



XII Riunione scientifica
POLITICA FISCALE, FLESSIBILITÀ DEI MERCATI E CRESCITA
Pavia, Collegio Ghislieri 6 - 7 ottobre 2000

DETERMINANTI DELL'EVASIONE DELL'IRPEG

F. Di Nicola - A. Santoro

Secit, Ministero delle finanze, sezione studi

Società italiana di economia pubblica

Dipartimento di economia pubblica e territoriale – Università di Pavia

DETERMINANTI DELL'EVASIONE DELL'IRPEG

F. Di Nicola - A. Santoro¹

Abstract

The purpose of this paper is to give, by means of multivariate statistical analysis, an accurate picture of the *propensity to evade* IRPEG, the Italian tax on profits earned by resident limited-liability companies (SPA's and SRL's).

The analysis is based on a comprehensive dataset including data about 450 Italian companies which form a representative sample of the limited-liability companies operating in Italy in 1997. This dataset includes, along with data about fiscal liabilities and balance sheets of the sampled companies, the outcomes of the audits conducted on the 450 companies forming the sample.

We construct an indicator of the propensity to evade which measures the evaded portion of the effective income, while taking into account the different situations in terms of fiscal liability. This indicator is calculated for every firm, including those whose evasion does not directly yield the immediate avoidance of the IRPEG tax, and therefore its estimates do not represent a measure of the tax-gap. The statistical analysis is conducted using different tools, such as the analysis of means within subgroups, the classification trees by the two-way analysis of variance and the analysis of multiple correspondences.

These different approaches all indicate the importance of certain factors in the evasion decision. Among these factors are the following: 1) the dimension of the firm, which seems to be inversely related to the propensity to evade; 2) some financial indicators (such as capital/debt ratio, directly related to propensity to evade); 3) some economic indicators (such as the share of stock in hand costs on total cost), which seem to be able to act as signals of peculiar kinds of evasion; 4) other

¹ Secit, Ministero delle finanze, sezione studi. E-mail: fdinicola@iol.it ; asantoro@finanze.it. Nell'ambito di un'impostazione comune, i paragrafi 1, 2, 4 e le considerazioni conclusive sono stati redatti congiuntamente, mentre il paragrafo 3 è stato redatto da A. Santoro ed il paragrafo 5 da F. Di Nicola. Sebbene prevalentemente di natura tecnica, le opinioni espresse in questo scritto sono attribuibili esclusivamente agli autori. Questa versione del lavoro ha beneficiato degli utili commenti di alcuni colleghi del SECIT e di alcuni partecipanti alla XIII conferenza annuale della SIEP; eventuali errori restano ovviamente responsabilità esclusiva degli autori.

elements directly related to evasion, as the recent constitution of the firm and its location in the south of Italy, which suggest the importance of institutional and social factors in the decision to evade.

On the basis of these results authors suggest some change in Italian strategy of audits and fiscal policy.

JEL classification: Tax Evasion (Compliance, Collection) (H260)

Keywords: tax evasion, audits, corporate taxation, fiscal policy.

INDICE

Abstract.....	1
1) Introduzione.....	3
2) Presentazione dei dati e degli indicatori scelti	6
2.1) L'indicatore dell'evasione relativa all'imponibile effettivo e le variabili descrittive	8
3) L'evasione osservata attraverso le medie incrociate.	12
4) L'evasione osservata attraverso l'analisi di segmentazione.....	21
4.1) I risultati ottenuti con la segmentazione ad albero	22
5) L'evasione osservata con l'analisi delle corrispondenze multiple.....	30
5.1) I risultati.....	31
Considerazioni conclusive	38
Bibliografia	43
Appendice 1: L'algoritmo "CHAID esaustivo" di segmentazione ad albero	45
Appendice 2: Richiami sull'analisi delle corrispondenze multiple	48

1) Introduzione

In una fase in cui appare importante migliorare efficienza ed efficacia delle strategie anti-evasione ai fini di una riduzione della pressione fiscale, o perlomeno di un'attenuazione del carico da aliquote legali reso possibile dall'allargamento delle basi imponibili e/o della platea di contribuenti, ci è parso utile valorizzare per quanto possibile la rara disponibilità di informazioni sulle società di capitali, riguardanti sia dati di bilancio e fiscali, sia l'evasione contestata-constatata nel corso di verifiche generali.

Questo lavoro costituisce perciò un approfondimento di uno studio effettuato dal Secit nel 1999², su incarico del Ministro delle Finanze, allo scopo di identificare e quantificare i diversi fattori (compresa l'evasione) che determinano il gettito dell'IRPEG.

Con un flow chart riportato in una recente pubblicazione³, si è tentato di rendere visibili la presenza ed il peso delle varie componenti che dal risultato economico "effettivo" portano all'imponibile fiscale. Tali componenti sono suddivise in due categorie:

- fattori che intervengono a monte del risultato economico dichiarato in bilancio, ossia l'evasione e l'elusione in senso stretto, qui intesa come la sola sottrazione di base imponibile riconducibile a tassazione secondo il dettato dell'art. 37 bis del DPR n.600/73⁴;
- fattori che intervengono a valle del risultato economico dichiarato in bilancio, ossia le variazioni fiscali, ulteriormente distinte a seconda della loro *ratio*: antielusiva, erosiva, di riporto delle perdite, di riallocazione temporale delle componenti positive o negative, ad evitare le doppie tassazioni, o di altro tipo.

L'approfondimento qui proposto riguarda la sola evasione intesa in senso lato, ovvero comprensiva dell'elusione come sopra specificata.

² "La tassazione delle società di capitali", di F.Di Nicola, S.Golino, V.Rinaldi (coordinatore), R.Stajano, A.Santoro.

³ Cfr Di Nicola F. et al. (2000).

⁴ Si esclude perciò il più generale concetto di elusione non equiparabile ad evasione e perciò non riportabile a tassazione a seguito di una verifica.

Più precisamente, l'obiettivo del presente lavoro è una migliore identificazione dei fattori aventi maggiore rilievo ai fini della descrizione della “**propensione all'evasione**”, come definita nel paragrafo 3.

Va da subito sottolineato che il tentativo si presenta piuttosto arduo, data la natura, di per sé sfuggente, del fenomeno evasivo. Tenendo conto di ciò, in questo lavoro si è tentato di attenuare l'indeterminatezza del fenomeno considerando diversi strumenti di analisi multivariata, aventi caratteristiche ed obiettivi tra loro diversi. Inoltre, per le stesse ragioni, si è evitato di utilizzare approcci limitanti o eccessivamente condizionati dalle convinzioni del ricercatore, quali ad es. la regressione lineare multivariata⁵.

Un significativo *caveat* riguarda il fatto che l'evasione qui considerata è quella contestata dalla Guardia di Finanza e dagli Uffici delle Entrate (come verrà spiegato nel prossimo paragrafo), che non necessariamente coincide con l'evasione poi definita a seguito della procedura innescata dall'atto di accertamento. Questo fenomeno di sovrastima (di cui si è tenuto conto in altra sede, dovendo quantificare l'evasione a livello aggregato) non può essere corretto a livello di microdati, in quanto al momento non è possibile avere informazioni circa gli esiti delle procedure innescate dagli accertamenti, che si definiscono spesso dopo diversi anni.

D'altronde, l'aver considerato l'evasione contestata ha portato ad ignorare anche l'evasione non intercettata dalle verifiche generali, portando ad una sottostima di alcune particolari tipologie evasive (quali, ad esempio, la falsa fatturazione⁶ o forme elusive sofisticate). Dato che le due approssimazioni sono di segno opposto e di non lontano ordine di grandezza, l'errore compiuto nelle elaborazioni si riduce in media, anche se determina distorsioni nelle analisi di specifici sottogruppi.

⁵ Va del resto considerato che in questi modelli (sia *à la* Allingham-Sandmo sia basati sulla teoria dei giochi) si assume che siano normalmente diverse le aliquote di prelievo, il che non si verifica nel caso dell'IRPEG; cfr. Andreoni et al. (1998), Bordinon e Zanardi (1997).

⁶ Fenomeno divenuto oggi di più difficile verifica a seguito della recente abolizione dell'elenco clienti e fornitori.

In sintesi dunque, l'analisi qui presentata presenta le seguenti peculiarità:

- si avvale di microdati recenti, riguardanti le dichiarazioni ed i bilanci risalenti all'anno di imposta 1997, per i quali il campionamento è stato di tipo rappresentativo⁷ ed il riporto all'universo è avvenuto sulla base della struttura di dati censuari forniti dalla Anagrafe Tributaria-SOGEI;
- presenta, per la prima volta a quanto è dato di sapere a chi scrive, un tentativo di descrivere a livello microeconomico l'evasione "contestata" dell'IRPEG alla luce di fattori sia di natura meramente descrittiva (quali la natura giuridica o la collocazione geografica della società) sia di natura economico-patrimoniale;
- non si basa su ipotesi restrittive concernenti la razionalità del comportamento dell'evasore, se non per aspetti marginali, puntando ad individuare determinanti del fenomeno evasivo e non alla conferma di modelli micro-economici predefiniti.

Gli strumenti analitici qui utilizzati sono, oltre all'analisi univariata e bivariata di medie e frequenze, l'analisi di segmentazione (o di "classificazione ad albero") e l'analisi delle corrispondenze multiple, quest'ultima corredata da un'analisi dei gruppi.

Ciascuno di essi si basa su fondamenti statistici a cui in questa sede si faranno solo brevi cenni, fornendo al lettore interessato più specifici rimandi nelle apposite appendici; ciò che merita di essere sottolineato è che queste tipologie di analisi sono fra loro simili, ma con contributi conoscitivi differenziati ed integrabili.

Il lavoro è organizzato come segue. Nel paragrafo 2 verranno descritti i dati utilizzati per l'analisi, nonché il metodo seguito per la ponderazione all'universo. Nel paragrafo 3 verranno descritti l'indicatore dell'evasione e le altre variabili utilizzate nelle diverse analisi. Nel paragrafo 4 si presenteranno, anche con finalità di descrizione dei dati, alcune tavole univariate e bivariate particolarmente significative, costruite considerando la relazione tra l'indicatore dell'evasione e ciascuna delle altre variabili (analisi bivariata). Nel paragrafo 5 verranno esposti e commentati i risultati ottenuti con l'analisi di segmentazione (per la metodologia si veda l'appendice 1). Nel paragrafo 6 si presentano i risultati dell'analisi delle corrispondenze multiple e della particolare analisi dei gruppi adottata (per i richiami metodologici si veda l'appendice 2).

⁷ Si pensi ad es. ai notevoli problemi di distorsione delle stime derivate dal campione Mediobanca delle principali imprese italiane o dai controlli effettuati correntemente dall'Amministrazione Finanziaria.

2) Presentazione dei dati e degli indicatori scelti

I dati qui raccolti si riferiscono, in origine, ad un campione rappresentativo di 500 società di capitali aventi fini di lucro, selezionate per rappresentare l'universo delle circa 500mila persone giuridiche dello stesso tipo effettivamente operanti in Italia e risultanti in Anagrafe Tributaria, ovvero di tutte le società di capitali esistenti con esclusione degli intermediari finanziari, delle cooperative, delle società formalmente e sostanzialmente cessate e di quelle interessate da procedure concorsuali⁸.

Per l'identificazione del campione rappresentativo, secondo la tecnica dell'estrazione casuale stratificata, sono stati utilizzati i seguenti criteri di stratificazione: il settore di attività (comprese le fondamentali attività terziarie), la dimensione per classe di ricavi e l'area geografica di collocazione della sede giuridica. Vale la pena di sottolineare il carattere rappresentativo del campione, sia per la sua migliore capacità di estensione dei risultati all'universo di riferimento, sia per la sua attitudine a minimizzare, *coeteris paribus*, il rischio di riproduzione delle convinzioni degli analisti.

Per ogni impresa del campione e con riferimento all'anno di imposta 1997 sono stati raccolti, controllati ed archiviati quattro tipi di dati:

- a) alcune informazioni di base raccolte mediante questionario strutturato dai verificatori;
- b) i dati di bilancio (di fonte CERVED) per l'anno in questione;
- c) le dichiarazioni dei redditi (mod.760/98);
- d) le quote di imponibile di cui è stata contestata l'evasione o l'elusione dalla Guardia di Finanza o dal Dipartimento delle Entrate a seguito di verifiche generali appositamente svolte.

Durante queste verifiche generali, inoltre, sono stati raccolti dati ulteriori concernenti alcune poste di bilancio (ad esempio i debiti verso soci) o elementi descrittivi della situazione patrimoniale (quali il grado di concentrazione delle quote) non rilevabili attraverso il bilancio o tramite il mod.760.

Dopo aver verificato l'appartenenza delle società campionate all'universo definito ed eliminato le società non utilizzabili, sono residue 449 società di capitali; i dati di queste imprese sono stati

⁸ L'universo non escludeva quindi le cosiddette società non operative, gli "evasori totali" (intesi qui nella particolare accezione di soggetti con partita IVA ma senza alcuna dichiarazione dei redditi), ed infine le società in ordinaria liquidazione.

quindi riportati all'universo, al fine sia di effettuare elaborazioni che correggessero le distorsioni del campione *ex post*⁹ rispetto alla struttura censuaria delle tre variabili note, sia di riportare all'universo gli ammontari complessivi e le numerosità, per facilità di lettura e confronto con aggregati nazionali di altra fonte.

