza delle osservazioni per ogni strato, si ritiene opportuno effettuare uno studio caso-controllo 1-3 per avere degli strati omogenei dal punto di vista della numerosità. Per 56 (su 59) aziende fallite si selezionano, sulla base delle variabili indicate, 168 aziende attive: all'interno di ciascuno strato per ogni azienda fallita si procede all'estrazione casuale di tre aziende sane.

3.2 Individuazione degli indici maggiormente associati con la variabile risposta e verifica della presenza di correlazione

Seguendo la metodologia proposta da Hosmer e Lemeshow (2000, cap. 7) si effettua un modello logistico univariato per tutte le variabili riportate in Tabella 2 (ad eccezione delle prime otto) in quanto si tratta dei dati identificativi di ogni società, al fine di individuare gli indici più rilevanti da includere nel modello

Questa prima analisi porta alla selezione di un gruppo iniziale di variabili le cui stime di massima verosimiglianza risultano significative in base al test χ^2 di Wald⁴ ($Pr > \chi^2$ inferiore a 0.05): DMC, DVB, EBITDAV, GIT, ID, IIF, IIB⁵ IL, RCCL, RCI, ROS. In Tabella 4 sono riportate le stime di massima verosimiglianza per le variabili sopra indicate, il numero di strati non informativi⁶ ed il numero di osservazioni eliminate automaticamente⁷ nella stima di ciascun modello. Da una prima analisi delle stime univariate dei parametri si può subito notare come le variabili IIB, DVB e DMC, avendo un coefficiente stimato positivo, sono legate alla probabilità di fallimento in maniera crescente, viceversa le altre con segno negativo. Il valore assunto dalla stima del parametro ci dà una prima misura dell'intensità del fenomeno del quale si andrà a cercare conferma con le analisi successive.

Tab. 4 - Variabili risultate significative nella logistica condizionata univariata

Parametro	Stima	Err. std.	χ^2	$Pr > \chi^2$	Strati non inf.	Freq. str. non inf	Oss. omesse.
DMC	0.006	0.002	11.13	0.0008	14	34	41
DVB	0.028	0.009	9.25	0.0024	24	57	57
EBITDAV	-0.038	0.010	14.67	0.0010	4	12	9
GIT	-5.669	1.223	21.48	<0.001	4	12	12
ID	-1.876	0.446	17.71	<0.001	1	3	3
IIB	1.615	0.822	3.86	0.0495	0	0	0
IIF	-0.059	0.014	17.75	<0.001	12	36	14
IIL	-1.615	0.822	3.86	0.0495	0	0	0
IL	-1.597	0.440	13.19	0.0003	1	3	3
RCCL	-0.412	0.143	8.24	0.0041	1	3	2
RCI	-0.731	0.242	9.10	0.0026	2	6	4
ROS	-0.041	0.014	8.71	0.0032	17	48	30

Le analisi fino ad ora eseguite non tengono conto dell'effetto congiunto delle variabili, pertanto, per evitare di inserire nel modello multiplo indici tra loro molto correlati e per ovviare i conseguenti problemi dovuti alla collinearità, si effettua uno studio della correlazione fra le variabili selezionate come potenziali esplicative. L'analisi risulta utile in quanto si individua una correlazione superiore a 0.5 per i seguenti gruppi di indici:

⁴ Test statistico utilizzato per valutare la significatività di ogni variabile esplicativa all'interno di un modello.

⁵ Fra IIB e IIL esiste una correlazione pari a -1, dovuta alla composizione dei due indici. Infatti, l'indice di indebitamento a breve termine (IIB) è pari al totale dei debiti a breve termine diviso il totale dei debiti; l'indice di indebitamento a lungo termine (IIL) è pari al totale dei debiti a lungo termine diviso il totale dei debiti.

⁶ Per ogni variabile si intendono non informativi gli strati dove in corrispondenza del caso non si hanno i corrispondenti valori dei controlli, o, viceversa, è omesso il valore del caso.

- EBITDAV con ROS;
- GIT con ID, IIF e IL;
- RCCL con RCI;
- ID con IIF.

3.3 Definizione di un modello logistico condizionato per studi caso-controllo

Sulla base dei risultati fino ad ora ottenuti si scelgono pertanto come variabili per l'analisi multivariata:

- ID al posto di GIT e IIF, per il maggior numero di strati informativi, e al posto di IL per un minor *p-value* nel test χ^2 di Wald;
- EBITDAV al posto di ROS per un minor *p-value* nel test χ^2 di Wald e per il numero di strati significativi;
- DMC perché non correlata con nessuno, pur considerando che ha tanti strati non informativi;
- DVB perché non correlata con nessuno, pur considerando che ha tanti strati non informativi;
- RCCL al posto di RCI per il maggior numero di strati informativi.

