

NOTA METODOLOGICA GLI INDICATORI SINTETICI

L'analisi del contributo fornito dalle relazioni tra manifattura e servizi alla recente performance dei settori industriali richiede la disponibilità di dati in grado di sintetizzare fenomeni complessi e articolati. Nell'ambito del presente Rapporto, nei casi in cui l'informazione statistica era disponibile, si è provveduto all'elaborazione di opportuni indicatori compositi. Di seguito si illustrano le motivazioni e la metodologia di costruzione alla base di ciascuno di essi.

1. Indicatore di integrazione verticale d'impresa

L'integrazione verticale viene definita come la combinazione, sotto una unica proprietà, di due o più stadi della produzione e/o della distribuzione che sono abitualmente separati. La letteratura economica ha individuato da tempo una pluralità di ragioni all'origine della scelta, da parte di un'impresa, di preferire un maggiore o minore grado di integrazione verticale, quali la ricerca di strutture aziendali più efficienti data la "dinamicità" del mercato di riferimento (Adelman 1955), la gestione del *trade-off* tra costi di organizzazione e costi di mercato teorizzata dall'approccio dei costi di transazione (Williamson, 1985), l'inserimento di un determinato grado di integrazione nell'ambito di una più ampia strategia di sostenibilità di lungo periodo (Hart, 1995; Russo and Fouts, 1997).

Anche a causa della ricchezza delle interpretazioni teoriche, ancora non si dispone di una misura empirica universalmente accettata di integrazione verticale. L'indicatore più noto e diffuso è costituito dal rapporto tra valore aggiunto e fatturato (Adelmann, 1955), tanto più vicino all'unità quanto più una impresa è verticalmente integrata. Allo stesso tempo, tuttavia, sono noti i principali limiti di tale misura, quali la necessità di cautela nei confronti intersettoriali (dal momento che l'indice varia anche a seconda che si esaminino settori a monte o a valle della filiera produttiva) o il fatto che per costruzione essa registra, oltre all'integrazione verticale, anche la maggiore capacità di una impresa di generare reddito a parità di integrazione (legata alla presenza di fattori specifici come la capacità manageriale).

La natura dei dati utilizzati nel presente Rapporto consente di elaborare un indicatore di integrazione verticale *à la* Adelman "aggiustato", finalizzato a sterilizzare l'effetto della maggiore redditività a parità di integrazione. In particolare, seguendo Buzzell (1983) si è proceduto a calcolare il seguente indicatore:

$$\frac{\text{Valore aggiunto} - \text{profitti netti} + (\text{tasso di profitto normale} \times \text{investimenti al valore contabile})}{\text{Fatturato} - \text{profitti netti} + (\text{tasso di profitto normale} \times \text{investimenti al valore contabile})} \quad [2]$$

La misura di Buzzell, a rigore, prevede l'utilizzo di un tasso di profitto "normale" unico per tutto il sistema produttivo; di conseguenza, per meglio tener conto della specificità dell'economia italiana (e del periodo considerato, caratterizzato

dalla seconda fase della crisi), in questo Rapporto l'indicatore *à la Buzzel* è stato ulteriormente modificato introducendovi due correzioni: *a)* attraverso l'integrazione della base dati Frame con i bilanci civilistici, in luogo dell'utilizzo di un unico tasso di profitto sono stati applicati i tassi medi di profitto relativi a ciascun settore Ateco a 2 digit; *b)* l'investimento al valore contabile è stato approssimato attraverso l'uso delle immobilizzazioni materiali, immateriali e finanziarie, in base al valore iscritto in bilancio.