In termini formali, noti i pesi W_{ijk} della popolazione di riferimento¹⁰ (di numerosità $N=488'000$, con i, j, k contrassegnanti le celle come individuate dalle modalità delle tre variabili di stratificazione), ed osservati i pesi w_{ijk} delle stesse variabili nel campione disponibile (di numerosità $n=449$), il coefficiente R_{ijk} di peso e riporto è stato calcolato come segue, per ogni terna i, j, k :

$$R_{ijk} = (W_{ijk} / w_{ijk}) \times (N / n)$$

dove la riponderazione è costituita dal primo fattore, mentre il secondo rappresenta semplicemente un coefficiente di riporto all'universo, cioè una costante che riporta i 449 casi osservati ai 488 mila di riferimento.

La fase dell'archiviazione è servita anche ad effettuare controlli di coerenza tra i dati riportati su questionario, quelli presenti nei 760 e quelli di bilancio.

In particolare, sono stati effettuati controlli automatici e documentali (Una specie di liquidazione delle dichiarazioni) volti ad appurare:

- la coerenza delle voci del conto economico riportato su 760;
- la coerenza tra indebitamento da 760 e quello classificato in forme diverse nel questionario, all'occorrenza osservando anche il bilancio;
- la coerenza tra utile civilistico netto e riporto nel quadro apposito del 760;
- la correttezza del riporto del reddito o perdita da fine quadro RA (variazioni) ad inizio quadro RG (calcolo dell'imponibile e dell'imposta).
- in qualche caso, il calcolo del risultato economico in bilancio, quando questo era riportato correttamente nel 760, ma era mal calcolato nei dati del bilancio.

In alcuni casi si sono evidenziati errori o comportamenti evasivi (per lo più di lieve entità) del contribuente nel riclassificare le voci del conto economico in bilancio per riportarle nel quadro RB; in qualche caso il riporto danneggiava il contribuente stesso. Nel complesso è parso evidente che la

⁹ Al netto cioè delle mancate rilevazioni e di un sovracampionamento operato sulle società di elevata dimensione.

ricompilazione del conto economico in dichiarazione è un'operazione che viene svolta superficialmente, tenendo d'occhio soprattutto il corretto riporto delle imposte d'esercizio e l'utile civilistico netto. Quando è stato possibile, sono state imputate le voci ricalcolate senza che ciò inficiasse l'analisi del percorso tra risultato economico ed imposta dovuta.

Sebbene il tentativo sia stato quello di non escludere casi per non perdere informazioni, in un caso bilancio e conto economico sono risultati talmente incomprensibili ed incoerenti da consigliare l'esclusione della società dalle elaborazioni. Infine, il prolungarsi dei tempi per la conclusione del controllo di una grande società per la quale erano in corso da parte della GDF approfondimenti in merito ad un sospetto caso di evasione mediante trasferimento di fondi infruttiferi all'estero finanziati attraverso indebitamento in Italia ha indotto ad esaminare con più attenzione l'aspetto controverso. La soluzione prescelta è stata quella di imputare provvisoriamente e cautelativamente una probabile contestazione, ritenuta dal gruppo di lavoro Secit fondata, piuttosto che distorcere le successive analisi escludendo l'impresa in questione.

2.1) L'indicatore dell'evasione relativa all'imponibile effettivo e le variabili descrittive

In questo lavoro la variabile obiettivo scelta è un indicatore di evasione relativa, avente natura di variabile scalare, che chiameremo *indeval*. Nella definizione di questo indicatore si è tentato di misurare l'evasione in termini relativi come rapporto di composizione tra evasione ed imponibile "effettivo", ovvero, in linea di principio, l'imponibile che sarebbe stato dichiarato in assenza di evasione.

Indeval misura quindi, in prima definizione, il rapporto percentuale tra l'evasione rilevata (contestata-constatata dai verificatori) e l'imponibile effettivo. Si è scelto di ragionare in questi termini in quanto l'ammontare di evasione (ovvero l'evasione assoluta) non tiene conto delle diverse situazioni di reddito ai fini fiscali, né della diversa dimensione aziendale, e non può perciò fornire indicazioni circa la propensione ad evadere.

Quanto al concetto di imponibile effettivo, cioè al denominatore del rapporto individuato come variabile da spiegare, esso è stato definito come la somma tra l'imponibile dichiarato nel 760 (anche

¹⁰ Società di capitale con esclusione del credito e delle assicurazioni, delle cooperative, e dei soggetti sottoposti alle procedure fallimentari concorsuali.

detto imponibile fiscale) e quello occultato a seguito dei comportamenti evasivi (come quantificati dai verificatori).

L'indicatore così definito, se concettualmente ci è parso idoneo a misurare l'intensità del fenomeno evasivo in termini relativi, ha comportato problemi di indeterminatezza in alcune situazioni limite. Infatti questo rapporto, che idealmente dovrebbe variare tra zero (in assenza di evasione) e 100% (nel caso in cui l'evasione consente di annullare l'intero imponibile effettivo), è in primo luogo indeterminato se il denominatore è nullo in quanto evasione ed imponibile sono entrambi nulli; in questo caso, comunque, l'assenza di evasione consente senza dubbi di approssimare anche l'indicatore a zero.

Se esiste evasione bisogna invece distinguere altri due casi particolari. Se l'imponibile effettivo è positivo ma inferiore all'evasione (cioè l'imponibile dichiarato risulta negativo, ma inferiore in valore assoluto all'evasione), il contribuente realizza un'evasione maggiore rispetto a quella che sarebbe stata necessaria per evitare il pagamento dell'Irpeg con riferimento al 1997. In termini analitici, l'indicatore supererebbe il 100% e tenderebbe all'infinito al tendere a zero dell'imponibile effettivo.

Se invece l'imponibile dichiarato è negativo e superiore in valore assoluto all'evasione, mantenendo la formulazione dell'indicatore enunciata in precedenza anche l'imponibile effettivo, e conseguentemente $indeval$, risulterebbero negativi, con $indeval$ sostanzialmente privo della capacità informativa per la quale era stato pensato¹¹.

Questi due ultimi casi presentano alcune similitudini, ma anche rilevanti differenze.

La similitudine sta nel fatto che in entrambi i casi l'evasione è più di quella necessaria all'azzeramento dell'imponibile effettivo nel periodo di imposta considerato; poiché nel nostro sistema fiscale non è prevista l'imposta negativa, siamo apparentemente di fronte ad un comportamento irrazionale del contribuente. Tuttavia, anche in questi casi il comportamento del contribuente può, a ben vedere, essere frutto di un calcolo razionale. Va ricordata infatti la possibilità, riconosciuta alle persone giuridiche, di riportare le perdite risultanti ai fini fiscali per i cinque anni successivi; inoltre l'eccedenza di evasione di IRPEG, in un determinato anno d'imposta, potrebbe essere funzionale all'evasione di altre imposte, quali IRAP ed IVA. Infine, senza considerare altri periodi d'imposta o altre imposte, il comportamento in questione potrebbe

¹¹ Infatti, in questo caso bisognerebbe ricorrere ai valori assoluti di $indeval$ per rendere lo stesso un indicatore della propensione all'evasione. In questo modo si creerebbe, però, un'ambigua similitudine con i casi in cui l'evasione comporta un effettivo risparmio d'imposta.

rispondere ai dettami della *bounded rationality*, e cioè che la minimizzazione dell'imposta deriva dalla soluzione di un problema di ottimizzazione vincolato dalla conoscenza di una serie di fattori relativi alla situazione della singola impresa e alla normativa fiscale ad essa applicabile.

L'obiettivo dell'analisi è quello di ottenere un quadro della propensione ad evadere l'Irpeg in un determinato periodo d'imposta. Per questa ragione, abbiamo deciso di non considerare nell'indicatore quella parte di evasione che riteniamo si riferisca ad altri periodi d'imposta o ad altre imposte. Per quel che riguarda il primo caso particolare (evasione superiore in valore assoluto all'imponibile dichiarato (negativo)) queste premesse ci portano a porre *indeval* pari a 100.

Nel secondo caso, invece, caratterizzato da un'evasione inferiore in valore assoluto all'imponibile dichiarato (negativo) ci è sembrato potenzialmente fuorviante ritenere che tutta l'evasione realizzata fosse imputabile ad evasione futura o di altre imposte. Se avessimo seguito quest'ultima strada, *indeval* sarebbe stato nullo in questi casi, il che appare come una perdita secca di informazione. Si è trattato perciò di stimare la parte di evasione non riferibile al futuro o ad altra imposta, e riconducibile a ragioni diverse. Abbiamo effettuato tale stima attraverso l'analisi delle corrispondenze, che si presta a "quantificare" le modalità di variabili categoriche sulla base del posizionamento sul piano fattoriale, a sua volta determinato dall'interazione con tutte le altre informazioni considerate¹². Classificate le imprese in sottogruppi omogenei rispetto al comportamento evasivo (*indeval* trasformata in una variabile categorica con le seguenti modalità: "no evasione", "evasione tra 1 e 50%", "evasione tra 51 e 99%", "evasione al 100%", "evasori totali", "evasione senza risparmio di imposta Irpeg"), ed avendo osservato la proiezione di quest'ultimo sottogruppo sul piano fattoriale delle corrispondenze in posizione intermedia tra alta e bassa evasione, si è ritenuto di poter assegnare a questa particolare categoria di imprese un tasso di evasione pari al 50%.

Infine, nel caso degli evasori totali, ossia delle poche imprese campionate che non risultavano avere presentato né la dichiarazione né il bilancio e alle quali è stata sempre contestata una qualche evasione, il valore dell'indicatore è stato posto senza dubbi pari a 100.

In termini formali *indeval* è stato quindi definito come segue:

$$\text{INDEVA1} = \text{EVAELU} * 100 / \text{IMPEFF} \quad \text{se } \text{EVAELU} > 0 \quad \text{IMPEFF} > 0 \text{ ed } \text{IMPDICH} \geq 0$$

¹² Analisi chiamata per questo motivo anche "delle variabili latenti". Per maggiori riferimenti metodologici si rimanda ai successivi paragrafi ed appendici, nonché alla bibliografia.

INDEVA1=100

se EVAELU>0 IMPEFF >0 ed IMPDICH<0

INDEVA1=100

se EVASORI TOTALI

INDEVA1=50

se EVAELU>0 e IMPEFF<0

INDEVA1=0

se EVAELU=0.

dove:

EVAELU = somma dell'evasione e dell'elusione in senso stretto contestata;

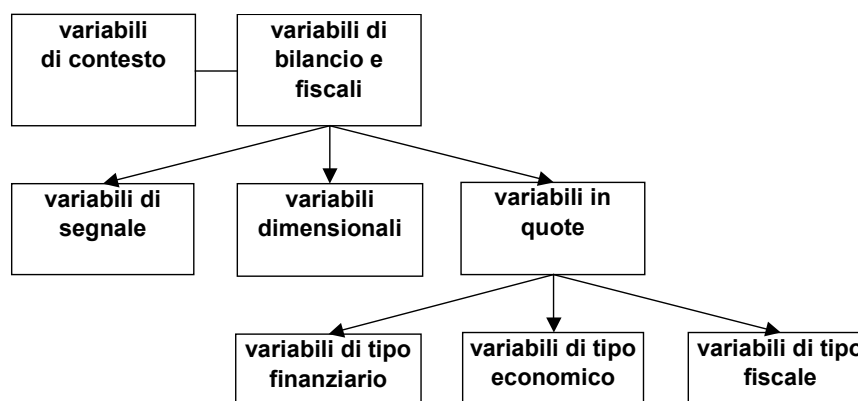
IMPEFF = imponibile effettivo = IMPDICH + EVAELU;

IMPDICH = imponibile dichiarato.

Va ribadito che *indeva1* ha la finalità di misurare il tasso di evasione per singola impresa, e quindi la sua media non può essere paragonata a stime macroeconomiche dell'evasione complessiva, ottenute attraverso il rapporto tra aggregati. Ne consegue che, anche a seguito delle attribuzioni di valore appena presentate, il tasso medio di evasione non rappresenta una stima della perdita di gettito originata dall'evasione¹³.

Per quanto concerne le altre variabili utilizzate, esse verranno descritte contestualmente alla presentazione dei risultati delle diverse analisi. La figura sottostante ne presenta un quadro generale.

Figura1: variabili descrittive



¹³ Del resto, vale la pena di ricordare che anche nel citato rapporto del SECIT *non* venivano presentate nuove stime macroeconomiche dell'evasione IRPEG, ma ci si riferiva invece al lavoro di Bernardi e Bernasconi (1996).

Per ***variabili di contesto*** si intendono quelle che indicano alcune caratteristiche di base dell'impresa, quali la sua natura giuridica, l'area geografica di collocazione della sede principale, il settore produttivo di appartenenza.

Le ***variabili di segnale*** sono variabili dicotomiche che segnalano la presenza di taluni fenomeni, quali, tra gli altri, l'appartenenza ad un gruppo di società, l'esistenza di perdite riportate ad esercizi successivi, la presenza di debiti verso soci.

Le ***variabili dimensionali*** sono variabili correlate alla dimensione assoluta dell'impresa: il valore della produzione, il totale dei ricavi e del patrimonio netto, la base imponibile ai fini IRAP. Tali variabili non sono state utilizzate in forma continua, ma invece categorizzata tenendo conto di alcuni valori rilevanti ai fini giuridici (per esempio, 1 miliardo è la soglia dei ricavi per la contabilità semplificata, 20 milioni che è il capitale sociale minimo per la costituzione di una SRL, 200 milioni che è quello per la SPA) ed anche della loro distribuzione nel campione riponderato.

