

L'AVVENTO DEI BIG DATA E DELLE NUOVE TECNOLOGIE: UN'OCCASIONE PER RIVEDERE I METODI DI STIMA DELL'EVASIONE FISCALE

di Roberta Corriere*

SOMMARIO: 1. Premessa. – 2. I metodi di stima indiretti dell'evasione fiscale e l'approccio di tipo top down adoperato per la predisposizione della Relazione sull'economia non osservata e sull'evasione fiscale e contributiva. – 3. I metodi di stima diretti dell'evasione fiscale. L'approccio di tipo bottom up. – 4. Effettività, attualità e precisione. Gli elementi disattesi nelle stime ufficiali poste a fondamento delle politiche fiscali adottate dal nostro Paese. – 5. Big data e Intelligenza Artificiale. Riflessioni conclusive in merito alla possibilità (non troppo remota) di un cambio di approccio nella stima del tax gap.

Abstract: Un'efficace lotta all'evasione fiscale richiede la reale conoscenza del fenomeno e suggerisce l'abbandono di metodi di calcolo, come quelli *top down*, basati su assunti eccessivamente teorici, per accogliere con maggior favore metodi di tipo *bottom up*. Questi infatti permetterebbero di acquisire informazioni utili a migliorare, dal punto di vista qualitativo e quantitativo, il processo di selezione dei contribuenti da sottoporre a verifica. Con l'avvento dell'era dei *Big data* e dell'Intelligenza Artificiale i tempi sembrano maturi per mettere a punto sistemi di rilevazione non scollegati dalla realtà fattuale, affinché le stime possano realmente essere d'ausilio per la messa a punto di politiche fiscali efficienti ed efficaci di contrasto all'evasione, plasmate su contesti economico-sociali effettivi ed attuali.

Abstract: An effective struggle against tax evasion requires an in-depth knowledge of the phenomenon, involving the abandonment of estimation methods, such as the top-down, based on excessively theoretical assumptions, to move to a bottom-up method, which allows acquiring useful information to improve, from a quantitative and qualitative point of view, the identification of those taxpayers to submit to tax audits. In the era of Big Data and A.I. the time has come to create estimation systems that are more closely linked to factual reality, to develop efficient and effective tax policies against tax evasion, based on real and current economic-social contexts.

* Professoressa a contratto di Diritto Tributario, presso l'Università degli Studi di Perugia, Dottore di ricerca in Diritto Tributario presso Sapienza Università di Roma.

1. Premessa

La misurabilità del fenomeno dell'evasione fiscale rappresenta un tema fortemente dibattuto, che tuttora divide la letteratura economica e i governi dei diversi Paesi.

Nel tentativo di quantificarne la portata, l'aspetto più delicato investe l'approccio metodologico che si ritiene più corretto adottare.

I metodi per la stima dell'evasione fiscale possono riassumersi all'interno di due macro-categorie; quelli di tipo deduttivo, definiti anche *top down* o indiretti, e quelli di tipo induttivo, definiti anche *bottom up* o diretti¹.

Sulla base di queste premesse, risulta determinante la scelta del sistema che meglio sia in grado di rappresentare la realtà. Differenti approcci, infatti, sono in grado di generare differenti risultati, sopra o sotto stimando il fenomeno che si intende analizzare. Si tratta di un aspetto di grande rilevanza, dal momento che i risultati ottenibili sono in grado di influenzare le decisioni del legislatore e vengono, il più delle volte, presi a giustificazione e a sostegno delle politiche fiscali che, di volta in volta, il Governo intende attuare.

In questo senso non convincono i metodi (indiretti) che sino ad oggi sono stati utilizzati dal nostro Paese nelle stime ufficiali, in linea peraltro con quelli della maggior parte dei Paesi europei.

Si tratta di approcci eccessivamente approssimativi e spesso disancorati dalla realtà fattuale cui il calcolo dovrebbe riferirsi.

In questo senso, le nuove tecnologie sembrano fornire l'occasione per rivedere le tecniche di misurazione del fenomeno in senso conforme alle logiche che sono alla base dei metodi diretti, indubbiamente idonei a fornire valutazioni più attendibili e precise.

2. I metodi di stima indiretti dell'evasione fiscale e l'approccio di tipo top down adoperato per la predisposizione della Relazione sull'economia non osservata e sull'evasione fiscale e contributiva

L'approccio di tipo *top down* è quello che a livello internazionale attualmente ha riscosso maggiore successo.

La facilità di reperimento del dato, la standardizzazione delle procedure e

¹ M. MARÈ, A. ALESINA, *Evasione e debito*, in A. MONORCHIO (a cura di), *La finanza pubblica italiana dopo la svolta del 1992*, Bologna, 1996, pp. 69-112; M. MARÈ, *L'evasione in Italia e nei paesi Ocse: evidenze empiriche, determinanti ed effetti economici*, in *Moneta e credito*, 1996, pp. 393-443; OECD, *Measuring the Non-Observed Economy: a Handbook*, Paris, 2002; R. ZIZZA, *Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano*, Tema di discussione n. 463, Banca d'Italia, 2002.

la più agevole applicazione² sono fattori che indubbiamente hanno contribuito alla sua predilezione rispetto a metodologie di tipo *bottom up*, più costose e particolareggiate.