2. Indicatore di efficienza d'impresa

La nuova base dati Frame predisposta dall'Istat, che riporta le principali variabili del conto economico e dati sull'input di lavoro per tutte le imprese italiane attive (4,4 milioni di unità, con un'occupazione pari a circa 17 milioni di addetti), permette di stimare una misura del grado di efficienza produttiva dell'impresa che, ai fini del presente Rapporto, è possibile ricondurre essenzialmente al livello di valore aggiunto prodotto data la dotazione dei fattori di produzione. In particolare, l'efficienza di un'impresa viene stimata a partire da un modello di frontiera di produzione stocastica (Aigner, Lovell e Schmidt, 1977; Meeusen e Van den Broeck, 1977), che valuta il divario tra il livello di output potenziale ricavabile dall'utilizzo ottimale di un dato insieme di fattori di produzione – nel nostro caso il livello massimo di valore aggiunto ottenibile – e il livello effettivamente prodotto (in tale ottica, pertanto, si tratta a rigore di una misura di *inefficienza tecnica*).

Più in dettaglio, la funzione di produzione viene definita nei seguenti termini funzionali:

$$y_i = f(\mathbf{x}_i; \boldsymbol{\beta}) \exp(v_i) TE_i \quad [1],$$

dove il valore aggiunto y_i viene generato da un insieme di input produttivi \mathbf{x}_i secondo la relazione definita dai parametri $\boldsymbol{\beta}$, mentre il termine $\exp(v_i) TE_i$ rappresenta lo scarto fra il potenziale produttivo e l'effettivo ammontare di valore aggiunto generato, dovuto all'insieme di fattori casuali ed inefficienza tecnica. Su queste basi l'inefficienza tecnica può essere definita come il rapporto tra la produzione potenziale e quella effettiva:

$$TE_i = \frac{y_i}{f(\mathbf{x}_i; \boldsymbol{\beta}) \exp(v_i)} \quad [2]$$

In termini di stima, il modello [1] diviene:

$$y_i = \alpha + \mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i \quad [3]$$

dove il termine di errore è composto da una componente erratica v_i ed una componente u_i di inefficienza tecnica, stimata separatamente.

In questo Rapporto è stata stimata una trasformazione logaritmica della funzione di produzione Cobb-Douglas che mette in relazione il valore aggiunto al numero di addetti e al livello degli ammortamenti, a sua volta approssimazio-

ne della dotazione di capitale (l'utilizzo di specificazioni funzionali diverse, quali trans-logaritmica e Fourier flessibile ha portato a risultati analoghi). A partire da tale stima, componendo gli errori sulla base del metodo di Battese e Coelli (Coelli, Rao e Battese, 1998; Kumbhakar e Lovell, 2000), è stato possibile confrontare il valore aggiunto potenziale dell'impresa con quello da essa effettivamente ottenuto e, dunque, determinare l'efficienza produttiva dell'impresa in termini di distanza tra le due quantità.

L'analisi è stata condotta su un dataset che, a partire dalla popolazione di circa 4,4 milioni di unità delle imprese attive in Italia nel 2011, considera solo le unità che: *a)* hanno fatturato superiore a 30.000 euro; *b)* impiegano almeno 1 addetto; *c)* presentano ammortamenti positivi; *d)* hanno valore aggiunto positivo; *e)* non appartengono ai settori produttivi dei tabacchi, delle attività finanziarie, assicurative e immobiliari.

Le condizioni *a)* e *b)* garantiscono la rilevanza economica delle imprese osservate; le ipotesi *c)* e *d)* sono necessarie alla definizione del modello (se sono negative non è possibile la trasformazione logaritmica). L'ipotesi *c)*, inoltre, definisce un'ulteriore condizione di rilevanza economica delle osservazioni, ovvero la presenza di capitale. L'ipotesi *e)* è funzionale alla coerenza con l'insieme di settori considerati dal Rapporto. In tal modo si circoscrive un insieme di oltre 2 milioni di imprese, che nel 2011 spiegavano circa il 75 per cento degli addetti complessivi del sistema e il 79 per cento di valore aggiunto totale. Su queste unità si sono effettuate due diverse tipologie di stima: la prima non tiene in considerazione la tecnologia prevalente nei diversi settori; la seconda, al contrario, include l'elemento tecnologico e stima l'efficienza tecnica all'interno di cluster settoriali, a un livello di aggregazione di Ateco-2 digit. Nel primo caso, dunque, la stima individua, per ogni osservazione, un livello di efficienza tecnica che consente un confronto tra le imprese di tutto il sistema produttivo; nel secondo caso, invece, viene individuato un livello di efficienza tecnica "relativo", circoscritto al settore in cui le unità operano. Per agevolare l'interpretazione dei risultati, infine, si è provveduto a una trasformazione dell'indicatore tale che il valore possa essere letto come la percentuale di valore aggiunto potenziale effettivamente ottenuta dall'impresa.

