

istat working papers

N. 7
2011

Utilizzo delle fonti amministrative nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

*M.C. Casciano, A. Cirianni, V. De Giorgi, T. Di Francescantonio, A. Mazzilli, O. Luzi,
F. Oropallo, M. Rinaldi, E. Santi, G. Seri, G. Siesto*

istat working papers

N. 7
2011

Utilizzo delle fonti amministrative nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

*M.C. Casciano, A. Cirianni, V. De Giorgi, T. Di Francescantonio, A. Mazzilli, O. Luzi,
F. Oropallo, M. Rinaldi, E. Santi, G. Seri, G. Sesto*

Comitato di redazione

Coordinatore: Giulio Barcaroli

Componenti:

Rossana Balestrino	Francesca Di Palma	Luisa Picozzi
Marco Ballin	Alessandra Ferrara	Mauro Politi
Riccardo Carbini	Angela Ferruzza	Alessandra Righi
Claudio Ceccarelli	Danila Filipponi	Luca Salvati
Giuliana Coccia	Cristina Freguja	Giovanni Seri
Fabio Crescenzi	Aurea Micali	Leonello Tronti
Carla De Angelis	Nadia Mignolli	Sonia Vittozzi

Segreteria:

Lorella Appolloni, Maria Silvia Cardacino, Laura Peci, Gilda Sonetti, Antonio Trobia

Istat Working Papers

Utilizzo delle fonti amministrative nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

N. 7/2011

ISBN 88-458-1689-3

Istituto nazionale di statistica
Servizio Editoria
Via Cesare Balbo, 16 – Roma

Utilizzo delle fonti amministrative nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

M.C. Casciano, A. Cirianni, V. De Giorgi, T. Di Francescantonio, A. Mazzilli,
O. Luzi, F. Oropallo, M. Rinaldi, E. Santi, G. Seri, G. Siesto

Sommario

Il presente lavoro analizza diverse fonti informative di natura amministrativa contenenti dati economici sulle imprese e fornisce una valutazione sperimentale sul loro possibile impiego nei processi di produzione statistica, in particolare per la rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI). L'obiettivo è quello di verificare la possibilità di alleggerire l'onere statistico sulle imprese e di fornire elementi utili volti ai fini di una riorganizzazione complessiva dei processi di produzione per le statistiche strutturali sulle imprese (SBS), con il passaggio dall'attuale strategia basata su processi di produzione "verticali" che utilizzano indipendentemente e spesso solo marginalmente il dato amministrativo, ad un sistema completamente integrato di cui le fonti amministrative costituiscano la base informativa di riferimento, opportunamente integrata da un sistema di indagini campionarie, per la produzione di tutte le stime SBS al livello di dettaglio richiesto dagli utenti interni ed esterni.

Parole chiave: statistiche strutturali sulle imprese, fonti amministrative, micro-integrazione, mancate risposte.

Abstract

In this paper the different administrative sources containing economic information on enterprises are analysed, and an experimental study for evaluating their possible use in statistical production processes is performed. In particular, the paper focuses on the Italian survey on Small and Medium Enterprises (SME). The main purpose is to verify the possibility of reducing the burden on enterprises, and of providing useful elements for the overall redesign of the structural business statistics (SBS) production process. More specifically, the objective is to move from the current "stove-pipe" production model, where administrative data are used independently in different survey processes, to a completely integrated production strategy, where the administrative information form the core of the overall SBS statistical production system, possibly integrated with a specifically-designed system of sample surveys for the production of the whole set of SBS estimates at the required detail levels.

Keywords: structural business statistics, administrative data, micro-integration, missing data.

M. C. Casciano, Istat, casciano@istat.it, par. 2.3, 3.2, 3.4.3; app. A.2, A.2.1, A.2.2, A.2.3

A. Cirianni, Istat, cirianni@istat.it, app. D

V. De Giorgi, Istat, degiorgi@istat.it, par. 3.2, 3.3.1, 3.3.2, 3.4.2, 3.4.3, 3.4.4; app. A.1, A.4, A.5, A.5.7, A.5.8, B

T. Di Francescantonio, Istat, difrance@istat.it, par. 3.2; app. A.5, A.5.5, A.5.6

A. Mazzilli, Istat, mazzilli@istat.it, app. A.5, A.5.9, A.5.10

O. Luzi, Istat, luzi@istat.it, par. 1, 2.1, 2.1.1, 2.1.2, 4.

F. Oropallo, Istat, oropallo@istat.it, par. 3.2, 3.4.1, 3.4.3; app. A.3, A.4, A.5, A.5.1, A.5.2, A.5.3, A.5.4

M. Rinaldi, Istat, mrmarina@istat.it, app. C

E. Santi, Istat, santi@istat.it, app. D

G. Seri, Istat, seri@istat.it, par. 2.2

G. Siesto, Istat, siesto@istat.it, par. 1, 3.1, 3.2, 3.4.2, 3.4.3, 3.4.4; app. A.2, A.2.2, A.2.4, A.3, A.5

Indice

	Pag.
1. Introduzione e obiettivi della ricerca	8
2. L'uso dei dati amministrativi per la produzione dell'informazione statistica sulle imprese	10
2.1. La qualità dei dati di fonte amministrativa.....	10
2.1.1. <i>La verifica della qualità dei dati di fonte amministrativa: controllo e correzione di errori e mancate risposte</i>	12
2.1.2. <i>L'uso dei dati amministrativi per il miglioramento della qualità delle statistiche economiche</i>	13
2.2. L'uso dei dati di fonte amministrativa nella stima degli aggregati SBS in Europa.....	15
2.3. Disegni di campionamento integrati all'uso di fonti amministrative.....	17
3. L'uso delle fonti amministrative per l'integrazione delle mancate risposte nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI): uno studio sperimentale	20
3.1. La rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni.....	20
3.2. Fonti statistiche e amministrative utilizzate nell'attività di ricerca e definizioni delle variabili.....	23
3.3. Copertura dell'archivio ASIA e del campione PMI con le fonti amministrative disponibili.....	24
3.3.1. <i>Copertura dell'archivio ASIA</i>	24
3.3.2. <i>Copertura del campione PMI</i>	28
3.4. La procedura di integrazione delle mancate risposte nella rilevazione PMI basata sull'uso delle fonti amministrative disponibili.....	33
3.4.1. <i>Confronti rispetto a PMI rispondenti e selezione delle fonti con le variabili che hanno maggiore coerenza con quelle di PMI2007</i>	33
3.4.2. <i>Ricostruzione del campione teorico</i>	41
3.4.3. <i>Analisi delle differenze tra campione della rilevazione PMI e campione ricostruito e valutazione dell'effetto delle mancate risposte totali</i>	41
3.4.4. <i>Un'applicazione dell'utilizzo delle fonti amministrative nella rilevazione PMI2008</i>	50
4. Conclusioni e prospettive future	51
Riferimenti bibliografici	52
Appendice A.: Fonti statistiche e amministrative sulle imprese disponibili per l'Istat ...	56
A.1. Archivio statistico delle imprese attive.....	56
A.2. Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni.....	56
A.2.1. <i>Il disegno campionario</i>	56
A.2.2. <i>L'integrazione con dati di fonte Bilanci civilistici</i>	61
A.2.3. <i>Procedura di stima</i>	64
A.2.4. <i>Principali variabili</i>	70
A.3. Bilanci civilistici delle società di capitale.....	71
A.4. Studi di settore.....	71
A.5. Modello Unico e modello Irap.....	71
A.5.1. <i>Modello Unico Persone fisiche -Quadro RE: Redditi di lavoro autonomo</i>	74

A.5.2. <i>Modello Unico Persone fisiche - Quadro RF: Redditi di impresa in contabilità ordinaria.....</i>	75
A.5.3. <i>Modello Unico Persone fisiche - Quadro RG: Redditi di impresa in regime di contabilità semplificata e regimi forfetari.....</i>	75
A.5.4. <i>Modello Irap Persone fisiche - Quadro IQ: Irap.....</i>	76
A.5.5. <i>Modello Unico Società di persone - Quadro RE: Redditi di lavoro autonomo derivanti dall'esercizio di arti e professioni.....</i>	77
A.5.6. <i>Modello Unico Società di persone - Quadro RF: Redditi di impresa in contabilità ordinaria.....</i>	77
A.5.7. <i>Modello Unico Società di persone - Quadro RG: Redditi di impresa in regime di contabilità semplificata.....</i>	77
A.5.8. <i>Modello Irap Società di persone - Quadro IQ: Irap.....</i>	78
A.5.9. <i>Modello Unico Società di capitale - Quadro RS: Prospetti vari.....</i>	79
A.5.10. <i>Modello Irap Società di capitale - Quadro IQ: Irap.....</i>	79
Appendice B.: Tavole dettagliate della copertura dell'archivio ASIA e del campione PMI con le fonti amministrative disponibili.....	81
Appendice C.: La stima preliminare del regolamento SBS: un caso pratico di integrazione fra dati di fonte statistica e amministrativa.....	86
Appendice D.: Gli adempimenti amministrativi delle imprese con un solo addetto senza dipendenti e delle imprese minime.....	89
D.1. <i>La classificazione delle microunità in base agli adempimenti amministrativi.....</i>	90
D.2. <i>Le imprese minime.....</i>	91
D.3. <i>Le imprese artigiane.....</i>	94
D.4. <i>Analisi descrittiva delle microunità per sezione di attività economica.....</i>	95
D.5. <i>Analisi descrittiva delle microunità in base agli adempimenti amministrativi ed al settore di attività economica di appartenenza.....</i>	96
<i>Conclusioni.....</i>	103

1. Introduzione e obiettivi della ricerca

La produzione di statistiche strutturali armonizzate a livello europeo è notevolmente cresciuta nel corso degli ultimi quindici anni, grazie all'introduzione del Regolamento sulle Statistiche Strutturali sulle imprese (SBS) n° 58/97 (approvato il 20 dicembre 1996), ai successivi emendamenti e al nuovo Regolamento n°295/2008 (approvato l'11 marzo 2008).

L'Istat, al fine di adempiere agli obblighi derivanti dai regolamenti europei e per soddisfare le esigenze della contabilità nazionale e di conoscenza delle caratteristiche del sistema delle imprese del Paese, ha progressivamente accresciuto il potenziale informativo di statistiche strutturali. Ha, inoltre, affrontato problemi di efficienza dei processi di costruzione delle informazioni statistiche, particolarmente rilevanti per il nostro Paese caratterizzato da una notevole consistenza di imprese di piccole e piccolissime dimensioni, che determinano una media di addetti per impresa molto bassa (3,9). Su circa 4,4 milioni di imprese attive in Italia nel 2007 - che assorbono 17,0 milioni di addetti, fatturano 2.961,4 miliardi di euro e creano 722 miliardi di euro di valore aggiunto - le imprese con meno di 10 addetti rappresentano, infatti, il 94,8% delle imprese del Paese, assorbono il 47,4% degli addetti totali, il 25,1% dei dipendenti, il 27,1% del fatturato ed il 32,5% del valore aggiunto; in questo segmento dimensionale di imprese quasi due terzi dell'occupazione è costituita da lavoro indipendente. Questa caratteristica determina oneri statistici rilevanti per l'Istat, nettamente superiori a quelli di altri Paesi europei in cui maggiore è la concentrazione di imprese nelle classi dimensionali medie e grandi, per l'elevata numerosità dei campioni necessari a garantire la qualità delle stime delle variabili con elevati livelli di disaggregazione (a livello settoriale, dimensionale e territoriale), previsti dal Programma Statistico Nazionale e dai Regolamenti comunitari.

D'altra parte, anche il sistema delle imprese sopporta oneri piuttosto elevati per la compilazione dei questionari d'indagine e, data la crescita delle esigenze informative sul sistema delle imprese, ciò produce un fastidio statistico per le imprese che l'Istat ha cercato di minimizzare soprattutto attraverso il coordinamento dei campioni delle varie rilevazioni.

Questo quadro rafforza, pertanto, la necessità di arrivare ad un miglior sfruttamento del complesso delle fonti informative di natura amministrativa esistenti sul sistema delle imprese, allo scopo di integrare e/o sostituire le rilevazioni dirette per quel che riguarda le informazioni economiche già raccolte dalla Pubblica Amministrazione a fini amministrativi e quindi alleggerire l'onere statistico sulle imprese. In prospettiva, si può realisticamente configurare una riorganizzazione complessiva dei processi di produzione per le statistiche economiche, con il passaggio dall'attuale strategia basata su processi di produzione "verticali" che utilizzano indipendentemente e spesso solo marginalmente il dato amministrativo, ad un sistema completamente integrato di cui le fonti amministrative costituiscano la base informativa di riferimento (opportunamente integrata da un sistema di indagini campionarie) per la produzione di tutte le stime SBS al livello di dettaglio richiesto dagli utenti interni ed esterni.

L'utilizzo dei dati previdenziali (occupazione e retribuzioni di fonte INPS) e dei dati di bilancio delle società di capitale (variabili del conto economico e dello stato patrimoniale di fonte Camere di Commercio) è ormai consolidato nell'Istituto, mentre rimangono da valutare ed analizzare più approfonditamente le opportunità di impiego del complesso delle dichiarazioni fiscali (fonte Ministero dell'Economia e delle finanze) ai fini statistici, specialmente nell'ottica di ridurre il carico statistico per le piccole e piccolissime imprese.

Sotto questo aspetto, la trasmissione telematica delle dichiarazioni fiscali delle persone fisiche e giuridiche al Ministero dell'economia e delle finanze è notevolmente aumentata nel corso degli anni diventando, peraltro, progressivamente obbligatoria per tutti i soggetti, e ciò ha permesso di ottenere la disponibilità di dati fiscali più aggiornati in tempi sensibilmente ridotti rispetto al passato, rendendo sempre più la fonte fiscale vicina alle esigenze statistiche per soddisfare le richieste comunitarie. Inoltre, non trascurabile è stato il miglioramento qualitativo dei dati inviati dalle imprese all'amministrazione fiscale.

L'Istat, grazie ad una convenzione stipulata con il Ministero dell'economia e delle finanze, ha la facoltà di richiedere anche dati individuali dell'Amministrazione fiscale e quindi ha la possibilità di acquisire, per finalità strettamente statistiche, le dichiarazioni informatizzate dei modelli *Unico Persone Fisiche* (che compilano le imprese individuali e familiari ed i lavoratori autonomi), *Unico Società di Persone* (che compilano le società semplici o di fatto, società in nome collettivo, società in accomandita semplice, studi associati, società cooperative a responsabilità illimitata), *Unico Società di Capitale* (che compilano le società per azioni, società a responsabilità limitata, società in accomandita per azioni, cooperative a responsabilità limitata e le cooperative a responsabilità per azioni) e degli *Studi di Settore* (che hanno la finalità di contrastare l'evasione fiscale attraverso metodologie di analisi statistiche ed economiche, grazie alle quali l'Amministrazione finanziaria è in grado di accertare con efficienza i redditi di impresa e di lavoro autonomo, tenendo conto di variabili sia contabili che extracontabili o strutturali).

Queste considerazioni spingono sempre più l'Istat verso l'utilizzo di dati di fonte amministrativa, il cui impiego è, tra l'altro incoraggiato dallo stesso regolamento SBS n°295/2008 che all'articolo 5 (acquisizione dei dati), comma 2, stabilisce che gli Stati membri, ispirandosi al principio della semplificazione amministrativa, possono procurarsi i dati necessari avvalendosi di una combinazione di fonti (rilevazioni dirette obbligatorie, fonti che siano equivalenti dal punto di vista della precisione e qualità oppure attraverso procedure di stima statistica nei casi in cui una parte delle caratteristiche non sia stata rilevata per tutte le unità). Al comma 3, il regolamento dispone, inoltre, che al fine di ridurre l'onere che grava sui rispondenti, le autorità nazionali e la Commissione (Eurostat) hanno accesso, entro i limiti e le condizioni stabiliti da ciascun Stato membro e dalla Commissione, nei rispettivi settori di competenza, alle fonti di dati amministrative che interessano le varie sfere di attività delle loro amministrazioni pubbliche, nella misura in cui questi dati siano necessari ai requisiti di precisione richiamati all'articolo 6 (precisione). Quest'ultimo articolo stabilisce al comma 1 che gli Stati membri adottino le misure atte a garantire che i dati trasmessi rispecchino la struttura della popolazione delle unità statistiche e al comma 2 che la valutazione della qualità è effettuata mettendo a confronto i vantaggi derivanti dalla disponibilità dei dati con il costo della raccolta e l'onere gravante sulle imprese, in particolare le piccole imprese.

Riguardo a questi aspetti l'Eurostat ha avviato un importante progetto volto a modernizzare nel periodo 2009-2013 le statistiche sulle imprese in ambito Europeo denominato MEETS (*Modernisation of European Enterprises and Trade Statistics*), approvato dal Consiglio e Parlamento Europeo il 16 dicembre 2008, che si propone di finanziare numerosi obiettivi strategici che riguardano l'individuazione delle aree di minore importanza e nuove aree di interesse, migliorare la coerenza negli aspetti legali e nelle metodologie, incoraggiare le azioni che possono ridurre il fastidio statistico sulle imprese, favorire l'uso delle migliori pratiche nell'utilizzo dei dati disponibili attraverso operazioni di linkage e supportare l'uso dei dati amministrativi nei processi di produzione dell'informazione statistica. Per quest'ultimo progetto l'Eurostat ha promosso la costituzione dell'ESSnet "*The Use of Administrative and Accounts Data for Business Statistics*" (*ESSNet Admin Data*) che sotto la guida dell'Istituto Statistico del Regno Unito (ONS) e con la partecipazione degli Istituti di statistica dei Paesi che hanno interesse a collaborare nello specifico argomento (Italia, Portogallo, Spagna, Germania, Olanda, Belgio e Lituania), si propone di individuare e sviluppare pratiche raccomandate nell'uso dei dati amministrativi per migliorare l'efficienza dei processi di produzione delle statistiche economiche sulle imprese, da condividere con tutti gli altri Paesi Europei, partecipanti o meno al progetto. Lo scopo è quello di conseguire i vantaggi legati alla sinergia fra Paesi condividendo le buone pratiche a livello di intero sistema statistico europeo (ESS European Statistical System).

Obiettivo del regolamento sulle statistiche strutturali è quello di istituire un quadro comune per la raccolta, l'elaborazione e la trasmissione dei dati al fine della predisposizione di statistiche armonizzate per valutare la struttura, l'attività e la competitività delle imprese nella Comunità. In sintesi, esso prevede la trasmissione all'Eurostat dei dati definitivi della rilevazione campionaria sulle piccole e medie imprese e sull'esercizio di arti e professioni - PMI (rivolta alle imprese con 1-99 addetti) e di quella totale sul sistema dei conti delle imprese - SCI (che copre le imprese con

100 addetti ed oltre) entro 18 mesi dalla fine dell'anno di riferimento dei dati, con un dettaglio molto articolato:

- a) dati nazionali a livello di 4 cifre della NACE,¹ senza distinzione di classe di addetti;
- b) dati nazionali a livello di 3 cifre della NACE² per classi dimensionali di addetti;
- c) dati regionale per NUTS II a livello di 2 cifre (3 per il commercio) della NACE³ senza distinzione di classe di addetti.

Con riferimento alla rilevazione PMI, la ricerca si propone di fare una prima valutazione e trovare possibili soluzioni per l'utilizzo dei dati contabili disponibili negli archivi amministrativi nell'ambito del processo di stima dei parametri obiettivo dell'indagine ai vari livelli di dettaglio richiesti. Le azioni che si intendono sviluppare in questo progetto riguardano l'acquisizione in forma sistematica e standardizzata delle informazioni di fonte amministrativa, l'analisi della copertura delle fonti in termini di unità della popolazione obiettivo e delle variabili amministrative utili a fini statistici, la verifica delle corrispondenze delle definizioni delle variabili amministrative con quelle statistiche, la valutazione della loro utilizzabilità ai fini dell'integrazione delle mancate risposte parziali e totali della rilevazione, della costruzione di modelli di stima, nonché la verifica della coerenza delle stime rispetto a quanto prodotto con i metodi tradizionali per il regolamento SBS e, infine, lo sviluppo di metodi di integrazione e di modelli di stima delle variabili SBS non presenti nella modulistica amministrativa.

Il non soddisfacente tasso di risposta alla rilevazione abbinato alla disponibilità di un considerevole patrimonio informativo di fonte amministrativa (*Bilanci civilistici delle società di capitale, Studi di Settore, Modello Unico*) ha orientato l'attività di ricerca nella direzione di avere una migliore copertura del campione teorico, attraverso l'utilizzo delle corrispondenti informazioni presenti nelle diverse fonti amministrative, al fine di pervenire ad una nuova stima delle principali variabili SBS (fatturato, costi per beni e servizi, valore aggiunto, costo del lavoro e margine operativo lordo) e valutare la possibile presenza di distorsione o *bias* dovuto alla mancata risposta.

Il lavoro è strutturato come segue. Nel paragrafo 2 è discusso in generale il problema della valutazione della qualità dei dati amministrativi con riferimento ai loro possibili utilizzi a fini statistici, con riferimenti alle principali esperienze a livello internazionale e ad alcune metodologie per l'uso di dati da fonti diverse nel settore delle statistiche economiche sulle imprese. Il paragrafo 3 contiene una descrizione sintetica della rilevazione PMI, delle principali fonti amministrative utilizzabili nell'ambito del processo di produzione delle stime PMI, e della relativa copertura rispetto alla popolazione di riferimento per la rilevazione. In particolare, il paragrafo 3 contiene la descrizione di obiettivi, struttura, metodologie e risultati di uno studio sperimentale condotto utilizzando i dati PMI 2007 al fine di valutare gli effetti statistici dell'utilizzo di dati di fonte amministrativa nel processo produttivo della rilevazione PMI. Nello stesso paragrafo sono illustrati alcuni risultati dell'utilizzo effettivo di alcune fonti amministrative (*Bilanci civilistici e Studi di Settore*) nella rilevazione PMI 2008.

2. L'uso dei dati amministrativi per la produzione dell'informazione statistica sulle imprese

2.1. La qualità dei dati di fonte amministrativa

In ogni processo di produzione statistica, la qualità dei risultati prodotti dipende largamente dalla qualità dei dati di input, siano essi ottenuti mediante indagine diretta oppure da fonti di natura amministrativa o da altra fonte. Risulta pertanto evidente che, anche nell'ambito delle statistiche

¹ Rev.1.1 fino al 2007 e Rev.2 a partire dal 2008

² Rev.1.1 fino al 2007 e Rev.2 a partire dal 2008

³ Rev.1.1 fino al 2007 e Rev.2 a partire dal 2008

economiche, la valutazione della qualità dei dati amministrativi utilizzati a fini statistici (eventualmente in modo integrato con dati di indagine) rivesta un ruolo cruciale nell'ambito del processo di produzione. Analogamente al caso delle indagini dirette, l'uso dei dati di fonte amministrativa implica una fase di valutazione dei requisiti di qualità necessari affinché la fonte stessa possa essere ritenuta effettivamente utilizzabile a fini statistici. Questo implica evidentemente la definizione del concetto di qualità per questo tipo di dati.

La varietà di definizioni adottate in letteratura (Wallgren et al, 2007; FMI, 2004; Thomas, 2005; ONS (2005); Daas and Fonville, 2007; Daas et al., 2008) rivelano la complessità di delineare un framework condiviso a livello internazionale in questo contesto. Nel presente lavoro si fa riferimento alle definizioni proposte a livello di Sistema Statistico Europeo (ESS). Nel caso dei dati di indagine, Eurostat (2003a) ha definito il concetto di qualità sulla base di sei *dimensioni* principali: *rilevanza o pertinenza* (capacità delle informazioni di soddisfare le reali esigenze degli utenti); *accuratezza* (accostamento tra la stima del parametro di interesse della popolazione target e il valore vero e ignoto dello stesso); *tempestività e puntualità*; *accessibilità e chiarezza* (facilità con cui gli utenti possono accedere alle informazioni e disponibilità di metadati sui dati stessi); *comparabilità* (confrontabilità dell'informazione nello spazio e nel tempo); *coerenza* (adeguatezza delle informazioni ad essere combinate in modi diversi e per vari usi, anche in caso di uso di statistiche da fonti diverse).

Tenendo conto delle specifiche peculiarità del dato amministrativo, Eurostat (2003b) ha esteso ed integrato tali definizioni sulla base di elementi aggiuntivi. Ad esempio, il concetto di *rilevanza* risulta completato da un ulteriore elemento legato all'adeguatezza delle definizioni (sia per le unità appartenenti alla popolazione di riferimento, sia ai fenomeni su di esse osservati) adottate per le finalità di natura amministrativa rispetto a quelle adottate per le specifiche finalità statistiche. Ancora, la valutazione dell'*accuratezza* implica l'analisi dei contenuti informativi della fonte amministrativa per la verifica della sua adeguatezza statistica alla stima della popolazione di interesse: tale componente può essere influenzata anche dall'eventuale presenza nei dati di fonte amministrativa di errori di natura non campionaria (in particolare, errori di misura), che quindi devono essere oggetto di appropriate azioni di individuazione (controllo) e trattamento (correzione) prima del loro utilizzo a fini statistici. Inoltre, *accessibilità e chiarezza* implicano la necessità di instaurare protocolli stabili con gli Enti produttori del dato amministrativo per la regolamentazione dei flussi informativi, ponendo come requisito di base la trasparenza e disponibilità di metadati sulle attività di raccolta e trattamento dei dati. Il requisito di *tempestività* deve includere la proprietà di regolarità e puntualità con cui i dati amministrativi vengono rilasciati all'Ente produttore di informazioni statistiche. I concetti di *coerenza e consistenza* vanno considerati rispetto agli specifici requisiti dello specifico contesto di produzione statistica, in termini non solo di definizioni adottate, ma anche di conformità e completezza delle informazioni di fonte amministrativa rispetto alle specifiche finalità statistiche. Ulteriori elementi che contribuiscono a determinare il livello qualitativo di una fonte amministrativa, e che contribuiscono in larga misura a determinare il suo grado di utilizzabilità a fini statistici sono la *completezza*, la *stabilità* e i *costi*. La *completezza* va valutata in termini sia di variabili, sia di unità (*coverage*) presenti nella fonte. Sul versante delle variabili, va valutato in particolare:

- quali variabili sono presenti, e con quale livello di coerenza rispetto alle variabili obiettivo;
- quante e quali variabili (richieste dagli utenti esterni) non sono presenti nella fonte e debbono quindi essere opportunamente "stimate".

Dal lato delle unità, il problema del *coverage* può comportare la non disponibilità di informazioni per specifiche sottopopolazioni di interesse statistico. A questo aspetto, si aggiunge quello dell'*identificabilità* delle unità di interesse statistico, legato sostanzialmente al già citato problema della coerenza delle definizioni adottate per le unità di riferimento. La *stabilità* nel tempo di una fonte amministrativa può essere riferita sia ai contenuti, sia alle unità. I *costi* (in termini sia di risorse impiegate, sia di tipo organizzativo-strutturale) riguardano non soltanto l'acquisizione dei dati, ma anche il loro trattamento (predisposizione di strumenti di individuazione e trattamento di errori e mancate risposte, stima di variabili non disponibili dalla fonte amministrativa) e la loro diffusione (predisposizione di strumenti adeguati a garanzia della riservatezza dei dati). Un ulteriore elemento

di qualità di una fonte amministrativa è relativo all'entità della riduzione del *fastidio statistico sui rispondenti* che il suo uso comporta.

In caso di integrazione fra più fonti amministrative, la qualità dei dati integrati, e quindi il loro grado di utilizzabilità, dipende anche da ulteriori elementi, che si aggiungono ai precedenti, legati principalmente alle seguenti problematiche (Eurostat, 2003b; Eurostat, 2009):

- qualità del linkage (qualità e confrontabilità dei codici identificativi, e quindi rischio di duplicazioni, mancati accoppiamenti, ecc.);
- copertura delle fonti in termini di unità (le diverse fonti possono fare riferimento a diverse sottopopolazioni di imprese);
- completezza delle fonti in termini di unità (presenza di “mancate risposte totali” dovute ad esempio alla incompletezza della fonte dovuta al momento in cui l'archivio viene richiesto, oppure alla non obbligatorietà della fonte utilizzata), ovvero in termini di variabili (presenza di “mancate risposte parziali” dovute al fatto che alcune variabili possono non essere presenti in alcuna fonte);
- periodicità, tempestività e periodi di riferimento delle fonti integrate;
- coerenza (in termini di relazioni statistiche in particolare di natura multivariata) fra i contenuti (variabili) delle fonti oggetto di integrazione;
- metodologie adottate per la raccolta, il trattamento ecc. dei dati delle varie fonti.

Nello specifico ambito delle statistiche economiche, al fine di perfezionare ulteriormente il framework di riferimento pervenendo ad un quadro definitorio maggiormente condiviso, Eurostat sta promuovendo attività di analisi e sviluppo congiunte a livello di ESS anche attraverso il finanziamento di progetti Europei. È questo il caso del già citato *ESSNet Admin Data*, in cui è previsto tra l'altro lo sviluppo di un insieme di indicatori (quantitativi e qualitativi) per la valutazione della qualità delle fonti amministrative utilizzabili nell'ambito delle statistiche economiche.

2.1.1. La verifica della qualità dei dati di fonte amministrativa: controllo e correzione di errori e mancate risposte

In questo paragrafo ci si sofferma sugli aspetti di *accuratezza* e *completezza* dei dati amministrativi nel contesto delle statistiche economiche. In particolare, si focalizza l'attenzione sulla fase di verifica della presenza nei dati stessi di errori di natura non campionaria (*editing*), basata in generale sull'analisi delle incoerenze statistiche fra le informazioni disponibili a livello di dati elementari o di aggregati (distribuzioni, stime).

Un aspetto importante in questo contesto è relativo al fatto che nel caso di dati di fonte amministrativa, sia il processo di cattura delle informazioni sia l'eventuale fase di individuazione e trattamento degli errori e delle mancate risposte sono effettuati dall'Ente produttore dell'informazione amministrativa. Non essendo tali processi (e le relative metodologie) sotto il controllo dell'Ente produttore dell'informazione statistica, può risultare problematico valutare la qualità dei dati acquisiti in termini sia di entità e tipologia degli errori originariamente presenti, sia di impatto statistico delle attività di controllo e correzione effettuate.

Indipendentemente dalle caratteristiche iniziali dei dati acquisiti, il loro utilizzo presuppone comunque il disegno di appropriate procedure di verifica della qualità dei dati per il ripristino della loro coerenza e completezza rispetto ai criteri adottati per le specifiche finalità statistiche. È quindi evidente che, nell'ambito di un processo di produzione statistica in cui sia previsto l'uso di dati di fonte amministrativa, da un lato si registrerà una riduzione di tempi, costi e fastidio sui rispondenti dovuta alla disponibilità diretta di informazioni e/o alla possibilità di loro utilizzo per l'ottimizzazione del processo di controllo e trattamento degli errori, dall'altro si avrà un aumento di tempi e costi di produzione dovuto alla necessità di effettuare *editing* aggiuntivo sulle fonti utilizzate.

In generale, nell'ipotesi di assenza di problemi definitivi, i dati di fonte amministrativa possono essere affetti dalle stesse tipologie di errore non campionario caratteristiche dei dati di indagine

(Daas, 2008; Hoogland, 2010; Haitzman, 2010)⁴: errori di misura sistematici o casuali, valori anomali (*outlier*), errori influenti, errori di non risposta (totale o parziale), errori di misura che danno luogo ad incoerenze statistico-matematiche fra i dati. La presenza di tali errori influenza sia il livello di affidabilità delle informazioni contenute nella fonte, sia la significatività delle relazioni statistiche fra le variabili disponibili (anche ai fini del loro utilizzo in fase di ricostruzione dei dati mancanti), sia la stima di variabili o domini di imprese non coperti dalle fonti amministrative utilizzabili.

Analogamente al caso delle indagini statistiche, i controlli di coerenza dei dati possono essere effettuati a livello di microdati (analisi delle relazioni fra variabili in una stessa fonte, oppure ottenute da fonti diverse), a livello di distribuzioni (univariate o multivariate) oppure a livello di stime per prefissati domini di interesse. In caso di integrazione di più fonti amministrative, nell'ipotesi di assenza di problemi di *linkage*, si pone l'ulteriore problema della verifica della coerenza statistica fra informazioni provenienti da fonti diverse, sia a livello di singole variabili eventualmente disponibili in più di una fonte, sia a livello multivariato, cioè in termini di coerenza fra fenomeni misurati in fonti diverse. È inoltre evidente che, oltre alla dimensione trasversale, la verifica dei dati amministrativi dovrebbe prevedere un'analisi di tipo longitudinale, a livello di dati elementari, di distribuzioni, di stime ai prefissati livelli di dettaglio.

La verifica della coerenza di dati di fonte amministrativa può essere effettuata utilizzando metodologie analoghe a quelle adottate nel caso di dati di fonte statistica. Hoogland (2010) illustra una strategia complessiva di micro e *macro editing* adottata presso Statistics Netherlands per la verifica della qualità di dati amministrativi utilizzati per la produzione di statistiche economiche, del tutto analoga a strategie per dati economici adottate in rilevazioni statistiche: in particolare, vengono proposti un approccio di *editing selettivo* (Granquist *et al.*, 1997) per l'individuazione di errori influenti, una specifica strategia per l'individuazione degli errori di natura sistematica, un approccio di tipo *macro-editing* per l'individuazione dei valori anomali rispetto alle stime di riferimento. La presenza di valori anomali, dovuta ad esempio a errori di misura in fase di compilazione e/o registrazione dei moduli amministrativi, rappresenta in effetti una tipologia di errore che può essere affrontato con metodologie analoghe a quelle utilizzate nelle indagini statistiche (vedere tra gli altri, Lorenz, 2010; Elliott, 2010; Zach, 2007). Yung and Lys (2007) descrivono la strategia generale di *editing* e imputazione adottata in alcune indagini economiche presso Statistics Canada.

2.1.2. L'uso dei dati amministrativi per il miglioramento della qualità delle statistiche economiche

Diversi approcci sono stati proposti ed adottati in vari Istituti Nazionali di Statistica per il miglioramento dell'efficienza dei processi di produzione delle statistiche economiche sulle imprese (in termini di accuratezza, completezza, tempi, costi e fastidio sui rispondenti) attraverso l'utilizzo di informazioni sulle popolazioni di riferimento già disponibili in altre fonti, di natura statistica o amministrativa (vedere anche i paragrafi 2.2 e 2.3). Questi approcci possono essere riassunti nelle seguenti categorie generali:

1. uso indiretto, per migliorare la qualità dell'informazione statistica prodotta (a livello micro e/o aggregato) attraverso l'ottimizzazione della fase di individuazione degli errori di coerenza (*editing*), a livello micro e aggregato, di un processo di indagine statistica;
2. uso diretto, per migliorare la qualità dell'informazione statistica prodotta (a livello micro e/o aggregato) attraverso l'ottimizzazione della fase di imputazione delle mancate risposte in un processo di indagine statistica;
3. uso indiretto, per la stima di variabili/domini di unità (*coverage*) non direttamente disponibili nelle fonti amministrative;

⁴ Per una definizione formale delle tipologie di errore non campionario individuate in fase di *editing* di indagini economiche, vedere ad esempio Eurostat (2007).

4. uso diretto, per il disegno di indagini campionarie che sfruttino la disponibilità di dati di fonte amministrativa per migliorare l'efficienza del processo di stima in termini di trade-off fra di dimensione del campione (e quindi burden sulle imprese) e livello di precisione delle stime).

2.1.2.1. Uso del dato amministrativo per l'ottimizzazione della fase di editing di un processo di indagine statistica

In questa situazione, le informazioni disponibili da una o più fonti amministrative vengono utilizzate al fine di ottimizzare la fase di controllo degli errori non campionari presenti nei dati di indagine statistica. Uno specifico ambito di applicazione è rappresentato dall'individuazione degli errori influenti (Eurostat, 2007), dovuti ad errori di misura in fase di compilazione e/o registrazione, o di errata interpretazione dei quesiti, i quali possono essere causa di aumento consistente di distorsione nelle stime dei parametri obiettivo. Le tecniche di individuazione degli errori influenti, in particolare l'editing selettivo, sono basate sull'uso di opportuni predittori dei valori ottenuti sulla base di informazioni endogene (altre variabili osservate) o esogene (ad esempio, dati storici) legate alle variabili potenzialmente affette da errore. È evidente che la disponibilità di predittori accurati consente di ottimizzare la ricerca degli errori influenti, cioè del sottoinsieme di valori/unità che, qualora errati, hanno un impatto significativo sulle stime dei parametri obiettivo in termini di distorsione aggiuntiva generata, e che quindi devono essere controllati (manualmente) al fine di garantire stime finali di prefissata accuratezza.

In questa ottica, i dati amministrativi possono essere utilizzati, direttamente o tramite l'uso di opportuni modelli (Lewis, 2010; Elliott, 2010), per ottenere valori attesi più accurati dei valori osservati, in base ai quali stilare la graduatoria delle unità influenti da sottoporre a revisione. Questo consente o di aumentare la qualità dei risultati, o di ottenere risultati di qualità analoga al caso di non utilizzo dell'informazione amministrativa, ma con costi e tempi inferiori.

2.1.2.2. Uso del dato amministrativo per l'ottimizzazione della fase di imputazione in un processo di indagine statistica

Nel caso in cui si utilizzino i dati di fonte amministrativa come predittori dei valori mancanti, si ottengono immediati vantaggi in termini sia di accuratezza delle imputazioni (e quindi dei risultati statistici), sia di tempestività, costi e riduzione del carico sui rispondenti. A causa di possibili discrepanze a livello di definizioni di variabili e/o unità osservate fra fonti statistiche e fonti amministrative, è preferibile procedere all'imputazione dei valori mancanti utilizzando non direttamente i dati amministrativi, bensì valori predetti sulla base di opportuni modelli stimati in base alle informazioni di fonte statistica e amministrativa. In genere, in letteratura vengono proposti modelli (robusti) di tipo regressivo (Lewis, 2010; Haitzmann, 2010; Zach, 2010; Lorenz, 2010).

2.1.2.3. Uso del dato amministrativo per la stima di variabili/domini non direttamente disponibili nelle fonti amministrative

Come già detto in precedenza, nel caso in cui si utilizzi come fonte primaria di informazione per la stima di fenomeni economici una o più fonti amministrative, si può porre sia un problema di non disponibilità nelle fonti (o di non utilizzabilità per insufficiente accuratezza) di particolari sottoinsiemi di variabili invece necessarie ai fini statistici, e/o di mancata copertura di specifiche sottopopolazioni (ad esempio, micro-imprese) della popolazione target (not covered units). In queste situazioni, è necessario integrare l'uso delle fonti amministrative con l'adozione di modelli e/o di indagini campionarie dirette per la stima delle informazioni non direttamente disponibili o dei sottodomini di unità non inclusi nel sistema di fonti utilizzate.

In questo contesto, in letteratura sono delineati, alcuni scenari diversi, in cui i dati amministrativi sono utilizzati in combinazione con l'uso di modelli e/o di (sistemi di) indagini dirette per la stima delle informazioni mancanti. Alcune esperienze (Elliott et al., 2010; Yung and Lys, 2007; Zach, 2007) sono basate sull'uso combinato di indagini ad hoc per la rilevazione di variabili non disponibili nelle fonti amministrative su parte della popolazione obiettivo: per i sottoinsiemi di imprese esclusi dalla rilevazione (Take none strata), le informazioni mancanti vengono previste sulla base

di modelli stimati sulla parte osservata della popolazione target. In altre applicazioni (Elliott et al., 2010) si propone la rilevazione diretta di tutte le unità della popolazione, ma alle unità appartenenti ad alcuni specifici strati (Ask some strata) non sono richieste alcune particolari informazioni non direttamente disponibili da fonti amministrative: anche in questo caso, i valori delle variabili non osservate in questi specifici strati sono predetti sulla base di modelli stimati utilizzando dati di indagine e dati amministrativi. Una significativa riduzione della frazione di campionamento (a fronte di stime accurate dei parametri obiettivo ai diversi livelli di dettaglio previsti) per le indagini dirette disegnate in modo complementare all'uso del dato amministrativo, può essere prospettata anche sulla base dell'utilizzo di tecniche di stima per piccole aree (Rao, 2003), in cui la stima dei parametri della popolazione per le variabili non direttamente disponibili dalle fonti amministrative avviene sfruttando in modo integrato i dati di indagine e i dati amministrativi.

2.1.2.4. *Uso del dato amministrativo per la stima efficiente dei parametri obiettivo*

In questo contesto (si veda ad esempio Falorsi et al., 2004), si propongono strategie di stima basate sulla progettazione integrata di disegni di indagine che tengano conto il più possibile della disponibilità di informazioni ausiliarie provenienti da una o più fonti amministrative. L'obiettivo è quello di ottenere una maggiore efficienza del processo di stima (riduzione della dimensione del campione a parità di precisione delle stime, riduzione delle variabili da rilevare, migliore coerenza del processo di stima rispetto ai totali noti (per una discussione più approfondita di questi approcci, vedere il paragrafo 2.3).

2.2. L'uso dei dati di fonte amministrativa nella stima degli aggregati SBS in Europa

In ambito europeo (nel quale includiamo anche i paesi di area EFTA) l'utilizzo dei dati amministrativi per la produzione degli aggregati SBS è ovviamente molto diffuso, ma in maniera piuttosto diversificata.⁵ Infatti, se pochi sono i paesi che utilizzano esclusivamente la fonte amministrativa (tra questi Portogallo e Spagna) o che usano esclusivamente la fonte d'indagine (tra cui ad esempio la Svizzera), la maggior parte dei paesi combina le due fonti ma quasi sempre con modalità e finalità specifiche di ciascun paese. Questo dipende, ovviamente, dalla disponibilità di fonti amministrative che varia di paese in paese, ma anche dalle scelte effettuate per definire il processo di produzione degli aggregati SBS. In particolare, in questa sede è interessante rilevare quanti paesi utilizzano i dati amministrativi come fonte primaria per la produzione delle statistiche strutturali sulle imprese e quanti come fonte ausiliaria per le fasi di *editing* e validazione dei risultati dell'indagine. Un'ulteriore distinzione può essere presa in considerazione in merito ai dati dei registri d'impresa (Business Register, BR) utilizzati come frame relativamente ad alcune variabili (come ad esempio la classificazione dell'attività economica) o per definire i pesi per il riporto all'universo.

Inclusi i tre citati in precedenza, circa un terzo dei paesi utilizza la fonte amministrativa come primaria almeno per una parte della rilevazione (tipicamente per le imprese di piccole dimensioni). Circa un quarto dei paesi (tra cui Germania e Regno Unito) dichiarano di utilizzare esclusivamente dati di registro e di indagine almeno per gli Annessi 1-4 del regolamento SBS. Un'analoga proporzione di paesi, tra cui l'Italia, utilizzano la fonte amministrativa come secondaria per integrare i risultati d'indagine nel trattamento delle mancate risposte e in fase di correzione/validazione dei dati rilevati. Per alcuni paesi non sono state ottenute informazioni sufficienti per classificarli.

Nella maggioranza dei casi le indagini sono esaustive per le imprese più grandi (rispetto a soglie variabili per i diversi paesi) e campionarie per le piccole o medio/piccole.

Negli ultimi anni, Eurostat ha sostenuto e promosso l'utilizzo delle fonti amministrative per la produzione statistica dei paesi del Sistema Statistico Europeo (ESS) sia attraverso finanziamenti diret-

⁵ Per il presente paragrafo sono stati consultati essenzialmente i Quality report e i National Methodologies disponibili sul sito Eurostat's Metadata Server, all'indirizzo <http://ec.europa.eu/eurostat/ramon/>. Per alcuni paesi le informazioni sono state tratte dai documenti presenti nei relativi siti degli Istituti Nazionali di Statistica.

ti a progetti di ricerca (ad esempio, gli ESSNet nell'ambito del programma MEETS)⁶ sia sostenendo iniziative di tipo seminariale sul tema dell'uso dei dati amministrativi nella produzione di informazione statistica. Da una rassegna, seppur non esaustiva, dei lavori presentati ad alcuni di tali convegni⁷ appare chiaro che l'uso del dato amministrativo nella produzione di statistiche sulle imprese sta evolvendo in senso positivo dal punto di vista sia qualitativo sia quantitativo. Sempre più Paesi, infatti, utilizzano (ad esempio Svezia,⁸ Finlandia,⁹ Portogallo¹⁰ o progettano di utilizzare (Francia)¹¹ la fonte amministrativa come primaria per le statistiche strutturali sulle imprese (non solo relativamente alle indagini SBS), relegando eventualmente le indagini dirette ad una porzione della popolazione o ad un uso specifico di controllo e/o di rilevazione di informazioni non presenti nella fonte primaria.

Si registra anche, in virtù dell'apprezzamento che una tale pratica ha comportato da parte sia dei produttori di statistiche ufficiali (migliore qualità e tempestività delle informazioni, riduzione dei costi) sia delle imprese fornitrici dei dati (riduzione del disturbo statistico), una tendenza ad un uso della fonte amministrativa ancora più integrato e regolato. Più esplicitamente, nei rapporti con l'esterno le sinergie fra i produttori di statistica e gli enti responsabili degli archivi amministrativi portano ad una maggiore uniformità di intenti e conseguentemente ad una prospettiva di un migliore (o più gestibile) adeguamento degli archivi amministrativi alle finalità statistiche in termini sia di definizioni che di contenuti. Tali sinergie possono spingersi, come nel caso del Portogallo, a delegare interamente la raccolta dei dati SBS all'amministrazione fiscale.¹²

Tale evoluzione ha richiesto non solo un'interazione tra i produttori di statistica e gli enti preposti alla raccolta dei dati amministrativi ma anche un adeguamento del contesto normativo (si pensi ad esempio alle norme per il trattamento e la protezione dei dati individuali). In Italia un passo importante è costituito dal decreto di riordino dell'Istituto, in fase di approvazione, che prevede esplicitamente il ruolo dell'Istat nel coordinamento delle modulistica amministrativa, che costituirà la base per interventi efficaci presso gli Enti titolari.

Relativamente all'organizzazione interna degli istituti di statistica, invece, si rilevano alcune proposte (Germania,¹³ Finlandia e lo stesso Portogallo) di organizzazione dei dati in un'unica struttura informativa (e informatica) che integri le fonti amministrative e quelle d'indagine per la produzione di tutte le statistiche sulle imprese, altrimenti demandate ad indagini diverse.

Una tale organizzazione viene vista come una possibilità di ottimizzare i processi produttivi (in termini sia di qualità, attraverso la riduzione delle mancate risposte, sia di riduzione dei costi e del fastidio statistico sui rispondenti) risolvendo al contempo problemi di coerenza di informazioni (uguali o simili) provenienti da indagini diverse.

La letteratura su questo aspetto appare, comunque, piuttosto limitata e orientata alla gestione dei registri di imprese piuttosto che alla conduzione di indagini statistiche mentre, naturalmente, si rivela più vasta quando si tratta di analizzare i processi produttivi di singole indagini.

⁶ Nell'ambito del Programme for the Modernisation of European Enterprises and Trade Statistics (MEETS) è previsto il finanziamento di diversi progetti ESSNet (ESS Network). Attualmente è in corso, con la partecipazione dell'Istat, il già citato ESSNet on the use of administrative and accounts data for business statistics. Un report dettagliato sull'uso dei dati amministrativi per la produzione di statistiche economiche è di prossima pubblicazione come risultato del Work Package 1 "Overview of MSs' Existing Practices in the Uses of Administrative Data for Business Statistics".

