

**RELAZIONE SULL'ECONOMIA NON OSSERVATA
E SULL'EVASIONE FISCALE E CONTRIBUTIVA
ANNO 2023**

(art. 10-bis.1 c. 3 Legge 31 dicembre 2009, n. 196)

APPENDICI METODOLOGICHE

INDICE

INDICE	1
I. METODOLOGIE DI STIMA DEI TAX GAP DEL CAPITOLO III	2
I.1 Metodologia per la stima del <i>gap</i> delle imposte IRES e IRPEF da lavoro autonomo e imprese e delle relative addizionali.....	2
I.2 Metodologia per la stima del <i>tax gap</i> IVA.....	10
I.3 Stima del <i>gap</i> IVA: metodologie a confronto	18
I.4 Metodologia per la stima del <i>tax gap</i> IRAP	21
I.5 Metodologia per la stima del <i>gap</i> relativo ai redditi da locazione	29
I.6 Metodologia per la stima del <i>tax gap</i> del canone RAI	30
I.7 Metodologia per la stima dell'evasione contributiva.....	32
I.8 Metodologia per la stima del <i>tax gap</i> delle imposte immobiliari	34
II. REVISIONE DELLE STIME TRA IL 2022 E IL 2023 DI ALCUNI TAX GAP	38
II.1 Revisioni del <i>gap</i> dell'IRPEF da lavoro autonomo e da impresa e dell'IRES.....	38
II.2 Revisioni del <i>gap</i> dell'IVA	39
II.3 Revisioni del <i>gap</i> dell'IRAP	40
II.4 Revisioni del <i>gap</i> delle locazioni	41
III. ALTRE METODOLOGIE	42
III.1 La stima <i>bottom-up</i> del <i>gap</i> IVA	42
III.2 La stima <i>bottom up</i> del <i>gap</i> IRAP	46
III.3 Valutazione <i>ex-post</i> dell'estensione della soglia di ricavi nel nuovo regime forfetario	50
III.4 Un modello a classi latenti per l'analisi delle frodi IVA.....	57
III.5 Derivazione del modello econometrico per l'eterogeneità degli effetti della fatturazione elettronica	60
IV. RIFERIMENTI BIBLIOGRAFICI	63

I. METODOLOGIE DI STIMA DEI TAX GAP DEL CAPITOLO III

I.1 METODOLOGIA PER LA STIMA DEL GAP DELLE IMPOSTE IRES E IRPEF DA LAVORO AUTONOMO E IMPRESE E DELLE RELATIVE ADDIZIONALI

Introduzione

In termini generali, le metodologie top-down impiegate per la stima del *tax gap* prevedono l'individuazione della base imponibile non dichiarata (BIND) mediante il confronto tra base imponibile potenziale (BIT), desumibile dai dati di contabilità nazionale (CN), resi omogenei con le classificazioni e le definizioni fiscali, e la base imponibile dichiarata (BID), contenuta negli archivi fiscali (i dati statistici sulle dichiarazioni fiscali sono pubblicati a cura del Dipartimento delle finanze).

Successivamente, si stima l'aliquota fiscale e si quantifica il *gap* d'imposta (TG).

A differenza dell'IVA, per le metodologie top-down applicate al calcolo del *gap* sulle imposte dirette non è disponibile una vasta letteratura in ambito internazionale¹. Le difficoltà di calcolo sono, in generale, riconducibili a due aspetti: in primo luogo il calcolo della base imponibile è più complesso, in quanto deve considerare una molteplicità di aspetti, molti dei quali hanno rilevanza ai fini fiscali ma non sono considerati negli indicatori macro; in secondo luogo, i sistemi di tassazione basati sulla progressività per scaglioni presuppongono la conoscenza anche della distribuzione dei sottostanti redditi e non solo del loro ammontare complessivo.

Nell'ambito delle imposte dirette, il termine di paragone naturale per le imposte sui redditi si rinviene nei Conti per Settore Istituzionale, che pongono in evidenza la ripartizione del reddito nazionale tra le diverse tipologie di soggetti. Tuttavia, tali dati presentano delle limitazioni in termini di disponibilità e comparabilità di dati compatibili con le finalità di stima (ad esempio, la disaggregazione per area geografica è disponibile solo per una parte della platea).

È inoltre opportuno sottolineare che, secondo la normativa fiscale italiana (come del resto in altri contesti), l'imposta sul reddito dovuta da ciascun contribuente varia a seconda della forma giuridica del soggetto considerato: l'IRES compete alle società di capitali ed agli enti non commerciali, mentre l'IRPEF è corrisposta dalle persone fisiche e, per tramite dei soci, dalle società di persone. Più precisamente, ai fini IRPEF, la base imponibile include la somma di tutti i redditi (con l'esclusione di quelli soggetti a tassazione sostitutiva o separata per particolari soggetti o per specifiche tipologie di reddito) al netto delle deduzioni spettanti. Tali redditi possono essere raggruppati in tre grandi categorie:

- redditi da lavoro dipendente pensioni e redditi assimilati (IRPEFdip);
- redditi da lavoro autonomo, impresa e partecipazione in società di persone (IRPEFaut);
- altri redditi (componente residuale che ammonta a circa il 5% del totale).

La Tabella I.1.1 evidenzia un primo quadro di raccordo macro tra la forma giuridica, rilevante ai fini fiscali, e la classificazione per settori istituzionali utilizzata nei conti nazionali.

Nella prima riga si trova il settore istituzionale delle società finanziarie, a cui possono appartenere sia le società di persone (SP) sia le società di capitali (SC). Questo settore è bene identificabile negli aggregati, ma come si spiegherà più diffusamente in seguito, per il momento è stato escluso dal calcolo.

¹ Tra i riferimenti più recenti si vedano Cabral et al. (2018), Gemmill e Hasseldine (2014), Ueda (2018).

Nel settore istituzionale delle società non finanziarie si collocano le società di persone e quelle di capitale non comprese nella prima categoria. In questo caso però non esiste una perfetta sovrapposizione tra le definizioni statistiche e quelle fiscali, poiché le società di persone con meno di cinque dipendenti in contabilità nazionale vengono inserite nelle famiglie produttrici, mentre fiscalmente sono società di persone. Per rendere omogeneo il calcolo è, quindi, necessario avere una rappresentazione quantitativa esauriente di questo segmento in modo di ricomporre il quadro di insieme delle società di persone assoggettate all'IRPEF.

Il settore istituzionale delle Amministrazioni Pubbliche si colloca, fiscalmente, nei due segmenti degli Enti non commerciali (ENC) e delle Amministrazioni Pubbliche (AAPP).

TABELLA I.1.1 CONFRONTO TRA SETTORI ISTITUZIONALI (CN) E SOGGETTI GIURIDICI						
		Soggetti giuridici a fini fiscali				
		PF	SP	SC	ENC	[AAPP]
Settori istituzionali (Contabilità nazionale)	Società finanziarie		[Escluse dall'analisi]	[Escluse dall'analisi]		
	Società non finanziarie		IRPEFaut	IRES		
	[Amministrazioni pubbliche]				[Escluse dall'analisi]	[Escluse dall'analisi]
	[Famiglie consumatrici]	[IRPEFdip]				
	Famiglie produttrici	IRPEFaut + "Minimi"	IRPEFaut			
	Istituzioni sociali private (ISP)				[IRES]	

In Contabilità nazionale, il settore istituzionale delle famiglie si divide in due:

- consumatrici, la cui fonte di reddito è rappresentata da: lavoro dipendente, pensioni e redditi assimilati;
- produttrici che percepiscono redditi da: lavoro autonomo, impresa e partecipazione in società di persone.

Le società di persone incluse nelle famiglie produttrici, come già menzionato, sono solo quelle con meno di cinque addetti. Nello stesso settore istituzionale si collocano anche i contribuenti che applicano regimi agevolati, sinteticamente indicati come "minimi", che saranno esclusi dal calcolo del *gap* in quanto assoggettati ad una tassazione sostitutiva.

La metodologia proposta in questa sede non considera il *gap* relativo ai redditi da lavoro dipendente (IRPEFdip). Inoltre vengono escluse dall'analisi le amministrazioni pubbliche, per le quali si ipotizza assenza di evasione, e le società finanziarie, in virtù delle peculiari modalità di determinazione del reddito che le caratterizzano. La procedura viene applicata a dati nazionali ripartiti per settori di attività economica.

Il calcolo del Risultato lordo di gestione come proxy della base imponibile IRES/IRPEF

Si assume come platea di riferimento dei contribuenti assoggettati all'IRES e all'IRPEF da lavoro autonomo e impresa, da ora in avanti indicata solo con IRPEF, l'insieme dei soggetti tenuti al pagamento dell'imposta regionale sulle attività produttive (IRAP). Tale scelta è dettata dalla finalità di sfruttare la disponibilità di informazioni sia per quanto riguarda la composizione della base imponibile sia per la disaggregazione per forma giuridica, area geografica e settore di attività economica.

Risultano, pertanto, esclusi dall'analisi i contribuenti che sono fuori dal campo di applicazione dell'IRAP, rappresentati dalle seguenti categorie: oltre ai già citati contribuenti "minimi", rientrano in tale casistica i professionisti privi di autonoma organizzazione, parte del settore agricolo e il settore dei servizi domestici².

² Per ulteriori dettagli si rimanda all'appendice metodologica relativa al calcolo del *gap* IRAP.

La platea sotto osservazione è ulteriormente limitata da problematiche relative alla metodologia di calcolo degli aggregati necessari alla determinazione del *gap*. In particolare si escludono completamente: il settore agricolo (incluso solo nella quota del *gap* IRES fino all'anno d'imposta 2014), la Pubblica Amministrazione e le società finanziarie.

Per la restante platea si dispone di due insieme di dati economici, omogenei, sia per quanto riguarda il campo di osservazione, sia per quel che concerne le definizioni degli aggregati, desunti uno dalle fonti fiscali e l'altro dai dati di contabilità nazionale.

Tali dati hanno consentito di mettere a confronto il valore aggiunto di contabilità nazionale, ovvero la produzione lorda IRAP potenziale, con la produzione lorda desunta dalle dichiarazioni, al fine di calcolare il *gap* IRAP. Dal momento che la base IRAP contiene al suo interno gli elementi reddituali che concorrono a determinare le basi IRPEF e IRES, il lavoro che segue consiste nell'isolare tali componenti in modo da verificare il loro contributo specifico alla determinazione del *gap* complessivo del sistema fiscale.

Il passo successivo che occorre compiere consiste, pertanto, nel derivare da entrambe le produzioni lorde un aggregato che approssimi il più possibile le basi imponibili IRES ed IRPEF. Tale aggregato è stato individuato, in prima istanza, nel risultato lordo di gestione (RLG), le cui modalità di calcolo sono illustrate nelle sezioni seguenti.

Risultato lordo di gestione dichiarato

Nel prospetto ³ è illustrato il quadro di raccordo tra le grandezze utilizzate nel calcolo della produzione lorda dichiarata IRAP e il reddito o la perdita di esercizio rilevante ai fini fiscali. Dal prospetto si ricava che se si sottrae alla produzione lorda il costo del lavoro si ottiene il risultato lordo di gestione (RLGd), che, concettualmente, rappresenta il punto di partenza per determinare il reddito o la perdita rilevante ai fini fiscali (RFd).

Per ottenere quest'ultimo è necessario sottrarre al risultato lordo di gestione una serie di voci, elencate in dettaglio nel prospetto 2, desumibili o da fonti fiscali o dal bilancio civilistico. L'insieme delle componenti che distinguono RLGd da RFd viene indicato con Δd .

Tale elemento, Δd , può essere calcolato per la componente dichiarata, ma, allo stato attuale delle conoscenze, risulta non stimabile per la componente non dichiarata. Pertanto, l'aggregato di riferimento che viene individuato in prima approssimazione come base del calcolo del *gap* IRPEF e IRES è il risultato lordo di gestione (RLG). Tale grandezza è calcolata, per la parte dichiarata, integrando i dati delle dichiarazioni IRAP, resi omogenei per quanto riguarda le definizioni e il campo di osservazione con la contabilità nazionale, con quelli relativi al costo del lavoro dipendente desunto dalle dichiarazioni 770 o dalle Certificazioni Uniche.

³ La relazione tra le grandezze illustrate fa riferimento alle voci di bilancio di un'impresa in contabilità ordinaria.

Prospetto 1
CALCOLO DEL RISULTATO LORDO DI GESTIONE DICHIARATO E RELAZIONE CON LA BASE IMPONIBILE

Componenti positivi: (A) Valore della produzione

Di cui: Ricavi delle vendite e delle prestazioni;

Variazioni delle rimanenze di prodotti in corso di lavorazione, semilavorati e finiti;

Variazione dei lavori in corso su ordinazione;

Incrementi di immobilizzazioni per lavori interni;

Altri ricavi e proventi.

(-) Componenti negativi: (B) Costi della produzione deducibili a fini IRAP, esclusi gli ammortamenti

Di cui: Acquisti di materie prime, sussidiarie, di consumo e merci;

Acquisti di Servizi;

Godimento beni di terzi;

Variazioni delle rimanenze di materie prime, sussidiarie di consumo e merci;

Oneri diversi di gestione.

(=) **Valore Aggiunto al Lordo degli ammortamenti = Produzione Lorda IRAP (PLd)**

(-) Costo del lavoro (Wd)

(=) **Risultato lordo di gestione (RLGd)**

(-) Ammortamenti

(=) Risultato netto di gestione dichiarato (RNGd)

(-) Ulteriori elementi contabili, non rilevabili da fonti fiscali

Di cui⁴: B10c: Altre svalutazioni delle immobilizzazioni;

B10d: Svalutazione dei crediti compresi nell'attivo circolante e delle disponibilità liquide;

B12: Accantonamenti per rischi;

B13: Altri accantonamenti;

C: Proventi e oneri finanziari;

D: rettifiche di valore delle attività finanziarie;

E: Proventi e oneri straordinari.

(+/-) Imposte sul reddito dell'esercizio (IRES/IRPEF; IRAP)

(=) **Risultato di esercizio (Utile o perdita) dichiarato**

(+/-) Variazioni in aumento e in diminuzione

(=) **Reddito (o perdita) rilevante a fini fiscali dichiarato (RFd)**

Risultato lordo di gestione potenziale

Il risultato lordo di gestione dichiarato (RLGd) deve essere messo a confronto con il Risultato lordo di gestione potenziale (RLGp) desunto dai dati della contabilità nazionale.

Per calcolare RLGp si assume come platea di riferimento quella degli operatori economici definita nella parte introduttiva del presente paragrafo e sottostante il calcolo di RLGd.

Il prospetto 2 illustra la procedura per il calcolo di RLGp a partire dal valore aggiunto di contabilità nazionale.

In primo luogo si operano le correzioni necessarie per rendere l'aggregato statistico coerente, nelle definizioni, all'aggregato fiscale della Produzione lorda, così come è definito nel prospetto 1. In questo caso, però, si tratta della produzione lorda potenziale, PLP, in quanto l'aggregato statistico di partenza include la componente di produzione occultata al fisco.

⁴ Le voci sono tratte dal bilancio civilistico.

Prospetto 2

CALCOLO DEL RISULTATO LORDO DI GESTIONE POTENZIALE (RLGp)

Valore aggiunto al costo dei fattori (VACF)

(+/-) Correzioni per adeguare le definizioni statistiche a quelle fiscali

* esclusione di: economia illegale, mance, ricerca e sviluppo, fitti, diritti d'autore, autoproduzioni;

* correzioni per divergenza di contabilizzazione sui SIFIM e i premi assicurativi;

* altre correzioni che non incidono sull'ammontare complessivo ma solo sulla ripartizione settoriale

(=) **Produzione lorda potenziale (PLP)**

(-) Costo del lavoro (Wt): regolare (Wd) + irregolare (Wnd; fonte ISTAT)

(=) **Risultato Lordo di Gestione potenziale (RLGp)**

Da PLP si sottrae il costo del lavoro, sia con riferimento alla sua componente regolare, Wd ovvero quello che è correttamente inserito nelle scritture contabili delle imprese, sia relativo alla componente irregolare, Wnd.

L'esclusione di quest'ultima componente è molto importante al fine di evitare duplicazioni con la stima delle componenti del *gap* relative all'IRPEF da lavoro dipendente irregolare e dei relativi contributi sociali.

Come risultato si ottiene RLGp che è omogeneo con RLGd sia riguardo alla platea di riferimento sia per la definizione dell'aggregato.

Determinazione della base imponibile non dichiarata (BIND)

La differenza tra RLGp e RLGd rappresenta il risultato lordo di gestione non dichiarato (RLGnd)

$$RLGnd = RLGp - RLGd \quad [1]$$

RLGnd si differenzia dalla base imponibile non dichiarata (BIND) in ragione di fattori analoghi a quelli che nel prospetto 2.1, con riferimento al calcolo della componente dichiarata, sono stati indentificati sinteticamente con Δd , e che vengono ora contrassegnati con Δnd .

$$BIND_{IRES+IRPEF} = RLGnd + \Delta nd \quad [2]$$

Come riportato in precedenza, allo stato attuale delle conoscenze, Δnd non è calcolabile, in virtù delle difficoltà insite nel ricostruire le sottostanti singole voci tramite gli aggregati statistici. Al fine di pervenire ad una stima del *gap* nell'imposta è pertanto necessario formulare delle ipotesi che lo mettano in relazione RLGnd. In particolare, si può supporre che:

- $\Delta nd = 0$: ciò equivale a ipotizzare che tutti gli elementi che distinguono RLG da RF non abbiano rilevanza ai fini della determinazione del *gap*, ad esempio che una impresa completamente in nero non abbia ammortamenti ovvero che un contribuente con una parte del risultato di gestione dichiarato ed una parte occultato scarichi tutti gli ammortamenti sulla prima componente;
- Δnd è proporzionale a RLGnd, con una proporzione che, in ciascun settore di attività economica, è analoga a quella riscontrata nel dichiarato.

Poiché l'ipotesi "a" appare poco realistica, nella prima fase di applicazione della presente metodologia si è deciso di adottare l'ipotesi "b". Tuttavia, a partire dall'edizione 2021 della relazione, si è deciso di orientarsi verso un'ipotesi intermedia (ipotesi "c") in virtù della quale si assume che alcuni elementi che compongono Δnd possono considerarsi privi di evasione mentre per altri si può mantenere l'ipotesi di proporzionalità. Al fine di approssimare maggiormente la base imponibile effettiva, in primo luogo si ricostruiscono gli ammortamenti potenziali e non dichiarati mantenendo per tale componente l'ipotesi di proporzionalità tra dichiarato ed evaso: non

è, infatti, possibile ricostruire l'ammontare complessivo degli ammortamenti a partire dai dati di contabilità nazionale in quanto le diverse definizioni di tale aggregato (statistica, civilistica e fiscale) non sono omogenee.

Ripartizione della BIND per tipologia di imposta

Per derivare il *gap* d'imposta è necessario, in primo luogo, ripartire $BIND_{IRES+IRPEF}$ tra le platee di contribuenti che pagano i due tributi, ovvero scomporlo in base alla forma giuridica.

I dati dichiarati sono per loro natura disaggregati per forma giuridica, in quanto i diversi contribuenti compilano campi o modelli distinti.

Per disaggregare il RLG potenziale, invece, si è utilizzato il quadro di raccordo illustrato nella Tabella I.1.1. Da quest'ultima si evince che esiste una corrispondenza sostanzialmente biunivoca tra i settori istituzionali ISTAT e la classificazione per forma giuridica che deve essere impiegata nel calcolo del *gap*. Fanno eccezione le società di persone con più di cinque addetti che sono collocate nell'ambito del settore istituzionale "società non finanziarie", e che devono essere collocate nel settore delle famiglie produttrici, i cui redditi sono prevalentemente assoggettati all'IRPEF⁵.

Utilizzando i conti per settore istituzionale, pertanto, è possibile scomporre il totale nazionale di RLG tra IRES e IRPEF. Grazie ad uno specifico gruppo di lavoro che ha visto la collaborazione dell'ISTAT, è stato possibile disporre di dati disaggregati sia per settore istituzionale sia per branca di attività economica, nonché di predisporre una matrice di transizione che garantisca omogeneità tra le classificazioni settoriali adottate nelle fonti statistiche e in quelle fiscali. Una volta ripartito RLG tra le due sotto-platee, il valore di RLGnd per ciascuna imposta e per ciascun settore di attività economica è calcolato per differenza, applicando la formula 1.

Determinazione dell'aliquota e del *gap* d'imposta

L'ultimo passaggio della stima consiste nella determinazione dell'aliquota da applicare alla base imponibile non dichiarata al fine di individuare l'imposta evasa.

Per l'IRES, la normativa vigente prevede un'aliquota unica, che dal 2017 è stata fissata al 24%. Tuttavia l'aliquota effettiva risulta inferiore rispetto all'aliquota legale, in virtù di agevolazioni ed eccezioni concesse a favore dei contribuenti. Pertanto, l'applicazione dell'aliquota legale alla base imponibile non dichiarata comporterebbe una sovrastima del *gap*.

Ciò rende necessario l'individuazione di un'aliquota per il non dichiarato e, nello specifico, si possono definire due diversi tipi di aliquota in corrispondenza delle prime due ipotesi proposte sulla tra RLGnd e $BIND$ ⁶:

a) $Aliq_{IRES} = IRESd / Rf d IRES$ [3] corrispondente all'ipotesi "a": $\Delta nd = 0$;

b) $Aliq_{IRES} = IRESd / RLGd IRES$ [4] corrispondente all'ipotesi "b": $\Delta nd = RLGnd$

dove $IRESd$ è l'imposta dichiarata dell'*i*-esimo settore di attività economica. Nell'ipotesi "c", invece, la base imponibile è determinata utilizzando un aggregato intermedio tra risultato lordo di gestione e reddito imponibile, come ad esempio il risultato netto di gestione ($RNG = RLG - \text{Ammortamenti}$) ovvero il reddito d'impresa.

Per l'IRPEF, invece, si riscontra un'ulteriore difficoltà, in quanto si prevede un sistema di aliquote crescenti per scaglioni di reddito imponibile⁷.

Di conseguenza, è necessario considerare che l'aliquota media sul non dichiarato può risultare diversa dall'aliquota media sul dichiarato. Per tenere conto di tale aspetto, si adotta l'aliquota

⁵ A tal fine è stata utilizzata una matrice di pesi fornita dall'ISTAT.

⁶ Si veda paragrafo 3.

⁷ Le attuali aliquote per scaglioni, in vigore dal 2007, sono: 23% per redditi fino a 15.000€; 27% da 15.001€ a 28.000€; 38% da 28.001€ a 55.000€; 41% da 55.001€ a 75.000€; 43% oltre 75.000€

implicita sul risultato lordo di gestione (*AliqI*) calcolata analogamente a quanto esposto in relazione all'IREs e si applica a tale aliquota un differenziale sull'evaso (*diffAliq*) ottenuto dalle stime bottom-up⁸, a sua volta corretto per tenere conto del fatto che l'aliquota considerata non è calcolata sulla base imponibile bensì sul risultato lordo di gestione:

$$Aliq_IRPEF^i = Aliq_IRES^i + diffAliq^i \quad [6]$$

dove *Aliq_IRPEF* è calcolata, secondo le modalità già esposte in relazione all'IREs, adottando una delle ipotesi illustrate nel paragrafo 3 (dall'edizione 2021, l'ipotesi "c").

L'imposta non dichiarata, a questo punto, è ottenuta moltiplicando la BIND di ciascun settore per la corrispondente aliquota:

$$IRESnd^i = BIND_IRES^i * Aliq_IRES^i \quad [7]$$

$$IRPEFnd^i = BIND_IRPEF^i * Aliq_IRPEF^i \quad [8]$$

Infine, per ottenere il *gap* complessivo, occorre aggiungere all'imposta relativa ad omessa dichiarazione, di cui alle formule [7] e [8], le somme non versate in virtù di errori nella compilazione dei modelli o a crisi di liquidità (*IRESmv* e *IRPEFmv*), desunte dagli esiti dei controlli automatizzati effettuati ai sensi dell'articolo 36-bis del DPR 600/1973:

$$GAP\ IRES = IRESnd + IRESmv \quad [9]$$

$$GAP\ IRPEF = IRPEFnd + IRPEFmv \quad [10]$$

Metodologia bottom-up per il calcolo del differenziale di aliquota tra dichiarato ed evaso Nella stima del *gap* IRPEF

La quantificazione del differenziale di aliquota tra dichiarato ed evaso utilizzato nel calcolo del *gap* IRPEF si fonda sull'utilizzo di una metodologia bottom-up basata sui dati degli accertamenti effettuati nell'ambito delle ordinarie attività di controllo: di conseguenza, poiché la selezione dei contribuenti da sottoporre ad accertamento non avviene in maniera casuale, si deve individuare una modalità per correggere il *selection bias* al fine di rendere i risultati ottenuti nel campione distorto rappresentativi dell'intera platea.

Nella definizione della popolazione oggetto di analisi, si è deciso di privilegiare la rimozione delle casistiche potenzialmente anomale rispetto al mantenimento di una platea più ampia e completa. Infatti, poiché l'obiettivo del calcolo è quantificare una differenza tra due rapporti, si è preferito evitare di introdurre elementi potenzialmente distorsivi. A tal fine, sono stati considerati i contribuenti Persone Fisiche con partita IVA attiva, in normale attività (sono esclusi fallimenti, liquidazioni e decessi), che abbiano presentato sia la dichiarazione IRAP sia la dichiarazione Unico (con l'esclusione del quadro LM per i soggetti in regimi agevolati), che non rientrino nei settori di attività economica esclusi dalla stima del *gap* IRPEF, per i quali la regione non sia ignota, che abbiano subito non più di un accertamento unificato (riguardante le imposte dirette) e che non abbiano subito altri tipi di controlli. Sono stati, inoltre, esclusi i casi in cui la dichiarazione risulta presentata in ritardo oppure è stata modificata successivamente al primo invio, ovvero in cui l'imposta lorda dichiarata risulta non compatibile con il reddito imponibile.

Innanzitutto, si è proceduto a valutare la qualità delle fonti di dati utilizzate. In primo luogo, sono stati analizzati gli importi dichiarati riportati nel database degli accertamenti, confrontandoli con i corrispondenti importi risultanti in dichiarazione. La coincidenza di tali importi nella quasi totalità dei casi (oltre il 99%) ha confermato come la maggior parte delle casistiche anomale siano state correttamente eliminate in fase di definizione della platea. In secondo luogo, sono stati

⁸ A partire dall'edizione 2021, è stata predisposta una procedura *ad hoc* di stima bottom-up illustrata nel Paragrafo 6.

effettuati controlli di coerenza sugli importi accertati al fine di garantire la corretta ricostruzione degli importi evasi.

Una volta identificate le variabili degli importi dichiarati (reddito e imposta) e, per i soggetti sottoposti ad accertamento unificato, le corrispondenti variabili degli importi accertati, è necessario imputare i valori (non osservati) delle variabili accertate anche per i soggetti non controllati, prima di procedere al calcolo e al confronto delle aliquote.

Si considerano due possibili soluzioni per la correzione del *selection bias*:

- 1) **Post-stratificazione:** in analogia a quanto effettuato con la precedente metodologia bottom-up, la platea viene suddivisa in strati in base alle variabili di collocazione geografica, settore di attività economica e fascia di ricavi. Per aumentare la numerosità campionaria, gli anni di imposta vengono considerati tutti insieme, di modo che il risultato ottenuto rappresenti una media ponderata delle annualità coinvolte nel calcolo. Gli importi dichiarati e accertati sono riportati all'universo moltiplicando il totale di strato per il fattore N_{tot_i}/n_{acc} , dove N_{tot_i} e n_{acc} rappresentano rispettivamente il totale dei soggetti e il numero di soggetti accertati nell'*i*-esimo strato. La scelta di utilizzare i valori stimati anche per i totali delle variabili dichiarate discende dal confronto con i valori effettivi (calcolati sommando gli importi sull'intera platea), che evidenzia come questa tecnica tenda a produrre valori sovra-stimati.
- 2) **Matching:** viene stimato un modello probit per la probabilità di essere accertato e il *propensity score* risultante dal modello viene utilizzato in un algoritmo K-NN per abbinare ciascun soggetto non accertato con K soggetti accertati (in questo caso, K=3), da cui mutuare l'importo di una o più variabili non osservate. Al fine di garantire risultati coerenti, si è deciso di non acquisire dai vicini il reddito accertato (onde evitare di ottenere un importo di evasione negativo qualora il vicino accertato abbia un reddito minore del soggetto non accertato) bensì il reddito evaso (ovvero la differenza tra reddito accertato e reddito dichiarato del k-esimo vicino) e di ottenere l'imposta applicando le aliquote per scaglioni al reddito ricostruito (ovvero alla somma del dichiarato del soggetto non accertato e dell'evaso medio dei 3 vicini più prossimi). Una volta attribuiti i valori accertati a tutti i soggetti della platea, gli importi vengono aggregati (sommati) in base alla medesima classificazione settoriale utilizzata nel calcolo dell'aliquota e del *tax gap*.

