

# ERRORS AND UNDERREPORTING OF TAX FORM'S AND INTERVIEW REPORT'S INCOME IN THE SURVEY ON ECONOMIC AND SOCIAL CONDITIONS OF FAMILIES AT MODENA\*

Massimo Baldini, Paolo Bosi, Michele Lalla

*Department of Political Economy, University of Modena and Reggio Emilia  
Via Jacopo Berengario 51 – 41100 Modena*

**Abstract.** The paper presents a review of measurement errors in the reports of earnings or income information, and the analysis of comparisons of individual income in an anonymous interview with the individual income recorded in the tax form produced by the Ministry of Finance. This kind of comparison is uncommon. On the one hand, our research corroborates the empirical evidence of other studies, that is, dependent workers tend to evade less than self-employed. On the other hand, the results showed some differences with respect to the assumption that the survey data on income collected in an anonymous context are more reliable than tax forms data of the Ministry of Finance. In fact, in many categories of interviewed individuals, the surveyed income was less than the tax form income, while other studies found the contrary, using aggregated and not individual data. It could be a specific or local case or it reconfirms the complexity of the cognitive aspects related to collection of income data, i.e., how respondent understand, think about, and form answers to income questions. A strange explanation focused the attention on the media noise about the sector's studies carried out to well calibrate the amount of taxes for the interested taxpayers because it could have influenced the behaviours respondents. Another questionable point concerns the total non-respondents because they could bias the estimates. It is important to know something about them.

JEL: C46, C81, D31, H26, Z13

**Keywords:** fiscal income, income amount reports, measurement errors, response bias, tax evasion, surveyed income, administrative records

## 1. Introduction

Le indagini sul reddito, sul patrimonio, sul risparmio, e sugli investimenti sono complesse e soggette a molteplici fonti di errore (Curtin, Juster, Morgan, 1989; Quintano, Lucev, 1990; Hurd, Juster, Smith, 2003). L'addestramento degli intervistatori è fon-

---

\* The study was developed for the National Research Project «Costruzione di un'indagine sulle famiglie e di un modello di microsimulazione per l'analisi delle politiche sociali e fiscali a livello locale», financed by MIUR (Ministry of Education, University, and Research, year 2001, prot. 2001135524), the Province, the Commune, and the Saving Bank Foundation of Modena. We are grateful to Giuliano Orlandi who coordinates the participation of municipal statistical office of Modena, which operated in some steps of survey process and Giovanni Bigi who carried out the survey sample and linked the fiscal, the census, and the survey databases. The work was jointly carried out by the authors, but the final paper is the result of separate efforts: Sections 3 and 4 were written mainly by Massimo Baldini, Sections 1 and 6 mainly by Paolo Bosi, and Sections 2 and 5 mainly by Michele Lalla.

The corresponding authors are Michele Lalla ([michele.lalla@unimore.it](mailto:michele.lalla@unimore.it)) for statistical questions and Massimo Baldini ([massimo.baldini@unimore.it](mailto:massimo.baldini@unimore.it)) for fiscal and tax evasion questions.

damentale sia per la qualità dei dati raccolti, sia per ottenere la partecipazione delle unità selezionate (Couper, Groves, 1992; Hox, de Leeuw, 2002), ma spesso è insufficiente perché il processo di raccolta dei dati è soggetto a vincoli temporali, logistici, e di risorse umane e finanziarie che generano limiti insormontabili ai miglioramenti del piano di campionamento e di raccolta dei dati: l'errore è un tratto fisiologico di ogni piano concreto. Per esempio, l'esperienza maturata nelle indagini sui bilanci delle famiglie (IBF) condotte dalla Banca d'Italia (2004) ha mostrato che l'attendibilità dei dati è migliore per le famiglie nelle quali il capofamiglia è giovane, ha un elevato titolo di studio, è un lavoratore dipendente.

Gli errori non campionari sono distinti generalmente in tre tipologie (Lessler, Kalsbeek, 1992). (i) Errori nella lista o errori di copertura: sono i più difficili da eliminare e derivano in prevalenza dalla incompletezza e imprecisione della lista. (ii) Errori da mancata risposta: derivano sia dall'impossibilità di procedere alla rilevazione per la non reperibilità o l'assenza al domicilio di alcune unità statistiche selezionate, sia dalla non partecipazione all'indagine delle unità statistiche selezionate e rintracciate –rifiuto totale–, sia dall'assenza di cooperazione su particolari domande del questionario –rifiuto parziale–. (iii) Errori di misurazione: derivano dalla non corrispondenza tra il valore rilevato e il valore «vero», risultante da qualche altra fonte, quando è disponibile, e sono generati da numerosi fattori che alterano il valore da osservare.

Gli errori di misurazione si sovrappongono, in parte, a quelli da mancata risposta perché l'assenza di una risposta o l'errore contenuto in essa potrebbero dipendere proprio da una formulazione ambigua o inadeguata. Una distinzione tipica si basa sulla causa che li ha prodotti. (1) Gli *errori di strumenti* sono riconducibili al questionario per domande mal formulate, ordinate in modo inadeguato, o con batterie di test non tarati bene, e così via. (2) Gli *errori di tecniche* derivano dalla metodologia e dalle modalità usate nella raccolta dei dati, come l'intervista auto-somministrata, il questionario postale, l'intervista telefonica con/senza ausilio di computer, la batteria di test. (3) Gli *errori dell'intervistatore* sono intrinseci alla dinamica dell'intervista; dipendono dalle sue caratteristiche personali (genere, istruzione, età, qualità della voce, abbigliamento) e dalle sue modalità di interazione che influenzano l'intervistato sia nell'incentivare o disincentivare la sua partecipazione, sia nel fornire o non fornire una risposta. (4) Gli *errori dell'intervistato*, come quelli dell'intervistatore, sono intrinseci al processo e connessi alle sue caratteristiche personali (genere, età, istruzione, carattere), alla sua capacità e modalità di comprensione o di ricordare gli eventi accaduti, alla sua idoneità e volontà di fornire risposte veritiere, all'ambiente che si crea durante l'intervista. (5) Gli *errori di «traduzione»* si generano nei passaggi successivi all'intervista che riguardano la codifica, la revisione, la registrazione, e l'elaborazione dei dati; questi sono non meno rilevanti dei precedenti, ma non coinvolgono intervistatori e intervistati, bensì il personale addetto a immagazzinare i dati.