Le ***variabili in quote*** sono un insieme di variabili, di natura diversa, accomunate dal fatto di essere originariamente formulate come quote di un totale al fine di segnalare l'intensità di determinati fenomeni. Più precisamente, le variabili *di tipo finanziario o di stock* sono quelle che esprimono la quota del patrimonio netto e dei debiti sul totale del passivo, nonché la ripartizione dell'attivo tra circolante ed immobilizzazioni; le variabili *di tipo economico o di flusso* sono quote che esprimono l'incidenza di talune componenti negative (positive) di reddito sul totale delle componenti negative (positive) di reddito; le variabili *di tipo fiscale* sono quote che esprimono l'incidenza delle variazioni fiscali sul risultato imponibile. Per tutte queste variabili sono state create apposite classi o categorie sulla base della numerosità osservata dal campione riponderato nel tentativo di evidenziare, ove possibile, tre modalità: un'incidenza particolarmente bassa, un'incidenza intermedia ed una particolarmente alta, identificate da soglie di quantile attorno al 25% e 75%.

3) L'evasione osservata attraverso le medie incrociate.

Un primo livello di analisi, di immediata comprensione, è consentito dalle tabelle a doppia entrata che presentano le medie di *indeval* rispetto a ciascuna delle variabili descrittive, opportunamente categorizzate. E' utile fornire innanzitutto un quadro generale:

Quadro generale: media di *indeval* per evasori e non evasori

Presenza evasione	Numero casi (% su totale)	Media <i>indeval</i>	Deviaz. standard
società con evasione	324.516 (66,5%)	62,8	34,5
società senza evasione	163.293 (33,5%)	0	0
Totale società	487.809 (100%)	41,8	40,9

Come si vede, le società che presentano evasione rappresentano i 2/3 del totale e tra esse la media dell'indicatore sale al 63% circa.

Riportiamo di seguito alcune delle tabelle più significative circa la relazione tra *indeval* e le variabili descrittive, iniziando da alcune variabili di contesto.

Tabella 1: media di *indeval* per natura giuridica

Natura giuridica	Media <i>indeval</i>	Numero casi	Deviaz.stand.
Srl	42,4	455'996	41,2
Spa	32,4	31'813	33,8
Totale	41,8	487'809	40,8

Tabella 2: media di *indeval* per anno di costituzione

Anno costituzione	Media <i>indeval</i>	Numero casi	Deviaz.stand.
Società vecchia	39,0	309'091	33,6
Società nuova	46,5	178'179	44,0
Totale	41,8	487'809	40,8

Tabella 3: media di *indeval* per area geografica

Area geografica	Media <i>indeval</i>	Numero casi	Deviaz.stand.
Nord-ovest	36,3	181'652	37,9
Nord-est	34,8	93'080	38,9
Centro	46,2	135'148	42,6
Sud	55,0	77'929	42,34
Totale	41,8	487'809	40,8

Tabella 4: media di *indeva 1* per settore attività

Settore di attività	Media <i>indeva 1</i>	Numero casi	Deviaz.stand.
Industria	42,3	120'741	38,7
Costruzioni	32	66'834	40,6
Commercio	52,8	142'492	41,2
Servizi	35,5	157'743	39,9
Totale	41,8	487'809	40,8

In generale, pare di poter dire che alcune variabili di contesto sembrano poter contribuire in misura piuttosto rilevante alla descrizione del fenomeno evasivo.

In relazione alla natura giuridica, va sottolineato come la differenza tra l'evasione media delle Srl e quella della Spa sembrerebbe essere un primo sintomo del fatto, su cui si tornerà in seguito, che la propensione all'evasione diminuisce all'aumentare delle dimensioni dell'impresa e del suo patrimonio.

Una simile constatazione potrebbe valere anche per le tabelle 3 e 4, tenendo conto della prevalenza di imprese di piccole dimensioni al sud e nel settore del commercio, anche se in questi casi la correlazione indiretta con la dimensione e patrimonializzazione è molto meno marcata, denotando quindi specifici nessi.

La tabella 4 riporta anche alcuni risultati inattesi, in particolare la bassa percentuale media di evasione nel settore delle costruzioni, normalmente considerato uno dei comparti a maggiore propensione evasiva, ma come emerge dall'analisi delle corrispondenze e dalla successiva spiegazione dei gruppi, il settore delle costruzioni sembra connotarsi per due modelli opposti di comportamento evasivo.

Per quanto concerne i servizi, l'alta varianza dipende dal fatto che la media dell'indicatore dell'evasione è estremamente differenziata all'interno del comparto: hanno valori ben superiori al 40%, con punte oltre il 70% tutta una serie di servizi quali istruzione, trasporti e sanità privata, mentre sulla media complessiva del comparto pesa il valore del 25% per i cosiddetti altri servizi, che comprendono, tra gli altri, molte attività professionali (il codice 74 ATECO91).

In ogni caso, già da queste prime osservazioni sembrerebbe opportuna una maggiore e più mirata articolazione per territorio e per settore dell'azione anti-evasiva svolta dall'amministrazione finanziaria.

Per quanto riguarda la tabella 2, va specificato che sono state considerate di vecchia costituzione le società costituite in anni antecedenti al 1993, mentre dal 1993 in avanti la società è stata considerata nuova. Le medie delle due categorie sembrano essere particolarmente stabili, con coefficienti di variazione ben inferiori all'unità; la spiegazione della differenza del 7% nella propensione all'evasione potrebbe ricondursi sia al fatto che società più nuove sono relativamente meno solide da un punto di vista economico e patrimoniale, sia al fatto che l'azione anti-evasiva non è sufficientemente dissuasiva nei confronti di chi entra sul mercato. Un'ulteriore possibile spiegazione, approfondita nel corso della successiva analisi di segmentazione, è la seguente: alcune società, specie nel centro-sud, vengono costituite allo scopo di evadere o di accedere a sostanziose agevolazioni, con l'intenzione di chiudere dopo poco tempo, ed è dunque più probabile che nei confronti di una società relativamente nuova e poco responsabilizzata vengano contestate evasioni di imponibile.

Tra le variabili cosiddette di segnale, alcuni risultati interessanti si ottengono considerando le dicotomiche che si riferiscano all'appartenenza a gruppi e all'esistenza di perdite riportate.

Tabella 5: media di *indeva 1* per appartenenza a gruppi

Appartenenza a gruppi	Media <i>indeva 1</i>	Numero casi	Deviaz.stand.
Si gruppi	29,1	72'772	36,6
No gruppi	44	415'037	41,1
Totale	41,8	487'809	40,8

Tabella 6: media di *indeva 1* per quantità perdite riportate

Perdite riportate	Media <i>indeva 1</i>	Numero casi	Deviaz.stand.
Non disp.(evasori totali)	100	15'196	0
Si riporto perdite	45,4	96'142	39,1
No riporto perdite	38,5	376'472	40,2
Totale	41,8	487'809	40,8

Riguardo alla tabella 5, va ricordato che anche nel nostro paese, dove pur non esiste una fiscalità diretta di gruppo, l'appartenenza ai gruppi consente di realizzare più agevolmente la pianificazione fiscale. Tale pianificazione può avvenire in molti modi, tra cui, in particolare, il trasferimento di contratti, la stipula di accordi di finanziamento interno, la cessione di rami d'azienda, ecc.. Si può perciò pensare che la maggiore propensione all'evasione da parte delle società non facenti parte di gruppi dipenda dall'impossibilità, per queste società, di ricorrere a tali strumenti elusivi. Se questa linea di spiegazione è valida allora il riporto delle perdite, sicuramente più agevole all'interno dei gruppi, non può essere considerato alla stregua di uno strumento di pianificazione fiscale (come pur sarebbe possibile pensare) in quanto la tabella 6 indica che l'evasione aumenta tra le società che riportano delle perdite. Piuttosto, l'aumento dell'evasione per le società che riportano perdite potrebbe dipendere dalla situazione di obiettiva maggiore difficoltà economica in cui queste società si trovano.

L'analisi delle più importanti variabili dimensionali (rispetto alla descrizione della propensione all'evasione) prende invece spunto dalle tabelle 7 e 8.

Tabella 7: media di *indeva1* per classi di patrimonio netto

Classi di patrimonio netto (valori limite)	Media <i>indeva 1</i>	Numero casi	Deviaz.stand.
Evasori totali	100	15'196	0
Nulla o negativo	42,1	48'824	43
1-250 mln	43	297'822	41,6
250 mln-2 mld	30,5	92'886	33,2
Più di 2 mld	34,5	33'081	35,2
Totale	41,8	487'809	40,8

Tabella 8: media di indeval per classi di valore della produzione

Classi di valore della produz. (valori limite)	Media <i>indeval</i>	Numero casi	Deviaz.stand.
<=100mln	32,6	158'263	43
101 mln-1 mld	50,9	159'889	41,6
1 mld-10 mld	43,4	138'080	33,2
10 mld-50 mld	35,4	28'165	35,2
Oltre 50 mld	23,8	3'412	29,2
Totale	41,8	487'809	40,8

Dalla lettura congiunta delle tabelle 7 e 8 emerge con una certa nettezza l'esistenza di una relazione inversa tra la dimensione dell'impresa, misurata dal suo patrimonio netto e dal suo valore della produzione, e la propensione all'evasione. L'eccezione riguarda le imprese con valore della produzione inferiore a 100 mln, ma alcune di queste sono società non operative e che quindi non producono alcun valore aggiunto potenzialmente occultabile. La lettura di questo dato può prestarsi a più di un'interpretazione, ed è comunque approfondito nel paragrafo sull'analisi delle corrispondenze.

In termini astratti, la decisione di evadere è presa più velocemente ed a minori costi in un'impresa di minori dimensioni. Infatti, queste imprese sono normalmente controllate da pochi soggetti le cui decisioni non sono messe in discussione né dagli azionisti né dagli organi di controllo.

In secondo luogo, in una grande impresa la corretta tenuta delle scritture contabili è lo strumento primario per il controllo interno: l'evasione risulta spesso controproducente dal punto di vista stesso dell'impresa perché impedisce al principale (i soci ovvero il livello superiore del *management*) di controllare l'agente (il livello superiore ovvero quello inferiore del *management*)¹⁴.

In terzo luogo, non va dimenticato che l'evasione oggetto di rilevamento in questo lavoro è l'evasione contestata: la grande impresa è normalmente in grado di realizzare l'evasione attraverso procedimenti ben più raffinati rispetto all'impresa di minori dimensioni, e perciò meno visibili tramite le normali verifiche.

¹⁴ Cfr. Lupi (1995), in particolare il cap.6.

Vera la relazione individuata, dal punto di vista delle politiche di accertamento la maggiore propensione evasiva delle piccole imprese suggerisce di guardare con particolare attenzione al concreto utilizzo degli studi di settore, che sono applicabili solamente alle società con un fatturato inferiore a 10 miliardi. E' necessario evitare che la "catastizzazione" dei ricavi e dei redditi inibisca eccessivamente l'ordinaria attività di controllo nei confronti di un insieme di soggetti che mostrano un'alta propensione evasiva.

Se è quindi certo che abbia rilievo l'entità del patrimonio, meno rilevante appare il ruolo del grado di concentrazione del capitale sociale: nel caso tale grado sia minoritario l'evasione è pari al 42%, mentre nel caso opposto si scende al 41%. Questo è un dato in qualche misura sorprendente: poiché in presenza di una pluralità di soci la decisione di evadere è più costosa (in termini di tempo per completare il necessario *bargaining process*), ci si aspetterebbe una correlazione positiva tra grado di concentrazione ed evasione. E' pur vero che anche in questo caso la compresenza di ulteriori comportamenti differenziati all'interno dello stesso sottogruppo potrebbe spiegare il fenomeno.

Per quanto riguarda, infine, le variabili originariamente formulate in quote, è opportuno analizzare le medie di *indeval* rispetto alle categorie appositamente definite mediante ricodifica delle quote, secondo i criteri esposti nel precedente paragrafo, in modo da identificare anche le relazioni non lineari.

Tra quelle di tipo finanziario, la più rilevante sembra essere la quota del patrimonio netto sul totale delle passività, la cui relazione con l'evasione è descritta nella tabella 9.

Tabella 9: media di *indeval* per classi della quota di patrimonio netto su passivo

Classi di quota del patrimonio netto sul passivo	Media <i>indeval</i>	Numero casi	Deviaz.stand.
Patrim.negativo	47,4	49907	43,8
Quota bassa	44,3	102178	42,8
Quota intermedia	41,4	204434	38,6
Quota alta	32,3	92886	33,2
Totale	40,4	473696	40,2

La tabella 9 lascia intuire l'esistenza di una forte relazione inversa tra quota del patrimonio netto ed evasione: le società relativamente più patrimonializzate, ovvero con una minore incidenza dei debiti, sono presumibilmente più stabili e meno bisognose di ricorrere all'evasione. Inoltre, anche in questo caso, il risultato potrebbe essere legato alla dimensione, in quanto è plausibile che il grado di patrimonializzazione aumenti all'aumentare della dimensione. A questo livello di analisi, non sembrano avere un particolare rilievo, viceversa, le variabili che esprimono le quote del passivo per debiti complessivi e neppure quelle che indicano la ripartizione dell'attivo tra circolante ed immobilizzazioni. Nei prossimi paragrafi vedremo che l'analisi di segmentazione e quella delle corrispondenze multiple consentono di avere qualche informazione in più al riguardo.

Un discorso a parte meritano i debiti verso soci: la media di *indeval* è superiore al 46% quando la quota appartiene alla categoria intermedia, mentre scende al 40% quando la quota è bassa e addirittura al 35% quando la quota è alta. Una possibile spiegazione di questo andamento "ad U" è la seguente: l'alta quota dei debiti verso soci è legata ad una situazione di evasione solo quando i debiti sono contratti solo formalmente allo scopo di giustificare un quadro contabile conseguente a ricavi non dichiarati. In effetti, un controllo incrociato a tre variabili fa rilevare che la media di *indeval* arriva quasi al 58% quando la quota dei debiti verso soci è alta ed è alta anche la quota di costi per magazzino. Quest'ultimo dato è spesso indicativo, come vedremo tra breve, di un'alta evasione.