⁷ IAOS - International Association for Official Statistics (IAOS is a section of the ISI, <http://isi.cbs.nl/iaos/>). ICES II - International Conference on Establishment Surveys (<http://www.eia.doe.gov/ices2/>). METIS Workshop on the Statistical Business Process and Case Studies, Lisbon, 11-13 March 2009, (<http://www.unece.org/stats/documents/2009.03.metis.htm>). UNECE - Work session on statistical data editing, Neuchâtel, 5-7 October 2009 (<http://www.unece.org/stats/documents/2009.10.sde.htm>). Reengineering of business statistics, 11-12 October 2007, Lisbon, (http://www.ine.pt/ine_novidades/semin/rbs/apresentacao.html). Q2008 - The European Conference on Quality in Official Statistics, 9 - 11 July, Rome (<http://q2008.istat.it/>). ESSW09 - Efficient methodology for producing high quality establishment statistics, 7-9 September 2009, Stockholm, (<http://www.eesw09.eu/>). Wiesbaden Group on BR - International Roundtable on Business Survey Frames (<http://circa.europa.eu/irc/dsis/businesssurvey/info/data/links.htm>). OECD - SBS Expert Meeting: "Towards better Structural Business and SME Statistics", 3-4 November 2005, Paris. "Using Administrative Data in the Production of Business Statistics - Member States Experiences" Istat Rome, 18-19 marzo 2010, <http://essnet-admin-data-seminar-rome.ine.pt/>.

⁸ cfr Berg et al., 2006; Wallgren e Wallgren, 2006.

⁹ cfr. Orjala, 2008; Tolkki, 2007.

¹⁰ cfr. Chumbau et al., 2008.

¹¹ cfr. Brion e Gros, 2009.

¹² cfr. Chumbau et al., 2010.

¹³ cfr. R. Sturm (2007).

2.3. Disegni di campionamento integrati all'uso di fonti amministrative

Come già accennato, l'Istat ha avviato già da alcuni anni un progressivo sfruttamento delle informazioni di natura amministrativa esistenti sul sistema delle imprese, allo scopo di integrare e/o sostituire le rilevazioni dirette per quel che riguarda le informazioni economiche già raccolte dalla Pubblica amministrazione (fra i più recenti, vedere ad esempio Biffignandi et al., 2009; Luzi et al., 2010).

Recentemente in ambito Istat sono stati proposti approcci alternativi all'utilizzo delle informazioni di fonte amministrativa, volti a definire nuove strategie di campionamento delle rilevazioni statistiche sulle imprese che tengano conto della molteplicità delle fonti di dati sulla popolazione di interesse; il contesto è quello in cui ciascuna fonte è relativa ad una specifica sottopopolazione e i dati in essa disponibili possono rappresentare sia direttamente le variabili di interesse, sia variabili ausiliarie utili per la specificazione di modelli statistici per la predizione dei valori delle variabili di interesse (Falorsi et al., 2005).

Volendo descrivere in sintesi la logica alla base delle strategie di campionamento che combinano l'approccio campionario tradizionale e l'uso di dati amministrativi (Falorsi et al., 2004), si consideri una fonte amministrativa a , che copre un sottoinsieme della popolazione di interesse U e fornisce con buona accuratezza un sottoinsieme V_a delle variabili d'interesse.

Si definiscano inoltre:

- U_a ($U_a \in U$) la sottopopolazione di imprese coperte da a ;
- $U_{\bar{a}}$ ($U_{\bar{a}} \in U$) il sottoinsieme complementare;
- y_k il valore della variabile di interesse per l'unità k ;
- $\mathbf{y}_{a,k}$ il vettore colonna ($k \in U_a$), contenente tutti i valori delle variabili d'interesse di V_a .

Se $y \in V_a$ il parametro di interesse Y si può esprimere come:

$$Y = \sum_{k \in U} y_k = \sum_{k \in U_a} y_k + \sum_{k \in U_{\bar{a}}} y_k = Y_a + Y_{\bar{a}} \quad (1)$$

in cui il valore y_k , per $k \in U_a$, può essere ottenuto direttamente dalla fonte amministrativa a .

Se $y \notin V_a$ il parametro di interesse Y si può esprimere come:

$$Y = \sum_{k \in U} y_k = \boldsymbol{\beta}' \mathbf{Y}_a + \sum_{k \in U_{\bar{a}}} \varepsilon_k + Y_{\bar{a}} \quad (2)$$

in cui: $\mathbf{Y}_a = \sum_{k \in U_a} \mathbf{y}_{a,k}$; $\boldsymbol{\beta}$ e ε_k rappresentano, rispettivamente, un vettore di regressori fissi incogniti e i residui casuali di un modello lineare.

Il modello precedente, per $k \in U_a$, mette in relazione le variabili di interesse y_k , con il vettore di predittori $\mathbf{y}_{a,k}$ dove:

$$y_k = \boldsymbol{\beta}' \mathbf{y}_{a,k} + e_k, \quad (\text{per } k \in U_a) \quad (3)$$

I valori attesi dei residui sono $E(e_k) = 0$, $E(e_k^2) = c_k \sigma^2$, $E(e_k^2, e_{k'}) = 0$ (per $k' \neq k$), essendo c_k uno scalare noto.

La (1) e la (2) consentono di definire la nuova strategia di campionamento le cui caratteristiche principali sono le seguenti:

- se $y \in V_a$ e $k \in U_a$, il valore y_k viene rilevato dalla fonte amministrativa a ;
- i dati campionari consentono di stimare β e $Y_{\bar{a}}$.

Se la variabile d'interesse **appartiene** a V_a , la stima del totale può essere ottenuta come:

$$\tilde{Y} = Y_a + \tilde{Y}_{\bar{a}} = Y_a + \sum_{k \in \mathcal{S} \cap U_{\bar{a}}} y_k / \pi_k \quad (4)$$

dove π_k indica la probabilità di inclusione dell'unità k . Si noti che la variabilità relativa alla componente Y_a è nulla.

Se la variabile d'interesse **non appartiene** a V_a , si può stimare la componente Y_a mediante il seguente approccio predittivo:

$$\tilde{Y} = \sum_{k \in \mathcal{S} \cap U_a} y_k + \tilde{\beta}' \sum_{k \in \mathcal{S} \cap U_a} \mathbf{y}_{a,k} + \tilde{Y}_{\bar{a}} \quad (5)$$

in cui $\mathcal{S} \cap U_a$ indica il sottoinsieme di U_a non incluso nel campione; $\tilde{\beta}$ è la stima campionaria di β .

L'implementazione della strategia illustrata prevede la ridefinizione del disegno campionario con l'adozione di due differenti criteri di stratificazione per le sottopopolazioni U_a e $U_{\bar{a}}$.

Poiché parte dei totali delle principali variabili di interesse non sarebbe affetta da errore di campionamento in quanto ottenuta direttamente da fonte amministrativa, l'uso congiunto delle due strategie (specificate rispettivamente su U_a e $U_{\bar{a}}$) permetterebbe un considerevole incremento dell'accuratezza delle stime ed un conseguente miglioramento della qualità dei dati.

In ambito internazionale, approcci interessanti (anche se parziali, in quanto applicabili solo in fase di stima) sono quelli che prevedono l'utilizzo di metodi di programmazione matematica (Bikker et al., 2010) per assicurare la coerenza tra stime su insiemi di dati elementari derivanti dall'integrazione di registri amministrativi e uno o più file di indagine.

Tra le tecniche già sperimentate, anche se finora solo nel campo delle statistiche sociali, il metodo dei *Repeated Weighting* (RW) (Houbiers et al., 2010), proposto dall'Istituto di Statistica Olandese, sfrutta le proprietà di calibrazione dello stimatore di regressione generalizzato (Greg) per garantire la consistenza numerica tra le stime relative a differenti domini, anche quando ciascuno di essi può essere stimato a partire da micro dati provenienti da indagini distinte (Van Duin et al., 2010).

In particolare, il RW è stato sviluppato per quelle situazioni in cui l'informazione su una o più variabili di interesse è disponibile da più di una fonte. Siano pertanto s_1 e s_2 i campioni relativi a due indagini estratte dalla stessa popolazione U ; viene denominato *data block* il sottoinsieme dei dati ottenuto dall'intersezione di U , s_1 e s_2 , ossia il sottoinsieme di unità campionarie per le quali si osservano le stesse variabili di interesse.

Il metodo del *Repeated Weighting* si caratterizza per l'uso iterativo dello stimatore di regressione; l'obiettivo è produrre un unico vettore di pesi finali w_i , a partire dai pesi diretti d_i , per la stima di una certa variabile di interesse su molteplici domini di studio; ciascun dominio è identificato

da una tabella marginale T_m (detta *table set* o *table margin*) definita sul data block. Il principio alla base del RW è quello di ricalibrare i pesi con cui una data tabella T_m è stimata, ponendo come vincoli le marginali delle tabelle T_p già stimate in precedenza.

Sia y una variabile il cui parametro di popolazione (totale o media) debba essere stimato, relativamente ad un tipo di dominio rappresentato dalla marginale T_m , attraverso un insieme di variabili esplicative \mathbf{x} disponibili sul data block. Come è noto, la stima Greg del totale di y è definita da:

$$\tilde{Y}_{Greg} = Y_p + \tilde{\beta}'_s (X_p - \tilde{X}_d) \quad (6)$$

in cui:

- X_p e Y_p sono i totali di popolazione rispettivamente di \mathbf{x} e di y ;
- \tilde{X}_d e \tilde{Y}_d sono le corrispondenti stime dirette;
- X_s è la matrice delle osservazioni campionarie sul set di variabili ausiliarie \mathbf{x} ;
- y_s è il vettore delle osservazioni campionarie sulla variabile di interesse y ;
- $\beta_s = (X'_s D_s X_s)^{-1} X'_s D_s y_s$ è il vettore dei coefficienti di regressione stimati, essendo $D_s = \text{diag}(d_1, \dots, d_n)$.

Anziché la forma precedente, la procedura del *Repeated Weighting* utilizza per il set di coefficienti di regressione l'espressione:

$$\beta_w = (Z'_s W_s Z_s)^{-1} Z'_s W_s y_s, \text{ con } W_s = \text{diag}(w_1, \dots, w_n) \quad (7)$$

dove Z_s è la matrice delle osservazioni campionarie sulle variabili ausiliarie contenute nella tabella marginale, essendo i totali delle variabili marginali \mathbf{z} già stimati in un passo precedente o essendo noti da registro. Indicando tali stime o totali noti con \tilde{Z}_{RW} , la stima RW di Y è definita da:

$$\tilde{Y}_{RW} = \tilde{Y}_{Greg} + \tilde{\beta}'_w (\tilde{Z}_{RW} - \tilde{Z}_{Greg}) \quad (8)$$

Si può dimostrare, sostituendo (6) in (7), che i pesi così ottenuti sono tali da rendere le stime per la tabella marginale T_m coerente con le stime già determinate per le tabelle marginali calcolate, relativamente a dati domini di studio, ai passi precedenti.

Se più indagini abbinate ad uno stesso registro amministrativo hanno variabili in comune, è opportuno creare un data block rettangolare dall'unione dei record delle stesse indagini, in quanto sarà possibile stimare con maggiore accuratezza, mediante RW, le tabelle costruite su tali variabili comuni (ovviamente dopo una fase di correzione ed armonizzazione dei dati raccolti dalle indagini).

Un punto critico della procedura RW è l'ordine con cui le tabelle vengono stimate, in quanto i pesi calcolati in corrispondenza di ciascuna tabella sono dipendenti dalla specificazione del modello sottostante lo stimatore; quest'ultimo, a sua volta, dipende dalle tabelle già stimate nei passi precedenti (e quindi dalla scelta del vettore \mathbf{z} ad ogni passo della procedura).

Un altro aspetto da considerare è legato alla presenza di celle vuote, ossia di strati o domini con nessun rispondente nel file di indagine, che non permette di trovare una soluzione al problema di stima del parametro di interesse in quello strato o dominio.

Il problema della coerenza delle stime tra un insieme di domini distinti, quando alcuni di essi presentano pochi o nessun rispondente, può essere affrontato (Singh e Mohl, 1996) con l'applicazione dello stimatore *Greg modificato* (M-Greg; Rao, 2003).

Tale stimatore rientra nella classe degli stimatori diretti; è sufficientemente facile da implementare ed è stato già sperimentato con successo in Istat (De Vitiis et al., 2009) per la costruzione degli aggregati di benchmark da fornire ad Eurostat, in base ai Regolamenti SBS e STS, per differenti variabili economiche e per diverse tipologie di dominio, in seguito alla riclassificazione delle imprese rispondenti alle indagini strutturali secondo la Nace Rev.2; rispetto a tale contesto applicativo, lo stimatore è caratterizzato da tre proprietà particolarmente utili:

- la stima è ottenibile per ciascun dominio definito secondo la Nace Rev.2, anche se non è presente nessuna unità campionaria nel dominio;
- lo stimatore è calibrato rispetto ai totali di archivio delle variabili ausiliarie per ciascuno dei domini T_2 definiti secondo la Nace Rev.2;
- per ciascuna variabile di interesse le stime calcolate con lo stimatore M-Greg a livello dei domini T_1 costruiti a un certo livello della Nace Rev.1 coincidono con le stime già diffuse a tale livello, a meno di un'approssimazione generalmente trascurabile.

Il calcolo dei coefficienti finali dello stimatore si ottiene mediante la soluzione di minimo vincolato (Deville e Särndal, 1992). In termini molto generali, il set dei coefficienti finali minimizzano nel loro complesso la distanza con i pesi iniziali garantendo che le stime sulle variabili di calibrazione siano uguali ai totali noti. Lo stimatore M-Greg usa una distanza appartenente alla classe delle distanze di tipo lineare; nel calcolo dei pesi finali possono sorgere problemi di natura computazionale, in quanto lo stimatore M-Greg compie una calibrazione per ogni combinazione presente a livello di popolazione tra i domini T_1 e T_2 . Per ridurre la mole di calcolo si può ridurre il sistema di vincoli sui totali che lo stimatore M-Greg deve rispettare.

3. L'uso delle fonti amministrative per l'integrazione delle mancate risposte nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI): uno studio sperimentale

In questo capitolo viene descritto uno studio volto a verificare la qualità e l'utilizzabilità di alcune fonti amministrative disponibili sulle imprese nell'ambito del processo di stima della rilevazione campionaria Istat sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI).

In particolare, sulla base delle caratteristiche dell'indagine (campo di osservazione, contenuti, ecc.) e dell'attuale processo di produzione della rilevazione PMI, con riferimento ad uno specifico anno di riferimento, vengono dapprima prese in esame le caratteristiche di alcune fonti amministrative (contenuti, copertura, ecc.), quindi viene simulata e valutata una procedura di stima basata sull'integrazione delle informazioni mancanti da fonte amministrativa.

3.1 La rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

La rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI), di natura campionaria, ha come campo di osservazione le imprese con 1-99 addetti appartenenti ai settori di attività economica industriale, commerciale e dei servizi alle imprese ed alle famiglie e risponde alle esigenze richieste dal regolamento comunitario sulle statistiche strutturali (SBS).

Le informazioni sono raccolte attraverso la compilazione di un questionario molto complesso per la natura ed il dettaglio delle variabili richieste (valore e costi della produzione in linea con la IV direttiva comunitaria, occupazione per categoria professionale e sesso, costo del lavoro del personale dipendente disaggregato nelle diverse voci componenti, personale esterno alle imprese con i

relativi costi, investimenti effettuati nell'esercizio per categoria di beni, spese per la protezione dell'ambiente e informazioni di natura prettamente qualitativa finalizzate a rilevare particolari aspetti emergenti) essendo esso utile a soddisfare il regolamento strutturale sulle imprese SBS, le esigenze della Contabilità Nazionale e a conoscere in forma più organica il contesto produttivo del Paese. Le imprese oggetto di rilevazione ricevono per posta la lettera di presentazione della rilevazione con il codice impresa e la password per accedere al sito Web dedicato, da cui è possibile scaricare il questionario in forma elettronica che può essere trasmesso all'Istat per via telematica.

Il disegno di campionamento adottato è ad uno stadio stratificato, con selezione delle unità con probabilità uguali (per un dettaglio maggiore vedere l'appendice A.2.); gli strati sono definiti dalla concatenazione delle modalità delle variabili 'regione', 'classi di attività economica' (Ateco a 4 cifre) e 'classi di addetti'. La lista delle unità campionate è estratta dall'archivio ASIA (Archivio Statistico delle Imprese Attive), costruito sulla base dell'integrazione di varie fonti, di carattere sia amministrativo sia statistico.

In questo lavoro è stata presa in esame la rilevazione PMI2007 per la quale è stato estratto un campione pari a 103.039 imprese; i dati delle imprese rispondenti (37.330) sono stati sottoposti a revisione, a controlli di coerenza e compatibilità, a tecniche di trattamento delle mancate risposte parziali e delle osservazioni anomale. La numerosità effettiva di imprese utilizzate in fase di stima è risultato, tuttavia, pari a 43.701 unità in quanto è stata effettuata la ricostruzione delle mancate risposte totali per le imprese delle fasce dimensionali superiori, in particolare da 20 a 99 addetti, utilizzando le informazioni disponibili dalla fonte *Bilanci civilistici delle società di capitale*, in base alle quale sono state integrate 6.371 imprese. Per la classe dimensionale da 1 a 19 addetti non si è effettuata l'integrazione con la fonte Bilanci civilistici (se non marginalmente in settori caratterizzati da una larga presenza di società di capitale) per la rilevante presenza di altre forme di impresa o di attività svolte nella forma di lavoro autonomo. Per questa classe dimensionale si è preferito, quindi, compensare la mancata risposta totale in fase di stima, mediante calibrazione, anche al fine di evitare un eccessivo effetto distorsivo sulle stime stesse derivante dall'utilizzo di tecniche di imputazione che avrebbero potuto produrre uno sbilanciamento del campione verso la forma giuridica di società di capitali.

Queste riflessioni hanno, tuttavia, stimolato l'attività di ricerca ad investigare su un più ampio utilizzo delle diverse fonti amministrative sulle imprese appartenenti alle classi dimensionali inferiori, così come è documentato nei successivi paragrafi.

Per quanto riguarda la fonte informativa dei Bilanci civilistici, la rilevazione PMI ha iniziato a utilizzare tale fonte a partire dall'anno di riferimento 2001, al fine di integrare le mancate risposte totali per le imprese in forma di società di capitale della sola fascia dimensionale 50-99 addetti comprese nel campione (Tavola 1). Tenuto conto dei buoni risultati di questa valutazione, dall'anno successivo il metodo di integrazione è stato applicato anche alle imprese non rispondenti appartenenti alla fascia dimensionale 20-49 e, marginalmente, a quelle presenti nella fascia dimensionale inferiore.

La fase di integrazione delle mancate risposte è basata sull'uso combinato di dati di fonte amministrativa (Bilanci civilistici) e della tecnica del donatore (sulla base di un criterio di vicinanza che si basa sull'attività economica, la fascia dimensionale e la localizzazione territoriale). In particolare, i dati dei Bilanci civilistici sono utilizzati per la ricostruzione delle principali voci del conto economico, ovvero il fatturato, le variazioni delle rimanenze di prodotti finiti, in corso di lavorazione e semilavorati, le variazioni dei lavori in corso su ordinazioni, gli incrementi di immobilizzazioni per lavori interni, gli altri ricavi e proventi, i costi per acquisto di materie prime, sussidiarie e di consumo, i costi per servizi, i costi per godimento di beni di terzi, i costi per il personale, le quote di ammortamento, le variazioni delle materie prime sussidiarie, di consumo e di merci, gli accantonamenti, gli oneri diversi di gestione e l'utile o la perdita di esercizio (per un dettaglio maggiore vedere l'appendice A.2.).

Nella rilevazione PMI2007 l'integrazione delle mancate risposte totali è stata effettuata per le imprese del campione della fascia dimensionale 10-99 addetti.

La metodologia utilizzata per il riporto dei dati campionari all'universo si basa sugli *stimatori di ponderazione vincolata* (per un dettaglio maggiore vedere l'appendice A.2.); la metodologia di stima consente di calcolare pesi finali che, sotto certe ipotesi, risultano correttivi delle mancate risposte totali residue e della sottocopertura della lista. Essi assicurano il rispetto di uguaglianza fra alcuni totali noti dell'universo di riferimento (imprese e addetti) e le stime campionarie.

Tavola 1 - Imprese integrate nella rilevazione PMI per classe di addetti - Anni 2001-2007

ANNO	1 – 19 addetti			20 – 49 addetti			50 – 99 addetti		
	Campione	Integrate	%	Campione	Integrate	%	Campione	Integrate	%
2001	96.187	-	-	13.329	-	-	6.355	2.090	32,9
2002	102.184	-	-	13.091	2.977	22,7	5.962	1.836	30,8
2003	103.038	-	-	13.465	3.551	26,4	5.933	2.134	36
2004	89.499	-	-	12.821	3.765	29,4	6.079	2.288	37,6
2005	89.168	-	-	12.892	3.227	25	6.090	1.991	32,7
2006	85.705	275	0,3	12.942	3.966	30,6	5.998	2.037	34
2007	84.143	458	0,5	12.751	3.751	29,4	6.145	2.162	35,2

Il riporto dei dati all'universo, legato anche alle esigenze del regolamento SBS, è realizzato per i seguenti domini:

- per classe di attività economica della classificazione Ateco2002 a quattro cifre senza distinzione per classi di addetti (circa 460 domini di stima);
- per gruppo di attività economica Ateco2002 a tre cifre e fascia dimensionale delle imprese (1-9, 10-19, 20-49 e 50-99 addetti per le imprese industriali e per le imprese che operano nell'intermediazione monetaria e finanziaria, nell'istruzione, nella sanità ed in altri servizi pubblici, sociali e personali; 1, 2-9, 10-19, 20-49 e 50-99 addetti per le imprese commerciali, alberghiere e di ristorazione, di trasporto e comunicazione, immobiliari, di noleggio, informatiche, professionali ed imprenditoriali) (circa 1.050 domini di stima);
- per divisione di attività economica Ateco2002 a due cifre e regione (circa 984 domini di stima).

Tavola 2 - Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni - Anno 2007

SEZIONI DI ATTIVITÀ ECONOMICA (ATECO2002)	Campione teorico	Imprese rispondenti con questionario valido e integrato da fonte bilanci civilistici	Tasso di risposta
C - Estrazione di minerali	1.027	425	38,5
D - Attività manifatturiere	39.215	16.845	39,4
E - Produzione e distribuzione di energia elettrica, gas e acqua	1.130	540	49,2
F - Costruzioni	4.882	2.066	40,8
G - Commercio	18.344	8.400	45,8
H - Alberghi e ristoranti	2.960	1.066	37,2
I - Trasporto	7.142	2.530	38,3
J (67) - Attività ausiliarie dell'intermediazione finanziaria e delle assicurazioni	1.453	598	49,2
K - Attività immobiliari, noleggio, informatica, ricerca e servizi alle imprese	16.483	7.202	46,8
M - Istruzione	1.828	674	41,4
N - Sanità e assistenza sociale	1.805	873	49,5
O - Altri servizi pubblici, sociali e personali	6.770	2.482	39,2
CLASSI DI ADDETTI			
1-9	67.058	24.778	42,5
10-19	17.085	5.983	44,0
20-49	12.751	8.872	39,5
50-99	6.145	4.068	37,3
Totale	103.039	43.701	42,1

Il grado di copertura della rilevazione è stato in media dell'1,0%, con 43.701 imprese utilizzate nell'inferenza (37.330 rispondenti e 6.371 integrate dalla fonte Bilanci civilistici) rispetto ad un universo di 4.457.395 imprese. Tenendo conto delle imprese cessate, inattive, fallite, in liquidazione, fuori campo di osservazione e quelle per le quali non è stato possibile instaurare un contatto (imprese trasferite, sconosciute, non trovate all'indirizzo, ecc.), sono state utilizzate informazioni sui risultati economici delle imprese per il 42,1 % delle unità selezionate dal campione.

In generale le variabili più significative richieste nei diversi moduli del regolamento SBS sono:

- numero di imprese
- fatturato
- valore della produzione
- margine lordo sui beni destinati alla rivendita
- valore aggiunto al costo dei fattori
- margine operativo lordo

- acquisto complessivo di beni e servizi
- acquisto di merci da rivendere senza trasformazione
- variazione delle scorte di beni e servizi
- variazione delle scorte di beni e servizi destinati alla rivendita
- costo del personale
- retribuzione lorda
- ore lavorate
- investimenti lordi in beni materiali
- numero di persone occupate
- numero di dipendenti.

3.2. Fonti statistiche e amministrative utilizzate nell'attività di ricerca e definizione delle variabili

Le fonti utilizzate nell'attività di ricerca, aggiornate al 2007, sono le seguenti:

Archivio statistico delle imprese attive ASIA (Istat): è costituito dalle imprese e unità economiche che esercitano arti e professioni nelle attività industriali, commerciali e dei servizi alle imprese e alle famiglie. L'archivio è aggiornato annualmente attraverso un processo di integrazione di informazioni provenienti da fonti di diversa natura, amministrative e statistiche, e fornisce informazioni anagrafiche (denominazione e localizzazione) e di struttura (attività economica nella classificazione Ateco2002 e Ateco2007, addetti, dipendenti e indipendenti, forma giuridica e fatturato) delle unità;

Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (Istat): è la rilevazione campionaria sulle imprese con 1-99 addetti, descritta nel paragrafo 3.1. e nell'appendice A.2.;

Bilanci civilistici delle società di capitale (Camere di commercio): è la banca dati dei Bilanci civilistici depositati dalle società di capitale presso le Camere di Commercio;

Studi di settore (Ministero economia e finanze): è lo strumento che l'Agenzia delle entrate utilizza per rilevare i parametri fondamentali di liberi professionisti, lavoratori autonomi e imprese, sia sotto l'aspetto delle caratteristiche strutturali che economico, al fine di valutare la loro capacità di produrre reddito. Essi si articolano in diversi quadri, di cui i più rilevanti ai fini dell'attività di ricerca (Cerroni e De Giorgi, 2008) sono quelli relativi ai dati economici contabili:

- *Quadro F* - Per i soggetti con attività manifatturiere, commercio e servizi;
- *Quadro G* - Per i professionisti.

Modello Unico persone fisiche, società di persone, società di capitale (Ministero economia e finanze): è un modello unificato delle dichiarazioni fiscali tramite il quale le persone fisiche e giuridiche possono presentare la dichiarazione dei redditi, dell'IVA, dell'Irap e la dichiarazione dei sostituti di imposta ed intermediari (modello 770). Il modello Unico è distinto in relazione ai soggetti dichiaranti, ognuno dei quali deve compilare degli appositi quadri, dei quali si riportano i più rilevanti ai fini della ricerca:

- modello Unico Persone Fisiche:
 - *Quadro RE* - Redditi di lavoro autonomo;
 - *Quadro RF* - Redditi di impresa in contabilità ordinaria;
 - *Quadro RG* - Redditi di impresa in regime di contabilità semplificata e regimi forfetari;
 - *Quadro IQ* - Irap
- modello Unico Società di persone:
 - *Quadro RE* - Redditi di lavoro autonomo;
 - *Quadro RF* - Redditi di impresa in contabilità ordinaria;
 - *Quadro RG* - Redditi di impresa in regime di contabilità semplificata e regimi forfetari;
 - *Quadro IP* - Irap
- modello Unico Società di capitale:
 - *Quadro RS* - Prospetti vari;
 - *Quadro IC* - Irap

Il lavoro di ricerca ha riguardato l'analisi del contenuto di ciascuna fonte e ha consentito di costruire degli aggregati comuni che sono poi stati confrontati a livello di singola impresa per valutare la coerenza fra gli aggregati amministrativi ed i corrispondenti aggregati statistici comunicati dalle imprese all'Istat per la rilevazione PMI. Al fine di isolare l'effetto della calibrazione sono state selezionate le variabili con maggiore coerenza e si sono confrontate le stime a livello macro fra quelle tradizionali calcolate sui rispondenti alla rilevazione e quelle ottenute sostituendo i dati rilevati con quelli provenienti dalle diverse fonti amministrative, dando un livello di priorità stabilito in relazione alla maggiore coerenza riscontrata in fase sperimentale.

L'attività successiva è costituita dalla ricostruzione del campione teorico, nel calcolo nuovamente dei pesi per l'espansione dei dati all'universo e dal confronto delle stime e a livello macro fra le variabili dei rispondenti con l'aggiunta di quelle di fonte amministrativa per i non rispondenti e fra le variabili dell'intero campione teorico ricostruito con solo le variabili di fonte amministrativa.

Di entrambe le analisi si rinvia ai rispettivi paragrafi.

3.3. Copertura dell'archivio ASIA e del campione PMI con le fonti amministrative disponibili

In questo paragrafo si è verificato quanto le fonti amministrative a disposizione (Bilanci Civiltistici, modello Unico, Studi di Settore) coprano l'universo delle imprese attive fornito dall'archivio ASIA (anno di riferimento 2007) e le imprese rilevate con l'indagine PMI 2007. In particolare si forniscono informazioni sui tassi di copertura degli universi di interesse in termini di attività economica (Ateco Rev 1.2), classe di addetti e forma giuridica.

3.3.1. Copertura dell'archivio ASIA

Lo studio della copertura di ASIA è stato effettuato, in termini sia di imprese sia di addetti, considerando il solo insieme delle imprese che rappresentano l'universo di riferimento della rilevazione PMI riferita all'anno 2007. In altri termini, sono state escluse le imprese attive meno di 6 mesi, quelle con attività economica fuori campo d'osservazione (divisioni 65, 66, 75, 91) e le imprese con 100 e più addetti.

Tavola 3 - Tassi di copertura con le fonti amministrative, in termini di imprese e addetti, dell'universo ASIA riferito al campo di osservazione dell'indagine PMI per attività economica e fonte amministrativa - Anno 2007

	Bilanci		SdS F		SdS G		Totale fonti	
	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti
C-Estrazione di minerali	49,9	69,8	39,5	23,9	0,1	0,0	89,5	93,8
D-Attività manifatturiere	22,5	54,5	64,8	37,9	0,0	0,0	87,3	92,4
E-Produzione e distribuzione di energia elettrica, gas e acqua	57,5	81,8	2,1	0,6	0,1	0,0	59,7	82,4
F-Costruzioni	14,3	33,4	72,9	56,5	0,1	0,0	87,2	89,9
G-Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	11,1	30,7	77,0	60,1	0,1	0,0	88,2	90,9
H-Alberghi e ristoranti	10,8	24,5	75,1	66,1	0,0	0,0	85,9	90,7
I-Trasporti, magazzinaggio e comunicazioni	16,3	47,6	68,7	39,9	1,1	0,3	86,1	87,8
J-Attività finanziarie	6,1	13,9	72,8	68,0	6,3	4,3	85,2	86,3
K-Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	13,9	31,5	23,9	22,8	49,9	34,6	87,7	88,9
M-Istruzione	19,2	45,6	22,6	15,9	1,5	0,5	43,3	62,0
N-Sanità e assistenza sociale	4,6	31,1	2,9	4,2	81,9	55,3	89,4	90,7
O-Altri servizi pubblici, sociali e personali	7,8	26,2	64,5	53,9	3,7	1,8	76,0	81,9
Totale	13,2	37,0	56,6	45,1	17,0	7,8	86,8	90,0

La tavola 3 riassume i tassi di copertura (calcolati sia in termini di imprese che di addetti) dei *Bilanci civilistici* delle società di capitale e degli *Studi di Settore*,¹⁴ rispetto all'universo delle imprese attive presenti nell'archivio ASIA. La copertura degli *Studi di Settore* è stata calcolata separatamente per le imprese (quadro F) e per i professionisti (quadro G).

¹⁴ Per ciò che riguarda i dati dei modelli Unico e Irap, l'Istat acquisisce le informazioni riguardanti le sole imprese del campione PMI

Il numero degli addetti considerato è quello presente in ASIA, i cui record sono stati abbinati via codice fiscale con quelli delle fonti amministrative.

Delle 4.458.006 imprese dell'universo di riferimento, l'86,8% delle imprese è presente in almeno una delle fonti amministrative considerate. Tale percentuale è così suddivisa: il 13,2% si trova nei *Bilanci civilistici* e il 73,6 negli *Studi di Settore*. Le fonti amministrative considerate rappresentano il 90% degli addetti dell'universo di riferimento (il 53% si trova negli *Studi di Settore*).

La tavola 4 riassume i tassi di copertura per tutte le fonti amministrative (*Bilanci civilistici e Studi di Settore*) in termini di imprese e addetti dell'universo di riferimento per attività economica e classi di addetti.

Le tavole relative alla copertura di ogni singola fonte, anche relativamente a disaggregazioni più dettagliate, sono presentate nell'Appendice B.

Tavola 4 - Tassi di copertura con le fonti amministrative, in termini di imprese e addetti, dell'universo riferito al campo di osservazione dell'indagine PMI per attività economica e classe dimensionale - Anno 2007

ATTIVITÀ ECONOMICA	Numero di addetti														Totale	
	1		2		3..5		6..9		10..19		20..49		50..99		Impr.	Add.
	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.		
C-Estrazione di minerali	77,4	77,4	83,8	83,9	92,7	93,4	95,4	95,3	96,4	96,6	93,8	93,6	91,5	91,4	89,5	93,8
D-Attività manifatturiere	81,1	81,0	86,3	86,4	89,8	90,0	92,5	92,6	94,3	94,4	94,7	94,7	94,1	94,1	87,3	92,4
E-Produzione e distribuzione di energia elettrica, gas e acqua	46,8	46,9	56,3	56,4	68,2	68,8	80,9	81,6	85,6	85,9	86,1	86,4	89,7	90,1	59,7	82,4
F-Costruzioni	85,2	85,1	87,3	87,5	89,5	89,7	91,7	91,8	93,4	93,4	92,7	92,5	90,0	89,9	87,2	89,9
G-Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	85,3	85,3	90,7	90,8	93,4	93,6	95,1	95,1	94,2	94,1	90,7	90,4	89,0	88,9	88,2	90,9
H-Alberghi e ristoranti	78,6	78,5	84,0	84,1	89,7	90,1	94,0	94,1	94,9	94,9	95,4	95,4	95,7	95,6	85,9	90,7
I-Trasporti, magazzinaggio e comunicazioni	85,4	85,3	84,8	5,0	86,4	86,5	89,2	89,3	90,8	90,8	88,8	88,9	86,9	86,7	86,1	87,8
J-Attività finanziarie	82,9	82,9	87,2	87,3	91,2	91,4	91,8	91,7	86,9	86,0	69,3	68,3	46,7	46,6	85,2	86,3
K-Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	86,9	86,9	89,5	89,6	90,2	90,4	91,7	91,7	91,5	91,5	89,6	89,5	88,9	89,0	87,7	88,9
M-Istruzione	25,2	25,4	60,9	61,2	61,3	61,0	55,8	55,5	64,3	65,3	80,5	81,0	84,0	83,4	43,3	62,0
N-Sanità e assistenza sociale	88,2	88,1	93,8	93,9	92,6	92,4	89,5	89,4	88,8	89,1	91,9	92,2	93,7	93,5	89,4	90,7
O-Altri servizi pubblici, sociali e personali	69,1	69,2	83,6	83,7	88,1	88,1	85,8	85,9	86,1	86,1	87,2	87,4	84,1	84,7	76,0	81,9
Totale	84,4	84,3	88,2	88,3	90,5	90,7	92,4	92,4	93,2	93,2	92,5	92,4	91,4	91,4	86,8	90,0

Le coperture in termini di addetti, che rispecchiano quasi sempre le coperture in termini di imprese, fatta eccezione per il totale delle sezioni E ed M (evidentemente influenzate dalla bassa copertura delle imprese nelle classi più piccole), vanno da un minimo di 25,2% (*Istruzione* con 1 addetto) ad un massimo di 96,4% (*Classe 10-19 addetti dell'Estrazione di minerali*). La classe di addetti meglio rappresentata risulta essere la 10-19 (93,2% in termini sia di imprese che di addetti). Scendendo nel dettaglio dell'attività economica, ed escludendo la percentuale 97,9% della classe 20-49 dei *Trasporti marittimi e per vie d'acqua* dove le imprese sono poche decine, la divisione meglio rappresentata è la 24 (*Fabbricazione di prodotti chimici e di fibre sintetiche e artificiali*) con il 96,9% di imprese sul totale e il 97,0% di addetti, mentre quella meno rappresentata è la 41 (*Raccolta, depurazione e distribuzione d'acqua*), con un quinto circa delle imprese e degli addetti dell'universo di riferimento.

Analisi più approfondite per ateco a 3 cifre rivelano che il gruppo peggio rappresentato è il 924 (*Attività delle agenzie di stampa*) con solo il 10,4% di imprese e il 21,5% di addetti, quello meglio rappresentato è il gruppo 331 (*Fabbricazione di apparecchi medicali e chirurgici e di apparecchi ortopedici*) con il 95,2 di imprese. Se si considera anche la classe di addetti per le imprese più piccole (fino a 10 addetti) i gruppi meglio rappresentati sono il 612 (*Trasporti per vie d'acqua interne, compresi i trasporti lagunari*) per le imprese con 1 addetto (95,1%), il 331 (*Fabbricazione di apparecchi medicali e chirurgici e di apparecchi ortopedici*) per le imprese con 2-5 addetti e il 523 (*Commercio al dettaglio di prodotti farmaceutici, medicinali, di cosmetici e di articoli di profumeria*) per le imprese con 6-9 addetti (entrambe con il 96,5%). Le imprese con 10-19 addetti sono meno rappresentate nel gruppo 671 (*Attività ausiliarie dell'intermediazione finanziaria, escluse le assicurazioni e i fondi pensione*) con il 49,4% e meglio rappresentate nel 202 (*Fabbricazione di fogli da impiallacciatura, compensato, pannelli stratificati, pannelli di truciolato ed altri pannelli di legno*) con il 98,8% delle imprese. Anche le imprese con 20-99 sono poco rappresentate nel gruppo 671, ma vengono rappresentate al 98,8% nel gruppo 286 (*Fabbricazione di articoli di coltelleria, utensili e oggetti diversi in metallo*) le imprese 20-49 e con il 98,5% nel gruppo 501 (*Commercio di autoveicoli*) le imprese 50-99. Considerando infine le classi di attività economica (Ateco2002 a 4 cifre), ed escludendo quelle con poche imprese, la classe meglio rappresentata è la 5231 (*Farmacie*) con il 96,2%, quella con la più bassa percentuale di imprese è la 9240 (*Attività delle agenzie di stampa*).

Analogamente può essere interpretata la tavola 5 per forma giuridica di cui se ne riporta in sintesi la copertura per classi di addetti.

Tavola 5 - Tassi di copertura con le fonti amministrative, in termini di imprese e addetti, dell'universo ASIA riferito al campo di osservazione dell'indagine PMI per forma giuridica e classe dimensionale - Anno 2007

FORMA GIURIDICA	Numero di addetti														Totale	
	1		2		3..5		6..9		10..19		20..49		50..99		Impr.	Add.
	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.	Impr.	Add.		
Imprenditore individuale, libero professionista e lavoratore autonomo	84,7	84,6	89,4	89,6	91,0	91,1	90,9	90,9	89,0	88,9	81,3	80,1	58,6	57,3	86,1	87,5
Società di persone	83,4	83,3	86,8	86,8	89,8	90,1	92,6	92,6	91,7	91,5	82,7	80,6	47,0	45,8	87,8	88,8
Società di capitali	85,2	85,0	88,3	88,6	92,5	92,7	94,7	94,7	95,6	95,7	95,4	95,4	94,5	94,4	90,1	94,1
Altre società (cooperative, consorzi, enti)	48,4	48,4	58,9	59,3	71,0	71,2	77,4	77,7	84,5	84,8	86,2	86,0	86,5	86,5	66,8	81,8
Totale	84,4	84,3	88,2	88,3	90,5	90,7	92,4	92,4	93,2	93,2	92,5	92,4	91,4	91,4	86,8	90,0

La forma giuridica meglio rappresentata è quella delle "società di capitale" (90,1% di imprese e 94,1% di addetti), quelle con la più bassa copertura, sia di imprese sia di addetti, sono le "altre società". Se si osserva la distribuzione per numero di addetti si nota che imprenditori individuali, liberi professionisti, lavoratori autonomi e società di persone sono poco rappresentati nelle imprese con 50-99 addetti. Le "altre società" sono poco rappresentate fra le imprese con 1-5 addetti.

La tavola 6 riassume la distribuzione percentuale per forma giuridica delle imprese presenti in ASIA, riferite al campo di osservazione PMI, mettendola a confronto con le numerosità delle fonti amministrative disponibili e le rispettive coperture, in termini di imprese e addetti.

Tavola 6 - Numero di imprese e addetti dell'universo ASIA e delle fonti amministrative disponibili per forma giuridica, riferiti al campo di osservazione dell'indagine PMI - Anno 2007

FORMA GIURIDICA	Asia		distribuzione %		Fonti amministrative		% copertura di ASIA	
	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti
Imprenditore individuale, libero professionista e lavoratore autonomo	2.882.609	4.556.903	64,7	35,5	2.482.194	3.987.550	86,1	87,5
Società di persone	814.401	2.951.999	18,3	23,0	715.167	2.622.222	87,8	88,8
Società di capitali	697.956	4.782.008	15,7	37,2	629.185	4.500.546	90,1	94,1
Altre società (cooperative, consorzi, enti)	63.040	561.540	1,4	4,4	42.105	459.199	66,8	81,8
Totale	4.458.006	12.852.450	100,0	100,0	3.868.651	11.569.516	86,8	90,0

Nella tavola si osserva che la struttura produttiva del Paese è costituita in larga prevalenza da piccolissime imprese con forma giuridica "imprenditore individuale, libero professionista e lavora-

tore autonomo” che rappresentano due terzi dell’universo delle imprese ASIA dando occupazione a circa 4,5 milioni di addetti che costituiscono il 35,5% degli addetti complessivi. Le società di persone ammontano a circa 814 mila e rappresentano il 18,3% delle imprese e il 23% degli occupati; le società di capitale sono circa 700 mila e costituiscono il 15,7% delle imprese e il 37,2% degli addetti mentre le altre società sono circa 63.000 e rappresentano solo l’1,4% delle imprese il 4,4% degli addetti. La tavola mostra, come già in precedenza descritto, che la copertura delle fonti amministrative disponibili (*Bilanci civilistici* e *Studi di settore*) rappresentano molto bene l’universo ASIA sia in termini di imprese (86,8%) che di addetti (90,0%). L’unica eccezione sono le “altre società” (con una copertura del 66,8% delle imprese di ASIA che sale, tuttavia, all’81,8% in termini di addetti) che comunque rappresentano una quota molto piccola delle imprese dell’archivio ASIA.

L’analisi dell’archivio ASIA per forma giuridica e classi dimensionali rileva la seguente struttura che è sostanzialmente confermata anche nella corrispondente analisi delle fonti amministrative disponibili:

- 1 addetto: 82.3% sono imprenditori individuali, liberi professionisti, lavoratori autonomi;
- 2 addetti: 56.0% sono imprenditori individuali, liberi professionisti, lavoratori autonomi e 31.9% sono società di persone;
- 3-5 addetti: 39.5% sono imprenditori individuali, liberi professionisti, lavoratori autonomi e 38.7% sono società di persone;
- 6-9 addetti: 36.9% sono società di persone e 36.1% sono società di capitale;
- 10-19 addetti: 55.9% sono società di capitale e 27.0% sono società di persone;
- 20-49 addetti: 74.8% sono società di capitale e 11.8% sono società di persone;
- 50-99 addetti: 81.7% sono società di capitale e 14.0% sono altre società.

Tavola 7 - Sintesi dei tassi di copertura, in termini di imprese e addetti, dell’universo ASIA riferito al campo di osservazione dell’indagine PMI: numero di celle per classi di copertura % e tipologia di disaggregazione - Anno 2007

COPERTURA %	Ateco 2 cifre – classe di addetti		Ateco 3 cifre – classe di addetti		Ateco 4 cifre		Forma giuridica - classe di addetti	
	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti
<=50	18	16	111	106	59	16	9	9
50 ... 70	24	24	91	88	50	28	2	2
70 ... 80	23	25	125	125	34	31	2	2
80 ... 90	113	109	342	342	177	111	13	12
90 ... 95	102	105	377	375	128	172	7	8
95 ... 100	49	50	296	306	17	107	2	2
Numero totale di disaggregazioni	329	329	1342	1342	465	465	35	35

La tavola 7 offre una sintesi dei tassi di copertura per i diversi livelli di disaggregazione (Ateco a 2 cifre per classi di addetti, Ateco a 3 cifre per classi di addetti, Ateco a 4 cifre per il complesso delle imprese con 1-99 addetti, forma giuridica e classe di addetti).

Nelle celle della tavola è indicato il numero di disaggregazioni (attività economica o forma giuridica e classi di addetti precedentemente indicate) con i livelli di copertura della prima colonna, sia in termini di addetti sia di imprese. A parte la prima classe (percentuale di copertura compresa fra 0 e 50%), che sicuramente andrebbe depurata delle disaggregazioni con pochissime unità, la classe modale è quasi sempre la classe 80-90, tranne nel caso dell’attività a 3 cifre con classe di addetti, dove la classe modale è 90-95 (modale anche per la copertura in termini di addetti della disaggregazione a 4 cifre).

In ogni caso, per ciò che riguarda le analisi sia per attività economica sia per forma giuridica, i motivi dell’eventuale bassa copertura sono da ricercare nella particolarità delle fonti a disposizione per il calcolo: i Bilanci civilistici evidentemente non possono rappresentare la prima tipologia di imprese di ASIA (imprenditori individuali, liberi professionisti e lavoratori autonomi) e analogamente sottorappresentano le società di persone e le imprese con 1 o 2 addetti, che invece sono rappresentate più o meno adeguatamente negli Studi di Settore, a seconda che si escluda o non si applichi lo studio di settore.

Per una più dettagliata analisi della copertura, si dovrebbe, infatti, restringere l’insieme di riferimento alle sole imprese che compilano il bilancio o che sono sottoposte a studio di settore, escludendo pertanto le attività economiche non soggette a studio di settore o con forma giuridica non

compatibile con il bilancio. L'analisi qui elaborata, però, si propone di testare se, e in che modo, le fonti amministrative ad oggi disponibili in Istat siano sufficienti a rappresentare l'universo di riferimento. Nell'analisi si dovrebbero considerare anche le altre fonti amministrative (*Modello Unico* e *Quadro Irap* che al momento della ricerca erano disponibili solo relativamente al campione della rilevazione PMI), come avviene nel successivo paragrafo in cui è stato effettuato lo studio della copertura del campione PMI.

3.3.2 Copertura del campione PMI

Nella tavola 8 mostra il numero di imprese, rispondenti e non rispondenti, presenti nel campione teorico della rilevazione PMI2007 ed in almeno una fonte, per fonte amministrative.