3. Indicatore sintetico di efficienza dei servizi acquistati dalla manifattura

Nel valutare il ruolo ricoperto dall'efficienza dei *business services* sulla performance delle imprese manifatturiere, la tipologia dei dati alla base di questo Rapporto ha suggerito di elaborare un indicatore sintetico secondo i passi seguenti. A partire dall'indicatore di efficienza individuale descritto in precedenza in questo stesso riquadro, si è ricavata una media settoriale a livello di Ateco-2 digit per il comparto dei servizi. In tal modo si è ottenuta una misura dell'efficienza e_s specifica di ciascun settore s dei servizi.

Successivamente, per collegare tale misura alla performance dei settori manifatturieri, si è provveduto a correggere e_s per il peso che l'acquisto del servizio s ricopre nell'ambito della struttura dei costi intermedi del comparto manifatturiero m , ottenendo l'indicatore seguente:

$$e_{sm} = e_s \times w_{sm}$$



dove il peso w_{sm} è ricavato sulla base delle tavole “use” di contabilità nazionale rapportando gli acquisti di s da parte del settore m sul totale dei costi intermedi di m .

L'indicatore e_{sm} misura di fatto quanto, dell'efficienza del servizio s , il comparto manifatturiero m “assorbe” sulla base dei rapporti produttivi esistenti tra i settori s e m .

Sulla base della classificazione dei servizi all'impresa riportata nel Riquadro “Una classificazione dei servizi all'impresa”, l'indicatore e_{sm} è stato ulteriormente aggregato in tre gruppi, uno relativo all'efficienza dell'insieme dei servizi ad alta intensità di conoscenza (KIBS), uno che sintetizza l'efficienza dei servizi operativi (SO), uno relativo all'efficienza dei servizi di rete (SR):

$$E_{mj} = \sum_{s=1}^n e_{sm}$$

Dove $j = \text{KIBS, SO, SR}$ e n è il numero dei settori Ateco a 2 digit per ciascuno dei tre gruppi identificati.

E_{mj} , pertanto, rappresenta una misura dell'efficienza media del gruppo j di servizi acquistati dal settore manifatturiero m .

NOTA METODOLOGICA PERFORMANCE MANIFATTURIERA E EFFICIENZA DEI SERVIZI: UN MODELLO LINEARE A EFFETTI MISTI (MULTILEVEL)

L'analisi empirica delle relazioni tra efficienza degli *input* di servizi acquistati dalle imprese manifatturiere e *performance* d'impresa è stata effettuata attraverso un modello di regressione per lo studio di dati dipendenti. L'analisi preliminare dei dati ha evidenziato una forte correlazione tra osservazioni individuali ripetute nel tempo (ad esempio, del fatturato o di altre variabili di *performance*), con il risultato che esse tendono a posizionarsi in maniera sistematica (al di sopra o al di sotto) rispetto alla media di settore (o dell'intero campione). Tale dipendenza può derivare dalla presenza di una struttura gerarchica nei dati; essa caratterizza la generalità delle basi di dati d'impresa, in cui le unità di rilevazione sono raggruppate in uno o più *cluster* di riferimento (coerentemente al disegno di rilevazione) e le corrispondenti variabili misurate in istanti successivi di tempo. Nel caso in esame, lo studio è condotto su una base di dati di imprese manifatturiere (con 20 addetti e più) relativa al periodo compreso tra il 2012 e il 2014. Adottando la terminologia dei modelli gerarchici, le imprese rappresentano le unità di rilevazione di primo livello, mentre i settori produttivi (identificati dai comparti Ateco a 2 digit) costituiscono il livello più aggregato (unità secondo livello).