Per le variabili di tipo economico, va rilevato che l'andamento della media di *indeval* non sembra essere correlato in modo lineare con la maggior parte delle quote di costo: la media di *indeval* presenta un valore massimo nelle categorie intermedie per gli ammortamenti, per la quota di costi del personale e per quella riferita al costo dei beni e servizi in godimento. Non sembra avere rilievo la presenza o meno di accantonamenti per svalutazioni di poste dell'attivo.

Qualcosa di più si può dire in relazione alla quota dei costi per il magazzino e a quella per oneri finanziari: sono significative al proposito le tabelle 10 e 11.

Tab. 10: media di *indeval* per classi della quota di costi di magazzino sul totale dei costi

Classi di quota dei costi di magazzino sul totale costi	Media <i>indeval</i>	Numero casi	Deviaz. stand.
Quota bassa	31,6	135'992	40,7
Quota intermedia	42,5	212'507	39
Quota alta	46,7	123'299	40,1
Totale	40,4	471'798	40,2

Tabella 11: media di *indeval* per classi della quota di oneri finanziari sul totale dei costi

Classi di quota degli oneri finanziari sul totale costi	Media <i>indeval</i>	Numero casi	Deviaz. stand.
Quota bassa	35	114'381	40,2
Quota intermedia	41,8	225'561	38,7
Quota alta	41,4	128'762	41,8
Totale	40	468'705	40

Per quanto concerne i costi di magazzino, in questo lavoro sono definiti come la somma dei costi relativi alle materie prime necessarie per la produzione e delle variazioni di rimanenze di materie prime e merci. In termini astratti, la significatività di questa variabile potrebbe essere in parte connessa a talune fattispecie tipiche di evasione quali le seguenti:

- utilizzazione scorretta dei criteri di valutazione delle scorte per la diminuzione dell'imponibile dovuto (evasione di costi);
- omessa fatturazione di vendite derivanti dalla trasformazione delle materie prime in prodotti finiti (evasione di ricavi da omessa fatturazione o sottofatturazione).

Un'analisi della tipologia evasiva contestata alle imprese con alti costi di magazzino sembra indicare la possibilità che entrambe le ipotesi formulate siano state riscontrate in quanto sono simili i valori medi dell'evasione da ricavi e da costi. Tali ipotesi, del resto, sono piuttosto grossolane e costituiscono violazioni facilmente riscontrabili in sede di verifica.

Tra la quota degli oneri finanziari e l'indicatore della propensione evasiva sembra esistere una relazione diretta, e ciò è congruente con un quadro in cui l'evasione è legata ad una situazione di sofferenza aziendale; in questo senso si potrebbe stabilire una similitudine con il ruolo giocato dal riporto delle perdite.

In merito alle variazioni fiscali, va segnalato che in presenza di variazioni erosive, la media dell'evasione è quasi doppia rispetto al caso opposto (42,2% contro 24,3%); ciò potrebbe indicare una maggiore capacità, da parte degli evasori, di sfruttare tutte le possibilità offerte dal sistema fiscale per la diminuzione del carico fiscale complessivo.

Riassumendo gli esiti principali dell'analisi delle medie incrociate possiamo dire che:

- emerge in primo luogo il ruolo dei fattori dimensionali, nel senso che la propensione all'evasione sembra essere decisamente più alta nelle imprese di minore dimensione, se si accetta di considerare il valore della produzione e il patrimonio come buone *proxies* della dimensione;
- presumibilmente correlate alle variabili dimensionali sono alcune variabili che descrivono il modello organizzativo prescelto, quali il tipo sociale o l'appartenenza a gruppi. Modelli organizzativi dove è più semplice prendere la decisione di evadere ovvero ricorrere ad alternative efficaci all'evasione sembrano essere prevalenti.
- emerge anche l'esistenza di un'evasione presente in situazioni di sofferenza, ovvero realizzata da imprese con alti oneri finanziari, bassa patrimonializzazione e perdite riportate da periodi precedenti;
- il ruolo di alcune variabili di contesto, quali il settore o l'area geografica, sembra preminente rispetto a quello delle variabili di bilancio, ed in particolare delle variabili di costo, il cui andamento non è leggibile a questo primo livello di segmentazione se non in relazione ad alcune situazioni-limite (come le operazioni evasive che si riflettono sulle quote di costi di magazzino). Da quest'ultima considerazione si origina l'esigenza di procedere ad un'analisi più approfondita.

4) L'evasione osservata attraverso l'analisi di segmentazione

In questo paragrafo si presenta l'analisi di segmentazione (o classificazione ad albero), condotta tramite l'algoritmo detto CHAID esaustivo, disponibile nel software SPSS-Answer Tree¹⁵. La segmentazione è una metodologia di identificazione a più fasi di sottogruppi (o “rami”) che massimizzano le differenze di una variabile dipendente individuando per ogni suddivisione, mediante un'analisi multivariata della varianza, i migliori predittori tra tutti quelli presi in considerazione.

¹⁵ Per un più formale richiamo del metodo si rimanda all'apposita appendice, contenente anche i richiami bibliografici.

Per questa ragione, può essere utilizzata in alternativa alla regressione multivariata “stepwise” (lineare o logistica)¹⁶. Rispetto all'analisi delle corrispondenze multiple e a quella di *cluster*, invece, la segmentazione, nella forma qui utilizzata, è più indicata per identificare le situazioni-limite del fenomeno oggetto di analisi (qui la propensione ad evadere) per due ragioni:

- mette direttamente in relazione una variabile dipendente con tutte le altre;
- effettua l'analisi non solo per la totalità dei dati, ma anche all'interno di sottoinsiemi particolarmente significativi.

Nel nostro caso, quindi, era plausibile che la segmentazione fosse più efficacemente in grado di indicare i sottogruppi con medie di *indeval* molto superiori o molto inferiori a quella generale.

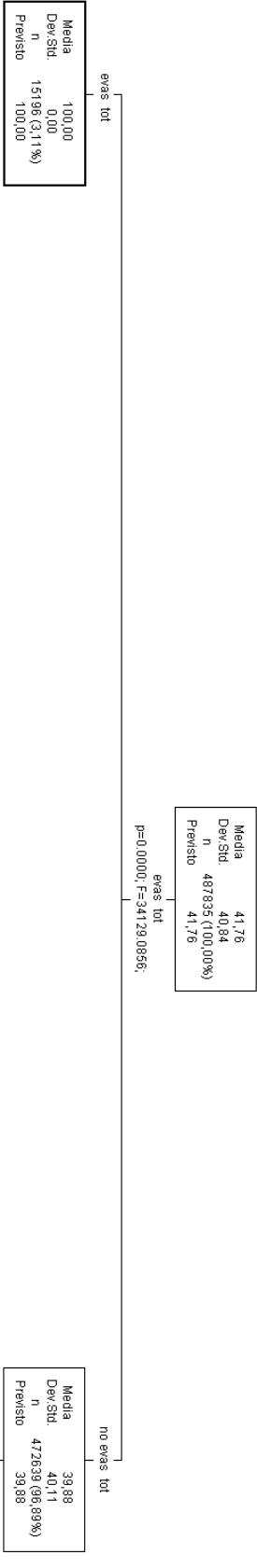
4.1) I risultati ottenuti con la segmentazione ad albero

L'albero ottenuto, limitatamente ai primi 5 livelli di segmentazione, è riportato nelle figure sottostanti.

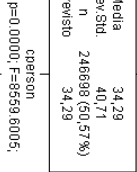
I risultati della segmentazione si prestano sia ad una lettura *verticale*, volta ad identificare i sottogruppi in cui la media della variabile obiettivo assume valori particolarmente alti o bassi, sia ad una lettura *orizzontale*, che consente di definire il ruolo di un certo predittore nei confronti della variabile obiettivo all'interno di un determinato sottogruppo. Il primo livello di segmentazione può essere ignorato, in quanto distingue semplicemente gli evasori totali (caratterizzati da variabili con valore prefissato a -2) dagli altri soggetti.

¹⁶ Dalla quale differisce peraltro in maniera sostanziale, in quanto non richiede di formulare particolari ipotesi circa le relazioni tra le variabili indipendenti, che nella segmentazione si chiamano predittori, né di quantificare effetti medi vevolevoli sull'intero set di dati.

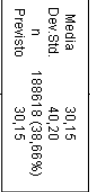
INDEVA1



qservizi bassa,qservizi alta,<mancante>

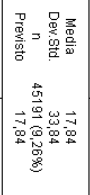


no dipendenti,qperson alta,<mancante>

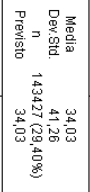


patrimet.

nulla o neg.,250 mln-2 mld.;più di 2 mld.

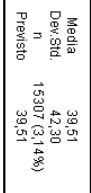


1-250 mln.

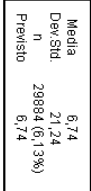


ciò che deb.

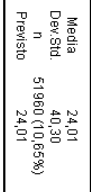
industria agric.,costituz.



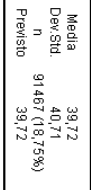
com.,servizi



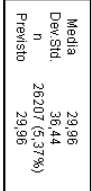
quota bassa,quota alta



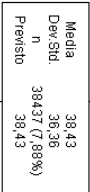
quota medio-alta



[2,83,9014]

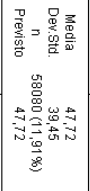


qmaggaz.

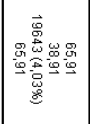


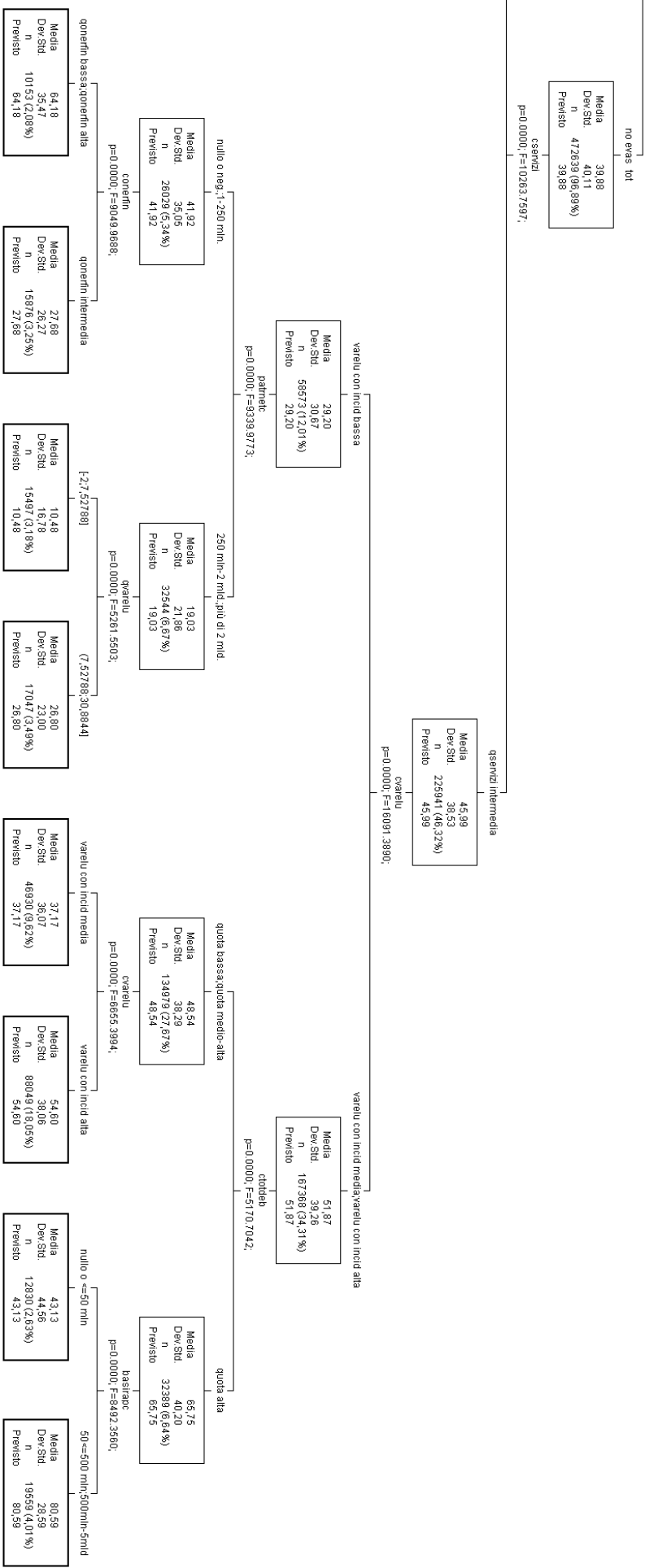
soc old

annocost



soc new





Nella lettura verticale dei risultati si evidenzia la maggior stabilità dei sottogruppi caratterizzati da valori di *indeval* particolarmente alti nei confronti di quella dei sottogruppi caratterizzati da valori di *indeval* particolarmente bassi. Rispetto alla media generale di *indeval* tra le società diverse dagli evasori non totali, pari al 39,9%, si evidenzia un sottogruppo di casi con media di *indeval* particolarmente elevata avente le seguenti caratteristiche:

quota di costi del personale intermedia, incidenza alta delle variazioni anti-elusive rispetto all'imponibile fiscale, quota alta di debiti sul totale delle passività, base imponibile irap compresa tra i 50 milioni e i 5 miliardi.

Questo gruppo rappresenta il 4% dei casi totali, ha una media di evasione pari all'80,6% (cioè più che doppia rispetto alla media generale tra i non evasori totali), l'estremo inferiore dell'intervallo di confidenza, con probabilità al 95%, è pari al 67,3%, l'estremo superiore è pari al 93,8%. Il dato pare quindi sufficientemente stabile, anche se la sua interpretazione non è agevole.