Il processo cognitivo nella dinamica domanda-risposta si sviluppa in più fasi: (1) comprensione della domanda, (2) recupero delle informazioni con la memoria, (3) valutazione della corrispondenza tra le informazioni richieste e le recuperate, (4) comunicazione (Tourangeau, Rips, Resinski, 2000). In queste fasi, molti fattori possono influenzare l'esito finale. La memoria genera diversi tipi di errori: l'errore telescopico per errata collocazione degli eventi (Sudman, Bradburn, 1973), l'errore di richiamo per incapacità a ricordare (*recall loss effect*) o per la numerosità eccessiva di eventi da riportare (*report loading effect*), l'errore di condizionamento per un decremento del numero di eventi riportati rispetto a quello reale nelle interviste ripetute nel tempo. Anche la desi-

derabilità sociale può avere importanza nell'accertamento del reddito e generare risposte normative con esiti sia negativi o sottodichiarazioni, sia positivi o sopradichiarazioni (Tourangeau, Rips, Resinski, 2000). Forme ulteriori di errori non campionari possono emergere in altre fasi del processo di indagine: durante la codifica, la revisione, la registrazione, e l'elaborazione dei dati.

Il trattamento degli errori non campionari richiede assunti sulle caratteristiche della popolazione, sulla natura, e sulla distribuzione degli errori: un'ampia rassegna si trova in Bound, Brown, e Mathiowetz (2001).

In questo lavoro ci si limita a considerare i problemi inerenti all'accertamento del reddito, che presenta frequenti errori dei rispondenti sulle fonti e sull'ammontare del reddito (Moore, Stinson, Welniak, 1997), la peculiarità di essere difficilmente misurabile con precisione per la sua complessità e di non avere un reale riscontro perché il dato dichiarato nella denuncia dei redditi si presume affetto da evasione, erosione, elusione. Tale presunzione, la complessità del lavoro richiesto per ottenere dati coerenti per il confronto, e la scarsa accessibilità dei dati fiscali non hanno incoraggiato, forse, le analisi dei due tipi di dati. Il filone che ne tratta concerne l'accertamento e la stima dell'evasione delle tasse tramite il metodo diretto che consiste nell'uso dei dati rilevati tramite una indagine campionaria, basata su partecipazione volontaria, o di attività di verifiche ufficiali delle autorità sui contribuenti. L'IBF, svolta dalla Banca d'Italia ogni due anni, rappresenta la fonte principale di informazione campionaria sui redditi individuali per l'Italia.

Cannari, Ceriani, e D'Alessio (1995) utilizzano i dati dell'IBF del 1989 e confrontano gli esiti campionari con le corrispondenti statistiche fiscali, supponendo che le rilevazioni dell'IBF siano più veritiere e tendenzialmente più alte di quelle dichiarate al fisco; pertanto, confrontano dati aggregati.

Marenzi (1996) stima il profilo distributivo dell'evasione attraverso un confronto tra i redditi rilevati dall'IBF della Banca d'Italia e quelli dichiarati al fisco, assumendo che i primi siano più veritieri dei secondi e escludendo la presenza di sottodichiarazioni (*underreporting*). I dati dichiarati al fisco consistono nelle distribuzioni di frequenza dei contribuenti per fasce di reddito complessivo, elaborate dall'Anagrafe tributaria. L'evasione è misurata dalla differenza tra il reddito rilevato nell'indagine e quello dichiarato al fisco. L'analisi è condotta per quattro tipologie di contribuenti: dipendenti, pensionati, liberi professionisti e imprenditori. Nel caso del lavoro dipendente, l'evasione si concentra di fatto solo nei primi due decili, dovuta probabilmente a posizioni lavorative non dichiarate. Più elevata è la propensione a evadere tra i pensionati, a causa soprattutto di posizioni lavorative autonome ignote al fisco. Molto elevata è la propensione a evadere tra gli autonomi: la quantità di reddito evaso è crescente in valore assoluto al crescere del reddito, mentre in termini percentuali è sostanzialmente costante nei vari decili della distribuzione del reddito degli autonomi, attorno al 40%. Nel caso dei liberi professionisti, infine, la propensione ad evadere è elevata soprattutto nei decili più bassi: l'evasione vale il 40% del reddito per il terzo decile, solo il 4% per il decimo (v. anche Bernasconi, Marenzi, 1997).

Bordignon e Zanardi (1997) hanno analizzato l'evasione fiscale utilizzando i risultati individuali dei controlli sui redditi soggetti a IRPEF, eseguiti negli anni 1987 e 1989 dall'amministrazione sugli autonomi. L'evasione sembra dipendere dal grado di progressività dell'imposta personale, dalle variabili territoriali, dal tipo di attività economica, e dalla discrepanza tra le tasse pagate e i servizi pubblici offerti.