In entrambi i casi, il differenziale di aliquota per ciascun raggruppamento viene calcolato, in base alla formula seguente:

$$\frac{\text{Totale imposta accertata}}{\text{Totale reddito accertato}} - \frac{\text{Totale imposta dichiarata}}{\text{Totale reddito dichiarato}}$$

e il differenziale effettivamente utilizzato nelle stime rappresenta la media geometrica dei risultati delle due tecniche proposte.

Ripartizione territoriale del *gap* IRPEF e calcolo del *gap* sulle addizionali

La ripartizione regionale del *gap* IRPEFaut viene effettuata replicando a livello regionale la procedura di stima del *tax gap* nazionale: in particolare, il risultato lordo di gestione potenziale viene ricavato a partire dai conti regionali delle famiglie produttrici (cui vengono apportate le stesse correzioni effettuate per il corrispondente importo nazionale) mentre il costo del lavoro non dichiarato è tratto dai conti territoriali.

Il *gap* delle addizionali regionali e comunali è calcolato moltiplicando la quota di *gap* IRPEFaut attribuita a ciascuna regione per un coefficiente che tiene conto della diversa incidenza delle addizionali rispetto all'IRPEF in ciascuna regione.

Trattandosi di una procedura ancora in fase sperimentale, si è ritenuto preferibile riportare i primi risultati in un *box* dedicato, senza tuttavia includerli nelle stime complessive del *gap*.

I.2 METODOLOGIA PER LA STIMA DEL TAX GAP IVA

Introduzione

L'imposta sul valore aggiunto (IVA) è un'imposta sui consumi che grava sulla maggior parte dei beni e servizi scambiati per l'uso finale. È la fonte principale di gettito tra le imposte indirette e fornisce circa il 28% del gettito fiscale totale. Da un punto di vista teorico, il tributo non è influenzato dalla lunghezza della filiera di produzione e distribuzione, poiché è versato in modo frazionato in ogni fase dello scambio in ragione del valore addizionale che si crea. Questo aspetto garantisce la neutralità dell'imposta rispetto al grado di integrazione verticale dei processi di produzione e di distribuzione che i beni e i servizi seguono prima di essere acquistati dai consumatori finali.

Tecnicamente, la neutralità fiscale si ottiene con il diritto alla detrazione dell'IVA sugli acquisti. Attraverso questo meccanismo il tributo è versato in modo frazionato dai contribuenti, partite IVA, che operano lungo la filiera, dal momento della produzione, e/o importazione, fino al consumo finale. Tramite la traslazione in avanti dell'imposta, da parte delle partite IVA, il soggetto inciso dal tributo è il consumatore finale.

Alcuni operatori economici non possono applicare questo meccanismo, in tutto o in parte, come ad esempio la Pubblica amministrazione o il sistema bancario, e rimanendo incisi dal tributo, agiscono come consumatori finali.

In ogni fase della filiera, l'importo dell'IVA versato dalle partite IVA, che operano come soggetti passivi, è determinato sulla base del valore delle cessioni imponibili che implicano un'IVA dovuta e degli acquisti, che danno diritto alla detrazione di tutta, o parte, dell'IVA assolta sugli stessi.

Da un punto di vista macroeconomico, le operazioni IVA coinvolgono, principalmente, tre attori: le famiglie, le Amministrazioni pubbliche e le imprese. Nel corso degli scambi di beni e servizi che si instaurano tra di loro, ogni volta che un acquirente non ha il diritto di detrarre l'IVA, parzialmente o totalmente, agisce come consumatore finale. In generale, le famiglie e le Pubbliche amministrazioni non possono detrarre l'IVA sui loro acquisti⁹, mentre alle imprese è concesso. Per queste ultime esistono, però, numerose eccezioni che insorgono quando le loro vendite sono esentate dall'applicazione dell'IVA, ovvero quando effettuano degli acquisti per i quali, la legislazione non permette la detrazione.

Le operazioni effettuate tra gli operatori economici possono portare a comportamenti non conformi alla normativa con riguardo sia al rapporto tra le imprese (BtoB), sia a quello tra le imprese e le famiglie (BtoC).

In teoria, il meccanismo di deduzione, negli scambi BtoB, è una forma di mutuo contralto tra i contribuenti business che, tuttavia, non rimuove l'incentivo a evadere l'IVA. Quest'ultimo, infatti, non è costituito in modo preminente dal mancato versamento dell'IVA, quanto, dalla mancata dichiarazione dei redditi "a valle", generati dalle transazioni occultate.

Le transazioni BtoC sono ancora più esposte al rischio di evasione, in quanto il consumatore finale non ha alcun incentivo a richiedere l'emissione della fattura al venditore.

L'Appendice è organizzata come segue: nel Paragrafo A.2 sono illustrati gli aspetti principali del metodo di calcolo adottato, mentre nella sezione successiva si illustrano i dettagli tecnici della procedura e la fonte dei dati utilizzati.

Una visione d'insieme della metodologia

Il *gap* IVA è stimato con un approccio top-down, mediante il confronto dell'imposta acquisita a titolo definitivo dalle casse dallo Stato, con quella che si sarebbe dovuta incassare se tutti

⁹ Fanno eccezione le famiglie quando operano come produttrici e le Pubbliche amministrazioni quando effettuano cessioni imponibili.

i contribuenti avessero operato in perfetta conformità alla legislazione esistente. Il gettito stimato nel caso di perfetta *compliance* è definito come "potenziale" o "teorico". Da un lato, si qualifica come "potenziale" poiché rappresenta il massimo delle entrate conseguibili in base alla normativa vigente. D'altra parte, è anche "teorico" in quanto rappresenta l'ammontare di imposta che lo Stato dovrebbe incassare in assenza di evasione.

Per calcolare il gettito "potenziale" si utilizzano i dati della Contabilità Nazionale prodotti dall'Istat. A tal fine si possono seguire due approcci: il primo si basa sui dati dell'offerta e il secondo sui dati della domanda. In particolare, il primo utilizza principalmente i dati del valore aggiunto prodotto dalle imprese, mentre il secondo considera le categorie di spesa, quali: i consumi finali, i consumi intermedi, gli investimenti, ecc.

Nelle pagine che seguono si illustra l'approccio dal lato della domanda, che è il metodo seguito dall'Italia¹⁰.

Al fine di ottenere una misura accurata dell'IVA potenziale (IVAT), abbiamo bisogno di identificare correttamente la sottostante base potenziale (BIT) e la corrispondente aliquota fiscale da applicare. A tal fine è necessario utilizzare i dati di contabilità nazionale ad un livello di disaggregazione molto dettagliato, al fine di considerare correttamente la complessità della normativa IVA e, quindi, rendere coerente le grandezze statistiche con le definizioni fiscali.

In particolare, da ogni aggregato dei Conti nazionali si scorporano le componenti esenti e fuori dal campo di applicazione del tributo. Alla parte residuale è applicata la specifica aliquota IVA di legge.

L'IVA acquisita a titolo definitivo dalle casse dello Stato è altresì denominata IVA effettiva di competenza economica (IVAEC). In particolare, si considera solo il tributo versato spontaneamente dal contribuente a seguito delle transazioni economiche effettuate in un anno solare. Ovvero IVAEC è calcolata in base allo stesso criterio di competenza economica adottato dalla contabilità nazionale per calcolare gli aggregati della domanda e dell'offerta.

La base IVA coerente con l'IVA effettiva di competenza economica (BID) si ottiene dividendo IVAEC per l'aliquota media calcolata sulla base delle dichiarazioni fiscali, distinguendo le componenti del mercato interno da quelle delle importazioni.

Il *gap* nella base IVA (BIND) e quello dell'imposta (IVAEV) sono calcolate tramite le formule [2.1] e [2.2].

$$BIND = BIT - BID \quad [2.1]$$

$$IVAEV = IVAT - IVAEC \quad [2.2]$$

Il *gap* con e senza complicità

Il *gap* IVA è stimato adottando due ipotesi di comportamento dei contribuenti: "con complicità" tra venditore e acquirente (entrambi si accordano, non viene emessa alcuna fattura né versata l'imposta); "senza complicità" tra le parti (l'imposta è applicata e incassata dal venditore ma non è versata). La differenza tra i due comportamenti è esemplificata nella Tabella I.2.1.

Nell'esempio si ipotizzano tre transazioni avvenute fra un soggetto IVA ed un consumatore finale.

Nella prima transazione siamo in assenza di evasione: si effettua una vendita per €1000, si emette regolare fattura e l'IVA incassata dal venditore è regolarmente versata. Il dato di Contabilità nazionale coincide con quello fiscale, ovvero in entrambe le basi dati si ritroverà un valore di base più imposta pari a €1.200.

¹⁰ Per approfondimenti sul metodo dal lato dell'offerta si rimanda a D'Agosto et al. (2016).

Nel secondo caso, quello di evasione "con complicità", il venditore si accorda con l'acquirente vende €1.000 di merce, non emette fattura e non versa l'IVA. In questa fattispecie la Contabilità nazionale registra un valore dei consumi finali pari a €1.000, mentre negli archivi fiscali sia la base che l'imposta sono uguali a zero.

Nella terza transazione, quella di evasione senza complicità, il venditore emette regolare fattura per €1.000, incassa €200 di IVA dall'acquirente ma poi non la versa. Come nel caso precedente, negli archivi fiscali non comparirà né la base, né l'imposta, ma il corrispondente aggregato di Contabilità nazionale relativo ai consumi finali sarà uguale a €1.200, ovvero base e imposta che, entrambi, concorrono a remunerare l'attività del venditore.

TABELLA I.2.1 ESEMPIO DI EVASIONE CON E SENZA COMPLICITÀ, RIFERITO A 3 TRANSAZIONI DI BENI DI CONSUMO FINALE CON ALIQUOTA IVA AL 20%

ID	Base IVA potenziale	IVA fatturata	IVA versata	Dato di Contabilità nazionale	Comportamento economico
1	1.000	200	200	1.200	Assenza di evasione
2	1.000	0	0	1.000	Evasione con complicità
3	1.000	200	0	1.200	Evasione senza complicità
Totale	3.000	400	200	3.400	

La disaggregazione illustrata in Tabella I.2.1 è ignota. Ipotizzando che le tre transazioni costituiscano una voce di spesa considerata per il calcolo del *gap* IVA, i dati a nostra disposizione sono solo:

- la base potenziale più l'IVA effettivamente versata dal venditore, pari a €3.400, ovvero il dato dei consumi finali;
- l'IVA effettivamente versata = € 200;
- la base IVA dichiarata = € 1.000.

Con riferimento all'esempio di Tabella I.2.1, la procedura per stimare il *gap* con e senza complicità è illustrata in Tabella I.2.2.

Se si assume che tutta l'evasione avviene con complicità, allora al dato di Contabilità nazionale (€ 3.400) dovrebbe essere sottratta l'IVA effettivamente versata (€ 200). Da ciò deriverebbe una base potenziale pari a € 3.200. Se, invece, si assume che tutto il *gap* provenga da un comportamento senza complicità, allora il dato di Contabilità nazionale dovrebbe essere diviso per l'aliquota media IVA, pari al 20%. In questo caso la base potenziale sarebbe uguale a €2.833.

Dalla Tabella I.2.1 emerge che la vera base non dichiarata (non osservabile) è uguale a €2.000, valore inferiore a quello stimato nell'ipotesi con complicità, € 2.200 riportato in Tabella I.2.2, e superiore a quello calcolato con l'ipotesi senza complicità, € 1.833. Dal momento che l'evasione IVA può avvenire sia con sia senza complicità, ma è ignoto quanta parte sia imputabile all'una e quanta all'altra, allora, quando si adotta il metodo con complicità il *gap* "vero" è sovra-stimato; al contrario, quando si segue il metodo senza complicità si sottostima il *gap*. Attualmente, non si è in grado di identificare l'entità di evasione per ogni comportamento, quindi si elaborano due stime. In primo luogo, si calcola il *gap* assumendo che tutta l'evasione avvenga senza complicità, quindi si determina il *gap* supponendo che avvenga tutta con complicità.

TABELLA I.2.2 PROCEDURA DI CALCOLO DELL'EVASIONE CON E SENZA COMPLICITÀ APPLICATE AI DATI ILLUSTRATE NELLA TABELLA I.2.1

Aggregati	Con complicità	Senza complicità
Dato di contabilità nazionale	3.400	3.400
IVA effettivamente versata (IVAEC)	200	
Base IVA potenziale (BIT)	3.200	3.400/1.2=2.833
Base IVA dichiarata (BID)	1.000	1.000
Stima del gap nella base (BIND)	2.200	1.833
	Sovra-stima	Sotto-stima

Il punto di partenza della procedura è rappresentato dai dati di Contabilità nazionale (NA). Si consideri NAG come un generico gruppo di prodotti di NA (ad esempio una voce dei consumi delle famiglie). Il nostro scopo consiste nello scomporre NAG come segue:

$$NAg = EXEg + BIDg + BINDg_{wout} + BINDg_{with} + IVAEVg_{wout} + IVAECg \quad [2.3]$$

dove, oltre alle variabili già definite, troviamo:

$BINDg_{wout}$ = Base IVA evasa senza complicità;

$BINDg_{with}$ = Base IVA evasa con complicità;

$EXEg$ = parte di NAG esente o fuori dal campo di applicazione dell'IVA, $0 \leq NAg \leq 1$;

$IVAEVg_{wout}$ = IVA fatturata, acquisita ma non versata dal venditore (senza complicità);

$IVAECg$ = IVA effettivamente fatturata e versata;

$BIDg$ = Base IVA dichiarata.

L'equazione [2.3] implica:

$$BINDg = BINDg_{wout} + BINDg_{with}$$

Subordinatamente all'ipotesi che l'aliquota IVA, τ_g , per ciascun NAG, sia uguale tanto per le transazioni dichiarate, quanta per quelle occultate al Fisco¹¹, la [2.3] si può riscrivere come:

$$NAg = (BIDg + BINDg_{wout})(1 + \tau_g) + BINDg_{with} + EXEg \quad [2.4]$$

Nella prima fase della procedura $EXEg$ è scorporato da NAG. Quindi, fondandosi sulla normativa fiscale e utilizzando specifici indicatori, si calcola il coefficiente $cexe$, che rappresenta una proxy del rapporto:

$$cexe = \left(\frac{EXEg}{(BIDg + BINDg_{wout})(1 + \tau_g) + BINDg_{with} + EXEg} \right)$$

Applicando $(1 - cexe)$ a NAG, si deriva una grandezza che rappresenta la base IVA potenziale, più l'IVA effettivamente versata, più l'IVA fatturata e non versata. In termini formali:

$$BITIVAECG = (BIDg + BINDg_{wout})(1 + \tau_g) + BINDg_{with} \quad [2.5]$$

Da qui in avanti la metodologia segue due approcci diversi a seconda che si voglia calcolare il gap con o senza complicità.

¹¹ Tale assunzione è tanto più realistica quanto più è dettagliato il livello di "g". Se "g" si riferisce ad un singolo bene, in fatti, l'aliquota TVA sarà identica per entrambe le transazioni.

Sotto l'ipotesi che tutta l'evasione avvenga con complicità, allora si sottrae dalla [2.5] l'IVA effettivamente versata (IVAEC), ottenendo la base potenziale IVA, $BITg_{with}$, come:

$$BINDg_{with} = BIDg + BINDg_{wout}(1 + \tau_g) + BINDg_{with} \quad [2.6]$$

Dalla [2.6] si ricava che la base potenziale è sovra-stimata in quanto include l'IVA fatturata ma non versata da coloro i quali pongono in essere l'evasione senza complicità, $BINDg_{with}(1 + \tau_g)$.

Se, invece, si ipotizza che tutta l'evasione avvenga senza complicità, allora per calcolare la base potenziale occorre dividere la [2.5] per $(1 + \tau_g)$. In termini formali:

$$BINDg_{wout} = \frac{(BIDg + BINDg_{wout})(1 + \tau_g) + BINDg_{with}}{(1 + \tau_g)} \quad [2.7]$$

La [2.7] sotto-stima la base IVA potenziale, $BITg$, in quanto divide $BINDg_{with}$ per $(1 + \tau_g)$, ovvero ipotizza che anche a questa base sia associate un valore dell'imposta che, invece, non è incluso negli aggregati di Contabilità nazionale.

Confrontando la [2.6] e la [2.7], si può concludere che il vero valore, non conosciuto, di $BITg$ compreso nel seguente intervallo:

$$BITg_{wout} \leq BITg \leq BITg_{with} \quad [2.8]$$

Al fine di calcolare l'IVA potenziale, $IVATg$, $BITg_{wout}$ e $BITg_{with}$ sono entrambe moltiplicate per la corrispondente aliquota, τ_g , ovvero:

$$IVATg_{wout} = BITg_{wout} \cdot \tau_g \quad [2.9a]$$

$$IVATg_{with} = BITg_{with} \cdot \tau_g \quad [2.9b]$$

Sulla base della [2.8], si ottiene:

$$IVATg_{wout} \leq IVATg \leq IVATg_{with} \quad [2.10]$$

Il valore di BIT e di IVAT per il totale dell'economia è ottenuto sommando $BITg$ e $IVATg$ per tutti i gruppi di prodotti "g" assoggettati al tributo.

Per completare il calcolo delle grandezze incluse nella [2.1] e nella [2.2] è necessario calcolare i valori della base dichiarata, BID, e dell'imposta effettivamente versata IVAEC. I dettagli tecnici di questo calcolo sono illustrati nel Paragrafo 3, mentre di seguito si segnalano due aspetti di carattere generale.

Per le caratteristiche tipiche dell'IVA, è difficile calcolare BID e IVAEC per prodotto, così come si è fatto per i valori potenziali di base e imposta. Infatti, l'IVA è versata dalle imprese che, di solito, vendono un paniere di prodotti che possono essere anche molto eterogenei tra di loro. Per questo motivo il calcolo di BID e di IVAEC è effettuato per il totale dell'economia.

Inoltre, al fine di includere nel *gap* anche la componente imputabile alle somme dichiarate e non versate, il calcolo di BID si effettua a partire da IVAEC. Ovvero, BID è calcolata dividendo IVAEC per l'aliquota media ponderata calcolata sulle dichiarazioni fiscali¹², τ_d .

$$BID = \frac{IVAEC}{\tau_d} \quad [2.11]$$

È importante notare che τ_d è differente da τ , cioè l'aliquota impiegata per calcolare l'IVA potenziale. La prima, infatti, è l'aliquota che grava effettivamente sui contribuenti che

¹² Per i dettagli relativi al metodo di calcolo si rimanda al Paragrafo 3.

versano l'imposta, mentre la seconda è quella che si sarebbe dovuta applicare in caso di perfetta *compliance*¹³.

Sulla base delle [2.6], [2.7], [2.9a], [2.9b] and [2.11], le [2.1] e [2.2], possono essere riscritte come:

$$BIND_{wout} = BIT_{wout} - BID \quad [2.12a]$$

$$BIND_{wth} = BIT_{wth} - BID \quad [2.12b]$$

$$IVAEC_{wout} = IVAT_{wout} - IVAEC \quad [2.13a]$$

$$IVAEC_{wth} = IVAT_{wth} - IVAEC \quad [2.13b]$$

Dove le [2.12a] e [2.13a] rappresentano il limite inferiore della stima, mentre le [2.12b] e [2.13b] quello superiore.

L'IVA regionale

In analogia con le stime nazionali, quelle territoriali¹⁴ del *gap* dell'IVA sono effettuate utilizzando sia i dati di Contabilità nazionale disaggregati su base regionale, sia alcuni specifici dati dichiarativi che consentono un'opportuna articolazione del gettito relativo.

Per identificare una misura regionale del *gap* dell'IVA, sia il gettito teorico che quello di competenza sono ripartiti territorialmente. In questa articolazione occorre tenere conto di due aspetti importanti: il concetto di IVA correlata al luogo di consumo, e quello di IVA riferita alla sede legale dell'impresa. Infatti, mentre il gettito teorico emerge e, quindi, viene stimato dove la spesa è effettuata, ovvero per luogo di consumo, diversamente, l'imposta viene versata sul territorio dove l'impresa ha la sede legale. Per poter effettuare un confronto opportuno tra le due grandezze territoriali si utilizza il quadro VT della dichiarazione IVA, che contiene informazioni sulla ripartizione del gettito dichiarato per la regione dove è stata effettuata la spesa. Dal raffronto tra le grandezze teoriche e quelle effettive derivano i livelli di *gap* dell'imposta e della sua propensione.

Fonti e metodi della procedura di calcolo

In questo paragrafo si descrivono dettagliatamente i passaggi sottostanti la metodologia delineata nel precedente paragrafo. L'elenco delle principali fonti dei dati impiegati e illustrate nella Tabella I.2.3.

¹³ I dati dell'IVA desunti dalle dichiarazioni fiscali non possono essere utilizzati per la stima di BID e di IVAEC, a causa della distorsione dovuta ai mancati versamenti dell'IVA dichiarata. La [2.11] sconta l'ipotesi che tale distorsione non influenzi l'aliquota media τ_d , cioè che i mancati versamenti non modificano la relazione tra la base e l'imposta indicata in dichiarazione (i mancati versamenti si distribuiscono proporzionalmente tra le differenti aliquote), e, pertanto, non inducano alcuna distorsione sull'aliquota media.

¹⁴ Per approfondimento sul metodo si veda D'Agosto, Marigliani, Pisani (2014).

TABELLA I.2.3 FONTE DEI DATI UTILIZZATI PER CALCOLARE IL GAP IVA

Aggregati	Fonte dei dati
IVA di competenza economica (IVAEC): • IVA lorda, rimborsi e compensazioni; • Aggiustamento per la competenza economica; • Variazione dello stock di crediti da portare all'annosuccessivo.	Bilancio delle Stato; Contabilità nazionale (Istat); Dichiarazioni IVA.
Base IVA dichiarata (BID): • IVAEC; • Aliquota media delle dichiarazioni	Calcolo dell'Agenzia delle entrate Dichiarazioni IVA.
Base IVA potenziale (BIT): • Consumi delle famiglie (HC); • Investimenti della Pubblica amministrazione (GGI); • Consumi intermedi della Pubblica amministrazione (GGIC); • Consumi intermedi delle imprese market (MEIC) • Specifiche voci di investimenti per le imprese market (MEI) • Tassi di esenzione per gruppo NACE e per prodotto	Dati di contabilità nazionale ed indagine Istat sui bilanci di famiglia. Istat, Conto consolidato della Pubblica amministrazione. Istat, Conto consolidato della Pubblica amministrazione. Dati di Contabilità nazionale per divisione Nace. Dati di Contabilità nazionale per tipologia di investimenti. Base di calcolo risorse proprie IVA (per HC, GGIC, GGI) Dichiarazioni IVA (MEIC, MEI).
IVA potenziale (IVAT): Aliquote di legge per tipologie di prodotti.	Base di calcolo risorse proprie IVA

La base IVA potenziale (BIT) è calcolata coerentemente con le classificazioni e le definizioni sancite dalla legislazione, applicate ai dati della Contabilità nazionale Istat. Questi ultimi possono essere raggruppati nelle seguenti 5 macro categorie:

- Consumi delle famiglie (261 voci di spesa);
- Investimenti della Pubblica amministrazione (12 voci di spesa);
- Consumi intermedi della Pubblica amministrazione (17 voci di spesa);
- Consumi intermedi delle imprese (58 settori di attività economica);
- Specifiche tipologie di investimenti delle imprese *market* (ad esempio autoveicoli).

Ai fini IVA, sia le Amministrazioni pubbliche che offrono servizi non di mercato (la predominante parte dei servizi pubblici lo sono), sia specifici segmenti di mercato delle imprese, agiscono come consumatori finali quando non recuperano l'imposta.

Le imprese di mercato non possono recuperare l'IVA in due situazioni: la prima è dovuta al tipo di merce acquistata, indicata come "indetraibilità oggettiva"; la seconda, indicata come "soggettiva", è legata al tipo di attività: imprese che vendono beni e servizi esenti non possono recuperare l'IVA pagata per gli acquisti "afferenti" tali operazioni.

Per la quantificazione di tali grandezze, si parte dall'assunto che i costi intermedi di CN (prezzi base) siano separabili in diverse componenti, di cui alcune sono note, altre vengono determinate sulla base di quote caratteristiche per ogni divisione di attività economica, determinate partendo da alcuni campi della dichiarazione annuale IVA. Tali rapporti, applicati all'aggregato dei costi intermedi netti noti, permettono di esploderlo nelle componenti rimanenti.

Ciascuna voce compresa nei punti da 1) a 5), rappresenta il generico aggregato di Contabilità nazionale indicate con, NAg , nel Paragrafo A.2 (si veda equazione [2.4]). Pertanto, applicando a ciascun NAg la procedura descritta nel Paragrafo A.2, formule da [2.3] a [2.7], si arriva a calcolare $BITg$. La corrispondente imposta $IVATg$ è calcolata moltiplicando $BITg$ per l'aliquota τ_g , formule [2.9a] e [2.9b].

Disponendo di $BITg$ e di $IVATg$, per ciascun g , è possibile calcolare una aliquota media ponderata, AMT , che rappresenta l'aliquota implicita IVA del sistema nell'ipotesi di completa *compliance* fiscale. AMT si ottiene tramite la seguente formula

$$AMT = \frac{IVAT_{wout}}{BIT_{wout}} = \sum_{g=1}^n \tau_g \frac{BITg_{wout}}{\sum_{g=1}^n BITg_{wout}} = \frac{IVAT_{wout}}{BIT_{with}} \quad [3.1]$$

Per ultimare il calcolo del *gap* è necessario calcolare l'imposta effettivamente versata, IVAEC, e la base coerente con essa.

IVAEC è calcolato secondo un criterio di competenza economica più stringente rispetto alle regole dettate dall'Eurostat ed adottate in Italia per calcolare i flussi di finanza pubblica. L'adozione di tale criterio è essenziale per allocare correttamente il *gap* in ciascun anno di riferimento. A tal fine IVAEC si deve riferire al gettito che scaturisce dalle transazioni, gravate di imposta, effettuate nell'anno solare di riferimento.

La formula adottata per calcolare IVAEC è la seguente¹⁵:

$$\text{IVAEC} = \text{IVA lorda da adempimento spontaneo} - (\text{Rimborsi} + \text{Compensazioni}) - \text{Aggiustamenti per competenza giuridica} - \text{variazione dello stock di crediti da portare all'anno successivo}^{16}$$

[3.2]

A destra del segno di uguaglianza della [3.2] troviamo le seguenti grandezze:

- A. IVA lorda da adempimento spontaneo = IVA versata spontaneamente dai contribuenti a seguito di operazioni svolte sul mercato nazionale o su merci importate; l'ammontare si indica come lordo in quanto non considera le somme richieste a rimborso, o in compensazioni, dai contribuenti; si tratta di somme da adempimento spontaneo poiché esclude le somme raccolte a seguito dell'attività di accertamento o tramite condoni;
- B. Rimborsi + Compensazioni = rimborsi richiesti e le compensazioni effettuate nell'anno;
- C. Aggiustamenti per competenza giuridica = correzioni per tenere conto degli sfasamenti temporali dei pagamenti¹⁷;
- D. Variazione nello stock di crediti da portare all'anno successivo = è una residua di IVA detraibile non impiegata nelle liquidazioni periodiche, né richiesta a rimborso, né portata in compensazione e che ogni anno i contribuenti possono annotare nella dichiarazione IVA in aggiunta allo stock preesistente. Tale importo potrà essere utilizzato l'anno successivo, o in compensazione o richiesto a rimborso; la variazione complessiva di questo stock misura l'incremento/decremento del credito IVA che è stato generato nel sistema economico in un determinato anno.

IVAEC differisce da quella calcolata dagli istituti statistici europei, in accordo con gli standard previsti dal SEC, prevalentemente per due ragioni: nella voce a) esclude le somme recuperate tramite accertamento, aggiunge la correzione della voce d) che non è considerata¹⁸.