Tavola 8 - Imprese delle diverse fonti amministrative presenti nel campione teorico dell'indagine PMI - Anno 2007

FONTE	Numero imprese	di cui rispondenti in Pmi	
		no	si
Bilanci	30.622	10.883	19.739
Studi di settore F	61.596	33.413	28.183
Studi di settore G	2.672	1.449	1.223
Unico PF - RG	18.700	11.428	7.272
Unico PF - RE	3.621	2.044	1.577
Unico SP - RG	6.243	3.626	2.617
Unico SC - RS	44.035	20.552	23.483
Unico SP - RE	40.196	18.224	21.972
Irap PF	24.146	14.644	9.502
Irap SP	20.552	8.151	12.401
Irap SC	18.259	10.262	7.997

Ovviamente nel calcolo delle imprese che hanno risposto a PMI, si tiene conto del solo fatto se siano o no presenti, a prescindere dal numero di fonti in cui ogni record compare.

Un'analisi sulle numerosità delle fonti e sulla validità statistica del contenuto delle variabili disponibili in ciascuna di esse (cfr. par. 3.4.1), ha evidenziato non solo la priorità di utilizzo delle fonti¹⁵ ma anche la non utilizzabilità di alcune di esse. Non sono stati, infatti, utilizzati nello studio, perché non sufficientemente affidabili in termini di numero di variabili e/o di validità dei dati, i modelli Unico persone fisiche (*Quadro RE*) e i modelli *Irap*.

Nella tavola 9 viene ricostruito il campione teorico utilizzando le priorità definite e includendo anche le fonti che successivamente verranno escluse dall'analisi (*Unico SP-RE* e modelli *Irap*). Come si può vedere le fonti escluse rappresentano solo l'1,3% del campione teorico. Ciascuna unità del campione viene recuperata dalla fonte disponibile con priorità più elevata.

I dati riferiti al numero degli addetti sono stati ricavati dall'indagine PMI, se il record è presente; sono invece quelli dell'archivio ASIA se l'impresa non ha risposto a PMI. I record definiti come "Residuo PMI" (1.251 casi) sono quei record che si riferiscono ad imprese di cui non si ha traccia in nessuna fonte amministrativa ma hanno risposto al questionario di indagine.¹⁶ I record non coperti da nessuna fonte amministrativa sono invece presenti nell'archivio ASIA dell'anno 2007; gli errori di lista sono, invece, le imprese con attività economica fuori campo di osservazione o con addetti maggiori di 100. Questi casi rappresentano in tutto il 13% circa (3,0% "Nessuna fonte"; 9,9% "Errori di lista"). La percentuale di errori di lista è piuttosto alta ma ciò è senz'altro legato al fatto che il campione della rilevazione PMI dell'anno 2007 è stato estratto dall'archivio ASIA aggiornato all'anno 2005 (t-2 rispetto all'anno di riferimento della rilevazione) mentre usualmente la lista viene estratta

¹⁵ Dare priorità maggiore ad una fonte piuttosto che ad un'altra significa attribuirle un valore (utilizzabilità) maggiore. Nel nostro caso dire che i Bilanci hanno priorità maggiore degli studi di settore significa che se esistono delle registrazioni in entrambe le fonti, si privilegerà il bilancio (per il dettaglio sulle priorità cfr. tavola 17).

¹⁶ Avendo agganciato i record delle fonti amministrative con il campione teorico mediante il codice impresa dell'archivio ASIA, è possibile anche che le imprese non si trovino in nessuna fonte perché nell'archivio risultano non attive.

dall'archivio aggiornato al tempo t-1. Tutti i confronti di copertura sono stati comunque effettuati rispetto all'archivio ASIA aggiornato al 2007. La presenza di errori di lista è comunque inevitabile. Resta eventualmente da approfondire quante imprese etichettate con "Nessuna fonte" sono effettivamente esonerate da adempimenti amministrativi e quante imprese classificate negli "Errori di lista" sono realmente nello stato di inattività, cessate, fuse, scorporate o in liquidazione.

Tavola 9 - Numero di imprese del campione teorico per fonte amministrativa disponibile e classe di addetti - Anno 2007

FONTE DISPONIBILE	Classi di addetti								Totale
	1	2	3..5	6..9	10..19	20..49	50..99	100+	
NUMERO DI IMPRESE									
Bilanci	2.549	1.443	3.186	3.029	7.742	8.155	4.005	0	30.109
Studi di settore F	16.956	7.956	7.562	3.194	4.679	1.882	224	0	42.453
Studi di settore G	1.939	281	174	55	90	27	0	0	2.566
Unico PF - RG	2.403	515	308	50	21	4	1	0	3.302
Unico PF - RE	1.179	40	7	3	1	0	0	0	1.230
Unico SP - RG	426	395	287	62	15	3	0	0	1.188
Unico SC - RS	1.484	583	886	575	1.030	1.179	648	0	6.385
Unico SP - RE	20	16	26	15	23	22	5	0	127
Irap PF	188	79	129	81	88	46	10	0	621
Irap SP	0	0	0	1	0	0	0	0	1
Irap SC	70	83	118	99	117	45	7	0	539
Residuo PMI	388	121	196	137	205	144	60	0	1.251
Nessuna fonte	1.051	399	501	322	344	281	151	0	3.049
Errori di lista	5.227	1.446	1.102	354	775	467	211	636	10.218
Campione teorico	33.880	13.357	14.482	7.977	15.130	12.255	5.322	636	103.039
% SUL TOTALE									
Bilanci	7,5	10,8	22,0	38,0	51,2	66,5	75,3	0,0	29,2
Studi di settore F	50,0	59,6	52,2	40,0	30,9	15,4	4,2	0,0	41,2
Studi di settore G	5,7	2,1	1,2	0,7	0,6	0,2	0,0	0,0	2,5
Unico PF - RG	7,1	3,9	2,1	0,6	0,1	0,0	0,0	0,0	3,2
Unico PF - RE	3,5	0,3	0,0	0,0	0,0	0,0	0,0	0,0	1,2
Unico SP - RG	1,3	3,0	2,0	0,8	0,1	0,0	0,0	0,0	1,2
Unico SC - RS	4,4	4,4	6,1	7,2	6,8	9,6	12,2	0,0	6,2
Unico SP - RE	0,1	0,1	0,2	0,2	0,2	0,2	0,1	0,0	0,1
Irap PF	0,6	0,6	0,9	1,0	0,6	0,4	0,2	0,0	0,6
Irap SP	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Irap SC	0,2	0,6	0,8	1,2	0,8	0,4	0,1	0,0	0,5
Residuo PMI	1,1	0,9	1,4	1,7	1,4	1,2	1,1	0,0	1,2
Nessuna fonte	3,1	3,0	3,5	4,0	2,3	2,3	2,8	0,0	3,0
Errori di lista	15,4	10,8	7,6	4,4	5,1	3,8	4,0	100,0	9,9
Campione teorico	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Tavola 10 - Tassi di copertura con le fonti amministrative in termini di imprese e addetti del campione teorico dell'indagine PMI per sezione di attività economica - Anno 2007

	Totale fonti amministrative + residuo Pmi	Nessuna fonte	Fuori copertura / errori di lista	Totale
C - Estrazione di minerali	76,1	2,7	21,2	100,0
D - Attività manifatturiere	82,2	4,0	13,7	100,0
E - Produzione e distribuzione di energia elettrica, gas e acqua	80,2	8,3	11,5	100,0
F - Costruzioni	87,6	4,2	8,1	100,0
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di	87,7	3,0	9,3	100,0
H - Alberghi e ristoranti	87,7	3,9	8,4	100,0
I - Trasporti, magazzinaggio e comunicazioni	84,9	6,8	8,3	100,0
J - Attività finanziarie	90,9	6,8	2,3	100,0
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	91,5	3,9	4,6	100,0
M - Istruzione	88,7	8,0	3,3	100,0
N - Sanità e assistenza sociale	90,1	4,2	5,7	100,0
O - Altri servizi pubblici, sociali e personali	88,8	5,3	5,9	100,0
Totale	85,9	4,2	9,9	100,0

La tavola 10 sintetizza la copertura del campione teorico con le fonti amministrative disponibili (solo quelle con priorità da 1 a 7 definite nella tavola 17) per sezione di attività economica; le tavole con dettaglio maggiore sono presentate in appendice B.

La tavola evidenzia che attraverso l'uso delle fonti amministrative si riesce a coprire l'86% del campione teorico. Nella tavola 11 è riportata la distribuzione per sezione di attività economica e per fonte e nella tavola 12 la distribuzione per divisione di attività economica e classi di addetti.

La tavola 11 evidenzia che i *Bilanci* e gli *Studi di Settore* rappresentano circa l'85% del campione teorico; l'apporto che dà l'utilizzo del modello Unico è molto limitato, se si escludono la sezione E ed O (dove rappresenta circa un terzo del campione teorico) e la sezione M (in cui rappresenta circa la metà), anche se la numerosità del campione in tali sezioni è piccola

Tavola 11 - Numero di imprese del campione ricostruito e distribuzione percentuale per attività economica e fonte amministrativa - Anno 2007

	Bilancio	SdS F	SdS G	Unico PF - RG	Unico PF - RE	Unico SP - RG	Unico SC - RS	Residuo PMI	Totale
NUMERO DI IMPRESE DEL CAMPIONE RICOSTRUITO									
C - Estrazione di minerali	440	392	0	12	0	3	45	3	895
D - Attività manifatturiere	13.453	15.899	17	1.264	5	366	2.258	462	33.724
E - Produzione e distribuzione di energia elettrica, gas e acqua	541	23	1	31	0	22	254	46	918
F - Costruzioni	1.132	2.861	2	73	0	9	141	25	4.243
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	4.894	10.202	7	347	2	95	757	134	16.438
H - Alberghi e ristoranti	541	1.581	3	103	0	62	161	26	2.477
I - Trasporti, magazzinaggio e comunicazioni	2.027	2.286	166	406	11	126	558	74	5.654
J - Attività finanziarie	115	920	31	46	6	8	48	61	1.235
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	4.664	6.066	1.280	486	314	281	1.061	200	14.352
M - Istruzione	387	151	28	148	247	96	284	74	1.415
N - Sanità e assistenza sociale	439	76	795	28	60	15	163	31	1.607
O - Altri servizi pubblici, sociali e personali	1.476	1.996	236	358	585	105	655	115	5.526
Totale	30.109	42.453	2.566	3.302	1.230	1.188	6.385	1.251	88.484
DISTRIBUZIONE PERCENTUALE									
C - Estrazione di minerali	49,2	43,8	0,0	1,3	0,0	0,3	5,0	0,3	100,0
D - Attività manifatturiere	39,9	47,1	0,1	3,7	0,0	1,1	6,7	1,4	100,0
E - Produzione e distribuzione di energia elettrica, gas e acqua	58,9	2,5	0,1	3,4	0,0	2,4	27,7	5,0	100,0
F - Costruzioni	26,7	67,4	0,0	1,7	0,0	0,2	3,3	0,6	100,0
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	29,8	62,1	0,0	2,1	0,0	0,6	4,6	0,8	100,0
H - Alberghi e ristoranti	21,8	63,8	0,1	4,2	0,0	2,5	6,5	1,0	100,0
I - Trasporti, magazzinaggio e comunicazioni	35,9	40,4	2,9	7,2	0,2	2,2	9,9	1,3	100,0
J - Attività finanziarie	9,3	74,5	2,5	3,7	0,5	0,6	3,9	4,9	100,0
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	32,5	42,3	8,9	3,4	2,2	2,0	7,4	1,4	100,0
M - Istruzione	27,3	10,7	2,0	10,5	17,5	6,8	20,1	5,2	100,0
N - Sanità e assistenza sociale	27,3	4,7	49,5	1,7	3,7	0,9	10,1	1,9	100,0
O - Altri servizi pubblici, sociali e personali	26,7	36,1	4,3	6,5	10,6	1,9	11,9	2,1	100,0
Totale	34,0	48,0	2,9	3,7	1,4	1,3	7,2	1,4	100,0

Se utilizziamo la copertura totale (85,9%) come copertura media del campione teorico, si può concludere che per le imprese più piccole la copertura è inferiore alla media: nelle imprese con 1 solo addetto, infatti, la copertura è dell'80,6% e in quelle con 2 addetti dell'84,9% (Tavola 12). La copertura aumenta per le classi dimensionali più grandi, tra le quali la meglio rappresentata è la 20-49 (93,0% in media). Ovviamente l'analisi qui condotta potrebbe essere effettuata per dettagli maggiori dell'attività economica: più si disaggrega, più le percentuali di copertura variano.

Tavola 12 - Numero di imprese delle fonti disponibili e del campione teorico e tassi di copertura per divisione di attività economica e classe di addetti - Anno 2007

	Classi di addetti							Totale
	1	2	3..5	6..9	10..19	20..49	50..99	
NUMERO DI IMPRESE DEL CAMPIONE RICOSTRUITO								
C - Estrazione di minerali	142	86	145	143	202	139	38	895
D - Attività manifatturiere	7.305	3.322	4.716	3.629	6.483	5.647	2.622	33.724
E - Produzione e distribuzione di energia elettrica, gas e acqua	323	73	119	86	124	133	60	918
F - Costruzioni	1.357	518	578	357	701	516	216	4.243
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	5.205	2.837	2.395	1.153	2.432	1.756	660	16.438
H - Alberghi e ristoranti	772	362	437	231	299	261	115	2.477
I - Trasporti, magazzinaggio e comunicazioni	1.665	755	1.054	550	721	643	266	5.654
J - Attività finanziarie	324	335	277	83	120	84	12	1.235
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	6.491	2.214	1.877	760	1.405	1.105	500	14.352
M - Istruzione	544	142	236	120	180	136	57	1.415
N - Sanità e assistenza sociale	743	109	115	99	204	212	125	1.607
O - Altri servizi pubblici, sociali e personali	2.453	581	657	397	626	583	229	5.526
Totale	27.324	11.334	12.606	7.608	13.497	11.215	4.900	88.484
NUMERO DI IMPRESE DEL CAMPIONE TEORICO								
C - Estrazione di minerali	275	131	184	149	231	155	47	1.176
D - Attività manifatturiere	10.676	4.326	5.694	3.811	7.286	6.133	2.829	41.008
E - Produzione e distribuzione di energia elettrica, gas e acqua	429	101	141	96	139	161	69	1.144
F - Costruzioni	1.601	601	650	382	768	581	233	4.841
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	6.314	3.208	2.661	1.195	2.679	1.895	719	18.740
H - Alberghi e ristoranti	928	400	480	230	348	287	122	2.824
I - Trasporti, magazzinaggio e comunicazioni	2.082	877	1.200	587	842	704	301	6.657
J - Attività finanziarie	358	351	295	89	146	97	20	1.359
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	7.071	2.418	2.037	783	1.534	1.209	535	15.688
M - Istruzione	603	163	270	134	205	152	61	1.595
N - Sanità e assistenza sociale	813	127	129	107	231	224	133	1.783
O - Altri servizi pubblici, sociali e personali	2.730	654	741	414	721	657	253	6.224
Totale	33.880	13.357	14.482	7.977	15.130	12.255	5.322	103.039
PERCENTUALE DI COPERTURA								
C - Estrazione di minerali	51,6	65,6	78,8	96,0	87,4	89,7	80,9	76,1
D - Attività manifatturiere	68,4	76,8	82,8	95,2	89,0	92,1	92,7	82,2
E - Produzione e distribuzione di energia elettrica, gas e acqua	75,3	72,3	84,4	89,6	89,2	82,6	87,0	80,2
F - Costruzioni	84,8	86,2	88,9	93,5	91,3	88,8	92,7	87,6
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	82,4	88,4	90,0	96,5	90,8	92,7	91,8	87,7
H - Alberghi e ristoranti	83,2	90,5	91,0	100,4	85,9	90,9	94,3	87,7
I - Trasporti, magazzinaggio e comunicazioni	80,0	86,1	87,8	93,7	85,6	91,3	88,4	84,9
J - Attività finanziarie	90,5	95,4	93,9	93,3	82,2	86,6	60,0	90,9
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	91,8	91,6	92,1	97,1	91,6	91,4	93,5	91,5
M - Istruzione	90,2	87,1	87,4	89,6	87,8	89,5	93,4	88,7
N - Sanità e assistenza sociale	91,4	85,8	89,1	92,5	88,3	94,6	94,0	90,1
O - Altri servizi pubblici, sociali e personali	89,9	88,8	88,7	95,9	86,8	88,7	90,5	88,8
Totale	80,6	84,9	87,0	95,4	89,2	91,5	92,1	85,9

La successiva tavola 13 rappresenta le coperture per forma giuridica e fonte amministrativa. Essa evidenzia la buona copertura degli Studi di settore per l'attività svolta dagli imprenditori individuali, liberi professionisti e lavoratori autonomi (73,9% con il Quadro F e 8,3% con il Quadro G) che trovano una buona rappresentazione anche nel modello Unico PF-RG (11,9%). Rilevante è anche la copertura degli Studi di Settore (Quadro F) per le società di persone con l'82,9%. Le società di capitale sono invece coperte per il 73,2% dai Bilanci civilistici, ma trovano ulteriore copertura per il 15,8% dagli Studi di Settore (Quadro F) e per il 10,4% dal modello Unico SC-RS. I dati delle altre società (cooperative, consorzi, enti) sono reperibili dai Bilanci civilistici (53,4%), Studi di Settore Quadro F (13,0%) e modello Unico SC-RS (30,3%). La tavola 14 riporta la distribuzione delle unità del campione ricostruito per forma giuridica e per classe di addetti.

Tavola 13 - Numero di imprese del campione teorico e distribuzione per forma giuridica e fonte - Anno 2007

	Bilancio	SdS F	SdS G	Unico PF-RG	Unico PF-RE	Unico SP-RG	Unico SC RS	Residuo PMI	Totale
NUMERO DI IMPRESE DEL CAMPIONE RICOSTRUITO									
Imprenditore individuale, libero professionista e lavoratore autonomo	0	20.497	2.290	3.302	1.230	0	0	419	27.738
Società di persone	44	15.427	276	0	0	1.185	1.255	420	18.607
Società di capitali	27.977	6.021	0	0	0	3	3.945	282	38.228
Altre società (cooperative, consorzi, enti)	2.088	508	0	0	0	0	1.185	130	3.911
Totale	30.109	42.453	2.566	3.302	1.230	1.188	6.385	1.251	88.484
DISTRIBUZIONE PERCENTUALE									
Imprenditore individuale, libero professionista e lavoratore autonomo	0,0	73,9	8,3	11,9	4,4	0,0	0,0	1,5	100,0
Società di persone	0,2	82,9	1,5	0,0	0,0	6,4	6,7	2,3	100,0
Società di capitali	73,2	15,8	0,0	0,0	0,0	0,0	10,3	0,7	100,0
Altre società (cooperative, consorzi, enti)	53,4	13,0	0,0	0,0	0,0	0,0	30,3	3,3	100,0
Totale	34,0	48,0	2,9	3,7	1,4	1,3	7,2	1,4	100,0

Tavola 14 - Numero di imprese del campione ricostruito e distribuzione percentuale per forma giuridica e classi di addetti - Anno 2007

	Classi di addetti							Totale
	1	2	3..5	6..9	10..19	20..49	50..99	
NUMERO DI IMPRESE DEL CAMPIONE RICOSTRUITO								
Imprenditore individuale, libero professionista e lavoratore autonomo	17.389	4.753	3.316	936	1.061	263	20	27.738
Società di persone	3.811	3.998	4.470	2.075	2.953	1.141	159	18.607
Società di capitali	5.583	2.388	4.338	3.644	8.918	9.037	4.320	38.228
Altre società (cooperative, consorzi, enti)	541	195	482	450	851	953	439	3.911
Totale	27.324	11.334	12.606	7.105	13.783	11.394	4.938	88.484
NUMERO DI IMPRESE DEL CAMPIONE TEORICO								
Imprenditore individuale, libero professionista e lavoratore autonomo	20.882	5.422	3.829	1.112	1.295	341	39	32.924
Società di persone	4.467	4.652	5.031	2.270	3.217	1.244	174	21.066
Società di capitali	7.650	2.995	4.996	4.013	9.579	9.569	4.600	43.928
Altre società (cooperative, consorzi, enti)	881	288	626	582	1.039	1.101	509	5.121
Totale	33.880	13.357	14.482	7.977	15.130	12.255	5.322	103.039
PERCENTUALE DI COPERTURA								
Imprenditore individuale, libero professionista e lavoratore autonomo	83,3	87,7	86,6	84,2	81,9	77,1	51,3	84,2
Società di persone	85,3	85,9	88,8	91,4	91,8	91,7	91,4	88,3
Società di capitali	73,0	79,7	86,8	90,8	93,1	94,4	93,9	87,0
Altre società (cooperative, consorzi, enti)	61,4	67,7	77,0	77,3	81,9	86,6	86,2	76,4
Totale	80,6	84,9	87,0	89,1	91,1	93,0	92,8	85,9

Dalla tavola emerge chiaramente che le coperture sono nella maggior parte dei casi superiori alla media, risultando per le società di persone e di capitali con più di 5 addetti anche superiori al 90%. Le basse coperture per la forma giuridica “imprenditore individuale, libero professionista e lavoratore autonomo” nelle classi di addetti superiori a 20 fa pensare, proprio per le caratteristiche di tali soggetti, ad errori di lista nell’attribuzione della forma giuridica piuttosto che a problemi di scarsa copertura. La forma giuridica peggio rappresentata risulta quella delle altre società (cooperative, consorzi, enti).

Si presenta di seguito la tavola 15 di sintesi della percentuale di copertura per i diversi livelli di disaggregazione (Ateco a 2 cifre per classi di addetti, Ateco a 3 cifre per classi di addetti, Ateco a 4 cifre per il complesso delle imprese con 1-99 addetti, forma giuridica per classe di addetti).

Tavola 15 - Sintesi delle percentuali di copertura in termini di imprese: numero di celle per classi di copertura percentuale e tipologia di disaggregazione - Anno 2007

COPERTURA %	Ateco 2 cifre - classe di addetti	Ateco 3 cifre - classe di addetti	Ateco 4 cifre	Forma giuridica - classe di addetti
<=50	12	54	6	3
50 ... 70	23	109	24	6
70 ... 80	33	133	100	6
80 ... 90	108	381	232	12
90 ... 95	125	371	74	8
95 ... 100	34	302	29	0
Numero totale di disaggregazioni	335	1.350	465	35

Le classi modali per la percentuale di copertura risultano sempre la classe 80-90 e la classe 90-95; nel caso dell'Ateco a 3 cifre un quinto delle celle hanno una copertura molto alta (almeno 95%). Anche nel caso dello studio della copertura del campione, i motivi dell'eventuale bassa copertura sono da ricercare nella particolarità delle fonti a disposizione per il calcolo.

3.4. La procedura di integrazione delle mancate risposte nella rilevazione PMI basata sull'uso delle fonti amministrative disponibili

In questo paragrafo viene descritta la procedura di integrazione delle mancate risposte nella rilevazione PMI2007 sulla base dell'uso delle fonti amministrative analizzate nei paragrafi precedenti. In particolare sono descritte le analisi preliminari delle fonti (paragrafo 3.4.1), la ricostruzione del campione teorico con i dati delle fonti amministrative più coerenti con quelli della rilevazione (paragrafo 3.4.2), l'analisi sperimentale dell'effetto della fonte e dell'effetto dell'integrazione (paragrafo 3.4.3) e una breve descrizione dell'applicazione di questa attività sperimentale alla rilevazione PMI2008.

3.4.1. Confronti rispetto a PMI rispondenti e selezione delle fonti con le variabili che hanno maggiore coerenza con quelle di PMI2007

L'armonizzazione delle informazioni contenute nelle diverse fonti amministrative, secondo gli schemi SBS utilizzati per l'indagine PMI, ha richiesto un consistente lavoro di analisi degli archivi amministrativi, sia dal punto di vista della loro copertura in termini di unità statistiche sia dal punto di vista della adattabilità del loro potenziale informativo. Il problema della copertura è stato affrontato attraverso un utilizzo gerarchico delle fonti definito in modo non arbitrario e che scaturisce dal grado di completezza e comparabilità con i dati dei rispondenti di PMI (test *Kolmogorov-Smirnoff* sulle distribuzioni delle variabili comuni).

Si utilizza, pertanto, prima la fonte dei *Bilanci* (per le società di capitale), poi si passa agli *Studi di Settore* (per le attività obbligate e con volumi di affari inferiori a 7,5 milioni di euro) e in via residuale al *Modello Unico*, distinto per forma giuridica di impresa.

Il secondo problema ha riguardato l'armonizzazione delle informazioni e la definizione di un set comune di variabili relative al conto economico dell'impresa (Tavola 16). Si tratta in totale di 23 variabili. Non tutte le fonti amministrative contengono le stesse informazioni. La più completa è quella dei *Bilanci*, poi c'è la fonte degli *Studi di Settore* (sia imprese, sia professionisti), ed infine quella dei *Modelli Unico* distinti per forma giuridica e regime di contabilità di impresa.

Tavola 16 - Descrizione delle informazioni del conto economico provenienti dalle fonti amministrative e armonizzate secondo lo schema dell'indagine PMI

Descrizione	Codice SBS	Codice Pmi	Indagine Pmi	Bilanci	Studi di Settore (a)	Modello Unico (b)
Totale ricavi correnti escl. IVA, al lordo delle imposte indir.	12110	11100	fatt_tot_pmi	fatt_tot_bil	fatt_tot_sdsx	fatt_tot_uyyzz
Variazioni delle rim. prodotti finiti, in corso di lav. e semilav.	13213	11200	var_rpfpcl_pmi	var_rpfpcl_bil		CRS271
Variazione rimanenze		11200-12600	var_riman_pmi	var_riman_bil	var_riman_sdsx	var_riman_uyyzz
Variazioni dei lavori in corso su ordinazione		11300	var_lavco_pmi	var_lavco_bil	var_lavco_sdsx	var_lavco_uyyzz
Incrementi di immobilizzazioni per lavori interni		11400	inc_immli_pmi	inc_immli_bil	inc_immli_sdsx	
Altri ricavi e proventi (non finanziari, non straordinari)		11500	ric_altri_pmi	ric_altri_bil	ric_altri_sdsx	ric_altri_uyyzz
Acquisti - totale		12100	acq_beni_pmi	acq_beni_bil	acq_beni_sdsx	acq_beni_uyyzz
Acquisti beni e servizi	13110	12100+12200	acq_bese_pmi	acq_bese_bil	acq_bese_sdsx	acq_bese_uyyzz
Acquisti di merci da rivendere in nome proprio senza trasf.	13120	12103				CRS321
Costi per servizi		12200	acq_ser_pmi	acq_ser_bil	acq_ser_sdsx	acq_ser_uyyzz
Godimento di beni di terzi		12300	acq_gdbt_pmi	acq_gdbt_bil	acq_gdbt_sdsx	acq_gdbt_uyyzz
Ammortamenti		12500	acq_amm_pmi	acq_amm_bil	acq_amm_sdsx	acq_amm_uyyzz
Variazioni delle rimanenze di materie prime e merci da riv.	13210	12600	var_rmpriv_pmi	var_rmpriv_bil		CRS331
Variazioni delle rimanenze - di merci da rivendere	13211	12602				var_rriv_uscrs (CRS341)
Accantonamenti		12700	acq_acc_pmi	acq_acc_bil	acq_acc_sdsx	acq_acc_uyyzz
Oneri diversi di gestione - totale		12900	acq_oneri_pmi	acq_oneri_bil	acq_oneri_sdsx	acq_oneri_uyyzz
Valore aggiunto (Va)	12150	14000 (c)	vagg_pmi	vagg_bil	vagg_sdsx	vagg_uyyzz
Retribuzioni	13320	41110+42110	ret_pmi	ret_bil		
costi per dirigenti, quadri e impiegati: oneri sociali	13330	42121	onerisoc_pmi	onerisoc_bil		
Trattamento di fine rapporto		42131	tfr_pmi	tfr_bil		
Personale	13310	44000	clav_pmi	clav_bil	clav_sdsx	clav2_uyyzz
Altri costi del lavoro		45000	altrclav_pmi	altrclav_bil		
Margine operativo lordo (Mol)	12170	15000 (d)	margope_pmi	margope_bil	margope_sdsx	margope_uyyzz

(a) Studi di Settore x= f, g

(b) Modello Unico yy= pf, sp, sc

zz= rg, re, rs, iq, ip, ic

(c) 14000 (Va) = 11100+11200-12600+11300+11400+11500-12100-12200-12300-12900

(d) 15000 (Mol) = 14000-44000

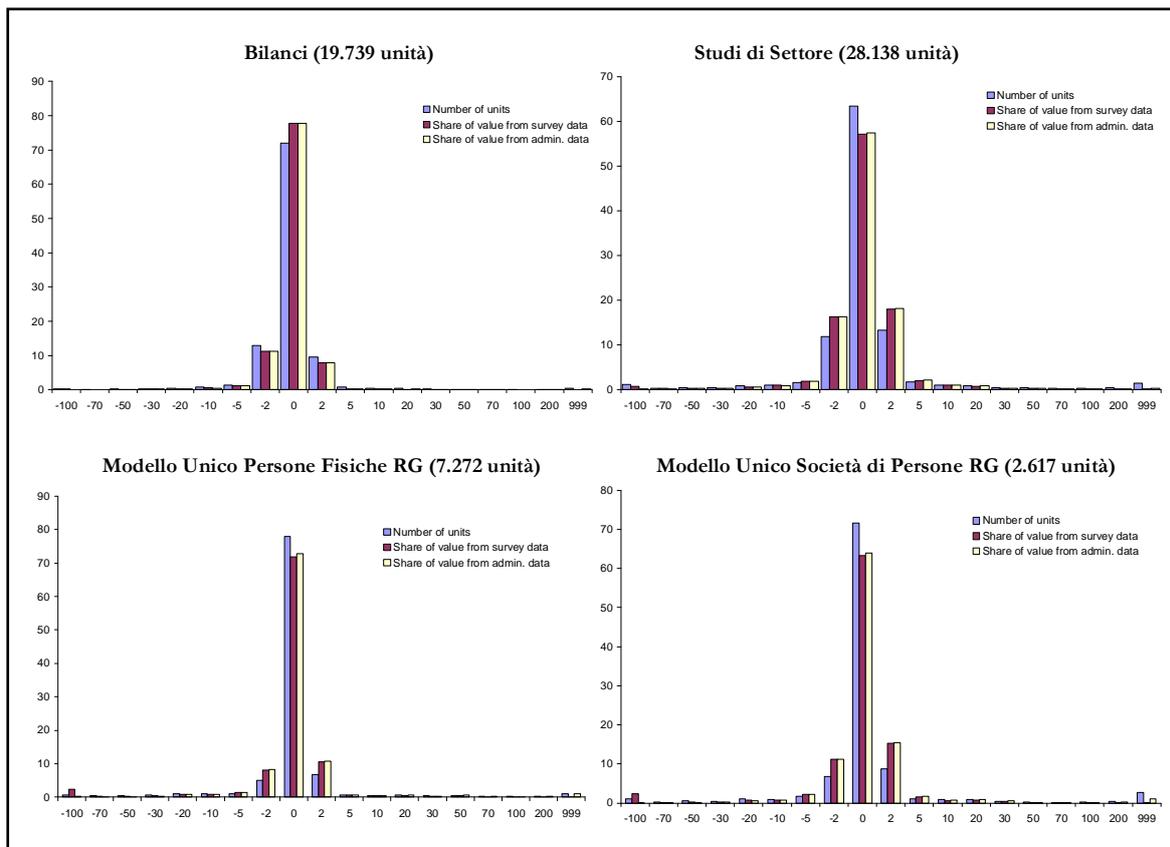
Successivamente sono state integrate, attraverso il codice impresa, le informazioni delle imprese rispondenti alla rilevazione PMI2007, con quelle disponibili nelle diverse fonti amministrative (*Bilanci civilistici, Studi di Settore, Modello Unico e Irap*). Da ciò ne è scaturito un data set contenente le 43.701 unità rispondenti.

È stato quindi possibile confrontare le distribuzioni delle variabili presenti in PMI con quelle corrispondenti ricostruite dalla fonte amministrativa, calcolando le differenze relative percentuali così definite:

$$diffV = \frac{V(\text{Amministrativa}) - V(\text{Pmi})}{V(\text{Pmi})} \cdot 100 \quad (9)$$

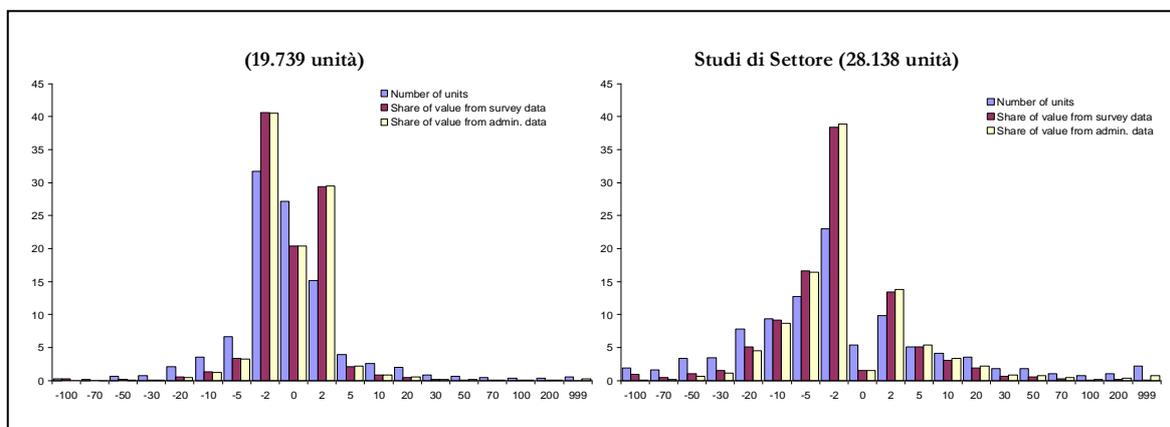
dove V indica la generica variabile utilizzata per il confronto mentre tra parentesi viene indicata la fonte dalla quale sono derivati i dati. Tali differenze sono state calcolate a livello di singola impresa e solamente nei casi in cui il valore espresso al denominatore sia diverso da zero. Nel caso in cui sia invece $V(\text{Amministrativa}) > 0$ e $V(\text{Pmi}) = 0$, la differenza $diffV$ è stata posta pari al valore convenzionale '999', mentre nel caso in cui $V(\text{Amministrativa}) = 0$ e $V(\text{Pmi}) = 0$ si è posto $diffV = 0$; infine, nei casi in cui $V(\text{Amministrativa}) = \text{mancante}$, i valori di $diffV$ sono stati esclusi dall'analisi. Per ognuna delle variabili considerate, le imprese sono state raggruppate in base ai valori di $diffV$ in classi di differenze percentuali, sia negative che positive: 0, 0-2%, 2-5%, 5-10%, 10-20%, 20-30%, 30-50%, 50-70%, 70-100%, 100-200%, oltre 200%. L'analisi si è concentrata sullo studio delle distribuzioni percentuali di queste differenze relative calcolate per ognuna delle variabili considerate.

Figura 1 - Distribuzione delle unità rispondenti e delle quote di fatturato per range di variazione della variabile ricavi delle vendite e delle prestazioni



Circa il 94% delle osservazioni, comuni alla fonte Bilanci, ha valori di fatturato compresi in un range di +/- 2%, con un 70% circa di imprese con valori identici. Anche per la fonte Studi di settore la situazione è soddisfacente con una percentuale dell'88% compresa nel range di +/- 2%, e con il 63% con valori identici. Lo stesso accade per le unità abbinato con i dati fiscali del modello Unico, sia nel caso del modello Unico persone fisiche (PF) sia nel caso del modello Unico società di persone (SP) in regime di contabilità semplificata (che compilano il quadro RG). In tutti i casi si osserva una distribuzione simmetrica delle differenze percentuali a testimonianza della presenza di una casualità, e non sistematicità, di queste differenze.

Figura 2 - Distribuzione delle unità rispondenti e delle quote di costo per acquisti di beni e servizi per range di variazione della variabile costo per acquisti di beni e servizi



Per quanto concerne i Costi per acquisti di beni e servizi si osserva una buona confrontabilità per le fonti dei Bilanci e degli Studi di Settore, mentre per i dati del quadro del modello Unico relativo alle imprese in contabilità semplificata non è possibile effettuare un confronto sui costi per l'assenza di informazioni contabili dettagliate. Relativamente ai Bilanci si osserva una concentrazione nell'intorno dei range di variazione di +/- 5% di circa l'85% delle unità abbinata. In termini di valore, poi, si arriva a circa il 96%. Meno concentrata è la situazione per gli Studi di Settore se si guarda alle unità statistiche, mentre in termini di valore si riscontra una percentuale del 75% circa tra i +/- 5% di differenze percentuali. Si osserva, inoltre, sempre per gli Studi di Settore una asimmetria positiva della distribuzione a testimonianza del fatto che sono più frequenti i casi in cui si verificano valori di costo inferiori a quelli rilevati da PMI. Infine solo in pochi casi, meno del 5%, si osservano valori identici di costo tra il dato amministrativo e PMI, ciò è dovuto alla diversa riclassificazione delle voci di costo nell'ambito del quadro contabile degli Studi di Settore. Comunque il grado di compensazione appare soddisfacente.

I confronti sulla variabile Costi del personale rivelano un buon grado di accostamento in particolare per la fonte Bilanci; buono è anche l'accostamento per gli Studi di Settore mentre meno buono risulta per il Modello Unico. In quest'ultima fonte si evidenziano costi del personale mediamente maggiori del relativo costo rilevato dall'indagine e ciò è legato al fatto che nella variabile fiscale sono inclusi i costi per i soggetti assimilati ai dipendenti e per lavoro autonomo. Per le fonti esaminate si osserva comunque una buona percentuale di casi perfettamente concordanti (65% circa nei Bilanci, 57% circa negli Studi di settore, 74% circa nel Modello Unico PF-RG e 67% circa nel Modello Unico SP-RG) e, più nel dettaglio, per i Bilanci la percentuale di casi con variazione +/- 10% è di circa il 93%, per gli Studi di Settore si avvicina al 90% e per il Modello Unico all'80%.

Infine i confronti sulla variabile Valore Aggiunto, che rappresenta una voce di saldo tra componenti positive e negative, e che pertanto è più difficile da ricostruire, si osserva una simmetria delle differenze osservate. Migliore è la situazione per i Bilanci, che non a caso già vengono utilizzati per la stima di questa variabile. Per gli Studi di Settore, considerata la difficoltà riscontrata sulle voci di costo, non si osserva quasi mai una concordanza precisa, con un 46% circa delle unità e un 55% del valore concentrati nelle classi di variazione +/-5%. Comunque il 61% delle unità e il 72% del valore si concentrano nelle classi di variazioni tra +/- 10%. In ogni caso, le imprese che compilano gli Studi di settore hanno mediamente un Valore aggiunto più basso delle imprese che compilano i bilanci. Per concludere, le analisi di confronto mostrano delle distribuzioni simmetriche intorno allo zero che possono essere assimilate a errori casuali e spiegano in definitiva l'ordine di priorità utilizzato per le fonti amministrative (tavola 17).

Figura 3 - Distribuzione delle unità rispondenti e delle quote di costo del personale per range di variazione della variabile Costi del personale

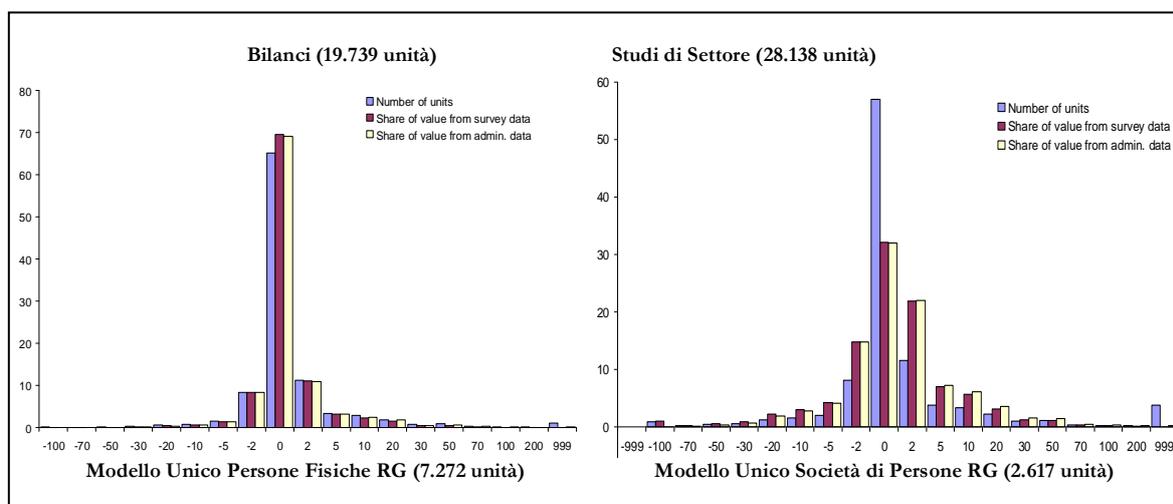


Figura 3 segue - Distribuzione delle unità rispondenti e delle quote di costo del personale per range di variazione della variabile Costi del personale

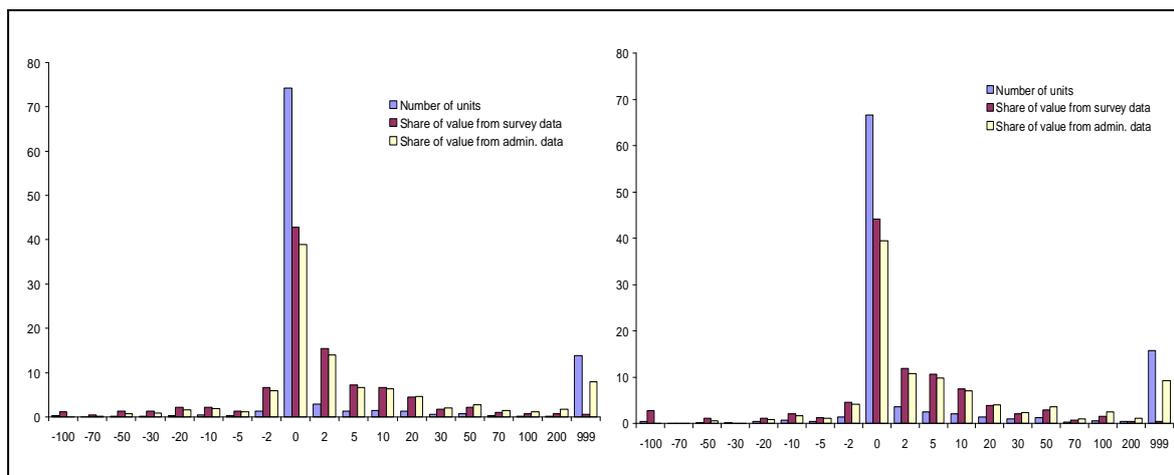
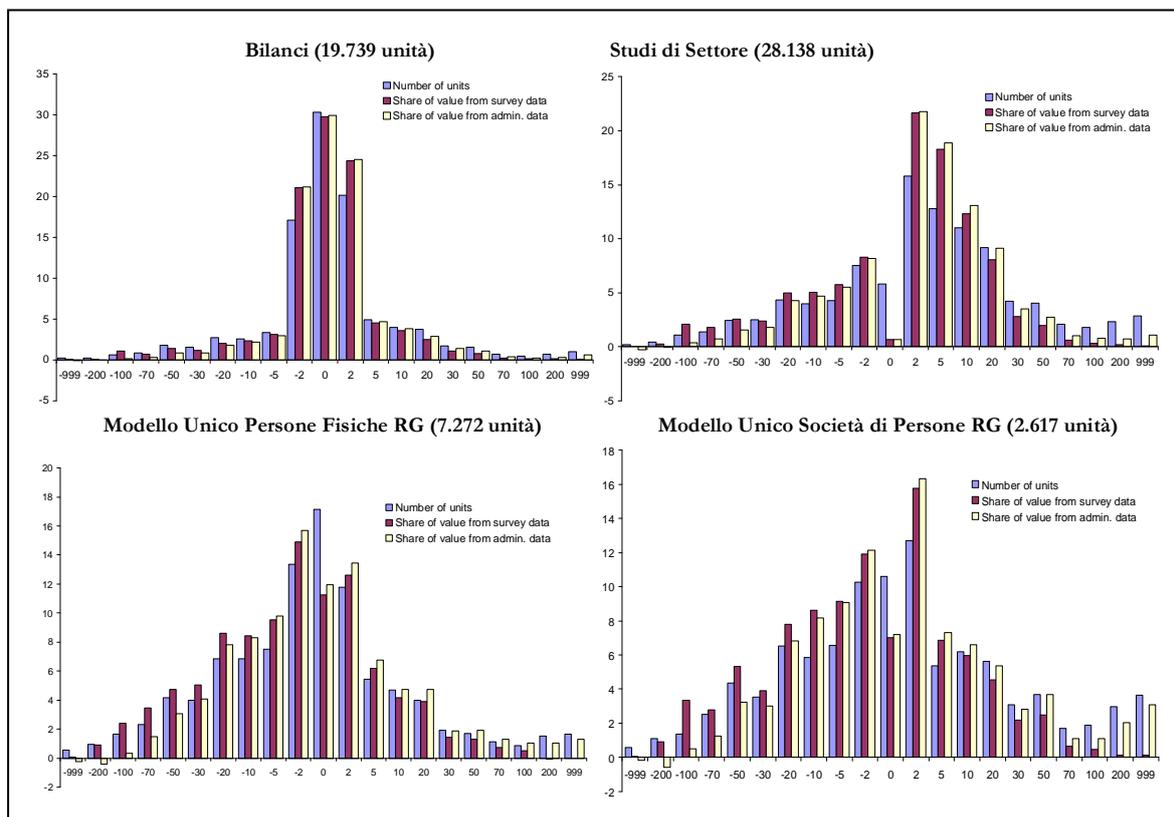


Figura 4 - Distribuzione delle unità rispondenti e delle quote di valore aggiunto per range di variazione della variabile valore aggiunto



Per valutare l'uguaglianza della forma delle distribuzioni dei dati del campione PMI con quelli di ciascuna fonte è stato effettuato il test di *Kolmogorov-Smirnoff*, i cui risultati sono evidenziati nella tavola 17.