Fiorio e d'Amauri (2005) analizzano l'evasione utilizzando il campione del 2000 con 22268 soggetti, relativo all'IBF della Banca d'Italia, e un campione di 250000 contribuenti costruito dalla SeCIT (Servizio Consultivo ed Ispettivo Tributario) del Ministero dell'Economia e delle Finanze e da essa elaborato. La supposta evasione è analizzata per i redditi da lavoro disaggregati in decili: il reddito da occupazione è evaso a livelli bassi e diventa trascurabile dopo la mediana; l'evasione degli autonomi è elevata nei primi decili e decresce, senza azzerarsi, al crescere dei decili.

Cannari e D'Alessio (2007) riportano le opinioni degli italiani sull'evasione fiscale. In base a un indicatore di propensione all'evasione fiscale rilevano una maggiore propensione tra i lavoratori indipendenti rispetto ai dipendenti; per questi ultimi, è maggiore tra gli operai, è minore tra gli impiegati e i dirigenti e direttivi. L'atteggiamento favorevole nei confronti dell'evasione fiscale diminuisce al crescere del titolo di studio e dell'età; inoltre, i meccanismi di riprovazione sociale siano inefficaci a ostacolare il problema.

In Italia rimane poco indagato il confronto a livello individuale tra i dati del reddito dichiarati al fisco e i dati rilevati con una indagine sul campo, che costituisce la parte rilevante e originale del presente lavoro. A livello internazionale, Jäntti (2004) ha analizzato dati simili e ha riscontrato un comportamento diverso da quello riportato nella maggior parte dei casi: l'ammontare dei redditi rilevati tramite interviste tendono a essere più bassi di quelli contenuti negli archivi amministrativi, come in questo lavoro.

In questo saggio si illustrano i risultati del confronto tra il reddito dichiarato nella denuncia dei redditi e il reddito rilevato nell'indagine sulle condizioni economiche e sociali delle famiglie a Modena (ICESMO1), svolta nel 2002. In particolare, sulla base delle informazioni socio-demografiche relative agli intervistati si sono individuati i fattori che determinano le differenze tra i due dati, ricavando indicazioni presuntive, ma di qualche interesse, sulle sopradichiarazioni o sottodichiarazioni.

Il testo è organizzato come segue. Nel paragrafo 2 si illustrano le caratteristiche salienti della rilevazione del reddito e dell'ICESMO1. Nel paragrafo 3 si descrivono alcuni aspetti della raccolta dei dati fiscali a Modena inerenti sia all'affidabilità dei dati sia al campione risultante dal processo. Nel paragrafo 4 si riportano sia la descrizione dei principali dati campionari sia i risultati salienti del confronto tra reddito dichiarato al fisco e reddito rilevato durante le interviste. Nel paragrafo 5 si analizzano le differenze esistenti nel segmento dei percettori a reddito fisso: operai, impiegati, insegnanti. Nel paragrafo 6 seguono le conclusioni.

## 2. Literature review

Il problema dell'affidabilità dei dati si è posto dagli inizi dell'uso di indagini campionarie per l'accertamento del reddito e del patrimonio, intorno agli anni quaranta (Bancroft, 1940). Le domande su tale oggetto presentano difficoltà enormi di tipo cognitivo, etico, e psicologico (Juster, Smith, 1997), che generano numerosi errori di misura da parte del rispondente: (1) *ignoranza* sui dati reali, (2) *riluttanza* a fornire informazioni su dati sensibili, (3) *timore* per i controlli delle istituzioni e le possibili fughe di notizie. Tali fattori introducono una distorsione sulla misura e si presume che sia tipicamente sottovalutata o «sotto-dichiarata» (*underreporting*) o falsa (*misreporting*). Si vedano, per esempio, Mandell e Lundsten (1978), Greenberg, Moffitt, e Friedmann (1981), Pedace e Bates (2000).

Per convalidare i dati rilevati nel processo di indagine, si usano indicatori multipli interni e contestuali della stessa grandezza o, come accade piú spesso, si usano misure esterne ottenute in altri contesti di accertamento. Il ricorso a fonti esterne deve assumere la trasferibilità del modello di misura esterno all'interno dell'indagine sicché è preferibile, quando è possibile, ricorrere solo a convalide interne.

Si distinguono due procedimenti. Uno è a livello *macro*, nel quale si confrontano le stime del campione con i corrispondenti valori della popolazione, stimati o veri, provenienti dalle fonti esterne. L'altro è a livello *micro*, nel quale si confrontano individuo per individuo il valore rilevato rispetto a quello della fonte esterna.

Il confronto a livello *micro* si può realizzare seguendo tre modalità: (1) *controllo diretto*, in cui le unità sono selezionate dalla base di dati amministrativa conservando le informazioni oggetto di studio in essa presenti e, quindi, sono intervistate rilevando, tra gli altri, gli stessi dati per verificare la corrispondenza tra quanto dichiarato alla amministrazione e quanto all'intervistato sicché il confronto è immediato; (2) *controllo indiretto*, in cui le unità sono prima intervistate, poi, si individuano le informazioni che hanno un corrispondente negli archivi amministrativi e si recuperano, quindi, in questi ultimi i dati utili per confermare il comportamento dell'intervistato; (3) *controllo completo*, in cui le unità sono selezionate dalla base di dati amministrativa conservando tutte le informazioni disponibili siano o no di interesse e, quindi, sono intervistate (Bound, Brown, Mathiowetz, 2001).

Il controllo diretto costruisce l'indagine sulla base di dati amministrativa e, dunque, è piú sensibile a fenomeni di sottovalutazione di eventi discreti (come il numero di incidenti o di ricoveri), mentre è meno vincolante per fenomeni di sopravvalutazione perché la fonte esterna può essere incompleta rispetto a tutti i possibili eventi che interessano il rispondente. Il controllo indiretto, invece, realizza un campione svincolato dalla base di dati amministrativa e sembra, quindi, piú adeguato per accertare fenomeni di sopravvalutazione, anche se i risultati dipendono da ciò che si sta misurando, da quali definizioni si utilizzano, e da come si attua la misura. Il controllo completo fornisce tutte le informazioni esistenti nella base di dati amministrativa e, quindi, consente una modellazione piú consistente con il fenomeno indagato, sia per la sottovalutazione e sia per la sopravvalutazione, anche se spesso i caratteri non sono rilevati tutti con la stessa accuratezza (Martini, 1990).

