

**ALLEGATO 2**

**Nota Tecnica e Metodologica**

**SG61F**

## **NOTA TECNICA E METODOLOGICA**

### **1. CRITERI PER LA COSTRUZIONE DELLO STUDIO DI SETTORE**

Di seguito vengono esposti i criteri seguiti per la costruzione dello studio di settore.

Oggetto dello studio è l'attività economica rispondente al codice ISTAT:

- 51.11.0 - Intermediari del commercio di materie prime agricole, di animali vivi, di materie prime tessili e di semilavorati.

La finalità perseguita è di determinare un “ricavo potenziale” tenendo conto non solo di variabili contabili, ma anche di variabili strutturali in grado di determinare il risultato di un'impresa.

A tale scopo, nell'ambito dello studio, vanno individuate le relazioni tra le variabili contabili e le variabili strutturali, per analizzare i possibili processi produttivi e i diversi modelli organizzativi impiegati nell'espletamento dell'attività.

Al fine di conoscere le informazioni relative alle strutture produttive in oggetto si è progettato ed inviato ai contribuenti interessati un questionario per rilevare tali informazioni (il codice del questionario relativo agli intermediari del commercio è SG61).

Il numero dei questionari inviati è stato pari a 3.829. I questionari restituiti sono stati 2.613, pari al 68,2% degli inviati.

Sui questionari sono state condotte analisi statistiche per rilevare la completezza, la correttezza e la coerenza delle informazioni in essi contenute.

Tali analisi hanno comportato, ai fini della definizione dello studio, lo scarto di 716 questionari, pari al 27,4% dei questionari rientrati.

I principali motivi di scarto sono stati:

- presenza di attività secondarie con un'incidenza sui ricavi complessivi superiore al 20%;
- settore merceologico prevalente e altri settori assenti e/o errati;
- volume delle provvigioni per tipologia delle aziende rappresentate non compilato;
- quadro F del questionario (modalità di espletamento dell'attività) non compilato;
- quadro G del questionario (elementi specifici dell'attività) non compilato;
- quadro M del questionario (elementi contabili) non compilato;
- errata compilazione delle percentuali relative alla tipologia di clientela (quadro F del questionario);
- ricavi dichiarati maggiori di 10 miliardi di lire;
- incongruenze fra i dati strutturali e i dati contabili contenuti nel questionario.

A seguito degli scarti effettuati, il numero dei questionari oggetto delle successive analisi è risultato pari a 1.897.

### ***1.1 IDENTIFICAZIONE DEI GRUPPI OMOGENEI***

Per segmentare le imprese oggetto dell'analisi in gruppi omogenei sulla base degli aspetti strutturali, si è ritenuta appropriata una strategia di analisi che combina due tecniche statistiche:

- una tecnica basata su un approccio di tipo multivariato, che si è configurata come un'analisi fattoriale del tipo *Analyse des données* e nella fattispecie come un'*Analisi in Componenti Principali*<sup>1</sup>;
- un procedimento di *Cluster Analysis*<sup>2</sup>.

L'utilizzo combinato delle due tecniche è preferibile rispetto a un'applicazione diretta delle tecniche di clustering.

In effetti, tanto maggiore è il numero di variabili su cui effettuare il procedimento di classificazione, tanto più complessa e meno precisa risulta l'operazione di clustering.

Per limitare l'impatto di tale problematica, la classificazione dei contribuenti è stata effettuata a partire dai risultati dell'analisi fattoriale, basandosi quindi su un numero ridotto di variabili (i fattori) che consentono, comunque, di mantenere il massimo delle informazioni originarie.

In un procedimento di clustering di tipo multidimensionale, quale quello adottato, l'omogeneità dei gruppi deve essere interpretata, non tanto in rapporto alle caratteristiche delle singole variabili, quanto in funzione delle principali interrelazioni esistenti tra le variabili esaminate che contraddistinguono il gruppo stesso e che concorrono a definirne il profilo.

Le variabili prese in esame nell'Analisi in Componenti Principali sono quelle presenti in tutti i quadri di cui si compone il questionario ad eccezione del quadro M che contiene i dati contabili presenti nella dichiarazione dei redditi.

---

<sup>1</sup> L'Analisi in Componenti Principali è una tecnica statistica che permette di ridurre il numero delle variabili originarie di una matrice di dati quantitativi in un numero inferiore di nuove variabili dette componenti principali tra loro ortogonali (indipendenti, incorrelate) che spieghino il massimo possibile della varianza totale delle variabili originarie, per rendere minima la perdita di informazione; le componenti principali (fattori) sono ottenute come combinazione lineare delle variabili originarie.