Il metodo si basa sul confronto tra i dati fiscali e un corrispondente aggregato macro-economico (solitamente rappresentato dai flussi di Contabilità nazionale), che incorpora al suo interno una stima dell'economia non osservata (sommersa).

Tale aggregato viene utilizzato al fine di costruire una base imponibile teorica onnicomprensiva, con la quale confrontare la base imponibile dichiarata dalla platea dei contribuenti.

Si tratta di metodologie facilmente applicabili, poiché basate su modelli statistici e sull'impiego di dati informativi facilmente reperibili³.

² Su questi aspetti si v. E. DE SIMONE, *Tax gap ed enti locali*, in *Osservatorio regionale sul Federalismo Fiscale – IRPET*, n. 1, 2017, p. 8; GRUPPO DI LAVORO “ECONOMIA NON OSSERVATA E FLUSSI FINANZIARI”, *Rapporto finale sull'attività*, disponibile al seguente indirizzo online http://www.tesoro.it/primopiano/documenti/2012/economia_non_osservata_e_flussi_finanziari_rapporto_finale, 2011; F. SCHNEIDER, D.H. ENSTE, *Shadow Economies: Size, Causes, and Consequences*, in *Journal of economic literature*, vol. 38, n. 1, march 2000.

³ L'approccio di tipo *top down* è contraddistinto da una pluralità di modelli, che utilizzano strumenti differenti per determinare la stima dell'economia non osservata. Tali modelli possono essere suddivisi in due categorie a seconda che si basino su approcci economico-statistici ovvero su approcci macroeconomici. I modelli economico-statistici sono stati oggetto di critica sia in sede Eurostat sia OCSE, in quanto fondati su assunzioni eccessivamente approssimative. Di questi fa parte ad esempio il *Currency Demand Approach* (CDA), anche detto approccio monetario; un metodo di stima econometrica della domanda di contanti fondato sull'assunto che le transazioni in nero avvengono verosimilmente con denaro contante, in modo tale da evitarne la tracciabilità. Tale metodo consentirebbe di quantificare l'evasione stimando l'eccesso della domanda di moneta, rispetto alla preferenza *standard* per la liquidità, sulla base dell'assunto che, all'incremento del sommerso, debba parallelamente corrispondere un incremento della domanda di circolante. Originariamente introdotto da Phillip Cagan (per il quale si v. P. CAGAN, *The Demand for Currency Relative to the Total Money Supply*, in *Journal of Political Economy*, vol. 66, 1958, pp. 303-328), tale metodo è stato in seguito affinato da Tanzi (per il quale si v. V. TANZI, *The underground economy in the United States: estimates and implications*, *Banca Nazionale del Lavoro Quarterly Review*, 135, 1980, pp. 427-453). Per le successive rielaborazioni del modello si v. J. FERWERDA, I. DELEANU, B. UNGER, *Reevaluating the Tanzi-Model to Estimate the Underground Economy*, *Tjalling C. Koopmans Research Institute, Discussion Paper Series, Utrecht School of Economics Working Papers*, 2010; G. ARDIZZI, C. PETRAGLIA, M. PIACENZA, G. TURATI, *L'economia sommersa fra evasione e crimine: una rivisitazione del Currency Demand Approach con una applicazione al contesto italiano*, in *Econpubblica Working Paper n. 155*, Università Bocconi, Milano, 2011. Rientra nei modelli economico-statistici anche il modello composito, conosciuto anche come *Model Approach*, teorizzato da Bruno Frey e Hannelore Weck-Hanneman (per i quali si v. B.S. FREY, H. WECK-HANNEMAN, *The Hidden Economy As An 'Unobserved' Variable*, in *European Economic Review*, 1984, pp. 33-53). Quest'approccio considera l'entità del sommerso come una variabile non direttamente osservabile, soffermandosi sui suoi legami di causa-effetto con una pluralità di determinanti e di indicatori monetari (come l'aumento delle transazioni in contante), della produzione e del mercato del lavoro.

Degli approcci macroeconomici fanno, invece, parte il modello basato sulla discrepanza tra

Gli indubbi punti di forza che contraddistinguono gli approcci di tipo *top down* hanno indotto anche il nostro Paese ad avvalersi di queste metodologie per la misurazione ufficiale del *tax gap* delle imposte dirette e delle imposte indirette.

In questo senso, vengono periodicamente pubblicate, da parte del Governo italiano, le stime dell'ammontare complessivo evaso riferibile ai principali tributi quali l'IRPEF, l'IRES, l'IVA, l'IRAP e l'IMU.

Allo stesso tempo, in attuazione del d.lgs. 24 settembre 2015, n. 60, il Ministero dell'Economia e delle Finanze presenta annualmente, in allegato alla Nota di Aggiornamento del Documento di Economia e Finanza, un *Rapporto sui risultati conseguiti in materia di misure di contrasto all'evasione fiscale*, in cui indica quali sono le strategie adottate per il contrasto all'evasione fiscale ed enumera i risultati raggiunti in relazione agli obiettivi attesi.

Per la redazione di tale *Rapporto* il Governo si avvale della *Relazione sull'economia non osservata e sull'evasione fiscale e contributiva*, predisposta da una apposita Commissione istituita con decreto del MEF.