### 2.1. *Measurement errors in income surveys carried out in USA*

Le verifiche empiriche sull'attendibilità e precisione delle dichiarazioni inerenti al reddito, raccolte tramite indagine campionaria con interviste a uno o piú membri della famiglia, sono state eseguite prevalentemente sui salari o stipendi e hanno riguardato la remunerazione annuale, la remunerazione abituale (rispetto a un dato periodo), l'ultima remunerazione, e la paga oraria. I dati ottenuti dalle interviste sono stati confrontati con i dati effettivi o veri, ricorrendo agli archivi del personale del luogo di lavoro o amministrativi e assumendo che questi ultimi fossero privi di errore; ma, in generale, anche questi potrebbero essere affetti da carenze che generano le discrepanze osservate. I risultati piú stabili e comuni alle indagini condotte hanno riguardato la precisione e l'attendibilità della rilevazione che sono piú elevate per la remunerazione annuale rispetto a quella oraria o settimanale o mensile. La media delle stime della remunerazione annuale è sembrata soggetta a un errore relativamente piccolo e le medie degli errori hanno mostrato segni opposti, indicando sia una sottostima, sia una soprastima. Non so-

no emerse forti correlazioni con le variabili demografiche e relative al capitale umano. Le altre remunerazioni (mensili, settimanali, orarie) sono soggette a complessi meccanismi di recupero delle informazioni che conducono a errori più elevati e dipendono, presumibilmente, dal tipo di lavoro e di lavoratore; per esempio, chi è remunerato con una paga mensile conoscerà meglio il dato mensile rispetto a quello orario, mentre chi è remunerato con paga oraria conoscerà con più approssimazione la remunerazione mensile e, pertanto, le risposte possono essere affette da errori notevoli.

Uno dei primi studi fu condotto dall'ESRI (*Employment Stabilization Research Institute*) del Minnesota agli inizi degli anni quaranta (1940–42) su un campione di disoccupati di Saint Paul: il salario dichiarato e quello effettivo mostrarono un coefficiente di correlazione pari a 0,90 per gli uomini e a 0,93 per le donne, ma non fu osservata alcuna distorsione sopra o sotto il valore vero (Keating, Paterson, Stone, 1950). Uno studio degli anni sessanta rivelò un coefficiente di correlazione pari a 0,95 e differenze tra uomini e donne, titoli di studio alti e bassi, lavoratori anziani e giovani, salari e orari raggruppati in alti e bassi; si osservò anche una tendenziale sopravvalutazione (*overreporting*) della remunerazione media (Borus, 1966). Un'indagine sui redditi bassi a Fort Wayne, Indiana, fu condotta con due metodi di accertamento del reddito annuale e i risultati furono confrontati con i dati degli archivi amministrativi riscontrando un errore medio dell'1,9% (Borus, 1970). La remunerazione annuale (mensile, oraria) rilevata in una indagine del 1977 non ha mostrato differenze significative in media; mentre differenze significative si sono osservate sia nel salario orario, settimanale, e mensile e sia per genere (Carstensen, Woltman, 1979). In generale, da tali studi si evince che i dati riportati dai lavoratori intervistati sono ben correlati con i dati forniti da altre fonti e le medie presentano valori con una moderata sopravvalutazione o sottovalutazione.

Negli anni ottanta furono condotte indagini simili, ma non risultarono differenze rilevanti in Yellow e Sider (1983), mentre l'esperimento di Gary mostrò che le persone trattate (beneficiari) tendevano a sottostimare la loro remunerazione più dei non trattati e l'esperimento di Seattle-Denver ottenne una leggera soprastima (Greenberg, Moffitt, Friedmann, 1981; Greenberg, Halsey, 1983). In uno studio di validazione longitudinale sulle dinamiche del reddito si rilevarono sopravvalutazioni e sottovalutazioni e gli errori si rivelarono più elevati per il reddito dell'anno precedente (Duncan, Hill, 1985).

Negli anni novanta il numero di lavori sul tema è cresciuto. Lo studio di Bound e Krueger (1991) ha mostrato che l'errore di misura nella rilevazione della remunerazione annuale ha una distribuzione simmetrica rispetto alla media, che è sostanzialmente zero, per uomini e donne, ma con una varianza doppia degli uomini rispetto a quella delle donne, sebbene queste ultime mostrassero una tendenza a sottodichiarare i loro redditi. Lo studio di Rodgers, Brown, e Duncan (1993) ha rilevato che le remunerazioni dichiarate e vere erano correlate e i valori diminuivano passando da quelle annuali, a quelle relative alle precedenti due settimane, a quelle relative a un periodo abituale; gli errori di misura erano correlati negativamente ai logaritmi delle remunerazioni reali e una proporzione elevata di errori seguiva la distribuzione normale, ma una proporzione minore proveniva da una distribuzione con una varianza maggiore e influenzava notevolmente le stime dei coefficienti che esprimevano le relazioni tra le variabili. In un altro studio, gli errori sono risultati negativamente correlati ai valori veri del reddito e le distorsioni sono state moderatamente più grandi per i cambiamenti della remunerazione e le remunerazioni orarie sono state riportate con minore affidabilità delle remunerazioni annuali (Bound, Brown, Duncan, Rodgers, 1994). Gli errori osservati nelle indagini trasversali sono risultati più elevati di quelli rilevati nelle indagini longitudinali, la relazio-

ne negativa tra gli errori di misura e l'ammontare della remunerazione annuale era guidata dalla sopradichiarazione dei percettori di redditi bassi, la mediana degli errori di risposta non era in relazione con la remunerazione annuale (Bollinger, 1998). In generale, il reddito annuale è riportato con un errore inferiore rispetto al reddito settimanale o mensile e alle ore lavorate.