Tavola 17 - Sintesi dei risultati dell'applicazione del test KS alle fonti amministrative disponibili - Anno 2007

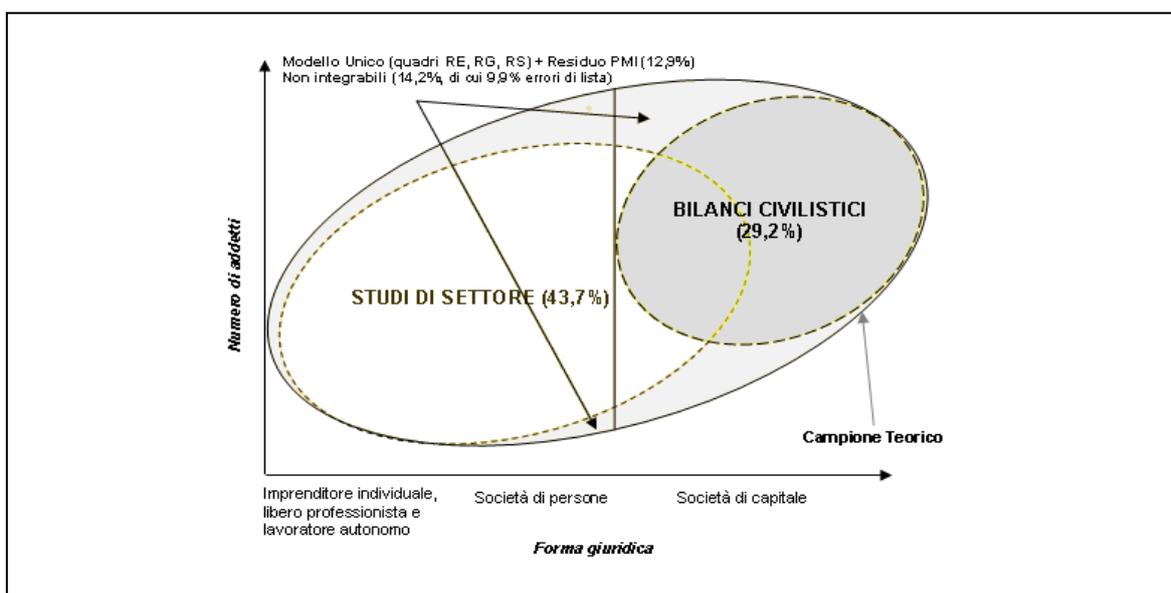
FORTE AMMINISTRATIVA	Variabili confrontabili	Test KS positivo	Priorità
Bilanci	21	13	1
Studi di settore F	15	8	2
Studi di settore G	13	7	3
Unico PF - RE	13	6	4
Unico PF - RG	14	6	5
Unico SP - RG	14	6	6
Unico SC - RS	16	2	7
Unico SP - RE	13	0	-
Irap PF	16	0	-
Irap SP	16	0	-
Irap SC	16	0	-

Nella seconda colonna della tavola 17 è indicato il numero di variabili che possono essere confrontate con quelle dell'indagine PMI, mentre nella terza il numero di test validi. Per la fonte *Bilanci*, delle 21 variabili ricostruite e confrontabili con quelle dell'indagine PMI, 13 variabili superano il test di uguaglianza delle distribuzioni empiriche osservate su fonti diverse e relative alle stesse unità statistiche. Esse riguardano tutte le principali variabili SBS, dal fatturato ai costi per beni e servizi, dal valore aggiunto al Mol, con l'eccezione delle scorte e del valore della produzione (a seguito della mancanza maggior dettaglio relativo alle merci da rivendere).

Dalla fonte degli *Studi di Settore* si ottengono 15 variabili confrontabili di cui più della metà (8 variabili) mostrano una distribuzione empirica simile. Esse riguardano il fatturato, gli incrementi di immobilizzazione, le variazioni dei lavori in corso, e aggregazioni di voci di costo quali i costi per beni e servizi, il costo del lavoro, gli ammortamenti e variabili di saldo quali le variazioni di rimanenze e il valore aggiunto.

Per i dati fiscali relativi al *modello Unico* si raggiunge un soddisfacente grado di confrontabilità per i quadri RG e RE (rispettivamente contabilità semplificata e professionisti) dei modelli sulle persone fisiche (PF) ed il quadro RG del modello per le società di persone (SP). Infine ulteriori informazioni, con un buon grado di dettaglio, relativamente alle merci da rivendere, possono essere desunte dal quadro RS del modello compilato dalle società di capitali (SC).

In base ai risultati esposti nella terza colonna della tavola 17, le fonti amministrative utilizzate sono quelle individuate con priorità da 1 a 7. Nella figura 5 è rappresentato il contributo delle diverse fonti amministrative alla copertura in termini di imprese del campione teorico PMI2007, come precedentemente riportato nel dettaglio nella tavola 9 (cfr. par. 3.3.2.).

Figura 5 - Copertura del campione teorico PMI2007 con le fonti amministrative utilizzate

Limitatamente alle sole imprese che si abbinano agli *Studi di Settore* (circa 19.000 unità), al netto dunque delle società di capitali abbinato con i *Bilanci*, sono state calcolate alcune misure sintetiche di confronto per le principali variabili di PMI.

Le differenze (non ponderate) mostrano valori leggermente inferiori per tutte le variabili -1% di ricavi e costi e una differenza maggiore sulle variabili di saldo: -3% sul valore aggiunto e -7% per il Mol. Il test di *Kolmogorov-Smirnoff* sull'uguaglianza delle distribuzioni, su questo sottoinsieme di osservazioni (non società di capitali), risulta, però, sempre positivo per tutte le variabili. Anche se in media dalla fonte amministrativa si stima un valore aggiunto più basso, a livello di valori medi la differenza percentuale è di solo 0,3%. Tale differenza aumenta sulle code della distribuzione per valori molto bassi rilevati in PMI (primo 5% della distribuzione) che negli *Studi di Settore* sono superiori del 98,4%. Si può affermare, anche alla luce delle analisi settoriali svolte e dei controlli effettuati, che esiste una compensazione, non sistematica, di differenze positive e negative.

Questa analisi di confronto sulle osservazioni comuni è stata propedeutica per valutare la coerenza fra le variabili calcolate nelle diverse fonti amministrative con quelle della rilevazione PMI2007 e soprattutto per assegnare delle priorità qualora si volessero utilizzare i dati di fonte amministrativa in sostituzione di quelli della rilevazione per l'insieme delle unità del campione iniziale non rispondenti. La tavola 17 evidenzia nell'ordine il livello di priorità delle diverse fonti.

Si è optato per la definizione di una lista di priorità per fonte piuttosto che per singola variabile in quanto definire le priorità per fonte lascia inalterata anche la caratteristica di coerenza tra le variabili, mentre l'utilizzo di variabili proveniente da fonti diverse non garantisce la coerenza complessiva.

Dove i dati di fonte amministrativa non si abbinavano con quelli della rilevazione, si è deciso di utilizzare in forma residuale (1.251 unità) i dati rilevati direttamente sulle imprese con la rilevazione. Il ruolo centrale è svolto dai Bilanci, grazie ai quali già si ricostruiscono 6.371 mancate risposte totali.

Tavola 18 - Misure di confronto (non pesate) tra gli Studi di Settore e l'indagine PMI (differenze in euro e percentuali)

	Ricavi di vendita	Acquisti di beni e servizi	Valore aggiunto	Costo del lavoro	Margine operativo lordo
	STATISTICHE				
Mean (euro)	-5.257	-3.398	-4.991	-147	-5.505
Mean (% diff)	-1,0	-1,0	-3,1	-0,2	-6,8
Standard Dev.	-2,0	-2,4	-6,6	0,2	-17,7
Coeff of variation	-1,0	-1,4	-3,7	0,3	-11,7
Skewness	-6,3	-9,9	-7,4	-0,9	-11,8
Kolmog. Smirn.(d)	0,8	1,0	1,2	1,6	1,6
	QUANTILI (DIFF. %)				
100% Max	-7,2	-10,4	-3,0	0,0	-4,0
99%	-1,7	-0,1	-8,0	0,4	-18,4
95%	-0,5	-0,5	-2,7	0,6	-10,2
90%	-1,2	-1,5	-2,7	-0,4	-4,8
75% Q3	-1,3	-1,1	-0,6	-1,1	-1,0
50% Median	0,5	-0,5	0,3	-	2,1
25% Q1	2,8	1,2	2,1	-	4,7
10%	6,2	9,9	12,7	-	37,2
5%	13,8	18,6	98,4	-	-

Numero di osservazioni: 18,788

(a) valore critico 1,6 al livello di 1% di probabilità di significatività

Tavola 19 - Imprese rispondenti PMI2007 con dati sostituibili con quelli di fonte amministrativa secondo l'ordine di confrontabilità - Anno 2007

Rispondenti PMI2007	43.701
Bilanci civilistici (Bil, sì in PMI)	6.371
Bilanci civilistici (Bil, no in PMI)	13.368
Studi di Settore quadro F (Sdsf)	17.798
Studi di Settore quadro G (Sdsg)	1.223
Unico persone Fisiche quadro RG (Upfgr)	990
Unico persone fisiche quadro RE (Upfre)	483
Unico società di persone quadro RG (Usprg)	378
Unico società di capitale quadro RS (Uscrs)	1.839
Residuo PMI	1.251

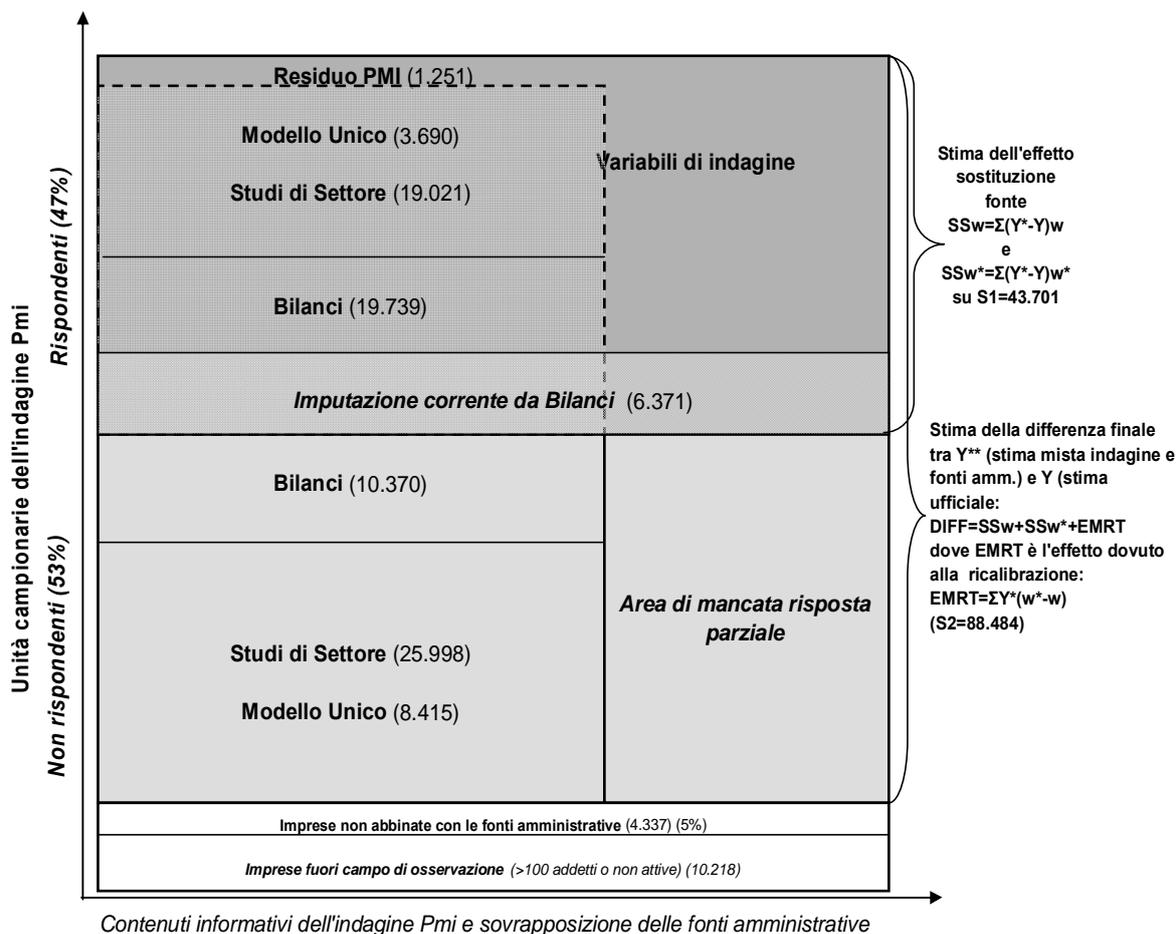
Alla luce dei confronti effettuati è stata condotta una ricostruzione delle informazioni dei non rispondenti, limitatamente ad un sottoinsieme di variabili comuni, utilizzando le fonti amministrative nell'ordine gerarchico della tavola appena mostrata.

Specularmente a quanto fatto per le unità rispondenti, è stato poi possibile abbinare la maggior parte delle unità non rispondenti a PMI alle diverse fonti amministrative: nell'ordine *Bilanci* (10.370 unità), *Studi di Settore* (25.998 unità) e *modello Unico* (8.415 unità residuali). L'estensione delle informazioni da fonte amministrativa a tutte le unità non rispondenti del campione iniziale ha permesso, come verrà analizzato in seguito, di ottenere un insieme più ampio su alcune variabili comuni e di ricalcolare i pesi finali (w^*) e confrontarli con i pesi precedentemente ottenuti solo sui rispondenti (w).

Di conseguenza, la differenza tra le due stime finali: quella ufficiale Y (solo da indagine) e quella mista Y^{**} (da indagine e da fonte amministrativa) può essere scomposta in differenze parziali riconducibili a due effetti principali: 1) effetto sostituzione fonte (parte alta della figura 6: SEw , SEw^*); 2) effetto ricalibrazione o distorsione da mancata risposta totale (parte bassa della figura 6: EMRT).

La figura 6 mostra in modo sintetico il processo di copertura delle diverse fonti (*Bilanci*, *Studi di Settore* e *Modello Unico*), con l'indicazione della numerosità, sia per la parte delle unità rispondenti del campione iniziale (47%), sia per la parte non rispondente (53%). I due punti problematici riguardano l'area di mancata risposta parziale, risolvibile attraverso un processo di imputazione delle variabili che non possono essere ricostruite direttamente dalla fonte amministrativa, e la parte inferiore del grafico relativa alle unità non riscontrabili in nessuna fonte, risolvibile attraverso il nuovo processo di ricalibrazione.

Figura 6 - Schema sintetico di integrazione del campione PMI con le fonti amministrative utilizzate per la stima alternativa – Anno 2007



3.4.2. Ricostruzione del campione teorico

La ricostruzione del campione teorico utilizzando tutte le fonti informative disponibili ha portato alla predisposizione di un data set di 88.484 unità, in cui alle 43.701 imprese della rilevazione PMI2007 sono state aggiunte 44.783 imprese integrate da fonte amministrativa.

Questa attività ha consentito la ricostruzione dell'85,9% del campione teorico: il 14,1% di imprese residue che non è stato possibile ricostruire è legato sia al cambiamento di stato per alcune di esse (imprese cessate) sia alla non disponibilità delle variabili di interesse in alcuni quadri della modulistica amministrativa (ad esempio per le società di persone, dove i relativi quadri riportano solo le variazioni in aumento e diminuzione del reddito) e sia alla non disponibilità di informazioni relative ad alcuni soggetti come ad esempio i contribuenti minimi (nel *Modello Unico* 2008 è stato, invece, introdotto uno specifico quadro CM, descritto nell'appendice D, che deve essere compilato da tali soggetti). Il contributo delle diverse fonti alla ricostruzione del campione teorico è sintetizzato nella tavola 20: alle 43.701 imprese rispondenti alla rilevazione, in cui sono tuttavia incluse 6.371 imprese che erano già state integrate dalla fonte *Bilanci* in prevalenza nella fascia dimensionale 20-99 addetti, si sono aggiunte ulteriori 10.370 imprese dalla fonte *Bilanci*, 24.655 imprese che hanno compilato il *Quadro F* dagli *Studi di Settore* e, quindi, altre imprese rispondenti ad altri quadri della modulistica amministrativa.

Tavola 20 - Imprese del campione teorico PMI2007: dati sostituibili con quelli di fonte amministrativa

Campione teorico PMI2007	103.039	
Pmi	43.701	} 88.484 Numerosità per l'inferenza
Bil	10.370	
Sdsf	24.655	
Sdsg	1.343	
Upfrg	2.312	
Upfre	747	
Usprg	810	
Uscrs	4.546	
Non integrabili	14.555	
44.783		

Alla luce di quanto detto, è stato costruito un file con tutte le informazioni di strato necessarie per l'esecuzione della procedura di stima mediante tecnica di calibrazione (ateco, regione amministrativa e addetti al momento dell'estrazione del campione e l'Ateco e addetti dichiarati dall'impresa o aggiornati dall'archivio per le imprese integrate) e sono state costruite le variabili su cui effettuare i successivi confronti.

3.4.3. Analisi delle differenze tra campione dell'indagine e campione ricostruito e valutazione dell'effetto delle mancate risposte totali

In questo paragrafo viene riportata una sintesi dei risultati dei confronti delle stime dell'indagine PMI per l'anno 2007 con le stime che si ottengono utilizzando le fonti amministrative disponibili, che permettono di ricostruire quasi tutto il campione teorico dell'indagine. In questo modo è possibile estendere la ricostruzione delle variabili per unità rispondenti e non rispondenti.

Il campione teorico può essere suddiviso in unità rispondenti (che sono il 47%) e unità non rispondenti (53%) così come rappresentato dalla figura 6. Con le informazioni desunte dagli archivi amministrativi riusciamo a coprire solo una parte delle informazioni contenute nel questionario di indagine. Tali informazioni possono essere confrontate con i dati dell'indagine al fine di ricavarne conclusioni riguardo la valutazione dell'effetto delle mancate risposte e dell'effetto delle fonti utilizzate.

Al fine di valutare l'effetto della fonte e delle mancate risposte totali dell'attuale indagine (43.701 unità) si è deciso di confrontare le stime desunte dai seguenti insiemi di dati:

1. 43.701 unità con i dati di indagine – utilizzando il vecchio sistema di pesi;
2. 43.701 unità ricostruite con i soli dati amministrativi – utilizzando il vecchio sistema di pesi;
3. 88.484 unità costituite dai dati dell'indagine (43.701 unità) e dai dati amministrativi (44.783 unità) – utilizzando il nuovo sistema di pesi (determinati con la stessa metodologia di calibrazione utilizzata tradizionalmente e descritta in appendice A.2.);
4. 88.484 unità ricostruite con i soli dati amministrativi – utilizzando il nuovo sistema di pesi.

La relazione che lega i confronti ottenibili è una scomposizione della differenza tra le stime del terzo insieme (88.484 unità costruite integrando i dati amministrativi ai dati di indagine con nuovo sistema di pesi) e le stime del primo insieme (43.701 unità da dati di indagine con vecchi pesi) in tre addendi: una parte relativa all'effetto di sostituzione fonte con vecchi pesi, una parte relativa all'effetto di sostituzione fonte con nuovi pesi e una parte relativa all'effetto delle mancate risposte totali.

Tenendo presente la schematizzazione proposta nella Figura 6, vengono di seguito illustrati i passaggi che portano alla formulazione finale di tale differenza.¹⁷

Indicati con S_1 l'insieme composto da 43.701 imprese e w i pesi abbinabili a esse (vecchio sistema di pesi), con S_2 l'insieme delle 88.484 imprese e w^* i pesi a esse abbinabili (nuovo sistema di pesi) e con y e y^* rispettivamente i dati di indagine e i dati amministrativi, possiamo rappresentare le stime iniziali nel seguente modo:

$$Y_1 = \sum_{S_1} y_k w_k \quad (10)$$

e cioè, supposto che Y sia la variabile da stimare, il suo valore calcolato sul primo insieme con i soli dati di indagine sarà dato dalla sommatoria in S_1 dei valori y moltiplicati per i pesi w .

Analogamente, se si considerano i soli dati amministrativi, il valore che scaturisce dal riporto all'universo sul quarto insieme è il seguente:

$$Y_4 = \sum_{S_2} y_k^* w_k^* \quad (11)$$

che scindendo l'insieme delle unità in S_1 e $S_2 - S_1$ equivale a

$$Y_4 = \sum_{S_1} y_k^* w_k^* + \sum_{S_2 - S_1} y_k^* w_k^* \quad (12)$$

Le stime prodotte dal terzo insieme vengono calcolate infine con la seguente formula:

$$Y_3 = \sum_{S_1} y_k w_k^* + \sum_{S_2 - S_1} y_k^* w_k^* \quad (13)$$

¹⁷ Per maggiori dettagli vedere il paragrafo 3 del documento "Experimental analysis in the estimation of SBS variables for small firms by using administrative data", Casciano, De Giorgi, Oropallo, Siesto, presentato al seminario Using Administrative Data in the Production of Business Statistics, Roma, Marzo 2010.

Si indica, inoltre, con Y_2 la stima effettuata sui dati amministrativi utilizzando i vecchi pesi:

$$Y_2 = \sum_{S1} y_k^* w_k \quad (14)$$

La differenza tra le stime del terzo insieme (88.484 unità con dati di indagine integrati con dati amministrativi e nuovi pesi) e le stime del primo insieme (43.701 unità con dati di indagine e vecchi pesi) è scomponibile nelle seguenti tre componenti:

$$DIFF = Y_3 - Y_1 = (Y_3 - Y_4) + (Y_4 - Y_2) + (Y_2 - Y_1) \quad (15)$$

dove:

- $(Y_3 - Y_4)$ indica l'effetto di sostituzione nuovi pesi
- $(Y_4 - Y_2)$ indica l'effetto MRT
- $(Y_2 - Y_1)$ indica l'effetto sostituzione vecchi pesi

Sostituendo le stime definite dai quattro insiemi si ottiene:

$$DIFF = Y_3 - Y_1 = \left(\sum_{S1} y_k^* w_k^* + \sum_{S2-S1} y_k^* w_k^* - \sum_{S1} y_k^* w_k^* - \sum_{S2-S1} y_k^* w_k^* \right) + \left(\sum_{S2} y_k^* w_k^* - \sum_{S1} y_k^* w_k^* - \sum_{S2-S1} y_k^* w_k^* \right) + \left(\sum_{S1} y_k^* w_k^* - \sum_{S1} y_k w_k \right) \quad (16)$$

dal momento che

$$\sum_{S2-S1} y_k^* w_k^* = 0 \quad (17)$$

si ha quindi che:

$$\begin{aligned} DIFF &= -\sum_{S1} (y_k^* - y_k) w_k^* + \sum_{S2} (w_k^* - w_k) y_k^* + \sum_{S1} (y_k^* - y_k) w_k = \\ &= -SSw^* + SSw + EMRT \end{aligned} \quad (18)$$

dove:

- $DIFF$ è la differenza le stime ottenute utilizzando in forma integrata i dati di indagine e quelli amministrativi (terzo insieme) e le stime dell'indagine (primo insieme);
- SSw è l'effetto di sostituzione fonte nell'insieme delle 43.701 unità con i pesi dell'indagine;
- $-SSw^*$ è l'effetto di sostituzione fonte nell'insieme delle 43.701 unità con il nuovo sistema di pesi;
- $EMRT$ è l'effetto delle mancate risposte totali nell'insieme delle 88.484 unità.

L'effetto di sostituzione fonte deriva dalla discrepanza che si osserva tra i dati dell'indagine ed i dati ricostruiti con la fonte amministrativa, tenendo costante la struttura dei pesi, e utilizzando prima i vecchi pesi poi i nuovi pesi. L'effetto singolo legato a vecchi e nuovi pesi ha come risultato un effetto fonte globale dato dalla differenza dei due: in sostanza essi agiscono in direzioni opposte.

All'effetto fonte va aggiunto l'effetto dovuto alle mancate risposte totali.

Le cause per cui l'effetto fonte non è nullo sono sostanzialmente legate a due ordini di motivazioni. La prima riguarda l'aspetto pratico della sperimentazione fin qui condotta, in particolare il fatto che i dati della fonte fiscale sono stati considerati nella loro interezza, nel senso che non è stata applicata, sebbene sia prevista in un prossimo futuro, così come lo è per i dati dell'indagine, una metodologia di studio dei valori anomali e influenti. Inoltre, i dati fiscali non sono stati opportunamente controllati con delle regole di check, valide invece per l'indagine, al fine di rendere i record coerenti.

La seconda riguarda più nello specifico l'aspetto definitivo delle voci considerate nell'analisi e la loro mancata riconciliazione sia in termini di definizione sia numerici. Va ribadito, infatti, che sono state utilizzate diverse fonti, le cui voci, sottostando ad una logica fiscale, sono definite con formulazioni differenti da quelle dello schema SBS. Questo porta a sottolineare non solo l'importanza di un'analisi delle definizioni più approfondita, ma anche l'importanza dell'Istat di partecipare alle operazioni di definizione delle caratteristiche della fonte amministrativa fin dal suo avvio, cioè di intervenire ex-ante, utilizzando schemi concettuali, definizioni e classificazioni del sistema statistico nazionale e dell'Unione europea.

Nelle tabelle successive si presentano in maniera sintetica i confronti (differenze percentuali delle stime di cella) per macro settori di attività economica e classi di addetti della rilevazione per le seguenti variabili:

- ricavi delle vendite e delle prestazioni
- costi per acquisti di beni
- costi per servizi
- costi per godimento di beni di terzi
- costi (acquisti di beni, servizi, godimento beni terzi, oneri diversi)
- costi del personale
- valore aggiunto

In tutte le tavole la differenza percentuale fra la stima dell'insieme delle unità costituite da i dati dell'indagine e dai dati amministrativi (amministrativa) utilizzando il nuovo sistema di pesi e la stima dei dati di indagine (Pmi) utilizzando il vecchio sistema di pesi è stata calcolata secondo la seguente formula:

$$diff\% = \frac{V(Ammministrativa) - V(Pmi)}{V(Pmi)} \cdot 100 \quad (19)$$

Nella tavola 21 è riportata l'analisi delle differenze percentuali fra le stime della variabile "Ricavi delle vendite e delle prestazioni", per macrosettori di attività economica e classi di addetti.

Per la variabile *Ricavi delle vendite e delle prestazioni* la differenza fra le stime con dati integrati e quelle dell'indagine è poco rilevante (solo lo 0,03%), con un effetto fonte che bilancia l'effetto delle mancate risposte totali: in altre parole l'indagine tende a sottostimare i dati (la differenza è a favore dei dati integrati) a causa dell'effetto delle mancate risposte totali, più alto e positivo nel caso delle imprese più grandi più basso e negativo nelle imprese fino a 19 addetti, che viene recuperato da un effetto fonte nella direzione opposta. Differenziati sono poi gli andamenti nei diversi settori, dove si nota che nel settore delle costruzioni l'effetto delle mancate risposte è piuttosto accentuato, soprattutto nelle classi dimensionali centrali.

L'analisi per le altre variabili (tavole 22-27) evidenziano soprattutto un effetto delle mancate risposte totali maggiore dell'effetto della fonte. Gli scostamenti fra le stime con dati integrati e quelle

dell'indagine risultano essere positivi per i *costi per acquisto di beni* (1,20%), *costi per servizi* (1,74%), *costi in complesso* (1,26%), mentre risultano negativi per i *costi per godimento beni di terzi* (-3,19%), *costi del personale* (-0,91%) e per il *valore aggiunto* (-4,50%). Le risultanze assumono una consistenza diversa analizzando le diverse combinazioni di macrosettori e classi di addetti. Ovviamente più si scende nel dettaglio dell'attività economica più le differenze e gli effetti si potranno accentuare o attenuare rispetto all'effetto medio della cella. Rimane comunque da sottolineare che nell'utilizzo dei dati integrati non si è provveduto ad effettuare un'analisi dei casi anomali che potrebbero avere un'influenza importante nelle stime finali.

Con le successive tavole 28, 29 e 30 si è cercato un modo per sintetizzare le differenze percentuali tra le stime con i dati dell'indagine integrati con i dati amministrativi e le stime dell'indagine (insieme 3 e insieme 1) per i diversi domini di stima (attività economica a 2 cifre e classe di addetti, attività economica a 3 cifre e classe di addetti, attività economica a 4 cifre).

La prima tavola (tavola 28) riguarda i domini di stima a 2 cifre di attività economica per classi di addetti. Se consideriamo come esempio la variabile "Ricavi delle vendite e delle prestazioni", su un totale di 189 celle, 77 di esse hanno una differenza percentuale fra valore stimato integrando il resto del campione con i dati amministrativi e quello dall'indagine nell'intervallo [-5%,+5%]. Tali celle corrispondono al 40% dei domini, ma rappresentano il 66,5% dell'aggregato totale. Lo stesso tipo di confronto può essere analizzato per le altre variabili nei diversi domini di stima.

Tavola 21 - Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Ricavi delle vendite e delle prestazioni

SETTORE	Differenza tra le due stime (valori percentuali)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	4,33	-0,89	-0,66	3,13	1,36
Costruzione	-6,92	-7,86	16,01	-3,48	-3,53
Servizi	-0,81	-0,04	0,31	3,99	0,06
Totale	-0,96	-1,25	1,15	3,21	0,03

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	-0,23	-3,30	-0,82	-0,54	-1,16
Costruzione	-0,67	-0,09	-2,62	0,60	-0,76
Servizi	-0,64	-1,21	-2,71	-0,70	-1,09
Totale	-0,59	-1,77	-1,89	-0,56	-1,07

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	0,43	1,41	0,74	0,40	0,73
Costruzione	0,46	0,23	3,29	-0,67	0,82
Servizi	0,03	0,93	2,15	0,59	0,60
Totale	0,15	1,02	1,65	0,44	0,66

SETTORE	Differenza per effetto pesi				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	4,13	1,00	-0,58	3,27	1,80
Costruzione	-6,71	-8,00	15,34	-3,40	-3,59
Servizi	-0,20	0,25	0,87	4,10	0,55
Totale	-0,53	-0,50	1,39	3,33	0,44

Tavola 22 - Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Costi per acquisti di beni

SETTORE	Differenza tra le due stime (valori percentuali)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	3,17	2,32	0,22	3,74	2,24
Costruzione	-4,61	-1,05	9,24	-1,25	-2,01
Servizi	0,09	2,71	-0,91	6,01	1,16
Totale	-0,08	2,29	0,10	4,69	1,20

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	-3,60	-2,39	-1,45	-1,27	-2,04
Costruzione	-2,51	-4,11	-6,42	-1,07	-3,21
Servizi	-5,42	-2,55	-2,28	-1,01	-3,75
Totale	-4,78	-2,63	-2,15	-1,14	-3,20

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	2,05	1,76	1,35	1,04	1,50
Costruzione	1,25	1,82	5,77	0,83	1,97
Servizi	1,69	1,57	2,17	0,85	1,65
Totale	1,71	1,65	2,04	0,94	1,64

SETTORE	Differenza per effetto pesi				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	4,72	2,96	0,32	3,97	2,78
Costruzione	-3,35	1,24	9,89	-1,01	-0,77
Servizi	3,81	3,69	-0,80	6,17	3,25
Totale	2,99	3,27	0,21	4,89	2,76

Tavola 23 - Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Costi per servizi

SETTORE	Differenza tra le due stime (valori percentuali)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	13,93	-6,31	-3,23	3,46	1,24
Costruzione	2,42	-17,60	11,19	-11,86	-1,87
Servizi	-0,39	8,82	5,77	6,49	3,22
Totale	2,13	-1,80	2,90	3,62	1,74

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	6,56	-6,70	1,83	2,05	0,96
Costruzione	0,98	2,32	0,33	1,64	1,24
Servizi	10,81	6,14	1,71	0,61	7,00
Totale	8,11	1,29	1,57	1,36	4,27

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	-3,25	0,75	-1,73	-1,73	-1,62
Costruzione	-0,56	-1,90	1,96	-1,51	-0,44
Servizi	-5,92	-2,47	-1,15	-0,40	-3,53
Totale	-4,26	-1,42	-0,93	-1,10	-2,43

SETTORE	Differenza per effetto pesi				Totale
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	
Industria	10,62	-0,36	-3,33	3,14	1,89
Costruzione	2,00	-18,02	8,90	-11,98	-2,66
Servizi	-5,29	5,15	5,20	6,29	-0,26
Totale	-1,72	-1,66	2,26	3,35	-0,10

Tavola 24: Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Costi per godimento beni di terzi

SETTORE	Differenza tra le due stime (valori percentuali)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	1,92	-0,70	7,89	-1,01	2,47
Costruzione	-15,21	-20,73	15,40	-3,70	-9,15
Servizi	-3,82	1,74	-12,85	-4,46	-4,38
Totale	-4,31	-2,13	-1,28	-2,95	-3,19

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-4,56	-1,23	0,82	-4,15	-2,19
Costruzione	-7,89	-0,02	2,16	-0,50	-3,66
Servizi	-10,39	-5,59	-3,59	-7,70	-8,40
Totale	-9,19	-3,37	-1,16	-5,57	-6,22

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	3,66	1,09	-0,89	4,32	1,84
Costruzione	4,83	0,30	-0,54	0,37	2,19
Servizi	4,09	4,11	2,96	7,08	4,29
Totale	4,13	2,72	0,77	5,31	3,39

SETTORE	Differenza per effetto pesi				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	2,82	-0,56	7,95	-1,18	2,82
Costruzione	-12,15	-21,01	13,78	-3,57	-7,68
Servizi	2,48	3,22	-12,22	-3,85	-0,28
Totale	0,75	-1,48	-0,89	-2,70	-0,37

Tavola 25: Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Costi (acquisto di beni, servizi, godimento beni di terzi, oneri diversi)

SETTORE	Differenza tra le due stime (valori percentuali)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	6,39	-0,18	-0,40	3,59	2,15
Costruzione	-1,36	-8,67	10,81	-6,01	-1,31
Servizi	-0,44	4,12	0,66	5,68	1,32
Totale	0,36	1,19	1,04	4,16	1,26

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-0,70	-3,82	-0,73	-0,52	-1,32
Costruzione	-1,55	-1,39	-2,24	0,17	-1,52
Servizi	-2,79	-0,84	-1,34	-0,85	-1,94
Totale	-2,31	-1,85	-1,16	-0,65	-1,72

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	0,54	1,50	0,68	0,43	0,75
Costruzione	0,67	0,58	3,33	-0,21	1,02
Servizi	0,80	0,78	1,33	0,74	0,87
Totale	0,74	0,97	1,23	0,55	0,86

SETTORE	Differenza per effetto pesi				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	6,56	2,14	-0,35	3,69	2,72
Costruzione	-0,49	-7,85	9,72	-5,96	-0,82
Servizi	1,55	4,19	0,67	5,78	2,39
Totale	1,94	2,06	0,97	4,25	2,12

Tavola 26 - Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Costi del personale

SETTORE	Differenza tra le due stime (valori percentuali)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	3,14	1,94	-1,16	-0,44	0,58
Costruzione	5,01	0,93	0,99	-2,95	2,30
Servizi	-3,35	-3,00	-3,33	-0,94	-2,91
Totale	-0,49	-0,54	-1,82	-0,86	-0,91

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	2,60	-1,14	0,18	0,13	0,30
Costruzione	2,24	-0,19	-0,29	-0,33	0,80
Servizi	1,65	0,44	-1,17	0,44	0,57
Totale	1,95	-0,24	-0,46	0,22	0,51

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-0,83	0,45	-0,19	-0,12	-0,13
Costruzione	-0,83	0,05	0,80	0,24	-0,16
Servizi	-0,98	-0,15	0,87	-0,32	-0,28
Totale	-0,93	0,12	0,38	-0,18	-0,21

SETTORE	Differenza per effetto pesi				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	1,38	2,63	-1,15	-0,45	0,41
Costruzione	3,60	1,07	0,49	-2,86	1,66
Servizi	-4,01	-3,28	-3,03	-1,05	-3,21
Totale	-1,50	-0,42	-1,75	-0,91	-1,21

Tavola 27 - Analisi della differenza tra stime: effetto fonte ed effetto mancate risposte totali: variabile Valore aggiunto.

SETTORE	Differenza tra le due stime (valori percentuali)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-3,24	-3,16	-2,37	1,90	-1,88
Costruzione	-8,16	-8,93	2,19	-1,51	-6,35
Servizi	-5,99	-5,62	-2,79	-3,91	-5,32
Totale	-5,95	-5,25	-2,03	-0,84	-4,50

SETTORE	Differenza per effetto sostituzione fonte (vecchi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-1,22	-3,30	-0,41	-0,58	-1,33
Costruzione	0,21	6,51	-2,49	-9,12	0,51
Servizi	1,53	-0,86	-5,00	1,13	0,26
Totale	0,94	-0,57	-2,62	-0,51	-0,15

SETTORE	Differenza per effetto sostituzione fonte (nuovi pesi)				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	1,06	1,76	0,31	0,55	0,89
Costruzione	-0,19	-3,08	2,30	10,02	0,38
Servizi	-0,89	1,20	3,56	-1,14	0,03
Totale	-0,50	0,83	1,94	0,57	0,34

SETTORE	Differenza per effetto pesi				
	1-9 addetti	10-19 addetti	20-49 addetti	50-99 addetti	Totale
Industria	-3,08	-1,61	-2,27	1,93	-1,43
Costruzione	-8,19	-12,35	2,38	-2,41	-7,24
Servizi	-6,63	-5,96	-1,35	-3,90	-5,62
Totale	-6,39	-5,51	-1,35	-0,90	-4,68

Tavola 28 - Sintesi per domini di stima a 2 cifre di ateco e classe di addetti della differenza tra stime con dati integrati e stime da indagine. Numero di domini e peso della variabile sul totale di cella per classi di differenza percentuale e variabile di studio

	Ricavi		Acquisto beni		Costi beni terzi		Costi servizi		Costi		Costo personale		Valore aggiunto	
	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%
<=-20	8	4,2	20	10,6	23	12,2	15	7,9	12	6,3	5	2,6	8	4,2
-20 , -10	20	10,6	15	7,9	35	18,5	16	8,5	22	11,6	12	6,3	27	14,3
-10 , -5	24	12,7	21	11,1	31	16,4	13	6,9	21	11,1	11	5,8	24	12,7
-5 , -2	18	9,5	22	11,6	12	6,3	17	9,0	17	9,0	29	15,3	36	19,0
-2 , +2	41	21,7	42	22,2	31	16,4	39	20,6	39	20,6	70	37,0	48	25,4
+2 , +5	18	9,5	18	9,5	12	6,3	14	7,4	14	7,4	29	15,3	19	10,1
+5 , +10	23	12,2	16	8,5	12	6,3	24	12,7	19	10,1	16	8,5	13	6,9
+10 , +20	17	9,0	14	7,4	10	5,3	23	12,2	19	10,1	12	6,3	6	3,2
>+20	20	10,6	21	11,1	23	12,2	28	14,8	26	13,8	5	2,6	8	4,2
Totale	189	100,0	189	100,0	189	100,0	189	100,0	189	100,0	189	100,0	189	100,0
-5 , +5	77	40,7	82	43,4	55	29,1	70	37,0	70	37,0	128	67,7	103	54,5

Percentuale di variabile sull'aggregato totale

	Ricavi	Acquisto beni	Costi beni terzi	Costi servizi	Costi	Costo personale	Valore aggiunto
<=-20	1,0	2,5	12,3	6,8	3,1	0,2	0,3
-20 , -10	5,1	3,4	26,2	9,0	3,8	8,4	13,2
-10 , -5	13,1	4,4	11,0	7,9	9,4	3,6	26,4
-5 , -2	11,1	21,4	5,6	10,2	11,1	18,2	30,9
-2 , +2	40,8	41,2	24,2	23,0	43,8	44,9	19,0
+2 , +5	14,5	13,0	3,5	16,7	4,2	18,9	6,6
+5 , +10	5,3	6,0	3,3	9,0	14,0	3,3	2,5
+10 , +20	5,1	4,8	4,6	8,6	6,1	2,4	0,7
>+20	4,0	3,4	9,2	8,7	4,5	0,1	0,5
Totale	100,0	100,0	100,0	100,0	100,0	100,0	100,0
-5 , +5	66,5	75,5	33,3	49,9	59,0	82,0	56,5

Tavola 29 - Sintesi per domini di stima a 3 cifre di ateco e classe di addetti della differenza tra stime con dati integrati e stime da indagine. Numero di domini e peso della variabile sul totale di cella per classi di differenza percentuale e variabile di studio

	Ricavi		Acquisto beni		Costi beni terzi		Costi servizi		Costi		Costo personale		Valore aggiunto	
	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%
<=-20	71	9,2	103	13,4	131	17,1	95	12,4	85	11,1	38	4,9	75	9,8
-20 , -10	93	12,1	91	11,8	121	15,8	73	9,5	97	12,6	58	7,6	120	15,6
-10 , -5	92	12,0	85	11,1	89	11,6	77	10,0	81	10,5	71	9,2	93	12,1
-5 , -2	84	10,9	57	7,4	60	7,8	62	8,1	68	8,9	115	15,0	97	12,6
-2 , +2	138	18,0	125	16,3	112	14,6	136	17,7	125	16,3	220	28,6	174	22,7
+2 , +5	53	6,9	54	7,0	35	4,6	47	6,1	49	6,4	88	11,5	58	7,6
+5 , +10	70	9,1	51	6,6	53	6,9	56	7,3	48	6,3	66	8,6	63	8,2
+10 , +20	65	8,5	72	9,4	58	7,6	75	9,8	89	11,6	55	7,2	43	5,6
>+20	102	13,3	130	16,9	109	14,2	147	19,1	126	16,4	57	7,4	45	5,9
Totale	768	100,0	768	100,0	768	100,0	768	100,0	768	100,0	768	100,0	768	100,0
-5 , +5	275	35,8	236	30,7	207	27,0	245	31,9	242	31,5	423	55,1	329	42,8

Percentuale di variabile sull'aggregato totale

	Ricavi	Acquisto beni	Costi beni terzi	Costi servizi	Costi	Costo personale	Valore aggiunto
<=-20	3,3	4,7	19,8	17,0	5,5	2,7	4,1
-20 , -10	17,9	21,2	27,5	7,5	16,0	8,6	22,3
-10 , -5	9,6	11,1	7,7	10,8	15,9	13,9	18,1
-5 , -2	16,5	7,0	3,8	13,6	8,6	15,9	13,3
-2 , +2	15,7	15,3	9,2	14,8	14,4	27,2	23,7
+2 , +5	7,9	9,2	6,6	4,8	7,9	16,2	6,4
+5 , +10	13,0	6,1	5,9	8,8	5,9	8,8	7,7
+10 , +20	9,6	15,6	7,5	8,6	16,2	4,6	3,1
>+20	6,5	9,9	11,9	14,1	9,6	2,1	1,3
Totale	100,0	100,0	100,0	100,0	100,0	100,0	100,0
-5 , +5	40,1	31,6	19,6	33,2	30,9	59,3	43,4

Tavola 30 - Sintesi per domini di stima a 4 cifre di Ateco della differenza tra stime con dati integrati e stime da indagine. Numero di domini e peso della variabile sul totale di cella per classi di differenza percentuale e variabile di studio

	Ricavi		Acquisto beni		Costi beni terzi		Costi servizi		Costi		Costo personale		Valore aggiunto	
	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%	Num. domini	%
<=-20	33	7,1	55	11,9	82	17,7	35	7,6	40	8,7	12	2,6	32	6,9
-20 , -10	57	12,3	64	13,9	81	17,5	57	12,3	62	13,4	42	9,1	75	16,2
-10 , -5	60	13,0	41	8,9	39	8,4	41	8,9	56	12,1	49	10,6	78	16,9
-5 , -2	49	10,6	35	7,6	38	8,2	35	7,6	31	6,7	63	13,6	65	14,1
-2 , +2	69	14,9	74	16,0	52	11,3	62	13,4	74	16,0	129	27,9	86	18,6
+2 , +5	40	8,7	26	5,6	23	5,0	39	8,4	28	6,1	61	13,2	41	8,9
+5 , +10	60	13,0	44	9,5	30	6,5	52	11,3	54	11,7	53	11,5	36	7,8
+10 , +20	47	10,2	58	12,6	53	11,5	67	14,5	58	12,6	25	5,4	31	6,7
>+20	47	10,2	65	14,1	64	13,9	74	16,0	59	12,8	28	6,1	18	3,9
Totale	462	100,0	462	100,0	462	100,0	462	100,0	462	100,0	462	100,0	462	100,0
-5 , +5	158	34,2	135	29,2	113	24,5	136	29,4	133	28,8	253	54,8	192	41,6
Percentuale di variabile sull'aggregato totale														
	Ricavi	Acquisto beni	Costi beni terzi	Costi servizi	Costi	Costo personale	Valore aggiunto							
<=-20	3,1	6,1	17,0	5,2	4,9	2,9	3,6							
-20 , -10	19,7	22,6	27,3	10,8	11,0	9,3	17,3							
-10 , -5	10,7	10,1	3,4	20,5	20,9	11,5	28,5							
-5 , -2	14,1	7,2	11,8	10,7	8,3	16,3	17,5							
-2 , +2	14,8	11,6	10,3	12,2	15,6	24,9	14,3							
+2 , +5	7,3	9,1	5,1	12,5	5,0	20,8	6,1							
+5 , +10	17,5	17,3	5,2	8,1	19,5	8,9	6,8							
+10 , +20	7,2	8,6	7,8	11,9	8,9	3,7	4,9							
>+20	5,6	7,4	12,0	8,1	6,0	1,7	1,0							
Totale	100,0	100,0	100,0	100,0	100,0	100,0	100,0							
-5 , +5	36,1	27,9	27,3	35,4	28,8	62,0	37,9							

3.4.4. Un'applicazione dell'utilizzo delle fonti amministrative nella rilevazione PMI2008

Sulla base dell'attività sperimentale condotta sulla rilevazione PMI2007, si è esteso l'utilizzo dei dati di fonte amministrativa (limitatamente ai *Bilanci civilistici* e agli *Studi di Settore* nei *Quadri F e G*) alla rilevazione PMI2008. Il campione di questa rilevazione, pari a 104.862 unità, è stato estratto da ASIA in modo da rispettare i domini di stima del nuovo regolamento SBS 295/2008 che richiede la diffusione di dati a livello di Nace Rev.2 (Ateco2007). Il numero di imprese rispondenti direttamente alla rilevazione con dati validi è stato di 36.584 unità. L'integrazione per la mancata risposta totale è stata effettuata, quindi, utilizzando come fonti i *Bilanci civilistici* e gli *Studi di Settore*. Per quest'ultima fonte è stato necessario introdurre una metodologia di imputazione basata sul donatore per disaggregare le variazioni complessive delle rimanenze di prodotti finiti, semilavorati, in corso di lavorazione, materie prime e merci nelle componenti variazioni delle rimanenze di prodotti finiti, semilavorati e in corso di lavorazione e nelle variazioni delle rimanenze di materie prime e merci. Inoltre la fonte fiscale non fornisce una disaggregazione del costo del lavoro nelle sue componenti e quindi, in questo caso, si è effettuata un'imputazione legata alle risultanze desumibili dalla rilevazione stessa.

Questo processo ha consentito di recuperare 22.511 imprese dalla fonte *Bilanci*, 22.209 imprese dalla fonte *Studi di Settore (Quadro F)* e 2.340 imprese dalla fonte *Studi di Settore (Quadro G)*. La metodologia per ricostruire tutte le informazioni del questionario di rilevazione ha ricalcato, con qualche accorgimento, quella utilizzata nelle precedenti occasioni di indagine (vedere l'appendice A.2).

Conseguentemente il campione su cui è stata applicata la procedura di calibrazione per la determinazione dei pesi finali è risultato pari a 81.726 imprese, che rappresentano il 77,9% del campione teorico, dal quale dovrebbero essere eliminate le imprese non eleggibili in quanto cessate, inattive, in liquidazione, fuori campo di osservazione e quelle eventualmente non risolte (ovvero quelle imprese non trovate all'indirizzo in quanto trasferite o con indirizzo errato).

L'applicazione effettuata ha mostrato come, nel breve periodo, l'uso dei dati amministrativi può rivelarsi cruciale per il miglioramento non solo della qualità dei dati (in particolare attraverso il recupero di informazioni su variabili chiave sui non rispondenti all'indagine), ma anche della tempestività dei risultati, con una riduzione del fastidio sui rispondenti e quindi dei costi di produzione. I risultati economici e strutturali della rilevazione PMI2008 saranno diffusi congiuntamente a quelli

della rilevazione SCI2008 nella statistica in breve sulla struttura e competitività del sistema delle imprese industriali e dei servizi e attraverso un datawarehouse in corso di predisposizione.