Tab. 5 - Stima dei coefficienti, errore standard, statistica di Wald e *p-value* per il modello contenente tutte le covariate

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
RCCL	0.53	0.69	0.60	0.44
DVB	0.07	4.00	3.53	0.06
DMC	0.02	0.01	2.74	0.10
EBITDAV	-0.02	0.04	0.33	0.56
ID	-6.31	2.78	5.15	0.02
IIB	17.53	10.57	2.75	0.10

Il modello logistico condizionato riportato in Tabella 5 conferma nei segni delle stime ciò che era stato anticipato dalla precedente analisi univariata. Analizzando i risultati del test di Wald si nota una non significatività nella stima di alcuni parametri (Pr > χ^2 superiore a 0.05), pertanto si effettua una scelta ragionata delle variabili da escludere dal modello. In questa fase, per ogni variabile, si tiene in considerazione sia il risultato del test di Wald sia il numero di dati omessi. Il modello che ne deriva (dettagli omessi), notevolmente migliorato rispetto al precedente, dà luogo a tutte variabili significative (Tabella 6). Il numero di osservazioni utilizzate e gli strati non informativi sono superiori a quelli del modello precedente (vengono utilizzate 221 osservazioni contro 147 in presenza di 6 strati non informativi contro 28).

⁷ Il numero di osservazioni eliminate automaticamente è rappresentato dai dati omessi, che possono dare luogo o non a strati non informativi, e dalle osservazioni non omesse appartenenti agli strati non informativi.

Tab. 6 - Stima dei coefficienti, errore standard, statistica di Wald e p-value per il modello che esclude DVB, DMC e IIB

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
RCCL	-0.69	0.24	8.03	0.0046
EBITDAV	-0.03	0.01	8.68	0.0032
ID	-3.29	0.87	14.29	0.0002

3.4 Verifica dell'assunzione di linearità

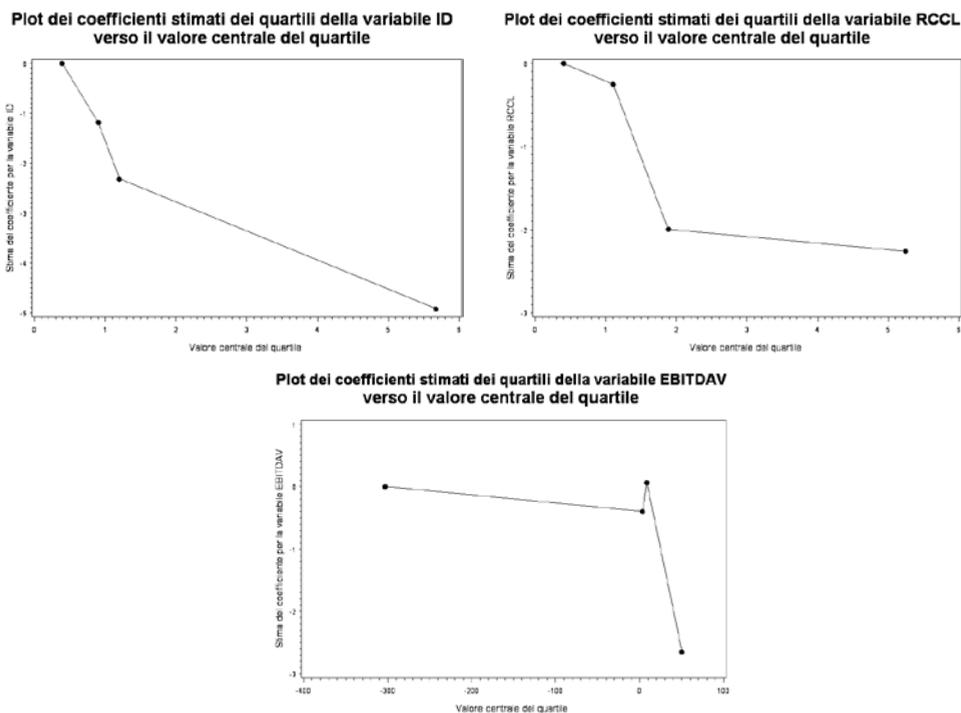
Prima di poter affermare con certezza la bontà del modello stimato è necessario verificare l'ipotesi che le variabili selezionate abbiano un effetto lineare sull'odds che una azienda sia fallita, come postulato dal modello logistico. Questo viene fatto trasformando ogni variabile continua in qualitativa raggruppando le unità in quattro classi sulla base dei quartili; la variabile così modificata viene inserita nel modello logistico, che utilizzerà la prima classe come livello di riferimento. La procedura vuole la sostituzione di una variabile alla volta con la sua trasformata qualitativa. Il grafico che pone sull'asse delle ascisse il punto centrale di ciascuna classe e sull'asse delle ordinate il valore dei coefficienti stimati permette di visualizzare l'andamento della variabile e di supportare o meno l'ipotesi di linearità. Le analisi condotte mostrano un andamento sicuramente lineare e decrescente per le variabili ID e RCCL⁸, mentre per la variabile EBITDAV si desume, sia dalla stima del modello (Tabella 7) che dal corrispondente grafico in Figura 1, che i coefficienti delle dummies della seconda e terza classe sono non significativi, mentre si ha un andamento significativamente decrescente nella quarta classe.