Al primo livello di segmentazione compare la variabile della quota dei costi per i servizi acquisiti dall'esterno e per i beni concessi in godimento: la maggiore propensione all'evasione è associata alle imprese che fanno registrare una quota media, mentre una minore propensione è legata alle società con quota bassa od alta. Trattandosi del primo livello di segmentazione tra le società diverse dagli evasori totali, il risultato corrisponde a quello evidenziato al termine del paragrafo 4: l'evasione sembra avere una relazione non lineare con questa quota di costi.

Considerando che nella categoria dei costi per i servizi e i beni da terzi rientrano i canoni di leasing e le collaborazioni coordinate o continuative od occasionali, ci si sarebbe aspettati che ad un innalzamento della quota corrispondesse una maggiore propensione all'evasione. Ciò in quanto la deducibilità dei canoni di leasing è spesso contestata dall'A.F quando si ricollega a contratti di finanziamento improprio, ovvero i cosiddetti accordi di *sale and lease back*. Inoltre, l'intenso ricorso alle collaborazioni dovrebbe segnalare una struttura d'impresa meno stabile e quindi una maggiore attitudine all'evasione. I dati non danno riscontro a queste ipotesi, sebbene vada segnalato che la variabile in questione definisca due gruppi tra i quali la media di *indeval* non differisce di tanto, ponendosi tra il 34,3% e il 46%, rispetto ad una media generale (per i non evasori totali) pari al 39,9%. Rimane comunque il dato di fatto che la quota di costi per servizi e beni in godimento è la più significativa (nel senso descritto nel paragrafo precedente) e che la sua relazione con l'indicatore di evasione da noi scelto non ci pare spiegabile agevolmente.

Una possibile soluzione di questo apparente paradosso potrebbe derivare dalla distinzione delle voci di costo effettivamente comprese nella quota considerata, ottenibile attraverso un collegamento

sistematico tra il database dell' A.T. e quello della Cerved, che pure non risolverebbe il problema di distinguere le collaborazioni dagli altri acquisti di servizi, anche professionali¹⁷.

Si presenta più agevole la spiegazione delle variabili di segmentazione successive. Una maggior quota di variazioni anti-elusive indica una maggior incidenza delle norme immesse nel sistema fiscale per contrastare le più grossolane manovre di sottrazione dell'imponibile, tra cui, ad esempio, la non corresponsione nell'esercizio di compensi liquidati agli amministratori, l'utilizzo promiscuo di telefoni od auto aziendali, eccetera. La relazione tra questa variabile e l'evasione contestata (*indeval* raggiunge il 52%) sembra indicare che la società che realizza l'evasione è anche pronta ad utilizzare tutti i mezzi elusivi consentiti dall'ordinamento fiscale.

Quando alla quota intermedia di costi per servizi e beni in godimento si associa, oltre ad una quota alta di variazioni anti-elusive, anche un'alta quota di debiti sul totale del passivo, l'indicatore della propensione all'evasione raggiunge quasi i due terzi dell'imponibile potenziale. La presenza di un'alta quota di debiti in questo specifico segmento è il fenomeno speculare di una bassa quota del patrimonio netto che, come abbiamo visto nel paragrafo precedente, per la generalità delle imprese è già legata in generale ad una maggiore propensione evasiva. La relazione logica tra le due variabili si ritrova quindi in questo sottoinsieme, che rappresenta il 6,7% dei casi totali.

L'indebitamento elevato può essere associato all'evasione, sia quando si tratta di un indebitamento fittizio volto a coprire ricavi non dichiarati, sia quando si tratta di un indebitamento reale che indica un'oggettiva difficoltà economica (e quindi anche finanziaria) della società che ricorre quasi per necessità all'evasione.

Il quinto ed ultimo livello di segmentazione sembra indicare che il maggior grado di evasione è associato all'indebitamento oneroso, piuttosto che a quello verso fornitori. La media dell'evasione arriva infatti all'80,6% quando la base imponibile dell'Irap è pari ad almeno 50 milioni, e a tale base contribuiscono in modo decisivo gli oneri finanziari. Più precisamente, tali oneri rappresentano il 30% della base imponibile IRAP in questo sottogruppo e solo il 17% a livello aggregato.

Riassumendo, la segmentazione identifica un gruppo di società con evasione media pari ai 4/5 dell'imponibile potenziale in relazione a cui i dati appaiono sufficientemente affidabili. Queste società, presumibilmente, tentano di occultare l'evasione realizzata attraverso un indebitamento oneroso fittizio, che genera oneri finanziari corrispondenti i quali contribuiscono ad abbattere

¹⁷ È per motivi di trasparenza e leggibilità dei bilanci, oltre che per maggiore omogeneità con le voci necessarie a costruire la base imponibile Irap, che appare opportuno agli scriventi procedere ad un'evoluzione delle voci classificatorie del bilancio in grado di distinguere le collaborazioni dall'acquisto di servizi.

ulteriormente l'imponibile ai fini IRPEG. E' pur vero che tali oneri sono compresi nella base imponibile IRAP, ma poiché tale base imponibile include anche l'utile, l'evasione dello stesso ai fini IRPEG deprime anche l'imponibile IRAP.

Come già detto, risultano meno affidabili le stime riferite ai sottogruppi con media di *indeval* più bassa. Tra questi, ne emergono due aventi le seguenti caratteristiche:

- società con variazioni antielusive inferiori al 7,5% dell'imponibile potenziale e con patrimonio netto pari almeno a 250 milioni. La media di *indeval* stimata per queste società, che rappresentano solo il 3,18% delle società totali, è pari al 10,5%, ma l'intervallo di confidenza, con probabilità al 95%, va dal 2,6% al 18,7% (gruppo *low 1*).
- società con una quota di costi per il personale nulla o alta, con patrimonio netto almeno pari a 250 milioni ovvero negativo ed operante nel settore del commercio o dei servizi. Questo sottogruppo rappresenta il 6,13% delle società di capitali italiane e ha una media di *indeval* pari al 6,74%; l'intervallo di confidenza andrebbe dallo 0% ad oltre il 14% (gruppo *low 2*).

Volendo per un attimo accantonare i dubbi concernenti la stabilità dei dati, va sottolineato che gruppo *low 1* ha quelle caratteristiche di stabilità patrimoniale e di minore propensione all'utilizzo delle "scappatoie" offerte dall'ordinamento fiscale la cui relazione inversa con l'evasione è già stata riscontrata e commentata.

Per quanto riguarda il gruppo *low 2* con i valori inferiori di *indeval*, esso ha delle caratteristiche di non facile interpretazione.

In particolare, non è semplice comprendere il ruolo della quota dei costi per il personale, che compaiono al terzo livello di segmentazione, evidenziando l'andamento non lineare già ricordato in precedenza, cioè segnalando una propensione all'evasione minore nei casi di quota bassa o alta dei costi per il personale, e maggiore nel caso di quota intermedia.

Si può supporre, infatti, che la quota di costi per il personale sul totale dei costi aumenti all'aumentare della stabilità dell'impresa, ed era lecito attendersi che al suo crescere si realizzasse una diminuzione dell'evasione contestata. Invece, l'algoritmo di segmentazione accorpa il sottoinsieme con quota bassa e di quota alta dei costi del personale, con una media di *indeval* pari al 30,15%. La particolarità della quota bassa di costi del personale è che essa è pari, in effetti, ad una quota nulla: il dato campionario, del tutto clamoroso, è che oltre il 42% delle società di capitali italiane non indicano, nel prospetto di bilancio riportato nel 760, alcun costo per il personale dipendente. Questo dato non pare essere frutto né di un errore campionario né di errori nella

dichiarazione: abbiamo avuto la possibilità di ottenere un riscontro del tutto simile analizzando un campione di quasi 5000 bilanci CERVED, anch'essi selezionati con criteri di rappresentatività.

Va ricordato che le società in questione sono quelle caratterizzate da una quota di costi per servizi o per godimento di beni di terzi bassa o alta. Tale correlazione suggerisce alcune possibili spiegazioni della relazione tra propensione evasiva e quota di costi per il personale:

- che possa esservi un'evasione o elusione strutturale di tipo contributivo che, in qualche misura, sostituisce l'evasione di imponibile fiscale. Più esplicitamente, per alcune aziende con risultato economico nullo o esiguo può risultare conveniente il totale occultamento di rapporti di lavoro dipendente effettivamente esistenti, anche se ciò impedisce la deduzione dei costi connessi, in misura tale da rendere superfluo il ricorso all'evasione.
- Per altre aziende, il risparmio contributivo che deriva dal mascheramento di tale tipo di rapporti attraverso un ricorso solo formale a contratti di collaborazione può rappresentare un sostituto sufficiente dell'evasione delle imposte dirette. Ciò spiegherebbe perché la media di *indeval* si abbassa quando la quota di costi del personale (dipendente) si azzerà; d'altronde, *indeval* si abbassa nuovamente quando la quota di costi del personale si eleva tra le società che effettivamente presentano quel maggior grado di stabilità a cui si faceva cenno in precedenza.
- Alcuni commercianti o artigiani potrebbero aver scelto di operare attraverso la schermatura di una società di capitali, ma di fatto con una condizione individuale o di impresa familiare (questo spiegherebbe una parte del fenomeno per le Srl, ma certo non il 16% di Spa senza dipendenti).

Ovviamente, quanto detto rappresenta solo un abbozzo di spiegazione di un fenomeno che si mantiene, per qualche verso, difficilmente interpretabile; per la verifica della tesi formulata sarebbe anche in questo caso importante poter ottenere un maggior dettaglio della voce dei costi per godimento di beni e servizi attualmente riportata nel 760.

Per quanto riguarda il quarto livello di segmentazione in *low 2*, va segnalato che, tra le società con quota di costi del personale bassa o alta, la presenza di un patrimonio netto nullo o negativo si associa ad una media di *indeval* bassa, contrariamente a ciò che accade per il complesso delle imprese. Un discorso simile vale per l'ultimo livello di segmentazione di questo gruppo: tra le

società che ne fanno parte e che appartengono al commercio, la propensione all'evasione è particolarmente bassa, contrariamente a quanto accade per l'insieme delle imprese di questo settore.

La lettura *orizzontale* dell'albero di segmentazione ottenuto porta inoltre a formulare le seguenti considerazioni, tenendo presente anche l'analisi delle tabelle a doppia entrata condotta nel paragrafo precedente:

- la maggior propensione all'evasione delle società di nuova costituzione è confermata e si segnala come particolarmente forte tra le società con quota dei costi del personale intermedia (ossia, secondo l'interpretazione data in precedenza, che tendenzialmente prediligono l'evasione IRPEG a quella contributiva e che non hanno una struttura occupazionale dominata dal lavoro dipendente). In questo sottoinsieme, infatti, le società costituite dal 1993 in avanti fanno registrare una media di *indeva1* vicina al 66%, mentre quelle costituite in precedenza si fermano al 38,5%.
- un ulteriore ed importante elemento interpretativo di questa differente propensione all'evasione tra società di nuova e vecchia costituzione è consentito dalla disaggregazione a livello di area geografica, riportata nella tabella 12.

Tabella 12: media di *indeva1* per area geografica e per anno di costituzione della società

area geografica	Anno di cost.	<i>indeva1</i>	Numero casi	Deviaz. stand.
Nord-ovest	Soc.vecchia	36,9	127'235	36,4
	Soc.nuova	34,9	54'417	41,4
Nord-est	Soc.vecchia	37,9	52'592	37,7
	Soc.nuova	30,9	40'488	40,1
Centro	Soc.vecchia	39,3	90'425	41,0
	Soc.nuova	60,2	44'723	42,3
Sud	Soc.vecchia	47,0	38'838	39,9
	Soc.nuova	63,0	39'091	43,2
Totale	Soc.vecchia	39,0	309'091	38,6
	Soc.nuova	46,5	178'719	44,1
	Totale società	41,8	487'809	40,8

- per quel che riguarda le società di vecchia costituzione sembra evidenziarsi il ricorso a forme di evasione appunto tradizionali, quali quelle attuate attraverso il magazzino:

quando la quota di costi per il magazzino è superiore all'80% la media di indeval, tra le società di vecchia costituzione e con quota dei costi del personale intermedia, raggiunge il 56,6%, mentre per la categoria residuale (con costi di magazzino fino all'80% dei costi totali) è di circa il 30%.

- la quota di costo per oneri finanziari si conferma un buon indicatore della propensione all'evasione: anche quando i casi esaminati rientrano in un gruppo dalle caratteristiche alquanto eterogenee (variazioni antielusive con incidenza bassa ma anche patrimonio netto basso) la quota alta di tali oneri è associato ad un alto livello di *indeval*. Scorporando infatti il dato della cella relativa, si ottiene che, malgrado la quota alta e quella bassa siano accorpate, alla prima corrisponde in effetti una media di oltre il 70% di *indeval*, mentre nella seconda *indeval* è pari al 55%. In questo sottogruppo, peraltro, si perde la relazione diretta tra quota di costo per oneri finanziari e *indeval*, in quanto quando la quota è intermedia *indeval* scende a meno del 30%; il risultato va però tarato alla luce della bassa numerosità dei casi compresi nella cella a cui ci si riferisce.

In conclusione, possiamo affermare che l'analisi di segmentazione conferma in linea di massima quanto ricavato con l'analisi delle medie incrociate, aggiungendovi la possibilità di identificare, con un certo grado di sicurezza, che società ad alta propensione evasiva sono tendenzialmente quelle che ricorrono all'indebitamento oneroso in misura particolarmente elevata.

Merita invece un approfondimento il ruolo della quota dei costi per il personale dipendente, che sembrerebbe poter segnalare una sorta di elasticità di sostituzione tra l'evasione di imponibile ai fini delle imposte dirette e l'elusione contributiva (ricorso ai collaboratori anziché ai dipendenti).