4. Conclusioni e prospettive future

La disponibilità di fonti di dati amministrativi sulle imprese sempre più stabili, ricche in termini di contenuti e copertura, e consistenti con le definizioni e gli obiettivi della Statistica Ufficiale, ha reso possibile per l'Istat l'avvio di progetti volti ad incrementare il loro utilizzo nei processi di produzione dell'informazione statistica.

In particolar modo nell'area delle statistiche economiche questa tendenza, che si trova in linea con la "visione" di medio-lungo periodo manifestata dall'Eurostat, implica investimenti importanti in termini sia di sviluppi metodologici, sia di gestione delle trasformazioni organizzative connesse ad una così radicale mutazione di impostazione del processo produttivo, in cui il ricorso alle indagini dirette assume un ruolo marginale rispetto allo sfruttamento massivo, coordinato e standardizzato delle fonti di informazione già esistenti sulle popolazioni di riferimento.

A questi "costi" metodologico/organizzativi fanno però da contrappeso i numerosi vantaggi che possono derivare da una siffatta impostazione del processo produttivo in termini di riduzione del fastidio sulle imprese, aumento della qualità dei dati (anche grazie al contenimento del fenomeno della mancata risposta totale), aumento del livello di coerenza delle stime prodotte per i vari domini di imprese per i diversi fenomeni sotto regolamento, riduzione dei tempi e dei costi di acquisizione, trattamento e rilascio dei dati.

Diverse esperienze in campo internazionale hanno mostrato come il dato amministrativo possa essere utilizzato, in maniera diretta o indiretta, per aumentare l'efficienza del processo di stima nell'area delle indagini economiche (sia congiunturali, sia strutturali). In questo lavoro l'attenzione è stata focalizzata sul settore delle statistiche strutturali sulle imprese (SBS), ed in particolare sull'indagine sulle piccole e medie imprese (PMI), al fine di verificare l'impatto sulla qualità delle stime obiettivo della rilevazione di un utilizzo più consistente delle fonti di dati amministrativi attualmente disponibili sulle imprese (in particolare, fonti fiscali).

Con riferimento all'anno 2007 e ad alcune variabili chiave dell'indagine, la sperimentazione condotta ha mostrato che: (i) i dati amministrativi (bilanci civilistici e dati fiscali) sono sufficientemente armonizzati a livello di definizioni, anche a seguito di opportuni pre-trattamenti e riclassificazione delle voci della modulistica analizzata; (ii) sono consistenti, a livello di distribuzioni, con i dati rilevati dall'indagine PMI; (iii) si ottengono risultati di accuratezza soddisfacente per la maggior parte dei domini di stima pianificati per la rilevazione, sia in caso di utilizzo dei dati fiscali in fase di integrazione delle mancate risposte totali dell'indagine (quindi nell'ipotesi di mantenimento dell'attuale assetto organizzativo della rilevazione), sia in caso di utilizzo dei dati fiscali come fonte primaria di informazione sulle imprese.

Tale sperimentazione ha permesso, per l'indagine PMI successiva, relativa all'anno 2008, di integrare quasi la metà del campione teorico iniziale riducendo notevolmente i tassi di mancata risposta totale, eliminando possibili distorsioni causate dall'autoselezione dei non rispondenti e utilizzando successivamente tecniche di imputazione statistiche per stimare le informazioni non direttamente desumibili dalle fonti amministrative.

Questi risultati incoraggiano a proseguire nella direzione di ulteriori sviluppi metodologici e alla realizzazione di ulteriori sperimentazioni volte a verificare la fattibilità di una riorganizzazione complessiva del processo di produzione delle statistiche che rispondono al regolamento SBS, volto alla realizzazione di un sistema (*frame*) di stime, coerenti sia trasversalmente sia longitudinalmente, basato sull'integrazione fra dati di fonte amministrativa e di fonte statistica, dove questi ultimi rivestano però un ruolo ancillare rispetto ai primi, essendo ottenuti sulla base di disegni di indagine e procedure di stima *condizionati* alla base informativa pre-esistente regolarmente alimentata dalle fonti amministrative.

Riferimenti bibliografici

- Agenzia delle Entrate. Studi di Settore 2008. Periodo d'imposta 2007. Modello per la comunicazione dei dati rilevanti ai fini dell'applicazione degli Studi di Settore. Istruzioni Parte Generale e Modelli da allegare alla dichiarazione dei redditi. <http://www.agenziaentrate.gov.it/>
- Agenzia delle entrate. Modelli di dichiarazione Unico 2008. Periodo d'imposta 2007. Istruzioni per la compilazione e Modelli di dichiarazione Unico PF, Unico SP, Unico SC, Unico Irap. <http://www.agenziaentrate.gov.it/>
- Bethel, J. (1989), Sample allocation in multivariate surveys. *Survey methodology*, 15 (1989): 47-57.
- S. Berg, C. Hertzman and M. Johansson (2006), Developing and implementing a New SBS at Statistics Sweden, European Conference on Quality in Official Statistics (Q2006), Cardiff (UK).
- Bikker R., Daalmans J., Mushkudiani N. (2010), *A multivariate Denton method for benchmarking large data sets*, discussion paper, Statistics Netherlands.
- Biffignandi S., Nascia L., Zeli A. (2009), Survey and Administrative Data Mix in a Business Survey, Proceedings of the European Establishment Statistics Workshop (EESW09), 7-9 September, Stockholm.
- Branchi F., Casciano M.C., Di Francescantonio T., Falorsi P.D., Filiberti S., Marsigliesi G., Sansone U., Santi E., Sanzo R., Siesto G., Zeli A. (2006), *Valutazione delle possibilità di uso di dati fiscali a supporto della rilevazione PMI*, Documenti, Istat, n.17.
- Branchi F., Casciano M.C., Di Francescantonio T., Falorsi P.D., Filiberti S., Marsigliesi G., Sansone U., Santi E., Sanzo R., Siesto G., Zeli A. (2008), *Messa a regime dell'uso dei dati fiscali (modello Unico) per l'integrazione delle mancate risposte totali e la riduzione del numero delle unità campione della rilevazione PMI*, Documenti, Istat, n.4.
- Brion P., Gros E. (2009), Methodological issues related to the reengineering of the French structural business statistics, Proceedings of the European Establishment Statistics Workshop (EESW09), Stockholm (Sweden).
- Casciano M.C., De Giorgi V., Oropallo F., Siesto G. (2010), Experimental Analysis in the estimation of SBS variables for small firms by using administrative data, paper presented at the Seminar on Using Administrative Data in the Production of Business Statistics - Member States Experiences, Rome, 16-18 Marzo, http://www.ine.pt/filme_inst/essnet/papers/Session3/Paper3.3.pdf
- Cerroni F., De Giorgi, V. (2008) The Tax Authority Source as an example of the use of an administrative source as a statistical one. Proceedings of the IAOS Conference on Reshaping Official Statistics 2008. Shanghai, 14-16 Ottobre. http://www.stats.gov.cn/english/specialtopics/iaos/Papers/CS_11_4_Giorgi_Cerroni.doc
- Commission Regulation on Statistical Units: n. 2186/93 and n. 696/93; on Structural Business Statistics: n. 58/97, 410/98, 2700/98, 2056/02, 1670/03; on Evaluation of Quality of Structural Business Statistics: n. 1618/99.
- Chumbau A., Neves C., Pereira H. (2008), Simplified business information: Improving quality by using administrative data in Portugal, European Conference on Quality in Official Statistics (Q2008), Rome.
- Chumbau A., Pereira H. J., Rodrigues S. (2010), Simplified Business Information (IES): Impact of Admin Data in the production of Business Statistics, presented at the Seminar on Using Administrative Data in the Production of Business Statistics - Member States Experiences, Rome, 18-19 Marzo, http://www.ine.pt/filme_inst/essnet/papers/Session3/Paper3.6.pdf.

- Daas P.J.H., Arends-Tòth J., Schouten B., Kuijvenhoven L. (2008), Quality Framework for the evaluation of Administrative data, Proceedings of the European Conference on Quality in Official Statistics (Q2008), Roma, 8-11 Luglio,
<http://www.pietdaas.nl/beta/pubs/pubs/21Daas.pdf>
- Daas P.J.H., and Fonville T.C. (2007), Quality control of Dutch Administrative Registers: An Inventory of Quality Aspects, paper presented at the Seminar on Registers in Statistics - methodology and quality, Helsinki.
- Deville, J.C., Särndal, C.E. (1992,) Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87 (1992): 376-382.
- De Vitiis C., Falorsi P.D. (2009), Rapporto conclusivo delle attività del “Gruppo di Lavoro per la definizione delle metodologie necessarie a gestire le problematiche connesse all’adozione della Ateco 2007”, Istat.
- Elliott, D., (2010), The potential use of additional VAT data in ONS business surveys, Proceedings of the European Conference on Quality in Official Statistics (Q2010), Helsinki, 4-6 May.
- Elliott, D., van Elswijk D., Luzi O., Siesto G., Redling B., Kavaliauskiene D. (2010), Methods of estimation for business statistics which cannot be obtained from administrative data sources, Proceedings of the European Conference on Quality in Official Statistics (Q2010), Helsinki, 4-6 Maggio, http://q2010.stat.fi/media/presentations/Elliott_Q2010_WP3_paper.doc
- Eurostat, European Commission (1999), *Use of Administrative Sources for Business Statistics Purpose, Handbook on good practices*, 1999 Edition.
- Eurostat (2003a), *Definition of Quality in statistics*. Eurostat Working Group on Assessment of Quality in Statistics, Luxembourg, 2-3 October.
- Eurostat (2003b), *Item 6 - Quality Assessment of Administrative Data for Statistical Purposes*, Luxembourg.
- Eurostat, European Commission (2003c) “*Manual for Structural Business Statistics*” Directorate D: Business statistics - Unit D-2: Structural business statistics - July 2003
- Eurostat (2007), *Recommended Practices for Editing and Imputation in Cross-Sectional Business Surveys*,
http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/RPM_EDIMBUS.pdf
- Eurostat (2009), *Insights on Data Integration Methodologies*,
http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-RA-09-005.
- Falorsi, P.D., Ballin, M., De Vitiis, C., Scepi, G. (1998), Principi e metodi del software generalizzato per la definizione del disegno di campionamento nelle indagini sulle imprese condotte dall’Istat. *Statistica Applicata*, 20, n. 2 (1998).
- P.D.Falorsi, S.Filiberti, S.Pallara (2004), L’uso di dati di fonte amministrativa per migliorare la qualità delle statistiche strutturali sulle imprese”, presentazione al convegno su Utilizzo delle nuove strategie di formazione dei dati e di informazione statistica per lo studio e il governo dei sistemi economici, Bergamo, 10 maggio.
- P.D.Falorsi, M.C.Casciano, S.Filiberti, D.Orsini, A.Pallara, A.Russo (2005), Definizione della strategia di indagine della rilevazione sui Conti economici delle Imprese mediante l’uso congiunto di dati statistici e amministrativi, in *L’integrazione di dati di fonti diverse*, Franco Angeli.
- Granquist L., and Kovar J. (1997), Editing of Survey Data: How Much is Enough?, in: *Survey Measurement and Process Quality*, ed. Lyberg, Biemer, Collins, De Leew, Dippo, Schwart, and Trewin, John Wiley & Sons, pp.415-435.

- Haitzmann M. (2010), Model based estimation of enterprises below thresholds in Structural Business Statistics, paper presented at the Seminar on Using Administrative Data in the Production of Business Statistics - Member States Experiences, Roma, 16-18 Marzo, http://www.ine.pt/filme_inst/essnet/papers/Session1/Paper1.4.pdf
- Hoogland J. (2010), Editing Strategies for VAT Data, paper presented at the Seminar on Using Administrative Data in the Production of Business Statistics - Member States Experiences, Rome, 16-18 Marzo, http://www.ine.pt/filme_inst/essnet/papers/Session1/Paper1.1.pdf
- Houbiers M., Knotterus P., Kroese A.H., Renssen R.H., Snijders V. (2010), *Estimating consistent table sets: position paper on repeated weighting*, discussion paper, Statistics Netherlands.
- IMF (2004), *The Fourth Review of the Fund's Data Standards Initiatives*, Supplement on the Data Quality Assessment Framework. <http://dsbb.img.org/dqrsindex.htm>
- ISTAT (2007), *Conti economici delle imprese - Anno 2003*.
- ISTAT (2010), *Conti economici delle imprese - Anno 2007*
- Lewis D. (2010) Integrating data from different sources, in the production of business statistics (WP5), Proceedings of the European Conference on Quality in Official Statistics, (Q2010), Helsinki, 4-6 May.
- Lorenz R. (2010), The integrated System of Editing Administrative Data for STS in Germany, paper presented at the Seminar on Using Administrative Data in the Production of Business Statistics - Member States Experiences, Roma, 16-18 Marzo. http://www.ine.pt/filme_inst/essnet/papers/Session1/Paper1.2.pdf
- Luzi O., Perani G., Seri G. (2009), The Use of Administrative Fiscal Data for the Production of Research and Development Statistics in Italy, Proceedings of the European Conference on Quality in Official Statistics, (Q2010), Helsinki, 4-6 Maggio.
- ONS (2005) *Guidelines for measuring statistical Quality*, version 3.0, London: Office of National Statistics.
- Orjala H. (2008), Potential of administrative data in business statistics – a special focus in improvements in short term statistics, Statistics Finland, - IAOS Conference on Reshaping Official Statistics - Shanghai, 14-16 Ottobre, http://www.stats.gov.cn/english/specialtopics/iaos/Papers/CS_16_4_Orjala.doc
- Rao, J. N. K. (2003), *Small Area Estimation*, New Jersey: John Wiley & Sons.
- Särndal, C.E., Swensson, B., Wretman, J. (1992), *Model Assisted Survey Sampling*. New York: Springer-Verlag, 1992
- Singh, A. C., Mohl, C. A. (1996), Understanding Calibration Estimators in Survey Sampling, *Survey Methodology*, vol. 22, n. 2, pp. 107-115.
- Sturm R. (2007), Business Register as a source for multi-source-analyses, DESTATIS, Wiesbaden Group on BR.
- Tolkki V. (2007), Finnish SBS System: use of administrative data, methods and process, presented at the Seminar on Reengineering of Business Statistics, Lisbona, 11-12 Ottobre.
- van Duin C., Snijders V. (2010), *Simulation studies of Repeated Weighting*, discussion paper, Statistics Netherlands.
- Yung, W., K. Cook and S. Thomas (2004), Use of GST Data by the Monthly Survey of Manufacturing, Proceedings of the Joint Statistics Meeting, Toronto, 5-12 Agosto.

Yung, W., Lys P. (2008), “Use of Administrative Data in Business Surveys - The Way Forward”
Statistics Canada - IAOS Conference on Reshaping Official Statistics - Shanghai, 14-16 Ot-
tobre, http://www.stats.gov.cn/english/specialtopics/iaos/Papers/CS_8_1_Yung_Lys.pdf

Wallgren A., Wallgren B. (2006), Register-Based economic statistics on enterprises – Editing Is-
sues, UNECE Work Session on Statistical Data Editing, Bonn (Germany).

Wallgren A., Wallgren B. (2007), *Register-based Statistics: Administrative Data for Statistical
Purposes*, John Wiley & Sons.

Zach S. (2007), Model based estimation of enterprises below thresholds in Austria, presented at the
Seminar Reengineering of Business Statistics, Lisbona, 11-12 Ottobre.

Appendice A: Fonti statistiche e amministrative sulle imprese disponibili per l'Istat

A.1. Archivio statistico delle imprese attive

L'archivio statistico delle imprese attive (ASIA) è costituito dalle unità economiche che esercitano arti e professioni nelle attività industriali, commerciali e dei servizi alle imprese e alle famiglie. L'archivio è aggiornato annualmente attraverso un processo di integrazione di informazioni provenienti da fonti di diversa natura, amministrative e statistiche, e fornisce informazioni anagrafiche (denominazione e localizzazione) e di struttura (attività economica nella classificazione Ateco2002 e Ateco2007, addetti, dipendenti e indipendenti, forma giuridica e fatturato) delle unità.

Le principali fonti amministrative utilizzate sono:

- gli archivi gestiti dall'Agenzia delle Entrate per il Ministero dell'Economia e delle Finanze, quali l'Anagrafe Tributaria, le dichiarazioni annuali delle imposte indirette, le dichiarazioni dell'imposta regionale sulle attività produttive (Irap), gli Studi di Settore;
- i registri delle imprese delle Camere di Commercio, Industria, Artigianato e Agricoltura e gli archivi collegati dei soci delle Società di Capitale delle "persone" con cariche sociali;
- gli archivi dell'Istituto Nazionale di Previdenza Sociale, relativamente alle posizioni contributive dei dipendenti delle imprese e a quelle degli artigiani e commercianti;
- l'archivio delle utenze telefoniche;
- l'archivio dei Bilanci consolidati e di esercizio;
- l'archivio degli Istituti di credito gestito dalla Banca D'Italia;
- l'archivio delle società di assicurazioni gestito dall'Isvap.

Le principali fonti statistiche utilizzate sono invece:

- l'indagine sulle unità locali delle grandi imprese (IULGI);
- le indagini strutturali che l'Istat effettua sulle imprese, in particolare quelle sul sistema dei conti delle imprese (SCI), sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI), sul commercio con l'estero e sulla produzione industriale;
- le indagini congiunturali che l'Istat effettua sulle imprese, in particolare quella sulla produzione, fatturato e ordinativi delle attività manifatturiere, sul fatturato del commercio e dei servizi, sui prezzi alla produzione e sull'occupazione della grande impresa.

L'archivio ASIA rappresenta il quadro di riferimento per l'estrazione dei campioni e per il rapporto all'universo delle diverse rilevazioni effettuate dall'Istat; rappresenta, inoltre, il riferimento per aggiornare alcune informazioni strutturali sulle imprese (*attività economica, addetti e dipendenti*) e per agganciare tutte le fonti amministrative disponibili attraverso il codice fiscale.

A.2. Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni

In questa appendice vengono fornite ulteriori informazioni sulla rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni ed, in particolare, sul disegno campionario (paragrafo A.2.1.), la procedura di integrazione con dati sui Bilanci civilistici (paragrafo A.2.2.), la metodologia di stima (paragrafo A.2.3.) e sulle principali variabili presenti nella rilevazione (A.2.4.).

A.2.1. Il disegno campionario

La rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni ha lo scopo di fornire stime dei totali e dei valori medi per alcune variabili di interesse riferite a tre tipi di domini, ottenuti come combinazioni delle modalità assunte dalle seguenti variabili:

- divisione di attività economica (Ateco a 2 cifre);
- gruppo di attività economica (Ateco a 3 cifre);
- classe di attività economica (Ateco a 4 cifre);
- regioni amministrative (considerando separatamente le province di Trento e Bolzano);
- classi dimensionali di addetti, stabilite in modo differente a seconda della divisione di attività economica.

Tavola 1 - Domini di studio e loro denominazione

TIPO DI DOMINIO	Descrizione	N° di domini
DOM1	Classe di attività economica	462
DOM2	Gruppo di attività economica × Classi addetti	830
DOM3	Divisione di attività economica × Regione	966

Le classi dimensionali in termini di addetti sono descritte nella tavola 2:

Tavola 2 - Definizione delle classi dimensionali di addetti

DIVISIONI DI ATTIVITÀ ECONOMICA (Ateco2002)	Classi di addetti	Etichette
Industria: 10-45	1-9	I1
	10-19	I2
	20-49	I3
	50-99	I4
Servizi: 50, 51, 52	1	SA1
	2-9	SA2
	10-19	SA3
	20-49	SA4
	50-99	SA5
Servizi: 55, 60, 61, 62, 63, 64, 67, 70, 71, 72, 73, 74	1	SB1
	2-9	SB2
	10-19	SB3
	20-49	SB4
	50-99	SB5
Servizi: 80, 85, 90, 92, 93	1-9	SC1
	10-19	SC2
	20-49	SC3
	50-99	SC4

Poiché il regolamento sulle statistiche strutturali richiede stime senza limiti di fascia dimensionale delle imprese per i domini del tipo DOM1 e DOM3, lo studio del disegno di campionamento e l'allocazione del campione è avvenuta considerando congiuntamente questa indagine con quella censuaria (rilevazione sul sistema dei conti delle imprese) sulle imprese con oltre 100 addetti (10.425 imprese censite).

Il disegno di campionamento adottato è ad uno stadio stratificato, con selezione delle unità con probabilità uguali; gli strati sono definiti dalla concatenazione delle modalità delle variabili 'regione', 'classi di attività economica' (Ateco a 4 cifre) e 'classi di addetti', secondo quanto descritto nei due prospetti precedenti.

La stratificazione adottata costituisce la partizione minima della popolazione che permette di ottenere i domini di stima come aggregazione di strati elementari. Il fatto che tutti i domini di interesse sono di tipo stratificato presenta alcuni considerevoli vantaggi dal punto di vista della progettazione dell'indagine. In particolare, ciò permette di allocare il campione pianificando i livelli di precisione attesa delle stime su tutti i domini di interesse.

Tuttavia, come si è già accennato, al momento della progettazione del campione, l'archivio può non risultare aggiornato all'anno di riferimento della rilevazione; ciò implica che, stratificando la lista di selezione secondo i criteri illustrati sopra, si possono determinare strati teorici in cui vengono incluse imprese che nella realtà si collocano in strati differenti; pertanto, la migrazione di unità tra gli strati dovuta a questa mancata coincidenza tra lista di selezione e popolazione d'interesse fa sì che i livelli di errore sui domini siano solo in parte pianificabili.

In generale, quando un'indagine ha l'obiettivo di fornire stime per più domini di studio e per diverse variabili di interesse, il campione deve essere progettato in modo tale da assicurare simultaneamente l'accuratezza delle stime ai diversi livelli di dettaglio richiesti.

Prima di illustrare il metodo utilizzato per risolvere questo problema di allocazione multivariata e multidominio, è utile introdurre alcune notazioni simboliche.

- ${}_p y$ = la generica caratteristica di interesse p ;
- d = il generico tipo di dominio di studio, $d=1, \dots, D$;
- j_d = il generico dominio di tipo d , $j_d = 1, \dots, J_d$ con J_d = numero di domini del tipo d ;
- h = il generico strato, $h=1, \dots, H$ dove H è il numero di strati in cui è suddivisa la popolazione d'indagine;
- H_{j_d} = il numero di strati che contengono il dominio j_d ;
- N_h = il numero di unità dell'universo che appartengono allo strato h ;
- n_h = il numero di unità del campione che appartengono allo strato h ;
- ${}_p y_{hi}$ = il valore della caratteristica p osservato sulla generica unità i appartenente allo strato h ;
- ${}_p Y_h$ = il totale di popolazione della caratteristica ${}_p y$ nello strato h ;
- ${}_p \bar{Y}_h = \sum_{i=1}^{N_h} \frac{{}_p y_{hi}}{N_h}$ la media di ${}_p y$ nello strato h ; (1)
- ${}_p S_h^2 = \sum_{i=1}^{N_h} \frac{({}_p y_{hi} - {}_p \bar{Y}_h)^2}{N_h}$ la varianza di ${}_p y$ nello strato h ; (2)
- ${}_p \tilde{Y}_h = \frac{N_h}{n_h} \sum_{i=1}^{n_h} {}_p y_{hi}$ la stima diretta (o di Horvitz-Thompson) di ${}_p Y_h$ nello strato h ; (3)
- $\text{Var}({}_p \tilde{Y}_h) = N_h^2 \frac{(1-f_h)}{n_h} {}_p S_h^2$ la varianza della stima (3) nello strato h , con $f_h = \frac{n_h}{N_h}$; (4)
- ${}_p \tilde{Y}_{j_d} = \sum_{h=1}^{H_{j_d}} {}_p \tilde{Y}_h$ la stima di Horvitz-Thompson del totale di ${}_p y$ nel dominio j_d ; (5)
- $\text{Var}({}_p \tilde{Y}_{j_d}) = \sum_{h=1}^{H_{j_d}} N_h^2 \frac{(1-f_h)}{n_h} {}_p S_h^2$ la varianza della stima (5) nel dominio j_d . (6)

È facile verificare che la varianza (6) della stima del totale della variabile ${}_p y$ nel dominio j_d , propria di un campione stratificato con selezione delle unità senza reimmissione e con probabilità uguali, può essere decomposta come segue:

$$\text{Var}({}_p \tilde{Y}_{j_d}) = {}_p V_{j_d} + {}_p V_{j_d}^0 = \sum_{h=1}^{H_{j_d}} \frac{N_h^2}{n_h} {}_p S_h^2 - \sum_{h=1}^{H_{j_d}} N_h {}_p S_h^2 \quad (7)$$

dove ${}_p V_{j_d}^0$ è la parte di varianza non influenzata dall'allocazione.¹⁸

¹⁸ Falorsi, P.D., e M. Ballin, e C. De Vitiis, e G. Scepi. "Principi e metodi del software generalizzato per la definizione del disegno di campionamento nelle indagini sulle imprese condotte dall'Istat". *Statistica Applicata*, 20, n. 2 (1998).

La ricerca dell'allocazione ottima prevede che sia definita una funzione C' con cui quantificare il costo della rilevazione; nel caso in esame, la funzione di costo C' è stata esplicitata come segue:

$$C' = C_o + C = C_o + \sum_{h=1}^{H_{jd}} C_h n_h \quad (8)$$

in cui C_o rappresenta il costo fisso dell'indagine che non dipende né dalla numerosità campionaria né dall'allocazione, C il costo variabile e C_h ($h=1, \dots, H$) il costo per unità campionaria relativo allo strato h (che è stato assunto pari ad 1 in tutti gli strati).

Nel determinare la numerosità da assegnare a ciascuno strato, un approccio consiste nel fissare, per ciascun ${}_p V_{j_d}$, un estremo superiore della varianza ${}_p V_{j_d}^*$ della stima del totale di ${}_p y$ e nel minimizzare la funzione di costo C sotto i vincoli ${}_p V_{j_d} \leq {}_p V_{j_d}^*$ ($p=1, \dots, P$; $d=1, \dots, D$; $j_d=1, \dots, J_d$).

Tale approccio è alla base della metodologia sviluppata dall'Istat per la determinazione della dimensione campionaria ottima in presenza di molteplici vincoli ed obiettivi: essa adotta infatti una generalizzazione della soluzione proposta da Bethel,¹⁹ in cui viene definito un problema di minimo vincolato con funzione obiettivo convessa e vincoli di tipo lineare.

Il problema consiste dunque nel minimizzare la funzione obiettivo:

$$f(nh) = \sum_{h=1}^{H_{jd}} C_h n_h = \min \quad (9)$$

sotto un sistema di vincoli che si esplicita nella forma:

$$\sum_{h=1}^{H_{jd}} \frac{N_h^2}{n_h} {}_p S_h^2 - \sum_{h=1}^{H_{jd}} N_h {}_p S_h^2 \leq {}_p V_{j_d}^* \quad (p=1, \dots, P; d=1, \dots, D; j_d=1, \dots, J_d) \quad (10)$$

L'algoritmo proposto da Bethel per il calcolo dell'allocazione individua la soluzione ottima in maniera iterativa, partendo da una soluzione iniziale che coincide con la soluzione ottima nel caso univariato per la prima variabile sul primo dominio. In ciascuno dei passi successivi, la numerosità campionaria viene aumentata minimizzando la funzione obiettivo fino al soddisfacimento di tutti i vincoli. Bethel dimostra che tale algoritmo converge.

Sotto l'aspetto applicativo, le unità di ASIA facenti parte dell'universo di indagine sono state classificate secondo le modalità concatenate dei caratteri 'regione', 'classi di attività economica' e 'classi di addetti': il numero degli strati così costruiti, contenenti almeno un'impresa, è risultato pari a 25.507, con una distribuzione per classi di numerosità delle imprese negli strati illustrata nella tavola 3.

¹⁹ Bethel, J. "Sample allocation in multivariate surveys". Survey methodology, 15 (1989): 47-57.

Tavola 3 - Distribuzione degli strati per numero di imprese

CLASSE DI AMPIEZZA DEGLI STRATI (n. di imprese)	Freq. Assoluta	Freq. Cumulata	Freq. %	Freq. % cumulata
1	4.735	4.735	18,56	18,56
2	2.492	7.227	9,77	28,33
3-5	3.799	11.026	14,89	43,23
6-10	2.922	13.948	11,46	54,68
Oltre 10	11.559	25.507	45,32	100,00

Poiché per i domini del tipo DOM1 e DOM3 il regolamento richiede stime senza limiti di fasce dimensionali, oltre ai 25.507 strati campionati si sono dovuti considerare, in fase di allocazione, anche gli strati cosiddetti censiti, ossia quelli in cui si collocano le 10.425 imprese della fascia 100 addetti ed oltre soggetti alla rilevazione Sci. Nella progettazione del campione si è pertanto tenuto conto della necessità di includere tali strati completamente osservati e di non superare, per ragioni logistiche e di costo, la soglia di 120 mila unità cui inviare il questionario.

La ricerca dell'allocazione ottima è avvenuta assumendo *fatturato* e *valore aggiunto* come variabili d'interesse.

Nell'impostazione del problema di ottimizzazione, definito dalle relazioni (9) e (10), si è preferito esprimere i vincoli di precisione sulle stime di tali variabili in termini di coefficienti di variazione anziché in termini di varianze attese; in tal modo il problema di ottimo è stato risolto utilizzando come vincoli un'espressione equivalente alla (10), data da:

$$\frac{\sqrt{\sum_{h=1}^{H_{jd}} \frac{N_h^2}{n_h} {}_p S_h^2 - \sum_{h=1}^{H_{jd}} N_h {}_p S_h^2}}{\tilde{Y}_{jd}} \leq {}_p CV_{jd}^* \quad (p=1, \dots, P; d=1, \dots, D; j_d=1, \dots, J_d) \quad (11)$$

dove si sono sostituite le quantità ${}_p V_{jd}^*$ nella (10) con i coefficienti di variazione:

$${}_p CV_{jd}^* = \frac{\sqrt{{}_p V_{jd}^*}}{\tilde{Y}_{jd}} \quad (12)$$

Il dimensionamento del campione è stato, dunque, ottenuto imponendo alle stime dei totali di *fatturato* e *valore aggiunto* gli errori massimi pari a 11% per il dominio DOM1, a 9% per il dominio DOM2 ed a 11% per il dominio DOM3 (Tavola 4). Questi vincoli hanno determinato un'allocazione complessiva pari a 113.648 imprese (incluse le 10.425 della rilevazione SCI).

Tavola 4 - Errore massimo imposto nella stima del fatturato e del valore aggiunto in ciascuna tipologia di dominio e dimensione del campione

ERRORE MASSIMO ATTESO	DOM1	DOM2	DOM3	Numerosità campionaria (imprese della fascia 1-99 addetti)	Numerosità complessiva (incluse le imprese censite)
${}_p CV_{jd}^*$	11%	9%	11%	103.223	113.648

La procedura di allocazione impiegata ha richiesto la specificazione a priori di medie e varianze per ogni strato della popolazione e per ognuna delle variabili di interesse. Tali quantità sono state stimate direttamente sulla base dei dati osservati in ciascuno strato nella precedente *rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni* (PMI, fino a 99 addetti) e nella rile-

vazione censuaria sul *sistema dei conti delle imprese* (SCI, 100 addetti ed oltre). In quegli strati dell'archivio nei quali non è stata possibile una stima diretta per mancanza di dati rilevati con le precedenti occasioni d'indagine si sono attribuite le stime dei parametri relative ad insiemi della popolazione che contenessero tali strati.

Come già spiegato, l'errore massimo imposto rappresenta il livello teorico di errore che le stime avrebbero se valessero due condizioni: che vi sia coincidenza perfetta tra lista di selezione e popolazione di indagine; che tutte le imprese selezionate restituiscano il modello debitamente compilato.

Tuttavia entrambe le condizioni non sono realizzate, sia per il disallineamento temporale fra l'archivio di selezione e la popolazione di interesse sia per il fenomeno della mancata risposta che può provocare la caduta di interi strati e determinare, per le stime sui domini costituiti da tali strati, un livello di precisione più basso di quello pianificato. L'incremento d'errore dovuto alla mancata risposta totale può essere attenuato dall'uso degli stimatori di ponderazione vincolata.

La procedura impiegata per la selezione del campione appartiene alla categoria di metodi, detti di selezione coordinata,²⁰ usualmente adottati per conseguire diversi obiettivi, tra cui:

il controllo della sovrapposizione di campioni relativi ad indagini longitudinali, ossia il controllo della rotazione delle unità campionarie;

la minimizzazione della sovrapposizione tra campioni provenienti dallo stesso archivio di estrazione e relativi ad indagini differenti.

La tecnica impiegata, nota come campionamento Jales,²¹ è basata sull'uso di Numeri Casuali Permanenti (Ncp): parte cioè dal presupposto che a ciascuna unità dell'archivio di selezione sia associato un numero casuale, generato da una variabile aleatoria uniforme definita nell'intervallo [0,1]; il Ncp resta assegnato all'impresa per l'intera sua esistenza e viene riportato invariato in tutte le successive versioni dell'archivio ASIA in cui l'impresa risulta essere presente.

A.2.2. *L'integrazione con dati di fonte bilanci civilistici*

In questo paragrafo viene descritto il processo di integrazione della rilevazione PMI con la fonte amministrativa dei Bilanci civilistici delle società di capitale, che è stato utilizzato fino alle stime ufficiali del 2007. Il processo di integrazione, che può essere esteso ad altre fonti amministrative (come è stato fatto per la rilevazione PMI2008 che ha utilizzato sia i bilanci civilistici che gli studi di settore), si articola in una fase di individuazione delle imprese che non hanno risposto alla rilevazione, attraverso l'incrocio fra la lista di spedizione ed il file di monitoraggio degli arrivi. Quest'ultimo registra, oltre alle imprese che hanno compilato il questionario, le imprese che non sono risultate reperibili all'indirizzo (indirizzo sbagliato, impresa trasferita), le imprese che hanno modificato lo stato di attività (imprese cessate, inattive, in liquidazione, soggette ad operazioni di fusione o scorporo, eccetera), le imprese fuori campo di osservazione e le imprese che hanno inviato una documentazione incompleta.

Individuati i casi di mancata risposta totale, si è cercato di ridurre gli effetti di distorsione sulle stime finali, in parte attraverso il processo di integrazione con dati amministrativi ed in parte mediante la metodologia degli stimatori di ponderazione vincolata.

La procedura individua per ogni impresa non rispondente un profilo economico ad essa applicabile, desumibile all'interno di un insieme di imprese omogenee in relazione all'attività economica esercitata (a livello di quattro cifre della classificazione Ateco), alla classe dimensionale in termini di addetti e alla localizzazione regionale. Nel caso in cui non si individuano le imprese donatrici in relazione alla stratificazione descritta, la procedura opera un progressivo collassamento degli strati fin quando non individua un gruppo di imprese donatrici: dei tre caratteri di stratificazione considerati si preferisce mantenere in forma prioritaria l'attività economica che spiega in misura maggiore l'omogeneità delle imprese dal punto di vista delle variabili economiche di interesse.

²⁰ Ballin, M., e S. Loriga. "Metodi di selezione coordinata nelle indagini campionarie sulle imprese". Rivista di statistica ufficiale, n. 2 (2000).

²¹ Ohlsson, E., e B.G. Cox, e D.A. Binder, e B.N. Chinnappa, e A. Christianson, e M.J. Kott, e P.S. Colledge. "Coordination of samples using permanent random numbers". In Business Survey Methods. New York: Wiley, 1995.

In pratica la procedura partendo dalla tipologia di stratificazione più ampia, classifica le imprese non rispondenti (mancate risposte totali) e quelle rispondenti (potenziali donatori) in una serie di strati via via più aggregati, in modo che la ricerca del donatore più vicino vada a buon fine. Le differenti tipologie di strato considerate sono le seguenti:

- strato_1 = (ate4 || regione || classe addetti)
- strato_2 = (ate4 || ripartizione || classe addetti)
- strato_3 = (ate4 || ripartizione)
- strato_4 = (ate4)
- strato_5 = (ate3 || regione || classe addetti)
- strato_6 = (ate3 || ripartizione || classe addetti)
- strato_7 = (ate3 || ripartizione)
- strato_8 = (ate3)
- strato_9 = (ate2 || regione || classe addetti)
- strato_10 = (ate2 || ripartizione || classe addetti)
- strato_11 = (ate2 || ripartizione)
- strato_12 = (ate2)

L'algoritmo calcola quindi la numerosità degli strati e determina per ognuno di essi il valore dell'indice $tasso_i$, definito come il rapporto tra il numero di unità donatrici sul numero di unità non rispondenti della medesima classe.

La variabile *strato*, che conterrà alla fine dell'iterazione il gruppo omogeneo dal quale ad ogni unità non rispondente verrà selezionata la sua corrispondente unità donatrice, viene inizializzata con la modalità "blank". Fissato un valore k , rappresentante il numero minimo di unità donatrici per unità non rispondente che deve essere presente all'interno di ciascun gruppo omogeneo, per ogni strato occorre verificare che

$$tasso_i \geq k$$

se il test è soddisfatto si pone $strato = strato_i$ e la ricerca del gruppo omogeneo per i non rispondenti inclusi in quella classe è da considerarsi conclusa.

In caso contrario, continua ad essere $strato = "blank"$ ma si passa all'aggregazione immediatamente successiva, ponendo $strato_i \leftarrow strato_{i+1}$ e verificando se in corrispondenza di questa aggregazione meno fine si osserva $tasso_{i+1} \geq k$.

La procedura ha termine quando non esiste alcuna classe con $strato = "blank"$, ossia quando a tutte le unità non rispondenti è stato assegnato un gruppo omogeneo di donatori.

Ciò può verificarsi prima che il programma arrivi all'ultima iterazione, nel caso in cui l'insieme dei donatori su cui formare le classi è sufficientemente numeroso. Viceversa, se in corrispondenza dell'ultima iterazione non è stato possibile individuare, per alcune unità, un gruppo omogeneo perché il numero di donatori risulta ancora insufficiente, la procedura viene forzata ad assegnare ad esse l'ultima classe (quella determinata in corrispondenza di $strato_{12}$) anche se il test non riesce ad essere soddisfatto. Si è preferito, infatti, fare in modo che tutti i non rispondenti fossero integrabili con un donatore, anche in considerazione del fatto che i casi di forzature sono comunque abbastanza rari.

È evidente l'importanza della scelta del criterio di omogeneità secondo il quale si definiscono gli strati e anche dell'ordine con cui si effettua il collassamento degli strati.

Quello adottato prevede il passaggio dalla regione amministrativa alla ripartizione territoriale, da più classi di addetti ad una sola, per poi cercare l'impresa donatrice considerando la sola attività economica a quattro cifre. Se anche in quest'ultimo caso non si riesce ad individuare il donatore, si itera la procedura secondo le stesse modalità ma partendo dalla classificazione dell'attività economica a tre cifre, e così via fino ad arrivare alla classificazione dell'attività economica a due cifre.

Al fine di evitare che fra i potenziali donatori possano esserci imprese con dati anomali, si effettua preliminarmente una ricerca di tali casi. Tradizionalmente, gli outliers vengono individuati con-

siderando la loro distanza relativa da un centro della distribuzione; se m ed s sono, rispettivamente, un indicatore di posizione e di scala dell'insieme ordinato di valori y_1, y_2, \dots, y_n , si definisce distanza relativa di un generico elemento y_i come la distanza tra y_i ed m diviso per s , ossia:

$$d_i = \frac{|y_i - m|}{s} \quad (13)$$

Se tale misura d_i eccede un predeterminato valore di soglia c , la corrispondente osservazione viene considerata un outlier. Alternativamente, si possono identificare come outliers tutte quelle osservazioni che cadono all'esterno di un dato intervallo di tolleranza di estremi $(m - c_l s, m + c_u s)$, con c_l e c_u valori prefissati ed eventualmente distinti. Media e varianza campionaria sono le statistiche più frequentemente utilizzate come indicatori di posizione e di scala, tuttavia tale scelta non si dimostra essere la più efficace in presenza di più outliers nella distribuzione, in quanto sia media che varianza sono a loro volta molto sensibili ai valori estremi.

Per evitare tale problema, si fa generalmente ricorso a statistiche robuste, assumendo come parametro di posizione m un quantile della distribuzione e come parametro di scala s la deviazione della stessa distribuzione da tale quantile. Nel nostro caso, si è scelto m pari alla mediana m_e ed s pari allo scostamento semplice medio dalla mediana S_{m_e} , definito da:

$$S_{m_e} = \frac{\sum_i |y_i - m_e|}{n} \quad (14)$$

Il problema a questo punto consiste nel fissare il valore della soglia c oltre il quale l'osservazione è da considerarsi un outlier; infatti, la scelta di una soglia c piccola può produrre l'esclusione di troppe unità, non necessariamente estreme, con il conseguente appiattimento della distribuzione; viceversa, un c ampio può non consentire l'identificazione di eventuali outliers.

C'è inoltre da considerare il fatto che valori particolarmente anomali per una o più osservazioni possono dar luogo ad un cosiddetto effetto maschera, ossia impedire l'individuazione di casi meno appariscenti anche se ugualmente influenti sulla variabilità della distribuzione.

Infine, se la ricerca dei valori estremi è condotta unicamente in base al criterio del percentile e l'insieme di dati è poco numeroso, può accadere che venga segnalata come outlier una osservazione esclusivamente per il fatto che risulta occupare l'ultimo posto, y_n , dell'insieme ordinato, indipendentemente dal grado di concentrazione della distribuzione: al limite, y_n può apparire come un outlier pur potendo risultare il suo valore coincidente con quello del dato mediano della distribuzione.

Per tali motivi si è preferito effettuare la ricerca implementando una procedura iterativa, che partendo dall'intera distribuzione dei dati identifica le potenziali unità anomale analizzando via via l'insieme delle osservazioni ridotto degli outliers individuati nei passi precedenti secondo un doppio criterio, basato sia sul percentile che sull'intervallo di tolleranza; la procedura termina quando, pur continuando ad eliminare i valori identificati come estremi al passo precedente, non si osservano guadagni apprezzabili nella riduzione della variabilità della distribuzione, ossia quando tale variabilità appare stabile.

Il criterio impiegato è piuttosto conservativo: infatti, scegliendo il 99° percentile anziché, ad esempio il 95°, si tende, ad ogni iterazione, ad individuare ed escludere come outliers il minor numero possibile di osservazioni. Si è preferito optare per il 99° percentile poiché il numero di osservazioni del data set disponibile su cui effettuare la ricerca a livello di ateco a più cifre è piuttosto esiguo: in queste condizioni è pertanto più opportuno non scartare tante osservazioni in blocco per non appiattare oltre il necessario la distribuzione.

Gli stessi motivi valgono per la scelta della soglia c : per la determinazione del livello di tale soglia e per la scelta del valore minimo dell'indice di riduzione, in base al quale si decide se terminare con le iterazioni, è opportuno effettuare un'analisi esplorativa dei dati lanciando una procedura per avere una misura della variabilità della distribuzione di partenza.

Effettuata anche questa fase, si procede all'estrazione casuale con ripetizione e probabilità uguale della impresa donatrice da ciascuno strato individuato. Tale operazione assegna a ciascuna impresa non rispondente una corrispondente impresa donatrice e genera un file che contiene le seguenti variabili:

- codice dell'impresa non rispondente (MRT);
- codice dell'impresa donatrice con dati validi;
- codice dello strato da cui è stata estratta l'impresa donatrice;
- numero di volte in cui è stata estratta la stessa impresa da quel determinato strato.

Il passo successivo consiste nel ricostruire i dati dell'impresa non rispondente a partire dai dati validati dell'impresa donatrice che vengono moltiplicati per un fattore rappresentato dal rapporto fra gli addetti dell'impresa da integrare aggiornati al tempo t e gli addetti dell'impresa donatrice relativi allo stesso tempo t .

In pratica si suppone che le performances dell'impresa sono correlate alla dimensione dell'impresa e quindi tutti i dati economici vengono modificati dal rapporto sopra citato (i dati dell'impresa di cui non si hanno avuto informazioni vengono ricostruiti a partire dai dati dichiarati all'impresa donatrice, modificati sulla base del rapporto fra addetti dell'impresa da integrare desunti dall'archivio aggiornato e gli addetti dichiarati dell'impresa donatrice). Per quanto riguarda la struttura occupazionale ed il costo del lavoro si procede al calcolo delle ore lavorate in media da ciascuna categoria professionale dell'impresa donatrice, della quota di occupazione femminile, delle retribuzioni medie per categoria professionale, del peso sul costo complessivo del lavoro di tutta una serie di voci (contributi sociali a carico dell'impresa, quote accantonate al trattamento di fine rapporto, provvidenze varie al personale). Si procede, poi, alla ricomposizione della struttura occupazionale per categoria professionale ed al calcolo, sulla base degli indicatori sopra citati, di tutte le variabili del questionario e, infine, si effettua la quadratura di tutte le variabili delle diverse sezioni del questionario.

La procedura si conclude con l'effettiva integrazione con i dati di bilancio delle società di capitale. In pratica si rimpiazzano i dati calcolati al passo precedente con quelli realmente dichiarati nel bilancio dalle società di capitale, effettuando una scomposizione dei totali nelle diverse voci parziali ed operando, in assenza di informazioni, secondo le regole del programma di correzione delle incompatibilità oppure sulla base di alcuni rapporti fra le variabili. Alla fine si effettua la quadratura di tutte le variabili delle diverse sezioni del questionario.

Il passo finale è finalizzato al controllo del processo di integrazione delle mancate risposte totali attraverso il calcolo di indicatori di posizione sul set completo di dati (imprese rispondenti e imprese integrate) calcolati sulla base dell'attività economica. Le principali variabili prese in esame sono il fatturato, il costo di acquisto di beni e servizi, il valore aggiunto, il costo del lavoro, gli investimenti e i seguenti indicatori:

- Fatturato per addetto;
- Valore aggiunto per addetto;
- Investimenti per addetto;
- Costo del lavoro per dipendente;
- Ore lavorate per dipendente;
- Costo del lavoro orario.

A.2.3. Procedura di stima

Il principio su cui è fondato qualsiasi metodo di stima campionaria è quello che il sottoinsieme delle unità della popolazione incluse nel campione deve rappresentare anche il sottoinsieme complementare costituito dalle rimanenti unità della popolazione stessa. Tale principio viene realizzato attribuendo a ciascuna unità inclusa nel campione un peso, che può essere visto come numero di elementi della popolazione rappresentati da detta unità.

In generale, per ottenere la stima di un totale si devono eseguire le tre seguenti operazioni:

1. determinare il peso da attribuire a ciascuna unità inclusa nel campione;
2. moltiplicare il valore relativo ad una data variabile oggetto di indagine, rilevata sulla generica unità inclusa nel campione, per il peso attribuito alla medesima unità;
3. effettuare la somma dei prodotti di cui al punto 2.

Nell'indagine PMI, basata su un disegno di campionamento complesso, ad uno stadio stratificato, il peso da attribuire a ciascuna unità è ottenuto in base ad una procedura articolata in più passi:

- in primo luogo, viene calcolato un peso iniziale, definito *peso diretto* (o *peso base*), determinato in funzione del disegno di campionamento come reciproco della probabilità di inclusione dell'unità campionata;
- successivamente, vengono calcolati dei fattori correttivi del peso base, che consentono sia di correggere (almeno parzialmente) la distorsione conseguente alla mancata risposta totale, sia di rispettare la condizione di uguaglianza tra alcuni parametri noti della popolazione e le corrispondenti stime campionarie;
- infine, viene determinato un peso, noto sotto il nome di *peso finale*, espresso come prodotto del peso base per i fattori correttivi.