Per ottenere la base imponibile IVA coerente con IVAEC (BID), si divide quest'ultima per un'aliquota media calcolata sulla base delle dichiarazioni fiscali, τ_d , si veda formula [2.11].

τ_d è calcolata come media ponderata dell'aliquota applicata sulle importazioni di beni e servizi, τ_{dimp} , e quella sulle transazioni interne, τ_{dom} . In termini formali:

$$\tau_d = \tau_{dimp} \cdot w_{imp} + \tau_{dom} \cdot w_{dom} \quad [3.3]$$

dove τ_{dimp} e τ_{dom} sono aliquote implicite IVA calcolate sui dati delle dichiarazioni fiscali e dove w_{imp} e w_{dom} sono i rispettivi pesi, uguali a:

¹⁵ Si veda anche Rodrigues (2015).

¹⁶ Al fine di richiedere il rimborso o la compensazione nell'anno successivo, i contribuenti devono indicare tale ammontare in uno specifico rigo della dichiarazione.

¹⁷ La correzione definita in ambito EU, coerentemente con quanto stabilito dal regolamento SEC.

¹⁸ È opportuno tenere in considerazione queste diversità quando si confronta la stima del *gap* italiano con quelle elaborate in ambito internazionale, si veda, ad esempio, CASE - CPB (2013-2016).

$$w_{imp} = \frac{\frac{IVAEC_{imp}}{\tau_{dimp}}}{\frac{IVAEC_{imp}}{\tau_{dimp}} + \frac{IVAEC_{dom}}{\tau_{dom}}}$$

$$w_{dom} = \frac{\frac{IVAEC_{dom}}{\tau_{dom}}}{\frac{IVAEC_{imp}}{\tau_{dimp}} + \frac{IVAEC_{dom}}{\tau_{dom}}}$$

Con $IVAEC_{imp} + IVAEC_{dom} = IVAEC$

Dalla procedura è pertanto possibile derivare tre differenti aliquote implicite IVA: AMT, formula [3.1], e l'aliquota IVA che si registrerebbe in caso di perfetta compliance; $\tau_d = ALBID$, formula [3.3], e l'aliquota IVA che risulta dalle dichiarazioni fiscali e $ALBIND = IVAEV/BIND$ rappresenta l'aliquota IVA che graverebbe sulle transazioni occultate al Fisco.

L'evidenza empirica mette in luce che sussiste, regolarmente, la seguente disuguaglianza:

$$ALBID \leq AMT \leq ALBIND$$

Ciò è dovuto al fatto che maggiore è l'aliquota legale IVA, più alta è la quota di transazioni evase.

I.3 STIMA DEL GAP IVA: METODOLOGIE A CONFRONTO

La Commissione Europea (CE) pubblica¹⁹ annualmente la stima del *gap* IVA in tutti i Paesi membri. L'approccio adottato è di tipo top-down in cui il *gap* IVA è dato dalla differenza tra l'IVA potenziale teorica e l'ammontare di imposta effettivamente raccolta e si basa sui dati della Contabilità nazionale, come la tavola delle risorse e degli impieghi (*SUT - Supply and Use Table*); ciò comporta sia dei benefici, in termini di comparabilità delle stime tra Paesi, sia dei costi, in termini di diversa affidabilità e qualità a livello nazionale.

Negli ultimi anni, non si registrano variazioni rilevanti della metodologia di calcolo, anche se, a fini di comparabilità con le stime prodotte dalle amministrazioni fiscali nazionali, la CE ha affiancato alle stime ufficiali anche delle *alternative estimates* che tengono conto di alcune specificità nazionali (e.g. Italia e Spagna nel 2019).

L'obiettivo di questo paragrafo è quello di fornire una rappresentazione delle differenze metodologiche esistenti tra le stime pubblicate dalla CE nel Report annuale di "*Study and Reports on the VAT Gap in the EU-28 Member States*" e quelle prodotte e pubblicate nella "Relazione sull'economia non osservata e sull'evasione fiscale e contributiva" dall'Agenzia delle entrate (AdE).

Metodologia di calcolo a confronto

In estrema sintesi²⁰, la metodologia top-down prevede che l'evasione nella base (BIND) sia data dalla seguente relazione:

$$BIND = BIT - BID \quad [1]$$

¹⁹ Le stime vengono effettuate da CASE (Center for Social and Economic Research, Warsaw).

²⁰ Si veda l'Appendice metodologica per maggiori dettagli.

Ovvero, sottraendo alla base potenziale (BIT) quella dichiarata (BID) si ottiene una misura della base imponibile sottratta agli obblighi di legge (*gap* di base IVA).

Il *gap* calcolato in termini di imposta permette di avere una misura più accurata dell'impatto che il *gap* ha sui saldi di finanza pubblica. Applicando alla BIT le aliquote di legge è possibile calcolare l'imposta potenziale che sarebbe generata dal sistema in assenza di evasione (IVAT). Sottraendo a quest'ultima l'IVAEC risulta agevole quantificare l'ammanto per le casse dello Stato, cioè l'IVA non versata (IVANV), ovvero il vero e proprio *gap* IVA. Dal rapporto tra IVANV e IVAT si ottiene un indicatore di *compliance* relativo all'imposta, che comprende sia l'imposta non dichiarata correttamente al fisco sia la parte dichiarata e non versata.

La BID, ovvero la Base Imponibile Dichiarata, coerentemente con le norme adottate in sede comunitaria, deriva dall'IVA di competenza economica²¹. Per ottenere tale aggregato si parte dal gettito IVA al lordo delle rettifiche indotte dai rimborsi e dalle compensazioni. Tenendo conto di queste correzioni e dell'aggiustamento dovuto allo slittamento tra il criterio di competenza giuridica e quello economico dell'imposta, si giunge all'IVA di competenza economica.

Nella definizione dell'IVA versata si rileva la prima e determinante differenza tra la metodologia di calcolo della CE e quella adottata dall'AdE. Infatti, in Italia, il flusso di finanza pubblica è rappresentato dall'IVA effettiva di competenza (IVAEC). Questo aggregato è ottenuto sottraendo all'IVA di competenza economica la variazione dello stock di crediti, per ottenere una rappresentazione dell'IVA effettiva di competenza (IVAEC) coerente con i criteri di contabilizzazione dei Conti nazionali (utilizzati nel metodo top-down). Con il termine *stock* di crediti si definisce l'ammontare di crediti che il contribuente può traslare all'anno di imposta successivo a quello della dichiarazione, indicato nell'apposita riga del quadro IVA. La variazione dello *stock* di crediti da un anno all'altro rappresenta esattamente l'ammontare di IVA detraibile non portata in detrazione nelle liquidazioni periodiche, né richiesta a rimborso né a compensazione entro l'anno.

Inoltre, un'ulteriore differenza riscontrabile nel calcolo dell'IVA versata è dovuta all'inclusione, nel caso della CE, dei proventi derivanti dall'attività di accertamento, che non vengono considerati, invece, nel calcolo dell'IVAEC.

Riassumendo, le seguenti equazioni mostrano il metodo di calcolo rispettivamente di CE ($IVAV_{CE}$) e AdE ($IVAEC_{AdE}$) relativamente all'IVA dichiarata:

$$IVAV_{CE} = IVA \text{ lorda} + \text{accertamento} - \text{rimborsi} - \text{compensazioni} - \text{slittamento} \quad [2]$$

$$IVAEC_{AdE} = IVA \text{ lorda} - \text{rimborsi} - \text{compensazioni} - \text{slittamento} - \Delta \text{stock crediti} \quad [3]$$

Nel medio periodo, le stime effettuate con i due diversi aggregati di gettito, tendono a coincidere, pur presentando degli sfasamenti temporali che possono assumere un'entità rilevante²². Per evitare tali problemi, la CE presenta anche annualmente delle *alternative estimates* che si differenziano per l'inclusione delle informazioni relative allo stock di credito e all'esclusione dei proventi derivanti dall'attività di accertamento. Queste stime permettono una comparabilità tra le due istituzioni preposti alla stima del *gap*.

Per il calcolo della Base Imponibile Teorica (BIT) non si riscontrano differenze metodologiche rilevanti e/o sistematiche nel tempo, anche se valori diversi di IVA potenziale vengono riscontrati annualmente. La fonte dei dati utilizzata in larga parte è la stessa trattandosi di dati di Contabilità nazionale; tuttavia, la non perfetta corrispondenza tra le componenti dell'IVA potenziale (ad esempio, consumi delle famiglie, conti della protezione sociale) non permette un raffronto puntuale delle macro-voci. Nell'ultimo biennio si è osservato uno scostamento tra le due BIT corrispondente a differenti metodologie di calcolo della componente di IVA potenziale da sottrarre ai

²¹ La procedura è quella definita in sede comunitaria in accordo con il regolamento SEC95 e successive modificazioni.

²² La Tabella I.3.1 riporta i flussi che concorrono alla formazione del gettito IVA di competenza. In particolare, le colonne 2 e 5 mostrano l'aggregato di IVA di competenza economica e IVA effettiva di competenza, mentre la colonna 4 la variazione dello stock di crediti.

consumi finali e legata all'attività di contribuenti in regime di vantaggio non tenuti al versamento dell'imposta, quali i minimi e i forfetari.

Sia la CE che l'AdE provvedono annualmente ad aggiornare e revisionare l'intera serie storica della BIT e della BID. Le revisioni possono essere dovute ad aggiornamenti nei dati di Contabilità nazionale pubblicati dall'Eurostat, nel Calcolo delle Risorse Proprie IVA, oppure nella revisione dei parametri che influiscono sull'IVA potenziale; queste revisioni, impattano tanto le stime pubblicate dalla CE, quanto quelle prodotte dall'Agenzia delle entrate.

Risultati a confronto

La Tabella I.3.1 mostra il confronto tra l'imposta dichiarata e l'imposta teorica calcolati dalla CE e dall'AdE per il periodo 2015 - 2019 e pubblicati nel 2020²³.

TABELLA I.3.1: CONFRONTO NELLE COMPONENTI DEL GAP IVA. ANNI 2015 – 2019, VALORI IN MILIONI DI EURO					
Commissione Europea – CASE					
	2015	2016	2017	2018	2019
IVATce	139.703	140.400	142.939	144.772	146.855
IVAVce	100.345	102.086	107.576	109.333	111.793
GAP IVA	39.358	38.314	35.363	35.439	35.062
GAP IVA % iva potenziale	28,2%	27,3%	24,7%	24,5%	23,9%
Agenzia delle entrate					
	2015	2016	2017	2018	2019
IVATade	132.429	133.776	135.562	138.731	138.449
IVAEcade	96.543	98.188	98.638	105.471	110.509
GAP IVA	35.887	35.588	36.923	33.259	27.940
GAP IVA %	27,1%	26,6%	27,2%	24,0%	20,2%
Differenze tra CE e AdE					
	2015	2016	2017	2018	2019
IVATce – IVATade	7.274	6.624	7.377	6.041	8.406
IVAVce – IVAEcade	3.802	3.898	8.938	3.862	1.284
GAP IVA(ce-ade)	3.471	2.726	-1.560	2.180	7.122
GAP IVA % (ce-ade)	1,1%	0,7%	-2,5%	0,5%	3,7%

Fonte: elaborazioni Agenzia delle entrate. *Study and Reports on the VAT Gap in the EU-28 Member States, Annual Report 2020*; Relazione sull'economia non osservata e sull'evasione fiscale e contributive, ed. Novembre 2020.

La Tabella I.3.1 mostra una differenza sistematica positiva sia per l'IVA teorica che per l'IVA versata. Ovviamente tali discrepanze si ripercuotono sul *gap*. Nonostante questo, il trend delle due propensioni è simile ad eccezione del 2017 e del 2019. La notevole discrepanza del 2019, anche in termini di *gap* in valore assoluto (circa 7 miliardi), è riconducibile alla provvisorietà dei dati utilizzati da entrambe le istituzioni, per cui non verrà approfondito in questa sede. Interessante, invece, è la divergenza del 2017 guidata dal picco nella differenza di IVA versata, come calcolata da CE, e di IVAEC, come calcolata da Ade, che implica un'inversione di tendenza nell'andamento del *gap*. Come spiegato nel paragrafo precedente (vedi equazione [2] e [3]), la differenza di calcolo è legata all'inclusione (esclusione) della variazione dello stock di credito e dei proventi dell'accertamento. Per questo motivo, si riporta di seguito il calcolo dell'IVAEC alternativa, ovvero coerente con l'aggregato utilizzato da CE:

$$IVAV_{alternativa} = IVAEC_{ade} + accertamento + \Delta stock\ crediti \quad [4]$$

La Tabella I.3.2 mostra che l'utilizzo dell'aggregato $IVAV_{alternativa}$ permette di avere una discrepanza trascurabile tra i due aggregati che determinano l'IVA dichiarata (versata), mentre permane la differenza in termini di *gap* riconducibile alle discrepanze nel calcolo dell'IVA potenziale.

²³ Ci si riferisce, in particolare a *Study and Reports on the VAT Gap in the EU-28 Member States, Annual Report 2019*; Relazione sull'economia non osservata e sull'evasione fiscale e contributive, ed novembre 2020.

TABELLA I.3.2: CALCOLO DELL'IVA VERSATA "ALTERNATIVA", COERENTE CON EQ. [4]. ANNI 2015 – 2019, VALORI IN MILIONI DI EURO

	2015	2016	2017	2018	2019
Delta stock crediti	853	852	4.522	-743	-3.598
$IVAV_{altern}$	99.782	101.673	106.572	109.032	111.418
$IVAV_{CE} - IVAV_{alternativa}$	563	413	1.004	301	375
GAP alternativo	32.648	32.103	28.990	29.699	27.031
GAP IVA %new	24,7%	24,0%	21,4%	21,4%	19,5%

Tuttavia, il vantaggio di utilizzare la stima di $IVAV_{alternativa}$ permette di ottenere valori di *gap* che presentano un andamento simile nel trend (vedi Tabella I.3.3). Infatti, le serie di $IVAV_{CE}$ e di $IVAV_{alternativa}$ registrano la stessa variazione sia in termini di segno che di impatto nel 2017, mentre nello stesso anno la serie ufficiale dell'Agenzia delle entrate presenta scostamenti leggermente positivi (+0,63 pp), spostando il recupero di *compliance* all'anno successivo (Tabella I.3.3).

TABELLA I.3.3: DIFFERENZE IN TERMINI PERCENTUALI TRA LE SERIE STORICHE DEL GAP %. ANNI 2016-2019

	2016	2017	2018	2019
diff y-o-y CE	-0,88	-2,55	-0,26	-0,60
diff y-o-y Ade	-0,50	0,63	-3,26	-3,79
diff y-o-y Ade(alternativo)	-0,66	-2,61	0,02	-1,88

La traslazione nel tempo della riduzione del *tax gap* dal 2017 (CE) al 2018 (Ade) è dovuta alla mancanza dello stock di credito nel calcolo del *gap* della CE. Lo stock di credito, difatti, subisce una forte riduzione nel 2018 - 2019, come evidenziato dalla Tabella I.3.2.

I.4 METODOLOGIA PER LA STIMA DEL TAX GAP IRAP

Introduzione

L'IRAP, Imposta Regionale sulle Attività Produttive, è caratterizzata da un'ampia platea di contribuenti (circa 3,5 milioni) e, essendo rivolta agli operatori economici che operano dal lato dell'offerta produttiva, permette di analizzare i fenomeni per settore di attività economica. È inoltre ripartita regionalmente in modo analogo agli aggregati di contabilità nazionale (CN), ed è quindi adatta per le analisi territoriali del *gap*.

L'Agenzia delle entrate calcola regolarmente il *gap* IRAP dal 2002²⁴. Le stime sono state pubblicate in diversi documenti istituzionali, tra i quali il rendiconto annuale della Corte dei Conti e l'Allegato alla nota di aggiornamento del documento di economia e finanza.

I principali aspetti della metodologia sono stati pubblicati prima sulla collana «Documenti di Lavoro dell'Ufficio Studi», poi su «Argomenti di discussioni», entrambe disponibili sul sito internet dell'Agenzia delle entrate. I metodi di calcolo sono stati affinati nel corso del tempo, con l'obiettivo di applicare al meglio la normativa fiscale ai dati di contabilità nazionale.

Nel corso del 2016 si è condotta una revisione straordinaria dettata principalmente dalle nuove serie di contabilità nazionale prodotte in accordo con i nuovi standard internazionali (SEC2010)²⁵. Oltre alle revisioni dovute alla variazione nei dati di base, sono state apportate anche delle innovazioni metodologiche grazie alla proficua collaborazione con l'ufficio di Contabilità Nazionale e la Direzione Centrale per le Statistiche Economiche ISTAT.

Nei paragrafi seguenti si definiscono i macro-aggregati (fiscali e di contabilità nazionale) di riferimento, poi si descrivono gli interventi di integrazione necessari alla definizione degli stessi, infine si riportano i dettagli relativi alla metodologia seguita per il passaggio dalla base evasa all'imposta.

²⁴ A proposito si veda: Braiotta et al. (2015), Braiotta et al (2013), Pisani (2014), Pisani S., Polito C (2006), Liberati et al (2002).

²⁵ Si veda Eurostat (2010)

Aggregati di riferimento

L'aggregato cardine di riferimento nella stima del *gap* IRAP è la Produzione lorda (PL). La PL corrisponde al valore aggiunto al costo dei fattori armonizzato con la legislazione fiscale; è un aggregato diverso rispetto alla base imponibile in quanto contiene il costo del lavoro e l'utile di impresa e, pertanto, fornisce un'indicazione indiretta del *gap* sulle altre imposte dirette.

Dal valore aggiunto al costo dei fattori (VACF) di CN si estrae la Produzione Lorda Potenziale (PLP). La PLP è la misura del VACF più omogenea con la PL riportata in dichiarazione. Mentre la PL riflette la realtà economica dichiarata dai contribuenti, la PLP offre una rappresentazione esauriente della realtà economica complessiva, cioè comprensiva anche della componente occultata al fisco. Se definito in maniera corretta, sterilizza gli effetti normativi (aliquote agevolate, deduzioni ecc.) che possono portare a distorcere il confronto tra dati statistici e dati fiscali. Dal confronto tra PLP e PL si ottiene una stima della base imponibile non dichiarata (PLND):

$$PLP-PL=PLND$$

Integrazione tra i dati IRAP e quelli di CN

L'unità elementare di analisi

La principale differenza che esiste tra i dati IRAP e quelli di CN è riconducibile al diverso criterio utilizzato nell'identificazione della unità elementare di analisi; la Contabilità Nazionale, infatti, si avvale del concetto di Unità di Attività Economica Locale (UAEL), mentre le dichiarazioni IRAP utilizzano il concetto di impresa. Come è noto l'impresa si definisce come l'esercizio professionale di un'attività economica avente il fine della produzione e dello scambio di beni e servizi. Nella maggior parte dei casi le imprese svolgono più di un'attività e, al fine di evidenziare le relazioni tecnico-economiche, in contabilità esse devono essere disgiunte. Ciò si ottiene utilizzando il concetto di UAEL che raggruppa tutte le parti dell'impresa che si trovano in un'unica sede e concorrono all'esercizio di un'attività al livello di classe (4 cifre). Pertanto, per le imprese che esercitano diverse attività si creano delle discrepanze poiché, ai fini dell'IRAP si classificano secondo l'attività prevalente, mentre in CN vengono suddivise in base alle singole attività economiche svolte.

Al fine di correggere tali differenze si è fatto ricorso ai dati dell'archivio statistico ISTAT FRAME-SBS²⁶. Utilizzando questa fonte è stato possibile costruire una matrice quadrata M , di 49 righe e 49 colonne, contenente il valore aggiunto classificato per impresa (sulle righe) e per unità funzionale (sulle colonne).

Dato il vettore V , di dimensioni $(49,1)$, contenente i dati di valore aggiunto di contabilità nazionale per 49 branche di attività, si ottiene il vettore V^* contenente il valore aggiunto di CN riclassificato per impresa tramite il prodotto matriciale

$$V^* = M_X V$$

Questa uguaglianza vale sotto l'ipotesi che l'unità funzionale approssimi la UAEL.

Da un'analisi dei microdati sui CF che hanno presentato dichiarazione IRAP per l'anno di imposta 2013 presenti anche nell'archivio statistico ISTAT FRAME-SBS è stato possibile costruire una matrice quadrata Q , le cui righe contengono il VA ISTAT aggregato per settore di attività economica e le cui colonne contengono il Valore della produzione IRAP aggregato per settore di attività economica.

²⁶ Il confronto è il frutto dell'attività svolta in un gruppo di lavoro congiunto ISTAT e Agenzia delle entrate.

La diagonale principale di Q misura l'ammontare di VA prodotto dalle imprese classificate analogamente da ISTAT e ADE; altrove le frequenze rappresentano l'ammontare prodotto da imprese per le quali non c'è analogia di classificazione.

Dato il vettore V^* , contenente i dati VA di CN per branche di attività, riclassificati per impresa, tramite il prodotto matriciale tra V^* e Q_n (matrice che si ottiene dopo aver normalizzato per i totali di riga Q)

$$Vc^* = Q_n \times V^*$$

si ottiene il vettore Vc^* contenente il VA di CN riclassificato per impresa e corretto per gli errori di classificazione.

Le modifiche ai dati IRAP

Per assicurare la coerenza nelle classificazioni e nelle definizioni utilizzate dall'amministrazione fiscale con quelle adottate dall'ISTAT si è reso necessario effettuare gli aggiustamenti sui dati fiscali che sono di seguito elencati.

Ammortamenti

Per quanto riguarda gli ammortamenti il confronto sul valore aggiunto evidenzia un effetto distorsivo in merito alla loro deducibilità. Dall'esame dei quadri IRAP, infatti, gli ammortamenti risultano esclusi dalla produzione lorda poiché iscritti tra le componenti negative, mentre in CN concorrono alla determinazione del valore aggiunto.

Dal momento che la definizione statistico-economica degli ammortamenti differisce sostanzialmente da quella fiscale, al fine di non introdurre una distorsione nel confronto con i dati di CN, le spese relative agli ammortamenti dichiarati dai soggetti vengono aggiunte alla PL riportata in dichiarazione, ovvero:

$$PL = \text{Componenti positive} - \text{Componenti Negative} + \text{Ammortamenti}$$

A titolo di esempio, per gli anni 2010-2014, in media l'ammontare degli ammortamenti dichiarati è stato pari a circa 113,5 miliardi di euro, il 13,2% della PL complessiva media dello stesso periodo.

GEIE e filiali di imprese estere

La seconda, correzione apportata sui dati fiscali si è resa necessaria per approssimare il concetto di residenza di Contabilità Nazionale. Il PIL, e conseguentemente il valore aggiunto, misura il reddito prodotto dalle unità residenti.

Al fine di approssimare il concetto di residenza seguito dalla Contabilità Nazionale, nei dati IRAP sono stati inclusi i Gruppi di Interesse Economico Europeo (GEIE) che hanno sede in Italia, mentre sono state escluse le filiali estere di imprese italiane (EE).

Servizi postali

Il VACF di Poste Italiane per l'attività di **servizi postali** viene classificato dalla CN nella branca di poste e telecomunicazioni, in IRAP invece in quello degli **Ausiliari del Credito**. Per non creare effetti distorsivi nel confronto tra PLP e PL questa posta, nell'aggregato di PL, viene sottratta alla branca dell'intermediazione e aggiunta a quella di Poste e Telecomunicazioni.

Il VACF di CN fuori dal campo di applicazione dell'IRAP

In questo paragrafo si illustrano le correzioni apportate ai dati di VACF al fine di renderli omogenei con quelli IRAP. Tali interventi riguardano l'esclusione dall'aggregato di CN di quelle voci che, coerentemente con la normativa fiscale, sono escluse dal campo di applicazione IRAP

(tali voci non concorrono alla determinazione della base imponibile per gli operatori economici pur partecipando alla determinazione del VACF nazionale).

Le esclusioni dall'aggregato di CN riguardano:

- **Agricoltura:** il VACF di CN di questa branca viene di fatto azzerato dal periodo di imposta 2016 in poi. Con la legge di stabilità 2016 è stata infatti prevista l'esenzione IRAP per tutto il settore agricolo, fatti salvi i produttori agricoli che svolgono attività dirette anche alla produzione di servizi (tipo agriturismi).
- **Economia illegale:** i proventi sono fuori dal campo di applicazione del tributo e si collocano nei settori del commercio e degli altri servizi.
- **Ricerca e Sviluppo** (la correzione riguarda tutti i settori); in contabilità nazionale sono considerate spese d'investimento e, quindi, concorrono a incrementare il valore aggiunto, mentre ai fini dell'IRAP sono costi intermedi riducendo la PLP.
- **Servizi domestici e mance:** i primi sono fuori dal campo di applicazione del tributo e le seconde non concorrono ad incrementare la PL.
- **Produzione per uso finale non market:** fuori dal campo di applicazione del tributo.
- **Fitti:** i fitti effettivi percepiti dalle famiglie (**branca 68**) e quelli figurativi (somme imputate alle famiglie che vivono in alloggi di proprietà, **branca 69**) non costituiscono base imponibile IRAP.
- **Diritti d'autore:** sono deducibili come costi ai fini IRAP e non vengono dichiarati da chi li percepisce. È, pertanto, necessario dedurre dai dati di CN la specifica voce.
- **Minimi e autonoma organizzazione:** il VA per settore di attività economica attribuito ai contribuenti minimi e a quelli che svolgono un'attività economica «non autonomamente organizzata» ossia un'attività in cui «non esistono capitali investiti e/o lavoro apportato da terzi che, coordinati dal titolare, consentono di produrre un reddito aggiuntivo rispetto al lavoro del titolare stesso» (D.lgs. 10/04/1998 n.137). L'Agenzia delle entrate ha approntato un'apposita stima per quantificare il fenomeno.
- **Non residenti:** il valore aggiunto generato dalle imprese non residenti alle quali è riconosciuto il presupposto di stabile organizzazione viene considerato nel VA di CN. Dal punto di vista fiscale le imprese non residenti possono avvalersi di specifici accordi internazionali per evitare la doppia imposizione fiscale. Tali accordi prevedono che «le imprese non residenti sono tassabili esclusivamente nello Stato contraente in cui è situata la sede della direzione dell'impresa». Le imprese non residenti, quindi, pur avendo stabile organizzazione in Italia, non generano PL, perché esentate dal presentare dichiarazione IRAP (ma sono presenti nel VA di CN). Si rende necessaria, pertanto, un'operazione di correzione della PLP per il VA di tali imprese presente nella banca dati FRAME-SBS. In particolare, per ciascun settore di attività economica, alla PLP va sottratto il VA di queste ultime.

Le modifiche ai dati di CN

In questo paragrafo si illustrano gli interventi di integrazione/riclassificazione di quelle voci che concorrono alla determinazione di base imponibile ma che in CN non vengono considerate o vengono classificate in maniera non coerente con quanto fatto dal lato fiscale.

Gli interventi di integrazione/riclassificazione riguardano:

- **Imposta IRAP:** il SEC classifica l'IRAP tra le imposte indirette sull'attività produttiva, conseguentemente la sottrae per calcolare il VACF. Per coerenza con il dato fiscale l'imposta versata per settore di attività economica è sommata al VACF per la determinazione della PLP da confrontare con la PL.
- **Servizi di intermediazione Finanziaria Indirettamente Misurati (SIFIM):** in CN una parte di questi servizi è considerata costo intermedio delle imprese non finanziarie. Dato che i SIFIM non sono inclusi nelle componenti negative che concorrono alla determinazione della PL, per rendere confrontabili gli aggregati, tale valore è aggiunto ai costi intermedi delle famiglie produttrici e delle società non finanziarie.

- **Servizi assicurativi:** il valore aggiunto di CN e la base imponibile IRAP delle imprese di assicurazioni si determina in maniera analoga. Tuttavia persiste una distorsione per quanto riguarda il trattamento della componente relativa ai **servizi assicurativi ramo danni**, che viene destinata da ISTAT agli impieghi intermedi delle imprese contraenti. Dato che tale imputazione non trova riscontro ai fini IRAP, tale componente viene aggiunta al VA settoriale di CN.
- **Lavoratori interinali:** le remunerazioni dei lavoratori interinali non sono base imponibile IRAP. In CN rappresentano un ricavo per le agenzie d'intermediazione (che operano nel settore servizi alle imprese) ed un uguale costo per le aziende che li utilizzano (che possono operare in tutti i settori). Tale metodo di contabilizzazione non genera distorsioni nel confronto tra gli aggregati complessivi, ma a livello di settore di attività economica (sovrastima della differenza tra PLP e PL nel settore dei servizi alle imprese e uguale sottostima nel confronto complessivo tra gli altri settori). Per tale motivo nel calcolo della PLP si sottrae il valore delle remunerazioni degli interinali dalla branca dei servizi alle imprese e si ridistribuisce proporzionalmente nelle altre branche.
- Al fine di eliminare i motivi di distorsione dovuti alle differenze nelle modalità di contabilizzazione, il VA di CN è stato corretto sostituendo il Valore Aggiunto non market con quello non market dichiarato ai fini IRAP.
- Analoga operazione, per gli stessi motivi, viene svolta per quanto riguarda la branca dell'intermediazione monetaria e finanziaria.