Si tratta di un passaggio di grande rilevanza, dal momento che è evidente come i risultati relativi alla stima del *tax gap* riportati nella *Relazione* siano il punto di partenza per la programmazione delle politiche di contrasto all'evasione, e rappresentino, non in ultimo, anche la base per giustificare, agli occhi dei contribuenti, politiche fiscali talvolta eccessivamente onerose.

La Relazione sull'economia non osservata e sull'evasione fiscale e con-

la spesa e il reddito; l'approccio dell'*input* lavoro; l'approccio della discrepanza tra dati statistici e dati fiscali. Il primo si basa sulla discrepanza tra le entrate e le spese attese. Il modello valorizza l'utilizzo delle stime del PIL dal lato della domanda, prendendo in considerazione i consumi finali delle famiglie. Tale aggregato macroeconomico, se confrontato con il reddito dichiarato dal lato dell'offerta di beni e servizi, permetterebbe di stimare il sommerso nello scarto risultante da tale confronto.

L'approccio dell'*input* lavoro si basa, invece, sul calcolo della discrepanza tra la forza lavoro risultante dalle stime ufficiali e quella effettiva. Prevede una stima dell'offerta di lavoro (differenziata per attività economiche e dimensioni delle imprese) mediante indagini dirette su imprese e famiglie, quali censimenti e indagini campionarie. Partendo dall'assunto secondo cui il totale della partecipazione della forza lavoro sia costante nel tempo, la riduzione della forza lavoro registrata nell'economia ufficiale rappresenterebbe un valido indicatore del corrispondente incremento di attività lavorativa svolta nell'ambito dell'economia sommersa. Si tratta di un metodo applicato per la prima volta dall'ISTAT e regolarmente utilizzato in Italia; analizzato da I. CIPOLLETTA, *Le illusioni dell'economia sommersa*, in *Economia Italiana*, 2003, pp. 77-90; M.G. REY, *La controversia sull'economia sommersa*, in *Economia Italiana*, 2003, pp. 9-51; ISTAT, *The underground Economy in italian economic accounts*, in *Annali di Statistica*, 1993.

Infine, quello della discrepanza tra dati statistici e dati fiscali è un metodo indiretto di stima basato su un approccio di tipo contabile, che viene sviluppato prendendo in considerazione i dati fiscali e i dati di contabilità nazionale. Tale metodo viene analizzato, tra i tanti, anche da E. GALLI, C. SCAGLIONI, *L'evasione fiscale in Italia tra aspetti teorici ed evidenze empiriche*, in *Riv. dir. trib. internaz.*, 2017, p. 99.

tributiva viene predisposta, per espressa previsione di legge⁴, con l'utilizzo di un metodo indiretto di rilevazione, basato sul confronto tra i dati della contabilità nazionale (che forniscono una descrizione quantitativa di tutte le attività economiche in un determinato periodo di tempo) e quelli acquisiti dall'Anagrafe tributaria e dall'INPS.

La stima prende, quindi, in considerazione il divario tra le basi fiscali e contributive *dichiarate* e quelle *teoriche* desumibili dagli aggregati di contabilità nazionale.

Il *tax gap* complessivo si ottiene dalla somma delle stime effettuate per ciascun tributo e consiste nel minor gettito rispetto a quello che si sarebbe dovuto riscuotere in base alla disciplina fiscale vigente in un regime di perfetto adempimento (in assenza cioè di evasione fiscale).

Nello specifico, il calcolo del *tax gap* relativo alle singole imposte viene effettuato secondo un approccio indiretto.

Si tratta di un risultato differenziale tra il gettito che si sarebbe ottenuto applicando la disciplina fiscale vigente alla base imponibile teorica (BIP), e quello ottenuto applicando la medesima disciplina alla base imponibile dichiarata.

Al fine di costruire la base imponibile teorica, per ciascun tributo viene scelto un aggregato di riferimento desunto dalla contabilità nazionale⁵.

Si pensi ad esempio ai dati contenuti nelle dichiarazioni IRAP adoperati per la stima del *tax gap* riferibile alle imposte sul reddito (IRPEF ed IRES). In questo caso, infatti, viene scelto come aggregato di riferimento della Contabilità nazionale, sul quale costruire la BIP, il risultato lordo di gestione potenziale (RLGp)⁶.

L'RLGp viene ottenuto a partire dal valore aggiunto al costo dei fattori di Contabilità nazionale, che a sua volta rappresenta la misura esauriente⁷ dei redditi generati nel processo di produzione di beni e servizi, depurato della quota ritenuta non imputabile ai soggetti IRES o IRPEF.

⁴ La Legge di contabilità e finanza pubblica (l. 31 dicembre 2009, n. 196) all'art. 10-*bis*.1, comma 4, let. b), prevede espressamente che per la stima dell'ammontare evaso il calcolo debba essere effettuato "con la massima disaggregazione possibile a livello settoriale, territoriale e dimensionale, utilizzando una metodologia di rilevazione, riferita a tutti i principali tributi, anche locali, basata sul confronto tra i dati della contabilità nazionale e quelli acquisiti dall'anagrafe tributaria, con criteri trasparenti, stabili nel tempo, e adeguatamente pubblicizzati".

⁵ Per maggiori approfondimenti sul metodo di stima utilizzato in Italia si rimanda a R. CORRIERE, *Le metodologie di stima dell'evasione fiscale*, in AA.VV., *L'evasione fiscale. Ricerca su natura giuridica e dimensione quantitativa*, a cura di P. Boria, Roma, 2022, p. 143 ss.