La metodologia di stima applicata dall'Istat rientra nella classe nota sotto il nome di *stimatori di ponderazione vincolata* che sono applicabili quando esistono totali noti di variabili ausiliarie correlate alle variabili oggetto di indagine. Tali stimatori, sotto ipotesi piuttosto generali, sono caratterizzati dalle seguenti proprietà:

- sono più efficienti dello stimatore diretto, essendo l'efficienza dei primi tanto maggiore quanto più è alta la correlazione tra le variabili ausiliarie e le variabili oggetto di indagine;
- sono approssimativamente non distorti rispetto al disegno di campionamento;
- portano a stime dei totali noti che coincidono con i valori noti di tali totali;
- attenuano l'effetto distorsivo dovuto alla presenza di mancate risposte totali;
- attenuano l'effetto distorsivo dovuto alla sottocopertura della lista da cui è selezionato il campione.

Si ricorda, infine, che nelle indagini campionarie condotte su larga scala lo stimatore diretto è distorto e scarsamente efficiente, essendo fortemente influenzato dai fenomeni di mancata risposta e di sottocopertura della lista di selezione del campione.

Prima illustrare gli aspetti matematici del metodo di stima utilizzato, è utile introdurre la seguente simbologia:

- U = l'insieme delle imprese oggetto d'indagine, ossia le imprese operanti ed attive nel corso del 2007;
- α = un generico dominio di interesse;
- U_α = l'insieme delle imprese oggetto d'indagine, ossia l'impres operanti ed attive nel corso del 2007, appartenenti al dominio α ;
- s^* = l'insieme delle unità selezionate nel campione con $s^* \subseteq U$;
- s = l'insieme delle imprese selezionate nel campione e rispondenti all'indagine con $s \subseteq s^*$;
- s_α = il sottoinsieme di imprese del campione s appartenenti ad U_α , con $s_\alpha = s \cap U_\alpha$;
- n_α = il numero di imprese appartenenti al campione s_α .

Inoltre, con riferimento alla generica impresa, si assume che la si possa identificare univocamente con un indice k , definito dalla concatenazione dell'indice h di strato e i di unità richiamati nel paragrafo A.2.1.

Indichiamo pertanto con le seguenti quantità:

- y_k = il valore assunto dalla caratteristica y oggetto d'indagine sull'unità k ;

- $\lambda_k(\alpha)$ = variabile dicotomica che assume valore 1 se l'unità k appartiene al dominio α e assume valore 0 altrimenti.

Mediante la simbologia appena introdotta è possibile definire, con riferimento alla generica variabile rilevata y , i seguenti totali oggetto di stima:

- a) il totale riferito alle imprese appartenenti al dominio α :

$$Y_\alpha = \sum_{U_\alpha} y_k = \sum_U y_k \lambda_k(\alpha) \quad (15)$$

in cui il simbolo $\sum_U(\cdot)$ indica, con riferimento ad un generico insieme U , la sommatoria estesa a tutti gli elementi ad esso appartenenti;

- b) il totale riferito a tutte le imprese operanti sul territorio nazionale:

$$Y = \sum_U y_k \quad (16)$$

In generale, la stima \tilde{Y}_α del totale di dominio Y_α definito dalla (15), ottenuta con il metodo di ponderazione vincolata, ha la seguente espressione:

$$\tilde{Y}_\alpha = \sum_{s_\alpha} y_k w_k = \sum_s y_k \lambda_k(\alpha) w_k \quad (17)$$

Analogamente, la stima \tilde{Y} del totale Y viene ottenuta come semplice somma dei totali stimati per dominio:

$$\tilde{Y} = \sum_s y_k w_k \quad (18)$$

Il peso finale w_k è ottenuto, quindi, mediante il prodotto di tre fattori:

$$w_k = d_k \beta_k \gamma_k \quad (19)$$

in cui:

- $d_k = \frac{N_h}{n_h^*}$ ($k \in s_h^*$, $h=1, \dots, H$) indica il peso diretto, espresso come reciproco della probabilità di inclusione, dove N_h è il numero di unità dell'universo che appartengono allo strato h , n_h^* è il numero di unità del campione che appartengono allo strato h e s_h^* è l'insieme delle unità del campione nello strato h ;
- β_k indica il fattore di aggiustamento per mancata risposta;
- γ_k indica il fattore di aggiustamento per post-stratificazione.

La situazione in cui una unità statistica inclusa in una indagine non fornisce risposta ad alcun quesito previsto nel modello di rilevazione viene indicata con il termine di *mancata risposta totale*. Essa può essere determinata da cause differenti, come l'impossibilità di reperire l'unità inclusa nella popolazione oggetto di studio, l'incapacità o il rifiuto da parte di quest'ultima a rispondere, errori nella formulazione delle liste da cui si estraggono gli appartenenti al campione eccetera.

La presenza di mancate risposte totali è un problema comune sia alle indagini campionarie che a quelle censuarie; si possono applicare tecniche più o meno sofisticate per ridurne l'intensità, pur risultando difficile, se non impossibile, eliminarne totalmente la presenza.

Gli effetti più rilevanti che hanno le mancate risposte totali sui risultati finali sono due:

- in generale, riducono la quantità di informazione disponibile; nel caso di indagini campionarie, la riduzione produce un incremento dell'errore di campionamento e di conseguenza una minore precisione delle stime;
- apportano distorsioni nelle stime delle variabili oggetto di indagine quando il meccanismo di risposta è non casuale (può esservi, ad esempio, una maggiore propensione alla risposta in taluni sottogruppi della popolazione indagata, caratterizzati da particolari valori delle variabili d'indagine).

Le distorsioni che la presenza delle mancate risposte inducono sui risultati del processo di produzione dell'informazione statistica vengono prevenute od attenuate intervenendo in diverse fasi del processo. In particolare, il trattamento delle mancate risposte in fase di stima finale, come si è già accennato, prevede la riduzione dell'effetto distorsivo sulle stime attraverso l'uso di opportuni pesi correttivi da applicare al campione di rispondenti.

L'uso dei pesi per effettuare la correzione della mancata risposta totale poggia sull'assunzione di un meccanismo aleatorio che governa il fenomeno della mancata risposta; tale meccanismo fa sì che il campione s^* si ripartisca nel sottoinsieme dei rispondenti, s , e del suo complemento, formato dalle unità non rispondenti; il peso attribuito ad ogni elemento di s sta ad indicare il numero di unità non rispondenti rappresentato da ciascuna unità che ha risposto all'indagine.²²

Per trattare tale problema ci si riconduce alla teoria del campionamento in due fasi,²³ considerando la selezione dei non rispondenti come un ulteriore passo di campionamento governato da un processo di tipo bernoulliano, ignoto al ricercatore.

Formalmente, sia r_k una variabile dicotomica uguale ad 1 se l'unità campionaria k (con $k \in s^*$) è rispondente ed uguale a 0 altrimenti; nell'approccio considerato si assume che le r_k siano variabili di Bernoulli mutualmente indipendenti con valore atteso pari a θ_k che indica la probabilità dell'unità k di essere inclusa in s ($\theta_k = pr(k \in s)$) ed è denominato *probabilità di risposta* di k .

Nel caso in cui le probabilità di risposta θ_k fossero note per tutti i rispondenti, una stima corretta del totale Y sarebbe ottenuta tramite lo stimatore:

$$\tilde{Y}_c = \sum_s y_k d_k \theta_k^{-1} \quad (20)$$

Essendo però le quantità θ_k generalmente ignote, occorre darne una stima introducendo un modello di risposta, ossia formalizzando una relazione funzionale tra le θ_k ed un insieme di variabili esplicative disponibili sia per i rispondenti che per i non rispondenti.

Il modello più semplice di mancata risposta è quello che assume la probabilità θ_k costante per tutte

²² Ballin, M., e P.D. Falorsi, e B. Moreschi, e A. Pavone, e A.C. Seeber, e T. Tuoto. "Aspetti metodologici connessi alla produzione delle stime per il Censimento delle Istituzioni non profit". In Istituzioni nonprofit in Italia -I risultati della prima rilevazione censuaria - anno 1999. Roma: Istat, 2001. (Collana Informazioni).

²³ Särndal, C.E., e B. Swensson, e J. Wretman. Model Assisted Survey Sampling. New York: Springer-Verlag, 1992.

le unità della popolazione; tale ipotesi è poco realistica (a causa della effettiva eterogeneità tra le unità delle probabilità di risposta θ_k) e non è realmente utile all'eliminazione della componente distorsiva.

Viceversa, la definizione di probabilità variabili a livello di singola unità può dar luogo a stimatori poco stabili e alla generazione di valori anomali, a causa della variabilità associata alle stime delle probabilità θ_k per i singoli rispondenti.

La strategia generalmente adottata per superare tali problemi consiste in una soluzione intermedia in cui si assume un modello di mancata risposta che prevede una probabilità costante dei θ_k per sottoinsiemi di unità del campione s^* .

Indicando con $s_{(g)}^*$, $g=1, \dots, G$, il generico sottoinsieme di una partizione di s^* , di dimensione pari a $n_{(g)}^*$, (essendo $s^* \equiv \bigcup_{g=1}^G s_{(g)}^*$; $s_{(g)}^* \cap s_{(g')}^* = \emptyset$, per $g \neq g'$) il modello si esplicita nel modo seguente:

$pr(k \in s s^*) = \theta_g > 0$	per ogni $k \in s_{(g)}^*$	(21)
$pr((k, l) \in s s^*) = pr(k \in s s^*) pr(l \in s s^*)$	per ogni $k \neq l$; $(k, l) \in s_{(g)}^*$	

Tutte le unità appartenenti allo stesso gruppo, o *classe di omogeneità*, $s_{(g)}^*$, hanno uguale probabilità di risposta, mentre gruppi distinti hanno probabilità differenti e le risposte sono indipendenti.

Una volta determinata la partizione $\{s_{(g)}^*\}$, $g=1, \dots, G$, condizionando a s^* , al vettore $\mathbf{n}^* = (n_{(1)}^*, \dots, n_{(g)}^*, \dots, n_{(G)}^*)'$, contenente il numero di unità del campione appartenenti a ciascuna cella di ponderazione, ed al corrispondente vettore osservato $\mathbf{n} = (n_{(1)}, \dots, n_{(g)}, \dots, n_{(G)})'$ di unità rispondenti in ciascuna cella, una stima delle probabilità di risposta per le unità appartenenti alla cella di ponderazione g -esima è data da:

$$pr(k \in s | s^*, \mathbf{n}^*, \mathbf{n}) = \theta_k = \beta_k^{-1} = \frac{n_{(g)}}{n_{(g)}^*} \quad \text{per } k = l \in s_{(g)}; \quad (g = 1, \dots, G) \quad (22)$$

Pertanto, valendo il modello (21), uno stimatore corretto del totale Y è dato da:

$$\tilde{Y} = \sum_{g=1}^G \frac{n_{(g)}^*}{n_{(g)}} \sum_{s_{(g)}} y_k d_k = \sum_s y_k d_k \beta_k \quad (23)$$

Nella presente indagine la definizione delle classi di omogeneità è avvenuta considerando come esplicative del fenomeno della mancata risposta quelle variabili che descrivono il settore di attività dell'impresa, la sua dimensione in termini di addetti e la sua localizzazione. Per la determinazione dei correttori si è quindi operato come segue:

- le unità del campione sono state suddivise negli strati elementari, definiti dalla combinazione delle modalità delle variabili di archivio *regione*, *classe di attività economica (Ateco a 4 cifre)*, *classe dimensionali di addetti*;
- a ciascuna unità del campione è stata attribuita una probabilità di risposta individuale calcolata, pari al tasso risposta osservato nello strato elementare. La presenza di poche unità ri-

spondenti in alcuni strati ha causato la determinazione di stime poco affidabili, se non pari a 0 nel caso di caduta dell'intero strato;

- sono state quindi definite 10 *celle di ponderazione* aventi come estremi i *decili* della distribuzione delle probabilità di risposta stimate al punto precedente;
- alle unità *rispondenti* di ciascuna cella di ponderazione è stato attribuito un fattore correttivo β_k , pari al reciproco del tasso di risposta della cella di appartenenza.

Per il calcolo dei pesi finali si è adottato uno stimatore di ponderazione vincolata. Si tratta di uno stimatore appartenente alla più vasta categoria degli stimatori di regressione generalizzata, i quali permettono di migliorare l'efficienza delle stime quando si dispone, relativamente alla popolazione oggetto d'indagine, di totali noti di variabili ausiliarie correlate con le variabili d'interesse. In particolare, le variabili ausiliarie rappresentano le variabili esplicative di un modello di regressione, sottostante allo stimatore, con cui si esplicita una relazione funzionale tra le prime e le variabili d'interesse.

Per la costruzione dello stimatore è importante chiarire il concetto di *gruppo di riferimento del modello*: con questo termine si intende una partizione

$$\left\{ U(p) : U_L \equiv \bigcup_{p=1}^P U(p); U(p) \cap U(p') = \emptyset, p \neq p' \right\}$$

della popolazione U_L rispetto alla quale:

- sono noti i totali di popolazione di un insieme di variabili ausiliarie;
- viene definito il modello di regressione su cui si costruisce lo stimatore.

Definiamo inoltre:

- $\mathbf{x}_k = (x_{k1}, \dots, x_{kl}, \dots, x_{km})'$ il vettore colonna di m variabili ausiliarie osservato sull'unità k ;
- $\mathbf{X}_{U(p)} = \sum_{U(p)} \mathbf{x}_k$ i totali delle m variabili ausiliarie riferiti al sottoinsieme $U(p)$ della popolazione.

Sia $s(p)$ il sottoinsieme del campione dei rispondenti, di numerosità pari a $n(p)$, relativo al gruppo p -esimo della partizione.

Per ciascun sottoinsieme $s(p)$, gli $n(p)$ pesi finali $\{(w_k = d_k \beta_k \gamma_k); \text{ per } k \in s(p)\}$ sono ottenuti come soluzione di un sistema di minimo vincolato definito nel modo seguente:

$$\begin{cases} \text{Min} \left[\sum_{s(p)} c_k G(d_k \beta_k; d_k \beta_k \gamma_k) \right] \\ \sum_{s(p)} \mathbf{x}_k d_k \beta_k \gamma_k = \mathbf{X}_{U(p)} \end{cases} \quad p=1, \dots, P \quad (24)$$

dove:

- c_k indica una costante nota connessa alla dimensione dell'unità (nel caso in oggetto tale variabile è stata posta uguale al numero di addetti dell'impresa risultanti dall'archivio di selezione);

- $G(d_k \beta_k; d_k \beta_k \gamma_k)$ è la funzione di distanza logaritmica²⁴ tra i pesi finali $\{(w_k = d_k \beta_k \gamma_k); \text{per } k \in s(p)\}$ e i pesi diretti corretti per la mancata risposta $\{d_k \beta_k; \text{per } k \in s(p)\}$ definita da:

$$G(d_k \beta_k; d_k \beta_k \gamma_k) = d_k \beta_k \gamma_k \ln(\gamma_k) - d_k \beta_k \gamma_k + d_k \quad (25)$$

L'obiettivo è quello di individuare un vettore di pesi finali $\{(w_k = d_k \beta_k \gamma_k); \text{per } k \in s(p)\}$ che consenta di rispettare il sistema di vincoli (24) e che contemporaneamente modifichi il meno possibile l'insieme dei pesi diretti corretti per la mancata risposta.

La soluzione del sistema di minimo vincolato definito dalle equazioni precedenti viene trovata adottando il metodo dei moltiplicatori di Lagrange.

A.2.4. Principali variabili

Le principali variabili presenti nella rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI) sono state elaborate per arrivare a costruire aggregati utili per effettuare confronti con analoghe variabili che è possibile costruire dalle altre diverse fonti amministrative. A partire, quindi, dalle variabili di base della rilevazione PMI2007, costituita da 43.701 unità, sono state quindi costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
add_pmi	Addetti (in media annua) = imprenditori, titolari e cooperatori + coadiuvanti famigliari + dirigenti, quadri e impiegati + operai e commessi + apprendisti + lavoratori a domicilio
dip_pmi	Dipendenti (in media annua) = dirigenti, quadri e impiegati + operai e commessi + apprendisti + lavoratori a domicilio
fatt_tot_pmi	Ricavi delle vendite e delle prestazioni
- fatt_riv_pmi	vendite di merci acquistate e rivendute senza trasformazione
var_rfpcl_pmi	Variazione delle rimanenze di prodotti finiti, in corso di lavorazione e semilavorati
var_lavco_pmi	Variazione dei lavori in corso su ordinazione
inc_imml_i_pmi	Incrementi di immobilizzazioni per lavori interni
ric_altri_pmi	Altri ricavi e proventi
acq_beni_pmi	Acquisti di beni
- acq_riv_pmi	acquisti di merci da rivendere senza trasformazione
acq_ser_pmi	Acquisti di servizi + compensi agli amministratori
acq_bese_pmi	= acq_beni_pmi + acq_ser_pmi
acq_gdbt_pmi	Costi per godimento beni di terzi
clav_pmi	Costi del personale = retribuzioni lorde dei dirigenti, quadri e impiegati + retribuzioni lorde degli operai e commessi + retribuzioni lorde degli apprendisti + retribuzioni lorde dei lavoratori a domicilio + contributi sociali a carico dell'impresa + quote accantonate nell'esercizio per il trattamento di fine rapporto e di quiescenza + provvidenze varie al personale
- ret_pmi	Retribuzioni lorde = retribuzioni lorde dei dirigenti, quadri e impiegati + retribuzioni lorde degli operai e commessi + retribuzioni lorde degli apprendisti + retribuzioni lorde dei lavoratori a domicilio
- onerisoc_pmi	contributi sociali a carico dell'impresa
- tfr_pmi	quote accantonate nell'esercizio per il trattamento di fine rapporto e di quiescenza
- altrclav_pmi	provvidenze varie al personale
var_rmpriv_pmi	Variazione delle rimanenze di materie prime, sussidiarie e di consumo
- var_rmp_pmi	variazione delle rimanenze di materie prime, sussidiarie e di consumo
- var_rriv_pmi	variazione delle rimanenze di merci da rivendere in nome proprio senza trasformazione
acq_amm_pmi	Ammortamento di beni materiali e immateriali
acq_acc_pmi	Accantonamenti per rischi e altri accantonamenti
acq_oneri_pmi	formazione del personale + altri oneri diversi di gestione + imposte indirette sui prodotti e sulla produzione
var_riman_pmi	Variazione delle rimanenze = var_rfpcl_pmi - var_rmpriv_pmi
vagg_pmi	Valore aggiunto = (fatt_tot_pmi + var_rfpcl_pmi + var_lavco_pmi + inc_imml_i_pmi + ric_altri_pmi) - (acq_beni_pmi + acq_ser_pmi + acq_gdbt_pmi + var_rmp_pmi + acq_oneri_pmi)
margope_pmi	Margine operativo lordo = vagg_pmi - clav_pmi
valprod_pmi	Valore della produzione = (fatt_tot_pmi + var_rfpcl_pmi + var_lavco_pmi + inc_imml_i_pmi + ric_altri_pmi) - (acq_riv_pmi + var_rriv_pmi)

²⁴ Falorsi, P.D., e S. Falorsi. "Un metodo di stima generalizzato per le indagini sulle famiglie e sulle imprese". Rapporto di ricerca, n. 13. Università di Bologna, 1995. (Quaderni CON PRI).

A.3. Bilanci civilistici delle società di capitale

Dal 2000 l'Istat acquisisce in formato elettronico i dati dei *bilanci civilistici* depositati dalle società di capitale presso le Camere di Commercio, ed in particolare tutte le poste relative al conto economico e allo stato patrimoniale mentre dalla nota integrativa sono estratte solo alcune informazioni (ad esempio i dipendenti dell'impresa).

Le informazioni disponibili dalla fonte Bilanci civilistici utili per la rilevazione PMI2007 sono rappresentate dalle principali voci che compongono il conto economico, che sono solo una parte di quelle richieste nel questionario di rilevazione, con una maggiore disaggregazione del costo del lavoro. Attraverso un link delle informazioni disponibili dalla fonte Bilanci civilistici con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte Bilanci civilistici rientrano nel campione teorico della rilevazione. Le principali variabili messe a disposizione della rilevazione attraverso questo processo sono state elaborate per arrivare a costruire aggregati utili per effettuare confronti di coerenza con le analoghe variabili che è possibile costruire dalle altre diverse fonti e in particolare con quelle della rilevazione PMI2007.

A partire, quindi, dalle variabili di base della fonte Bilanci civilistici, riferita all'anno 2007, costituita da 30.109 unità che si accoppiano con le unità del campione teorico della rilevazione PMI, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_bil	Ricavi delle vendite e delle prestazioni
var_rpfpcl_bil	Variazione delle rimanenze di prodotti finiti e semilavorati
var_lavco_bil	Variazione lavori in corso
inc_imkli_bil	Incrementi delle immobilizzazioni per lavori interni
ric_altri_bil	Altri ricavi e proventi
acq_beni_bil	Costi acquisto materie prime e merci
acq_ser_bil	Costi acquisto servizi
acq_bese_bil	acq_beni_bil + acq_ser_bil
acq_gdbt_bil	Costi per godimento beni di terzi
clav_bil	Costo del personale
- ret_bil	retribuzione lorda per salari e stipendi
- onerisoc_bil	contributi sociali a carico dell'impresa
- tfr_bil	quote accantonate nell'esercizio per trattamento di fine rapporto + quote accantonate nell'esercizio per trattamento di quiescenza
- altrclav_bil	provvidenze varie al personale
var_rmpriv_bil	Variazione delle rimanenze di materie prime e merci
acq_amm_bil	Quote di ammortamento di beni e di svalutazione crediti – Svalutazione crediti
acq_acc_bil	Quote di accantonamenti per rischi + Altre quote di accantonamenti per oneri e spese future
acq_oneri_bil	Oneri diversi di gestione
var_riman_bil	Variazione delle rimanenze = var_rpfpcl_bil + var_rmpriv_bil
vagg_bil	Valore aggiunto = (fatt_tot_bil + var_rpfpcl_bil + var_lavco_bil + inc_imkli_bil + ric_altri_bil) – (acq_beni_bil + acq_ser_bil + acq_gdbt_bil + var_rmpriv_bil + acq_oneri_bil)
margope_bil	Margine operativo lordo = vagg_bil – clav_bil

A.4. Studi di settore

Gli *Studi di Settore* sono uno strumento che l'Agenzia delle Entrate utilizza per rilevare i parametri fondamentali di liberi professionisti, lavoratori autonomi e imprese, sia sotto l'aspetto delle caratteristiche strutturali sia sotto l'aspetto economico, al fine di valutare la loro capacità di produrre reddito. La disciplina degli Studi di Settore è stata introdotta nel 1993 (D.L. 30 agosto 1993, n. 331, convertito dalla legge 29 ottobre 1993 n. 427) e si è progressivamente estesa a moltissimi settori di attività economica in prevalenza nel campo dei servizi alle imprese e alle famiglie. Per consentire ai contribuenti e ai consulenti fiscali di gestire la varietà e la complessità dei dati richiesti per ciascun tipo di studio di settore, ogni anno viene messa a disposizione una procedura software, chiamata Gerico. L'obiettivo della procedura è quello di stimare il ricavo presunto del singolo con-

tributente in funzione di variabili strutturali (che sono un mix di variabili di scala e qualitative, di input e output, di capitale e di processo) da porre a confronto con quello dichiarato. Attraverso una procedura di clusterizzazione l'Agenzia delle entrate individua i gruppi di contribuenti omogenei rispetto alle variabili strutturali rilevate e mediante l'analisi discriminante associa ciascun contribuente a uno o più cluster con determinate probabilità di appartenenza.

Il contribuente in relazione al ricavo presunto ha la facoltà di adeguare i propri ricavi a quanto dovrebbe essere congruo in relazione al modello organizzativo che si è dato. Nel caso in cui non ritiene congruo il ricavo stimato dall'Agenzia delle entrate può confermare le proprie risultanze economiche, ma in questo caso andrà incontro a un accertamento induttivo da parte dell'Agenzia delle entrate.

Gli Studi di Settore rappresentano una delle fonti amministrative di maggiore interesse per le analisi statistiche, soprattutto per la fascia delle piccole e medie imprese che svolgono attività economiche nel campo industriale, commerciale e dei servizi e dei professionisti: sono infatti obbligati a compilare tale modulistica tutti i soggetti che realizzano un fatturato non superiore alla soglia di 7.500.000 euro.

I questionari sugli studi di settori sono specifici per ciascuna attività economica esercitata (codificata, a partire dal 2008, secondo la classificazione Istat Ateco2007) e sono caratterizzati da una struttura comune che raccoglie informazioni sull'occupazione, unità locali, beni strumentali, elementi contabili e altre informazioni.

Attraverso un aggancio con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte Studi di Settore (quadro F) rientrano nel campione teorico della rilevazione. A partire dalle informazioni di base del quadro F, per le 61.596 unità che si sono accoppiate con quelle del campione teorico della rilevazione PMI2007 sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_sdsf	Ricavi di cui alle lettere a) e b) dell'art. 85, comma 1, del TUIR + Ricavi derivanti dalla vendita di generi soggetti ad aggio o ricavo fisso
fatt_aggi_sdsf	Ricavi derivanti dalla vendita di generi soggetti ad aggio o ricavo fisso
riman_fin_sdsf	Rimanenze finali merci, prodotti finiti, materie prime, sussidiarie, semilavorati, servizi non di durata ultrannuale + Rimanenze finali relative a prodotti soggetti ad aggio o ricavo fisso
riman_ini_sdsf	Esistenze iniziali merci, prodotti finiti, materie prime, sussidiarie, semilavorati, servizi non di durata ultrannuale + Esistenze iniziali relative a prodotti soggetti ad aggio o ricavo fisso
var_riman_sdsf	Variazione delle rimanenze = riman_fin_sdsf - riman_ini_sdsf
var_lavco_sdsf	Rimanenze finali relative a opere, forniture e servizi di durata ultrannuale - Esistenze iniziali relative a opere, forniture e servizi di durata ultrannuale
inc_imkli_sdsf	Incrementi di immobilizzazioni per lavori interni
ric_altri_sdsf	Altri proventi considerati + Altri proventi e componenti positivi
acq_beni_sdsf	Costi per l'acquisto di materie prime, sussidiarie, semilavorati e merci + Costi per l'acquisto di prodotti soggetti ad aggio o ricavo fisso
acq_aggi_sdsf	Costi per l'acquisto di prodotti soggetti ad aggio o ricavo fisso
acq_ser_sdsf	Costo per la produzione di servizi + Spese per acquisti di servizi + Altri costi per servizi + spese per prestazioni rese da professionisti esterni + spese per personale di terzi distaccato presso l'impresa, interinale, somministrazione lavoro + spese per collaboratori coordinati e continuativi
acq_bese_sdsf	= acq_beni_sdsf + acq_ser_sdsf
acq_gdbt_sdsf	Costo godimento beni di terzi (canoni di leasing, per beni immobili, royalties)
clav_sdsf	Spese per lavoro dipendente e per altre prestazioni diverse da lavoro dipendente - spese per prestazioni rese da professionisti esterni - spese per personale di terzi distaccato presso l'impresa, interinale, somministrazione lavoro - spese per collaboratori coordinati e continuativi
var_aggi_sdsf	Rimanenze finali relative a prodotti soggetti ad aggio o ricavo fisso - Esistenze iniziali relative a prodotti soggetti ad aggio o ricavo fisso (fine - inizio)
acq_amm_sdsf	Ammortamenti
acq_acc_sdsf	Accantonamenti
acq_oneri_sdsf	Oneri diversi di gestione + Altri componenti negativi - Utili spettanti agli associati in partecipazione con apporti di solo lavoro
vagg_sdsf	Valore aggiunto = (fatt_tot_sdsf + var_riman_sdsf + var_lavco_sdsf + inc_imkli_sdsf + ric_altri_sdsf) - (acq_beni_sdsf + acq_ser_sdsf + acq_gdbt_sdsf + acq_oneri_sdsf)
margope_sdsf	Margine operativo lordo = vagg_sdsf - clav_sdsf

Analogamente, anche per la fonte Studi di Settore (quadro G sui professionisti), per le 2.672 unità che si accoppiano con quelle del campione teorico PMI2007, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_sdsg	Compensi dichiarati
ric_altri_sdsg	Altri proventi lordi
acq_beni_sdsg	Consumi
acq_ser_sdsg	Spese per personale con contratto di lavoro interinale o di somministrazione lavoro + Spese per prestazioni di collaborazione coordinata e continuativa + Compensi corrisposti a terzi per prestazioni afferenti l'attività professionale e artistica + Altre spese
acq_bese_sdsg	= acq_beni_sdsg + acq_ser_sdsg
acq_gdbt_sdsg	non disponibile
clav_sdsg	Spese per prestazioni di lavoro dipendente - spese per personale con contratto di lavoro interinale o di somministrazione lavoro
acq_amm_sdsg	Ammortamenti
acq_acc_sdsg	non disponibile
acq_oneri_sdsg	Altri componenti negativi
var_riman_sdsg	non disponibile
vagg_sdsg	Valore aggiunto = (fatt_tot_sdsg + ric_altri_sdsg) - (acq_beni_sdsg + acq_ser_sdsg + acq_oneri_sdsg)
margope_sdsg	Margine operativo lordo = vagg_sdsg - clav_sdsg

A.5. Modello Unico e modello Irap

Il *modello Unico* è un modello unificato delle dichiarazioni fiscali tramite il quale le persone fisiche e giuridiche possono presentare la dichiarazione dei redditi, dell'IVA, dell'Irap e la dichiarazione dei sostituti di imposta ed intermediari (modello 770); i contribuenti nei confronti dei quali si applicano gli *Studi di Settore* o i parametri, sono tenuti altresì a presentare il relativo modello con i dati rilevanti ai fini dell'applicazione degli stessi.

A partire dall'anno di imposta 2007, tutti i contribuenti sono obbligati alla presentazione del modello Unico per via telematica direttamente o tramite gli intermediari abilitati.

Il modello Unico è distinto in relazione ai soggetti dichiaranti, ognuno dei quali deve compilare degli appositi quadri:

- **modello Unico Persone Fisiche:** i soggetti obbligati alla presentazione di tale modello sono i soggetti obbligati alla tenuta delle scritture contabili, i lavoratori dipendenti ed in genere chi ha conseguito redditi nel periodo d'imposta. Con riferimento ai redditi di impresa e di lavoro autonomo i relativi quadri da compilare, di interesse per l'attività di studio, sono i seguenti:
 - Quadro RE - Redditi di lavoro autonomo;
 - Quadro RF - Redditi di impresa in contabilità ordinaria;
 - Quadro RG - Redditi di impresa in regime di contabilità semplificata e regimi forfetari.
- **modello Unico Società di persone ed equiparate:** i soggetti obbligati alla presentazione di tale modello sono le società e le associazioni residenti nel territorio dello Stato di seguito elencate:
 - società semplici;
 - società in nome collettivo e in accomandita semplice;
 - società di armamento (equiparate alle società in nome collettivo o alle società in accomandita semplice, a seconda che siano state costituite all'unanimità o a maggioranza);
 - società di fatto o irregolari (equiparate alle società in nome collettivo o alle società semplice a seconda che esercitano o meno attività commerciale);
 - associazioni senza personalità giuridica costituite fra persone fisiche per l'esercizio in forma associata di arti e professioni;
 - aziende coniugali se l'attività è esercitata in società fra coniugi (coniugi contestatari della licenza ovvero coniugi entrambi imprenditori);
 - gruppi europei di interesse economico GEIE

Si considerano residenti le società e le associazioni che per la maggior parte del periodo di imposta hanno la sede legale o la sede amministrativa o l'oggetto principale dell'attività (determinato

in base all'atto costitutivo, se esiste in forma di atto pubblico o di scrittura privata autenticata, e in mancanza, in base all'attività effettivamente esercitata) nel territorio dello Stato.

Tali soggetti sono tenuti alla compilazione dei seguenti quadri di interesse per l'attività di studio:

- *Quadro RE* - Redditi di lavoro autonomo;
- *Quadro RF* - Redditi di impresa in contabilità ordinaria;
- *Quadro RG* - Redditi di impresa in regime di contabilità semplificata.

- **modello Unico Società di capitale, enti commerciali ed equiparati:** i soggetti obbligati alla presentazione di tale modello sono i seguenti:

- società per azioni e in accomandita per azioni, società a responsabilità limitata, società cooperative (comprese le società cooperative che abbiano acquisito la qualifica ON-LUS e le cooperative sociali), società di mutua assicurazione, le società europee di cui al regolamento CE n.2157/2001 e le società cooperative europee di cui al regolamento CE n.1453/2003 residenti nel territorio dello Stato;
- enti commerciali (enti pubblici e privati diversi dalle società, trust che hanno per oggetto esclusivo o principale l'esercizio di attività commerciali) residenti nel territorio dello Stato;
- società di ogni tipo (tranne società semplici, società e associazioni ad esse equiparate ai sensi dell'art. 5 del TUIR) nonché enti commerciali non residenti nel territorio dello Stato, compresi i trust, che hanno esercitato l'attività nel territorio dello Stato mediante stabile organizzazione.

Tali soggetti sono tenuti alla compilazione del seguente quadro di interesse per l'attività di studio: Quadro RS - Prospetti vari.

Tutti i soggetti sopra descritti sono altresì tenuti alla compilazione del **modello Irap** di dichiarazione dell'imposta regionale sulle attività produttive, che è così articolato:

- modello Irap Persone Fisiche: Quadro IQ;
- modello Irap Società di persone: Quadro IQ;
- modello Irap Società di capitale: Quadro IQ.

A.5.1. Modello Unico Persone fisiche – Quadro RE: Redditi di lavoro autonomo

Il *quadro RE del modello Unico sulle persone fisiche* deve essere utilizzato per dichiarare i redditi derivanti **dall'esercizio di arti e professioni** indicati nel comma 1 dell'art.53 del TUIR, rientranti nel regime analitico, i redditi rientranti nel regime agevolato di cui agli artt. 13 e 14 della legge 23 dicembre 2000 n.388, nonché i proventi percepiti per prestazione di volontariato o cooperazione rese ad organizzazioni non governative riconosciute idonee ai sensi dell'art. 28 della legge 26 febbraio 1987, n.49, qualora dette prestazioni discendano dall'assunzione di obblighi riconducibili ad un rapporto di lavoro autonomo.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello Unico persone fisiche (quadro RE) rientrano nel campione teorico della rilevazione.

Per l'anno di riferimento 2007, la fonte modello Unico persone fisiche (Quadro RE) copre 3.621 unità del campione teorico della rilevazione PMI. A partire, quindi, dalle variabili presenti nel quadro RE sono state costruite le seguenti principali variabili.

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_upfre	Compensi per attività professionale e artistica
ric_altri_upfre	Altri proventi lordi
acq_beni_upfre	non disponibile
acq_ser_upfre	Spese relative agli immobili + Compensi corrisposti a terzi + Consumi (telefono, energia, carburante)+ Spese per prestazioni alberghiere: addebitate ai clienti + Spese per prestazioni alberghiere: altre spese + Spese di rappresentanza: sostenute + (50% spese per convegni, congressi e aggiornamento) *2
acq_bese_upfre	= acq_beni_upfre + acq_ser_upfre
acq_gdbt_upfre	Canoni di locazione finanziaria di beni mobili + Canoni di locazione non finanziaria e/o di noleggio
clav_upfre	Spese personale dipendente e assimilato
acq_amm_upfre	Quote ammortamento e acquisti inferiori a 1 milione
acq_acc_upfre	non disponibile
acq_oneri_upfre	Altre spese documentate
var_riman_upfre	non disponibile
vagg_upfre	Valore aggiunto = (fatt_tot_upfre + ric_altri_upfre) - (acq_beni_upfre + acq_ser_upfre + acq_gdbt_upfre + acq_oneri_upfre)
margope_upfre	Margine operativo lordo= vagg_upfre - clav_upfre

A.5.2. Modello Unico Persone fisiche – Quadro RF: Redditi di impresa in contabilità ordinaria

Il *quadro RF del modello Unico sulle persone fisiche* deve essere compilato dagli **esercenti attività commerciali in regime di contabilità ordinaria** e da quelli che, pur potendosi avvalere della contabilità semplificata e determinare il reddito ai sensi dell'art. 66 del TUIR, hanno optato per il regime ordinario. Per come esso è strutturato non può essere utilizzato come fonte informativa, in quanto riporta le variazioni in aumento ed in diminuzione del reddito.

A.5.3. Modello Unico Persone fisiche – Quadro RG: Redditi di impresa in regime di contabilità semplificata e regimi forfetari

Il *quadro RG del modello Unico sulle persone fisiche* deve essere compilato dagli **esercenti attività commerciali in regime di contabilità semplificata** di cui all'art. 18 del D.P.R. n.500 del 1973. Tali soggetti determinano il reddito ai sensi dell'art.66 del TUIR se nel periodo d'imposta precedente hanno conseguito ricavi per un ammontare non superiore a 309.874,14 euro per le imprese aventi per oggetto prestazioni di servizi e non superiore a 516.456,90 euro per le imprese aventi per oggetto altre attività, sempre che non abbiano optato per il regime di contabilità ordinaria.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello Unico persone fisiche (quadro RG) rientrano nel campione teorico della rilevazione.

Il modello Unico persone fisiche (Quadro RG) copre 18.700 unità del campione teorico della rilevazione PMI2007: dalle informazioni di base contenute nel quadro RG sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_upfrg	Ricavi
var_riman_upfrg	Variatione delle rimanenze = Rimanenze finali prodotti finiti, materie prime, merci – Rimanenze iniziali merci
var_lavco_upfrg	Rimanenze finali opere, forniture e servizi – Rimanenze iniziali opere
ric_altri_upfrg	Altri proventi
acq_beni_upfrg	Costi per acquisto materie prime, sussidiarie, semilavorati e merci
acq_ser_upfrg	non disponibile
acq_bese_upfrg	= acq_beni_upfrg + acq_ser_upfrg
acq_gdbt_upfrg	Canoni locazione finanziaria di beni mobili strumentali
clav_upfrg	Spese per lavoro dipendente, assimilato e per lavoro autonomo
acq_amm_upfrg	Quote ammortamento + Spese acquisto beni strumentali non superiori a 516,46 euro
acq_acc_upfrg	non disponibile
acq_oneri_upfrg	Spese ed altri componenti negativi per operazioni con soggetti fiscalmente privilegiati + Altri componenti negativi
vagg_upfrg	Valore aggiunto = (fatt_tot_upfrg + var_riman_upfrg + var_lavco_upfrg + ric_altri_upfrg) – (acq_beni_upfrg + acq_ser_upfrg + acq_gdbt_upfrg + acq_oneri_upfrg)
margope2_upfrg	Margine operativo lordo = vagg_upfrg - clav_upfrg

A.5.4. Modello Irap Persone fisiche – Quadro IQ: Irap

Il modello Irap sulle Persone fisiche (quadro IQ) deve essere compilato dalle persone fisiche per la dichiarazione dell'imposta regionale sulle attività produttive. L'imposta è determinata applicando al valore della produzione netta l'aliquota del 4,25%, salvo quanto previsto dal comma 1 dell'art.45 del D.lgs. n.466, nonché da leggi regionali o delle Province autonome. Il quadro IQ deve essere compilato dagli esercenti attività commerciale (titolari di redditi d'impresa ai sensi dell'art.55 del TUIR), esercenti arti e professioni (titolari di redditi di lavoro autonomo ai sensi dell'art.53, comma 1, del TUIR) e dai produttori agricoli (titolari di reddito agrario ai sensi dell'art.32 del TUIR).

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello Irap persone fisiche (quadro IQ) rientrano nel campione teorico della rilevazione. A partire dalle variabili di base del quadro IQ, costituito da 24.146 unità che si accoppiano con quelle del campione teorico PMI2007, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_upfiq	Ricavi delle vendite e delle prestazioni
var_rpfpc_l_upfiq	Variatione delle rimanenze prodotti finiti, semilavorati, in corso di lavorazione
var_lavco_upfiq	Variatione di lavori in corso su ordinazione
inc_immli_upfiq	Incrementi di immobilizzazioni per lavori interni
ric_altri_upfiq	Altri ricavi e proventi
acq_beni_upfiq	Costi per materie prime, sussidiarie e di consumo
acq_ser_upfiq	Costi per servizi
acq_bese_upfiq	= acq_beni_upfiq + acq_ser_upfiq
acq_gdbt_upfiq	Costi per godimento beni di terzi
clav_upfiq	Retribuzioni complessive per dipendenti e assimilati operanti in Italia
clav2_upfiq	Retribuzioni complessive per dipendenti e assimilati operanti all'estero + Retribuzioni complessive per dipendenti e assimilati operanti in Italia
acq_amm_upfiq	Ammortamento delle immobilizzazioni materiali + Ammortamento delle immobilizzazioni immateriali
acq_acc_upfiq	Accantonamenti
var_rmpriv_upfiq	Variatione delle rimanenze di materie prime e merci
acq_oneri_upfiq	Oneri diversi di gestione
var_riman_upfiq	Variatione delle rimanenze = var_rpfpc_l_upfiq - var_rmpriv_upfiq
vagg_upfiq	Valore aggiunto = (fatt_tot_upfiq + var_rpfpc_l_upfiq + var_lavco_upfiq + inc_immli_upfiq + ric_altri_upfiq) - (acq_beni_upfiq + acq_ser_upfiq + acq_gdbt_upfiq + var_rmpriv_upfiq + acq_oneri_upfiq)
margope_upfiq	Margine operativo lordo = vagg_upfiq - clav_upfiq
margope2_upfiq	Margine operativo lordo = vagg_upfiq - clav2_upfiq

A.5.5. Modello Unico Società di persone – Quadro RE: Redditi di lavoro autonomo derivanti dall'esercizio di arti e professioni

Il quadro RE del modello Unico sulle società di persone deve essere utilizzato dalle società semplici e dalle associazioni tra artisti e professionisti per dichiarare i redditi di lavoro autonomo derivanti dall'esercizio di arti e professioni.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa sono state individuante 40.916 unità della fonte modello Unico società di persone (quadro RE) che rientrano nel campione teorico della rilevazione: dalle informazioni di base del quadro RE sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_uspre	Compensi per attività professionale e artistica
ric_altri_uspre	Altri proventi lordi
var_lavco_uspre	non disponibile
acq_beni_uspre	non disponibile
acq_ser_uspre	Canoni di locazione finanziaria di beni mobili + Canoni di locazione non finanziaria e/o di noleggio + Spese relative agli immobili + Compensi corrisposti a terzi + Consumi (telefono, energia, carburante)+ Spese per prestazioni alberghiere: addebitate ai clienti + Spese per prestazioni alberghiere: altre spese + Spese di rappresentanza: ammontare sostenuto + (50% spese per convegni, congressi e aggiornamento)*2
acq_bese_uspre	acq_beni_uspre + acq_ser_uspre
acq_gdbt_uspre	non disponibile
clav_uspre	Spese personale dipendente e assimilato
acq_amr_uspre	Quote ammortamento e acquisti di costo unitario non superiore a 516,46 euro
acq_acc_uspre	non disponibile
acq_oneri_uspre	Altre spese documentate
var_riman_uspre	non disponibile
vagg_uspre	Valore aggiunto = (fatt_tot_uspre + var_riman_uspre + var_lavco_uspre + ric_altri_uspre) - (acq_beni_uspre + acq_ser_uspre + acq_oneri_uspre)
margope_uspre	Margine operativo lordo = agg_uspre - clav_uspre

A.5.6. Modello Unico Società di persone – Quadro RF: Redditi di impresa in contabilità ordinaria

Il quadro RF del modello Unico sulle società di persone deve essere utilizzato dalle società in nome collettivo e in accomandita semplice e dai soggetti equiparati obbligati alla tenuta della contabilità ordinaria e da quelli che, pur potendosi avvalere della contabilità semplificata e determinare il reddito ai sensi dell'art.66 del TUIR, hanno optato per il regime ordinario. Il quadro deve essere utilizzato anche dai Gruppi Europei di Interesse Economico (GEIE), indipendentemente dall'attività svolta.

Per come esso è strutturato non può essere utilizzato come fonte informativa, in quanto riporta le variazioni in aumento ed in diminuzione del reddito.

A.5.7. Modello Unico Società di persone – Quadro RG: Redditi di impresa in regime di contabilità semplificata

Il quadro RG del modello Unico sulle società di persone deve essere utilizzato dalle società di persone esercenti attività commerciali in regime di contabilità semplificata di cui all'art.18 del D.P.R. n.600 del 1973. Sono ammesse alla tenuta della contabilità semplificata quindi, determinano il reddito ai sensi dell'art.66 del TUIR, le società che non hanno optato per il regime di contabilità ordinaria e che, nel periodo d'imposta precedente, hanno conseguito ricavi per un ammontare non superiore a 309.874,14 euro per le imprese aventi per oggetto prestazioni di servizi e non superiore a 516.456,90 euro per le imprese aventi per oggetto altre attività.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello Unico società di persone (quadro RG) rientrano nel campione teorico della rilevazione.

A partire dalle informazioni di base del quadro RG, per le 6.243 unità che si sono accoppiate con il campione teorico della rilevazione PMI2007, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_usprg	Ricavi
ric_altri_usprg	Altri proventi considerati ricavi + Altri componenti positivi
var_lavco_usprg	Rimanenze finali opere, forniture e servizi - Rimanenze iniziali opere
acq_beni_usprg	Costi per acquisto materie prime, sussidiarie, semilavorati e merci
acq_ser_usprg	non disponibile
acq_bese_usprg	acq_beni_usprg + acq_ser_usprg
acq_gdbt_usprg	Canoni locazione finanziaria di beni mobili strumentali
clav2_usprg	Spese per lavoro dipendente, assimilato e per lavoro autonomo
acq_amm_usprg	= Quote ammortamento + Spese per l'acquisto di beni strumentali di costo unitario non superiori a 516,46 euro
acq_acc_usprg	non disponibile
acq_oneri_usprg	Spese ed altri componenti negativi per operazioni con soggetti fiscalmente privilegiati + Altri componenti negativi
var_riman_usprg	Rimanenze finali prodotti finiti, materie prime, merci - Rimanenze iniziali merci
vagg_usprg	Valore aggiunto = (fatt_tot_usprg + var_riman_usprg + var_lavco_usprg + ric_altri_usprg) - (acq_beni_usprg + acq_ser_usprg + acq_gdbt_usprg + acq_oneri_usprg)
margope2_usprg	Margine operativo lordo = vagg_usprg - clav2_usprg

A.5.8. Modello Irap Società di persone – Quadro IQ: Irap

Il modello Irap delle società di persone (quadro IQ) deve essere compilato dalle società di persone ed equiparate per la dichiarazione dell'imposta regionale sulle attività produttive. L'imposta è determinata applicando al valore della produzione netta l'aliquota del 4,25%, salvo quanto previsto dal comma 1 dell'art.45 del D.lgs. n.466, nonché da leggi regionali o delle Province autonome. Il quadro IQ deve essere compilato dalle società semplici, in nome collettivo e in accomandita semplice e quelle ad esse equiparate a norma dell'art.5 del TUIR, comprese le associazioni costituite tra persone fisiche per l'esercizio in forma associata di arti e professioni.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello Irap società di persone (quadro IQ) rientrano nel campione teorico della rilevazione.