## 2.2. *Quality of data: tax form's and interview reports of income*

La qualità dei dati, in Italia, è stata esaminata in alcune indagini che riguardano proprio le variabili concorrenti alla formazione del reddito, del patrimonio, e della ricchezza e i caratteri coinvolti nella loro rilevazione perché possono determinarle. L'affidabilità delle informazioni raccolte tramite l'IBF della Banca d'Italia è stata analizzata tramite la componente longitudinale (panel) delle famiglie. Le variabili socio-demografiche, che sono invariabili nel tempo (genere, anno e luogo di nascita) o a variazioni limitate (titolo di studio), mostrano un errore abbastanza stabile dell'ordine del 2%, eccetto il genere che si è ridotto sempre più nel tempo. Il titolo di studio presenta una variazione dell'ordine del 10%, ma le diminuzioni, che presumibilmente denotano un errore, variano tra il 2% e il 5%; mentre il tipo di titolo di studio presenta, specie per il diploma, una maggiore variabilità. Tra le variabili quantitative, la superficie dell'abitazione di residenza presenta differenze che stanno entro 5 metri quadri nel 44,5% dei casi, mentre l'anno di costruzione degli immobili mostra scarti inferiori a 5 anni nel 73% dei casi (Biancotti, D'Alessio, Neri, 2004; D'Alessio, Faiella, 2002). Le componenti del reddito sono soggette a variabilità temporale e presentano una affidabilità, calcolata con il metodo di Heise (1969), che varia con il tipo di componente. I redditi da pensione e le retribuzioni da lavoro dipendente sono rilevati con elevata affidabilità e mostrano indici intorno a 0,95. Il reddito da lavoro autonomo, il reddito da capitale e le integrazioni non monetarie del reddito presentano una minore affidabilità con indici che vanno a 0,74 a 0,35. Anche tra le componenti della ricchezza si osserva una elevata eterogeneità: vi sono elementi con una elevata affidabilità (immobili e titoli di stato) e elementi con scarsa affidabilità (depositi e altri titoli, passività finanziarie). Lo studio di D'Aurizio, Faiella, Iezzi, e Neri (2006) analizza il fenomeno della sottodichiarazione (*under-reporting*) degli intervistati in merito alle attività finanziarie e propone una procedura di correzione delle stime.

Per quanto concerne il reddito non sono frequenti le analisi che utilizzano i dati provenienti dall'archivio fiscale abbinati al singolo intervistato. Gli studi che accertano le differenze esistenti tra i dati raccolti tramite indagini dirette e i dati amministrativi di tipo fiscale sono importanti sia per capire i meccanismi sottostanti le risposte nelle indagini sul campo, sia per appurare la portata dei dati amministrativi fiscali. Se i dati dell'archivio fiscale sono affidabili almeno quanto quelli delle interviste, allora i comuni e le regioni potrebbero analizzarli direttamente per avere un quadro della situazione reddituale delle famiglie oppure utilizzarli per realizzare un campione più preciso, se l'oggetto di interesse è più esteso di quello che fornisce la base fiscale. L'accesso ai dati fiscali, seppure in forme limitate, può essere utile per sviluppare una migliore conoscenza delle strutture relazionali sottostanti ai dati stessi e con altre variabili nelle indagini che si conducono sul campo.

### 2.3. Survey on social and economic conditions of family at Modena

L'indagine sulle condizioni economiche e sociali delle famiglie della provincia di Modena (ICESMO1) è stata condotta nel 2002 e si basa su un piano di campionamento a due stadi, con stratificazione del territorio secondo i distretti sociosanitari, nel quale i comuni sono le unità di primo stadio e le famiglie sono le unità di secondo stadio. Si sono distinti i comuni in due categorie, come in altre indagini di questo tipo (Banca d'Italia, 2004). I comuni Auto-Rappresentativi (AR) sono i «capoluoghi» del distretto. I comuni Non Auto-Rappresentativi (NAR) sono tutti gli altri. Il numero di comuni NAR, per distretto, è proporzionato al numero di famiglie residenti: una unità NAR ogni 10000 famiglie. La dimensione campionaria provinciale è stata ripartita proporzionalmente, prima, tra i distretti secondo la numerosità delle famiglie (dimensione campionaria distrettuale); poi, tra i comuni inclusi nel campione del distretto, sempre in base alla numerosità delle famiglie (dimensione campionaria comunale). Nella provincia di Modena, escluso il comune di Modena, il numero di famiglie intervistate nel campione è pari a 789; ma tale campione non è considerato in questo ambito.

Nel comune di Modena si è estratto un campione con una precisione delle stime circa uguale a quella della provincia senza il capoluogo. Si è utilizzato un piano di campionamento stratificato secondo l'ampiezza delle famiglie (1, 2, 3, 4 o più componenti), l'età del capofamiglia (fino a 34 anni, 35-49 anni, 50-64 anni, 65-74 anni, 75 anni e più), e il genere. Con l'allocazione proporzionale si è determinato il numero di unità per strato da selezionare. Il numero di famiglie campione intervistate è stato 589 (Baldini, Bosi, Silvestri, 2004). I dati analizzati nel seguito si riferiscono esclusivamente al campione del comune di Modena.