⁶ Ciò, tuttavia, induce ad escludere dall'analisi i soggetti IRPEF esentati dal pagamento dell'IRAP, ossia i professionisti privi di autonoma organizzazione e i contribuenti rientranti nei regimi agevolati soggetti a tassazione sostitutiva.

⁷ Comprensiva della componente sommersa.

Viene così confrontato il gettito ottenuto in considerazione della base imponibile dichiarata (BID) e quello che si sarebbe dovuto ottenere alla luce della base imponibile potenziale (BIP).

3. I metodi di stima diretti dell'evasione fiscale. L'approccio di tipo bottom up

I metodi di tipo *bottom up* sono approcci di matrice microeconomica, sostanzialmente basati su indagini svolte presso i contribuenti (famiglie e imprese) e sui risultati delle attività di verifica fiscale.

Tali sistemi di valutazione hanno il pregio di fornire informazioni dettagliate, poiché basati su indagini a campione differenziate per categorie di soggetti⁸.

Le rilevazioni si basano, in particolare, sull'utilizzo di fonti informative di cui l'Amministrazione è in possesso e di micro dati di origine statistica, cioè i dati derivanti da indagini svolte presso i contribuenti e i risultati delle attività di verifica fiscale.

Si tratta di tecniche dirette di analisi del fenomeno evasivo, che – come dimostrato da innumerevoli studi – sono da preferire a quelle indirette; consentono, infatti, di ottenere informazioni molto dettagliate e, conseguentemente, di elaborare stime affidabili.

Nell'ambito di tali approcci, il più famoso ed applicato è quello dell'*Auditing fiscale*, basato sulla discrepanza tra le entrate dichiarate e quelle accertate con le verifiche a campione⁹, ossia sulle informazioni raccolte in sede di

⁸ Sebbene, come evidenziato da I. PALMIERI, *Metodologie utilizzate per quantificare l'evasione fiscale con particolare riferimento ai metodi fondati sugli accertamenti*, Agenzia delle Entrate – Ufficio Studi, 2004, 2, 8, possano non garantire una rappresentazione corretta del dato reale laddove il campione preso in considerazione non sia rappresentativo della popolazione.

⁹ Rientrano nella categoria dei metodi di stima diretti anche il metodo di Franz e le indagini campionarie su bilanci, spese e forza lavoro. Il primo metodo (proposto da Alfred Franz in A. FRANZ, *Estimates Of The Hidden Economy In Austria On The Basis Of Official Statistics*, in *Review of Income and Wealth*, vol. 31 (4), 1985, pp. 325-336) viene utilizzato per stimare l'evasione delle piccole imprese, e si basa sul confronto tra il reddito percepito dai lavoratori non dipendenti (imprenditori) e quello medio dei lavoratori dipendenti impiegati in imprese con caratteristiche simili. La stima si fonda sul presupposto che vi sia un vincolo di coerenza tra ricavi e costi sostenuti dall'impresa, per cui il reddito dei lavoratori non dipendenti (oggetto di verifica) non può mai essere inferiore a quello medio percepito dai lavoratori dipendenti. Si tratta di un approccio attualmente utilizzato anche dall'ISTAT per correggere i valori che risultano falsati da fenomeni di sotto-dichiarazione del fatturato o sovra-dichiarazione dei costi. Per la critica a tale approccio si v. R. ZAZZA, *Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano*, in *Banca d'Italia*, n. 463, dicembre 2002, p. 21, nota 10.

Il metodo delle indagini campionarie è, invece, sviluppato a livello disaggregato su famiglie

accertamento da parte delle autorità fiscali, tramite le quali è possibile effettuare una stima media del reddito evaso.

Questo metodo ha riscosso il suo maggior successo negli Stati Uniti¹⁰.

I dati ottenuti con i *fiscal auditing programs* permisero di acquisire informazioni utili alla conoscenza del fenomeno e migliorare – dal punto di vista qualitativo e quantitativo – il processo di selezione dei contribuenti da sottoporre a controllo¹¹.

Si tratta di una metodologia particolarmente virtuosa dal punto di vista del risultato ottenibile, sebbene finora di difficile attuazione e realizzazione poiché molto onerosa anche (e soprattutto) in termini di costi per lo Stato.

In particolare, affinché possa giungersi ad una stima che rappresenti il più fedelmente possibile la realtà fattuale che si intende analizzare, sarebbe necessario processare una mole di dati talmente elevata e dettagliata da tradursi in un'opera eccessivamente dispendiosa.

La sua riuscita richiederebbe, infatti, l'effettuazione di controlli non circoscritti a campioni casuali di contribuenti o a segmenti parziali della popolazione (come sino ad oggi è stato fatto), ma indagini uniformi e su larga scala, con assegnazione di un considerevole numero di risorse all'espletamento di tali attività.