<sup>2</sup> La Cluster Analysis è una tecnica statistica che, in base ai fattori dell'analisi in componenti principali, permette di identificare gruppi omogenei di imprese (cluster); in tal modo le imprese che appartengono allo stesso gruppo omogeneo presentano caratteristiche strutturali simili.

Tale scelta nasce dall'esigenza di caratterizzare le imprese in base ai possibili modelli organizzativi, alle diverse modalità di espletamento dell'attività, etc.; tale caratterizzazione è possibile solo utilizzando le informazioni relative alle strutture operative, al mercato di riferimento e a tutti quegli elementi specifici che caratterizzano le diverse realtà economiche e produttive di una impresa.

I fattori risultanti dall'Analisi in Componenti Principali vengono analizzati in termini di significatività sia economica sia statistica, al fine di individuare quelli che collegano i diversi aspetti strutturali delle attività oggetto dello studio.

La Cluster Analysis ha consentito di identificare quattro gruppi omogenei di imprese. I principali aspetti strutturali delle imprese considerati nell'analisi sono:

- settore merceologico;
- tipologia di azienda rappresentata;
- tipologia della clientela;
- modalità organizzativa dell'impresa (essenzialmente impresa individuale con assenza sia di personale dipendente che di sottorete).

## **1.2 DESCRIZIONE DEI GRUPPI OMOGENEI**

Di seguito vengono riportate le descrizioni di ciascuno dei gruppi omogenei (cluster).

### **Cluster 1 - Intermediari dell'industria e di altri intermediari**

**Numerosità: 873**

Il raggruppamento è costituito da intermediari dell'industria che rappresentano mediamente 5 aziende distribuite fra: media industria (29%), piccola industria (23%), grande industria (20%) e altri agenti di commercio (15%).

Il portafoglio clienti è minimo (76), composto prevalentemente da industrie (43%), privati (27%) e grossisti (16%).

I consumi di carburante sono superiori alla media così come le spese di viaggio e di albergo ed i giorni di partecipazione a fiere (3).

Gli spazi destinati ad ufficio risultano presenti solo nel 24% dei casi con una superficie media di 40 mq, mentre il 50% fa un uso promiscuo dell'abitazione.

Le merceologie trattate sono relative ad alimenti zootecnici ed additivi (per il 51%), fibre tessili e tessuti in genere (per il 16%), cereali secchi e conservati (per il 7%).

### **Cluster 2 - Intermediari con clientela dettaglio**

**Numerosità: 349**

La caratteristica di questo raggruppamento è la tipologia della clientela rappresentata per l'81% da punti vendita al dettaglio. Il portafoglio clienti è mediamente ampio (114).

Gli intermediari di questo cluster rappresentano mediamente tre aziende riconducibili a: grossisti (26%), media industria (22%), grande industria (21%) e piccola industria (17%).

I consumi di carburante sono leggermente superiori alla media, le spese di viaggio sono limitate mentre quelle per alberghi e ristoranti sono superiori alla media.

La struttura non presenta particolare ampiezza, gli uffici sono presenti solo nel 22% dei casi (22 mq), mentre risultano prevalenti gli uffici ad uso promiscuo.

Le merceologie sono diverse e coerenti con la tipologia della clientela: fertilizzanti (25%), alimenti zootecnici e additivi (22%), antiparassitari (17%), bulbi-piante e fiori (14%), prodotti per il giardinaggio (9%).

### **Cluster 3 - Intermediari con clientela privati**

#### **Numerosità: 384**

Il cluster raggruppa gli intermediari che vendono direttamente ai privati (che in relazione alle merceologie trattate potrebbero essere individuati come gli agricoltori e gli allevatori) per il 75% e che gestiscono un portafoglio clienti molto ampio (214).

La tipologia di committenza è prevalentemente costituita da grossisti (41%) e servizi (23%).

I consumi di carburante sono limitati, così come le spese di viaggio, di albergo e ristorante.

Il personale dipendente è presente nel 16% dei casi, i collaboratori familiari nel 15% ed i soci con occupazione prevalente si rilevano nel 14% dei casi.

Gli spazi dedicati ad ufficio sono invece presenti, anche se in misura limitata (24 mq) nel 43% delle imprese, mentre si rilevano nel 23% dei casi 421 mq destinati a deposito.

Le merceologie rilevate sono: fertilizzanti (34%), alimenti zootecnici ed additivi (27%), antiparassitari (26%).