Tab. 7 - Stima dei coefficienti, errore standard, statistica di Wald e p-value per il modello con EBITDAV ripartita in classi secondo i quartili

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
EBITDAVQ2	-0.40	0.67	0.35	0.5554
EBITDAVQ3	0.06	0.64	0.01	0.9196
EBITDAVQ4	-2.65	0.93	8.11	0.0044
RCCL	-0.73	0.22	11.15	0.0008
ID	-3.64	0.88	17.03	<0.0001

⁸ Lo studio delle variabili ID e RCCL in quartili viene omessa ed è disponibile a richiesta.

Fig. 1 - Plot dei coefficienti stimati dei quartili delle variabili ID, RCCL e EBITDAV.



Tutto ciò suggerisce la possibilità di una forma quadratica per la variabile considerata e pertanto si procede alla stima del modello che include come covariate RCCL, ID e la forma quadratica sopra menzionata. In Tabella 8 si riportano i risultati del modello così determinato, (denominato Modello 1).

Tab. 8 - Stima dei coefficienti, errore standard, statistica di Wald e p-value per il modello con forma quadratica per EBITDAV (Modello 1)

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
EBITDAV	-0.0408	0.0130	9.91	0.0016
EBITDAV ²	-0.00006	<0.0001	5.69	0.0171
RCCL	-0.7344	0.2540	8.36	0.0038
ID	-3.5293	0.9413	14.06	0.0002

Si è andata successivamente a verificare la possibilità di una interazione con le variabili ATECO e Forma giuridica, scelte per formare gli strati. L'interazione, se presente, permette di cogliere un diverso effetto delle variabili esplicative negli strati. L'analisi ha evidenziato una non significatività delle interazioni con entrambe le variabili. Nell'analisi non si è potuto tenere in considerazione la variabile Provincia, pur giudicandola di interesse conoscitivo, a causa del basso numero di fallite nella provincia di Terni.

4. Ulteriori approfondimenti

L'obiettivo di una analisi statistica è quello di fornire un quadro il più possibile completo ed esaustivo del fenomeno in studio. Pertanto, è sempre opportuno affiancare ad una prima analisi un insieme di modelli alternativi che abbiano comparabili capacità interpretative. E' infatti dalla discussione critica delle diverse ipotesi e dalla loro compatibilità che si accresce la conoscenza. Per tale motivo, in questa sezione si illustrano ulteriori analisi effettuate per ampliare e talvolta sostituire il modello precedente.

Prendendo come punto di partenza le analisi univariate, si è proceduto ad ampliare l'insieme iniziale di indicatori, ammettendo nel modello multiplo tutte le variabili che avevano una probabilità del χ^2 inferiore a 0.25 nel test di Wald. Fra queste si sono selezionate le variabili con un maggior numero di strati informativi. Questa analisi ci porta ad inserire nel modello, in aggiunta alle variabili precedentemente selezionate, anche le variabili ICIF, ICIP e OF. Inizialmente si è valutato se l'aggiunta al Modello 1 (Tabella 8) di queste nuove variabili, una alla volta, portasse ad evidenziare eventuali effetti rilevanti. Questo studio non ha dato risultati significativi e pertanto, per questo insieme più ampio di covariate si effettua un'analisi logistica condizionata per studi caso-controllo, con selezione automatica del modello in modalità backward, forward e stepwise. La selezione stepwise dà come unica covariata significativa la variabile IIF; la backward seleziona un modello con quattro variabili (utilizza 205 osservazioni ed ha 8 strati non informativi con una frequenza di 28 osservazioni omesse); la forward restituisce un modello con tre variabili (utilizza 208 osservazioni, di queste ulteriori 28 non vengono utilizzate per la presenza di 12 strati non informativi). Si riportano le stime degli ultimi due modelli in Tabella 9 e 10.

Tab. 9 - Stima dei coefficienti, errore standard, statistica di Wald e p-value utilizzando il metodo di selezione backward

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
GIT	-5.15	1.09	6.79	0.0092
ID	-2.90	1.03	7.88	0.0050
IIB	8.37	3.48	5.77	0.0163
OF	0.38	0.13	8.95	0.0028

Tab. 10 - Stima dei coefficienti, errore standard, statistica di Wald e p-value utilizzando il metodo di selezione forward

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
IIB	5.47	2.46	4.94	0.0263
IIF	-0.10	0.03	12.69	0.0004
RCI	-1.92	0.76	6.29	0.0122

In base alla metodologia già illustrata, si procede ad una verifica dell'ipotesi di linearità per le variabili incluse. Nell'ambito delle analisi effettuate⁹, per il modello di selezione backward si evidenzia la necessità di effettuare analisi più approfondite sulla variabile GIT. Queste conducono alla trasformazione di quest'ultima in qualitativa sulla base dei terzili e alla definizione del Modello 2 (Tabella 11). Per quanto riguarda il modello forward, è stato invece necessario effettuare una trasformazione quadratica della variabile IIF che ha portato alla formulazione del Modello 3 (Tabella 12).