Alla luce degli ingenti costi che si dovrebbero sostenere, in sede di bilan-

e imprese. Si basa sul confronto tra i dati relativi al reddito e al numero dei lavoratori emergenti dalle dichiarazioni ufficiali e i dati comunicati in sede di intervista. Tale approccio si basa sul presupposto che le informazioni fornite in sede di intervista siano più veritiere di quelle riportate nella dichiarazione dei redditi. Il metodo è stato estensivamente utilizzato in Norvegia (da A.J. ISACHSEN, J. KLOVLAND, S. STROM, *The Hidden Economy in Norway*, in V. Tanzi (ed.), *The Underground Economy in the United States and Abroad*. Lexington, D.C. Heath, 1982, pp. 209-231; A.J. ISACHSEN, S. STROM, *The Size and Growth of the Hidden Economy in Norway*, in *Review of Income and Wealth*, 31 (1), 1985, pp. 21-38) e in Danimarca (da G.V. MOGENSEN, H.K. KVIST, E. KÖRMENDI, S. PEDERSEN, *The Shadow Economy in Denmark 1994: Measurement and Results*, Study n. 3, Copenhagen: The Rockwool Foundation Research Unit, 1995). Per un'analisi dei vantaggi e degli svantaggi derivanti dall'utilizzo di questa metodologia di stima si v. G.V. MOGENSEN, H.K. KVIST, E. KÖRMENDI, S. PEDERSEN, *The Shadow Economy in Denmark 1994: Measurement and Results*, cit.

¹⁰ Dove l'IRS (amministrazione fiscale americana) per molti anni ha adottato il cd. *Taxpayer Compliance Measurement Program* (TCMP), analizzato da J.S. FEINSTEIN, *Approaches for estimating non compliance: examples from federal taxation in the United States*, in *The Economic Journal*, 1999, pp. 360-369; M. MARÈ, *L'evasione in Italia e nei paesi OCSE: evidenze empiriche, determinanti ed effetti economici*, in *Moneta e credito*, 1996, pp. 393-443.

¹¹ Ciò fu possibile grazie all'implementazione di una funzione discriminante (DIF). Tale funzione fu sviluppata analizzando i dati provenienti dal TCMP in relazione alle differenti *audit classes* in cui venne suddivisa la platea dei contribuenti. Questi ultimi, infatti, furono distinti per classi a seconda della fonte e dell'entità del reddito. La DIF permetteva di indirizzare l'azione accertatrice, circoscrivendola a determinate classi di contribuenti, poiché in base alle loro caratteristiche era in grado di calcolare la probabilità che si potesse verificare uno scostamento tra reddito dichiarato e reddito accertato.

cio si preferisce allora destinare le risorse verso controlli mirati, abbandonando altri tipi di approcci¹².

È questo il motivo per cui la metodologia di stima attualmente utilizzata è di tipo *top down*: è più semplice, è standardizzata, e si basa su dati macroeconomici facilmente reperibili.

4. Effettività, attualità e precisione. Gli elementi disattesi nelle stime ufficiali poste a fondamento delle politiche fiscali adottate dal nostro Paese

Come anticipato nelle pagine che precedono, i metodi di tipo *top down* peccano di innumerevoli margini di approssimazione, ed inevitabilmente danno luogo ad errori di rilevanza non trascurabile.

In questo senso, numerosi studi, condotti sia a livello nazionale sia a livello internazionale, hanno dimostrato la non convenienza nell'utilizzo di tali approcci per le misurazioni ufficiali. Il rischio è quello di generare stime lontane dalla realtà fattuale, impedendo la sua effettiva comprensione e l'elaborazione di strategie efficaci di contrasto all'evasione fiscale.

Nonostante gli evidenti punti di forza, l'approccio attualmente adoperato rifletterebbe numerose ed inevitabili imprecisioni, e ciò è vero in considerazione di molteplici fattori.

Innanzitutto il metodo *top down* non si mostra particolarmente indicato per il calcolo del *tax gap* delle imposte dirette, perché, come è stato correttamente osservato, «i dati di contabilità nazionale sui redditi sono spesso prodotti in maniera non del tutto indipendente rispetto alle informazioni sul dichiarato [e ciò li renderebbe] non utilizzabili per la stima della base imponibile teorica»¹³.

In secondo luogo, le definizioni degli indici di forza economica utilizzate ai fini della contabilità nazionale (come, ad esempio, il reddito), non coincidono con quelle adottate ai fini fiscali¹⁴.

Non in ultimo, un significativo punto di debolezza del metodo si ravvisa nell'impossibilità di procedere ad una disaggregazione dei risultati ottenuti in relazione alle differenti articolazioni dell'evasione fiscale. L'approccio macro non coglie, infatti, l'eterogeneità del fenomeno evasivo a livello locale o di gruppi, che invece potrebbe essere meglio analizzata se si ricorresse a tecniche di analisi di tipo *bottom up*, fondate sull'utilizzo di informazioni più dettagliate. In altri termini, non permettendo di cogliere le peculiarità della

¹² I. PALMIERI, *Metodologie utilizzate per quantificare l'evasione fiscale con particolare riferimento ai metodi fondati sugli accertamenti*, cit., p. 22 ss.

¹³ E. DE SIMONE, *Tax gap ed enti locali*, cit., p. 9.

¹⁴ Vedi V. VISCO, *Colpevoli evasioni*, Milano, 2017, p. 62.

specifica realtà alla quale la stima si dovrebbe riferire, il metodo *top down* impedirebbe di mettere a punto politiche di contrasto mirate e circoscritte a determinate fasce della popolazione ovvero a determinate aree geografiche del nostro Paese.