Le principali variabili presenti nel modello Irap persone fisiche (Quadro IQ) utili per la sperimentazione di integrazione con dati di fonte amministrativa, sono state elaborate per arrivare a costruire aggregati utili per effettuare confronti con analoghe variabili che è possibile costruire dalle altre diverse fonti amministrative. A partire, quindi, dalle variabili di base del modello Irap società di persone (quadro IQ), costituito da 20.552 unità che si accoppiano con il campione teorico PMI2007, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_ustpip	Ricavi delle vendite e delle prestazioni
var_rpfpcpl_ustpip	Variazione delle rimanenze prodotti finiti, semilavorati, in corso di lavorazione
var_lavco_ustpip	Variazione di lavori in corso su ordinazione
inc_imml_i_ustpip	Incrementi di immobilizzazioni per lavori interni
ric_altri_ustpip	Altri ricavi e proventi
acq_beni_ustpip	Costi per materie prime, sussidiarie e di consumo
acq_ser_ustpip	Costi per servizi
acq_bese_ustpip	acq_beni_ustpip + acq_ser_ustpip
acq_gdbt_ustpip	Costi per godimento beni di terzi
clav_ustpip	Retribuzioni complessive per dipendenti e assimilati operanti in Italia
clav2_ustpip	Retribuzioni complessive per dipendenti e assimilati operanti all'estero + Retribuzioni complessive per dipendenti e assimilati operanti in Italia
acq_amm_ustpip	Ammortamento delle immobilizzazioni materiali + Ammortamento delle immobilizzazioni immateriali
acq_acc_ustpip	Accantonamenti
var_rmpriv_ustpip	Variazione delle rimanenze di materie prime e merci
acq_oneri_ustpip	Oneri diversi di gestione
var_riman_ustpip	Variazione delle rimanenze = var_rpfpcpl_ustpip - var_rmpriv_ustpip
vagg_ustpip	Valore aggiunto = (fatt_tot_ustpip + var_rpfpcpl_ustpip + var_lavco_ustpip + inc_imml_i_ustpip + ric_altri_ustpip) - (acq_beni_ustpip + acq_ser_ustpip + acq_gdbt_ustpip + var_rmpriv_ustpip + acq_oneri_ustpip)
margope_ustpip	Margine operativo = vagg_ustpip - clav_ustpip
margope2_ustpip	Margine operativo2= vagg_ustpip - clav2_ustpip

A.5.9. Modello Unico Società di capitale – Quadro RS: Prospetti vari

Il quadro RS del modello Unico sulle società di capitale deve essere utilizzato dai seguenti soggetti ai fini dell'imposta sul reddito delle società (IRES):

- società per azioni e in accomandita per azioni, società a responsabilità limitata, società cooperative (comprese quelle che abbiano acquisito la qualifica ONLUS e le cooperative sociali), società di mutua assicurazione, società europee di cui al regolamento CE n.2157/2001 e società cooperative europee di cui al regolamento CE n.1453/2003 residenti nel territorio dello Stato;
- enti commerciali (enti pubblici e privati diversi dalle società, trust che hanno per oggetto esclusivo o principale l'esercizio di attività commerciali) residenti nel territorio dello stato;
- società di ogni tipo (tranne società semplici, società e associazioni ad esse equiparate ai sensi dell'art. 5 del TUIR) nonché enti commerciali non residenti nel territorio dello Stato, compresi i trust, che hanno esercitato l'attività nel territorio dello Stato mediante stabile organizzazione.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Attraverso il codice impresa si sono individuate 44.035 unità della fonte modello Unico società di capitale (quadro RS) che rientrano nel campione teorico della rilevazione. Dalle variabili di base del quadro RS, riferito all'anno 2007, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_uscrs	Ricavi delle vendite e delle prestazioni
ric_altri_uscrs	Altri ricavi e proventi della produzione
var_rpfpc_l_uscrs	Variazione delle rimanenze prodotti finiti, semilavorati e in corso di lavorazione
var_lavco_uscrs	Variazione dei lavori in corso su ordinazione
acq_beni_uscrs	Costi della produzione per materie prime, sussidiarie e di consumo + Costi della produzione per merci
acq_ser_uscrs	Costi della produzione per servizi e godimento beni di terzi - Costi per canoni di locazione finanziaria (comprende una parte dei beni di godimento di terzi)
acq_bese_uscrs	acq_beni_uscrs + acq_ser_uscrs
acq_gdbt_uscrs	Costi per canoni di locazione finanziaria (comprende solo una parte)
clav_uscrs	Costi della produzione per il personale
acq_amm_uscrs	Ammortamenti e svalutazioni delle immobilizzazioni immateriali e materiali
acq_acc_uscrs	Accantonamenti per rischi + Altri accantonamenti
acq_oneri_uscrs	Oneri diversi di gestione
var_rmpriv_uscrs	Variazione delle rimanenze di materie prime, sussidiarie e di consumo + Variazione delle rimanenze di merci
var_rriv_uscrs	Variazione delle rimanenze di merci
var_riman_uscrs	Variazione delle rimanenze= var_rpfpc_l_uscrs - var_rmpriv_uscrs
vagg_uscrs	Valore aggiunto = (fatt_tot_uscrs + var_rpfpc_l_uscrs + var_lavco_uscrs + inc_imml_i_uscrs + ric_altri_uscrs) - (acq_beni_uscrs + acq_ser_uscrs + var_rmpriv_uscrs + acq_oneri_uscrs)
margope_uscrs	Margine operativo lordo = vagg_uscrs - clav_uscrs

A.5.10. Modello Irap Società di capitale – Quadro IQ: Irap

Il modello Irap sulle società di capitale (quadro IQ) deve essere compilato dalle società di capitale, enti commerciali ed equiparati per la dichiarazione dell'imposta regionale sulle attività produttive. L'imposta è determinata applicando al valore della produzione netta l'aliquota del 4,25%, salvo quanto previsto dal comma 1 dell'art.45 del D.lgs. n.466, nonché da leggi regionali o delle Province autonome. Il quadro IQ deve essere compilato dalle società semplici, in nome collettivo e in accomandita semplice e quelle ad esse equiparate a norma dell'art.5 del TUIR, comprese le associazioni costituite tra persone fisiche per l'esercizio in forma associata di arti e professioni.

Attraverso un link delle informazioni disponibili con l'archivio ASIA, tramite il codice fiscale o la partita IVA, si associano ad ogni impresa le informazioni di archivio rappresentate dal codice impresa, l'attività economica, gli addetti e dipendenti aggiornati all'anno di riferimento della rilevazione. Dopodiché attraverso il codice impresa si individuano quante imprese della fonte modello

Irap società di capitale (quadro IQ) rientrano nel campione teorico della rilevazione. L'accoppiamento tra le fonti ha riguardato 18.259 unità, dalle quali, a partire dalle informazioni di base quadro IQ, sono state costruite le seguenti principali variabili:

Codifica delle variabili	Corrispondenza con le voci della fonte utilizzata
fatt_tot_uscic	Ricavi delle vendite e delle prestazioni
var_rpfpc_uscic	Variazione delle rimanenze prodotti finiti, semilavorati, in corso di lavorazione
var_lavco_uscic	Variazione di lavori in corso su ordinazione
inc_imml_uscic	Incrementi di immobilizzazioni per lavori interni
ric_altri_uscic	Altri ricavi e proventi
acq_beni_uscic	Costi per materie prime, sussidiarie e di consumo
acq_ser_uscic	Costi per servizi
acq_bese_uscic	= $acq_beni_uscic + acq_ser_uscic$
acq_gdbt_uscic	Costi per godimento beni di terzi
clav_uscic	Retribuzioni complessive per dipendenti e assimilati operanti in Italia
clav2_uscic	Retribuzioni complessive per dipendenti e assimilati operanti all'estero + Retribuzioni complessive per dipendenti e assimilati operanti in Italia
clav3_uscic	Retribuzioni ed altri compensi
acq_amm_uscic	Ammortamento delle immobilizzazioni materiali Ammortamento delle immobilizzazioni immateriali
acq_acc_uscic	Accantonamenti
var_rmpriv_uscic	Variazione delle rimanenze di materie prime e merci
acq_oneri_uscic	Oneri diversi di gestione
var_riman_uscic	<i>Variazione delle rimanenze</i> = $var_rpfpc_uscic - var_rmpriv_uscic$
vagg_uscic	<i>Valore aggiunto</i> = $(fatt_tot_uscic + var_rpfpc_uscic + var_lavco_uscic + inc_imml_uscic + ric_altri_uscic) - (acq_beni_uscic + acq_ser_uscic + acq_gdbt_uscic + var_rmpriv_uscic + acq_oneri_uscic)$
margope_uscic	<i>Margine operativo lordo</i> = $vagg_uscic - clav_uscic$
margope2_uscic	<i>Margine operativo lordo2</i> = $vagg_uscic - clav2_uscic$
margope3_uscic	<i>Margine operativo lordo3</i> = $vagg_uscic - clav3_uscic$

Appendice B - Tavole dettagliate della copertura dell'archivio Asia e del campione PMI con le fonti amministrative disponibili

Tavola B.1 - Percentuale di copertura in termini di imprese e addetti dell'universo di riferimento dell'indagine PMI per fonte, attività economica e classe di addetti - Anno 2007

	Classi di addetti																Totale	
	1		2		3..5		6..9		10..19		20..49		50..99		Imprese	Addetti		
	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti				
Bilanci	C - Estrazione di minerali	26,8	26,8	32,6	32,7	45,0	45,6	59,0	59,2	67,9	68,4	83,2	84,7	91,5	91,4	49,9	69,8	
	D - Attività manifatturiere	6,9	7,0	9,0	8,9	18,2	19,1	34,8	35,6	54,6	55,7	80,1	82,1	92,2	92,3	22,5	54,5	
	E - Produzione e distribuzione di energia elettrica, gas e acqua	44,0	44,1	53,1	53,3	66,1	66,6	79,9	80,7	85,0	85,3	85,6	86,1	89,7	90,1	57,5	81,8	
	F - Costruzioni	8,2	8,2	8,2	8,2	16,9	17,7	33,4	34,1	52,8	54,1	75,5	77,0	86,6	86,6	14,3	33,4	
	G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	4,6	4,6	8,8	8,7	20,6	21,6	40,0	40,9	62,4	63,5	80,4	81,7	88,0	88,0	11,1	30,7	
	H - Alberghi e ristoranti	5,0	5,1	5,1	5,0	9,9	10,4	20,9	21,5	41,2	42,8	66,2	68,0	84,8	85,0	10,8	24,5	
	I - Trasporti, magazzino e comunicazioni	4,5	4,6	11,9	11,8	26,0	27,2	41,7	42,0	61,0	62,2	75,9	77,4	84,3	84,1	16,3	47,6	
	J - Attività finanziarie	2,2	2,3	8,0	7,9	12,9	13,0	20,3	20,8	33,7	34,6	53,6	53,6	46,7	46,6	6,1	13,9	
	K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	9,7	9,8	16,0	15,9	25,7	26,6	44,8	45,5	62,9	64,0	78,4	79,2	84,6	84,9	13,9	31,5	
	M - Istruzione	9,1	9,2	16,2	16,1	23,6	24,5	36,3	36,6	59,7	61,0	80,5	81,0	84,0	83,4	19,2	45,6	
	N - Sanità e assistenza sociale	0,6	0,6	3,5	3,5	13,6	15,0	43,4	44,6	73,7	74,9	89,2	89,9	92,3	92,4	4,6	31,1	
O - Altri servizi pubblici, sociali e personali	3,6	3,6	4,7	4,7	13,2	14,1	33,5	34,6	57,8	58,8	76,7	78,2	81,4	82,4	7,8	26,2		
Totale	6,5	6,6	9,3	9,2	18,5	19,4	35,4	36,1	56,1	57,3	78,5	80,2	88,8	89,0	13,2	37,0		
SdS F	C - Estrazione di minerali	50,1	50,2	51,2	51,2	47,7	47,8	36,4	36,1	28,5	28,2	10,6	8,9	-	-	39,5	23,9	
	D - Attività manifatturiere	74,1	73,9	77,2	77,4	71,5	70,9	57,8	57,0	39,7	38,7	14,7	12,6	2,0	1,8	64,8	37,9	
	E - Produzione e distribuzione di energia elettrica, gas e acqua	2,6	2,5	3,1	3,2	2,1	2,2	1,0	0,9	0,5	0,6	0,5	0,3	-	-	2,1	0,6	
	F - Costruzioni	76,9	76,7	79,1	79,3	72,6	72,0	58,3	57,6	40,6	39,3	17,2	15,5	3,3	3,2	72,9	56,5	
	G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	80,7	80,5	81,9	82,1	72,8	72,0	55,1	54,3	31,9	30,6	10,3	8,7	1,0	0,9	77,0	60,1	
	H - Alberghi e ristoranti	73,5	73,2	78,9	79,1	79,8	79,7	73,1	72,6	53,7	52,1	29,2	27,5	10,9	10,6	75,1	66,1	
	I - Trasporti, magazzino e comunicazioni	79,1	78,9	72,5	72,7	60,2	59,2	47,5	47,3	29,8	28,6	12,9	11,6	2,6	2,5	68,7	39,9	
	J - Attività finanziarie	73,6	73,5	70,3	70,4	75,7	76,0	69,8	69,3	52,9	51,1	15,6	14,7	-	-	72,8	68,0	
	K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	20,4	20,4	40,6	40,8	34,2	33,7	23,0	22,6	15,4	15,2	9,2	8,6	4,3	4,0	23,9	22,8	
	M - Istruzione	13,3	13,4	44,3	44,7	37,7	36,4	19,4	18,7	4,6	4,3	-	-	-	-	22,6	15,9	
	N - Sanità e assistenza sociale	1,6	1,6	4,6	4,6	9,0	9,4	13,6	13,6	7,6	7,3	2,1	1,8	1,3	1,2	2,9	4,2	
O - Altri servizi pubblici, sociali e personali	60,0	60,1	77,0	77,1	74,7	73,8	52,3	51,2	28,3	27,3	10,6	9,2	2,7	2,3	64,5	53,9		
Totale	53,2	53,3	69,5	69,6	65,5	65,1	53,8	53,2	35,7	34,6	13,8	12,1	2,6	2,4	56,6	45,1		
SdS G	C - Estrazione di minerali	0,4	0,4	-	-	-	-	-	-	-	-	-	-	-	-	0,1	0,0	
	D - Attività manifatturiere	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	-	-	-	-	-	-	0,0	0,0	
	E - Produzione e distribuzione di energia elettrica, gas e acqua	0,3	0,3	-	-	-	-	-	-	-	-	-	-	-	-	0,1	0,0	
	F - Costruzioni	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	-	-	-	-	-	-	0,1	0,0	
	G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	0,1	0,1	0,0	0,0	0,0	0,0	-	-	-	-	-	-	0,1	0,0	0,1	0,0	
	H - Alberghi e ristoranti	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	-	-	-	-	0,0	0,0	
	I - Trasporti, magazzino e comunicazioni	1,8	1,8	0,5	0,5	0,1	0,1	0,0	0,0	0,0	0,0	-	-	-	-	1,1	0,3	
	J - Attività finanziarie	7,1	7,1	9,0	9,0	2,6	2,4	1,7	1,7	0,3	0,3	-	-	-	-	6,3	4,3	
	K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	56,8	56,7	33,0	33,0	30,3	30,1	23,9	23,5	13,2	12,3	2,0	1,6	0,1	0,0	49,9	34,6	
	M - Istruzione	2,8	2,8	0,4	0,4	0,1	0,1	0,2	0,2	-	-	-	-	-	-	1,5	0,5	
	N - Sanità e assistenza sociale	86,0	85,9	85,7	85,8	69,9	68,0	32,5	31,2	7,6	6,9	0,6	0,5	-	-	81,9	55,3	
O - Altri servizi pubblici, sociali e personali	5,6	5,5	1,9	1,9	0,2	0,2	0,0	0,0	0,0	0,0	-	-	-	-	3,7	1,8		
Totale	24,6	24,4	9,5	9,5	6,4	6,2	3,2	3,1	1,4	1,3	0,2	0,2	0,0	0,0	17,0	7,8		
Asia	C - Estrazione di minerali	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	D - Attività manifatturiere	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	E - Produzione e distribuzione di energia elettrica, gas e acqua	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	F - Costruzioni	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	H - Alberghi e ristoranti	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	I - Trasporti, magazzino e comunicazioni	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	J - Attività finanziarie	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	M - Istruzione	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
	N - Sanità e assistenza sociale	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
O - Altri servizi pubblici, sociali e personali	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0		
Totale	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0		

Tavola B.2 - Percentuale di copertura in termini di imprese e addetti dell'universo di riferimento dell'indagine PMI per fonte, forma giuridica e classe di addetti - Anno 2007

	Classi di addetti														Totale		
	1		2		3..5		6..9		10..19		20..49		50..99		Imprese	Addetti	
	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti	Imprese	Addetti			
Bilanci																	
Imprenditore individuale, libero professionista e lavoratore autonomo	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Società di persone	0,1	0,1	0,0	0,0	0,1	0,1	0,1	0,1	0,3	0,3	1,5	1,7	6,0	6,1	0,1	0,3	
Società di capitali	63,1	63,1	78,4	78,8	86,9	87,4	90,6	90,7	92,5	92,6	93,6	93,7	93,8	93,8	78,8	90,2	
Altre società (cooperative, consorzi, enti)	33,8	33,8	47,0	47,5	61,7	62,0	70,8	71,4	80,9	81,2	84,3	84,2	85,1	85,1	57,5	77,9	
Totale	6,5	6,6	9,3	9,2	18,5	19,4	35,4	36,1	56,1	57,3	78,5	80,2	88,8	89,0	13,2	37,0	
SdS F																	
Imprenditore individuale, libero professionista e lavoratore autonomo	54,9	55,0	75,4	75,6	80,4	80,7	83,8	83,8	83,6	83,6	80,5	79,3	57,6	56,5	61,0	68,1	
Società di persone	82,6	82,5	81,6	81,6	84,0	84,4	88,4	88,5	88,6	88,5	79,7	77,4	40,8	39,6	83,6	84,4	
Società di capitali	22,0	21,9	9,9	9,8	5,6	5,4	4,0	4,0	3,1	3,1	1,8	1,7	0,6	0,6	11,3	3,9	
Altre società (cooperative, consorzi, enti)	14,5	14,5	11,7	11,7	9,2	9,2	6,5	6,3	3,6	3,6	1,9	1,8	1,4	1,4	9,2	3,8	
Totale	53,2	53,3	69,5	69,6	65,5	65,1	53,8	53,2	35,7	34,6	13,8	12,1	2,6	2,4	56,6	45,1	
SdS G																	
Imprenditore individuale, libero professionista e lavoratore autonomo	29,8	29,7	14,0	14,0	10,7	10,4	7,1	7,1	5,5	5,2	0,8	0,8	1,0	0,8	25,1	19,4	
Società di persone	0,8	0,8	5,2	5,2	5,7	5,7	4,2	4,1	2,8	2,7	1,6	1,5	0,2	0,1	4,1	4,1	
Società di capitali	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	-	-	-	-	-	-	0,0	0,0	
Altre società (cooperative, consorzi, enti)	0,1	0,1	0,1	0,1	0,0	0,0	-	-	0,0	0,0	-	-	-	-	0,0	0,0	
Totale	24,6	24,4	9,5	9,5	6,4	6,2	3,2	3,1	1,4	1,3	0,2	0,2	0,0	0,0	17,0	7,8	
Asia																	
Imprenditore individuale, libero professionista e lavoratore autonomo	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
Società di persone	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
Società di capitali	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
Altre società (cooperative, consorzi, enti)	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	
Totale	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	

Tavola B.3 - Numero di imprese e percentuale di copertura del campione teorico per fonte, attività economica e classe di addetti - Anno 2007

	Numero di imprese									Percentuale di copertura									
	Classi di addetti								Totale	Classi di addetti								Totale	
	1	2	3.5	6.9	10.19	20.49	50.99	1		2	3.5	6.9	10.19	20.49	50.99				
Bilanci																			
C - Estrazione di minerali	31	19	53	75	115	112	35	440	11,3	14,5	28,8	50,3	49,8	72,3	74,5	37,4			
D - Attività manifatturiere	492	346	1.084	1.397	3.681	4.240	2.213	13.453	4,6	8,0	19,0	36,7	50,5	69,1	78,2	32,8			
E - Produzione e distribuzione di energia elettrica, gas e acqua	121	33	82	58	89	109	49	541	28,2	32,7	58,2	60,4	64,0	67,7	71,0	47,3			
F - Costruzioni	67	39	74	100	326	357	169	1.132	4,2	6,5	11,4	26,2	42,4	61,4	72,5	23,4			
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	356	321	636	454	1.325	1.259	543	4.894	5,6	10,0	23,9	38,0	49,5	66,4	75,5	26,1			
H - Alberghi e ristoranti	53	19	44	55	136	145	89	541	5,7	4,8	9,2	23,9	39,1	50,5	73,0	19,2			
I - Trasporti, magazzinaggio e comunicazioni	177	130	350	243	468	448	211	2.027	8,5	14,8	29,2	41,4	55,6	63,6	70,1	30,4			
J - Attività finanziarie	2	6	18	10	29	42	8	115	0,6	1,7	6,1	11,2	19,9	43,3	40,0	8,5			
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	1.089	430	579	383	979	819	385	4.664	15,4	17,8	28,4	48,9	63,8	67,7	72,0	29,7			
M - Istruzione	29	26	59	43	102	82	46	387	4,8	16,0	21,9	32,1	49,8	53,9	75,4	24,3			
N - Sanità e assistenza sociale	8	8	28	34	124	144	93	439	1,0	6,3	21,7	31,8	53,7	64,3	69,9	24,6			
O - Altri servizi pubblici, sociali e personali	124	66	179	177	368	398	164	1.476	4,5	10,1	24,2	42,8	51,0	60,6	64,8	23,7			
Totale	2.549	1.443	3.186	3.029	7.742	8.155	4.005	30.109	7,5	10,8	22,0	38,0	51,2	66,5	75,3	29,2			
SdS																			
C - Estrazione di minerali	91	53	81	57	86	24	0	392	33,1	40,5	44,0	38,3	37,2	15,5	-	33,3			
D - Attività manifatturiere	5250	2431	3.048	1.671	2.427	986	103	15.916	49,2	56,2	53,5	43,8	33,3	16,1	3,6	38,8			
E - Produzione e distribuzione di energia elettrica, gas e acqua	16	1	3	0	2	2	0	24	3,7	1,0	2,1	-	1,4	1,2	-	2,1			
F - Costruzioni	1206	449	472	222	362	137	15	2.863	75,3	74,7	72,6	58,1	47,1	23,6	6,4	59,1			
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	4406	2360	1634	558	976	263	12	10.209	69,8	73,6	61,4	46,7	36,4	13,9	1,7	54,5			
H - Alberghi e ristoranti	570	281	329	126	150	106	22	1584	61,4	70,3	68,5	54,8	43,1	36,9	18,0	56,1			
I - Trasporti, magazzinaggio e comunicazioni	975	458	510	198	188	100	23	2.452	46,8	52,2	42,5	33,7	22,3	14,2	7,6	36,8			
J - Attività finanziarie	292	291	227	56	71	14	0	951	81,6	82,9	76,9	62,9	48,6	14,4	-	70,0			
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	4.150	1446	1014	231	312	159	34	7.346	58,7	59,8	49,8	29,5	20,3	13,2	6,4	46,8			
M - Istruzione	72	34	42	14	14	3	0	179	11,9	20,9	15,6	10,4	6,8	2,0	-	11,2			
N - Sanità e assistenza sociale	639	81	54	22	48	23	4	871	78,6	63,8	41,9	20,6	20,8	10,3	3,0	48,9			
O - Altri servizi pubblici, sociali e personali	1228	352	322	94	133	92	11	2.232	45,0	53,8	43,5	22,7	18,4	14,0	4,3	35,9			
Totale	18.895	8.237	7.736	3.249	4.769	1.909	224	45.019	55,8	61,7	53,4	40,7	31,5	15,6	4,2	43,7			
Unico PF-RG, PF-RE, SP-RG, SC-RS																			
C - Estrazione di minerali	18	14	11	3	7	4	3	60	6,5	10,7	6,0	2,0	3,0	2,6	6,4	5,1			
D - Attività manifatturiere	1465	496	498	255	413	466	300	3.893	13,7	11,5	8,7	6,7	5,7	7,6	10,6	9,5			
E - Produzione e distribuzione di energia elettrica, gas e acqua	161	36	29	21	26	23	11	307	37,5	35,6	20,6	21,9	18,7	14,3	15,9	26,8			
F - Costruzioni	77	27	25	11	21	28	34	223	4,8	4,5	3,8	2,9	2,7	4,8	14,6	4,6			
G - Commercio all'ingrosso e al dettaglio, riparazione di autoveicoli, motocicli e di beni personali e per la casa	405	142	109	59	146	238	102	1.201	6,4	4,4	4,1	4,9	5,4	12,6	14,2	6,4			
H - Alberghi e ristoranti	141	59	57	25	23	16	5	326	15,2	14,8	11,9	10,9	6,6	5,6	4,1	11,5			
I - Trasporti, magazzinaggio e comunicazioni	484	161	180	69	85	87	35	1.101	23,2	18,4	15,0	11,8	10,1	12,4	11,6	16,5			
J - Attività finanziarie	27	35	24	6	7	8	1	108	7,5	10,0	8,1	6,7	4,8	8,2	5,0	7,9			
K - Attività immobiliari, noleggio, informatica, ricerca, servizi alle imprese	1.156	312	262	83	133	128	68	2.142	16,3	12,9	12,9	10,6	8,7	10,6	12,7	13,7			
M - Istruzione	421	72	122	50	55	46	9	775	69,8	44,2	45,2	37,3	26,8	30,3	14,8	48,6			
N - Sanità e assistenza sociale	86	20	30	29	34	41	26	266	10,6	15,7	23,3	27,1	14,7	18,3	19,5	14,9			
O - Altri servizi pubblici, sociali e personali	1051	159	141	79	117	101	55	1.703	38,5	24,3	19,0	19,1	16,2	15,4	21,7	27,4			
Totale	5.492	1.533	1.488	690	1.067	1.186	649	12.105	16,2	11,5	10,3	8,6	7,1	9,7	12,2	11,7			

Tavola B.4 - Numero di imprese e percentuale di copertura del campione teorico per fonte, forma giuridica e classe di addetti - Anno 2007

		Numero imprese							Totale	Numero imprese							Totale
		Classi di addetti								Classi di addetti							
		1	2	3..5	6..9	10..19	20..49	50..99		1	2	3..5	6..9	10..19	20..49	50..99	
Bilanci	Imprenditore individuale, libero professionista e lavoratore autonomo	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-
	Società di persone	0	0	4	1	13	18	8	44	-	-	0,1	0,0	0,4	1,4	4,6	0,2
	Società di capitali	2.407	1.386	3.004	2.807	7.187	7.496	3.690	27.977	31,5	46,3	60,1	69,9	75,0	78,3	80,2	63,7
	Altre società (cooperative, consorzi, enti)	142	57	178	221	542	641	307	2.088	16,1	19,8	28,4	38,0	52,2	58,2	60,3	40,8
	Totale	2.549	1.443	3.186	3.029	7.742	8.155	4.005	30.109	7,5	10,8	22,0	38,0	51,2	66,5	75,3	29,2
SdS	Imprenditore individuale, libero professionista e lavoratore autonomo	13594	4156	2940	841	997	243	16	22787	65,1	76,7	76,8	75,6	77,0	71,3	41,0	69,2
	Società di persone	3115	3373	3875	1818	2597	860	65	15703	69,7	72,5	77,0	80,1	80,7	69,1	37,4	74,5
	Società di capitali	2.097	658	827	516	1.085	724	114	6.021	27,4	22,0	16,6	12,9	11,3	7,6	2,5	13,7
	Altre società (cooperative, consorzi, enti)	89	50	94	74	90	82	29	508	10,1	17,4	15,0	12,7	8,7	7,4	5,7	9,9
	Totale	18.895	8.237	7.736	3.249	4.769	1.909	224	45.019	55,8	61,7	53,4	40,7	31,5	15,6	4,2	43,7
Unico PF-RG, PF-RE, SP-RG, SC-RS	Imprenditore individuale, libero professionista e lavoratore autonomo	3582	555	315	53	22	4	1	4532	17,2	10,2	8,2	4,8	1,7	1,2	2,6	13,8
	Società di persone	618	564	504	194	250	226	84	2440	13,8	12,1	10,0	8,5	7,8	18,2	48,3	11,6
	Società di capitali	1.013	331	481	300	599	753	471	3.948	13,2	11,1	9,6	7,5	6,3	7,9	10,2	9,0
	Altre società (cooperative, consorzi, enti)	279	83	188	143	196	203	93	1.185	31,7	28,8	30,0	24,6	18,9	18,4	18,3	23,1
	Totale	5.492	1.533	1.488	690	1.067	1.186	649	12.105	16,2	11,5	10,3	8,6	7,1	9,7	12,2	11,7
Campione PMI	Imprenditore individuale, libero professionista e lavoratore autonomo	213	42	61	42	42	16	3	419	1,0	0,8	1,6	3,8	3,2	4,7	7,7	1,3
	Società di persone	78	61	87	62	93	37	2	420	1,7	1,3	1,7	2,7	2,9	3,0	1,1	2,0
	Società di capitali	66	13	26	21	47	64	45	282	0,9	0,4	0,5	0,5	0,5	0,7	1,0	0,6
	Altre società (cooperative, consorzi, enti)	31	5	22	12	23	27	10	130	3,5	1,7	3,5	2,1	2,2	2,5	2,0	2,5
	Totale	388	121	196	137	205	144	60	1.251	1,1	0,9	1,4	1,7	1,4	1,2	1,1	1,2
Fuori copertura, errori di lista, nessuna fonte	Imprenditore individuale, libero professionista e lavoratore autonomo	3493	669	513	176	234	78	19	5186	16,7	12,3	13,4	15,8	18,1	22,9	48,7	15,8
	Società di persone	656	654	561	195	264	103	15	2459	14,7	14,1	11,2	8,6	8,2	8,3	8,6	11,7
	Società di capitali	2.067	607	658	369	661	532	280	5.700	27,0	20,3	13,2	9,2	6,9	5,6	6,1	13,0
	Altre società (cooperative, consorzi, enti)	340	93	144	132	188	148	70	1.210	38,6	32,3	23,0	22,7	18,1	13,4	13,8	23,6
	Totale	6.556	2.023	1.876	872	1.347	861	384	14.555	19,4	15,1	13,0	10,9	8,9	7,0	7,2	14,1
Totale	Imprenditore individuale, libero professionista e lavoratore autonomo	20882	5422	3829	1112	1295	341	39	32924	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	Società di persone	4467	4652	5031	2270	3217	1244	174	21066	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	Società di capitali	7.650	2.995	4.996	4.013	9.579	9.569	4.600	43.928	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	Altre società (cooperative, consorzi, enti)	881	288	626	582	1039	1101	509	5.121	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	Totale	33.880	13.357	14.482	7.977	15.130	12.255	5.322	103.039	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

APPENDICE C: La stima preliminare del regolamento SBS: un caso pratico di integrazione fra dati di fonte statistica e amministrativa

La stima preliminare dei principali aggregati economici delle imprese dell'industria e dei servizi, effettuata annualmente entro 10 mesi dalla fine dell'anno di riferimento, rientra nel quadro degli obblighi comunitari imposti agli Istituti Nazionali di Statistica dal regolamento sulle statistiche strutturali SBS n°58/97 e dal successivo n°295/2008. Le variabili oggetto di stima sono il fatturato, il valore della produzione, l'acquisto di beni e servizi, la retribuzione lorda e il numero di addetti medi; il livello di dettaglio richiesto, meno articolato rispetto a quello delle stime definitive SBS, corrisponde alle 3 cifre (cioè al gruppo) della classificazione Nace.

L'esigenza di produrre stime più tempestive rispetto a quelle definitive viene soddisfatta grazie all'utilizzo integrato di fonti amministrative e dati di indagine disponibili all'epoca dell'elaborazione. Il calcolo delle stime preliminari si basa su un approccio metodologico consolidato in oltre dieci anni e sempre più robusto grazie alla mole crescente di dati utilizzati. Per ottenere le stime preliminari delle variabili obiettivo nell'anno t si stimano, sulla base dei dati disponibili per il biennio $(t-1, t)$, le variazioni annue delle stesse variabili per gruppo Nace \times classe di addetti (1-99, 100 e oltre); la stima delle variazioni viene calcolata anche per classe di addetti – dettaglio non richiesto dai regolamenti – per tenere conto dell'influenza della dimensione delle imprese sulla dinamica delle variabili. Le variazioni calcolate, si applicano quindi ai rispettivi totali SBS stimati al tempo $t-1$, ottenendo la stima al tempo t dei totali delle variabili obiettivo per gruppo Nace \times classe di addetti. Sommando i totali ottenuti rispetto alla classe dimensionale, si ottengono infine le stime delle variabili per gruppo Nace, come prescritto dagli obblighi comunitari.

Le basi di dati utilizzate per la produzione delle stime preliminari relative all'anno di riferimento t sono le seguenti:

- ASIA (Archivio Statistico delle Imprese Attive) relativo all'anno $t-1$;
- stime SBS definitive per l'anno base $(t-1)$;
- *Bilanci civilistici* delle società di capitale depositati presso le Camere di Commercio; dati relativi al biennio $(t-1, t)$;
- dati INPS sulle imprese con dipendenti per il biennio $(t-1, t)$;
- dati delle imprese rispondenti all'indagine Sistema dei Conti delle Imprese (SCI) in corso (anno t), se disponibili.

La descrizione dell'archivio ASIA è stata oggetto del paragrafo A.1 dell'appendice. L'archivio ASIA disponibile al momento dell'elaborazione è quello aggiornato all'anno $t-1$ e viene utilizzato per definire il campo di osservazione della stima e per collegare ai database di fonte amministrativa le informazioni anagrafiche.

Le stime SBS definitive sono il risultato di due indagini Istat rivolte alle imprese operanti nei principali settori dell'industria e dei servizi: la rilevazione campionaria sulle piccole e medie imprese e sull'esercizio di arti e professioni della classe di addetti 1-99 (PMI), già illustrata in dettaglio nel paragrafo A.2, e la rilevazione sul sistema dei conti delle imprese (SCI), che censisce le imprese del restante sottoinsieme. Al momento in cui si eseguono le elaborazioni per il calcolo delle stime preliminari per l'anno t , le stime definitive SBS per l'anno $t-1$ sono state già trasmesse ad Eurostat; queste ultime, calcolate per diversi domini e comunque ad un livello di dettaglio maggiore rispetto a quello delle 3 cifre Nace richiesto per le corrispondenti stime preliminari, costituiscono – come si accennava – i livelli a cui applicare la variazione annua delle variabili, stimata mediante il complesso delle altre fonti.

I *bilanci civilistici* contengono tutte le variabili economiche per cui viene calcolata la stima preliminare. Dall'insieme dei bilanci forniti all'Istat, vengono esclusi le duplicazioni, le imprese non appartenenti ad ASIA $(t-1)$ e quelle per cui i dati di bilancio non sono presenti in entrambi gli anni di esercizio $t-1$ e t . In questo modo, nell'anno di riferimento 2008, si è pervenuti ad un panel *Bilanci* contenente 276.643 record.

Il database INPS è costituito da tutte le imprese con dipendenti che nel biennio $(t-1, t)$ hanno fornito all'Istituto di Previdenza le dichiarazioni mensili riguardanti i versamenti contributivi e le

informazioni sul numero e la tipologia degli occupati, sulle retribuzioni e sulle ore di lavoro. I dati mensili vengono abbinati con l'archivio ASIA ($t-1$) ed elaborati allo scopo di ottenere le informazioni anagrafiche sulle imprese ed i totali trimestrali delle variabili: numero di dipendenti, numero di ore lavorate e retribuzioni totali. Rispetto al database iniziale, vengono considerate le sole unità che hanno la variabile "dipendenti totali" diversa da zero in almeno un mese del primo semestre di ciascun anno e che – come già visto per i dati di bilancio – sono presenti nell'archivio ASIA ($t-1$). Infine, il panel INPS viene abbinato al panel *Bilanci* utilizzando come variabile chiave il codice fiscale, in modo da selezionare le sole unità compresenti nei due database *Bilanci* e INPS. Risultato dell'abbinamento è il panel ASIA-*Bilanci*-INPS, che nell'anno di riferimento 2008 era costituito da 229.211 unità e che costituisce la base di dati iniziale da cui stimare la dinamica delle variabili richieste tra l'anno base e l'anno di riferimento delle stime preliminari. In particolare, la variazione annua delle variabili economiche viene stimata mediante i dati di bilancio, mentre le variabili occupazionali utilizzate nei calcoli sono di fonte INPS.

Negli ultimi anni, all'informazione per l'anno t proveniente dai dati di bilancio e da quelli previdenziali, si è aggiunta quella dei dati dell'indagine SCI in corso, relativa all'anno t .²⁶ Ciò è stato reso possibile dall'introduzione, nel processo di produzione statistica di questa rilevazione, di nuove tecniche di indagine che, riducendo i tempi di disponibilità dei dati, li rendono in parte utilizzabili già all'epoca dell'elaborazione delle stime preliminari. In questo caso, la base di dati utilizzata per stimare la dinamica delle variabili obiettivo è costituita dall'unione del panel ASIA-*Bilanci*-INPS con l'insieme dei rispondenti all'indagine SCI in corso. Quando per un'impresa sono disponibili sia i dati di fonte *Bilanci*-INPS, sia quelli dell'indagine SCI, si preferisce utilizzare i dati di indagine ai fini della stima delle variazioni annue di tutte le variabili, economiche ed occupazionali.

Preliminarmente all'utilizzo del database ASIA-*Bilanci*-INPS – eventualmente integrato con i dati SCI – per stimare le variazioni delle variabili in ($t-1$, t), si procede ad analizzare l'informazione in esso contenuta, al fine di selezionare un sottoinsieme di record utili a produrre una stima robusta delle variabili di interesse. Una prima selezione di unità viene effettuata escludendo tutte le imprese che presentano una variazione ritenuta anomala in ($t-1$, t) di almeno una delle variabili oggetto di stima (fatturato, valore della produzione, acquisti di beni e servizi, retribuzioni e numero di addetti medi) e di altre variabili rilevanti (costo del lavoro, valore aggiunto e numero di dipendenti).

Un'ulteriore controllo, volto all'eliminazione di possibili outlier dal database ottenuto viene realizzata mediante il seguente procedimento:

- per ciascuna impresa del database, si calcolano i valori assunti dai principali indicatori economici nei due anni $t-1$ e t (ad esempio: fatturato per addetto, valore aggiunto per addetto, costo del lavoro per dipendente, retribuzione per dipendente);
- sempre per impresa, si calcola il rapporto tra i valori assunti da ciascun indicatore rispettivamente negli anni t e $t-1$;
- si suddivide la popolazione di imprese del database in sottopopolazioni corrispondenti a livelli di dettaglio variabili (dalla sezione al gruppo Nace, concatenate per la classe di addetti 1-99, 100 e oltre); la scelta del dettaglio dipende dalla distribuzione di frequenza delle imprese nei diversi settori di attività economica;
- per ciascuna sottopopolazione, si calcolano predeterminati quantili dei rapporti di cui al punto (b), in genere corrispondenti ai valori che lasciano su ciascuna coda della distribuzione l'1%, il 5% o il 10%;
- per ogni impresa, ognuno dei rapporti calcolati in (b) viene confrontato con una data coppia di percentili (ad esempio nel calcolo 2008 sono stati utilizzati il 10° e il 90° percentile) relativa alla sottopopolazione a cui l'impresa appartiene;

²⁶ Ad esempio, per l'anno di riferimento 2006 è stato possibile sfruttare l'informazione proveniente da 2.332 imprese rispondenti all'indagine SCI 2006, i cui dati erano utilizzabili già ad ottobre 2007, epoca del calcolo delle stime preliminari

- vengono eliminate dal panel le imprese che hanno almeno uno dei rapporti calcolati in (b) nelle “code” della distribuzione.

Le imprese che rimangono nel database dopo questa procedura di controllo sono quelle effettivamente utilizzate per stimare la dinamica delle variabili oggetto di stima.

Per ottenere la stima preliminare SBS della generica variabile obiettivo V nell'anno t utilizzando congiuntamente l'informazione amministrativa e da indagine presente nel database e la stima definitiva SBS relativa all'anno $t-1$, si procede nel modo seguente:

- si suddivide la popolazione delle imprese del database in due sottopopolazioni, distinte per la classe di addetti (1-99, 100 e oltre);
- per ciascuna sottopopolazione, si calcolano i totali della variabile V per gruppo Nace nei due anni – rispettivamente $V_{A,t-1}$ e $V_{A,t}$, dove l'indice A denota l'utilizzo della fonte prevalentemente amministrativa;
- sempre a livello di gruppo Nace \times classe di addetti, si calcola il rapporto $V_{A,t}/V_{A,t-1}$;
- dalle stime definitive SBS in $(t-1)$ si calcola mediante somma la stima finale $V_{SBS,t-1}$ della variabile V a livello di gruppo Nace \times classe di addetti;
- La stima preliminare della variabile V al tempo t , allo stesso dettaglio, è data dalla seguente:

$$\hat{V}_{SBS,t} = V_{SBS,t-1} \times \frac{V_{A,t}}{V_{A,t-1}}$$

- da cui, sommando in ciascun gruppo Nace i due totali stimati per le classi di addetti “1-99” e “100 e oltre”, si ottiene la stima provvisoria della variabile V al tempo t .

I risultati delle stime preliminari SBS, confrontati ex post con le corrispondenti stime definitive, hanno finora condotto a risultati accettabili, con differenze tra le due stime che tendono a ridursi man mano che il calcolo viene effettuato a un livello più aggregato (fino al macrosettore di attività economica). Resta tuttavia aperto il problema degli effetti distorsivi potenzialmente indotti nella stima delle variabili dalla peculiare base di dati utilizzata per la stima della dinamica delle variabili; infatti, questa è costituita da società di capitale e/o imprese con almeno 100 addetti che non possono dirsi rappresentative dell'intera popolazione di imprese a cui si riferisce la stima.

APPENDICE D: Gli adempimenti amministrativi delle imprese con un solo addetto senza dipendenti e delle imprese minime

L'obiettivo è quello di analizzare le imprese italiane con un solo addetto in media annua senza dipendenti (di seguito "microunità"), estratte dall'archivio ASIA riferito all'anno 2007 in classificazione Ateco2002. al fine di individuarne le caratteristiche strutturali e di classificarle per attività economica svolta e forma giuridica. La decisione di investigare su questo segmento di imprese è stato condizionato solo in parte dal regolamento sulle statistiche strutturali, in quanto esso richiede, per il settore dell'industria e delle costruzioni, di fornire delle stime per una classe di addetti più ampia (imprese con 1-9 addetti, attive nel corso dell'anno per più di 6 mesi, a livello di Ateco2002 a 3 cifre) mentre, per il settore dei servizi, di produrre delle stime per tutte le imprese con 1 solo addetto in media annua (a prescindere dalla presenza o meno di dipendenti), sempre a livello di Ateco2002 a 3 cifre.

L'analisi si concentrerà, quindi, sulle imprese minime (introdotte con la legge finanziaria del 2008 e presenti nel quadro CM del modello unico 2009, riferito all'anno di imposta 2008) e sulle imprese artigiane, che, in base al processo produttivo adottato, presentano, a parità di classe dimensionale e attività economica, caratteristiche differenti dalle altre imprese.

Le microunità oggetto di analisi, che comprendono tutte le imprese con un addetto in media annua e sono senza dipendenti, possono avere oltre all'imprenditore individuale che lavora presso l'impresa, un lavoratore indipendente (socio oppure coadiuvante familiare, socio lavoratore nel caso di cooperative) che lavora meno di 6 mesi nell'impresa.

Questa tipologia di impresa comprende anche le unità che, costituite sotto forma di società di persone o di capitali, hanno soci non lavoratori che apportano esclusivamente capitale²⁷ o si tratta di microunità che fanno ricorso a lavoratori esterni (quali consulenti, lavoratori con contratto di somministrazione lavoro, parasubordinati, come lavoratori con contratti a progetto,²⁸ associati in partecipazione).²⁹

Per quanto concerne le cooperative, il socio lavoratore può avere un rapporto di lavoro in forma subordinata o autonoma o altra forma (quale ad esempio il lavoro parasubordinato).³⁰ Nel caso in cui nessun socio lavoratore sia iscritto come lavoratore dipendente e gli altri come lavoratori autonomi o parasubordinati, le microunità includerebbero anche queste cooperative. I soci inquadrati come lavoratori autonomi, infatti, sono imprenditori che svolgono nella cooperativa solo attività di indirizzo e di coordinamento e ricevono un compenso che viene integrato con l'attività lavorativa prevalente, costituita dalla figura di imprenditore nell'impresa associata alla cooperativa.

I soci lavoratori delle società di persone, considerati dipendenti e inquadrati come lavoratori subordinati, possono esserci solo per le società in accomandita semplice per la categoria di soci che non possono gestire la società e, quindi, sono, di fatto, subordinati all'altra categoria di soci (vale sempre il principio che un socio non può essere subordinato di se stesso).

I soci lavoratori delle società di capitali, oltre al caso delle società in accomandita per azioni, si possono avere per dipendenti remunerati esclusivamente con azioni della società o nel caso di soci di minoranza di fatto subordinati al socio di maggioranza che ha potere dominante.

Le imprese minime, invece, non possono avere lavoratori esterni ma possono essere anche imprese familiari e, quindi, non appartenere al campo di osservazione delle microunità, se hanno, oltre al titolare, almeno un coadiuvante familiare che lavora per più di 6 mesi.

²⁷ I soci lavoratori, remunerati solo in base alla partecipazione agli utili, rientrano, invece, nella definizione statistica di indipendenti.

²⁸ I lavoratori con contratto a progetto sono parasubordinati, che sostituiscono i co.co.co. in quanto non ci può essere lavoro parasubordinato senza che ci sia un progetto ben definito.

²⁹ Gli associati in partecipazione sono soci che apportano esclusivamente lavoro, vengono remunerati mediante la partecipazione agli utili e sono trattati come lavoratori esterni.

³⁰ Art 1, comma 3 della 'revisione della legislazione in materia cooperativistica, con particolare riferimento alla posizione del socio lavoratore', pubblicata in GU n. 94 del 23 aprile 2001. Nelle cooperative, il socio è sempre lavoratore in quanto le cooperative hanno scopo mutualistico (nel senso che sono costituite per aumentare i ricavi o ridurre i costi delle imprese associate) e non di lucro (e, quindi, non prevedono la distribuzione degli utili). Ciò significa che i soci non possono essere remunerati mediante la distribuzione degli utili, anche se la partecipazione alla cooperativa implica vantaggi indiretti per l'impresa associata.

Alcune variabili di interesse statistico richieste dal Regolamento possono essere soddisfatte, in generale, da quelle amministrative in misura significativamente differente, a seconda del regime fiscale adottato dalle imprese (minimo, semplificato e ordinario) e della loro forma giuridica (lavoratore autonomo, ditta individuale, società di persone o di capitali).

D.1. La classificazione delle microunità in base agli adempimenti amministrativi

Le microunità (definite come imprese con un addetto medio senza dipendenti) hanno differenti adempimenti amministrativi che dipendono dalle seguenti variabili:

- forma giuridica;
- classe di fatturato;
- settore economico di appartenenza;
- altre variabili economiche non presenti nell'archivio ASIA.