5) L'evasione osservata con l'analisi delle corrispondenze multiple

Insieme alle principali statistiche descrittive bivariate ed all'identificazione di sottogruppi di imprese massimizzanti le differenze dei tassi di evasione, ci è parso opportuno seguire anche la strada di una esplorazione più generale, che non ponesse in diretta relazione una variabile con tutte le altre ma consentisse uno sguardo d'insieme sulle affinità e diversità tra modalità. Si è scelto così di effettuare un'analisi multivariata delle "corrispondenze multiple" (integrata da una particolare forma di *cluster analysis* che ci ha consentito di quantificare il peso delle aree omogenee di evasione identificate).

L'analisi delle corrispondenze consiste essenzialmente in una proiezione grafica ottimale delle modalità prese in esame (nel nostro caso quelle in cui sono state classificate le citate variabili di bilancio e di base delle società) su un piano scelto in modo tale che modalità vicine sono tendenzialmente possedute dalle stesse unità (le imprese del campione)¹⁸.

Tale analisi può perciò essere utilizzata come un'estensione multivariata dello studio delle interdipendenze statistiche bivariate (sarebbe estremamente difficile e laborioso cogliere i nessi tra molte variabili attraverso tabelle di contingenza multiple), oppure come un mezzo alternativo alle interpolazioni per cogliere relazioni tra caratteristiche non modellabili facilmente con curve teoriche (limitazione presente nella maggior parte dei casi), o infine come strumento per l'emersione di fattori "latenti" altrimenti non identificati.

5.1) I risultati

I risultati possono essere osservati principalmente sul grafico rappresentante il primo piano delle corrispondenze (cioè il piano identificato dalle prime due variabili fattoriali), che è riportato nella pagina che segue.

Partendo dalle aree meglio definite e circoscritte, quella in alto a destra identifica certamente un'area omogenea caratterizzata dalle grandi aziende (patrimonio netto, valore della produzione, risultato economico e valore aggiunto¹⁹ massimi).

¹⁸ Per la metodologia si rimanda all'appendice n.2, contenente anche i richiami bibliografici.

¹⁹ Quest'ultimo approssimato dalla base imponibile Irap.

Le associazioni più significative, cioè le modalità proiettate nelle vicinanze di quest'area, sono quelle dell'essere appartenenti a gruppi, cioè controllati o controllanti come società, della presenza di accantonamenti d'esercizio, di alta quota dei costi per personale dipendente, ma soprattutto da tasso di evasione inferiore al 50%. L'appartenenza a gruppi appare essere più significativa, ai fini di una bassa evasione, di quanto appaiono in generale essere la concentrazione del capitale nelle mani del socio di maggioranza e la concentrazione familiare del capitale sociale; evidentemente l'azione di controllo richiede con più forza un bilancio abbastanza veritiero.

Nelle vicinanze di quest'area si posizionano il settore industriale, un patrimonio netto intermedio, e, significativamente, la tipologia di prevalente evasione da costi. Quest'ultima variabile, che tenta di distinguere tra evasione da costi o ricavi, indica in maniera chiarissima (piano delle corrispondenze, cluster, incroci) che l'evasione più alta si associa ad occultamento di ricavi, mentre il gonfiamento di costi è una pratica più di tipo "elusivo", e perciò più adatto alle grandi imprese, che emerge e viene registrata come evasione solo per la parte contestata in sede di verifica. In altre parole, ed è un risultato importante perché verificato sul campo, l'evasione è essenzialmente di ricavi, salvo il fenomeno penalmente rilevante delle false fatturazioni, del resto qui non adeguatamente riscontrabile.

Dall'altra parte, nella sinistra del piano, emerge una seconda area che pare caratterizzarsi per poca o nessuna attività: valore della produzione minimo o nullo, nessun ammortamento, nessun dipendente, bassa quota di debiti, ricavi e magazzino. Come modalità caratterizzanti quest'area, che potrebbe essere definita delle società di comodo o non operative, si trovano in posizione contigua l'assenza di evasione ed una quota dei debiti verso soci elevata.

Questa tipologia sembra segnalare il fatto che esiste un'area di società di comodo caratterizzate probabilmente dalla volontà di schermatura nei riguardi del fisco, ma non interessata direttamente ad evadere; tale area è sostanzialmente non operativa, e tra le caratteristiche maggiormente presenti emergono il totale controllo azionario ed il settore di attività terziario (associazioni del tutto comprensibili).

Un'altra area di omogeneità, contigua a quella delle società di comodo ma soprattutto vicina al baricentro del gruppo caratterizzato da bassa evasione, presenta indicatori di patrimonializzazione ed investimento: quota alta di patrimonio netto sul totale delle passività, ammortamenti alti (in rapporto al complesso dei costi), immobilizzazioni alte, circolante basso, oneri finanziari bassi, quota costi alta per acquisto di servizi (collaborazioni comprese) e godimento beni di terzi. L'estrema vicinanza tra la modalità "nessuna evasione", oneri finanziari bassi e acquisto di servizi

alto fa pensare che la tendenza all'esternalizzazione di attività ed un basso indebitamento spingano meno a comportamenti evasivi, in linea con quanto già osservato nelle analisi precedenti.

In quest'area tendono a posizionarsi anche le imprese con totale controllo azionario, il settore dei servizi, le imprese con variazioni basse in rapporto all'imponibile ed all'interno di esse con bassa quota di variazioni da noi definite "antielusive".

La presenza del settore servizi in quest'area può essere indicativa solo di una tendenza, stante l'estrema varietà di comportamenti, ed allo stesso tempo può associarsi anche al fatto che nel settore dei servizi gli strumenti di controllo fiscale hanno minore efficacia.

Interessante appare il ruolo delle variazioni complessive e di quelle antielusive, che sembrano essere correlate al tasso di evasione; il dato, confermato anche dalle medie incrociate di gruppo (riportate nelle tre tabelle che seguono), può essere spiegato come una tendenza di chi evade ad utilizzare anche gli strumenti contigui dell'elusione, di cui queste variazioni sono un indicatore indiretto di presenza.

CVARIAZ qvariaz quota variazioni in classi		CLUSTER a 3 gruppi			Totale
		1 evas bassa	2 evas media	3 evas alta	
evasori totali	numero			15'196	15'196
	% colonna			10.3	3.1
1 variaz.tot. <0	numero	66'295	17'944	34'088	118'327
	% colonna	26.0	20.9	23.1	24.3
2 variaz.tot. medio-basse	numero	81'007	19'912	20'224	121'143
	% colonna	31.8	23.2	13.7	24.8
3 variaz.tot >40%	numero	107'406	47'842	77'896	233'144
	% colonna	42.2	55.8	52.8	47.8
Totale	numero	254'708	85'698	147'404	487'810
	% colonna	100	100	100	100

CVARELU qvarelu quota variaz antielusive in classi		CLUSTER a 3 gruppi			Totale
		1 evas bassa	2 evas media	3 evas alta	
evasori totali	numero			15'196	15'196
	% colonna			10.3	3.1
1 varelu con incid bassa	numero	108'030	24'731	31'731	164'492
	% colonna	42.4	28.9	21.5	33.7
2 varelu con incid media	numero	50'095	14'049	27'176	91'320
	% colonna	19.7	16.4	18.4	18.7
3 varelu con incid alta	numero	96'582	46'919	73'300	216'801
	% colonna	37.9	54.7	49.7	44.4
Totale	numero	254'707	85'699	147'403	487'809
	% colonna	100	100	100	100

CVARERO qvarero quota variaz erosive in classi		CLUSTER a 3 gruppi			Totale
		1 evas bassa	2 evas media	3 evas alta	
evasori totali	numero			15'196	15'196
	% colonna			10.3	3.1
1 varero presenti	numero	245'948	83'829	130'762	460'539
	% colonna	96.6	97.8	88.7	94.4
2 varero nulle	numero	8'760	1'869	1'445	12'074
	% colonna	3.4	2.2	1.0	2.5
Totale	numero	254'708	85'698	147'403	487'809
	% colonna	100	100	100	100

Nella zona centrale tendono a cadere o i punti-modalità non correlati agli altri fenomeni, o quelli mal rappresentati dallo schiacciamento operato dall'interpolazione. E' in quest'area che cadono il nord-ovest ed il nord-est, insieme alle società con oltre 4 anni di anzianità, ed il grado di evasione tende ad essere un mix tra quella media, quella bassa e quella degli "evasori senza risparmio Irpeg", che confermano così la loro atipicità. Da notare che accanto a quest'ultima tipologia si collocano le aziende "familiari", cioè quelle a totale controllo azionario da parte di membri di una famiglia, ma non è chiaro perché ciò si correli ad una condizione di già difficile interpretabilità.

Nella parte bassa del piano, a destra troviamo l'area dell'evasione medio-alta, mentre a sinistra quella alta, mediamente vicina al 100% del reddito imponibile effettivo.

L'area di evasione medio-alta si caratterizza per imprese di dimensione media (valore aggiunto, risultato economico e valore della produzione intermedi²⁰), ma anche con indicatori di bilancio intermedi (immobilizzazioni, oneri finanziari, ammortamenti, costi del personale, quota dei ricavi sul totale delle componenti positive) e quota intermedia delle variazioni. Le peculiarità da evidenziare sono le modalità rappresentanti il settore del commercio, un'evasione sia da costi (gonfiati) che ricavi (occultati), ma soprattutto una quota alta dei costi di magazzino, già emersa come un fattore indicatore di possibile occultamento di ricavi, ed un'alta quota di circolante, forse indicatrice di minori investimenti ed immobilizzazioni, e dunque minore convinzione nel futuro.

Anche la presenza di perdite riportate da periodi precedenti ci dice che questa condizione non è neutrale rispetto all'evasione, come è confermato dalle medie incrociate per gruppo: oltre il 26% delle imprese del gruppo con evasione alta hanno riportato perdite precedenti, mentre nei gruppi a bassa e media evasione tale percentuale scende al 17-18%. Sia la presenza di questa relazione, sia l'esistenza di un'area con evasione non immediatamente giustificata da immediato risparmio Irpeg,

²⁰ Si noti in particolare che il valore della produzione compreso tra 1 e 10 miliardi fa cadere queste imprese nell'area di applicazione degli studi di settore.

sembrano suggerire l'esistenza di un atteggiamento favorevole ad evadere o per motivi "culturali – di abitudine", o più probabilmente per tesaurizzare i vantaggi dell'evasione, attuandola cioè quando è possibile per utilizzarla, eventualmente ce ne fosse bisogno, mediante la riportabilità delle perdite²¹.

Infine, in basso a sinistra va analizzata l'area dell'alta evasione, dove si collocano cioè sia la modalità dell'evasione al 100% dell'imponibile effettivo o potenziale, sia il cluster dell'alta evasione, con tasso medio di evasione del 95%.

E' qui che vanno a cadere le modalità legate alla bassa dimensione e patrimonializzazione delle imprese (patrimonio netto nullo o piccolo, o comunque minoritario rispetto ai debiti, valore della produzione piccolo, risultato economico mediamente basso), ma anche bassi costi per collaborazioni e servizi, immobilizzazioni, ammortamenti. La quota dei ricavi sul totale entrate qui risulta elevato, così come sono alte le variazioni complessive e la parte di esse di carattere antielusivo, nonché presenti le variazioni classificate come "erosive".

Qui l'evasione è essenzialmente da occultamento ricavi, mentre tra le variabili strutturali compaiono l'essere una società nata di recente e comunque non controllata, la collocazione al sud o al centro, la natura giuridica di Srl, l'assenza di accantonamenti.

Per quanto riguarda la collocazione delle società "nuove" nell'area di massima evasione, più che ad una tendenza al deterioramento dei comportamenti fiscali si può pensare al fatto che quella parte di contribuenti "mordi e fuggi", cioè quelli che aprono e chiudono rapidamente imprese allo scopo di effettuare non solo evasione, ma anche altri progetti che necessitano di prestanome, assenza di patrimonio, scarsa reperibilità degli amministratori, vanno a cadere per forza di cose tra le aziende nuove, peggiorando la media di gruppo (Indeva1 al 47% per le nuove contro il 39% per le vecchie). Appare convincente anche che difficoltà iniziali di redditività non specificatamente prese in considerazione dal fisco vengono affrontate dagli operatori con una maggiore tendenza ad evadere, specie nel centro-sud, dove la citata presenza di forti incentivi all'apertura di nuove imprese senza reali controlli nel medio periodo pare paradossalmente incoraggiare la precarietà delle imprese e l'intensificazione dell'evasione, come si è visto anche nella tabella 12.

²¹ È d'altronde noto agli addetti ai lavori che non raramente i consulenti fiscali d'azienda si trovano a dover consigliare a propri assistiti di emettere regolarmente scontrini e fatture nell'ultimo periodo dell'anno, allo scopo di non esagerare con le perdite e di non far emergere anomalie negli indicatori di bilancio.

Un'ultima annotazione riguarda il ruolo positivo di una buona patrimonializzazione, che emerge come uno dei fattori più potenti per ridurre l'evasione; oltre alle tendenziali migliori *performance* aziendali ed alle grosse dimensioni, variabili correlate che non spingono ad evadere (ma che non compaiono essere più importanti della patrimonializzazione), sembra esserci una componente specifica ed addizionale di maggiore deterrenza dell'azione di controllo e recupero proprio a causa del maggiore patrimonio, che renderebbe meno conveniente e fattibile chiudere l'impresa e rendere inesigibili le pretese tributarie. Se questo elemento è vero, appare positiva la spinta data dal fisco alla patrimonializzazione delle imprese attraverso le riforme dell'Irap e della Dit-Superdit.