Al contrario, il particolarismo che contraddistingue le informazioni utilizzate nell'ambito degli approcci diretti garantirebbe una migliore aderenza al dato reale.

Ne è stato esempio l'*auditing fiscale* adottato negli Stati Uniti dall'IRS per stimare l'economia sommersa mediante il confronto tra le entrate fiscali dichiarate e quelle accertate con le verifiche a campione.

La letteratura più autorevole evidenzia l'importanza dell'utilizzo di micro-dati di fonte amministrativa, in grado di fornire informazioni analitiche sui contribuenti, per supportare e indirizzare l'azione di contrasto verso categorie specifiche di soggetti, identificando «con precisione chi, dove e con quale entità evade reddito [...] al fine di disegnare politiche di contrasto sempre più efficaci»¹⁵.

Nonostante la maggiore onerosità, i metodi di tipo *bottom up* rappresentano l'unica via che permette di ottenere «stime puntuali dei tassi di evasione per livelli di reddito e categorie di individui/famiglie»¹⁶.

Finanche all'interno della Relazione sull'economia non osservata e sull'evasione fiscale e contributiva allegata alla NADEF si evidenzia, peraltro, l'opportunità di pervenire all'utilizzo di rilevazioni dirette, proprio perché idonee a garantire una maggiore aderenza al dato reale.

La loro adozione è invero l'unica via che consentirebbe di ottenere «stime puntuali dei tassi di evasione per livelli di reddito e categorie di individui/famiglie sulla base di specifiche caratteristiche personali, familiari e territoriali [e al contempo di] produrre risultati da utilizzare come *input* nei modelli di micro simulazione *tax-benefit* per sviluppare analisi distributive e valutazioni di *policy* che considerano l'evasione fiscale»¹⁷.

Ne è riprova la tendenza, che negli ultimi anni ha iniziato a farsi strada, ad includere, nella stessa Relazione, anche i risultati ottenuti dall'applicazione (ancora in via sperimentale) di una metodologia di tipo *bottom up*.

Il tema non è di poco conto, anche perché non si deve dimenticare che la stima dell'evasione è poi a sua volta in grado di dirigere l'azione politica di contrasto al fenomeno e la pianificazione dei relativi strumenti da utilizzare, ed oltretutto è in grado di giustificare, agli occhi dei contribuenti, le politiche fiscali che il Governo intende perseguire.

¹⁵ P. DI CARO, M. MANZO, *Conoscere l'evasione fiscale può contribuire a ridurla*, in *Argomenti – Rivista di economia, cultura e ricerca sociale*, n. 17, 2020, p. 133.

¹⁶ In questi termini, Allegato alla Nota di Aggiornamento del Documento di Economia e Finanza 2020, p. 86.

¹⁷ È quanto si legge nella stessa Relazione allegata alla Nota di Aggiornamento del Documento di Economia e Finanza 2020, p. 86.

Affinché dunque tali azioni possano rilevarsi efficaci è necessario che gli studi siano condotti nel modo più accurato possibile, basandosi su dati attendibili, dettagliati e riferibili – soprattutto – al momento storico nel quale devono essere declinati.

Si tratta di aspetti che, in base a quanto emerge dai documenti ufficiali, sembrerebbero fortemente disattesi.

Le stime sulle quali oggi, nel 2023, si incentrano le politiche fiscali e quelle di contrasto all'evasione, risalgono – addirittura – all'anno 2020.

Ciò vorrebbe dire che la metodologia ufficiale, attualmente applicata, si basa su rilevazioni rese disponibili con grande ritardo rispetto al periodo di *data recording*.

È allora evidente come il dato preso in considerazione sia del tutto inattendibile, in primo luogo perché inattuale e, non in ultimo, perché relativo ad un periodo eccezionale.

L'analisi si riferisce a tre anni fa e ad un momento storico (quello dell'esplosione della pandemia da Covid-19) in cui, a causa del susseguirsi di ripetuti *lock down* e di restrizioni alla circolazione, la stragrande maggioranza delle transazioni per l'acquisto di beni e servizi non poteva che avvenire *online*, con inevitabile tracciabilità di tutti i pagamenti.

La misurazione più recente del *tax gap* (contenuta nella Relazione pubblicata nel 2022) poggia, dunque, le sue fondamenta su dati riferibili ad un momento stra-ordinario, nel quale, in considerazione della inevitabile contrazione delle transazioni in nero, si sarebbe dovuta registrare una considerevole riduzione dell'evasione fiscale; ciononostante, talvolta viene rilevata una riduzione del valore evaso pressoché irrilevante, talaltra addirittura un suo incremento.

In questo senso, viene spontaneo domandarsi come sia possibile che l'ammontare del *tax gap* IRPEF relativo ad imprese e professionisti riferibile all'anno 2020 ammonti (solo) a circa 4,428 miliardi in meno rispetto a quello registrato per l'anno precedente; come pure non si comprende come sia possibile che il *tax gap* IRES relativo al 2019 sia pari a 8,720 miliardi e quello registrato per il 2020 sia addirittura superiore (8,917 miliardi).

Il dato non convince e la sua verosimiglianza vacilla sul piano della ragionevolezza.