### 3. Profiles of tax form's and interview reports of income at Modena

L'Ufficio statistico del comune di Modena, che aveva partecipato alla costituzione del campione per l'ICESMO1, è venuto in possesso nel 2006 dell'archivio delle dichiarazioni dei redditi del 2002 e indicato nel seguito con il sintagma «archivio fiscale». Le dichiarazioni sono accessibili con il Sistema di Interscambio dell'Anagrafe Tributaria degli Enti Locali (SIATEL): un sistema di collegamento telematico, a copertura nazionale, realizzato dal Ministero dell'Economia e delle Finanze per consentire l'interscambio delle informazioni anagrafiche e tributarie con gli Enti Locali, i quali stipulano una convenzione per la visione e l'utilizzo dei dati in esso raccolti nell'esercizio delle proprie funzioni. L'anno di riferimento è assai vicino a quello del censimento e, pertanto, si sono collegati i dati delle unità campionarie sia a quelle dell'archivio fiscale, sia a quelle del censimento.

L'abbinamento delle informazioni delle tre basi di dati ha fornito l'opportunità di eseguire controlli interessanti sul comportamento delle diverse tipologie di unità statistiche osservate nel processo di campionamento: le unità partecipanti all'indagine, le unità irreperibili, le unità che hanno rifiutato di partecipare, le unità di riserva che non sono state contattate perché quelle designate hanno partecipato. In questo lavoro, come si è detto, ci si limita a verificare l'attendibilità e la precisione delle dichiarazioni degli intervistati nell'ICESMO1, utilizzando i dati fiscali corrispondenti. L'uso dell'archivio fiscale può presentare qualche limite inerente sia alla equivalenza delle voci del questionario e della denuncia dei redditi, sia alla natura amministrativi dei dati, sia alla per-

fetta corrispondenza del periodo temporale tra la rilevazione e l'anno fiscale. Tali limiti riguardano, tuttavia, anche altri tipi di indagini simili che l'hanno utilizzato per confrontarne gli esiti con quelli dell'IBF della Banca d'Italia.

L'equivalenza dei due periodi temporali di riferimenti è il limite più critico perché l'ICESMO1 ha rilevato informazioni relative ai redditi *netti* percepiti nei dodici mesi precedenti l'intervista per avere un lasso di tempo il più piccolo possibile tra l'intervista e il periodo di riferimento della produzione del reddito sicché non ha fissato come periodo di riferimento l'anno solare, che è quello dell'archivio fiscale. La scelta è sembrata idonea a fornire anche i dati più recenti, nonostante richieda all'intervistato una stima del reddito senza poter contare su documenti esistenti che lo rilevano in termini completi; infatti, le interviste sono iniziate nel mese di febbraio del 2002 e sono terminate nel mese di ottobre dello stesso anno. I redditi riguardano, quindi, un periodo variabile a cavallo tra il 2001 e il 2002.

L'abbinamento ha dato esiti molto soddisfacenti, anche se non sono mancati i problemi. Per esempio, 84 soggetti dell'archivio fiscale su 1388 avevano il genere mancante, ossia circa il 6%. Il reddito fiscale mancava in 357 soggetti, il 25,7% del totale; ma in 254 (18,3%) casi mancavano entrambi i redditi perché soggetti non tenuti a presentare denuncia. Gli errori eventuali si hanno, pertanto, nella diagonale secondaria della [Tabella 1](#) dove ogni cella riporta la presenza di uno solo dei due redditi e, in totale, ammontano al 10,4%. I casi abbinati e confrontabili sono 987 e tale numero costituisce l'ampiezza campionaria delle analisi successive.

**Table 1** – Number of cases (*n*), mean ( $\bar{y}$ ), standard deviation (*SD*), and percentages (%) on total of tax form's report (*FR*) and interview report (*IR*) of income (*Y*) by their corresponding binary variables: not missing/missing

	IR of Y not missing			IR of Y missing			Total <i>n</i>
	<i>n</i>	$\bar{y}$	<i>SD</i>	<i>n</i>	$\bar{y}$	<i>SD</i>	
FR of Y not missing							
FR of income	987	23867.2	23674.7	44	4089,1	4656,5	1031
IR of income		20595.7	19381.2		0	0	
% on total	71.1			3.2			
FR of Y missing							
FR of income	103	0	0	254	—	—	357
IR of income		14080.6	15377.9		—	—	
% on total	7.4			18.3			100
Total	1090			298			1388

I modelli risultanti sono stati tre: 730, Unico, e 770. Il modello 730 contiene voci molto analitiche sia per il reddito percepito dal soggetto (dominicali, agrari, dei fabbricati, da lavoro dipendente e assimilati, altri, complessivo, imponibile), sia per le deduzioni (crediti d'imposta sui dividendi, oneri deducibili, deduzione per l'abitazione principale), sia per le detrazioni (dovute per coniuge a carico, per familiari a carico, per lavoro dipendente o assimilati, per redditi da pensione, per gli oneri di cui alla sezione I e III e IV del quadro E), sia per le imposte sul reddito delle persone fisiche o IRPEF (imposta lorda, credito d'imposta per il riacquisto della prima casa e per nuove assunzioni, totali di detrazioni e crediti d'imposta, imposta netta, credito di imposta sui dividendi e sulle imposte pagate all'estero e per canoni non percepiti, ritenute, differenza, eccedenza dell'IRPEF risultante dalla precedente dichiarazione, acconti versati, residui del credito d'imposta per il riacquisto della prima casa e per nuove assunzioni da utilizzare in

compensazione, totale spese sanitarie da riportare nelle successive dichiarazioni dei redditi).