Considerazioni conclusive

Il recente dibattito di politica economica sull'utilizzo del "dividendo" da recupero di gettito avvenuto negli ultimi anni mostra quanto sia rilevante un allargamento della base imponibile che derivi da recupero di evasione e da deterrenza piuttosto che da fattori più congiunturali²². Tale dividendo può essere infatti utilizzato sia a fini redistributivi che per un abbassamento strutturale delle aliquote legali volto a favorire maggiore crescita e competitività.

In questo contesto, nel presente lavoro si è tentato di individuare e spiegare alcune relazioni significative tra evasione osservata, come constatata da verifiche generali su società di capitali, ed informazioni di natura "demografica", economica e patrimoniale, osservabili agevolmente e perciò capaci di meglio orientare l'attività di controllo.

Gli strumenti di analisi usati sono stati scelti al fine di valorizzare le interdipendenze tra più variabili osservate, senza vincolarsi al rispetto di modelli troppo rigidi o da definire a priori.

I risultati mostrano, in estrema sintesi, che pur permanendo la difficoltà di riscontro di un fenomeno, l'evasione, la cui caratteristica intrinseca è di non essere facilmente intuibile, sono rintracciabili indicatori, specie se considerati simultaneamente, in grado di migliorarne la comprensione.

Le medie incrociate, l'analisi di segmentazione e quella delle corrispondenze multiple mettono in evidenza molti fattori comuni, tra i quali la relazione inversa tra evasione da un lato e dimensione aziendale, patrimonializzazione e base imponibile Irap (*proxy* del valore aggiunto) dall'altro.

Altri importanti fattori associati ad un'alta propensione ad evadere sono la piccola dimensione aziendale, alte quote di oneri finanziari e costi di magazzino, la non appartenenza a gruppi, la presenza di perdite negli esercizi precedenti, l'alto grado di indebitamento, la presenza di variazioni antielusive consistenti, la recente data di costituzione se con sede legale nel centro-sud.

L'analisi di segmentazione ad albero individua anche le modalità che si associano alle punte del fenomeno in questione, ossia ai sottogruppi caratterizzati da una propensione all'evasione particolarmente alta ovvero particolarmente bassa. La massima propensione evasiva, oltre l'80%

²² Come ad es. le nuove forme di conciliazione che hanno anticipato riscossioni altrimenti diluite nel tempo, l'anticipazione di liquidazioni relative ad anni recenti, il fisco telematico che ha favorito il riallineamento degli imponibili fiscali e contributivi, ma anche un andamento congiunturale migliore del previsto o previsioni di entrate prudenti.

dell'imponibile effettivo, viene raggiunta quando convivono quote elevate di variazioni antielusive, indebitamento, oneri finanziari e costi per il personale. Tuttavia, i fattori già elencati di correlazione con l'evasione risultano significativi anche negli altri sottogruppi. Per quanto concerne la minore propensione evasiva, la segmentazione individua la possibile esistenza di una relazione di sostituibilità tra evasione o elusione contributiva ed evasione delle imposte dirette.

L'analisi delle corrispondenze, come confermata anche dalla *cluster analysis*, individua aree omogenee costituite da:

- società di comodo o non operative, con bassa evasione;
- un'area di bassa evasione correlata a basso circolante e bassi oneri finanziari, insieme ad alta quota di patrimonio netto, basso indebitamento, alte immobilizzazioni ed ammortamenti, alta quota di costi per servizi e collaborazioni;
- un'area di media evasione, caratterizzata da alte dimensioni aziendali e patrimonializzazione, evasione da costi più che da ricavi, società con oltre quattro anni di anzianità, settore industriale, variazioni antielusive basse, appartenenza a gruppi societari, area geografica nord-ovest e nord-est;
- un'area di alta evasione, correlata a bassa e media dimensione aziendale, bassa patrimonializzazione, società di recente costituzione, area geografica del centro-sud, alto indebitamento, variazioni alte con particolare riferimento a quelle antielusive ed erosive.

I risultati dell'analisi meritano, innanzitutto, di essere considerati all'interno di un disegno complessivo di politica fiscale.

Da questo punto di vista vale la pena di notare, in primo luogo, che politiche di incentivo fiscale alla **patrimonializzazione** (quali quelle legate alla Dit e alla Superdit) sembrano poter avere una qualche efficacia anche sul fronte della riduzione dell'evasione; in effetti, i risultati mostrano come l'alto indebitamento sia legato a comportamenti distorsivi tra cui rientra anche l'evasione, mentre la maggior patrimonializzazione aumenta il grado di stabilità dell'impresa e ne diminuisce la propensione evasiva.

Per quanto concerne le società di comodo, poi, i provvedimenti *ad hoc* presi con la legge 662_96 sembrano avere avuto un netto effetto disincentivante, anche se una parte delle imprese si è limitata a modificare i suoi parametri di attività (ad es. spostando sulla società di comodo qualche attività svolta altrove) in modo da non essere colpiti da un ricavo, reddito e tassazione forfetari.

D'altronde, gli studi di settore dovrebbero superare quasi totalmente anche quest'approccio, forfetizzando in base a parametri contabili ed extracontabili i ricavi e quindi i redditi anche per queste società.

In terzo luogo, sembra potersi proporre un discorso simile anche per le politiche tendenti a creare i presupposti per una **fiscalità diretta dei gruppi**: l'appartenenza ai gruppi fa infatti diminuire l'evasione contestata. Al riguardo, va peraltro precisato che tale minor evasione potrebbe essere collegata alla possibilità di ricorrere a molte forme di elusione e di erosione attraverso specifiche manovre intra-gruppo e quindi la definizione di una fiscalità di gruppo può rappresentare anche l'occasione per limitare talune di queste manovre. Una base di partenza, al riguardo, potrebbe essere fornita dall'art.12 della proposta di legge presentata nel 1994 dall'attuale Ministro del Tesoro Visco.

Per altro verso, i risultati ottenuti richiamano l'attenzione sulla necessità di interrogarsi circa le politiche di incentivo alla creazione di **nuove imprese**; la maggiore propensione all'evasione osservata per queste società potrebbe essere legata all'evasione *mordi e fuggi* e rappresentare, in qualche modo, un dato di sistema. In effetti, gli ulteriori incroci presentati nella tabella 12 hanno mostrato che la maggiore evasione delle società nuove rispetto alle altre è presente solo nel centro-sud. Questo fenomeno potrebbe essere legato anche alle numerose leggi di incentivazione alla creazione di nuove imprese vigenti in questa zona del paese, tutte praticamente prive di meccanismi di riscontro dell'effetto netto di lungo periodo su attività ed occupazione²³, nonché di filtri che impediscano alle stesse persone di beneficiare più volte di ingenti aiuti chiudendo società ed aprendone altre. Questa interpretazione, la cui fondatezza andrebbe verificata con dati sugli incentivi utilizzati dalle singole imprese, non appare in contrasto con l'esistenza di un maggiore *turnover* di imprese nel sud riscontrato nel lavoro del Secit citato in precedenza.

Ancora, l'esistenza di una relazione statistica significativa tra le alte **perdite riportate** e l'evasione suggerisce di inserire anche questo elemento nella discussione sulla revisione, in chiave europea, di questi istituti.

²³ Solo di recente, e solo a seguito di ripetute sollecitazioni della UE, l'esigenza generale di valutazione ex-ante ed ex-post delle politiche pubbliche ha trovato un riscontro normativo con la previsione dell'istituzione presso le Amministrazioni centrali e regionali di apposite unità tecniche di monitoraggio e valutazione (art. 1 L. 144/99).

I risultati suggeriscono inoltre diverse considerazioni utili per la definizione di migliori **strategie di controllo ed accertamento**.

In primo luogo, approcci come quello proposto potrebbero costituire un'integrazione delle metodologie di programmazione dei controlli in tema di evasione dell'Irpeg. Ovviamente, un'utilizzazione su ampia scala dell'analisi di segmentazione richiederebbe un affinamento delle variabili utilizzate ed una maggiore robustezza statistica dei nessi qui individuati. Tenendo conto anche dei tassi di caduta degli accertamenti a seguito di concordati e contenzioso tributario, sarebbe utile potersi riferire all'evasione "definita" anziché a quella contestata.

Si potrebbero estendere con sistematicità le analisi multivariate all'intera base informativa sulle attività di controllo, accertamento e definizione svolte annualmente, aumentando l'affidabilità dei risultati, grazie alla maggiore numerosità campionaria, ma perdendo qualcosa in termini di rappresentatività: non va dimenticato, infatti, che i controlli vengono programmati sin dall'inizio con l'obiettivo di massimizzare la proficuità in termini di maggiore imposta accertata e riscossa, e rappresentano perciò un campione fortemente distorto dell'insieme delle aziende.

Un secondo ordine di considerazioni, generato dalla maggiore evasione riscontrata su piccole e medie aziende con valore della produzione fino a 10 miliardi, riguarda la recente innovazione degli **studi di settore**, ai quali l'amministrazione finanziaria potrebbe in futuro affidare anche funzioni parzialmente sostitutive dei controlli ordinari, e più in generale della contabilità analitica.

Sebbene gli strumenti statistici per pervenire alla quantificazione dei ricavi siano molto sofisticati e non troppo differenti da quelli qui utilizzati, ed il metodo di concertazione preventiva con le categorie interessate sia basato sul buon senso e sul consenso, resta il fatto che le basi dati di natura censuaria utilizzate negli studi di settore non osservano direttamente l'evasione.

Più in dettaglio, essendo gli studi di settore finalizzati alla costruzione di gruppi che risultano omogenei rispetto a talune variabili, contabili e non, dichiarate dai contribuenti²⁴, non è detto che queste variabili rappresentino anche quelle significative ai fini della stima dell'evasione vs la stima delle diversità di redditività aziendale. E' quindi teoricamente possibile che in un determinato gruppo omogeneo tutte le imprese risultino avere ricavi congrui semplicemente perché hanno lo stesso comportamento evasivo; oppure, ed è questo forse il caso di maggiore interesse, è possibile

²⁴ Per una sintesi degli studi di settore si veda Brunello (2000).

che parte delle imprese con ricavi incongrui siano effettivi evasori, mentre altre siano semplicemente imprese con redditività inferiore a quella media del gruppo.

Per ottenere una migliore discriminazione tra evasione e *performance* aziendali, ed una migliore selezione delle variabili utili per le stesse stime e per i controlli ordinari, si potrebbe fare ricorso ad un approccio come quello qui proposto, adattato ovviamente alla realtà degli studi di settore ed esteso agli esiti dei controlli annuali realizzati dall'amministrazione finanziaria. Poiché la logica del reddito normale, e le esigenze di miglioramento del rapporto tra fisco e contribuenti, lasciano prevedere che il contribuente con ricavi congrui non subisca controlli di alcun tipo, la congruità dovrebbe essere stabilita a livelli di precisione sufficientemente alti; si tratta quindi di trovare un buon compromesso tra le esigenze di semplificazione del sistema fiscale e di miglioramento dei rapporti tra il fisco e i contribuenti, e quelle contrapposte di mantenimento di un certo livello di rispondenza tra il risultato economico dichiarato e quello effettivo²⁵.

A tal fine, e tenuto conto che l'evasione sembra derivare principalmente da ricavi occultati nelle piccole imprese, pare opportuno non abbandonare il ruolo dei cosiddetti "controlli strumentali", cioè quelli svolti a campione dalla Guardia di Finanza sulla corretta emissione di scontrini e ricevute, attività finora relegata di fatto in un ruolo marginale ed a rischio di abbandono nel futuro; a chi scrive appare che controlli strumentali a campione abbiano un'utilità insostituibile ai fini sia dell'essenziale funzione di deterrenza, sia dell'attendibilità dei ricavi dichiarati presi a riferimento anche dagli studi di settore.

²⁵ Dell'esigenza di salvaguardare l'accertamento analitico è convinta anche una parte significativa del mondo produttivo, cfr. Osservatorio fiscale (2000).

Bibliografia

- Andreoni J., Erard B. e Feinstein J.** (1998), *"Tax compliance"*, Journal of Economic Literature, 36, 818-60.
- Bernardi L. e Bernasconi M.** (1996), *"L'evasione fiscale: evidenze empiriche"*, Il Fisco, suppl.38.
- Bernardi L.** (2000), *"Note sull'evoluzione recente e sulle prospettive future dei sistemi tributari"*, Studi e note di economia, n.1_2000, Ed. Banca Toscana, Firenze.
- Biggs D., De Ville B. e Suen E.** (1991), *"A method of choosing multiway partitions for classification and decision trees"*, Journal of Applied Statistics, 18,1, 49-62.
- Bolasco S. e Coppi R.** (1983), *"Analisi dei dati e sue applicazioni nella ricerca economica"*, Klim Ed., Roma.
- Bordignon M. e Zanardi A.** (1997), *"Tax evasion in Italy"*, Giornale degli Economisti e Annali di Economia, vol.56, n..3-4, pag.169-210.
- Bourouche J. M. e Saporta G.** (1984), *« L'Analyse des Données »*, éd. Paris: P.U.F., traduzione italiana *"L'Analisi dei dati"*, a cura di D'Alfonso, CLU Ed., Napoli.
- Brunello M.** (2000) *"Il contesto nel quale si collocano gli studi di settore"*, Atti del convegno *"I nuovi studi di settore"*, Il Fisco,25, alleg.31.
- CCIAA Milano - Osservatorio Fiscale** (2000), *"La fiscalità d'impresa dal 1973 al 2000. Analisi e proposte per un nuovo patto tra fisco e contribuente"*, Camera di Commercio, Industria, Artigianato e Agricoltura della provincia di Milano.
- Coppi R.** (1986), *"Appunti di statistica metodologica: analisi lineari dei dati"*, Dip. Statistica, Probabilità e statistiche applicate, Univ. Roma *"La Sapienza"*.
- Di Majo A.** (2000), *"Il sommerso si batte con i fatti"*, Sole 24 Ore 8.9.2000.
- Di Nicola F., Golino S., Rinaldi V., Santoro A. e Stajano R.** (2000), *"La pressione Irpeg sulle società di capitali"*, Il Fisco n.12,2000, pagg. 3250-3254.
- Giusti F.** (1983), *"Introduzione alla statistica"*, Loescher ed., Torino.
- Greenacre M.J.** (1984), *"Theory and applications of correspondence analysis"*, Academic Press, London.
- Lupi R.** (1995), *"La contabilità aziendale spiegata ai giuristi"*, Cedam, Padova.