Tab. 11 - Stima dei coefficienti, errore standard, statistica di Wald e p-value per il modello con GI ripartita in classi secondi e terzili (Modello 2)

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
GITT2	-2.10	0.73	8.31	0.0039
GITT3	-3.46	1.02	11.56	0.0007
ID	-1.05	0.69	2.30	0.1297
IIB	4.80	1.86	6.85	0.0089
OF	0.05	0.05	1.73	0.1880

Tab. 12 - Stima dei coefficienti, errore standard, statistica di Wald e p-value per il modello forward con forma quadratica per IIF (Modello 3)

Parametro	Stima	Err. std.	χ^2 Wald	Pr > χ^2
RCI	-0.7867	0.3737	4.43	0.0353
IIB	4.1017	1.5719	6.81	0.0091
IIF	-0.0848	0.0240	12.49	0.0004
IIF ²	0.0009	0.0004	5.46	0.0195

5. L'interpretazione economico-aziendale

La capacità di una azienda di fare fronte alle proprie obbligazioni debitorie è conseguenza di due fenomeni connessi: da un lato la componente competitiva del prodotto e del servizio offerto sul mercato, dall'altro la sostenibilità della struttura finanziaria (Roggi e Giannozzi, 2008). La prima caratteristica, ovvero la competitività dell'impresa, può essere valutata attraverso due aspetti: la potenzialità di sviluppo del core business e la capacità di generare rendimenti positivi sul capitale investito. La prima è colta nell'evoluzione e nella dimensione dei dati di bilancio connessi al turnover, al totale delle immobilizzazioni e delle risorse patrimoniali; la seconda è de-

⁹I dettagli vengono omissi in questo contesto ma sono disponibili su richiesta.

scritta dai numerosi indicatori di performance, di breve o di lungo termine, elaborati nelle discipline di analisi di bilancio (Montrone, 2005).

La sostenibilità delle fonti di finanziamento contratte riguarda essenzialmente tre aspetti: la consistenza dell'indebitamento, la capacità di far fronte alle scadenze periodali e l'esposizione al rischio di variazioni del costo del debito. Gli indici che mettono in relazione la quantità di fonti di finanziamento di terzi con le misure della dimensione aziendale, quali il capitale proprio o il fatturato, consentono di valutare la propensione del manager al ricorso a fonti di finanziamento esterne. In altri termini, un più elevato grado di indebitamento comporterà, verosimilmente, maggiori difficoltà di onorare gli impegni presi nei confronti dei finanziatori e, parallelamente, una maggiore sensibilità alle variazioni dei tassi di interesse. Inoltre, l'osservazione degli indicatori di liquidità consente di valutare la capacità dell'impresa di far fronte alle necessità di cassa di breve termine e di evitare crisi di liquidità.

Sono queste le considerazioni che hanno condotto alla prima selezione degli indicatori riportati in Tabella 2. L'obiettivo delle analisi statistiche è quello di identificare quali fra questi siano maggiormente efficaci per l'analisi dello stato di salute dell'azienda e di stimare per ciascuno di essi un coefficiente che ne rifletta l'importanza nel prevedere l'insolvibilità.

I risultati ottenuti in questo studio sono utili sia nella costruzione di modelli di rating, che nel fornire un orientamento nell'amministrazione aziendale. In altri termini, i coefficienti stimati delle variabili possono costituire una solida base sia per la costruzione di modelli volti alla classificazione delle aziende in base al rischio di default, sia per indicare gli obiettivi di successo nella gestione aziendale.

Nella trattazione che segue vengono analizzati i risultati dell'analisi in chiave economica. Dapprima si descrivono le variabili chiave identificate nei modelli statistici (5.1) e se ne commentano le relazioni e la capacità di modificare la probabilità di insolvenza (5.2), successivamente si evidenziano le performance del modello nel discriminare le aziende sane da quelle in default (5.3) ed infine si commenta la seconda fase di ottimizzazione del modello (5.4).

5.1 Le variabili chiave identificate nei modelli

L'osservazione dei risultati consente di rilevare la riduzione dimensionale del problema: il rischio di default può essere stimato attraverso un numero circoscritto di variabili. L'analisi empirica identifica inizialmente tre variabili significative: RCCL, EBITDAV e ID (Tabella 8), che rappresentano rispettivamente la rotazione del capitale circolante lordo, il rapporto fra EBITDA e le vendite e l'indice di disponibilità. Più precisamente l'indice di rotazione del capitale circolante lordo è dato dal rapporto tra il totale dei ricavi delle vendite e prestazioni e l'attivo circolante, mentre EBITDAV è

pari al rapporto tra EBITDA e il totale dei ricavi delle vendite e delle prestazioni più altri ricavi; l'indice di disponibilità infine è pari al rapporto tra l'attivo circolante e il totale dei debiti a breve termine. L'attenzione ad eliminare ridondanza di informazioni all'interno delle variabili osservate appare premiante sia dalle risultanze statistiche sia dalla constatazione che le tre misure individuate appartengono a tre famiglie di indicatori diversi. Più precisamente l'indice di rotazione del capitale circolante lordo (RCCL) rappresenta un indice della gestione corrente, EBITDAV è un indice di redditività e l'indice di disponibilità (ID) appartiene agli indici di tipo finanziario.