Un approccio tradizionalmente utilizzato per l'analisi empirica è rappresentato dal modello a effetti fissi: questi ultimi spiegano la variabilità tra i gruppi (*between*); tuttavia, i restanti parametri del modello sono costanti sia tra gli individui sia tra i gruppi.

Il modello lineare a effetti misti (Snijders e Bosker, 1999) rappresenta una generalizzazione di un modello lineare ad effetti fissi poiché questi ultimi includono fattori riferibili sia alle unità statistiche di interesse (unità di primo livello) sia ai corrispondenti *cluster* di appartenenza (unità statistiche di secondo livello). Per quanto concerne gli effetti non osservabili, la varianza totale del modello è scomposta in una o più componenti, ciascuna corrispondente a uno specifico livello della struttura gerarchica della base di dati (effetti *random*).

Nel caso di un modello lineare per la variabile dipendente y_{ij} (impresa i , settore j) e una sola variabile esplicativa,

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \zeta_{ij} \quad (1)$$

il termine complessivo di errore ζ è suddiviso in due componenti: una specifica per ciascun settore ζ_j ($\zeta_j \sim N(0, \psi)$); una idiosincratia ε_{ij} ($\varepsilon_{ij} \sim N(0, \theta)$), specifica per ciascuna unità produttiva. Sotto l'ipotesi $\text{Cov}(\zeta_j, \varepsilon_{ij}) = 0$, la varianza totale $\text{Var}(y_{ij}) = \psi + \theta$ (ψ è detta varianza *between*, θ è la varianza *within*). Il rapporto tra la varianza *between* e la varianza totale viene definito come coefficiente di correlazione intragruppo e fornisce una misura della variabilità complessiva attribuibile ad eterogeneità settoriali.

L'estensione a un modello lineare con effetti misti si ottiene includendo una o più variabili esplicative x_{ij} nel modello (1) (equazione di primo livello), dove l'intercetta β_{oj} e il parametro β_{lj} presentano una struttura stocastica del tipo

$$\begin{aligned}\beta_{oj} &= \gamma_{2o} + u_{2j} \\ \beta_{lj} &= \gamma_{1l} + u_{1j}\end{aligned}\quad (2)$$

(equazioni di secondo livello) dove i parametri γ rappresentano gli effetti fissi, mentre le componenti stocastiche u si assumono con media zero, varianze costanti e indipendenti dal termine di errore ξ . Sostituendo la (2) in (1) si ottiene la specificazione in forma ridotta di un modello lineare con effetti misti,

$$y_{ij} = \gamma_{2o} + \gamma_{1l}x_{ij} + u_{2j} + u_{1j}x_{ij} + e_{ij}\quad (3)$$

Questo modello permette di specificare una componente stocastica per l'intercetta (u_{2j}) e una per il parametro corrispondente alla variabile x_{ij} di primo livello (u_{1j}), che fornisce una stima della variabilità per x_{ij} a livello di settore (non spiegata dagli effetti fissi).

Tale tipologia di modelli permette un utilizzo efficiente dell'informazione statistica utilizzata per questo lavoro: le unità elementari di rilevazione sono rappresentate dalle oltre 30.000 imprese manifatturiere con almeno 20 addetti; i 23 settori manifatturieri costituiscono le variabili di secondo livello. La specificazione adottata per l'analisi empirica è la seguente,

$$\begin{aligned}y_{ij} = & \gamma_{2o} + \gamma_{1,1}prod_{ij} + \gamma_{1,2}prod_{ij}^2 + \gamma_{1,3}clav_{ij} + \gamma_{1,4}valfatt_{ij} + \gamma_{1,5}cladd_{ij} + \gamma_{1,6}dexp_{ij} \\ & + \gamma_{1,7}ie1_j + \gamma_{1,8}ie1_j * dexp_{ij} + \gamma_{1,9}ie2_j + \gamma_{1,10}ie2_j * dexp_{ij} + [u_{2,1} + u_{1,1}ie1_j + u_{1,2}ie2_j + e_{ij}]\end{aligned}\quad (4)$$

dove y_{ij} è la variazione cumulata del fatturato in valore osservata nel periodo 2012-2014.