Meulman J. e Heiser W. J. (2000), “*SPSS Categories*”, SPSS Inc. Ed., Chicago, USA.

Mood A., Graybill F. e Boes D.(1974), “*Introduction to the theory of statistics*”, Mc Graw-Hill, New York, 3rd edition.

Rizzi A., (1985), “*Analisi dei dati ed applicazione dell’informatica alla statistica*”, NIS editore, Roma.

Rizzi A., (1981), “*L’analisi dei gruppi – Cluster analysis*”, 2^a edizione, La Goliardica Ed., Roma.

Shyr J., (1999), “*Tree-Structured Data Analysis*”, www.sinter.com.tw/spss/WPINDEX.html/Tree-Structured.htm.

SPAD (1998), “*Système pourtable pour l’analyse des données*”, CISIA Ed., Saint Mandè (France).

SPSS (2000), “*SPSS Answer Tree*”, Ed. SPSS Inc., Chicago, USA.

T. H. Wonnacott e R. J. Wonnacott (1990), “*Introduzione alla statistica*”, Ed. Franco Angeli, Milano.

Appendice 1: L'algoritmo “CHAID esaustivo” di segmentazione ad albero

Si tratta di un' analisi che descrive il comportamento di una determinata variabile obiettivo utilizzando k variabili indipendenti dette predittori (P_i con $i=1, \dots, k$)

I predittori e la variabile obiettivo possono essere:

i) variabili qualitative o nominali:

ii) variabili continue o di scala;

iii) variabili categoriali o ordinali. In questo caso chi compie l'analisi definisce, per ciascuna variabile categoriale P_i , un numero k_i di categorie (o “modalità”).

Nel lavoro descritto nei paragrafi successivi la variabile obiettivo è continua, mentre i predittori sono variabili categoriali.

Per comprendere il funzionamento dell'algoritmo CHAID esaustivo²⁶, ipotizziamo per semplicità espositiva di trovarci al primo livello di segmentazione della variabile obiettivo quantitativa, con k predittori categorici già definiti, ed assumiamo che $k_i=2 \forall P_i$, ovvero che tutti i predittori siano distinti solo in due categorie.

L'algoritmo agisce in 2 fasi :

- a) nella prima fase testa l'ipotesi, per ogni predittore, che la media della variabile obiettivo sia uguale nelle due categorie del predittore;
- b) nella seconda fase sceglie come predittore quello in relazione a cui tale ipotesi è rigettata con maggior probabilità, cioè quello che differenzia al massimo le medie calcolate su ognuno dei due sottogruppi per variabile identificati dalle due categorie ipotizzate.

L'analisi dell'ipotesi dell'uguaglianza delle medie avviene tramite la cosiddetta *one-way analysis of variance* ²⁷. In sostanza, denotando con x la variabile obiettivo e con P_{ij} l'insieme dei

²⁶ Per una descrizione tecnica si può consultare Biggs et al, (1991). Per una descrizione più generale sui diversi algoritmi di segmentazione utilizzabili con AnswerTree, si veda J.Shyr (1999).

²⁷ Si veda, Mood, Graybill e Boes,(1974), pp.435-8.

casi in cui il predittore P_i appartiene alla categoria j del predittore P_i ($j=1,2$), la *null hypothesis* per ciascun predittore P_i , è la seguente:

$$H_0: E(x|x \in P_{i1}) = E(x|x \in P_{i2})$$

mentre l'ipotesi alternativa è

$$H_1: E(x|x \in P_{i1}) \neq E(x|x \in P_{i2}).$$

Se la distribuzione della popolazione in tutte le categorie è normale, allora la statistica utilizzabile per il test è una *F-statistics*, che assumerà un valore determinato per ogni predittore P_i . Se denotiamo questo valore con F_i , e con $\{F_P\}$ l'insieme delle statistiche ottenute (ovvero $\{F_P\} = \{F_1, \dots, F_n\}$) si ha che il predittore selezionato dall'algoritmo è $P_i: F_i = \max \{F_P\}$.

Nel caso le categorie selezionate da chi utilizza il programma siano più di 2, prima di effettuare i passaggi precedenti il programma effettua un test di similitudine con l'obiettivo di individuare i raggruppamenti di categorie tra le quali la media della variabile obiettivo è vicina. Tale operazione può portare, ad esempio, a raggruppare 4 categorie originarie in due sole classi, in relazione a cui vanno ripetute le operazioni descritte in precedenza.

E' possibile che i raggruppamenti individuati siano superiori a 2, oppure che non sia effettuato alcun raggruppamento e che venga mantenuta la distinzione in più di 2 categorie data originariamente; in questi casi le ipotesi (nulla ed alternativa) vengono riformulate in relazione al numero dei raggruppamenti o delle categorie originarie.

Una volta scelto il predittore P_i per il primo livello di segmentazione, si individuano anche i primi rami dell'albero, cioè i primi sottogruppi, in numero pari alle categorie originarie o ai raggruppamenti individuati. Le operazioni descritte in precedenza vengono quindi ripetute per ciascuno di questi rami e, aspetto importante, in maniera indipendente dagli altri, per tutti i livelli richiesti; poiché ogni ramo dà origine ad una nuova segmentazione, è probabile che in corrispondenza di uno stesso livello di segmentazione vengano selezionati predittori diversi per rami diversi, così come è possibile che una stessa variabile venga selezionata a livelli diversi e quindi con valori diversi.

La scelta del numero di livelli è normalmente effettuata dall'utilizzatore in modo da non diminuire eccessivamente la numerosità campionaria delle celle di casi così determinate. Il programma richiede, inoltre, di selezionare la numerosità campionaria minima dei nodi dell'albero.

In questo lavoro, la variabile obiettivo è stata fissata nell'indicatore della propensione all'evasione indeval, mentre i predittori sono le variabili descrittive definite in precedenza (cfr.

paragrafo 3). Il numero dei livelli di profondità massima dell'albero è stato fissato a 5, mentre la numerosità minima dei nodi-genitore è stata fissata a 20000 casi (cioè circa il 5% della numerosità dell'universo) e quella dei nodi-figlio a 10000 casi: ciò implica che la segmentazione viene interrotta se il nodo da segmentare ha meno di 20000 casi ovvero se non tutti i nodi che ne risulterebbero hanno una numerosità superiore a 10000 casi.

Nell'interpretazione dei risultati, si potrà tenere conto della natura campionaria delle stime ottenute, dato il piano di campionamento rappresentativo adottato. In particolare, è possibile costruire, a partire dall'errore standard della stima campionaria, l'intervallo di confidenza che, con un certa probabilità, contiene il parametro di interesse. Con riferimento alla media stimata $\hat{\mu}$ tale intervallo assume la seguente forma

$$\Pr\{\hat{\mu} - k\hat{\sigma}_{\hat{\mu}} \leq \mu \leq \hat{\mu} + k\hat{\sigma}_{\hat{\mu}}\} = P$$

dove $\hat{\sigma}_{\hat{\mu}}$ è la stima dell'errore di campionamento assoluto, mentre il valore di k dipende dalla probabilità P . L'errore standard della stima, nel caso di campionamento con ripetizione o senza ripetizione, ma con numerosità della popolazione di riferimento sufficientemente elevata²⁸, è pari a

$$\hat{\sigma}_{\hat{\mu}} = \frac{\sqrt{Var(x)}}{\sqrt{n}}$$

dove n è la numerosità campionaria e x è indeval. Nel prosieguo si farà riferimento a $P=0.95$ e quindi il valore rilevante di k è 1.96.

²⁸ Si veda ad es. Giusti (1983), pag. 504 e seguenti.

Appendice 2: Richiami sull'analisi delle corrispondenze multiple

L'analisi delle corrispondenze multiple²⁹ è una metodologia statistica proposta in Francia negli anni '70 ed affermata per lo studio di collettivi con informazioni qualitative sconnesse (o quantitative, se ricodificate mediante aggregazione in classi), e consiste nel costruire nuove variabili dette fattoriali, fittizie perchè non osservate ma ottenute mediante una combinazione lineare delle variabili originarie ottimizzata col criterio di minimizzazione delle distanze; questi fattori, che possono essere visti anche come rette interpolanti, sono tra loro incorrelati, consentendo una efficace riproduzione grafica su piano cartesiano dell'insieme di modalità oggetto di studio³⁰.

A causa della proiezione (o "schiacciamento") della nuvola originaria di punti-modalità sul piano fattoriale individuato dai soli primi due fattori si perde ovviamente una parte dell'informazione originaria, teoricamente rappresentata dalla variabilità dei punti nell'iperspazio originario di rappresentazione; è possibile però misurare la perdita di tale variabilità mediante un rapporto che prende il nome di "inerzia spiegata", e che può essere considerato un indicatore del tutto analogo al più noto indice di determinazione (varianza spiegata su varianza totale).

Punti-modalità vicini tra loro implicano la presenza congiunta di quelle modalità da parte degli stessi casi (le imprese nel nostro caso), mentre nel caso di lontananza sul piano fattoriale (quello individuato dai primi due fattori) si può dedurre che le imprese con una certa modalità raramente ne posseggono un'altra posizionata in un'altra area del piano. Ne consegue che un gruppo di modalità vicine tende ad identificare aree omogenee di comportamento, il che risulta particolarmente efficiente in alternativa all'esame di un enorme numero di tabelle a molte entrate, ed anche efficace come spunto per ulteriori approfondimenti.

Tra le limitazioni vanno ricordati i vincoli di dover categorizzare anche variabili di natura quantitativa, la natura non inferenziale del metodo e la rinuncia a conclusioni sulla dipendenza di una sola variabile dalle altre (obiettivo per il quale abbiamo qui sperimentato l'analisi della segmentazione in alternativa al modello logistico).

²⁹ Per approfondimenti relativi all'analisi delle corrispondenze, alla cluster analysis ed al raccordo tra le due si vedano Bouroche e Saporta (1984), Rizzi (1985), SPAD (1998), Meulman e Heiser (2000), Greenacre (1984), Coppi (1986), Bolasco e Coppi (1983).

³⁰ Perdendo ovviamente una parte dell'informazione originaria a causa dell'interpolazione; è possibile però misurare la perdita della variabilità dei punti mediante un rapporto che prende il nome di "inerzia spiegata", e che può essere considerato un indicatore della stessa natura del coefficiente di determinazione.

Nel nostro caso, l'alto numero di modalità di risposta da proiettare sul piano fattoriale delle corrispondenze e l'esigenza di metterle in relazione alle altre ci hanno consigliato una riclassificazione ragionata in poche risposte del tipo “alto, medio, basso”, in particolare per categorizzare le variabili quantitative (ad es. quota del patrimonio netto sul totale delle passività) dopo aver osservato l'intera distribuzione.

La proiezione di ogni modalità sul piano fattoriale è stata sfruttata, come detto, anche per assegnare una misura all'indicatore di evasione relativa (“indeval”) per la citata categoria degli “evasori senza risparmio d'imposta”³¹; il tasso di evasione Indeval risultava infatti in quei casi indeterminato o senza un senso coerente con l'obiettivo perseguito.

Dato l'utilizzo essenzialmente visivo dell'analisi delle corrispondenze, l'osservazione del piano fattoriale principale è stata integrata da un'analisi dei gruppi³² a partire dalle prime tre coordinate fattoriali e dall'indicatore Indeval, il cui inserimento è servito a dare maggior peso all'intensità dell'evasione nel processo di raggruppamento.

In questo modo sono stati conseguiti i seguenti vantaggi:

- consentire una *cluster analysis* pensata per variabili quantitative, ma applicata invece a modalità di variabili categoriche;
- individuare tipologie analoghe a quelle individuate con l'analisi delle corrispondenze, e ciò è tornato utile in sede di interpretazione dei risultati; non va dimenticato infatti che la cluster analysis ha il limite di dover allocare anche i "punti-frontiera" tra due gruppi, indebolendo così l'interpretabilità e l'omogeneità interna di essi.

Il processo di raggruppamento gerarchico ascendente utilizzato ha generato anche il cosiddetto *dendrogramma*, cioè una rappresentazione grafica della distanza media tra baricentri da superare per unire due “sottogruppi” durante il processo di aggregazione progressiva; in questo modo è stato possibile individuare un numero soddisfacente, che è stato fissato pari a tre, di gruppi finali da spiegare, risultati fortemente caratterizzati rispettivamente da tasso di evasione basso (Indeval mediamente pari al 6%), medio (56%) e alto (95%).

³¹ cioè quelli con presenza di evasione, ma anche con imponibile effettivo negativo, che non dovrebbe portare a comportamenti evasivi.

³² Il cosiddetto legame medio tra gruppi, di natura gerarchica ascendente ed operante sulle distanze quadratiche euclidee, utilizzato di default dal programma statistico SPSS, al quale si rimanda per ulteriori dettagli.

I baricentri di questi gruppi, grazie al fatto che anche le imprese hanno coordinate fattoriali a seguito dell'analisi delle corrispondenze, hanno potuto essere proiettati sul piano delle corrispondenze, accrescendo la capacità interpretativa.