Fra gli indici di gestione corrente è stato individuato quello maggiormente influenzato dai risultati dell'attività tipica dell'impresa e dalla capacità del manager di ridurre il più possibile il fabbisogno di finanziamento dell'attivo circolante. Inoltre, fra gli indici di redditività è stato individuato quello in maggior misura connesso al core business aziendale mentre risultano meno significativi gli indici influenzabili dalla gestione non tipica. Tra gli indici finanziari è risultato più significativo quello in grado di rilevare la capacità dell'impresa di finanziare la propria attività con il ricorso a fonti di finanziamento esterne di medio/lungo termine. Sono risultati invece meno significativi gli indici della struttura finanziaria di lungo termine, evidenza che può essere spiegata in considerazione della prassi, osservata tra le aziende del campione, di ricorrere fortemente al finanziamento esterno prediligendo in particolare orizzonti temporali di breve termine. Tale atteggiamento appare coerente con quanto emerge anche da analisi del tessuto industriale italiano contraddistinto da imprese di piccole e medie dimensioni caratterizzate da una spiccata tendenza alle fonti di indebitamento a breve termine sia per il soddisfacimento dei fabbisogni di liquidità che per il supporto dello sviluppo (Berti e Marchiori, 1998, pag. 54; Iannuzzi e Berardi, 2006, pag. 25).

5.2 I coefficienti stimati delle variabili

Il segno e l'entità delle stime dei coefficienti esprimono l'influenza di ogni indicatore sulla probabilità di default. I risultati ottenuti sono coerenti con le aspettative e le ricerche degli studi di economia aziendale. (Luppi e altri, 2007; Roggi e Giannozzi, 2008). È possibile affermare che lo stato di salute dell'azienda è diretta conseguenza della sua capacità di generare rendimenti attraverso le vendite dei beni e servizi prodotti, della capacità di far fronte alle proprie obbligazioni senza ricorrere a debiti di breve termine e della possibilità di disporre di adeguate risorse proprie in grado di finanziare l'attivo circolante.

La stima dei coefficienti per il Modello 1 (Tabella 8) mostra valori negativi per tutti gli indici risultati significativi: l'azienda che si caratterizza per valori elevati di tali indici avrà minore probabilità di default. Riguardo

alla influenza specifica degli indici sulla probabilità di default, i coefficienti stimati mostrano una netta preponderanza dell'indice di disponibilità, pari a -3.53. L'unità di misura delle variabili è in questo caso paragonabile trattandosi di rapporti fra poste di bilancio.

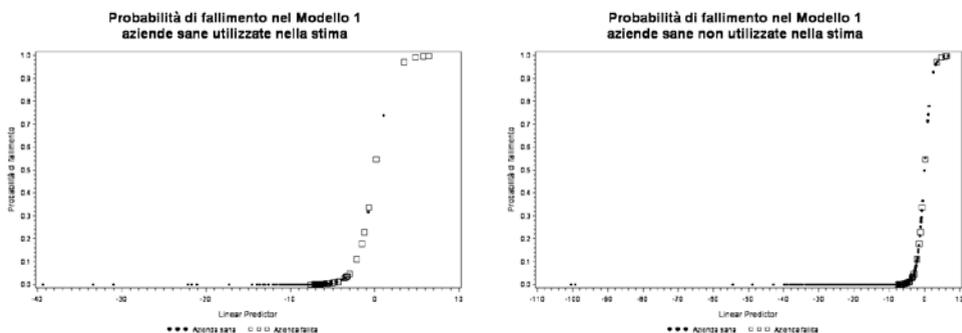
Il coefficiente di ID è circa cinque volte superiore a quello della variabile RCCL (pari a -0.73) e molto superiore a quello (non lineare ma leggermente quadratico) della variabile EBITDAV (pari a -0.0408 nella componente lineare); ciò fornisce una indicazione dell'elevato impatto di questa dimensione sulla probabilità di default.

5.3 Performance del modello

La capacità predittiva del modello è valutabile tramite l'analisi dell'andamento delle probabilità stimate di default rispetto al Linear Predictor (LP). Questo è ottenuto prendendo gli indicatori di bilancio delle aziende e moltiplicandoli per i parametri stimati del modello. Si riporta in Figura 2 l'andamento sia per le aziende utilizzate nella stima del modello (Figura 2 sinistra, indicatore quadrato per le aziende in default e indicatore punto per le aziende sane) che per le aziende sane non utilizzate nella stima del modello (Figura 2 destra). Si noti che per costruzione il modello logistico per studi caso-controllo non fornisce la stima dell'intercetta¹⁰.

Nell'interpretare questi grafici, di conseguenza, è opportuno tenere conto che il valore stimato della probabilità di default serve a fornire l'ordinamento fra le aziende (interpretazione ordinale, che non risente dell'intercetta) mentre può non essere accurato nella sua interpretazione cardinale.

Fig. 2 - Plot dell'andamento della probabilità stimata di default rispetto al Linear Predictor



⁹ Per la costruzione dei grafici l'intercetta è stata stimata attraverso il modello logistico classico.

Lo studio della rappresentazione grafica consente di rilevare la bontà del modello specificato, soprattutto nell'individuazione delle aziende a rischio. Le aziende fallite, rappresentate con indicatori quadrati, sono infatti posizionate nella parte destra dei grafici dove è più elevata la probabilità di default. L'obiettivo di discriminare le aziende sane dalle fallite è dunque raggiunto. Tale capacità si riscontra non solo nell'analisi del campione utilizzato ai fini di stima (Figura 2 sinistra), ma anche quando il modello prende in considerazione le aziende sane escluse dal processo di formazione degli strati (Figura 2 destra).