L'ipotesi, plausibile, alla base di queste ultime due correzioni è relativa al fatto che tali settori possono essere considerati immuni da evasione.

Analisi di coerenza sui risultati ottenuti

A seguito delle varie correzioni fin qui apportate, si possono riscontrare in alcuni casi risultati incoerenti (il dato IRAP risulta superiore a quello di CN). Tale risultato anomalo si osserva in particolare in alcune branche del settore dell'industria. Ciò fa presupporre che nelle attività dell'industria si annidino attività secondarie che non è stato possibile scorporare tramite le correzioni illustrate in precedenza. Per ovviare a tale inconveniente, si è reso necessario operare una stima alternativa dell'evasione che consiste nel sostituire, solamente nei settori in cui sono presenti dati anomali, i dati IRAP con quelli ISTAT sull'evasione da lavoro irregolare. Data la validità della stima dell'evasione per il totale dell'economia, le nuove stime così ottenute non si aggiungono alle precedenti, ma vengono compensate da diminuzioni proporzionali dell'evasione nei settori rimanenti.

Un riscontro quantitativo delle correzioni apportate

La Tabella I.4.1 riporta la quantificazione dell'impatto (assoluto e in percentuale) di ciascuna delle correzioni descritte nei paragrafi precedenti. Oltre al valore è importante considerare anche il segno. Le correzioni fatte per i SIFIM, i premi supplementari delle assicurazioni e l'imposta IRAP accrescono l'aggregato di partenza. La sostituzione del dato fiscale con quello statistico, per il settore non market e quello del credito, non ha un segno predefinito in termini di impatto sul VACF. Infatti, l'impatto è negativo se il dato di CN è maggiore di quello fiscale (come è accaduto nel 2011), positivo viceversa. Per quanto riguarda le altre correzioni (seconda parte della Tabella I.4.1), essendo tutte voci che concorrono a generare valore aggiunto, ma sono escluse dal campo di osservazione IRAP, hanno tutte un impatto negativo e vanno quindi a ridurre il VACF.

TABELLA I.4.1. RISCONTRO QUANTITATIVO DELLE CORREZIONI APPORTATE AL VALORE AGGIUNTO DI CONTABILITÀ NAZIONALE PER OTTENERE PLP

		VACF SEC 2010 (ed Sett.2016)	Impatto %
	VACF	1.432.941	
+	Sifim e premi supplementari	34.759	2,4%
+	Imposta IRAP	32.669	2,3%
Δ	Non Market	-119.452	8,3%
Δ	Credito	-45.787	3,2%
		-97.811	-6,8%
-	Agricoltura (minimi e autoproduzione)	-5.173	0,4%
-	Droga e contrabbando	-11.937	0,8%
-	Prostituzione	-3.549	0,2%
-	R&S	-10.825	0,8%
-	Servizi domestici, mance e diritti d'autore	-23.987	1,7%
-	Prod. per uso finale	-1.241	0,1%
-	Fitti figurativi	-129.015	9,0%
-	Fitti effettivi	-38.306	2,7%
-	Minimi e "autonoma organizzazione"	-11.709	0,8%
-	Trasporti imprese estere	-1.816	0,1%
		-237.288	16,6%
	PLP	1.097.842	76,6%

Dati in milioni di euro

Nel 2018 l'aggregato PLP è risultato pari a 1.129.678, milioni di euro che corrisponde al 73% del VACF. Questo significa che l'impatto complessivo degli aggiustamenti ha comportato una riduzione del VACF di circa il 27%.

La regionalizzazione dei dati

I dati di fonte fiscale

Un indiscutibile punto di forza della base informativa IRAP risiede nella disponibilità di informazioni a livello regionale. Nel modulo di dichiarazione, infatti, si richiede a tutte le imprese pluriregionalizzate di ripartire l'ammontare della base IRAP tra i differenti stabilimenti in base ai dati del costo del lavoro, ottenendo, in tal modo, una ripartizione territoriale analoga a quella effettuata dalla CN. Il valore aggiunto al costo dei fattori derivante dalle dichiarazioni IRAP è stato, quindi, regionalizzato utilizzando la voce "valore della produzione al netto delle deduzioni". Quest'ultima rappresenta il valore che, a livello territoriale, approssima maggiormente il valore della produzione lorda (assunto come base di calcolo del valore aggiunto fiscale). Si discosta da esso solo in virtù delle deduzioni per contributi per le assicurazioni obbligatorie, spese relative agli apprendisti, deduzioni per cooperative sociali e deduzioni per lavoratori dipendenti. L'ipotesi adottata è che queste spese presentino una distribuzione omogenea sul territorio. Per le imprese localizzate in differenti regioni, gli ammortamenti sono stati disaggregati in base al corrispondente valore della produzione.

I dati di fonte statistica

Il primo problema che si deve affrontare riguarda il fatto che a livello regionale l'Istat difonde unicamente il valore aggiunto ai prezzi base. Per realizzare il confronto con i dati IRAP è stato necessario richiedere una elaborazione ad hoc alla contabilità nazionale relativa al valore aggiunto al costo dei fattori. Inoltre, i conti regionali di CN presentano una ventunesima regione denominata extra-regio. Questa comprende, essenzialmente, la produzione delle piattaforme petrolifere marine. Il valore aggiunto corrispondente è stato assegnato alla regione prospiciente le piattaforme stesse. Inoltre, si è reso necessario riportare a livello regionale alcune delle poste correttive evidenziate per il calcolo nazionale. Tale disaggregazione è stata effettuata utilizzando degli indicatori disponibili a livello regionale, in particolare:

- per i fitti (sia imputati che effettivi) si è utilizzato il dato derivato dal patrimonio abitativo presente nel 1999;
- per il VA delle attività non market la CN ha fornito il dato disaggregato per regione, questo dato è stato ripartito per branca, all'interno di ciascuna regione, utilizzando la distribuzione nazionale.

La procedura di quadratura dei dati

Il livello di dettaglio scelto per effettuare il calcolo regionale è quello massimo consentito dalle stime di contabilità nazionale, cioè 22 branche di attività economica. Considerando 20 regioni, sono state quindi, prodotte 440 stime dell'evasione per ciascun anno preso in considerazione. Sulle stime ottenute si è riscontrato un certo numero di risultati non plausibili, cioè di branche che all'interno delle regioni presentavano un valore IRAP superiore a quello di Contabilità Nazionale. Il problema è analogo a quello evidenziato in precedenza, e, anche in questo caso, si è reso necessario correggere i dati anomali. A tal fine è stata approntata una specifica procedura articolata nei seguenti passi:

- a) si sostituisce il dato anomalo con una stima "iniziale" dell'evasione;
- b) si vincolano le stime dell'evasione totale nazionale per branca e per regione;
- c) si applica la quadratura che garantisce l'assenza di risultati implausibili e che consente di rispettare i vincoli imposti al punto precedente;
- d) le correzioni proposte dall'algoritmo sono applicate al dato di contabilità nazionale, lasciando inalterati i dati IRAP.

Le stime così ottenute costituiscono la base di partenza per le elaborazioni provinciali.

Il passaggio dalla produzione lorda non dichiarata al *gap* IRAP complessivo²⁷

In precedenza si sono illustrati i passaggi per determinare la produzione lorda potenziale, PLP, partendo dai dati di contabilità nazionale. In base all'identità illustrata prima sappiamo che la PLP è uguale alla produzione dichiarata (PL) più la produzione non dichiarata (PLND). Nella presente sezione si illustrano i metodi di calcolo per derivare le imposte che scaturiscono dalle due produzioni, ovvero quella riportata in dichiarazione (ID) e quella occultata al fisco (TGND).

L'imposta dichiarata (ID) è calcolata dal contribuente applicando l'aliquota di legge (q) alla base imponibile, in termini formali:

$$ID=(PL - \beta_1 PL - \beta_2 PL)q \quad 1)$$

dove (β_1) rappresenta la quota degli ammortamenti, rispetto a PL, e (β_2) quella delle deduzioni sul medesimo aggregato.

Analogamente TGND potrebbe essere calcolata come:

$$TGND=(PLND - \lambda_1 PLND - \lambda_2 PLND)q \quad 2)$$

dove, in base alle caratteristiche del tributo, l'aliquota q è uguale a quella utilizzata per il dichiarato, (λ_1) e (λ_2) sono le rispettive quote relative agli ammortamenti e alle deduzioni che si potrebbero portare a riduzione della PLND.

La 1) si può riscrivere:

²⁷ Le operazioni descritte nel documento fanno riferimento a valori misurati o osservati ad un dettaglio regionale (20 regioni italiane) e settoriale (6 settori di attività economica).

$$ID = PL \cdot q \cdot (1 - \beta_1 - \beta_2) \quad 3)$$

L'aliquota implicita media utilizzata nel calcolo, α , è pari al rapporto tra Imposta dichiarata e la produzione lorda dichiarata.

$$\alpha = \frac{ID}{PL} \quad 4)$$

TGND è pari alla produzione lorda non dichiarata (PLND) moltiplicata per α

$$TGND = \alpha \cdot PLND \quad 5)$$

Considerando le espressioni 3) e 4) possiamo riscrivere la 4) come:

$$\begin{aligned} TGND &= \frac{ID}{PL} \cdot PLND = \\ TGND &= \frac{PLND}{PL} \cdot PL \cdot q \cdot (1 - \beta_1 - \beta_2) = \\ TGND &= PLND \cdot q \cdot (1 - \beta_1 - \beta_2) \end{aligned} \quad 6)$$

L'ultima espressione della 6), concettualmente simile alla 3), ci dice che non esistono possibili effetti distorsivi nel calcolo di TGND se e solo se $\lambda_1 = \beta_1$ e $\lambda_2 = \beta_2$. L'ipotesi è plausibile, soprattutto quella relativa alle deduzioni: è difficile, ad esempio, che un contribuente tenda a chiedere deduzioni relative a base non dichiarata (ho un dipendente in nero, non lo dichiaro, ma ne chiedo le deduzioni), anche in virtù del fatto che se gli viene accertata l'evasione quelle deduzioni gli vengono comunque riconosciute.

L'utilizzo dell'aliquota implicita α implica anche che qualsiasi normativa che agisca sulle deduzioni e qualsiasi variazione dell'incidenza degli ammortamenti sulla base imponibile IRAP non si traduca in una variazione di PL o PLP, ma bensì si rifletta in una variazione dell'aliquota.

A titolo di esempio, la tabella I.4.2 riporta i valori dell'aliquota implicita media calcolata per il totale dell'economia negli anni 2010-2014.

TABELLA I.4.2 ALIQUOTE IMPLICITE MEDIE. TOTALE ECONOMIA ANNI 2010-2014					
Anno	2010	2011	2012	2013	2014
Aliquota	3,7%	3,8%	3,7%	3,6%	3,3%

L'effetto dell'innovazione normativa relativa all'introduzione delle deduzioni del costo del lavoro che, entrando in vigore a partire da novembre 2014, hanno iniziato a produrre effetti in termini di minore imposta dichiarata (circa 4 miliardi di euro), si riflettono nella riduzione dell'aliquota implicita del 2014. La variazione dell'aliquota implicita sterilizza gli effetti distorsivi che si avrebbero se si utilizzasse l'aliquota di legge su una PLND che è comprensiva anche della quota degli ammortamenti e delle deduzioni.

La somma di TGND e di ID misura il valore dell'imposta potenziale (IP).

$$IP = TGND + ID \quad 7)$$

Esiste tuttavia un altro fenomeno che bisogna considerare per quantificare l'evasione complessiva: l'imposta regolarmente dichiarata, ID, è uguale alla somma dell'imposta dichiarata e versata (IV) più quella dichiarata e non versata (INV). Non tutta l'imposta dichiarata, infatti, viene poi regolarmente versata. Alla luce di questa evidenza è necessario calcolare l'imposta potenziale totale, IP, tramite la 8):

$$IP = TGND + IV + INV \quad 8)$$

TGND più INV rappresenta il di *Gap* IRAP complessivo.

$$GAP\ IRAP = TGND + INV$$

La componente INV si calcola sulla base delle risultanze dei controlli automatizzati delle dichiarazioni.

I.5 METODOLOGIA PER LA STIMA DEL GAP RELATIVO AI REDDITI DA LOCAZIONE

I proventi da locazione intervengono nel calcolo del reddito complessivo IRPEF²⁸ in misura pari al 95% dei canoni percepiti (85% fino all'anno d'imposta 2011), con l'eccezione dei redditi per i quali si è optato per il regime della "Cedolare secca" (si veda oltre). Ai fini IRES, invece, il canone di locazione può essere ridotto fino ad un massimo del 15% dello stesso, in relazione a spese documentate di manutenzione ordinaria (art. 90 del TUIR).

A decorrere dall'anno 2011²⁹, il legislatore ha apportato una consistente innovazione con l'introduzione della "Cedolare secca", un regime opzionale ai fini IRPEF applicabile ai canoni di locazione percepiti in virtù di contratti aventi ad oggetto immobili ad uso abitativo (e le relative pertinenze locate congiuntamente all'abitazione). Nella formulazione originaria, pertanto il regime non era applicabile alle locazioni di unità immobiliari effettuate nell'ambito dell'esercizio di una attività d'impresa o di arti e professioni.

Tale regime prevede l'applicazione di un'aliquota fissa (quella ordinaria è fissata al 21%) al totale dei canoni percepiti. L'imposta che ne risulta sostituisce, oltre all'IRPEF e relative addizionali, anche le imposte di registro e di bollo (ivi comprese quelle dovute in caso di risoluzione e proroga del contratto, mentre resta dovuta l'imposta di registro per la cessione del contratto di locazione). Per i contratti a canone concordato (L. 431/98) è prevista l'applicazione di un'aliquota agevolata, il cui importo è stato ridotto nel corso del tempo: 19% (2011-2012); 15% (2013); 10% (2014-2019).

Il provvedimento ha dunque coniugato semplificazione e riduzione del peso fiscale (in quanto l'aliquota ordinaria è inferiore rispetto all'aliquota media pagata dai percettori di redditi di locazione). Inoltre, l'intervento normativo complessivo era associato ad un inasprimento delle sanzioni in caso di evasione, in quanto la previsione originaria, poi modificata a seguito dei pronunciamenti giudiziari³⁰, introduceva la possibilità della registrazione unilaterale del contratto da parte dell'inquilino, con contestuale determinazione di un canone ridotto *ex lege*.

Per determinare il *gap* relativo ai canoni di locazione, si procede al confronto degli importi dichiarati con il dato della produzione Istat per la branca 68 ("Affitto e gestione di immobili propri"), che corrisponde ai canoni di affitto percepiti dai locatori, disaggregato per settore istituzionale (società non finanziarie, famiglie e Amministrazioni pubbliche).

Coerentemente con l'approccio di calcolo top-down, la base potenziale è desunta dai dati Istat; a quest'ultima si sottraggono gli importi dichiarati al fisco al fine di ottenere il *gap* nella base imponibile. Tale confronto viene effettuato limitatamente alle persone fisiche, in quanto per le società il relativo *gap* è incluso nel calcolo del *gap* da imposte sul reddito, IRES o IRPEF. La distinzione tra PF titolari e non titolari di Partita IVA in questa sede viene omessa in quanto l'imprenditore individuale può locare fabbricati di sua proprietà al di fuori dell'attività d'impresa.

La base dichiarata (BID) è ricavata dai dati fiscali già disaggregata per forma giuridica del soggetto dichiarante. La percentuale di abbattimento effettiva (ovvero il rapporto tra reddito imponibile e corrispondente canone) è calcolata sulla base dei dati dei canoni dichiarati distinti per tipologia di utilizzo e codice canone. La stessa percentuale di abbattimento viene applicata

²⁸ Art. 37 del TUIR.

²⁹ Decreto legislativo 14 marzo 2011, n. 23.

³⁰ Da ultimo, la Sentenza 50/2014 della Corte Costituzionale.

ai canoni non dichiarati per ottenere i corrispondenti redditi (ad eccezione dell'ipotesi di applicazione della cedolare secca, nel qual caso il canone coincide con la base imponibile).

I canoni non dichiarati (CND) sono ottenuti come differenza tra produzione ISTAT (canoni totali) e canoni dichiarati dalle PF. La metodologia attuale, pertanto, non distingue tra fabbricati residenziali e non residenziali.

Il *gap* di imposta relativo ai canoni di locazione percepiti dalle persone fisiche è calcolato moltiplicando i canoni evasi per l'aliquota. Con l'introduzione della cedolare secca, a partire dall'anno 2011, si formulano due ipotesi alternative per il calcolo del *gap*: la prima è che ai canoni non dichiarati si applichi l'aliquota implicita calcolata sul dichiarato, che risulta essere una media ponderata tra l'aliquota della cedolare secca e l'aliquota media calcolata sulle sole dichiarazioni IRPEF contenenti redditi da locazioni; la seconda è che a tutti i canoni evasi si applichi il regime della cedolare secca, in quanto più conveniente. In questo modo si costruisce un intervallo delle stime dove la prima ipotesi rappresenta il limite superiore del *gap* e la seconda quello inferiore. Al fine di offrire una sintesi delle due stime, i risultati esposti nella relazione rappresentano la media tra le due ipotesi.

Per calcolare la corrispondente propensione al *gap*, si considera come imposta dichiarata la quota di imposta IRPEF riconducibile ai redditi da locazione, cui dal 2011 si somma l'importo della cedolare secca.

I.6 METODOLOGIA PER LA STIMA DEL TAX GAP DEL CANONE RAI

Introduzione

Chiunque detenga uno o più apparecchi atti o adattabili alla ricezione dei programmi televisivi deve per legge pagare il canone RAI³¹. Trattandosi di un'imposta sulla detenzione dell'apparecchio, il canone deve essere pagato indipendentemente dall'uso del televisore o dalla scelta delle emittenti televisive. Esistono due tipi di canone: uno denominato "ordinario", dovuto da chi detiene apparecchi televisivi in ambito familiare³², e uno denominato "speciale", dovuto da quei soggetti che detengono tali apparecchi in locali aperti al pubblico al di fuori dell'ambito familiare o comunque li utilizza a scopo di lucro diretto o indiretto³³.

Per la riscossione del canone la RAI opera sulla base di un rapporto convenzionale con l'Agenzia delle entrate, e la competenza esclusiva in materia è della Direzione provinciale 1 di Torino - Ufficio canone TV. Alla Direzione canone della RAI è demandato, invece, il compito di accertare il rispetto degli obblighi di legge provvedendo, nel contempo, alla gestione dell'utenza e al recupero della morosità. Gli abbonamenti alla televisione ad uso privato vengono monitorati attraverso statistiche aggiornate con periodicità mensile e ogni anno si disegna un quadro generale della situazione fornendo (a livello nazionale, regionale, provinciale e comunale) la consistenza relativa al numero di abbonati³⁴.

Di seguito si riporta la metodologia seguita per la stima del *tax gap* del solo canone ordinario.

La procedura di stima del *tax gap*

Per determinare la consistenza del *tax gap* relativo al canone ordinario, si opera attraverso una metodologia tesa ad individuare la base potenziale di famiglie soggette al pagamento del canone e relativa imposta dovuta a cui poi vengono sottratti gli effettivi abbonamenti iscritti a ruolo e il canone versato.

³¹ Regio Decreto Legge 21 febbraio 1938, n. 246.

³² Il canone si paga una sola volta per famiglia anagrafica (a condizione che i familiari abbiano la residenza nella stessa abitazione), indipendentemente dal numero di apparecchi detenuti e di abitazioni a disposizione della famiglia stessa.

³³ Art.27 Rdl 21 febbraio 1938, n.246, convertito nella Legge 4 giugno 1938, n.880, e art.2 D.L.Lt. 21 dicembre 1944, n.458.

³⁴ A tal fine la Direzione Canone si è avvalsa della collaborazione scientifica del Prof. Bruno Chiandotto.

Le principali grandezze prese in considerazione sono:

- Popolazione, distinta in numero di abitanti e di famiglie;
- Utenti soggetti a canone - numero degli utenti che in teoria possiedono un apparecchio TV;
- Abbonati iscritti a ruolo - numero degli utenti per i quali è stato accertato il possesso di un apparecchio TV;
- Paganti - numero di abbonati in regola con i pagamenti nel periodo di competenza considerato;
- Morosi - numero di abbonati non in regola con i pagamenti nel periodo di competenza considerato.

Alla stima delle famiglie soggette al pagamento del canone si perviene sottraendo al dato sulle famiglie residenti in abitazione al primo gennaio di ogni anno, le famiglie che hanno dichiarato di non possedere un apparecchio televisivo nell'anno precedente (dati forniti dall'Istat) e il numero di famiglie coabitanti, calcolate considerando le percentuali accertate alla data del censimento 2011.

In termini formali:

$$FSC_t = FRE_t * (1 - FCO_{11}) * FTV_{t-1}$$

dove:

FSC_t = Famiglie soggette a canone nell'anno t

FRE_t = Famiglie residenti nell'anno t

FCO_{11} = Tasso famiglie coabitanti al censimento 2011

FTV_{t-1} = Tasso di famiglie che possiedono una tv nell'anno t-1

Riguardo ai dati utilizzati occorre precisare che:

- i dati sul possesso di un apparecchio TV sono forniti a livello regionale;
- il dato sulle famiglie residenti in abitazione al 1° gennaio di ogni anno risente delle mancate cancellazioni anagrafiche, una sovrastima che si accentua allontanandosi dall'anno di censimento;
- allontanandosi dall'anno di censimento, anche le percentuali di famiglie coabitanti accertate al censimento possono subire delle variazioni.

La differenza tra il numero delle famiglie soggette a canone e il numero degli abbonati iscritti a ruolo fornisce il numero di evasori.

$$FEV_t = FSC_t - FIR_t$$

dove:

FEV_t = Famiglie che evadono il canone nell'anno t

FSC_t = Famiglie soggette a canone nell'anno t

FIR_t = Famiglie iscritte a ruolo nell'anno t

Per passare dal numero di famiglie evasori al *gap* monetario, si calcola dapprima un canone annuo medio, dato dal rapporto fra il canone complessivamente versato e il numero di utenti paganti (dato già in possesso della RAI). Tale importo viene poi moltiplicato per il numero di famiglie che evadono il canone nell'anno t (comprendendo anche le famiglie morose), pervenendo in tal modo al valore monetario del *gap*.

Infine, si calcola la propensione al *gap*, data dal rapporto fra *gap* e gettito potenziale, che è la somma del canone effettivamente versato dagli utenti e il *gap* stesso.

Le modifiche introdotte nel 2016 per la riscossione del canone

Per quanto concerne la modalità di pagamento del canone, fino al 2015 la più diffusa era il bollettino postale. Si poteva, tuttavia, optare anche per altre modalità: in tabaccheria, in banca, tramite telefono o internet e, in presenza di determinati requisiti, mediante addebito diretto sulla pensione. Dal 2016 (articolo 1, commi da 152 a 159, della Legge n. 208 del 2015):

- è stata introdotta la presunzione di detenzione dell'apparecchio televisivo nel caso in cui esista un'utenza per la fornitura di energia elettrica nel luogo in cui una persona ha la propria residenza anagrafica;
- i titolari di utenza elettrica per uso domestico residenziale effettuano il pagamento del canone mediante addebito nella fattura dell'utenza di energia elettrica in dieci rate mensili di pari importo, da gennaio a ottobre di ogni anno (per il 2016, anno di avvio del nuovo sistema di riscossione, nel mese di luglio sono state addebitate cumulativamente le prime sette rate del canone);
- per le famiglie in cui nessun componente della famiglia anagrafica è titolare di utenza elettrica residenziale e per gli utenti per i quali l'erogazione dell'energia elettrica avviene nell'ambito di reti non interconnesse con la rete di trasmissione nazionale (isole), il pagamento del canone deve essere effettuato mediante modello F24 (Regolamento n. 94 del 2016).

Per stimare il numero di famiglie soggette a canone RAI, dal 2016 è stata impiegata la stessa metodologia utilizzata nel 2015: sono state considerate quindi le famiglie in abitazioni depurate dalle coabitazioni e dal non possesso di un apparecchio televisivo³⁵. È necessario, però, riformulare le statistiche in base ai nuovi flussi di dati a disposizione, per cui al numero degli utenti TV effettivi si perviene sommando le varie categorie coinvolte: titolari di utenza di fornitura elettrica ad uso domestico residenziale (abbonati tv elettrici), F24 individuali, esenzioni *over 75*, pensionati rateizzati, c/c SAT; si ottiene così il nuovo universo di riferimento degli utenti riconducibile alla dicitura "Iscritti a ruolo". Tale dato sconta il numero delle dichiarazioni di non detenzione di un apparecchio TV (DND quadro A) e delle dichiarazioni sostitutive di non addebito del canone per la presenza di altra utenza elettrica (quadro B).

I.7 METODOLOGIA PER LA STIMA DELL'EVASIONE CONTRIBUTIVA

Il calcolo dei contributi evasi si basa sull'applicazione di apposite aliquote contributive alla base imponibile potenziale, definita come indicato nel seguito. Per i contributi a carico del datore di lavoro si utilizzano le aliquote contributive implicite nel calcolo delle stime dei redditi da lavoro dipendente (regolari) di Contabilità nazionale. Tali aliquote implicite vengono calcolate a partire dai dati di base (statistiche strutturali sulle imprese contenute nel registro statistico sui risultati economici delle imprese - *Frame-SBS*) prima delle riclassificazioni delle voci del costo del lavoro secondo i criteri del SEC e rispecchiano, quindi, le aliquote effettive registrate nei conti delle imprese. È da notare che tale stima costituisce un'approssimazione rispetto all'insieme di elementi che possono di fatto incidere sul livello delle aliquote³⁶.

La stima dei contributi a carico del lavoratore si basa sulle aliquote di legge, in quanto tali contributi non vengono registrati nei conti delle imprese e non si dispone, quindi, di informazioni dirette. Peraltro, la relativa base imponibile non corrisponde esattamente alla retribuzione; è

³⁵ Al riguardo va precisato che dal 2016, disponendo di dati aggiornati, non sono state considerate le famiglie accertate al 1° gennaio ma si è proceduto al calcolo della media del numero delle famiglie residenti in abitazione registrate al 1° gennaio e al 31 dicembre e che il dato sul possesso di un televisore non è quello riferito all'anno precedente ma quello stimato dall'Istat per l'anno corrente.

³⁶ Le aliquote effettive possono differire da quelle di legge per due motivi: i contributi assicurativi variano anche a parità di settore di attività economica e possono essere influenzati da accordi aziendali. Inoltre, le retribuzioni contabili non sono esattamente uguali all'imponibile contributivo.

stato, pertanto, stimato un fattore di correzione (che nel seguito sarà indicato con K) per tenere conto di tale differenza³⁷.

Le due stime di evasione contributiva si basano su due ipotesi alternative riguardo alla retribuzione lorda imponibile, che definiscono un intervallo di valori plausibili. La prima consiste nell'ipotizzare che nel caso in cui i lavoratori irregolari fossero regolarizzati essi dovrebbero essere pagati quanto i lavoratori regolari, per i quali esistono minimi contributivi al di sotto dei quali il datore di lavoro non può scendere. Di conseguenza i contributi vengono stimati applicando le aliquote di legge a un costo del lavoro orario per il lavoratore irregolare regolarizzato posto uguale a quello regolare, a parità di settore di attività economica e di classe dimensionale di impresa. Tale assunzione implica un aumento di costo del lavoro per il datore di lavoro e costituisce, pertanto, una soglia di massima del livello dell'evasione contributiva.

Nel dettaglio, sotto tale ipotesi, la procedura di calcolo è la seguente:

$$base_{b,c} = r_{bc} * hi_{bc}$$

dove:

$base_{b,c}$ = base retributiva imponibile stimata; r_{bc} = valore orario della retribuzione lorda dei regolari; hi_{bc} = numero di ore lavorate dai lavoratori irregolari (gli indici si riferiscono alla branca di attività e alla classe dimensionale)³⁸.