Il modello 730 riporta anche le voci che riguardano le imposte per gli enti locali (reddito imponibile, addizionale regionale (AR), AR risultante dalla certificazione, eccedenza dell'AR risultante dalla precedente dichiarazione, addizionale comunale (AC), AC risultante dalla certificazione, eccedenza dell'AC risultante dalla precedente dichiarazione). L'ultima parte della dichiarazione riporta i versamenti e i rimborsi a cura del sostituto d'imposta (importi IRPEF da versare e da rimborsare, importi AR IRPEF da versare e da rimborsare, importi AC IRPEF da versare e da rimborsare, prima e seconda o unica rata di acconto IRPEF per l'anno precedente, acconto del 20% sui redditi soggetti a tassazione separata, richiesta di rateizzazione dei versamenti di saldo e degli eventuali acconti).

Il modello Unico presenta due varianti (con e senza imposta sul valore aggiunto — IVA), una maggiore sinteticità rispetto al 730, e due sezioni. La prima mostra il totale del reddito imponibile, l'imposta netta, il reddito imponibile per le imposte locali, l'AR, l'AC, e l'imposta regionale sulle attività produttive (IRAP). La seconda riporta il reddito secondo le fonti: terreni, fabbricati, lavoro dipendente e assimilati, lavoro autonomo, partecipazione in società di persone e assimilate, capitali, altre fonti, soggetti a imposta sostitutiva. Unico con IVA fornisce, in aggiunta alle precedenti, anche informazioni sull'IVA: volume di affari, totale imposta sulle operazioni imponibili, totale acquisti e importazioni, totale imposta su acquisti e importazioni imponibili, IVA ammessa in detrazione, e imposta a credito.

Le rendite da terreni e da fabbricati sono denominati «quadro A» e «quadro B», rispettivamente, e la descrizione di ogni singola rendita è identica sia in 730 e sia Unico. Il quadro A fornisce il valore della rendita (domenicale e agricola), il titolo di godimento, i giorni di possesso, e l'ammontare del possesso espresso in termini percentuali. Il quadro B fornisce il valore della rendita (rivalutata del 5%), l'utilizzo del fabbricato (abitazione principale, «tenuta disposizione» per la quale si applica l'aumento di un terzo, locata con o in assenza di regime legale di determinazione del canone, di pertinenza dell'abitazione principale anche se iscritta in catasto con autonoma rendita, di società semplici o di società a esse equiparate che producono redditi da fabbricati, altri casi particolari), l'indicazione dei giorni di possesso, e la percentuale di tale possesso.

In Unico manca qualsiasi riferimento alle deduzioni; infatti, se si sommano, in ogni singola dichiarazione Unico, i redditi divisi per tipologie, il loro ammontare è spesso superiore all'importo del reddito imponibile fornito, ossia mancano quelli i passaggi intermedi che portano dal reddito complessivo al reddito imponibile. Ciò vale anche per l'imposta lorda: dato presente in 730, ma del tutto omissivo in Unico. Nei due modelli ci sono le voci sull'addizionale comunale all'IRPEF, anche se nel comune di Modena l'imposta è stata introdotta soltanto nel 2002. Il numero di paganti l'imposta nel campione è basso sicché si può presumere che il loro reddito sia maturato in comuni dove già nel 2001 era entrata in vigore.

Le dichiarazioni congiunte sono maggiormente presenti nei 730, mentre i le famiglie mononucleari (*singles*) preferiscono presentare la propria dichiarazione in modalità Unico. I figli che lavorano, ma fanno ancora parte del nucleo familiare di origine, prediligono il modello 730. La proprietà immobiliare è diversamente ripartita tra 730 e Unico. Il primo è caratterizzato da un patrimonio immobiliare costituito prevalentemente dall'abitazione principale e dalla pertinenza, ovvero garage e immobili destinati a impieghi simili. Al contrario, Unico è caratterizzato da individui titolari di patrimoni im-

mobiliari più estesi e differenziati. La proprietà fondiaria è presente in casi alquanto rari in ambedue i modelli.

Il modello 770 include due versioni, 770 e 770s, e riguarda quegli individui che hanno esclusivamente redditi da lavoro dipendente e non hanno alcuna proprietà immobiliare; quindi, non presentano alcuna dichiarazione, eccetto quella effettuata dal proprio sostituto di imposta. Si sono incontrate difficoltà a acquisire alcuni dati: se usufruivano di determinate deduzioni come quelle sull'affitto per l'abitazione principale, se percepivano le detrazioni e per il coniuge e per i familiari a carico e in quale percentuale, l'ammontare di spese come quelle mediche o per l'istruzione o a favore di ONLUS che sono soggette al regime di detrazione al 19%. Nonostante, il contenuto della cella Redditi Percepiti (Allegato VII) fornisca il codice fiscale del datore di lavoro, con cui è stato possibile visionare anche il prospetto dell'azienda, sostituto d'imposta, non si è riusciti a aumentare la quantità di informazione sugli individui sottoposti a tale regime dichiarativo. Le voci delle singole fonti di reddito e delle imposte non seguono, inoltre, la stessa logica e la stessa nomenclatura nei diversi modelli, eccetto le parti dei dati anagrafici, burocratici, e dei redditi da terreni e fabbricati. Ogni singola dichiarazione risulta un *unicum*, quindi, per la strutturata delle informazioni in ciascun modello.

The sub-sample used in the subsequent analyses is summarized in Table 2, where the number of cases, the means, and the standard deviations are reported for the total income reported in the tax form's corresponding to the year 2002, for the sake of brevity denoted by «Fiscal Income» (FI), and the total income reported in the interview of the survey appropriately adjusted and denoted, for the sake of brevity, «Surveyed Income» (SI). They are disaggregated by the type of tax's form and gender. The means of the two incomes proved to be different between the modalities of tax's form and gender. Moreover, the difference between the fiscal and surveyed incomes proved to significantly greater than zero for women and showed only borderline p-value for men in the tax form's type «Unico», while it yielded a borderline p-value for women in the tax form's type 730. However, these evidences could be spurious and a more complex model is necessary to ascertain the impact of different factors.