Tali imprese, inoltre, hanno sempre la facoltà di optare per il regime fiscale ordinario, anche se rientrano in quello dei minimi o semplificato.

L'analisi considera, inoltre, l'iscrizione delle imprese all'albo degli artigiani in quanto l'impresa artigiana ha un modello di processo produttivo (non in serie oppure in serie ma semi-automatizzato) che, a parità di settore e classe di fatturato, la potrebbe differenziare dalle altre imprese operanti sul mercato, anche per ragioni legate a leggi agevolate e contributi statali. Le imprese artigiane, infatti, possono accedere a diversificate forme di finanziamento che ne agevolano l'avvio di una nuova attività o favoriscono la crescita di quelle già esistenti. Tali aiuti, tuttavia, sono diversificati in base al territorio di appartenenza (si tratta spesso di leggi regionali che regolamentano gli incentivi per lo sviluppo delle imprese artigiane), al settore economico, alla classe dimensionale, al sesso (riguardano solo l'imprenditoria femminile) e alla nuova imprenditorialità o imprenditorialità giovanile.

In base agli adempimenti amministrativi e soprattutto in relazione alla classe di fatturato, le microunità possono essere raggruppate nel seguente modo:

A) microunità con classe di fatturato minima, ovvero con un volume d'affari non superiore a 30.000 euro.

- A1) ditta individuale minima
- A2) lavoratore autonomo minimo
- A3) SRL a socio unico minime
- A4) SRL minime
- A5) società di persone in contabilità semplificata minime
- A6) SPA a socio unico minime

È importante sottolineare che questo gruppo di imprese appartiene al regime dei minimi solo se utilizza la forma di ditta individuale o lavoratore autonomo. Le imprese minime, inoltre, non devono avere esportazioni, lavoratori esterni e non devono aver effettuato, negli ultimi tre anni, investimenti o affitti di beni strumentali (compreso il leasing) di importo superiore a 15.000 euro. Le imprese minime, inoltre, possono essere imprese familiari che non rientrano nel dominio delle microunità.

Per omogeneità di comparazione all'interno del gruppo A), sono stati esclusi alcuni settori economici³¹ per i quali non si applica il regime dei contribuenti minimi.

Le società di persone hanno una più esaustiva modulistica fiscale (Unico, Irap, *Studi di Settore*) così come le società di capitali (SRL a socio unico, SRL e società per azioni a socio unico) che possono redigere il bilancio anche in forma abbreviata, che contiene variabili più omogenee, anche in termini definitivi, con quelle richieste dal Regolamento SBS.

³¹ I settori sono i seguenti: 1) vendita Sali e tabacchi; 2) editoria; 3) gestione di servizi di telefonia pubblica; 4) agenzie di viaggio e di turismo; 5) agriturismo; 6) vendite a domicilio; 7) rivendita di beni usati, di oggetti d'arte, d'antiquariato o da collezione; 8) agenzie di vendita all'asta di oggetti d'arte, antiquariato o da collezione.

Le società di persone e di capitale rientrano in questa casistica se ci sono anche soci non lavoratori.

Va infine notato che questo gruppo di imprese possiede caratteristiche strutturali simili, anche se l'appartenenza o meno al settore artigiano e la presenza di lavoratori autonomi³² rappresenta una discriminante importante.

B) microunità con classe di fatturato media, ovvero con un volume d'affari compreso tra 30.000 euro e 309.874,14 se appartenenti al settore dei servizi e compreso tra 30.000 e 516.456,9 se appartenenti al settore dell'industria e delle costruzioni.

- B1) ditta individuale in contabilità semplificata media
- B2) lavoratore autonomo in contabilità semplificata medio
- B3) SRL a socio unico media
- B4) SRL media
- B5) società di persone in contabilità semplificata medie
- B6) SPA a socio unico media

Il lavoratore autonomo, a prescindere dal volume d'affari, rientra naturalmente nel regime semplificato, a meno che non abbia i requisiti per rientrare nel regime dei minimi.

C) microunità con classe di fatturato grande, ovvero con un volume d'affari superiore a 309.874,14 se appartenenti al settore dei servizi e superiore a 516.456,9 se appartenenti al settore dell'industria e delle costruzioni.

- C1) ditta individuale in contabilità ordinaria grande
- C2) lavoratore autonomo in contabilità semplificata grande
- C3) SRL a socio unico grande
- C4) SRL grande
- C5) società di persone in contabilità ordinaria grande
- C6) SPA a socio unico grande

Questo gruppo è il più completo in base alle variabili amministrative di interesse statistico. I lavoratori autonomi, a prescindere dal volume d'affari, devono redigere la contabilità semplificata, a meno che optino per quella ordinaria.

D) altre forme giuridiche

- D1) cooperative
- D2) consorzi
- D3) imprese estere operanti in Italia
- D4) altre forme giuridiche

Si tratta di gruppi di imprese distinti dagli altri, non solo per forma giuridica e relativi adempimenti amministrativi, ma anche per la funzione economica che svolgono nel mercato. Le cooperative ed i consorzi, infatti, sono di solito aggregazioni tra micro e piccole imprese che, superando il limite dimensionale, perseguono obiettivi specifici, dando luogo, ad esempio, a cooperative di consumo, consorzi fidi, cooperative export, per l'acquisto di materie prime oppure per attività pubblicitarie.

Le imprese estere operanti in Italia, appartenendo a gruppi stranieri spesso di grandi dimensioni, possono essere considerate non omogenee con le altre imprese appartenenti al gruppo delle microunità.

D.2. Le imprese minime

La legge finanziaria del 2008 ha introdotto il regime dei contribuenti minimi al fine di agevolare le imprese di piccolissime dimensioni, costituite sotto forma di ditta individuale o lavoratore autonomo.

³² I lavoratori autonomi, infatti, si differenziano dalle imprese in quanto per le prime è prevalente, nell'esercizio dell'attività economica, la capacità individuale del titolare rispetto all'organizzazione del lavoro e del capitale

Tale regime prevede l'applicazione solo di una imposta sostitutiva del 20% sul reddito, calcolato come differenza tra ricavi o compensi e spese sostenute, comprese le plusvalenze e le minusvalenze dei beni relativi all'impresa o alla professione. Il calcolo del reddito avviene applicando il principio di cassa e ciò significa che i costi sostenuti, anche relativi a spese pluriennali come i beni strumentali, possono essere sottratti dal reddito dell'esercizio in quanto pagati in questo periodo di riferimento.

Il principio di cassa, generalmente applicato per i lavoratori autonomi, viene applicato anche nel caso delle imprese minime. Per quest'ultime, ciò implica che le rimanenze di merci non hanno rilevanza fiscale in quanto il costo delle stesse deve essere attribuito al periodo in cui avviene il pagamento.

Nel caso delle imprese commerciali, il ruolo delle rimanenze delle merci è particolarmente rilevante per la definizione del margine commerciale di competenza dell'anno.

La definizione statistica di margine commerciale (variabile 12.13.0 del Regolamento SBS) è in linea con il bilancio civilistico ed è la seguente:

$$MC_{(voce_12.13.0)} = Ricavi_merci - Acquisto_merci_{(voce_13.12.0)} - Rimanenze_merci_{(voce_13.12.1)}$$

dove

$$Rimanenze_merci_{(voce_13.12.1)} = rimanenze_iniziali - rimanenze_finali$$

Abbiamo due casi:

- rimanenze delle merci > 0. Ciò significa che le rimanenze iniziali sono maggiori di quelle finali e, quindi, sono state vendute merci che si riferivano ad anni precedenti. I ricavi delle merci non si riferiscono quindi solo alle merci acquistate nell'esercizio ma anche a quelle provenienti dal magazzino che sono state vendute in questo esercizio. Il margine commerciale di competenza risulterà quindi minore.
- Rimanenze delle merci < 0. Ciò significa che le rimanenze finali di merci sono maggiori di quelle iniziali e, quindi, una parte delle merci acquistate in questo periodo non sono state vendute. Gli acquisti delle merci di competenza sono rettificati per tenere conto della quota invenduta che resta in magazzino e il costo delle merci di competenza dell'esercizio viene stornato ai periodi successivi per tenere conto della merce invenduta che giace in magazzino. Il margine commerciale di competenza risulterà maggiore in quanto una parte delle merci acquistata non fa parte dei costi relativi all'acquisto di merci di competenza del periodo corrente.

Tavola 1 - Quadro CM dei contribuenti minimi (modello Unico periodo di imposta 2008)

Variabili	Definizioni	Note
Codice fiscale Codice tipologia di impresa	Impresa; lavoratore autonomo; impresa familiare	Questa variabile è importante per selezionare le imprese familiari che sono fuori dal dominio di studio delle microunità e per distinguere le imprese dai lavoratori autonomi che hanno caratteristiche strutturali differenti.
CM1 – codice attività CM2 – totale componenti positivi	Ateco 2007 Ammontare dei ricavi dovuti alla vendita dei beni e alla prestazione dei servizi e dai corrispettivi delle cessioni di materie prime e sussidiarie, di semilavorati e di altri beni mobili (esclusi quelli strumentali) acquistati o prodotti per essere impiegati nella produzione. Vanno indicati anche i compensi in denaro e in natura derivanti dall'attività professionale o artistica, compresi quelli derivanti da attività svolte all'estero. Vanno indicate le sopravvenienze attive e le plusvalenze attive. Le fatture devono essere emesse senza l'addebito dell'IVA.	Le sopravvenienze e le plusvalenze non rientrano nella definizione statistica di ricavi. Le fatture sono emesse senza l'addebito dell'IVA
CM3 – rimanenze finali	Vanno indicate le rimanenze finali relative al periodo d'imposta precedente che vanno sottratte al valore di CM2 fino alla soglia. L'eventuale eccedenza va inserita nel rigo CM20. ³³	Questa voce corrisponde alle rimanenze iniziali.
CM4 – differenza (CM2-CM3) CM5 – totale componenti negativi (colonna 2)	Ammontare degli acquisti di materie prime, sussidiarie, semilavorati e merci pagate nel presente esercizio. Comprende le spese relative a beni ad uso promiscuo che sono deducibili al 50%. Comprende i canoni di leasing, spese per omaggi, vitto e alloggio, costo dei beni strumentali pagati nell'esercizio. Include le sopravvenienze passive e le altre spese sostenute nel periodo considerato. L'imposta sugli acquisti non può essere detratta come IVA a credito e si trasforma in un costo deducibile dal reddito.	non di interesse statistico. Potrebbe corrispondere ai costi intermedi di beni e servizi. Va escluso il costo dei beni strumentali a fecondità ripetuta (investimenti), oltre alle sopravvenienze passive. Il costo dei beni ad uso promiscuo (privato e per attività di impresa) va considerato se effettivamente corrisponde al 50%. L'IVA sugli acquisti non può essere scorporata e viene considerata un costo accessorio.
CM7 – Contributi previdenziali e assistenziali (colonna 1)	Ammontare dei contributi versati per legge, comprensivi di quelli corrisposti ai collaboratori dell'impresa familiare	Questi contributi rientrano tra i costi dei servizi nella voce relativa al costo del lavoro del personale esterno.
CM20 – eccedenza delle rimanenze di cui al rigo CM3 CM21 Eccedenza dei contributi previdenziali	CM3 - CM2 CM7 (colonna 1) - CM7 (colonna 2)	Variazione di rimanenze Si tratta della differenza tra contributi complessivi (colonna 1) e l'ammontare degli stessi che uguagliava il reddito lordo (CM6). La voce non è rilevante ai fini statistici.

Nel caso dei contribuenti minimi, adottando il principio di cassa, le rimanenze di merci non vengono considerate in quanto non hanno rilevanza fiscale e gli acquisti considerano tutte le merci che sono state pagate nell'esercizio. Pertanto, nel caso di merci invendute, il margine commerciale di una impresa che adotta il regime dei contribuenti minimi risulterà più basso di quello di una impresa che adotta un altro regime contabile basato sul criterio di competenza: la divergenza definitoria della variabile margine commerciale andrebbe misurata per evitare di introdurre una distorsione sistematica.

Le rimanenze finali di prodotti finiti non partecipano al reddito di impresa come componente positiva in quanto si tratta di prodotti per i quali non sono stati ancora incassati i ricavi.

I contribuenti minimi non devono pagare l'IVA. Ciò significa che nel caso delle vendite le fatture devono essere emesse senza l'addebito dell'IVA e nel caso degli acquisti l'IVA non può essere detratta si trasforma, quindi, in un costo accessorio.

Da un punto di vista statistico, si pone, pertanto, il problema dell'eterogeneità dei costi intermedi di una impresa che adotta il regime dei contribuenti minimi rispetto alle altre: la definizione del

³³ Le rimanenze finali di prodotti, riferite al precedente periodo di imposta (prima del regime dei minimi), venivano tassate in base al principio della competenza in quanto facevano parte del valore della produzione. Passando al regime dei minimi e applicando il criterio di cassa, si possono avere due casi. Il primo caso è relativo a $RF(t-1)=70$ e ricavi=100. I ricavi incassati per 100 al periodo t fanno riferimento a 70 dell'anno precedente che vanno decurtati. Il secondo caso è relativo a ricavi incassati per 50. In questo caso, si decurtano tutti i ricavi fino ad annullarli e nella voce CM20 si inseriscono $rimanenze=70-50=20$ che vanno decurtate nell'esercizio (t+1). Se l'impresa non ha rimanenze dell'anno precedente, si considerano solo i ricavi incassati mentre i beni prodotti e non venduti non sono rilevati nel quadro CM ma andranno rilevati solo quando si trasformeranno in vendite. La voce CM3 corrisponde alle rimanenze iniziali e la voce CM20 alla variazione delle rimanenze iniziali rispetto ai ricavi, considerando che rimanenze finali non ci possono essere in quanto gli acquisti di merci vengono imputati tutti nell'esercizio per il principio di cassa.

valore aggiunto subisce modifiche dovute al regime dei contribuenti minimi in quanto l'IVA sugli acquisti rientra tra i costi intermedi mentre è un credito nei confronti dell'erario nel caso di una impresa con altri regimi contabili.

I contribuenti minimi devono compilare gli *Studi di Settore*, escludendo solo i quadri contabili, mentre non devono compilare le dichiarazioni annuali IVA e dell'Irap.

I requisiti per rientrare nel regime dei minimi sono i seguenti:

- hanno conseguito ricavi o compensi in misura non superiore a 30.000 euro;
- non hanno effettuato cessioni all'esportazione;
- non hanno sostenuto spese per lavoro dipendente o assimilato. I dipendenti assimilati sono quelli con contratto di somministrazione lavoro, lavoro a progetto, associati in partecipazione, consulenti esterni che, ai fini statistici, rientrano nella definizione di lavoratori esterni;
- non hanno acquistato negli ultimi tre anni, anche mediante contratto di appalto e di locazione (compreso il leasing), beni strumentali di valore complessivo superiore a 15.000 euro;
- non svolgono le seguenti attività economiche: a) agricoltura e attività connesse e pesca; b) vendita di sali e tabacchi; c) commercio dei fiammiferi; d) editoria; e) gestione di servizi di telefonia pubblica; f) rivendita di documenti di trasporto pubblico e di sosta; g) intrattenimenti, giochi; h) agenzie di viaggi e turismo; i) agriturismo; l) vendite a domicilio; m) rivendita di beni usati, di oggetti d'arte, d'antiquariato o da collezione; n) agenzie di vendite all'asta di oggetti d'arte, antiquariato o da collezione.

D.3. Le imprese artigiane

La tutela e la promozione dell'artigianato è sancita dall'articolo 45 della Costituzione, che afferma che 'la legge provvede alla tutela e allo sviluppo dell'artigianato'. Il codice civile fornisce una definizione di piccolo imprenditore (art. 2083 c.c.) ma la definizione di imprenditore artigiano è regolamentata da una specifica legge (legge quadro n. 443/1985).

Il codice civile definisce piccolo imprenditore l'artigiano, il commerciante ed il coltivatore diretto del fondo che svolgono l'attività professionale organizzata prevalentemente con il lavoro proprio e dei componenti della famiglia.

Il carattere distintivo dell'artigiano è, quindi, la prevalenza del lavoro proprio e dei suoi familiari sugli altri fattori produttivi (lavoro dei dipendenti e capitale investito). La prevalenza deve essere qualitativa e funzionale e non solo quantitativa.

Secondo la legge quadro n. 443/1985, si effettua una distinzione tra imprenditore artigiano (art. 2) e impresa artigiana (art. 3).

L'imprenditore artigiano è colui che esercita personalmente, in modo continuativo e in qualità di titolare l'attività artigianale, svolgendo in misura prevalente il proprio lavoro nel processo produttivo: l'imprenditore artigiano è, quindi, una ditta individuale.

L'impresa artigiana (art. 3) è quella esercitata in forma societaria (società di persone o di capitali) nei limiti dimensionali richiesti in base al settore di appartenenza e al sistema produttivo adottato.

L'imprenditore artigiano deve svolgere attività di produzione di beni o prestazione di servizi, escludendo le attività agricole, del commercio e di somministrazione al pubblico di alimenti e bevande (bar e ristoranti).

L'impresa non deve eccedere i seguenti limiti dimensionali nelle diverse forme di attività:

- lavorazione non in serie: massimo di 18 dipendenti, compresi gli apprendisti in numero non superiore a 9; il numero può essere elevato a 22 dipendenti, a condizione che le unità di lavoro aggiuntive siano apprendisti;
- lavorazione in serie non del tutto automatizzata: massimo 9 dipendenti, compresi gli apprendisti in numero non superiore a 5; il limite può essere aumentato a 12 a condizione che le unità aggiuntive siano apprendisti;

- per le attività delle lavorazioni artistiche, tradizionali e dell'abbigliamento su misura:³⁴ massimo 32 dipendenti, compresi gli apprendisti in numero non superiore a 16; il numero di dipendenti può essere innalzato fino a 40 a condizione che le unità aggiuntive siano apprendisti;
- imprese di trasporto: un massimo di 8 dipendenti;
- imprese edili: un massimo di 10 dipendenti, compresi gli apprendisti in numero non superiore a 5; il numero massimo dei dipendenti può essere elevato a 14 a condizione che le unità aggiuntive siano apprendisti.

La scelta della dimensione aziendale varia, quindi, non solo in base al settore economico ma anche in base al processo produttivo impiegato.

La produzione in serie riguarda il processo produttivo in cui è prevalente l'uso della macchina (bene strumentale) e minimo l'intervento dell'uomo, che, comunque, è necessario per la conduzione dei macchinari e per l'effettuazione diretta di alcune fasi della lavorazione in quanto la totale meccanizzazione o automazione del processo produttivo comporta il passaggio da impresa a carattere artigianale a quella a carattere industriale.

La produzione non in serie, invece, implica che il ruolo dei beni strumentali sia marginale all'interno del processo produttivo rispetto a quello dei lavoratori. Questo processo produttivo è più elastico, nel senso che si adatta alle modifiche delle commesse e delle tipologie di materie prime da utilizzare. I macchinari possono essere utilizzati in modo differenziato e variare in base al differente risultato produttivo. Le mansioni dei dipendenti possono mutare in base alle variazioni del processo produttivo.

Le forme giuridiche dell'impresa artigiana sono quelle della ditta individuale, delle società di persone (società in nome collettivo e società in accomandita semplice), le società a responsabilità limitata, le società cooperative e la società per azioni a socio unico. La condizione è che la maggioranza dei soci partecipi con il proprio lavoro al processo produttivo e che il lavoro sia predominante sul capitale.

L'impresa iscritta all'albo degli artigiani, infine, si differenzia dalle altre per il fatto di poter spesso disporre di leggi agevolate, in termini di finanziamenti e contributi, che sono tuttavia distinte in base al territorio di appartenenza (si tratta di leggi regionali che regolamentano gli incentivi per lo sviluppo dell'artigianato), alla classe dimensionale, al settore, al sesso (incentivi per lo sviluppo dell'imprenditoria femminile), alla nuova imprenditorialità oppure all'imprenditorialità giovanile.

Queste diversificate forme di finanziamento agevolano l'avvio di una nuova attività o favoriscono la crescita di quelle già esistenti.

In conclusione, un'impresa artigiana, a parità di classe dimensionale e attività economica svolta, potrebbe avere caratteristiche economiche differenti dalle altre imprese operanti nel settore per il fatto di utilizzare un differente processo produttivo (non in serie oppure in serie non del tutto automatizzato), una diversa organizzazione del lavoro (basata sul ricorso obbligatorio ad una quota significativa di apprendisti) e per il fatto di ottenere forme di finanziamento agevolato, anche se erogate in modo differenziato a livello regionale o rivolte solo all'imprenditoria femminile, alla neoimprenditorialità, alle piccole imprese o a quelle operanti in specifici settori.

Per tale motivo, la predetta analisi descrittiva delle microunità, distinte in base agli adempimenti amministrativi, tiene separate le imprese artigiane dalle altre unità operanti nel settore.

D.4. Analisi descrittiva delle microunità per sezione di attività economica

L'analisi descrittiva delle microunità mostra che, nel 2007, esse erano oltre 2 milioni e mezzo su circa 4 milioni di imprese italiane (tavola 2).

³⁴ Le lavorazioni artistiche sono le creazioni, le produzioni e le opere di elevato valore estetico. Le lavorazioni tradizionali sono le produzioni e le attività di servizi realizzate secondo tecniche e modalità consolidate e tramandate nei costumi e nelle consuetudini a livello locale. Il settore dell'abbigliamento su misura considera le attività di confezione e lavorazione di abiti, capi accessori ed articoli di abbigliamento, realizzati su misura ed in base alle esigenze del cliente o del committente. La legge 288 del 2001 ha fornito un elenco molto dettagliato delle attività artigiane rientranti tra le lavorazioni artistiche, tradizionali e dell'abbigliamento su misura.

Le microunità sono concentrate soprattutto nel settore dei servizi e, in particolare, nelle attività immobiliari, informatica, ricerca e servizi alle imprese e nel settore commerciale, dove, la quota di fatturato è la più alta di tutte le sezioni economiche analizzate.

Il settore delle costruzioni rappresenta una quota di imprese (11,9%) e di fatturato (14,5%) maggiore rispetto all'intero comparto dell'industria in senso stretto.

Tavola 2 - Analisi delle microunità per sezione di attività economica in Ateco2002³⁵ (valori assoluti per numero di imprese e valori percentuali per composizione in termini di imprese, addetti e fatturato)

Sezione economica – Ateco 2002	numero imprese	composizione percentuale		
		imprese	addetti	fatturato
C – estrazione minerali	693	0,0	0,0	0,1
D – attività manifatturiere	165.483	6,5	6,5	9,8
E - produzione e distribuzione di energia elettrica, gas e acqua	1.507	0,1	0,1	1,5
F – costruzioni	303.425	11,9	11,9	14,5
G - commercio all'ingrosso e al dettaglio, riparazione auto, vendita carburante	727.226	28,5	28,4	34,4
H - alberghi e ristoranti	73.356	2,9	2,8	1,4
I - trasporti, magazzinaggio e comunicazioni	82.348	3,2	3,3	3,1
J - ausiliari finanziari (divisione 67)	42.548	1,7	1,7	0,7
K – attività immobiliari, informatica, ricerca, servizi alle imprese	828.259	32,5	32,7	30,2
M – istruzione	9.333	0,4	0,4	0,1
N – sanità e assistenza sociale	174.999	6,9	6,9	1,9
O - altri servizi pubblici, sociali e personali	138.124	5,4	5,4	2,3
Totale	2.547.301	100	100	100

Fonte: ASIA, anno 2007

D.5. Analisi descrittiva delle microunità in base agli adempimenti amministrativi ed al settore di attività economica di appartenenza

L'analisi delle microunità, in base agli adempimenti amministrativi, viene effettuata distintamente per il settore dell'industria, delle costruzioni e dei servizi, secondo la classificazione delle attività economiche Ateco2002.

L'obiettivo dell'analisi è quello di verificare il contributo, misurato in termini di addetti e di volume d'affari, delle microunità in base ai diversi adempimenti amministrativi, in modo da stabilire se le variabili amministrative sono in grado di soddisfare innanzitutto la domanda di informazioni statistiche richieste dal Regolamento sulle statistiche strutturali (SBS), anche se la domanda di informazioni proveniente dalla Contabilità Nazionale riveste un'enorme importanza nell'impostazione delle attuali rilevazioni statistiche annuali sui conti economici delle imprese condotte dall'Istat.

L'analisi svolta, inoltre, considera anche la classe di fatturato (minima, media e grande come descritta nel paragrafo 1 di questa appendice), che è fondamentale non solo per definire i gruppi amministrativi ma anche per rendere omogenee, in base alla struttura economica, le microunità.

L'analisi viene completata considerando anche l'appartenenza o meno dell'impresa all'albo degli artigiani, in quanto le differenti caratteristiche del sistema produttivo, le leggi agevolate e i contributi statali ricevuti da questa tipologia di impresa ne potrebbero differenziare, a parità di condizione, i risultati economici.

³⁵ Sono escluse le seguenti divisioni economiche: 1) intermediazione monetaria e finanziaria (J65); 2) assicurazioni e fondi pensione (J66); 3) amministrazione pubblica (L75); 4) attività di organizzazioni associative (O91).

Alcune considerazioni introduttive sono necessarie per una corretta interpretazione dei risultati mostrati nelle tavole seguenti:

- le microunità, costituite da un addetto medio e senza dipendenti, possono costituirsi anche sotto forma di società di persone e di capitale in quanto ci possono essere soci non lavoratori o che lavorano meno di 6 mesi l'anno. La presenza delle società di persone implica lo studio di differenti variabili amministrative rispetto a quelli delle persone fisiche (ditte individuali e lavoratori autonomi) e delle società di capitali;
- i gruppi amministrativi considerati non possono tenere conto della facoltà di optare per un altro regime contabile che spetta alle microunità;
- i contribuenti minimi sono sovrastimati in quanto non sono misurabili i seguenti parametri economici: a) non effettuano esportazioni; b) non hanno lavoratori esterni; c) non hanno effettuato, negli ultimi 3 anni, investimenti o non hanno affittato beni strumentali (anche in leasing) per un valore superiore a 15.000 euro;
- i lavoratori autonomi hanno sempre la contabilità semplificata, a prescindere dal volume d'affari, a meno che optino per quella ordinaria o rientrino nel regime dei minimi;
- le cooperative, anche di artigiani, hanno almeno tre soci lavoratori, che possono essere equiparati a subordinati, lavoratore autonomo o parasubordinati. Se non sono subordinati, rientrano tra lavoratori esterni e quindi sono esclusi dalla definizione di addetti.

Tavola 3 - Analisi delle microunità in base agli adempimenti amministrativi - Industria (Ateco2002)

GRUPPO IN BASE AD ADEMPIMENTI AMMINISTRATIVI	Artigiani				Non artigiani				Totale			
	Numero microunità	Composizione %			Numero microunità	Composizione %			Numero microunità	Composizione %		
		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato
Gruppo A – Classe di fatturato minima												
A1 – ditta individuale	58.910	35,1	34,7	3,2	7.764	4,6	4,4	0,3	66.674	39,8	39,1	3,5
A2 – lavoratore autonomo	155	0,1	0,1	0,0	76	0,0	0,0	0,0	231	0,1	0,1	0,0
A3 – SRL a socio unico	66	0,0	0,0	0,0	429	0,3	0,2	0,0	495	0,3	0,3	0,0
A4 – SRL dimensioni	190	0,1	0,1	0,0	3.656	2,2	2,1	0,2	3.846	2,3	2,2	0,2
A5 - società di persone in contabilità semplificata	2.401	1,4	1,4	0,1	2.707	1,6	1,6	0,1	5.108	3,0	3,0	0,3
A6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	115	0,1	0,1	0,0	115	0,1	0,1	0,0
Gruppo B – Classe di fatturato media												
B1 – ditta individuale in contabilità semplificata	55.668	33,2	33,8	12,8	6.441	3,8	3,8	1,6	62.109	37,0	37,6	14,4
B2 – lavoratore autonomo in contabilità semplificata	124	0,1	0,1	0,0	74	0,0	0,0	0,0	198	0,1	0,1	0,1
B3 – SRL a socio unico	236	0,1	0,1	0,1	1.237	0,7	0,7	0,6	1.473	0,9	0,9	0,7
B4 – SRL	1.003	0,6	0,6	0,6	10.050	6,0	6,0	4,7	11.053	6,6	6,6	5,3
B5 - società di persone in contabilità semplificata	6.474	3,9	3,9	2,0	4.640	2,8	2,8	1,6	11.114	6,6	6,7	3,6
B6 – SPA (anche a socio unico)	1	0,0	0,0	0,0	311	0,2	0,2	0,2	312	0,2	0,2	0,2
Gruppo C – Classe di fatturato grande												
C1 – ditta individuale in contabilità ordinaria	217	0,1	0,1	0,7	118	0,1	0,1	0,5	335	0,2	0,2	1,1
C2 – lavoratore autonomo in contabilità semplificata	1	0,0	0,0	0,0	0	0,0	0,0	0,0	1	0,0	0,0	0,0
C3 – SRL a socio unico	28	0,0	0,0	0,1	328	0,2	0,2	6,9	356	0,2	0,2	7,0
C4 – SRL	126	0,1	0,1	0,4	1.674	1,0	1,0	15,0	1.800	1,1	1,1	15,5
C5 - società di persone in contabilità ordinaria	94	0,1	0,1	0,3	181	0,1	0,1	0,9	275	0,2	0,2	1,2
C6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	210	0,1	0,1	7,0	210	0,1	0,1	7,0
Gruppo D – Altre forme giuridiche												
D1 - cooperative	86	0,1	0,0	0,0	909	0,5	0,5	1,0	995	0,6	0,6	1,0
D2 – consorzi	15	0,0	0,0	0,0	484	0,3	0,3	3,7	499	0,3	0,3	3,8
D3 - imprese estere operanti in Italia	0	0,0	0,0	0,0	407	0,2	0,2	34,3	407	0,2	0,2	34,3
D4 – NC (altre forme giuridiche)	0	0,0	0,0	0,0	77	0,0	0,0	0,9	77	0,0	0,0	0,9
Totale	125.795	75,0	75,2	20,5	41.888	25,0	24,8	79,5	167.683	100	100	100

Fonte: ASIA, anno 2007

La tavola 3 mostra che, nel 2007, le microunità operanti nel settore dell'industria, secondo la classificazione Ateco2002, erano circa 167.000 e, di queste, il 75,0% erano a carattere artigianale. Le microunità artigiane, tuttavia, coprivano solo il 20,5% del fatturato complessivo del settore riferito alle unità in esame mentre le circa 400 microunità estere operanti in Italia, a carattere non artigianale, coprivano il 34,3% del fatturato del settore industriale. Tali microunità, appartenendo a gruppi esteri, sono in realtà di grandi dimensioni sotto l'aspetto economico, in quanto, probabilmente, esternalizzano alle altre imprese del gruppo molte fasi del processo produttivo.

All'interno della classe di fatturato minima, le ditte individuali artigiane sono il 35,1% delle microunità industriali anche se non si può stabilire, per carenza di informazioni, quante di esse rientrano nel regime dei minimi. La possibilità di integrare le informazioni relative a questa classe dimensionale con quelle artigiane provenienti dalle società di persone in contabilità semplificata o dalle società di capitali (con un *set* molto più completo di variabili amministrative di interesse statistico), è molto limitata.

I lavoratori autonomi con classe di fatturato minima, invece, non possono essere considerati omogenei con le microunità in quanto per questa tipologia di attività è prevalente la creatività e l'intelletto dell'individuo rispetto all'organizzazione del lavoro dei dipendenti o collaboratori.

Nel gruppo delle microunità artigiane con classe di fatturato medio, le ditte individuali in contabilità semplificata coprono il 33,8% degli addetti ed il 12,8% del fatturato complessivo dell'industria, ma non è trascurabile il contributo delle società di persone (3,9% in termini di occupazione e 2,0% in termini di fatturato).

È importante sottolineare come le microunità industriali artigiane siano rappresentate soprattutto da ditte individuali (cosiddetti imprenditori artigiani).

La presenza di un'elevata quota di ditte individuali in contabilità semplificata nella classe di fatturato medio implica la possibilità di acquisire maggiori informazioni di interesse statistico dalla fonte amministrativa, anche se la quota degli imprenditori artigiani minimi non è trascurabile.

La presenza dell'informazione sul quadro CM del modello unico (contribuenti minimi riferito al periodo di imposta 2008) è quindi fondamentale per stabilire la quota di addetti e di fatturato che gli imprenditori artigiani minimi coprono nell'industria e, quindi, se la presenza di ditte individuali con regime di contabilità semplificata con classe di fatturato minimo possa sopperire alla carenza di variabili amministrative di interesse statistico presenti nel regime forfetario.

Con riferimento alle microunità a carattere non artigianale esse, pur costituendo solo il 25,0% delle microunità industriali, hanno una maggiore importanza economica rappresentando il 79,5% del fatturato complessivo delle unità prese in esame, concentrato prevalentemente nelle microunità estere operanti in Italia (34,3% di fatturato).

Per questa tipologia di microunità, il contributo delle ditte individuali minime può essere considerato trascurabile (circa il 4% di occupazione su un totale del 24% dell'occupazione a carattere non artigianale).

Le variabili amministrative di interesse statistico, quindi, sono rilevanti per questa tipologia di microunità.

In conclusione, il problema dell'introduzione del regime dei contribuenti minimi potrebbe essere rilevante soprattutto per le microunità artigiane, anche se non costituisce un dominio di stima del Regolamento SBS.

Tavola 4 - Analisi delle microunità in base agli adempimenti amministrativi - Costruzioni (Ateco2002)

GRUPPO IN BASE AD ADEMPIMENTI AMMINISTRATIVI	Artigiani				Non artigiani				Totale			
	Numero microunità	Composizione %			Numero microunità	Composizione %			Numero microunità	Composizione %		
		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato
Gruppo A – Classe di fatturato minima												
A1 – ditta individuale	138.713	45,7	45,0	6,5	8.123	2,7	2,6	0,3	146.836	48,4	47,6	6,8
A2 – lavoratore autonomo	149	0,0	0,0	0,0	156	0,1	0,1	0,0	305	0,1	0,1	0,0
A3 – SRL a socio unico	38	0,0	0,0	0,0	619	0,2	0,2	0,0	657	0,2	0,2	0,0
A4 – SRL dimensioni	151	0,0	0,0	0,0	5.719	1,9	1,9	0,2	5.870	1,9	1,9	0,2
A5 - società di persone in contabilità semplificata	1.382	0,5	0,4	0,1	2.152	0,7	0,7	0,1	3.534	1,2	1,1	0,1
A6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	67	0,0	0,0	0,0	67	0,0	0,0	0,0
Gruppo B – Classe di fatturato media												
B1 – ditta individuale in contabilità semplificata	98.361	32,4	33,0	17,0	5.156	1,7	1,7	1,3	103.517	34,1	34,7	18,3
B2 – lavoratore autonomo in contabilità semplificata	84	0,0	0,0	0,0	83	0,0	0,0	0,0	167	0,1	0,1	0,0
B3 – SRL a socio unico	213	0,1	0,1	0,1	1.606	0,5	0,5	0,8	1.819	0,6	0,6	0,9
B4 – SRL	1.012	0,3	0,3	0,5	16.591	5,5	5,5	8,1	17.603	5,8	5,9	8,5
B5 - società di persone in contabilità semplificata	3.734	1,2	1,3	1,0	3.748	1,2	1,3	1,4	7.482	2,5	2,5	2,4
B6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	181	0,1	0,1	0,1	181	0,1	0,1	0,1
Gruppo C – Classe di fatturato grande												
C1 – ditta individuale in contabilità ordinaria	399	0,1	0,1	0,9	210	0,1	0,1	0,6	609	0,2	0,2	1,5
C2 – lavoratore autonomo in contabilità semplificata	2	0,0	0,0	0,0	5	0,0	0,0	0,0	7	0,0	0,0	0,0
C3 – SRL a socio unico	48	0,0	0,0	0,1	776	0,3	0,3	5,3	824	0,3	0,3	5,4
C4 – SRL	164	0,1	0,1	0,4	8.043	2,7	2,7	34,7	8.207	2,7	2,8	35,1
C5 - società di persone in contabilità ordinaria	93	0,0	0,0	0,2	618	0,2	0,2	2,2	711	0,2	0,2	2,4
C6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	196	0,1	0,1	2,9	196	0,1	0,1	2,9
Gruppo D – Altre forme giuridiche												
D1 - cooperative	301	0,1	0,1	0,2	2.413	0,8	0,8	4,7	2.714	0,9	0,9	4,9
D2 – consorzi	130	0,0	0,0	0,2	1.819	0,6	0,6	8,6	1.949	0,6	0,6	8,8
D3 - imprese estere operanti in Italia	0	0,0	0,0	0,0	120	0,0	0,0	1,3	120	0,0	0,0	1,3
D4 – NC (altre forme giuridiche)	0	0,0	0,0	0,0	50	0,0	0,0	0,4	50	0,0	0,0	0,4
Totale	244.974	80,7	80,6	27,2	58.451	19,3	19,4	72,8	303.425	100	100	100

Fonte: ASIA, anno 2007

Le microunità delle costruzioni erano, nel 2007, oltre 300.000 (tavola 4) e rappresentavano l'11,9% delle microunità totali ed il 14,5% del fatturato delle unità prese in esame.

Le microunità a carattere artigianale (tavola 3) erano l'80,7% del totale delle microunità di costruzioni ma rappresentavano solo il 27,2% del fatturato.

Con riferimento alle microunità artigianali minime in termini di fatturato, esse si costituiscono prevalentemente sotto forma di ditta individuale (45,7% delle imprese e solo il 6,5% del fatturato).

La mancanza di informazioni sui contribuenti minimi, presenti nella popolazione, impedisce comunque di stabilire quanto le variabili amministrative riescano a stimare le variabili statistiche del Regolamento SBS.

Le ditte individuali artigiane in contabilità semplificata con classe di fatturato medio rappresentano il 32,4% delle microunità e coprono il 17,0% del fatturato ed hanno un insieme più completo di variabili amministrative di interesse statistico.

Il problema, quindi, delle microunità artigiane del settore delle costruzioni è quello di verificare se è realistico ipotizzare che la dimensione aziendale, misurata in termini di fatturato, può essere trascurata nell'integrazione delle ditte individuali minime con quelle medie.

Con riferimento alle microunità di costruzioni a carattere non artigianale, esse, pur essendo solo il 19,3% delle microunità, coprono il 72,8% di fatturato e sono costituite prevalentemente sotto forma di società di capitali, di cooperative e consorzi.

Questo significa che le microunità non artigianali hanno un set completo di variabili amministrative di interesse statistico, omogeneo, in termini di definizioni, al Regolamento SBS.

Anche per questo settore economico, come per l'industria, bisogna verificare quanti sono effettivamente i contribuenti minimi artigiani (quadro CM del modello unico) e quale è il contributo delle microunità alla stima del dominio della classe di addetti 1-9, richiesto, a livello di gruppo economico, dal Regolamento SBS.

Tavola 5 - Analisi delle microunità in base agli adempimenti amministrativi - Servizi (Ateco2002)

GRUPPO IN BASE AD ADEMPIMENTI AMMINISTRATIVI	Artigiani				Non artigiani				Totale			
	Numero microunità	Composizione %			Numero microunità	Composizione %			Numero microunità	Composizione %		
		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato		Micro unità	Addetti	Fatturato
Gruppo A – Classe di fatturato minima												
A1 – ditta individuale	138.262	6,7	6,7	1,1	207.608	10,0	9,7	1,6	345.870	16,7	16,4	2,7
A2 – lavoratore autonomo	4.612	0,2	0,2	0,0	572.788	27,6	27,3	3,9	577.400	27,8	27,5	4,0
A3 – SRL a socio unico	44	0,0	0,0	0,0	5.173	0,2	0,2	0,0	5.217	0,3	0,2	0,0
A4 – SRL dimensioni	195	0,0	0,0	0,0	38.261	1,8	1,8	0,3	38.456	1,9	1,8	0,3
A5 - società di persone in contabilità semplificata	4.030	0,2	0,2	0,0	46.945	2,3	2,2	0,4	50.975	2,5	2,4	0,4
A6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	374	0,0	0,0	0,0	374	0,0	0,0	0,0
Gruppo B – Classe di fatturato media												
B1 – ditta individuale in contabilità semplificata	89.843	4,3	4,4	3,1	274.870	13,2	13,4	10,6	364.713	17,6	17,8	13,7
B2 – lavoratore autonomo in contabilità semplificata	2.955	0,1	0,1	0,1	411.720	19,8	20,2	14,3	414.675	20,0	20,3	14,4
B3 – SRL a socio unico	172	0,0	0,0	0,0	11.934	0,6	0,6	0,7	12.106	0,6	0,6	0,7
B4 – SRL	809	0,0	0,0	0,1	92.121	4,4	4,5	5,3	92.930	4,5	4,5	5,4
B5 - società di persone in contabilità semplificata	7.629	0,4	0,4	0,3	86.616	4,2	4,2	4,0	94.245	4,5	4,6	4,3
B6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	1.571	0,1	0,1	0,1	1.571	0,1	0,1	0,1
Gruppo C – Classe di fatturato grande												
C1 – ditta individuale in contabilità ordinaria	879	0,0	0,0	0,3	12.895	0,6	0,6	4,7	13.774	0,7	0,7	4,9
C2 – lavoratore autonomo in contabilità semplificata	47	0,0	0,0	0,0	5.425	0,3	0,3	1,6	5.472	0,3	0,3	1,6
C3 – SRL a socio unico	20	0,0	0,0	0,0	4.835	0,2	0,2	7,4	4.855	0,2	0,2	7,4
C4 – SRL	150	0,0	0,0	0,1	29.433	1,4	1,4	19,1	29.583	1,4	1,4	19,2
C5 - società di persone in contabilità ordinaria	216	0,0	0,0	0,1	6.896	0,3	0,3	2,8	7.112	0,3	0,3	2,8
C6 – SPA (anche a socio unico)	0	0,0	0,0	0,0	1.920	0,1	0,1	5,1	1.920	0,1	0,1	5,1
Gruppo D – Altre forme giuridiche												
D1 - cooperative	158	0,0	0,0	0,0	7.602	0,4	0,4	0,7	7.760	0,4	0,4	0,8
D2 – consorzi	220	0,0	0,0	0,0	4.622	0,2	0,2	2,4	4.842	0,2	0,2	2,5
D3 - imprese estere operanti in Italia	0	0,0	0,0	0,0	1.737	0,1	0,1	8,4	1.737	0,1	0,1	8,4
D4 – NC (altre forme giuridiche)	0	0,0	0,0	0,0	606	0,0	0,0	1,4	606	0,0	0,0	1,4
Totale	250.241	12,1	12,1	5,2	1.825.952	87,9	87,9	94,8	2.076.193	100	100	100

Fonte: ASIA, anno 2007

Le microunità di servizi (tavola 5) erano, nel 2007, oltre 2 milioni e rappresentavano l'81,5% del complesso delle microunità. Esse operavano prevalentemente nel settore delle attività immobiliari, informatica, ricerca e servizi alle imprese e nel settore commerciale.

Le microunità di servizi a carattere artigianale (tavola 4) sono solo il 12,1% in quanto i settori commerciali, dei pubblici esercizi (bar e ristoranti) e dei servizi pubblici non possono avere imprese artigiane. Nell'ambito di questa tipologia di microunità, le ditte individuali minime e quelle in contabilità semplificata medie coprono la quasi totalità degli occupati e del fatturato.

L'uso dei dati amministrativi ai fini statistici dipende, quindi, da quanti sono i contribuenti minimi artigiani e dalla possibilità di integrare i dati provenienti dalle imprese medie con quelle minime, costituite sotto forma di ditta individuale.

Con riferimento alle microunità di servizi a carattere non artigianale, le ditte individuali minime sono solo il 10,0% e, quindi, le variabili amministrative, presenti per gli altri gruppi considerati, possono essere rilevanti per le stime statistiche richieste.

I lavoratori autonomi di dimensioni minime, invece, potrebbero essere stimati solo con quelli medi e non può esserne trascurata la quota rappresentata, in termini di microunità e di addetti.

Per gli artisti e professionisti di minori dimensioni andrebbe verificata la quota che compila il quadro CM del modello unico.

Conclusioni

La classificazione delle imprese con un addetto medio e senza dipendenti in base agli adempimenti amministrativi ha mostrato come, per i settori dell'industria e delle costruzioni, le imprese artigiane, costituite sotto forma di ditta individuale con classe di fatturato minima abbiano un'importante presenza economica ed è, quindi, fondamentale valutare quante di esse rientrino nel regime dei contribuenti minimi. L'analisi empirica del quadro CM dei contribuenti minimi, presenti nel modello Unico dal periodo di imposta 2008, è fondamentale per valutare se le fonti amministrative sono in grado di soddisfare la domanda di informazioni statistiche, anche se il dominio di studio richiesto dal Regolamento SBS per i settori dell'industria e delle costruzioni è rappresentato dal gruppo di attività economica e dalle imprese della classe 1-9 addetti.

Nel settore dei servizi, invece, il dominio di studio per il regolamento SBS è relativo alle microunità a livello di gruppo di attività economica. Il settore dei servizi, a livello aggregato, è caratterizzato da una bassa quota di ditte individuali minime e questo potrebbe incoraggiare un maggior uso di variabili amministrative, anche se la valutazione andrebbe effettuata a livello di gruppo economico, considerando il contributo dei contribuenti minimi.

Infine, si segnala che le imprese familiari possono adottare il regime dei contribuenti minimi e queste imprese rientrano in una classe di addetti diversa rispetto a quella delle microunità.

Informazioni per gli autori

La collana è aperta ad autori dell'Istat e del Sistema statistico nazionale, e ad altri studiosi che abbiano partecipato ad attività promosse dal Sistan (convegni, seminari, gruppi di lavoro, ecc.). Da gennaio 2011 essa sostituirà Documenti Istat e Contributi Istat.

Coloro che desiderano pubblicare sulla nuova collana dovranno sottoporre il proprio contributo alla redazione degli Istat Working Papers inviandolo per posta elettronica all'indirizzo iwp@istat.it. Il saggio deve essere redatto seguendo gli standard editoriali previsti, corredato di un sommario in italiano e in inglese; deve, altresì, essere accompagnato da una dichiarazione di paternità dell'opera. Per la stesura del testo occorre seguire le indicazioni presenti nel foglio di stile, con le citazioni e i riferimenti bibliografici redatti secondo il protocollo internazionale 'Autore-Data' del *Chicago Manual of Style*.

Per gli autori Istat, la sottomissione dei lavori deve essere accompagnata da una mail del proprio dirigente di Servizio/Struttura, che ne assicura la presa visione. Per gli autori degli altri enti del Sistan la trasmissione avviene attraverso il responsabile dell'ufficio di statistica, che ne prende visione. Per tutti gli altri autori, esterni all'Istat e al Sistan, non è necessaria alcuna presa visione. Tutti i lavori saranno sottoposti al Comitato di redazione, che valuterà la significatività del lavoro per il progresso dell'attività statistica istituzionale. La pubblicazione sarà disponibile su formato digitale e sarà consultabile on line.

Gli articoli pubblicati impegnano esclusivamente gli autori, le opinioni espresse non implicano alcuna responsabilità da parte dell'Istat. Si autorizza la riproduzione a fini non commerciali e con citazione della fonte.