I contributi a carico del datore di lavoro sono calcolati come:

$$CD1 = \sum_{b,c} (base_{bc} * qd_{bc})$$

dove:

qd_{bc} = aliquote contributive effettive medie a carico dei datori di lavoro per branca e classe dimensionale.

I contributi a carico del lavoratore sono pari a:

$$CL1 = \sum_{b,c} (base_{bc} * ql_b * k_{bc})$$

dove:

ql_b = aliquote contributive di legge a carico del lavoratore per branca di attività economica, k_{bc} è il fattore di correzione applicato alla retribuzione per calcolare l'imponibile contributivo.

La seconda ipotesi assume che nel caso in cui i lavoratori irregolari fossero regolarizzati, il datore di lavoro cerchi di contenere il livello del proprio costo del lavoro e quindi mantenga la stessa retribuzione oraria di un dipendente irregolare, pagando però su di essa i relativi oneri sociali; sul lavoratore gravano le imposte e i contributi dovuti, che quindi riducono la sua retribuzione netta. Dal punto di vista del calcolo dei contributi evasi, si tratta ragionevolmente dell'ipotesi minima che considera solo un ammontare che è comunque evaso in presenza di lavoro irregolare.

Secondo questa ipotesi, il calcolo dei contributi a carico del datore e quelli a carico dei lavoratori si esprime come:

$$CD2 = \sum_{b,c} (Ri_{bc} * qd_{bc})$$

dove:

Ri_{bc} = complesso delle retribuzioni irregolari per branca, classe dimensionale ($Ri_{bc} = ri_{bc} * hi_{bc}$)

³⁷ Il fattore di correzione per passare da retribuzioni lorde stimate da fonti contabili a imponibile contributivo è misurato a partire da dati sulle retribuzioni di fonte INPS Tale fattore correttivo impatta per il 5% circa sul calcolo dei contributi a carico del lavoratore.

³⁸ Il livello di disaggregazione è il medesimo specificato in precedenza, riguardo al calcolo delle retribuzioni dei lavoratori irregolari.

e

$$CL2 = \sum_b c(Ri_{bc} * ql_b * k_{bc}).$$

I.8 METODOLOGIA PER LA STIMA DEL TAX GAP DELLE IMPOSTE IMMOBILIARI

Il calcolo del *tax gap* delle imposte immobiliari è stato sottoposto ad un'importante revisione metodologica, che ha riguardato, per coerenza, l'intera serie storica, dall'anno d'imposta 2016 all'anno d'imposta 2021.

La revisione metodologica è duplice: da un lato, il miglioramento nella quantificazione del *compliance gap*, basato sulla differenza tra gettito teorico ed effettivo, non più misurati ad aliquota di base, ma applicando le aliquote deliberate dai singoli comuni; dall'altro lato, una migliore approssimazione del *policy gap*, ovvero l'erosione fiscale connessa alle esenzioni e agevolazioni fiscali previste a legislazione vigente. La prima revisione implica un incremento del *tax gap*, mentre la seconda revisione, riducendo il gettito teorico, va nella direzione opposta.

Come noto, fino alla Relazione dello scorso anno, il calcolo del *tax gap* era effettuato secondo l'aliquota di base, o aliquota *standard*, definita dal legislatore su base nazionale, pari a 8,6 per mille per i fabbricati diversi dall'abitazione principale (7,6 per mille fino al 2019). Pertanto, mentre il gettito IMU teorico era immediatamente derivabile sulla base dei dati degli archivi catastali, una maggiore complessità era connessa alla c.d. procedura di standardizzazione del gettito effettivamente riscosso, la quale consisteva nel ricondurre il gettito effettivo riscosso, ad aliquota deliberata, al gettito IMU effettivo standard, valutato ad aliquota di base. Con la revisione metodologica, la complessità risulta capovolta. Mentre il gettito IMU effettivo è immediatamente corrispondente al gettito riscosso ad aliquota deliberata, il gettito IMU teorico necessita di importanti elaborazioni sui dati derivanti dagli archivi catastali in grado di tenere conto delle differenti aliquote deliberate dai comuni. Tale revisione metodologica rende, tuttavia, il calcolo del *tax gap* delle imposte immobiliari coerente con il calcolo del *tax gap* effettuato per le altre imposte.

La determinazione del gettito IMU teorico è basata sui dati dell'archivio catastale dell'anno di riferimento (l'ultima annualità disponibile è il 2021). In tutta la serie storica considerata, l'IMU gravante sulle abitazioni principali è limitata alle categorie catastali A1, A8 e A9, i c.d. immobili di lusso. Considerato l'ammontare esiguo³⁹, si è trascurato il *tax gap* relativo agli immobili adibiti ad abitazione principale. Inoltre, le analisi effettuate fanno riferimento a tutti gli immobili esclusi i terreni, le aree fabbricabili e i fabbricati rurali strumentali che richiederebbero uno sforzo rilevante di omogeneizzazione delle basi imponibili (ad esempio, per le aree fabbricabili la base imponibile è data dal valore di mercato in luogo della rendita catastale).

Le analisi presentate impongono cautela nell'interpretazione dei risultati. Occorre, infatti, considerare alcune peculiarità del *tax gap* immobiliare. Innanzitutto, la tassazione immobiliare è, per varie ragioni, maggiormente soggetta a forme fisiologiche di bassa propensione all'adempimento dei contribuenti (c.d. *tax non compliance*). Si pensi, ad esempio, alla mancata percezione da parte dei contribuenti non residenti delle quote di proprietà degli immobili diversi dall'abitazione principale, che genera forme di inadempimento fiscale. In secondo luogo, si consideri la dinamica della *tax compliance* nel caso degli immobili adibiti ad uso produttivo, particolarmente colpiti dalla crisi economica, che rappresenta una conseguenza diretta dei problemi di liquidità dei contribuenti durante la recessione.

Infine, il calcolo del *tax gap* delle imposte immobiliari risente anche di altri fattori marginali, ma non meno importanti: gli effetti dovuti agli errori dei versamenti da parte dei contribuenti

³⁹ In termini generali, l'ammontare del *tax gap* per le abitazioni principali è piuttosto basso, come risulta anche dalle quantificazioni effettuate per l'anno d'imposta 2012 e riportate nelle precedenti edizioni della Relazione. Ancora più limitato è l'ammontare del *tax gap* se l'analisi viene ristretta alle abitazioni principali "di lusso".

(con un effetto *spillover* positivo verso alcuni Comuni che hanno riscosso somme non dovute, e negativo verso altri Comuni, privati di risorse altrimenti spettanti); gli errori e le omissioni che ancora sono presenti negli archivi catastali, ad esempio quelle derivanti dal disallineamento tra il territorio amministrativo dei Comuni e gli archivi catastali; le approssimazioni metodologiche dovute alla difficoltà di integrare puntualmente le informazioni sulle rendite catastali con quelle di fonte dichiarativa relativa alle destinazioni d'uso degli immobili, al fine di attribuire in modo accurato il corretto utilizzo degli immobili stessi.

Tuttavia, il nuovo approccio ha fornito l'opportunità di procedere al miglioramento della quantificazione delle esenzioni e agevolazioni fiscali previste a legislazione vigente, in virtù del fatto che l'imposta è calcolata applicando alla base imponibile le aliquote deliberate dal comune, tenuto conto sia della tipologia di immobile e dell'utilizzo dell'immobile stesso, sia delle caratteristiche specifiche del soggetto proprietario. L'affinamento metodologico si fonda, quindi, sull'acquisizione di informazioni provenienti da diverse fonti per cercare di specificare nel modo migliore possibile l'utilizzo dell'immobile e la tipologia del proprietario.

Particolare attenzione si è posta nella ricerca di algoritmi per l'individuazione delle tipologie di utilizzo dell'immobile da parte di soggetti la cui attività non è commerciale, o lo è solo parzialmente. Pertanto, al fine di ottenere una maggiore coerenza tra la base dati e la normativa IMU di riferimento, si è provveduto ad integrare le informazioni provenienti dall'archivio catastale e dalle dichiarazioni, con le ulteriori informazioni che è stato possibile acquisire. In particolare, le fonti utilizzate riguardano: i) il modello IMU Enti non commerciali (IMU-ENC), il quale risulta obbligatorio da compilare per ciascun contribuente interessato da specifiche esenzioni IMU; ii) i versamenti F24 effettuati dai singoli soggetti; iii) la base dati dei codici fiscali dell'Anagrafe degli Enti della Pubblica Amministrazione pubblicata dalla Ragioneria generale dello Stato.

In dettaglio, i passaggi seguiti nella revisione metodologica sono i seguenti.

Innanzitutto, la metodologia adottata ha previsto la quantificazione, per ogni singolo soggetto e all'interno di ogni comune, del gettito derivante dall'applicazione piena delle aliquote deliberate alla propria base catastale.

In un secondo momento, attraverso l'utilizzo delle informazioni prima menzionate, si è provveduto a monitorare le caratteristiche dei proprietari degli immobili. In particolare, attraverso le informazioni disponibili, i soggetti sono stati classificati separatamente in diversi sottogruppi seguendo, per quanto possibile, la normativa di riferimento sulle varie esenzioni dall'imposta. La distinzione ha riguardato i seguenti gruppi di contribuenti: ATER, Enti Pubblici, Enti non commerciali, Costruttori, Altri soggetti regolatori, Altre persone non fisiche (PNF), Persone fisiche (PF).

Successivamente sono stati stimati separatamente e, ove possibile, individualmente, gli effetti delle eventuali esenzioni attraverso la definizione congiunta di utilizzo dell'immobile e tipologia di proprietario, soprattutto con riferimento agli immobili di categoria catastale B⁴⁰, differenziati secondo le modalità di seguito riportate.

Utilizzo della dichiarazione IMU-ENC

Gli immobili utilizzati dai soggetti non commerciali, nell'ambito di funzioni istituzionali, possono essere considerati esenti dal pagamento dell'IMU, previa comunicazione attraverso il modello IMU-ENC. Lo stesso modello, da compilare per ciascun comune in cui risiede l'immobile, prevede che il soggetto interessato riporti il suo patrimonio immobiliare e ricalcoli l'imposta da versare al netto delle esenzioni dichiarate.

Per ogni unità così determinata è stata agganciata l'informazione presente nella dichiarazione IMU-ENC, espressa nel quadro C, relativa alla liquidazione calcolata. È stato, quindi, effettuato

⁴⁰ Gli immobili di tipo B sono: B/1 - collegi, orfanotrofi, conventi, seminari, ricoveri, ospizi, caserme -; B/2 - ospedali e case di cura senza fine di lucro -; B/3 - riformatori e prigionieri -; B/4 - uffici pubblici -; B/5 - laboratori scientifici e scuole -; B/6 - accademie, gallerie, musei, pinacoteche -; B/7 - oratori e cappelle non destinate all'esercizio pubblico del culto -; B/8 - magazzini sotterranei per depositi di derrate -.

un confronto tra l'importo teorico da catasto, l'importo calcolato da dichiarazione e l'importo versato.

Nel caso in cui non risulta alcuna dichiarazione IMU-ENC, il gettito teorico è assunto pari al gettito derivante dal catasto. Nel caso in cui il versamento corrisponde al gettito da catasto, allora si considera che il gettito teorico sia esattamente il gettito da catasto. Nel caso in cui il versamento è pari a quanto dichiarato dal quadro C, si ipotizza che tale ammontare sia equivalente al gettito teorico, con *tax gap* nullo. Nel caso in cui il versamento è pari al gettito da catasto al netto degli immobili di categoria catastale B, il gettito teorico si ipotizza pari al massimo tra il gettito dichiarato nel quadro C e il gettito versato. Nel caso in cui risulta un gettito dichiarato e versato pari a zero, si ipotizza che il gettito teorico e il gettito effettivo siano entrambi nulli. In tutti gli altri casi, il gettito teorico è assunto pari a quanto dichiarato nel quadro C.

In tutti i casi, comunque, sono stati riattribuiti i versamenti ai singoli immobili, utilizzando come fattore di riproporzionamento il gettito teorico opportunamente corretto.

Analisi sugli Enti pubblici

Nel caso degli Enti pubblici è stata effettuata un'analisi puntuale e "prudenziale" delle esenzioni previste a legislazione vigente. Al fine di non sovrastimare le esenzioni, sono stati, infatti, considerati solo gli immobili che presentano determinate caratteristiche individuabili da fonte catastale. In particolare, sono stati considerati solo gli immobili di tipo B, D04 ovvero A10, per alcune categorie di soggetti e contribuenti. Ulteriore condizione è che gli stessi immobili non risultino locati.

Sono stati considerati esenti gli immobili: i) di categoria B03 - riformatori e prigioni-, istituzionali per definizione; ii) con categoria catastale B02, B04, B05 oppure D04 - ospedali e case di cura con fine di lucro - nel caso in cui il soggetto proprietario è dato da "Aziende ospedaliere pubbliche", "Aziende ospedaliero-universitarie e Policlinici universitari pubblici", "Aziende sanitarie locali"; iii) con categoria catastale B01, B04, B06 oppure A10, nel caso in cui il proprietario è dato dalla "Presidenza del Consiglio dei ministri e Ministeri", dalle "Regioni", dalle "Province e Città metropolitane", nonché dalle "Agenzie fiscali"; iv) con categoria catastale B02, B04 e B05, nel caso in cui il soggetto è dato dagli "Istituti di ricovero e cura a carattere scientifico pubblici"; v) con categoria catastale B04 nel caso in cui il soggetto è dato dagli "altri Enti e Istituzioni pubbliche locali".

ATER

Questi soggetti sono esenti dall'imposta relativa agli immobili aventi le caratteristiche degli alloggi sociali. Il quadro normativo è molto complesso e spesso foriero di contenziosi presso il TAR o le Commissioni tributarie. Non è, quindi, possibile individuare precisamente, a partire dal catasto, quali immobili siano esenti o meno. Resta, comunque, ferma la condizione necessaria ma non sufficiente, ai fini dell'esenzione, che tali immobili siano abitativi.

Per stimare la quota di esenzione, è stata calcolata da catasto l'imposta teorica relativa alle proprietà e divisa tra la componente del patrimonio immobiliare abitativo e non abitativo.

La procedura implementata si basa sul presupposto che il versamento riguardi innanzitutto la componente del patrimonio immobiliare non abitativo (che non gode di esenzione) e, nel caso di eccesso di gettito, la componente residua viene attribuita al patrimonio abitativo.

In ultima analisi, il gettito teorico corrisponde alla base imponibile del patrimonio non abitativo (da catasto) e dalla componente di gettito versato in eccesso, attribuito al patrimonio abitativo⁴¹.

⁴¹ Per estendere la stima anche su coloro che non hanno versato, si è calcolata la percentuale di gettito versato sull'abitativo (rispetto al totale da catasto). Tale percentuale è stata, quindi, applicata anche ai soggetti che non hanno versato, ipotizzando che la struttura del patrimonio immobiliare abitativo e la ripartizione degli stessi fra esenti e non esenti sia la stessa per tutti gli ATER e in tutti i comuni.

Altri Regolatori

Esistono una serie di soggetti classificati nella divisione ATECO “84 - Servizi collettivi delle Amministrazioni pubbliche”, non inclusi nella lista fornita dalla Ragioneria generale dello Stato. Per questi soggetti si è supposto che gli immobili A10 - uffici - e B siano utilizzati nell’ambito dell’attività istituzionale. Pertanto, l’imposta teorica di questi immobili è considerata esente.

Costruttori

Per quanto concerne i soggetti la cui attività è quella del settore delle “Costruzioni”, occorre distinguere tre casi: i) il caso in cui l’immobile viene utilizzato per l’attività economica; ii) il caso in cui l’immobile viene locato; iii) il caso dei c.d. “beni merce”, ovvero gli immobili destinati alla vendita. I c.d. “beni merce” sono esenti. Tuttavia, non disponendo di informazioni sufficienti a individuare in modo dettagliato i beni merce, si è supposto, nel caso di versamento, che il soggetto abbia effettuato correttamente il calcolo. Pertanto, nel caso dei soggetti che hanno effettuato il versamento, sono stati distinti gli immobili locati (non esenti), quelli abitativi non locati (probabili beni merce) e tutti gli altri immobili residuali. Infine, si è confrontato il gettito versato con il gettito teorico non esente e, nel caso di differenziale positivo, è stato riassegnato proporzionalmente il gettito teorico alla categoria degli immobili residuali.

II. REVISIONE DELLE STIME TRA IL 2022 E IL 2023 DI ALCUNI TAX GAP

II.1 REVISIONI DEL GAP DELL'IRPEF DA LAVORO AUTONOMO E DA IMPRESA E DELL'IRES

La metodologia di stima del *gap* dell'IRPEF da lavoro autonomo e di impresa e dell'IRES non ha subito revisioni rispetto alla precedente edizione. Pertanto, le differenze osservate sono imputabili esclusivamente all'aggiornamento delle fonti di dati utilizzate nel calcolo. In particolare, con riferimento al *gap* IRES per l'anno d'imposta 2020, si evidenzia come la revisione delle stime sia riconducibile ad una correzione al ribasso sia dell'imposta potenziale sia dell'imposta dichiarata, mentre l'importo relativo ai mancati versamenti, tuttora oggetto di estrapolazione per indisponibilità della fonte amministrativa, risulta in linea con quanto previsto nell'edizione precedente. Ne consegue che, in termini di propensione, la componente non dichiarata subisca una variazione più contenuta rispetto a quella del dichiarato e non versato.

TABELLA II.1.1: GAP IRPEFaut (IN MLN DI EURO) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020

Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IRPEF EDIZIONE ATTUALE in mln di euro (IRPEFNV)	Non dichiarato	31.497	31.411	30.657	30.263	26.336
	Dichiarato e non versato	1.860	1.916	2.266	2.213	1.876
	Totale non versato	33.357	33.327	32.923	32.476	28.212
Gap IRPEF EDIZIONE 2022 in mln di euro (IRPEFNV)	Non dichiarato	31.497	31.411	30.657	30.263	26.476
	Dichiarato e non versato	1.862	1.906	2.280	2.188	1.855
	Totale non versato	33.359	33.318	32.938	32.451	28.331
Differenze tra il Gap IRPEF attuale e l'edizione 2022	Non dichiarato	0	0	0	0	-140
	Dichiarato e non versato	-1	9	-14	25	21
	Totale non versato	-1	9	-14	25	-119

TABELLA II.1.2: PROPENSIONE AL GAP IRPEFaut (IN PERCENTUALE) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020

Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IRPEF EDIZIONE ATTUALE in % dell'imposta potenziale (IRPEFNV/IRPEFp)	Non dichiarato	62,7%	64,1%	62,8%	64,5%	65,0%
	Dichiarato e non versato	3,7%	3,9%	4,6%	4,7%	4,6%
	Totale non versato	66,4%	68,0%	67,5%	69,2%	69,7%
Gap IRPEF EDIZIONE 2022 in % dell'imposta potenziale (IRPEFNV/IRPEFp)	Non dichiarato	62,7%	64,1%	62,8%	64,5%	65,2%
	Dichiarato e non versato	3,7%	3,9%	4,7%	4,7%	4,6%
	Totale non versato	66,4%	68,0%	67,5%	69,1%	69,7%
Differenze tra il Gap IRPEF attuale e l'edizione 2022	Non dichiarato	0,0%	0,0%	0,0%	0,0%	-0,1%
	Dichiarato e non versato	0,0%	0,0%	0,0%	0,1%	0,1%
	Totale non versato	0,0%	0,0%	0,0%	0,1%	-0,1%

TABELLA II.1.3: GAP IRES (IN MLN DI EURO) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020						
Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IRES EDIZIONE ATTUALE in mln di euro (IRESNV)	Non dichiarato	8.789	7.156	6.394	7.063	6.776
	Dichiarato e non versato	1.498	1.476	1.512	1.917	1.711
	Totale non versato	10.287	8.632	7.906	8.980	8.487
Gap IRES EDIZIONE 2022 in mln di euro (IRESNV)	Non dichiarato	8.789	7.156	6.394	7.063	7.424
	Dichiarato e non versato	1.497	1.507	1.560	1.674	1.658
	Totale non versato	10.286	8.662	7.954	8.738	9.082
Differenze tra il Gap IRES attuale e l'edizione 2022	Non dichiarato	0	0	0	0	-648
	Dichiarato e non versato	1	-31	-48	243	54
	Totale non versato	1	-31	-48	243	-594

TABELLA II.1.4: PROPENSIONE AL GAP IRES (IN PERCENTUALE) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020						
Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IRES EDIZIONE ATTUALE in % dell'imposta potenziale (IRESNV/IRESp)	Non dichiarato	22,7%	19,6%	17,3%	18,7%	19,8%
	Dichiarato e non versato	3,9%	4,0%	4,1%	5,1%	5,0%
	Totale non versato	26,6%	23,7%	21,4%	23,8%	24,9%
Gap IRES EDIZIONE 2022 in % dell'imposta potenziale (IRESNV/IRESp)	Non dichiarato	22,7%	19,6%	17,3%	18,7%	19,7%
	Dichiarato e non versato	3,9%	4,1%	4,2%	4,4%	4,4%
	Totale non versato	26,6%	23,7%	21,6%	23,2%	24,0%
Differenze tra il Gap IRES attuale e l'edizione 2022	Non dichiarato	0,0%	0,0%	0,0%	0,0%	0,2%
	Dichiarato e non versato	0,0%	-0,1%	-0,1%	0,6%	0,6%
	Totale non versato	0,0%	-0,1%	-0,1%	0,6%	0,8%

II.2 REVISIONI DEL GAP DELL'IVA

La metodologia di stima del *gap* dell'IVA non ha subito revisioni rispetto alla precedente edizione. La revisione al ribasso della stima del *gap* per l'anno di imposta 2020 è riconducibile ad una correzione dell'imposta potenziale, legata sia ad un aggiornamento delle fonti dati (in particolare ai dati di contabilità nazionale sui costi intermedi) sia ad aggiustamenti effettuati su approssimazioni precedentemente effettuate per tenere conto dei provvedimenti adottati per la pandemia.

Le differenze osservate (Tabella II.2.1) sono anche riconducibili all'aggiornamento degli archivi dei controlli automatizzati delle dichiarazioni: vale la pena ricordare che per le ultime due annualità considerate, i mancati versamenti vengono stimati sulla base della dinamica dell'IVA evasa.

TABELLA II.2.1: GAP IVA IN MLN DI EURO: CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020						
Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IVA EDIZIONE ATTUALE in mln di euro (IVANV)	Non dichiarato	24.274	26.694	20.942	18.073	15.391
	Dichiarato e non versato	9.888	10.130	10.216	8.817	7.508
	Totale non versato	34.162	36.824	31.158	26.889	22.899
Gap IVA EDIZIONE 2022 in mln di euro (IVANV)	Non dichiarato	24.764	25.362	21.479	18.718	17.011
	Dichiarato e non versato	9.894	10.208	10.344	9.014	8.192
	Totale non versato	34.659	35.570	31.823	27.732	25.203
Differenze tra il Gap IVA attuale e l'edizione 2022	Non dichiarato	-490	1.332	-537	-646	-1.620
	Dichiarato e non versato	-6	-78	-127	-197	-684
	Totale non versato	-496	1.253	-665	-843	-2.304

TABELLA II.2.2: PROPENSIONE AL GAP IVA IN PERCENTUALE: CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020

Aggregato	Tipologia	2016	2017	2018	2019	2020
Gap IVA EDIZIONE ATTUALE in % dell'imposta potenziale (IVANV/IVAT)	Non dichiarato	18,4%	19,8%	15,4%	13,4%	12,9%
	Dichiarato e non versato	7,5%	7,5%	7,5%	6,5%	6,3%
	Totale non versato	25,9%	27,3%	22,9%	19,9%	19,2%
Gap IVA EDIZIONE 2022 in % dell'imposta potenziale (IVANV/IVAT)	Non dichiarato	18,7%	19,0%	15,7%	13,7%	14,0%
	Dichiarato e non versato	7,5%	7,6%	7,6%	6,6%	6,7%
	Totale non versato	26,2%	26,6%	23,3%	20,4%	20,7%
Differenze tra il Gap IVA attuale e l'edizione 2022	Non dichiarato	-0,3%	0,8%	-0,3%	-0,4%	-1,1%
	Dichiarato e non versato	0,0%	-0,1%	-0,1%	-0,1%	-0,4%
	Totale non versato	-0,3%	0,7%	-0,4%	-0,5%	-1,5%

II.3 REVISIONI DEL GAP DELL'IRAP

Dal punto di vista dell'analisi di qualità delle banche dati, per l'anno d'imposta 2020, si segnala la disponibilità dei dati dichiarativi definitivi come elemento determinante della revisione del valore del *gap* nell'imposta. Tale aggiornamento non modifica la dinamica già registrata nella precedente versione delle stime, sia con riferimento al *gap* non dichiarato in valore assoluto sia per quanto riguarda la propensione al *gap*. Tuttavia, l'ammontare della riduzione per il 2020 è risultato leggermente meno marcato del previsto.

La serie dei mancati versamenti ha subito dei cambiamenti rispetto all'ultima versione rilasciata, dovuti all'aggiornamento degli archivi dei controlli automatizzati delle dichiarazioni. In particolare, si segnala sul 2018 la più rilevante riduzione nell'ammontare dichiarato e non versato (che passa da 1.103 milioni di euro a 1.063 milioni di euro). Tale decremento produce una riduzione del *trend* del *tax gap* IRAP complessivo (totale non versato), che risulta quindi in minore aumento tra il 2017 e il 2018.

TABELLA II.3.1: GAP IRAP (IN MLN DI EURO) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020.

	Anno	2016	2017	2018	2019	2020
Gap IRAP VERSIONE ATTUALE	Non dichiarato	3.955	4.118	4.134	3.925	3.613
	Non versato	997	991	1.063	1.101	1.000
	Complessivo	4.952	5.109	5.197	5.026	4.613
Gap IRAP VERSIONE 2022	Non dichiarato	3.955	4.118	4.134	3.925	3.506
	Non versato	991	999	1.103	1.098	1.008
	Complessivo	4.946	5.117	5.237	5.023	4.514
Differenze tra il Gap IRAP attuale e l'edizione 2022	Non dichiarato	0	0	0	0	107
	Non versato	6	-8	-40	4	-8
	Complessivo	6	-8	-40	4	99

Le revisioni del *tax gap* in valore hanno comportato, naturalmente, revisioni anche in termini di propensione (Tabella II.3.2). Si segnala che, rispetto alla precedente *release*, si è osservato una riduzione della propensione al *gap* per il 2020 sul 2019 un poco meno ampia del previsto (in maniera analoga a quanto osservato per il *gap* in valore assoluto).

TABELLA II.3.2: PROPENSIONE AL GAP IRAP (IN PERCENTUALE) CONFRONTO TRA L'EDIZIONE 2022 E L'ATTUALE. ANNI 2016-2020.

	Anno	2016	2017	2018	2019	2020
Gap IRAP EDIZIONE ATTUALE in % dell'imposta potenziale (IRAPNV/IRAP) ¹	Non dichiarato	15,0%	15,1%	14,7%	14,2%	13,9%
	Dichiarato e non versato	3,8%	3,6%	3,8%	4,0%	3,8%
	Complessivo	18,8%	18,8%	18,5%	18,2%	17,7%
Gap IRAP EDIZIONE 2022 in % dell'imposta potenziale (IRAPNV/IRAP) ¹	Non dichiarato	15,0%	15,1%	14,7%	14,2%	13,4%
	Dichiarato e non versato	3,8%	3,7%	3,9%	4,0%	3,9%
	Complessivo	18,8%	18,8%	18,6%	18,2%	17,3%
Differenze tra il Gap IRAP in % dell'imposta potenziale attuale e l'edizione 2022	Non dichiarato	0,0%	0,0%	0,0%	0,0%	0,4%
	Dichiarato e non versato	0,0%	0,0%	-0,1%	0,0%	0,0%
	Complessivo	0,0%	0,0%	-0,1%	0,0%	0,4%

II.4 REVISIONI DEL GAP DELLE LOCAZIONI

Per quanto riguarda il *gap* relativo alle locazioni, non sono intervenute modifiche metodologiche né aggiornamenti delle fonti informative che abbiano determinato una revisione delle stime rispetto all'edizione precedente.

III. ALTRE METODOLOGIE

III.1 LA STIMA BOTTOM UP DEL GAP IVA

La metodologia di tipo *bottom-up* permette di ottenere una stima del *gap* IVA ad elevato livello di disaggregazione sulla base degli attributi dei soggetti passivi.