#### **Cluster 4 - Intermediari dell'ingrosso per la frutta**

##### **Numerosità: 268**

Questo è il raggruppamento meno numeroso con tipologia di committenza prevalentemente individuabile nei grossisti (83% delle provvigioni); rappresentano mediamente 7 aziende.

Gestiscono un portafoglio clienti limitato (83), costituito da punti vendita all'ingrosso (86%).

Risultano particolarmente elevate le spese per utenze telefoniche e fax, (tale dato porta a ritenere che all'interno di questo cluster vi siano anche i mediatori).

Gli spazi destinati ad ufficio, per quanto mediamente ampi (39 mq) si rilevano solo nel 28% dei casi, mentre il 30% utilizza uffici ad uso promiscuo.

Fra le tipologie merceologiche emerge la frutta fresca, secca e conservata presente per il 69%, oltre agli alimenti zootecnici.

### 1.3 **DEFINIZIONE DELLA FUNZIONE DI RICAVO**

Una volta suddivise le imprese in gruppi omogenei è necessario determinare, per ciascun gruppo omogeneo, la funzione matematica che meglio si adatta all'andamento dei ricavi delle imprese appartenenti al gruppo in esame. Per determinare tale funzione si è ricorso alla *Regressione Multipla*<sup>3</sup>.

La stima della “funzione di ricavo” è stata effettuata individuando la relazione tra il ricavo (variabile dipendente) e alcuni dati contabili e strutturali delle imprese (variabili indipendenti).

E' opportuno rilevare che prima di definire il modello di regressione si è proceduto ad effettuare un'analisi sui dati delle imprese per verificare le condizioni di “normalità economica” nell'esercizio dell'attività e per scartare le imprese anomale; ciò si è reso necessario al fine di evitare possibili distorsioni nella determinazione della “funzione di ricavo”.

In particolare sono state escluse le imprese che presentano costi e spese dichiarati nel quadro M superiori ai ricavi dichiarati.

Successivamente sono stati utilizzati degli indicatori economico-contabili specifici delle attività in esame:

- resa oraria per addetto =  $\text{ricavi}/(\text{numero addetti}^4 * 312 * 8)$

dove:

- numero addetti = 1 + numero dipendenti a tempo pieno + numero (ditte individuali) dipendenti a tempo parziale + numero apprendisti

---

<sup>3</sup> La Regressione Multipla è una tecnica statistica che permette di interpolare i dati con un modello statistico-matematico che descrive l'andamento della variabile dipendente in funzione di una serie di variabili indipendenti relativamente alla loro significatività statistica.

<sup>4</sup> Le frequenze relative ai dipendenti sono state normalizzate all'anno in base alle giornate retribuite.

+ numero assunti con contratto di formazione e lavoro o a termine e lavoranti a domicilio + numero collaboratori coordinati e continuativi che prestano attività prevalentemente nell'impresa + numero collaboratori dell'impresa familiare e coniuge dell'azienda coniugale + numero associati in partecipazione che apportano lavoro prevalentemente nell'impresa

- numero addetti = Numero dipendenti a tempo pieno + numero dipendenti a tempo parziale + numero apprendisti + numero assunti con contratto di formazione e lavoro o a termine e lavoranti a domicilio + numero collaboratori coordinati e continuativi che prestano attività prevalentemente nell'impresa + numero associati in partecipazione che apportano lavoro prevalentemente nell'impresa + numero soci con occupazione prevalente nell'impresa + numero amministratori non soci

- resa del capitale = ricavi/valore dei beni strumentali.

Per ogni gruppo omogeneo è stata calcolata la distribuzione ventile di ciascuno degli indicatori precedentemente definiti e poi sono state selezionate le imprese che presentavano valori degli indicatori contemporaneamente all'interno di un determinato intervallo per costituire il campione di riferimento.

Per tutti i cluster è stato utilizzato l'intervallo compreso fra l'estremo superiore del 4° ventile e l'estremo superiore del 19° ventile. Tale intervallo è valido solo per la resa oraria. Per la resa del capitale sono state selezionate le

imprese che presentano un valore della resa del capitale al di sotto dell'estremo superiore del 18° ventile.

Così definito il campione di imprese di riferimento, si è proceduto alla definizione della “funzione di ricavo” per ciascun gruppo omogeneo.

Per la determinazione della “funzione di ricavo” sono state utilizzate sia variabili contabili (quadro M del questionario) sia variabili strutturali. La scelta delle variabili significative è stata effettuata con il metodo stepwise. Una volta selezionate le variabili, la determinazione della “funzione di ricavo” si è ottenuta applicando il metodo dei minimi quadrati generalizzati, che consente di controllare l'eventuale presenza di variabilità legata a fattori dimensionali (eteroschedasticità).