Il modello per gli effetti fissi include le seguenti covariate di primo livello:

- produttività del lavoro, in termini di valore aggiunto per addetto (*prod*);
- costo del lavoro per addetto (*clav*);
- indice di integrazione verticale (*valfatt*, si veda la Nota metodologica "Gli indicatori sintetici");
- tre variabili *dummy* corrispondenti alle classi dimensionali d'impresa (*cladd*);
- variabile *dummy* che indica se l'impresa è esportatrice (*dexp*).

La specificazione degli effetti fissi include anche esplicative di secondo livello: si tratta delle variabili *ie1-ie2*, che costituiscono gli indici sintetici di efficienza (si veda la Nota metodologica "Gli indicatori sintetici") per due raggruppamenti dei settori dei servizi alle imprese: *knowledge intensive business services* (KIBS - *ie1*), Servizi di rete (SR - *ie2*).

Gli effetti fissi, infine, includono le interazioni tra gli indici di efficienza e la variabile *dexp* per tener conto della possibile eterogeneità tra le imprese domestiche e le esportatrici circa il contributo alla *performance* d'impresa fornito dall'efficienza degli *input* di servizi acquistati. Tutte le variabili continue sono trasformate ai logaritmi.

La specificazione degli effetti *random* (riportati in parentesi quadre nell'equazione (4)) permette di stimare per ciascun indice di efficienza una misura dell'eterogeneità non-osservata dovuta a specificità dei singoli settori manifatturieri.

Circa la scelta di un modello lineare con effetti *random*, evidenze preliminari fornite dal test di Hausman hanno confermato l'assenza di distorsioni sistematiche tra i parametri dei modelli stimati, rispettivamente, con effetti fissi ed effetti *random*. Il test del rapporto di verosimiglianza, con cui si confronta il potere esplicativo del modello a effetti misti con quello (*nested*) corrispondente alla sola equazione di primo livello (1), rifiuta l'ipotesi nulla confermando la maggiore capacità esplicativa della specificazione del modello non vincolato.

Nell'equazione (4) si assume che i fattori *sector specific* per gli indici di efficienza siano uguali per tutte le imprese manifatturiere. È tuttavia presumibile che caratteristiche osservabili delle imprese possano presentare differenti variabilità in risposta all'efficienza media dei servizi acquistati da ciascun settore della manifattura industriale. L'analisi ha riguardato la tipologia di impresa "esportatore/non-esportatore": la specificazione degli effetti *random* è stata estesa esplicitando un modello per la stima di effetti *random* specifici per le imprese esportatrici e non-esportatrici, entrambi associati ad eterogeneità settoriali di efficienza per l'input di servizi.

La specificazione degli effetti *random* del modello (4) è modificata come segue,

$$[u_{2,1} + u_{1,1}iel_j * D_{ij} + u_{1,2}iel_j * E_{ij} + e_{ij}] \quad (5)$$

dove gli effetti *random* sono ottenuti attraverso le interazioni di ciascun indicatore settoriale di efficienza con le *dummy* esportatore ($E_{ij}(I=1)$) e non-esportatore ($D_{ij}(I=1)$). Il modello (5), in particolare, riporta la specificazione utilizzata per condurre inferenza sugli effetti *random* differenziati per la variabile *iel* ($u_{1,1}, u_{1,2}$).

Rispetto a questa specificazione finale, il coefficiente di correlazione infragruppo è risultato particolarmente basso (0,02), sia per la bassa incidenza delle eterogeneità non osservata a livello settoriale (varianza *between*), sia per l'elevata varianza *within* non spiegata. Per valutare la bontà delle stime, una misura del coefficiente di determinazione multipla per un modello a lineare a due livelli è fornito dalla riduzione proporzionale della varianza totale dei residui del modello completo rispetto a un analogo modello senza covariate: tale indice è risultato pari a circa il 24 per cento in larga misura per il contributo delle variabili esplicative di primo livello.