In entrambi i casi, tuttavia, è riscontrabile la presenza di alcuni errori di secondo tipo (ovvero errori dovuti all'associare ad aziende sane una elevata probabilità di default) in quanto nella zona della curva con più alta probabilità di fallimento si ritrovano anche aziende sane. Ai fini della valutazione del rischio di credito e dello sviluppo economico in senso lato, questo tipo di errore comporterebbe il mancato finanziamento di progetti che nel tempo si sarebbero potuti rivelare sostenibili. Tale errore è tuttavia sicuramente preferibile a quello di primo tipo che porta ad accettare una azienda fallita tra le sane.

5.4 Ottimizzazione del modello

Riguardo alla seconda fase del processo di analisi, nel metodo di selezione backward (Tabella 9) sono risultate significative il grado di indebitamento da terzi (GIT), l'indice di disponibilità (ID), l'indice di indebitamento a breve termine (IIB) e il rapporto degli oneri finanziari rispetto al fatturato (OF). Adottando il metodo di selezione forward (Tabella 10), invece, sono emerse come variabili significative l'indice di indebitamento a breve termine (IIB), l'indice di indipendenza finanziaria (IIF) e l'indice di rotazione del capitale investito (RCI).

Il modello di selezione backward evidenzia l'importanza della riduzione delle fonti di finanziamento a breve termine. Il grado di indebitamento è dato dal rapporto tra patrimonio netto e totale dei debiti. Gli indici di disponibilità, di cui si è già parlato precedentemente, e quello di indebitamento a breve termine evidenziano la necessità di ridurre la dipendenza della azienda da questa forma di finanziamento. Infine, il rapporto tra oneri finanziari e fatturato segnala l'importanza di mantenere un livello di sostenibilità dell'azienda rispetto ai costi del debito.

L'ottimizzazione del modello backward comporta l'introduzione della variabile GIT ripartita in terzili (Tabella 11 - Modello 2). I coefficienti di stima evidenziano la relazione positiva tra l'indice di indebitamento a breve termine e la probabilità di default (pari a 4.80), parallelamente mostrano la relazione inversa per il grado di indebitamento da terzi: minore è il grado di indebitamento e minore è la probabilità di default (si osservi l'elevata

riduzione della probabilità di default per le imprese nel secondo e terzo terzile della distribuzione). Anche l'indice di disponibilità è caratterizzato da una relazione inversa ma con minore intensità rispetto alle precedenti variabili (-1.05).

Il modello backward restituisce variabili estremamente concentrate nel comparto degli indici finanziari. La distribuzione della variabile GIT mostra una preoccupante sottocapitalizzazione delle aziende da noi analizzate: se è verosimile attendersi valori ridotti della variabile nelle aziende fallite, sembra non essere coerente e ragionevole la distribuzione delle aziende sane in cui il valore massimo dell'indice è pari ad appena il 15%.

Anche il modello ottenuto con la selezione forward (Tabella 12 - Modello 3) appare coerente con i risultati precedenti. L'indice di indipendenza finanziaria IIF risulta simile, in struttura, all'indice GIT essendo il rapporto tra patrimonio netto e totale dell'attivo. Infine, l'indice di rotazione del capitale investito RCI è pari al rapporto tra totale dei ricavi da vendite e prestazioni e totale dell'attivo; esso appare simile all'indice di rotazione del capitale circolante lordo RCCL individuato nella prima formulazione del modello. La stima dei coefficienti sottolinea ancora la relazione positiva tra l'indice di indebitamento a breve termine e la probabilità di default (il coefficiente stimato è pari a 4.10) mentre appaiono meno importanti i coefficienti stimati per le altre variabili (pari a -0.79 per l'indice di rotazione del capitale investito; -0.08 e 0.001 rispettivamente per l'indice di indipendenza finanziaria e la sua trasformazione quadratica).

6. Conclusioni

L'incessante evoluzione della regolamentazione finanziaria, l'attenzione delle autorità di vigilanza e, più recentemente, le ripercussioni della crisi finanziaria hanno imposto una modifica dei modelli di valutazione del merito creditizio da parte degli intermediari a favore di metodi quantitativi basati su solidi processi di credit scoring statistico (Palazzi, 2007). In questo lavoro si è individuato un modello di valutazione dello stato di salute delle aziende, tenendo in considerazione la localizzazione territoriale e la specializzazione di settore, similmente a quanto un intermediario finanziario è chiamato a operare nei processi di valutazione di impiego.