Affinché il modello di regressione non risentisse degli effetti derivanti da soggetti anomali (outliers), sono stati esclusi tutti coloro che presentavano un valore dei residui (R di Student) al di fuori dell'intervallo compreso tra i valori -2,5 e +2,5.

Nell'allegato 2.A vengono riportate le variabili ed i rispettivi coefficienti della “funzione di ricavo”.

## **2. APPLICAZIONE DEGLI STUDI DI SETTORE ALL'UNIVERSO DEI CONTRIBUENTI**

Per la determinazione del ricavo della singola impresa sono previste due fasi:

- l'*Analisi Discriminante*<sup>5</sup>;
- la stima del ricavo di riferimento.

---

<sup>5</sup> L'Analisi Discriminante è una tecnica che consente di associare ogni impresa ad uno dei gruppi omogenei individuati per la sua attività, attraverso la definizione di una probabilità di appartenenza a ciascuno dei gruppi stessi.

Nell'allegato 2.B vengono riportate le variabili strutturali risultate significative nell'Analisi Discriminante.

Non si è proceduto nel modo standard di operare dell'Analisi Discriminante in cui si attribuisce univocamente un contribuente al gruppo di massima probabilità; infatti, a parte il caso in cui la distribuzione di probabilità si concentri totalmente su di un unico gruppo omogeneo, sono considerate sempre le probabilità di appartenenza a ciascuno dei gruppi omogenei.

Per ogni impresa viene determinato il ricavo di riferimento puntuale ed il relativo intervallo di confidenza.

Tale ricavo è dato dalla media dei ricavi di riferimento di ogni gruppo omogeneo, calcolati come somma dei prodotti fra i coefficienti del gruppo stesso e le variabili dell'impresa, ponderata con le relative probabilità di appartenenza.

Anche l'intervallo di confidenza è ottenuto come media degli intervalli di confidenza, al livello del 99,99%, per ogni gruppo omogeneo ponderata con le relative probabilità di appartenenza.

## **ALLEGATO 2.A**

**Variabili e coefficienti della funzione di ricavo**

**COEFFICIENTI DELLE FUNZIONI DI RICAVO****SG61F**

<b>VARIABILI</b>	<b>CLUSTER 1</b>	<b>CLUSTER 2</b>	<b>CLUSTER 3</b>	<b>CLUSTER 4</b>
Costo del venduto	0,5003	-	-	0,9723
Costo per la produzione di servizi	0,8911	1,7691	1,4969	1,0158
Spese per acquisti di servizi	1,9788	1,2165	1,6150	1,5155
Valore dei beni strumentali	0,6986	0,4807	0,2345	-
Logaritmo in base 10 del valore dei beni strumentali	6.975,8189	7.898,6490	8.295,2714	10.562,1056
Spese per lavoro dipendente e per altre prestazioni diverse da lavoro dipendente	0,9937	0,8273	1,2929	1,1341
Soci e associati in partecipazione con occupazione prevalente (numero)	61.673,9361	44.093,7624	38.451,1522	78.656,9613
Collaboratori dell'impresa familiare e coniuge dell'azienda coniugale (numero)	35.947,5858	-	34.307,6804	28.842,0556

- Le variabili contabili vanno espresse in migliaia di lire.

- Il logaritmo in base 10 è calcolato per i soli valori maggiori di zero della variabile cui si riferisce.

## **ALLEGATO 2.B**

### **Variabili dell'analisi discriminante**

**Quadro A:**

- Numero giornate retribuite dipendenti a tempo pieno

**Quadro B:**

- Mq deposito destinati esclusivamente all'attività
- Mq esposizione destinati esclusivamente all'attività

**Quadro F:**

- Volume delle provvigioni derivanti da grossisti
- Volume delle provvigioni derivanti da agenti
- Volume delle provvigioni derivanti da servizi
- Volume totale delle provvigioni
- Tipologia della clientela: commercio all'ingrosso
- Tipologia della clientela: media/grande distribuzione
- Tipologia della clientela: commercio al dettaglio
- Tipologia della clientela: consumatori privati

**Quadro G:**

- Settore merceologico prevalente<sup>6</sup>
- Primo settore merceologico secondario<sup>7</sup>
- Secondo settore merceologico secondario<sup>8</sup>
- Terzo settore merceologico secondario<sup>9</sup>
- Spese di viaggio

---

<sup>6, 7, 8, 9</sup> Indicare i codici della tabella allegata al questionario SG61, corrispondenti al settore merceologico trattato

- Spese di pubblicità, propaganda e rappresentanza