Poiché l'approccio adottato è di tipo predittivo, si è reso necessario individuare degli *output* (oggetto della previsione) da utilizzare per definire ogni variabile *target*, che, concettualmente, può essere identificata con il *gap* IVA da attribuire al singolo contribuente. Questa scelta, tutt'altro che banale, ha portato ad optare per due diversi approcci di stima, speculari nella metodologia ma diversi sia nel risultato che nell'utilizzo finale. Questi due approcci, che chiamiamo globale e tematico, sono speculari nella metodologia ma differenti negli scopi.

Quando la variabile *target* è la maggiore imposta iva accertata (MIA) contestabile al singolo contribuente, il bottom-up stimato è chiamato *globale*. Quando invece procediamo alla stima delle singole componenti IVA contestabili al contribuente, parliamo invece di bottom-up *tematico*. Invero, è possibile estrapolare dalle nostre banche dati le singole voci della dichiarazione IVA che sono state rettificata in sede di accertamento: ciò ci permette di ricostruire nel dettaglio tutto il profilo evasivo attuato dal contribuente, e ovviamente anche di studiarlo in aggregato, sia nel sottoinsieme osservato che nella platea allargata su cui si effettua la stima. A livello metodologico i due approcci sono identici, con la differenza che nel globale il *matching* viene effettuato su una sola variabile, la MIA, mentre per il tematico viene effettuato su un vettore di osservazioni, nello specifico la maggiore imposta determinata nei quadri VE, VF, VJ. La differenza metodologica fra i due approcci è di natura esclusivamente tecnica, e relativa solamente all'algoritmo di *matching* utilizzato: univariato nel caso globale e multivariato nel tematico.

I due approcci differiscono invece sensibilmente per quanto riguarda lo scopo. Il bottom-up globale può essere utilizzato per una caratterizzazione immediata del fenomeno evasivo, rendendo possibile lo studio complessivo della maggiore imposta. Con la stima globale del bottom-up è infatti possibile costruire mappe del rischio e tabelle che affrontino la questione nel suo insieme, senza poter distinguere tra costi, ricavi, aliquote e detrazioni indebite. Queste distinzioni sono invece possibili grazie all'approccio tematico, attualmente in fase sperimentale su una platea ristretta, e dunque soggetto a possibili migliorie in futuro.

III.1.1 La stima globale

La scelta della variabile *target* si è orientata sulla maggiore imposta accertata (MIA), definibile come l'importo della pretesa erariale determinato dall'Amministrazione e "notificato" al contribuente. Per una discussione più approfondita sulla scelta della variabile obiettivo, si veda la Relazione sull'economia non osservata e sull'evasione fiscale e contributiva - anno 2022.

Al fine di operare in un quadro sufficientemente completo e, conseguentemente, rappresentativo dell'insieme dei controlli aventi ad oggetto un determinato anno, sono stati considerati, per la prima fase di stima, solo gli accertamenti relativi agli anni d'imposta 2014 e 2015 e, per la seconda fase, quelli relativi agli anni dal 2014-2017 in poi.⁴²

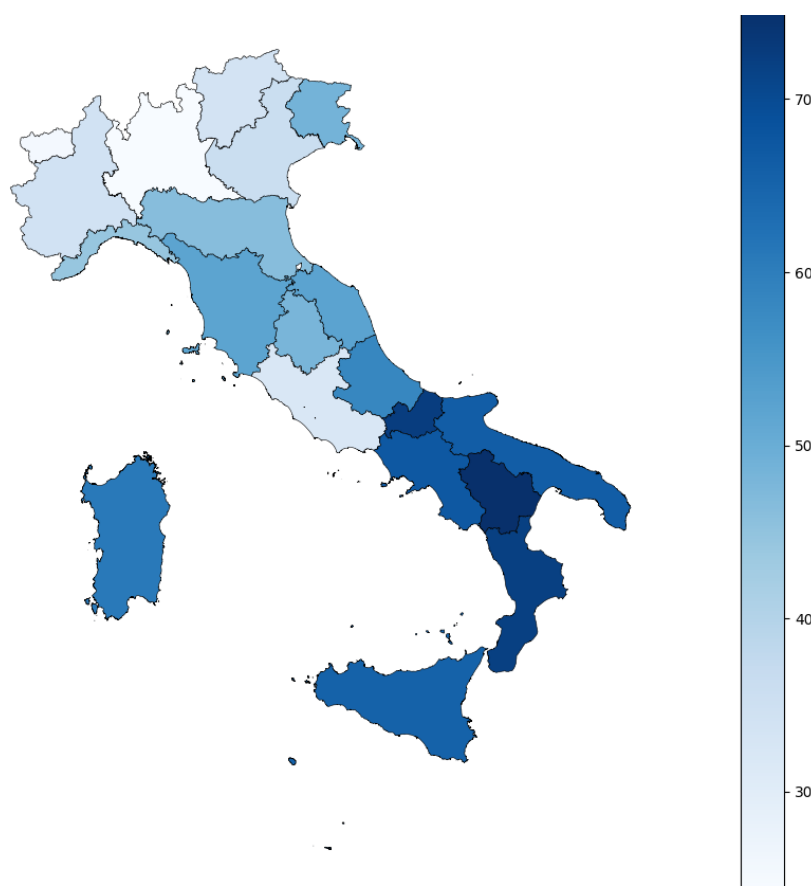
⁴² La metodologia di stima consta di due fasi distinte. Per una spiegazione dettagliata delle due fasi, si rimanda alla Relazione sull'economia non osservata e sull'evasione fiscale e contributiva - Anno 2022.

Il campione così identificato, subordinatamente ad opportune ipotesi di stabilità delle relazioni tra predittori e variabili risposta, è potenzialmente utilizzabile per stimare il *tax gap* riguardante gli anni dal 2014 in poi. Nella sperimentazione condotta, come sopra anticipato, l'anno di interesse è il 2021. Per l'illustrazione dettagliata della metodologia impiegata, si rimanda alla relazione 2022.

Come già sottolineato, lo scopo principale della metodologia sviluppata è quello di permettere la disaggregazione e la scomposizione del *gap* IVA sulla base dei diversi domini di interesse, tra cui quello geografico (ad esempio, regionale, provinciale, comunale), settoriale e dimensionale. In linea di principio, il metodo *bottom up* consente di ottenere previsioni del *gap* IVA e della propensione ad evadere a livello di singolo contribuente, e pertanto derivare stime su qualsiasi dominio di interesse per aggregazione delle previsioni micro. Occorre tenere presente, tuttavia, che all'aumentare del dettaglio richiesto, si riduce l'accuratezza della stima. Inoltre, il risultato a livello micro è importante non tanto per una valutazione del livello di evasione del singolo contribuente, quanto per il valore informativo che tale disaggregazione implica.

A titolo di esempio, si riportano i risultati del *gap* e della propensione a livello regionale (Figura III.1.1.1). I risultati evidenziano una propensione minore al nord, e maggiore al sud e nelle isole. Inoltre, tramite il modello *bottom-up* è possibile incrociare i diversi domini di interesse per effettuare analisi dettagliate sul fenomeno evasivo.

FIGURA III.1.1.1: PROPENSIONE REGIONALE AL GAP IVA



III.1.2 La stima tematica per singole componenti

Come già menzionato, la metodologia di tipo *bottom-up* descritta nella relazione evasione 2022 è stata applicata per distinguere le diverse componenti del *gap* IVA, in funzione delle diverse operazioni effettuate dai contribuenti (ad esempio cessioni o acquisti). Analizzando i singoli campi

di dichiarazione, infatti, si può verificare l'origine della maggiore imposta accertata, e ricostruire quella totale.

A differenza di quanto fatto per la metodologia *bottom-up* totale, poiché si intende studiare contemporaneamente le singole componenti della maggiore imposta accertata, la variabile di *output* (MIA totale) è ora da considerare come un vettore, le cui componenti, che andranno a comporre il totale della maggiore imposta accertata, sono state stimate simultaneamente.

In particolare, i due aspetti principali che contribuiscono alla formazione del *gap* IVA sono la sotto-dichiarazione delle vendite e la sovra-dichiarazione dei costi, che rappresentano, dal lato dell'imposta, la MIA dal lato attivo e la MIA da lato passivo, rispettivamente. La prima è stata ricavata dalle informazioni presenti in dichiarazioni e accertamenti relativi al quadro VE, rigo 26 mentre la seconda proviene da quelle del quadro VF, rigo 71. Inoltre, come approfondimento sono stati considerati due ulteriori fenomeni: l'evasione da aliquota, sia sul lato attivo che su quello passivo, e la scomposizione della sotto-dichiarazione dei ricavi nelle componenti B2B e B2C mediante il quadro VT.

Riguardo la formulazione dell'evasione da aliquota, in questa fase di analisi ne è stato definito un limite inferiore. Infatti, la maggiore imposta da aliquota è stata stimata unicamente nei casi in cui, in fase di accertamento, è stata attribuita solamente una maggiore imposta e non un maggior imponibile, poiché si tratta dei soli casi in cui l'evasione deriva sicuramente da un'aliquota errata. Tale calcolo è stato effettuato sia per gli acquisti che per le cessioni, dunque è stata ottenuta una maggiore imposta accertata da aliquota sia attiva che passiva, che rappresentano una parte del totale della maggiore imposta attiva e passiva, rispettivamente.

Oltre alla stima dell'evasione da aliquota, è stata calcolata la ripartizione della MIA totale attiva nella sua quota B2B, cioè derivante dalla cessione di beni e servizi ad altre imprese, e B2C, relativa alle operazioni destinate al consumo finale. Infatti, la quota della maggiore imposta B2C, oltre ad essere una stima di per sé informativa, potrebbe fornire una stima più realistica del *gap* IVA aggregato. L'ipotesi di base è che, una volta non dichiarata l'IVA in un'operazione B2B, essa non verrà né dichiarata e detratta nelle operazioni successive, né pagata dal consumatore finale. In questo caso, il mancato gettito per lo Stato sarebbe rappresentato dall'omesso pagamento dell'IVA ascrivibile al consumatore finale, più le componenti non detraibili come i costi d'impresa. Dunque, considerare la somma delle maggiori imposte IVA individuali porterebbe ad una sovrastima del *gap* IVA totale in quanto potrebbe essere contato più volte lo stesso ammontare di imposta evasa. Rimane tuttavia importante la stima del *bottom-up* globale, che come già evidenziato risponde ad esigenze differenti.

Da un punto di vista metodologico, la stima del *gap* IVA tematica è stata realizzata estendendo il modello predittivo già applicato per la stima *bottom-up* del *gap* IVA. Nel dettaglio, il primo dei due stadi di analisi, ossia la stima della probabilità di selezione per accertamento, è stato riproposto esattamente con le stesse modalità descritte nella Relazione *evasione 2022*.

Il secondo stadio, relativo alla predizione della maggiore imposta accertata IVA, è stato sottoposto a delle modifiche. In particolare, il *bagging* di alberi di regressione è stato utilizzato per stimare separatamente le previsioni della MIA lato attivo e lato passivo⁴³. Infine, per associare ad ogni contribuente non accertato il contribuente accertato con valori previsti più prossimi (*nearest neighbour*) è stata utilizzata una funzione di distanza multivariata (distanza di Mahalanobis), che permette di tenere conto contemporaneamente sia di entrambe le variabili di output principali che di eventuali correlazioni tra di esse. Una volta individuato, attraverso la previsione della MIA lato attivo e della MIA lato passivo, il contribuente accertato più prossimo, a ciascun soggetto non accertato vengono assegnate anche la maggiore imposta da aliquota e quella attribuibile alla componente B2C, oltre alle già citate MIA attiva e passiva.

Di seguito si riportano i risultati ottenuti dall'applicazione della metodologia descritta per la sola platea delle ditte individuali per l'anno 2019.

⁴³ Nel lato attivo consideriamo anche il quadro VJ, relativo alle operazioni in *reverse charge*, anche se in questa trattazione lo escludiamo per motivi di parsimonia espositiva.

TABELLA III.1.2.1: QUOTE DI MAGGIORE IMPOSTA ATTIVA E PASSIVA SU BASE REGIONALE

	MIA attiva	MIA passiva
PIEMONTE	57,8%	42,2%
VALLE D'AOSTA	54,4%	45,6%
LOMBARDIA	59,3%	40,7%
TRENTINO-ALTO ADIGE	53,5%	46,5%
VENETO	59,9%	40,1%
FRIULI VENEZIA GIULIA	58,6%	41,4%
LIGURIA	59,8%	40,2%
EMILIA ROMAGNA	59,9%	40,1%
TOSCANA	62,1%	37,9%
UMBRIA	60,6%	39,4%
MARCHE	59,1%	40,9%
LAZIO	62,2%	37,8%
ABRUZZO	58,3%	41,7%
MOLISE	54,8%	45,2%
CAMPANIA	59,2%	40,8%
PUGLIA	53,4%	46,6%
BASILICATA	51,6%	48,4%
CALABRIA	58,6%	41,4%
SICILIA	56,6%	43,4%
SARDEGNA	54,4%	45,6%
TOTALE	58,5%	41,5%

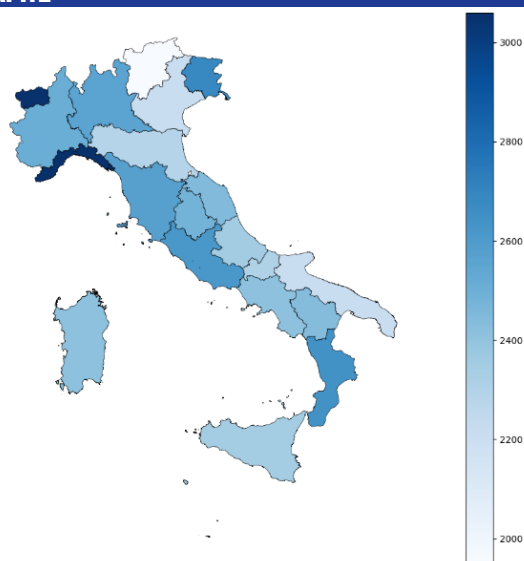
In Tabella III.1.2.1 è riportato la quota delle componenti della MIA: quella da sotto-dichiarazione dei ricavi (MIA attiva) e quella da sovra-dichiarazione dei costi (MIA passiva). Nella platea selezionata, il 58,5% della maggiore imposta totale predetta è ascrivibile al lato attivo, mentre il 41,5% al lato passivo. La distribuzione tra lato attivo e passivo è abbastanza omogenea tra le regioni. In tutti i casi la MIA attiva costituisce più del 50% della totale, con punte di oltre il 60% per Lazio, Toscana e Umbria.

Per quanto riguarda l'evasione da aliquota, questa può essere considerata unicamente come limite inferiore, nel senso che riguarda i casi in cui il mancato gettito deriva esclusivamente da una aliquota errata. Questa riguarda almeno l'1% dell'evasione da sotto-dichiarazione dei ricavi (lato attivo) e l'1,8% dell'evasione lato passivo.

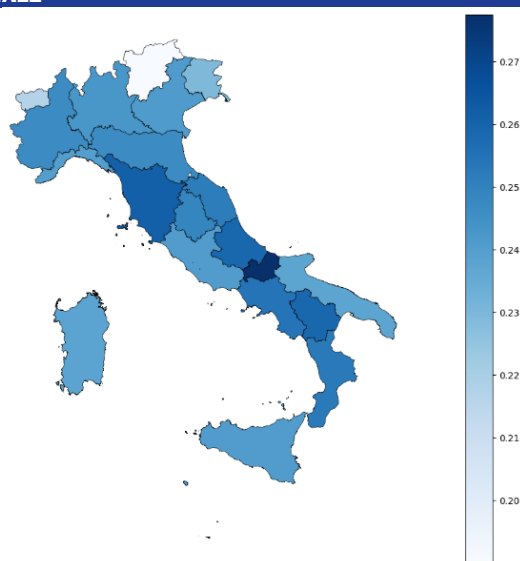
Nella Figura III.1.2.1 è riportata la distribuzione regionale della maggiore imposta pro capite derivante dalla sotto-dichiarazione dei ricavi riferita alle transazioni B2C. Poiché si considerano le sole operazioni B2C, il *gap* è imputato alla regione in cui vi è la sede legale delle ditte che effettuano l'operazione che coinvolge il consumatore finale. Le regioni che presentano i valori più elevati sono la Valle d'Aosta e la Liguria; il Trentino-Alto Adige si conferma tra le regioni a *gap* IVA pro capite più basso. I risultati ottenuti per il Trentino-Alto Adige risultano essere coerenti con quelli relativi alla propensione al *gap* IVA (Figura III.1.2.1b): questa, infatti, risulta essere la regione con la propensione al *gap* IVA inferiore. Al contrario, la regione in cui si evidenzia la più alta propensione al *gap* è il Molise, con un valore eccedente lo 0,27. Infine, per quanto riguarda il resto delle regioni sembra esserci una distribuzione regionale piuttosto uniforme.

Qualche considerazione sulle differenze visibili fra la Figura III.1.2.1a, rappresentativa dell'approccio globale, e la III.1.2.1b, che illustra i risultati dell'approccio tematico: nel primo caso, la mappa può essere considerata un'indicazione del comportamento evasivo individuale dei contribuenti di quella regione; nel secondo caso, invece, l'obiettivo è valutare il mancato gettito effettivo, focalizzandosi però su un particolare aspetto del fenomeno. Questo confronto mette bene in luce la differenza nello scopo dei due approcci considerati.

**FIGURA III.1.2.1a: GAP IVA B2C REGIONALE
PROCAPITE**



**FIGURA III.1.2.1b: PROPENSIONE GAP IVA B2C
REGIONALE**



III.2 LA STIMA BOTTOM UP DEL GAP IRAP

L'approccio *bottom-up* adottato per stimare il *tax gap* IRAP delle persone fisiche e società, si basa sull'applicazione di modelli econometrici che utilizzano sia le informazioni dei dati dichiarati al Fisco, sia quelle dedotte dalle fonti di accertamento fiscale, per il sottoinsieme di contribuenti sottoposti a tale processo. Per quanto riguarda la popolazione di stima, il modello prende a riferimento solo i contribuenti che hanno ricevuto un accertamento unificato sull'anno di imposta oggetto di studio, in quanto tale tipologia di controllo prevede uno specifico processo di selezione finalizzato alla ricostruzione della situazione patrimoniale e contabile e quindi dell'eventuale *gap* tra quanto dichiarato e quanto dovuto in termini di base imponibile e di imposta.

L'applicazione del modello riguarda la stima della Base Imponibile Potenziale (BIT) IRAP⁴⁴ sia per i soggetti accertati che per coloro che non hanno subito un accertamento, attraverso la proiezione su tutta la popolazione dei contribuenti dei risultati del modello statistico. Questa stima è importante poiché se confrontata con la base imponibile dichiarata (BID), consente una quantificazione della mancata base imponibile IRAP: il *gap* nella base (BIND).

La platea di studio è composta da Persone Fisiche, Società di Persone e Società di Capitali, che hanno effettivamente dichiarato reddito da impresa e/o da lavoro autonomo ovvero quei contribuenti che hanno dei dati IVA significativi. L'anno di imposta di riferimento per la stima dei modelli è il 2016.

Al fine di indentificare la platea di stima per il modello, sono stati identificate ed escluse le casistiche fuori contesto e non utili ai fini dello studio. Dalla popolazione iniziale sono stati esclusi i soggetti non passivi dell'imposta regionale sulle attività produttive⁴⁵ e i soggetti per i quali non si dispone di nessun dato reddituale. Per la definizione della popolazione di stima, state individuate ed escluse ulteriori casistiche specifiche dell'accertamento e sono stati applicati criteri per verificare e garantire l'affidabilità del dato. In seguito alla fase di *preprocessing*, la platea di riferimento è stata partizionata in due sotto-gruppi, all'interno dei quali si proceduto, in modo separato alla stima della variabile obiettivo: un gruppo per Persone Fisiche e Società di Persone studiate congiuntamente e uno per Società di Capitali.

⁴⁴ L'aggregato di riferimento nella stima del gap IRAP è il valore della Produzione Netta.

⁴⁵ I criteri specifici per la definizione della reale popolazione IRAP si basano sulle informazioni che derivano dai quadri del modello IRAP e per le sole Persone Fisiche anche sul concetto di autonoma organizzazione.

III.2.1 Il modello a due stadi di Heckman

L'approccio di tipo inferenziale alla base della metodologia è incentrato sul processo di stima a due stadi introdotto da Heckman negli anni Settanta, espresso da due equazioni differenti ma tra loro relazionate, ideate nell'ottica del superamento del problema del *selection bias* dovuto in questo caso a eventuali specificità dei Soggetti sottoposti ad accertamento fiscale, spesso frutto di un processo di selezione mirato da parte degli Enti preposti.

Tale metodologia si applica, infatti, nei casi in cui si debba stimare un modello econometrico a partire da un campione le cui osservazioni sulla variabile endogena sono disponibili solo se si verifica una data condizione che determina l'appartenenza del soggetto a un gruppo e non al suo complemento.

Nel caso in esame la variabile obiettivo, la Base Imponibile IRAP Potenziale, viene stimata tenendo conto della possibile distorsione dovuta al fatto che i contribuenti sottoposti ad accertamento fiscale, proprio per tale eventualità, potrebbero differire in maniera sistematica da coloro che non hanno subito tale processo.

Primo stadio: modello di selezione

Il primo modello, riconducibile ad un modello probit, ha come risposta una variabile dicotomica dove il valore 1 indica che il contribuente è stato sottoposto ad accertamento unificato e 0 in caso contrario; si evidenzia che il modello stimato non è utilizzato per fini previsivi, ma solo per determinare la probabilità di essere selezionati per un accertamento fiscale sull'anno d'imposta oggetto di studio⁴⁶.

Le probabilità stimate dal modello di selezione sono utilizzate per calcolare un fattore λ , denominato Lambda di Mills, coincidente con l'inverso del rapporto di Mills, da inserire come ulteriore covariata nella successiva fase di stima del Modello al secondo stadio con l'intento di catturare l'effetto sulla variabile obiettivo anche dell'eventualità di essere accertato e tener conto nelle stime quindi della distorsione legata alla selezione non casuale dei contribuenti.

Secondo stadio: modello sostanziale

Il modello di Heckman al secondo stadio prevede un modello di regressione per la stima della Base Imponibile Potenziale IRAP, eseguita solamente sul gruppo dei contribuenti sottoposti a controllo unificato sull'anno d'imposta 2016 e corretta per il coefficiente di distorsione (Lambda di Mills) ottenuto al primo stadio con il modello di selezione.

Dal punto di vista metodologico l'approccio di Heckman è incentrato sull'esistenza di correlazione tra i termini d'errore delle due equazioni supponendo l'influenza di elementi non osservabili direttamente ma che agiscono sia sull'aspetto della selezione che su quello strutturale.

Se i termini d'errore delle due equazioni sono incorrelati, il meccanismo di selezione può essere considerato ignorabile e il metodo di stima dei minimi quadrati ordinari (OLS) fornisce stimatori corretti: la selezione dipende perciò solo dalle variabili ausiliarie che sono poi impiegate per trattare la distorsione.

Se invece i termini d'errore sono correlati la probabilità di selezione è una funzione monotona della variabile obiettivo e quindi la selezione dipende dalla variabile oggetto dell'indagine. Pertanto, lo stimatore OLS è affetto da una distorsione dovuta alla selezione campionaria.

⁴⁶ È importante sottolineare che questa selezione riguarda diversi periodi temporali sui quali hanno insistito diversi momenti economici e diverse direttive dell'Agenzia: i controlli sull'anno d'imposta 2016 sono frutto di selezioni effettuate negli anni successivi alla presentazione della dichiarazione.

Assunzione del modello

L'impostazione fin qui illustrata fa riferimento a un contesto di tipo statistico-inferenziale che presuppone un insieme di assunzioni, non ultimo il fatto stesso di considerare l'insieme degli accertati un campione dell'universo dei dati di analisi e in particolare, un campione distorto.

Il modello di Heckman fornisce una rappresentazione parametrica sia della funzione strutturale che di selezione, assumendo indipendenza tra gli errori e i regressori e la normalità bivariata della distribuzione dei termini d'errore.

Risultati del modello

Le variabili utilizzate nella stima, sia nel primo che nel secondo stadio, derivano dalle informazioni dichiarate fiscalmente dal contribuente in materia IRPEF, IVA e IRAP e da informazioni presenti negli archivi anagrafici che ne caratterizzano il profilo, come la tipologia di soggetto fiscale, la regione del domicilio fiscale, il settore economico.

Una descrizione dettagliata delle variabili incluse nei modelli è riportata nella seguente tabella (Tabella III.2.1.1)⁴⁷:

TABELLA III.2.1.1. VARIABILI UTILIZZATE	
VARIABILI	SIGNIFICATO
Ricavi delle vendite e delle prestazioni	Ricavi delle vendite e delle prestazioni (dichiarazione Unico)
Attività totali	Totale compensi e componenti positive (dichiarazione Unico)
Redditi da lavoro 770	Totale redditi per i quali è possibile fruire della detrazione
Reddito lordo	Reddito lordo (dichiarazione Unico)
Presenza di ulteriori deduzioni	Dummy pari a 1 se il contribuente presenta ulteriori deduzioni; 0 in assenza di deduzioni.
Costi funzionamento	Costi funzionamento (dichiarazione IRAP)
Costi materie prime	Costi materie prime (dichiarazione IRAP)
Costi per servizi	Costi per servizi (dichiarazione IRAP)
Dimensione aziendale	Dummy dimensionali calcolate in base al numero di addetti.
IVA sulle operazioni effettuate nell'anno	Ammontare complessivo dell'IVA sulle operazioni effettuate nell'anno
IVA a credito	Totale IVA a credito (VL39)
Altri acquisti ed importazioni IVA	Altri acquisti ed importazioni IVA
Valore aggiunto fiscale	Differenza tra il volume d'affari dichiarato e l'ammontare complessivo degli acquisti effettuati nell'anno.
Tipologia Soggetto	Dummy pari a 1 se il contribuente è una Società di Persone; 0 altrimenti.
Regime fiscale	Dummy pari a 1 se è presente il reddito d'impresa semplificata; 0 altrimenti
Settore economico	Dummy macrosezione attività prevalente (Industria, Costruzioni, Commercio, Servizi Pubblici e Privati)
Ripartizione Geografica	Dummy Geografiche (Nord Ovest, Nord Est, Centro, Sud e Isole)

I risultati ottenuti dai modelli di Heckman (modello congiunto Persone Fisiche-Società di Persone e modello Società di Capitali), sono molto soddisfacenti, in quanto le correlazioni tra gli errori delle equazioni al primo e secondo stadio sono alte e significative, confermando il vantaggio dell'utilizzo della metodologia di Heckman per correggere la distorsione da selezione dei soggetti sottoposti ad accertamento.

I modelli stimati al primo e secondo stadio, sono caratterizzati da una buona efficacia interpretativa, in particolare: per quanto riguarda l'equazione al primo stadio, influenzano significativamente e risultano discriminanti nel processo di selezione le caratteristiche strutturali del contribuente, quali la ripartizione geografica di appartenenza, il settore di attività economica e la dimensione aziendale in base al numero di addetti. Da un punto di vista dichiarativo, i predittori con un peso maggiore nella stima della probabilità di ricevere un accertamento unificato, sono le attività totali da Modello Unico, il "Costo delle materie prime" dichiarate ai fini IRAP e la variabile

⁴⁷ Nella specificazione dei modelli analizzati, la trasformazione considerata per le variabili continue è quella della Radice Quadrata.

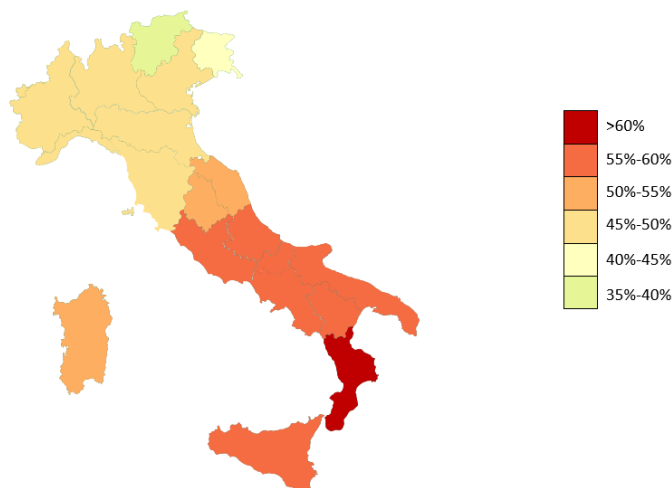
“Altri acquisti e importazioni⁴⁸”, informazione presente nella dichiarazione IVA dove possono essere trascritte anche tutte le operazioni di compensazione a seguito di fatturazioni stornate.