L'analisi effettuata ha preso in considerazione le aziende umbre con ricavi delle vendite positivi nell'anno 2006, verificandone lo stato giuridico al maggio 2010. I dati hanno evidenziato l'evento fallimento come un evento raro e pertanto lo studio si è focalizzato verso l'implementazione di uno strumento statistico di tipo caso-controllo. Ciò ha portato ad associare ad ogni azienda fallita tre aziende attive ad essa simili per codice ATECO e forma giuridica. Attraverso un procedimento statistico rigoroso si è giunti alla individua-

zione di tre distinti modelli che forniscono una buona visione d'insieme e consentono una interpretazione della situazione aziendale. Ciò ha consentito di ridurre l'iniziale paniere di indicatori, composto da oltre 40 rapporti ad appena 8 variabili significative. I modelli così definiti appaiono particolarmente parsimoniosi nella numerosità delle variabili considerate se paragonati con precedenti analisi empiriche condotte in riferimento ad aziende italiane (Cannata e altri, 2002; Luppi e altri, 2007; Roggi e Giannozzi, 2008).

Il primo modello ha preso in considerazione un indice della gestione corrente, uno di redditività ed uno di tipo finanziario; il secondo ed il terzo modello, molto simili fra loro dal punto di vista interpretativo, mostrano come il grado di indipendenza finanziaria dell'azienda sia rilevante nell'individuazione delle sue prospettive di permanenza sul mercato. In particolare il primo modello premia la capacità del manager di ridurre il fabbisogno di finanziamento dell'attivo circolante, la redditività derivante dal core business e la capacità di attrarre finanziamenti di medio/lungo termine; il secondo modello, più spostato sugli indicatori finanziari, gratifica la riduzione del peso dei finanziamenti a breve termine sul totale delle passività e la sostenibilità dell'azienda; il terzo modello, specularmente rispetto al precedente, premia le imprese con elevato peso di finanziamenti a lungo termine ed evidenzia l'importanza sia di un elevato rapporto fra patrimonio netto e totale dell'attivo patrimoniale che del rapporto tra il totale dei ricavi e l'attivo di bilancio.

Il rigoroso processo statistico di selezione degli indicatori di bilancio più significativi ha portato a conclusioni coerenti con gli studi di economia aziendale precedenti e contemporaneamente ad importanti risultati nell'ambito delle discipline scientifiche coinvolte. L'approccio adottato ha consentito l'individuazione di indici di riferimento, la determinazione della loro relazione positiva o negativa e la quantificazione dell'effetto sulla probabilità di default. La disponibilità di un coefficiente che sintetizzi tale impatto permette agli operatori di disporre di una chiave di lettura precisa del fenomeno aziendale: gli intermediari finanziari hanno la possibilità di costruire un processo di valutazione del rischio coerente e performante e le imprese possono disporre di una bussola in grado di indirizzare le scelte aziendali nella direzione migliore. Il processo di valutazione del merito creditizio si compone di diversi moduli e fra questi quello a maggiore carattere quantitativo ne rappresenta solo una parte. La coerenza con le discipline aziendali permette di inserire il modulo statistico all'interno di un processo costruttivo e non come imposizione regolamentare obbligatoria da correggere attraverso informazioni esterne del ciclo economico o di tipo qualitativo fornite dall'imprenditore.

I risultati delle nostre elaborazioni mostrano un'elevata capacità previsionale del modello nel discriminare le aziende sane da quelle in fallimento. Tuttavia mediante lo studio sul campione di convalida, si trova che ad

alcune aziende sane viene assegnata sulla base del modello una elevata probabilità di fallimento. Ogni metodo statistico, inoltre, presenta un limite dovuto alla qualità dei dati. In questo contesto ciò significa disporre di informazioni attendibili e tempestive per prevedere eventuali criticità sia a livello di singola azienda che di sistema economico. La qualità dei dati tende generalmente a deteriorarsi alla diminuzione della dimensione aziendale. Ciò rende più complessa l'applicazione e lo sviluppo dei modelli nei contesti caratterizzati da imprese medio piccole. In ultimo, il nostro lavoro evidenzia la necessità di competenze specifiche e passaggi successivi indispensabili ad una saggia applicazione dei modelli statistici.

In ogni caso, la costruzione dei modelli di rating a base statistica deve essere preceduta da una approfondita analisi delle distribuzioni delle aziende e da uno studio sulla forma funzionale che meglio rappresenta l'influenza degli indicatori di bilancio sulla probabilità di fallimento. L'approccio seguito nell'ambito di questo lavoro rappresenta, in tal senso, un utile esempio di indagine. Occorre inoltre rilevare che la elevata quantità di informazioni omesse nel dataset ha richiesto un notevole sforzo di implementazione e non ha consentito di valutare le capacità esplicative di indicatori di bilancio di potenziale interesse.

*Francesca Pierri,
Università di Perugia,
francesco.pierri@stat.unipg.it*

*Alberto Burchi,
Università di Perugia,
burchi@unipg.it*

*Elena Stanghellini,
Università di Perugia,
elena.stanghellini@stat.unipg.it*

Sommario

Obiettivo di questo lavoro è proporre un processo di sviluppo di un modello statistico rigoroso per la valutazione del rischio di credito. Il lavoro è volto ad individuare la capacità predittiva degli indici di bilancio e indicare quale sottoinsieme sia in grado di misurare più accuratamente lo stato di salute di una azienda. Attraverso uno studio caso-controllo, gli autori propongono un metodo di analisi in grado di individuare le variabili più rilevanti in considerazione delle caratteristiche territoriali e settoriali del campione di analisi. Il lavoro presenta caratteristiche innovative riguardo al metodo di analisi adottato, alla numerosità degli indici analizzati, all'individuazione di un numero ristretto di variabili chiave e alla quantificazione degli effetti delle variazioni degli indici sulla misura della probabilità di default.