5. Big data e Intelligenza Artificiale. Riflessioni conclusive in merito alla possibilità (non troppo remota) di un cambio di approccio nella stima del tax gap

Alla luce delle criticità sinora evidenziate, quali l'eccessiva approssimazione della stima e la sua inattualità, si mostra necessario un cambio di paradigma.

I tempi sembrano maturi per procedere in questa direzione; basti pensare all'avvento dell'era dei *Big data* e dell'Intelligenza Artificiale.

In particolare, con il termine *Big data* si fa riferimento ad un enorme volume di dati eterogenei fra loro, per fonte e per formato, analizzabili in tempo reale.

Le tre caratteristiche che li contraddistinguono sono, infatti, il volume (trattasi di una grande mole di dati), la velocità (i dati sono prodotti ed elaborati molto velocemente, ossia in tempo reale) e l'eterogeneità (in quanto provengono da un elevatissimo numero di fonti differenti).

Quando invece si fa riferimento all'Intelligenza Artificiale si vuole intendere in termini pressoché sommari un «sistema automatico (*machine learning*) in grado di apprendere/imparare dagli esempi e dall'esperienza, al fine di individuare una “funzione matematica predittiva”»¹⁸.

In considerazione delle caratteristiche che contraddistinguono questi dati e della possibilità di ricorrere all'ausilio dell'Intelligenza Artificiale per processarli e rielaborarli, sembra non essere troppo azzardato predire un cambio di paradigma nella scelta delle metodologie di stima da utilizzare. Dettaglio, eterogeneità e precisione sono, infatti, concetti che richiamano le fonti informative poste alla base degli approcci *bottom up*.

In quest'ottica, il ricorso a nuove tecniche di stima basate sull'elaborazione di funzioni e algoritmi che siano in grado di leggere questi dati e di individuare correlazioni di natura probabilistica, tendenze e/o elaborare modelli, potrebbe essere determinante per l'abbandono degli strumenti che fino ad oggi sono stati utilizzati per le stime ufficiali.

Tutto ciò permetterebbe di ovviare ai limiti che caratterizzano i metodi *top down* (indiretti) e, al contempo, consentirebbe l'applicazione efficiente di quello *bottom up* che fino a pochi anni fa probabilmente risultava essere di difficile applicazione, in considerazione di tutta una serie di ostacoli – che sono emersi anche grazie all'esperienza americana – oggi superabili con le nuove tecnologie. Come evidenziato nei §§ 3 e 4 l'applicazione pratica del modello

¹⁸ Questa definizione si rinviene in O. SIGNORILE, *La ricerca di dati su fonti aperte come nuovo strumento delle indagini fiscali*, in AA.VV., *Fisco digitale. Cripto-attività, protezione dei dati, controlli algoritmici*, a cura di G. Ragucci, Torino, 2023, p. 124. La definizione di Intelligenza Artificiale è in verità più complessa, come osserva M. FASOLA, *Le analisi del rischio di evasione tra selezione dei contribuenti da sottoporre a controllo e accertamento “algoritmico”*, in AA.VV. *Fisco digitale. Cripto-attività, protezione dei dati, controlli algoritmici*, cit., nota 6, ove l'A. precisa che «con la locuzione “intelligenza artificiale” ci si riferisce a una estrema varietà di strumenti informatici, che spaziano da sistemi fondati sulla valutazione dell'avveramento o meno di predeterminate condizioni (noti come algoritmi “condizionali” o “deterministici”), i quali possono comunque raggiungere elevati livelli di complessità, a sistemi che, operando in situazioni in cui il numero delle condizioni rilevanti non è quantificabile, le estrapolano autonomamente attraverso processi di “autoapprendimento” (è il caso, ad esempio, degli algoritmi di “*machine learning*”)».

dell'*Auditing fiscale* sebbene, da un lato, abbia consentito di acquisire informazioni utili alla conoscenza del fenomeno evasivo e migliorare – dal punto di vista qualitativo e quantitativo – il processo di selezione dei contribuenti da sottoporre a verifica, dall'altro, ne ha evidenziato gli inevitabili punti di debolezza. La riuscita di questo metodo richiede, infatti, l'effettuazione di controlli non a campione (come fino a quel momento poteva essere fatto), ma uniformi e su larga scala.

Da queste considerazioni emerge come l'approccio di tipo *bottom up* rappresenti una tecnica di stima particolarmente virtuosa dal punto di vista del risultato ottenibile, ma di difficile attuazione e realizzazione, perché molto onerosa in termini sia di tempo sia di costi per lo Stato.

Quest'ultimo, infatti, sarebbe tenuto a mobilitare tutta una serie di risorse e di personale per reperire, in un primo momento (nel modo più uniforme possibile) i dati mediante le verifiche fiscali, e, in un momento successivo, per processarli e rielaborarli.

Tali limiti, ad oggi, potrebbero essere superati proprio grazie alle nuove tecnologie, che oltretutto consentirebbero di ovviare al problema, evidenziato nelle pagine precedenti, dell'attualizzazione delle stime.

La maggiore velocità nel reperimento del dato rappresenta, a ben vedere, un aspetto non trascurabile e di fondamentale importanza affinché la visione d'insieme della realtà socio-economica, cui si dovrebbero riferire le politiche fiscali e quelle di contrasto all'evasione, non risulti distorta.