Per quanto riguarda l’equazione di regressione al secondo stadio (modello sostanziale), le variabili con maggior peso e più importanti nella stima della Base Imponibile Potenziale sono riconducibili a grandezze economiche inerenti il volume delle attività dell’impresa, siano esse desunte dalla dichiarazione IVA sia dalla dichiarazione dei redditi Unico.

I modelli di Heckman stimati sono stati utilizzati per ottenere una stima della Base Imponibile Totale e del Gap IRAP delle Società e Ditte Individuali per gli anni che vanno dal 2016 in poi. Naturalmente, partendo dalle previsioni micro relative ai singoli contribuenti, è possibile ottenere una stima aggregata della Base Imponibile Potenziale e del Gap IRAP per qualsiasi dominio di interesse.

Di seguito, si riportano i risultati della propensione al Gap⁴⁹ IRAP a livello regionale⁵⁰ (Figura III.2.1.1). Dalla figura è possibile desumere una scarsa variabilità regionale all’interno della stessa macroarea ma anche una discreta contrapposizione delle regioni settentrionali rispetto all’area meridionale. In sintesi, nelle regioni meridionali, si evince una forte propensione all’evasione rispetto al resto del territorio, con la Calabria con un valore superiore al 60%.

FIGURA III.2.1.1. PROPENSIONE REGIONALE AL GAP IRAP



Questi risultati sono abbastanza allineati con quelli ottenuti tramite la metodologia top down, e questo ci conforta sulla possibilità di utilizzare tali studi al fine di ottenere una maggiore granularità dei dati ufficiali.

⁴⁸ La variabile “Altri acquisti e importazioni imponibile” dal punto di vista dichiarativo riporta tutti gli acquisti di beni e servizi che non siano beni ammortizzabili, beni strumentali non ammortizzabili o beni e/o servizi acquistati e funzionali alla attività economica (rigo VF 27 col 4)

⁴⁹ La propensione al Gap IRAP si ottiene come rapporto dell’ammontare della Base Imponibile Non Dichiarata (gap) su quello della Base Imponibile Potenziale.

⁵⁰ La figura riporta il valore medio della propensione al Gap IRAP calcolato sull’arco temporale 2016-2020 per ripartizione territoriale in livelli.

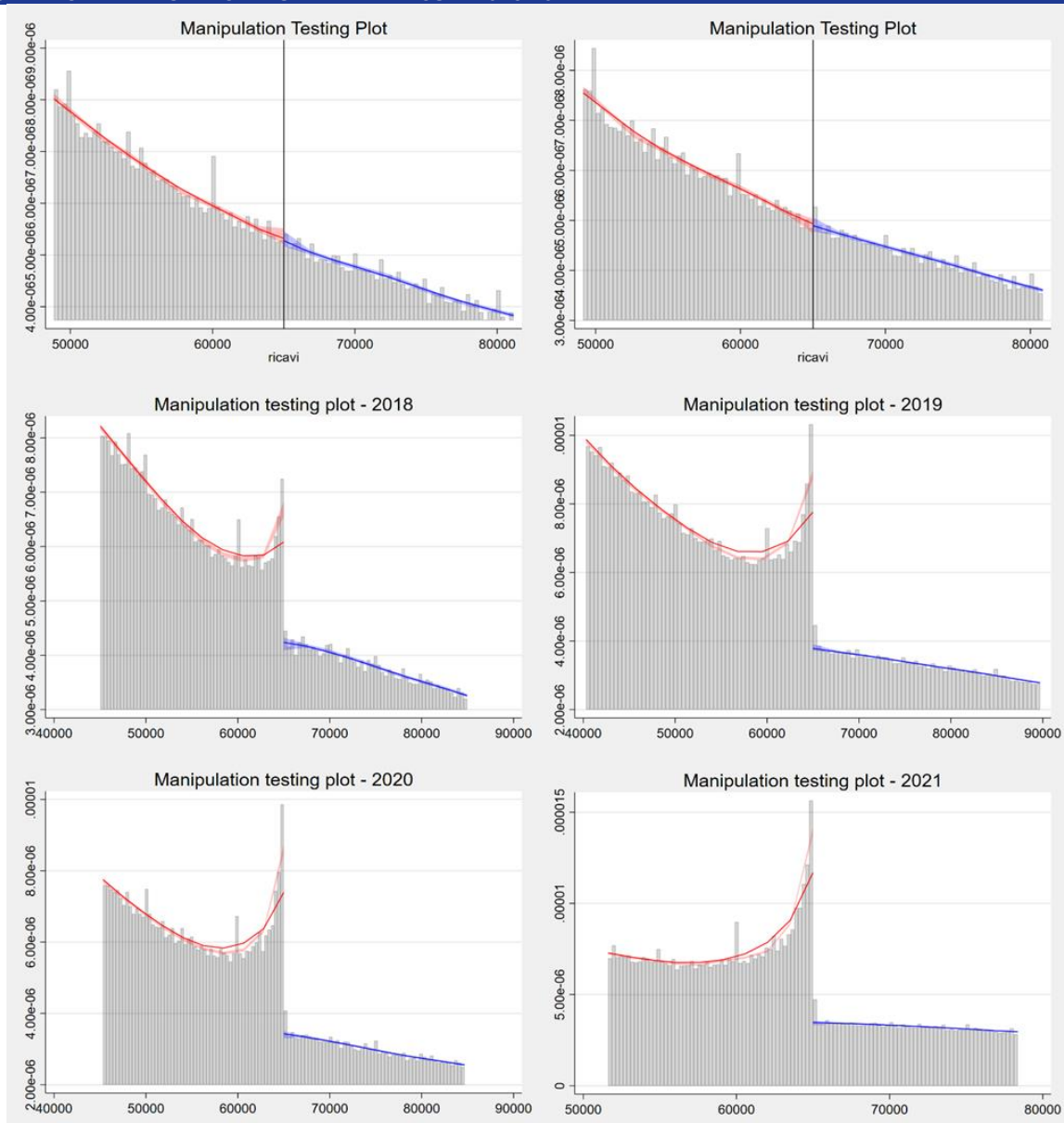
III.3 VALUTAZIONE EX-POST DELL'ESTENSIONE DELLA SOGLIA DI RICAVI NEL NUOVO REGIME FORFETARIO

Analisi della discontinuità nella distribuzione dei ricavi e stima dell'elasticità rispetto alla variazione dell'aliquota marginale

L'analisi si basa sui dati relativi a tutti i contribuenti che compilano i quadri RE (i redditi derivanti dall'esercizio di arti e professioni rientranti nel regime analitico), RG (redditi derivanti dall'esercizio di attività commerciali in contabilità semplificata) e LM (contribuenti che aderiscono al regime di vantaggio e al regime forfetario) negli anni di imposta dal 2016 al 2021.

La Figura III.3.1 mostra la distribuzione dei ricavi stimata con il *manipulation test*, attraverso lo stimatore basato su *kernel density* e regressioni polinomiali locali. L'analisi conferma e rafforza i risultati già emersi sulla base della semplice ispezione grafica e identifica una forte discontinuità nella distribuzione dei ricavi intorno alla soglia dei 65 mila euro per gli anni di imposta 2018-2021. Tale fenomeno, come atteso, non si verifica nel 2016 e nel 2017, anni precedenti all'introduzione della norma.

FIGURA III.3.1 MANIPULATION TEST ANNI D'IMPOSTA 2016-2021



L'analisi di *bunching*, attraverso una stima non parametrica, fornisce una evidenza empirica nelle risposte comportamentali agli incentivi di tipo fiscale. Nella fattispecie, questa tecnica è in grado di analizzare fenomeni in cui la distribuzione di una variabile è soggetta a discontinuità a causa dell'implementazione di una *policy* che genera un incentivo/disincentivo fiscale, alterando i comportamenti degli individui. Il *bunching* è, quindi, particolarmente indicato per analizzare dati amministrativi (perché questi presentano un'ampia numerosità campionaria e bassi errori di misurazione) ed è stata spesso utilizzata per analizzare fenomeni relativi all'offerta di lavoro e all'evasione fiscale.

Si suddividono, quindi, i contribuenti per gruppi di ricavi ("*bins*"); successivamente, si analizza la distribuzione della variabile di interesse e si stima un controfattuale artificiale dato dalla distribuzione che si sarebbe osservata intorno alla soglia in assenza di *policy*, sotto l'ipotesi di assenza di discontinuità (Chetty et al. 2011). Una volta stimata la distribuzione controfattuale artificiale, questa si confronta con la distribuzione osservata nei dati e si calcolano: i) la differenza di densità rispetto al controfattuale artificiale (*excess mass*), cioè il numero di soggetti in eccesso, che si localizzano al di sotto della soglia, rispetto a quello che si osserverebbe in assenza di *policy*; ii) l'elasticità dei ricavi, cioè la variazione percentuale dei ricavi dovuta alla variazione di aliquota, che viene stimata localmente intorno al punto di discontinuità (*kink*).

L'analisi attorno alla soglia dei 65 mila euro è stata eseguita considerando i soggetti con ricavi compresi fra 30 mila euro e 100 mila euro, che sono stati divisi in 140 gruppi corrispondenti a classi di ricavi ordinate in misura crescente e di ampiezza pari a 500 euro. Per il calcolo dell'elasticità, sono state considerate l'aliquota del 15% per i soggetti al di sotto della soglia dei 65 mila euro e l'aliquota IRPEF del 38% per i soggetti con ricavi superiori a 65 mila euro⁵¹.

Nella Figura III.3.2 si riportano i risultati dell'analisi principale⁵². La curva bordeaux rappresenta la distribuzione teorica controfattuale che si sarebbe osservata in assenza di *policy* e di discontinuità, mentre la linea nera rappresenta la distribuzione empirica dei ricavi. L'analisi grafica conferma i risultati del *Manipulation test* e mostra assenza di autoselezione intorno alla soglia dei 65 mila euro (indicata dalla linea verticale) negli anni di imposta 2016 e 2017. Al contrario, si evince una forte autoselezione dei soggetti al di sotto dei 65 mila euro negli anni in cui la normativa era in vigore (dal 2018 al 2021). Nella Figura III.3.2 vengono inoltre riportati i valori della "excess mass", *b*, dell'elasticità, *e*, ed i corrispondenti *bootstrapped standard errors*.

È importante considerare che il valore dell'elasticità che osserviamo approssima solo in parte quella strutturale del fenomeno perché: i) è calcolata localmente nell'intorno della soglia; ii) ci possono essere delle "optimization frictions", cioè dei costi di aggiustamento o inerzia che alterano la reazione degli individui all'incentivo determinando una elasticità osservata che è minore di quella strutturale del fenomeno; iii) la scelta dei parametri (per esempio l'ampiezza delle classi di ricavi, l'area di *bunching*, etc) può influire sulla stima. Al fine di tenere conto di quest'ultimo aspetto, nella Tabella III.3.1 si mostrano le stime di *excess mass* e dell'elasticità al variare dei parametri.

Osservando i valori delle *excess mass*, nelle diverse specificazioni, si nota come nel 2019 in corrispondenza della soglia ci siano fra il 52 e il 69% di contribuenti in più rispetto a quelli che si sarebbero osservati nello scenario controfattuale (cioè, in assenza di *policy* e di discontinuità nella distribuzione). Questo valore cresce nel 2020 e si attesta a valori compresi fra l'85 e il 102% nel 2021. Al contrario, nel 2016 e nel 2017 (quando la soglia dei 65 mila euro non era in vigore) la distribuzione empirica non presenta punti di discontinuità e si sovrappone a quella teorica non presentando significative *bunching mass* in corrispondenza della soglia.

Anche l'elasticità dei ricavi alla soglia cresce dopo l'introduzione della norma attestandosi ad un valore compreso fra il 2.6% e il 3.4% nel 2021.

Nella Figura III.3.3 si mostra, per alcuni settori economici, l'analisi *bunching* relativa alle soglie del vecchio regime forfetario, in vigore fino al 2018. L'analisi è implementata considerando l'aliquota del 15% per i soggetti con ricavi al di sotto della vecchia soglia e quella del 27% per i soggetti con ricavi superiori rispetto alle soglie previgenti⁵³. L'analisi conferma i risultati ottenuti con il *manipulation test*, indicando una significativa autoselezione degli individui al di sotto delle soglie dei ricavi massimi, in particolare per i settori ATECO del "Commercio all'ingrosso e al dettaglio", delle "Attività Professionali" e delle "Altre attività economiche", per cui osserviamo rispettivamente il 15%, il 68% e il 57% di contribuenti in più rispetto al controfattuale artificiale che si osserverebbe in caso di assenza di *policy*. Al contrario, per il settore delle "Attività dei servizi di alloggio e di ristorazione" si osserva una *excess mass* del 5% non significativa.

⁵¹ Per ragioni di ordine pratico, non è stata apportata una distinzione fra soggetti con partita IVA in essere da più di 5 anni (aliquota del 15%) e nuove partite IVA (aliquota del 5%), quindi l'elasticità osservata è da considerarsi leggermente sottostimata rispetto a quella strutturale del fenomeno. La scelta dell'aliquota del 38% dipende dal fatto che applicando il coefficiente di redditività a ricavi pari a 65 mila euro, la maggior parte dei contribuenti sarebbe soggetta ad un'aliquota IRPEF del 38%.

⁵² L'analisi del *bunching* viene sviluppata in R (si veda Mavrokonstantis, P. 2022, Introduction to the bunching Package). La stima controlla per i "round numbers" pari a 1000, cioè per valori di ricavi in corrispondenza dei quali spesso si creano dei punti di discontinuità legati all'arrotondamento dei ricavi dichiarati ma che non costituiscono un punto di discontinuità determinato dalle *policy*.

⁵³ La scelta dell'aliquota del 27% dipende dal fatto che applicando il coefficiente di redditività a ricavi pari alle soglie del precedente Regime forfetario, tutti i contribuenti sarebbero soggetti ad un'aliquota IRPEF del 27%.

FIGURA III.3.2: ANALISI BUNCHING AL KINK DI 65 MILA EURO

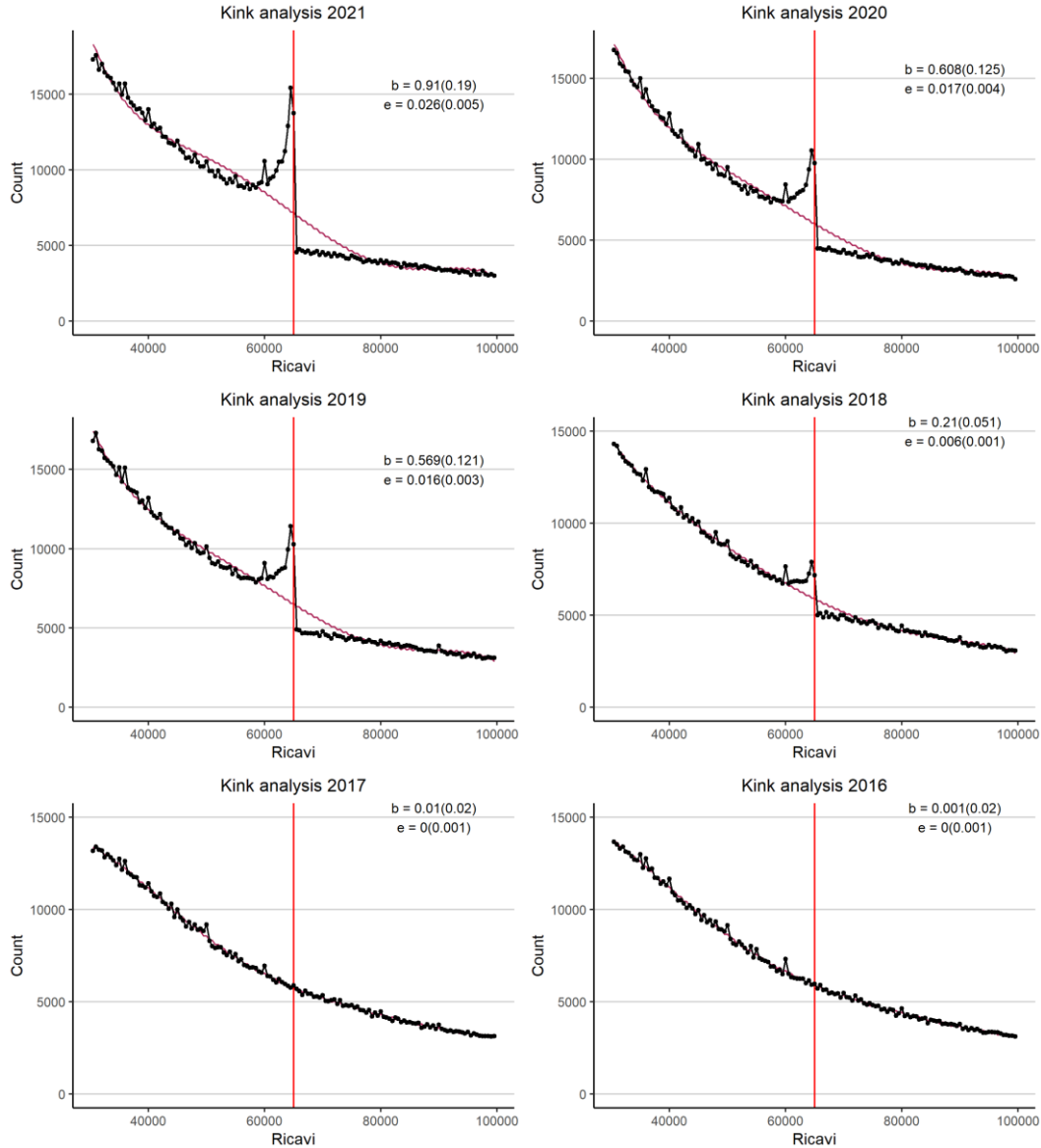
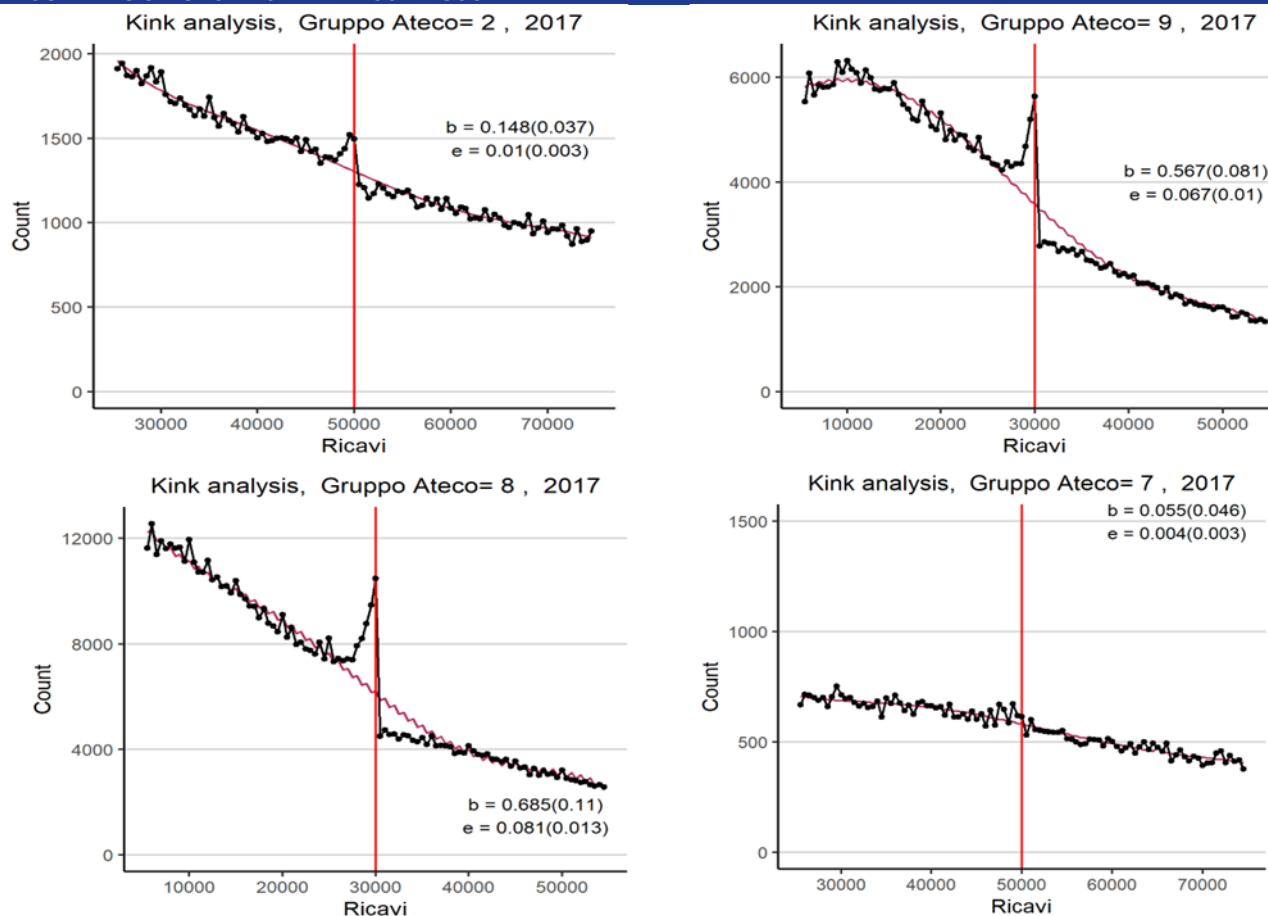


TABELLA III.3.1: ANALISI BUNCHING AL KINK DI 65 MILA EURO – ANALISI DI SENSIVITÀ

N. bins	Ampiezza bins		2016	2017	2018	2019	2020	2021
200	350 euro	Excess mass	0.031 (0.022)	0.014 (0.023)	0.249 (0.08)	0.687 (0.181)	0.731 (0.185)	1.021 (0.304)
		Elasticità	0.001 (0)	0 (0)	0.005 (0.002)	0.014 (0.004)	0.015 (0.004)	0.02 (0.006)
140	500 euro	Excess mass	0.01 (0.02)	0.001 (0.02)	0.21 (0.051)	0.569 (0.121)	0.608 (0.125)	0.91 (0.19)
		Elasticità	0 (0.001)	0 (0.001)	0.006 (0.001)	0.016 (0.003)	0.017 (0.004)	0.026 (0.005)
100	700 euro	Excess mass	0.008 (0.023)	0.018 (0.022)	0.211 (0.055)	0.523 (0.131)	0.569 (0.134)	0.85 (0.208)
		Elasticità	0 (0.001)	0.001 (0.001)	0.008 (0.002)	0.021 (0.005)	0.023 (0.005)	0.034 (0.008)

Nota: In parentesi vengono riportati i bootstrapped standard errors.

FIGURA III.3.3: BUNCHING ALLE VECCHIE SOGLIE

Nota: Settore Ateco 2: "Commercio all'ingrosso e al dettaglio"; Settore Ateco 7: "Attività dei servizi di alloggio e di ristorazione"; Settore Ateco 8: "Attività Professionali"; Settore Ateco 9: "Altre attività economiche".

Stima econometrica dell'impatto del nuovo regime forfetario sulla variazione dei ricavi dichiarati

Ai fini della stima econometrica si è costruito un panel bilanciato per gli anni d'imposta 2016-2021 e si è circoscritta l'analisi ai contribuenti con ricavi inferiori a 100 mila euro e maggiori di 0. Il dataset si riduce a circa 1,2 milioni di contribuenti per ciascuna annualità. L'approccio metodologico utilizzato si basa su una stima panel ad effetti fissi.

Nella stima panel ad effetti fissi la variabile dipendente è rappresentata dal logaritmo dei ricavi dichiarati e la variabile esplicativa di interesse è costituita da una variabile dummy "Forfetari Nuovo Regime" (FNR) che assume valore uno per i contribuenti che aderiscono al nuovo regime forfetario negli anni d'imposta 2019-2021 e zero altrimenti⁵⁴. I risultati delle stime sono riportati nella Tabella III.3.2.

Nella Colonna (1) della Tabella III.3.2 risulta un incremento positivo e significativo, pari al 5,9%, dei ricavi dichiarati da parte dei contribuenti che aderiscono al nuovo regime forfetario rispetto agli altri contribuenti. Nell'ipotesi in cui siano validi gli effetti di variazione della base imponibile dei ricavi dichiarati su entrambe le soglie, ovvero un effetto di incremento associato al venir meno delle soglie precedenti e un effetto di riduzione corrispondente all'istituzione della nuova soglia, tale risultato sembra suggerire la prevalenza del primo effetto rispetto al secondo. In altri termini, la riduzione dei ricavi dichiarati attorno alla soglia dei 65 mila euro risulterebbe

⁵⁴ A tal proposito, occorre sottolineare che l'analisi ad effetti fissi fornisce una stima della variazione dei ricavi tenendo conto delle caratteristiche "time invariant" proprie di ogni contribuente.

più che compensata dall'aumento dei ricavi dichiarati attorno alle precedenti soglie di ricavi, con una prevalenza degli effetti di emersione di base imponibile o di stimolo all'attività economica innescata dall'estensione del regime forfetario.

Tuttavia, nella stima riportata nella Colonna (2) emerge una differenziazione degli effetti rispetto ai vari settori di attività economica. L'impatto risulta, infatti, positivo e significativo soprattutto per le attività professionali (pari al 14,5%), gli intermediari del commercio e il settore delle costruzioni e delle altre attività immobiliari. Viceversa, si ottiene un risultato opposto soprattutto nel caso del commercio ambulante (in media del 28,5%) e del settore dell'alloggio e della ristorazione (pari al 18,5%).

Molto interessanti ai fini dell'analisi differenziata per le soglie di ricavi sono i risultati della stima riportata nella Colonna (3), che confermano gli effetti ipotizzati sul piano teorico e desumibili dall'analisi della discontinuità nella distribuzione dei ricavi. Le classi di ricavi sono state costruite nell'anno d'imposta 2017, ovvero prima che il nuovo regime forfetario producesse i suoi effetti. In particolare, risulta un effetto positivo e significativo di incremento dei ricavi per le classi di ricavi da zero a 20 mila euro (con un aumento del 12,7%) e da 20 mila a 40 mila euro (con un aumento del 4,3%). Viceversa, per le classi di ricavi più elevate si osserva un effetto opposto, ovvero una riduzione di ricavi dichiarati del 6,6% per i contribuenti con ricavi compresi tra 40 mila e 65 mila euro, del 25,5% per i contribuenti con ricavi dichiarati compresi tra 65 mila e 85 mila euro e addirittura del 47,3% per i contribuenti con ricavi dichiarati da 85 mila fino a 100 mila euro. Vala la pena, a tale proposito, precisare che l'impatto sulle classi di ricavi più basse, pur essendo inferiore, risulta nel complesso più rilevante alla luce della maggiore numerosità dei contribuenti che aderiscono al nuovo regime forfetario all'interno di ciascuna classe.

Tuttavia, la stima riportata nella Colonna (4) risulta più esaustiva e affidabile in quanto consente un confronto immediato tra i contribuenti che aderiscono o meno al nuovo regime forfetario. Nella classe di ricavi più bassa si osserva una variazione positiva maggiore per i contribuenti che aderiscono al nuovo regime forfetario; viceversa, nella classe di ricavi compresi tra 20 mila e 40 mila euro la diminuzione dei ricavi dichiarati è inferiore nel caso dei contribuenti che aderiscono al nuovo regime forfetario rispetto agli altri contribuenti. Al contrario, nelle classi di ricavi più elevate la riduzione dei ricavi è inferiore nel caso dei contribuenti che non aderiscono al nuovo regime forfetario.

Infine, dalla stima riportata nella Colonna (5) si evince che anche nel caso delle attività professionali l'effetto di incremento dei ricavi si concentra nelle classi di ricavi inferiori, mentre per le classi di ricavi attorno alla nuova soglia del regime forfetario si osserva sempre e comunque una riduzione dei ricavi dichiarati.