Abstract

The aim of this work is to propose a rigorous statistical procedure to identify which subset of balance sheet indicators is able to accurately measure the default probability of a firm and to assess their predictive capacity. Through a case-control study on firms operating in the Italian region of Umbria, and after performing a matching based on industry and legal status, the authors propose a statistical procedure of analysis which allows to identify the most influential variables and then study the functional form linking their effect to the probability of default. The paper presents innovative features with respect to the number of indicators considered, the identification of a limited number of key variables and the quantification of their effect on the default probability.

Classificazione Jel: G21, G33, C25, D22

Parole chiave (Keywords): analisi di bilancio, rischio di credito, regressione logistica condizionata, caso-controllo

Bibliografia

Altman E. I.; Hotchkiss E. (2006). *Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt*. Wiley Finance Series, Hoboken, NJ.

Altman E.; Marco G.; Varetto F. (1994). *Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the italian experience)*. *Journal of Banking and Finance*, 18(3)(3), 505–529.

Altman E. I. (1968). *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. *The Journal of finance*, 23(4), 589–609.

Altman E. I.; Sabato G. (2007). *Modelling credit risk for smes: Evidence from the u.s. market*. *Abacus*, 43(3), 332–357.

Beaver W. H. (1966). *Financial ratios as predictors of failure*. *Empirical Research in Accounting: Selected Studies 1966*, Vol. 4, 41 (71–111).

Berti A.; Marchiori M. (1998). *Intermediari creditizi relazioni di clientela con le PMI*. *Piccola Impresa / Small Business*, n. 3, 1998.

Bracalente B. (2010). *Caratteri strutturali e scenari di sviluppo regionale. L'Umbria verso il 2020*. Franco Angeli.

Cannata F.; Fabi F.; Laviola S. (2002). *Rating interni e gestione del rischio di credito: la misurazione delle probabilità di insolvenza*. *Bancaria*, 4, 45–62.

Cox D. R.; Hinkley D. V. (1974). *Theoretical Statistics*. Chapman and Hall, London, UK.

De Laurentis G.; Maino R. (2009). *I rating a base statistica*. Bancaria Editrice.

Friedman C.; Huang J.; de Servigny A.; Salinas E. (2003). *A utility based private firm default probability model*. Standard & Poor's. Risk Solutions.

Giampaoli A. (2006). *La previsione delle insolvenze*. *Piccola Impresa / Small Business*, n. 3.

Hosmer D. W.; Lemeshow S. (2000). *Applied logistic regression*. Wiley Series in probability and statistics. Wiley-Interscience Publication.

Huyhebaert N.; Gaeremynck A.; Roodhooft F.; Van de Gucht L. M. (2000). *New firm survival: The effects of start-up characteristics*. *Journal of Business Finance & Accounting*, 27(5-6), 627–651.

Iannuzzi E.; Berardi M. (2006). *Interpretazione critico-empirica del modello Modigliani-Miller: il problema del default delle PMI. Una comparazione anglo-Italiana*. *Piccola Impresa / Small Business*, n.1, 2006.

King G.; Zeng L. (2001). *Explaining rare events in international relations*. *International Organization*, 55(3), 693–715.

Luppi B.; Marzo M.; Scorcu A. (2007). *A credit risk model for Italian SMEs*. Università degli studi di Modena, Dipartimento di Economia Politica, Working Paper Series, 600.

Lusignani G.; Sironi A. (1999). *L'innovazione nel finanziamento delle PMI*. In Forestieri G. e Onado M. (a cura di). *Banche, mercati e strutture di intermediazione*. EGEA.

Montrone A. (2005). *Il sistema delle analisi di bilancio per la valutazione dell'impresa*. *Economia - Monografie - Economia e gestione aziendale*.

Moody's (2000). *RiskCalc Private Model: Moody's Default Model For Private Firms*. May 2000.

Palazzi E. (2007). *Vincoli e opportunità di Basilea 2 per lo sviluppo delle pmi: le opinioni degli operatori economici*, *Piccola Impresa / Small Business*, n. 2.

Pompe P. P.; Bilderbeek J. (2005). *The prediction of bankruptcy of small-and medium-sized industrial firms*. *Journal of Business Venturing*, 20(6), 847 – 868.

Roggi O.; Giannozzi A. (2008). *Predicting sme default risk. Does regional model make sense?* 8th Global Conference on Business and Economics.

Sironi A.; Resti A. (2008). *Rischio e valore nelle banche*. EGEA.

Standard & Poor's (2002). *Credit risk tracker italy, technical document*. Relazione tecnica.

Stanghellini E. (2009). *Introduzione ai metodi statistici per il credit scoring*. Springer Verlag.

Watson J. (2003). *The potential impact of accessing advice on sme failure rates*. 16th Annual Conference of Small Enterprise Association of Australia and New Zealand.