Si tratta di una logica non lontana da quanto in ultimo previsto con il piano di riforme finanziate dell'Unione europea nell'ambito del PNRR, volte alla digitalizzazione della Pubblica Amministrazione, tra le quali rientra l'adozione di misure volte al «completamento del processo di pseudonimizzazione¹⁹ e analisi dei *big data* nell'intento di rendere più efficace l'analisi del rischio inerente alla selezione dei contribuenti da sottoporre a verifica»²⁰.

¹⁹ Ai sensi dell'art. 4(5) del Regolamento UE 2016/679 (GDPR), per pseudonimizzazione si intende «il trattamento dei dati personali in modo tale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive, a condizione che tali informazioni aggiuntive siano conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile».

²⁰ È ciò che si legge nell'*Allegato riveduto della decisione di esecuzione del Consiglio relativa all'approvazione della valutazione del piano per la ripresa e la resilienza dell'Italia*. Per maggiori approfondimenti sul tema si rinvia a M. FASOLA, *Le analisi del rischio di evasione tra selezione dei contribuenti da sottoporre a controllo e accertamento "algoritmico"*, cit., p. 80 ss. ove l'A. osserva come tali misure rappresentino «la frontiera tecnologicamente più avanzata delle innovazioni ivi descritte. Esse consentiranno infatti, attraverso sofisticati sistemi di intelligenza artificiale (inclusi sistemi di tipo non deterministico come gli algoritmi di "*machine learning*"), di analizzare il ricco patrimonio di dati oggi a disposizione del fisco e di esternare informazioni rilevanti ai fini dei controlli e della *compliances*».

Il problema principale che emergerebbe qualora si facesse ricorso all'utilizzo dei *Big data* come strumento di stima dell'evasione sarebbe quello legato alla tutela della *privacy*²¹.

Il tema dell'interconnessione tra informatica, diritto tributario e tutela del contribuente non è nuovo, ed è oggetto di grande attenzione da parte della più autorevole dottrina²².

In questo senso bisognerebbe lavorare per studiare sistemi di *data mining*, (cioè metodologie e tecniche di estrazione e analisi di informazioni da un grande mole di dati) che siano compatibili con la tutela dei diritti individuali.

Probabilmente sarà questa la sfida del prossimo futuro: da un lato, elaborare funzioni, algoritmi, modelli che siano in grado di ricercare, all'interno di una vasta base di dati, informazioni di natura reddituale, patrimoniale e finanziaria che possano essere validamente utilizzate ai fini della stima, e, dall'altro, condurre quest'opera di interconnessione dei risultati ottenuti senza trascurarne gli effetti sui diritti individuali.

Per cui la raccolta dei dati dovrebbe avvenire, innanzitutto, mettendo il contribuente nella condizione di essere consapevole della loro raccolta e del loro sfruttamento.

A questo fine, per l'utilizzo di alcune tipologie di dati, si potrebbe fare ricorso al rilascio del consenso da parte del contribuente a fronte del riconoscimento, ad esempio, di meccanismi premiali – in ipotesi del suo rilascio – ma non penalizzanti, in caso contrario.

²¹ Si tratta di un aspetto peraltro già analizzato in relazione all'utilizzo di dati derivanti dallo scambio automatico di informazioni in ambito internazionale, per i cui approfondimenti si rimanda a A. CONTRINO, *Banche dati tributarie, scambio di informazioni fra autorità fiscali e "protezione dei dati personali": quali diritti e tutele per il contribuente?*, in *Riv. dir. trib.*, supplemento online, 29 maggio 2019, il quale evidenzia come «in questi ultimi anni è divenuta centrale la questione del contrasto dei fenomeni di elusione ed evasione attuati su scala sovranazionale, che ha portato ad ampliare e rafforzare il ricorso agli strumenti di scambio di informazioni, la cui efficacia è grandemente cresciuta grazie al progresso tecnologico: esso non solo ha reso possibile la formazione di imponenti "archivi" con dati immediatamente accessibili, ma ha anche facilitato l'evoluzione dei processi di trasmissione di dati fra le autorità fiscali dei diversi Stati [...]. La nuova rilevanza della protezione dei dati personali in materia fiscale fa sorgere il problema dei limiti, dandone per scontata la necessitata esistenza, che il diritto alla protezione dei dati personali pone alla acquisizione/utilizzo/conservazione dei dati da parte dell'Amministrazione finanziaria, e ciò nella prospettiva di un corretto bilanciamento tra esigenze di contrasto all'evasione fiscale e tutela del diritto alla protezione dei suddetti dati, tanto a livello interno quanto a livello sovranazionale».

²² Si vd., fra tutti, G. RAGUCCI, *Investigazioni fiscali tra strumenti informatici e mancanza di confini*, in *Boll. trib.*, 2020, p. 333 ss.; ID., *L'analisi del rischio di evasione in base ai dati dell'Archivio dei rapporti con gli intermediari finanziari: prove generali dell'accertamento "algoritmico"?*, in *Riv. dir. trib.*, supporto online, 4 settembre 2019; F. PAPARELLA, *L'ausilio delle tecnologie digitali nella fase di attuazione dei tributi*, in *Riv. dir. trib.*, 2022, I, p. 617 ss.; A. CARINCI, *Fisco e privacy: storia infinita di un apparente ossimoro*, in *Il fisco*, 46, 2019, p. 4407 ss.