Nel complesso, comunque, i risultati, sebbene molto differenziati per classi di ricavi e settore di attività economica, non sembrano suggerire un effetto di riduzione significativa della base imponibile, in quanto l'impatto positivo che si registra attorno alle precedenti soglie, caratterizzate da una maggiore numerosità di contribuenti, sembra in grado di compensare l'effetto negativo che si osserva generalmente attorno alla nuova soglia di ricavi prestabilita. Ciò nonostante, occorre interpretare questi risultati con molta cautela, in quanto la variabile esplicativa considerata, nonché la strategia econometrica impiegata, sono in grado di catturare soltanto parzialmente l'impatto della riforma, essendo difficile isolare tutti gli altri fattori confondenti.

TABELLA III.3.2: EFFETTO DEL NUOVO REGIME FORFETARIO SUI RICAVI DICHIARATI – STIMA PANEL EFFETTI FISSI -

	(1)	(2)	(3)	(4)	(5)
Forfetari Nuovo Regime (FNR)	0.05883*** 0.001061	-0.01854*** 0.001846			FNR - AMB_ALI - CL_0-20 -0.2362*** 0.032452
FNR - Industrie Alimentari (ALIM)		-0.07083*** 0.017723			FNR - AMB_ALI - CL_20-40 -0.2705*** 0.026913
FNR - Commercio Ingrosso (COMM)		-0.03402*** 0.003776			FNR - AMB_ALI - CL_40-65 -0.4611*** 0.042339
FNR - Comm. Ambulante Alim. (AMB_ALI)		-0.28155*** 0.019805			FNR - AMB_ALI - CL_65-85 -1.0845*** 0.242921
FNR - Comm. Ambulante Altro (AMB_ALT)		-0.29298*** 0.006846			FNR - AMB_ALI - CL_85-100 -0.7333*** 0.130075
FNR - Costruzioni Immobiliare (COSTR)		0.11113*** 0.00324			FNR - AMB_ALT - CL_0-20 -0.2969*** 0.009297
FNR - Intermediari del Commercio (INT)		0.08759*** 0.003536			FNR - AMB_ALT - CL_20-40 -0.3008*** 0.009693
FNR - Alloggio e Ristorazione (RIST)		-0.185*** 0.008405			FNR - AMB_ALT - CL_40-65 -0.4214*** 0.022962
FNR - Attività Professionali (PROF)		0.14509*** 0.002123			FNR - AMB_ALT - CL_65-85 -0.6028*** 0.05249
FNR - Altre Attività Economiche (ALTR)					FNR - AMB_ALT - CL_85-100 -0.8724*** 0.16794
FNR - Classe Ricavi 0-20k (CL_0-20)			0.127478*** 0.001592	0.084039*** 0.001817	FNR - RIST - CL_0-20 -0.1168*** 0.017053
FNR - Classe Ricavi 20k-40k (CL_20-40)			0.04291*** 0.001344	-0.00307*** 0.001502	FNR - RIST - CL_20-40 -0.1971*** 0.010899
FNR - Classe Ricavi 40k-65k (CL_40-65)			-0.06573*** 0.001877	-0.1142*** 0.001987	FNR - RIST - CL_40-65 -0.2864*** 0.015606
FNR - Classe Ricavi 65k-85k (CL_65-85)			-0.25549*** 0.005727	-0.31548*** 0.005778	FNR - RIST - CL_65-85 -0.522*** 0.065676
FNR - Classe Ricavi 85k-100k (CL_85-100)			-0.47262*** 0.01856	-0.54886*** 0.018579	FNR - RIST - CL_85-100 -0.9349*** 0.242693
Non Forfetari - Classe Ricavi 0-20k (CL_0-20)				0.069079*** 0.002061	FNR - PROF - CL_0-20 0.2075*** 0.002174
Non Forfetari - Classe Ricavi 20k-40k (CL_20-40)				-0.03475*** 0.001281	FNR - PROF - CL_20-40 0.1023*** 0.001776
Non Forfetari - Classe Ricavi 40k-65k (CL_40-65)				-0.0788*** 0.001211	FNR - PROF - CL_40-65 -0.0102*** 0.002294
Non Forfetari - Classe Ricavi 65k-85k (CL_65-85)				-0.15012*** 0.001553	FNR - PROF - CL_65-85 -0.1817*** 0.006902
Non Forfetari - Classe Ricavi 85k-100k (CL_85-100)				-0.23256*** 0.002837	FNR - PROF - CL_85-100 -0.3474*** 0.020674
Costante	10.03064*** 0.000503	10.0311*** 0.006301	10.03064*** 0.000502	10.03064*** 0.000503	10.037*** 0.006239
Dummy Anno	√	√	√	√	√
Dummy Ateco		√	√	√	√
Osservazioni	1,216,058	1,216,058	1,216,058	1,216,058	1,216,058
R-quadro	0.000	0.0203	0.0161	0.0571	0.0051

Nota: La variabile dipendente è espressa in logaritmi. La stima nella Colonna (5) riporta soltanto i coefficienti associati alle principali variabili esplicative di maggior interesse. Standard errors robusti in corsivo. *** p<0.01, ** p<0.05, *p<0.1.

III.4 UN MODELLO A CLASSI LATENTI PER L'ANALISI DELLE FRODI IVA

III.4.1 Premessa

Nella presente appendice viene illustrata la metodologia che l'Agenzia delle Entrate sta sviluppando per l'analisi del fenomeno delle frodi IVA, in particolare da *missing trader*. Si tratta di un approccio di tipo non supervisionato che prevede l'utilizzo di modelli a classi latenti volti a profilare i contribuenti in termini di propensione all'evasione IVA ad "alto rischio", e si affianca all'approccio utilizzato finora, di tipo supervisionato.

La metodologia supervisionata, i cui risultati sono presentati nella Relazione Evasione, si basa sull'utilizzo di dati di controllo derivanti da verifiche e accertamenti unificati relativi ai cinque anni precedenti l'anno di analisi, e risulta costituita da una platea di riferimento che include tutti i soggetti che non presentano dichiarazione IVA. Al contrario, il metodo non supervisionato non utilizza le informazioni relative ai dati di controllo, consentendo così da un lato di aumentare la tempestività della stima e, dall'altro, di considerare anche i meccanismi fraudolenti di coloro che presentano dichiarazione IVA ma non versano l'IVA all'Erario. L'approccio non supervisionato consente quindi una rappresentazione dei fenomeni ad elevato rischio di frode sull'intera platea delle partite IVA, profilandosi come un efficace strumento di analisi del rischio.

Per meglio catturare e analizzare la pluralità di comportamenti evasivi, sono stati formalizzati due diversi modelli a seconda della presenza della dichiarazione (versamento) IVA; in particolare:

- *Modelli di mistura per variabili miste*, per analizzare i soggetti che non hanno presentato una dichiarazione IVA o che non hanno effettuato versamenti nell'anno di imposta oggetto dello studio;
- *Modelli mixtures of experts*, per analizzare i soggetti che hanno presentato la dichiarazione IVA o che hanno effettuato versamenti.

I modelli sono stati implementati in R utilizzando i pacchetti *clustMD* (McParland and Gormley, 2015) e *MoEClust* (Murphy and Murphy, 2020).

In ogni caso, la metodologia impiegata si avvale di tecniche di cluster *analysis* di tipo *model-based*: le diverse componenti del modello mistura sono associate ad altrettanti cluster (classi di contribuenti caratterizzate da diversi livelli di rischio) corrispondenti a distinte distribuzioni (eventualmente condizionate) di probabilità nell'ambito di una stessa famiglia parametrica (in questo caso Gaussiana).

III.4.2 Il modello

La metodologia proposta ha l'obiettivo di stimare il valore dell'IVA che dovrebbe essere dichiarata e/o versata in caso di perfetta *compliance* sulla base delle informazioni desumibili principalmente dai dati della fatturazione elettronica, dei corrispettivi inviati telematicamente e delle bollette doganali. A causa della complessità della normativa fiscale e dell'eventuale incompletezza dei dati disponibili, non è tuttavia sempre possibile effettuare una ricostruzione esatta dell'IVA dovuta o a credito; viene pertanto adottato un approccio statistico a più stadi.

Nel primo stadio, l'imposta teorica (l'IVA che dovrebbe essere dichiarata e/o versata in caso di perfetta *compliance*) è espressa, mediante un modello di regressione, in funzione dell'imposta ricostruita da fatture elettroniche e corrispettivi telematici.

In termini formali, il modello di primo stadio può essere specificato come segue:

$$Y^* = \alpha + \beta X + U,$$

dove Y^* è l'imposta teorica che dovrebbe essere dichiarata e/o versata in caso di *compliance*, X è l'IVA ricostruita da fatture e corrispettivi, U è il termine di errore che si assume seguire una distribuzione Gaussiana a media zero e varianza σ^2 , e α e β sono i parametri da stimare.

Nel secondo stadio, si ipotizza che l'insieme dei contribuenti sia suddiviso in due sottoinsiemi: un sottoinsieme composto da contribuenti *compliant*; un altro sottoinsieme in cui i contribuenti adottano comportamenti evasivi che si concretizzano in omesse o infedeli dichiarazioni e/o mancati o insufficienti versamenti. Si assume, inoltre, che quest'ultimo insieme si suddivida a sua volta in gruppi G_k ($k = 1, \dots, K - 1$), ciascuno corrispondente ad un determinato livello di *intensità* di rischio fiscale.

Per ogni contribuente, la probabilità π_k ($k = 0, \dots, K - 1$) di appartenere al k -esimo cluster G_k (probabilità *a priori*) può dipendere, a seconda che il contribuente dichiari (versi) oppure no, da un insieme di variabili di profilazione Z attraverso una funzione logistica secondo la relazione: $\log(\pi_k/\pi_0) = b'Z$, dove b' è un vettore di parametri e $k = 0$ identifica l'assenza di evasione; in tal modo, il cluster G_0 rappresenta il gruppo di riferimento.

Il modello risultante è un modello di tipo *logit-multinomiale*, in cui le variabili I_k , indicatrici di appartenenza ai diversi cluster ($I_k = 1$ in caso di appartenenza a G_k e $I_k = 0$ altrimenti), non sono osservate (variabili latenti).

Il terzo e ultimo stadio della formulazione del modello consiste nella specificazione della distribuzione dell'imposta effettivamente dichiarata (versata) condizionatamente al corrispondente valore dell'importo corretto. Tale distribuzione condizionata ha carattere semi-continuo a causa dell'inflazione di zeri determinata dai casi di omessa dichiarazione o omesso versamento; per questo motivo, vengono implementati due diversi modelli a seconda della presenza di valori dichiarati/versati.

Specificamente, i dati relativi ai soggetti non dichiaranti (o che non versano) vengono modellati mediante una opportuna distribuzione di probabilità per variabili miste (cioè sia quantitative che categoriche) che, previa ricodifica delle variabili stesse, può essere ricondotta ad una mistura finita di Gaussiane.

Per i contribuenti per cui sono presenti i valori dichiarati o versati, si specifica un "modello di evasione" in cui l'IVA dichiarata/versata viene espressa in funzione dell'imposta che dovrebbe essere dichiarata/versata in caso di perfetta *compliance*:

$$Y = \delta_k + \gamma_k Y^* + \epsilon_k,$$

dove ϵ_k è una variabile aleatoria distribuita secondo una gaussiana a media zero e varianza ξ_k^2 , indipendente da U , e δ_k e γ_k sono parametri da stimare.

Dalle assunzioni fatte, discende che la distribuzione delle variabili "manifeste" Y può essere espressa come mistura di K distribuzioni gaussiane:

$$f(Y|x, z; \theta) = \sum_{k=0}^{K-1} \pi_k(z; w) N(y; \mu_k, \sigma_k^2) \quad \text{con} \quad \sum \pi_k = 1$$

dove θ è il vettore dei parametri del modello ($\alpha_k, \beta_k, \sigma_k^2$) e $\pi_k(z; w)$, detti *mixing weights*, sono le probabilità a priori determinate dal modello *logit-multinomiale* specificato sopra.

Nel caso di omessa dichiarazione (o versamento), le variabili manifeste Y sono le stesse variabili che caratterizzano i diversi profili di rischio, mentre i pesi π_k sono costanti da stimare (non ci sono variabili Z).

Nel caso di importo dichiarato (versato) diverso da zero, sarà proprio quest'ultimo a giocare il ruolo di variabile manifesta Y e i *mixing weights*, $\pi_k(z; w)$, saranno funzioni delle variabili Z . Inoltre, in questo caso, le medie μ_k della variabile Y nei cluster saranno funzioni lineari del valore dell'imposta x ricostruito da fatture e corrispettivi (*mistura di regressioni*): $\mu_k = \alpha_k + \beta_k x$.

In entrambi i casi citati, i parametri del modello sono stimati con il metodo della massima verosimiglianza mediante l'*algoritmo EM (Expectation-Maximization)*. Quest'ultimo si compone di due passi: nel primo passo viene calcolato il valore atteso del logaritmo della verosimiglianza utilizzando i valori correnti dei parametri (*Expectation*), nel secondo passo viene massimizzato tale valore atteso e dunque viene aggiornata la stima dei parametri (*Maximization*).

Il numero totale di cluster viene determinato utilizzando il *BIC* (*Bayesian Information Criterion*), un indicatore della bontà di adattamento del modello, e sulla base di scelte che rendano possibile l'interpretazione dei risultati.

Al fine di analizzare in maniera più appropriata le caratteristiche del fenomeno delle frodi IVA, sono addestrati modelli distinti per dichiarazione IVA attiva (cessioni) e passiva (acquisti), e per i versamenti.

Esempio: modello per l'IVA sulle operazioni attive

A titolo d'esempio, si consideri il caso in cui la variabile target Y^* è l'imposta che dovrebbe essere dichiarata dal contribuente relativamente alle operazioni attive. Per quanto riguarda il modello per non dichiaranti, sono state selezionate le seguenti variabili di profilazione:

- totale IVA esposta in fattura e corrispettivi, espresso in scala logaritmica;
- rapporto tra l'IVA relativa agli acquisti e quella relativa alle cessioni;
- variabile dicotomica relativa alla data di attribuzione della partita IVA (tale variabile assume il valore 1 se la partita IVA è stata aperta negli ultimi due anni rispetto all'anno in esame e il valore 0 altrimenti);
- variabile dicotomica relativa all'iscrizione alla banca dati VIES (tale variabile assume il valore 1 se la partita IVA è presente in VIES e 0 altrimenti);
- differenza tra l'importo dell'IVA periodica dovuta e versata.

Per quanto riguarda invece il modello per coloro che hanno presentato la dichiarazione IVA, sono state considerate le seguenti variabili di profilazione Z :

- rapporto tra l'IVA relativa agli acquisti e quella relativa alle cessioni;
- variabile dicotomica relativa alla data di attribuzione della partita IVA;
- variabile dicotomica relativa all'iscrizione alla banca dati VIES.

Inoltre, come variabile X è stata utilizzata l'IVA sulle operazioni attive ricostruita sulla base delle informazioni derivanti dalla fatturazione elettronica e dai corrispettivi; invece come variabile Y è stato utilizzato il totale dell'imposta dichiarata sulle operazioni attive (quadro VE26 del modello dichiarativo).

III.4.3 Gli output

La clusterizzazione di tipo *model-based* dei soggetti analizzati consente di attribuire al soggetto i -esimo, la probabilità a posteriori di appartenere al k -esimo cluster G_k (classificazione Bayesiana) ovvero al gruppo associato al k -esimo livello di rischio:

$$\tau_k(y_i | z_i, x_i; \theta) = \frac{\pi_k(z_i) N_k(y_i)}{\sum_{h=0}^{K-1} \pi_h(z_i) N_h(y_i)}, \quad k = 0, \dots, K-1$$

dove le variabili Y assumono il significato appropriato a seconda che siano presenti i valori dichiarati oppure no, e, in quest'ultimo caso, le variabili Z sono omesse.

Partendo dalla specificazione del modello e utilizzando le regole algebriche della teoria della distribuzione Normale multivariata, è possibile ricavare i valori attesi condizionati della variabile target Y^* , cioè il valore corretto che il contribuente i -esimo dovrebbe dichiarare o versare:

$$E(Y_i^* | y_i, z_i, x_i; \theta) = \sum_{k=0}^K \tau_k(y_i, z_i, x_i; \theta) E(Y_i^* | y_i, z_i, x_i, i \in G_k; \theta).$$

Sulla base del valore atteso della variabile target Y^* è possibile ottenere una stima (previsione) del *gap totale individuale*

$$gap_i \equiv \pm(E(Y_i^* | y_i, z_i, x_i; \theta) - Y_i)$$

e delle componenti del gap associate ai vari cluster

$$gap_i^k \equiv \pm(\tau_k(y_i|z_i, x_i; \theta)E(Y_i^*|y_i, z_i, x_i, i \in G_k; \theta) - Y_i).$$

Nelle precedenti formule, il segno \pm dipende dalla natura della variabile Y_i che stiamo considerando: segno positivo (+) nel caso dell'IVA sulle operazioni attive (cessioni), segno negativo (-) nel caso dell'IVA sulle operazioni passive (acquisti).

La predetta profilazione può essere effettuata relativamente all'imposta dichiarata sulle operazioni attive o passive, ovvero all'imposta versata. I singoli profili possono essere unificati in modo da ottenere una profilazione globale del contribuente.

III.5 DERIVAZIONE DEL MODELLO ECONOMETRICO PER L'ETEROGENEITÀ DEGLI EFFETTI DELLA FATTURAZIONE ELETTRONICA

In questa appendice si descrive il modello utilizzato per l'analisi riportata al paragrafo V.I.3. In particolare, riprendendo il modello utilizzato nell'edizione 2020 della Relazione per valutare l'impatto di fatturazione elettronica (FE) sull'IVA esigibile, se ne illustrano gli sviluppi realizzati. Il modello, nella formulazione originaria, è rappresentato formula che segue:

$$(Z_2 - Z_1) = \alpha + \gamma q + \lambda'U + \varepsilon \quad (1)$$

dove Z_1 e Z_2 rappresentano (in scala logaritmica) l'IVA esigibile dichiarata nelle LIPE del primo semestre del 2018 e 2019 rispettivamente, U rappresenta un insieme di variabili di controllo⁵⁵, q identifica il rapporto tra il totale dell'IVA associata alle cessioni di beni e/o prestazioni di servizi di tipo B2B e il totale dell'IVA esigibile risultante dalle comunicazioni IVA relative allo stesso semestre.

Per quanto concerne la variabile q , in particolare, essa può assumere valori compresi tra 0 e 1, dove per valori pari, o prossimi, all'unità si è in presenza di un comportamento dichiarativo orientato verso la fatturazione elettronica mentre, per valori pari, o prossimi, allo zero si è in presenza di un gettito IVA del contribuente derivante da operazioni certificate tramite corrispettivi. Il termine ε , infine, rappresenta il residuo casuale che soddisfa le usuali ipotesi di incorrelazione e omoschedasticità, e $(\alpha, \gamma, \lambda)$ l'insieme di parametri da stimare.

Rispetto al modello di cui sopra, sono state apportate alcune modifiche al fine di verificare se il comportamento delle diverse tipologie di contribuenti fosse omogeneo. I risultati di tali estensioni sono sintetizzati nel paragrafo V.I.3 e dettagliati nel seguito.

La prima modifica apportata ha riguardato l'introduzione di una specifica variabile K per poter individuare i soggetti IRPEF, ossia persone fisiche e società di persone. Ciò ha, quindi, permesso di poter distinguere due platee di soggetti (IRPEF, noIRPEF) e di analizzarne, separatamente, il comportamento a seguito dell'introduzione della FE. Più precisamente, la variabile K è una variabile dicotomica, o *dummy*, che assume valore 1 in presenza di soggetti IRPEF e 0 altrimenti. Considerata la seguente riformulazione del modello:

$$(Z_2 - Z_1) = \alpha + \gamma q + \lambda'U + \delta K + \eta (K * q) \varepsilon \quad (2)$$

il coefficiente δ rappresenta la differenza di intercetta dei soggetti IRPEF rispetto ai soggetti noIRPEF, mentre il coefficiente η associato all'interazione tra K e q rappresenta la differenza di

⁵⁵ Le variabili di controllo sono inserite per sterilizzare l'effetto indotto da fattori esogeni alla FE quali, ad esempio, il settore di attività economica di appartenenza o alla localizzazione geografica.

pendenza. In altri termini quest'ultimo parametro identifica il diverso comportamento dei soggetti IRPEF in relazione alla fatturazione elettronica.

I risultati così ottenuti sono riportati nella Tabella III.5.1 che riporta i valori delle stime dei parametri, con i relativi errori standard, per le due sotto-platee di soggetti (IRPEF, noIRPEF).

TABELLA III.5.1: STIME DEI PARAMETRI DEL MODELLO 2 PER LA SEZIONE G		
Parametri	Coefficiente	Std.Error
q	0,0299***	0,0033
q*K	-0,0226***	0,0032
Costante	-0,0329***	0,0058
K	0,0037	0,0023
dummy settoriale	si	Si
dummy regionale	si	Si
N	393.147	

Variabile dipendente: IVA esigibile
K: flag che vale 1 per i soggetti Irpef e 0 altrimenti

Dai risultati emersi si desume un differente effetto della FE per le due sotto platee. Nello specifico, poiché il parametro associato al flag *K* risulta non significativo, i due gruppi condividono la stessa intercetta. D'altro canto, il parametro relativo all'interazione tra la variabile *q* e la variabile *K*, che, come richiamato in precedenza rappresenta la differenza tra il valore stimato del coefficiente della variabile *q* (parametro γ , positivo e circa pari a 0,03) e quello dei soggetti IRPEF, risulta negativo e pari -0,023. Ciò si traduce in un extraggettito pari all' 1,04% e al 0,6% per i soggetti, rispettivamente, non IRPEF e IRPEF. I risultati ottenuti permettono, dunque, di concludere che **l'extraggettito IVA riscontrato nella relazione evasione 2020, pari al 2%, sia da ascrivere principalmente al comportamento dei soggetti "noIRPEF"**.

Successivamente, entrambi i modelli sopra definiti nelle equazioni (1) e (2) sono stati utilizzati anche per valutare l'impatto della FE sul reddito da impresa e lavoro autonomo dichiarato, al fine di comprendere se il comportamento dell'intera popolazione considerata, e delle due sotto-platee individuate (IRPEF, noIRPEF), fosse coerente tra le due imposte considerate (IVA-IRPEF). La formulazione dei due modelli rimane, quindi, invariata, con l'unica eccezione della variabile di output *Z* che in questo caso non rappresenta l'IVA esigibile bensì il reddito da impresa e lavoro autonomo dichiarato⁵⁶.

Le Tabelle III.5.2 e III.5.3 riportano le risultanze del modello applicato sul reddito da impresa e lavoro autonomo. Nel dettaglio, i risultati in Tabella III.5.2 definiscono il comportamento dell'intera popolazione considerata, mentre nella Tabella III.5.3 si riportano i parametri per le singole sotto platee.

TABELLA III.5.2: STIME DEI PARAMETRI DEL MODELLO 1 PER LA SEZIONE G		
Parametri	Coefficiente	Std.Error
Q	-0,1096***	0,0174
Costante	-0,2023***	0,0459
dummy settoriale	si	Si
dummy regionale	si	Si
N	393.147	

Variabile dipendente: Reddito da impresa e lavoro autonomo

⁵⁶ Si tenga presente come l'utilizzo del reddito da impresa e lavoro autonomo, e non dell'imposta, come variabile di output nel modello sull'IRPEF non distorce il confronto con l'IVA. In quest'ultimo caso, infatti, poiché l'aliquota media e mediana IVA per i due anni considerati (2018-2019) rimane pressoché costante e pari al 22%, replicare il modello sulle basi produrrebbe risultati equivalenti. D'altra parte, invece, l'utilizzo dell'imposta nel modello sui redditi rischierebbe di essere influenzato da altri fattori (deduzioni e detrazioni) non riguardanti l'attività d'impresa.

TABELLA III.5.3: STIME DEI PARAMETRI DEL MODELLO 2 PER LA SEZIONE G

Parametri	Coefficiente	Std.Error
q	-0,0081	0,0279
q*K	-0,0724 **	0,0266
Costante	-0,4249***	0,0492
K	0,2349***	0,0193
dummy settoriale	si	si
dummy regionale	si	si
N	393.147	

Variable dipendente: Reddito da impresa e lavoro autonomo
K: flag che vale 1 per i soggetti Irpef e 0 altrimenti

I risultati ottenuti per l'intera platea mostrano un effetto FE sul reddito da impresa e lavoro autonomo contrario rispetto a quello sull'IVA esigibile (Tabella III.5.2). Il coefficiente della variabile q assume valore negativo e pari a -0,11, al quale corrisponde una contrazione del reddito da impresa e lavoro autonomo pari al 6,58%.

I risultati riportati nella Tabella III.5.3 indicano che l'effetto FE negativo riscontrato sull'intera platea è in realtà riconducibile al solo comportamento dei soggetti IRPEF. Il coefficiente relativo alla variabile q risulta essere non significativo, mentre quello dell'interazione con variabile *dummy* assume valore negativo determinando un effetto FE pari a -0,0805 che si traduce in una contrazione del reddito da lavoro autonomo e di impresa pari al 3,31%.

IV. RIFERIMENTI BIBLIOGRAFICI

Braiotta A., Carfora A., Pansini R.V., (2015), "Tax Gap and Redistributive Aspects across Italy", Argomenti di Discussione dell'Agenzia delle Entrate, n.2.

Braiotta A., Pisani S., Pisano E. (2013), "Evasione fiscale e distribuzione primaria del reddito", Rivista di diritto finanziario e scienza delle finanze, n. 2.

D'Agosto E., Marigliani M, Pisani S., (2016) "A general framework for measuring Vat compliance in Italy". Argomenti di discussione Agenzia delle Entrate, N.02/2016.

D'Agosto E., Marigliani M, Pisani S., (2014) "Asymmetries in the territorial VAT gap". Argomenti di discussione Agenzia delle Entrate, N.02/2014.

Das-Gupta A., Mookherjee D. (2000), Incentives and institutional reform in tax enforcement an analysis of developing country experience, Lavoisier.

European Commission (2020), "Study and reports on the VAT gap in the EU-28 Member States 2020 final report".

Eurostat (2010) European System of Account - ESA 2010, Luxembourg.

Fiscalis Tax Gap Project Group (2018), The concept of tax gaps - Report II: Corporate Income Tax Gap Estimation Methodologies, Directorate General Taxation and Customs Union, European Commission.

Fiscalis Tax Gap Project Group (2018), "The concept of tax gaps - Report III: MTIC Fraud Gap estimation methodologies", Directorate General Taxation and Customs Union, European Commission

Gallucci M., Pansini R.V., Pisani S. (2020), Direct taxes gap estimates: methodology and preliminary results, Argomenti di Discussione dell'Agenzia delle entrate, n. 02/2020.

Heckman, James (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models". *Annals of Economic and Social Measurement*, 5 (4): 475-492.

HM Revenue & Customs (2021), "Measuring tax gaps 2021 edition - tax gap estimates for 2019 to 2020", March, London

Imperioli L., Manzo M., Monteduro M. T. (2015), Il tax gap sulle imposte immobiliari e la perequazione fiscale, XXXVI Conferenza Italiana di Scienze Regionali.

Liberati C.; Pisani S.; Serra A. (2002) Nota metodologica per effettuare un confronto tra i dati IRAP e di Contabilità Nazionale nel settore privato Documento di lavoro dell'Ufficio Studi, n. 2, Agenzia delle Entrate.

Little, Roderick J. A. (1988). "Missing-Data Adjustments in Large Surveys". *Journal of Business & Economic Statistics*. 6 (3): 287-296.

Mavrokonstantis P.; Seibold A. (2022), "Bunching and adjustment costs: Evidence from cypriot tax reforms", *Journal of Public Economics* 214: 104727

McParland D., Gormley I. C. (2015). "Model based clustering for mixed data: clustMD".

Murphy K., Murphy T. B. (2020). “Gaussian parsimonious clustering models with covariates and a noise component”. *Advances in Data Analysis and Classification*. 14 (2): 293-325.

Pisani S., (2014), “An approach to assess how the activity of the Italian Revenue Agency affects compliance”, *Argomenti di Discussione dell’Agenzia delle Entrate*, n.1.

Pisani S., Polito C (2006), “Metodologia di integrazione tra i dati IRAP e quelli di Contabilità Nazionale”, *Agenzia delle Entrate, Ufficio Studi, documenti di lavoro*

Reckon, (2009), “Study to quantify and analyze the VAT gap in the EU-25 Member States”, *Working paper Reckon*, 21 September, London, UK.

Ueda J. (2018), *Estimating the Corporate Income Tax Gap: the RA Gap Methodology*, IMF Technical Notes and Manual, August, Washington.