**Table 2** – Number of cases (*n*), mean ( $\bar{y}$ ), standard deviation (SD) of fiscal income (FI) and surveyed income (SI) by type of form and by gender

Type of Form	Type of Income	Men			Women			Total <i>n</i>
		<i>n</i>	$\bar{y}$	SD	<i>n</i>	$\bar{y}$	SD	
M-730	FI	241	26413.6	14691.1	262	17577.4	9356.6	503
	SI		25955.7	18510.5		15971.8	10053.1	
M-Unico	FI	145	39070.2	45381.3	132	20311.8	19035.3	277
	SI		30532.3	32636.3		13979.3	15277.7	
M-770	FI	87	16758.8	7565.5	120	11950.7	6913.8	207
	SI		15514.3	6750.0		11273.5	7990.0	
Total		473			514			987

Per fornire qualche elemento descrittivo ulteriore del campione, la Tabella 3 riporta il numero di casi, il valore medio, e la deviazione standard, del reddito fiscale e rilevato per titolo di studio e per genere. Si notano subito sia la notevole differenza tra uomini e donne dell'ammontare dei redditi percepiti, sia il marcato scarto tra i valori medi dei redditi fiscali e dei redditi rilevati. I primi risultano, in genere, più alti dei secondi. Si osserva il contrario solo per gli uomini con nessun titolo di studio o con il di-

ploma universitario; ma questa ultima modalità conta solo tre casi. Tali differenze vanno esaminate, tuttavia, in un contesto più ampio perché la distribuzione del reddito non è normale, nelle celle si osserva una elevata variabilità, e vi sono altre variabili che possono influenzare la differenza, come l'età del percettore. L'analisi della covarianza, con l'età come covariata, mostra solo una significativa differenza di genere.

I limiti delle entità dei redditi riportati nelle dichiarazioni fiscali sono noti e riguardano, essenzialmente, la loro affidabilità o veridicità perché si presume che siano soggetti a evasione e elusione; pertanto, non possono rappresentare il valore reale di confronto con quello dichiarato, come nelle altre indagini. Sia per la loro presunta falsità e sia per la loro inaccessibilità, i redditi delle dichiarazioni fiscali non sono stati analizzati diffusamente e non sono mai stati confrontati a livello individuale con quelli dichiarati nelle interviste, in genere, ritenuti di qualità superiore: validi, attendibili, e precisi. Tale supposizione non sembra comprovata da evidenze empiriche sufficientemente solide; perciò, il confronto è importante perché può fornire indicazioni sulla loro qualità, sulla differenza di comportamento nelle dichiarazioni alle interviste e al fisco.

**Table 3** – Number of cases ( $n$ ), mean ( $\bar{y}$ ), and standard deviation ( $SD$ ) of fiscal income ( $FI$ ) and surveyed income ( $SI$ ) by educational level and by gender

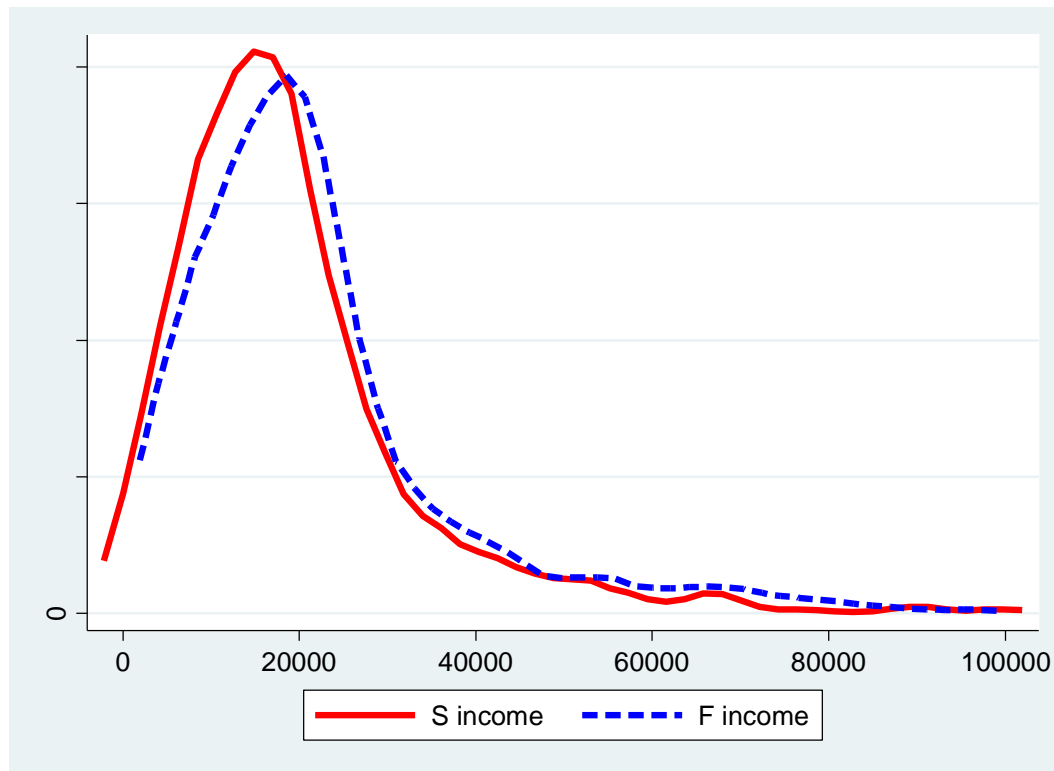
Educational level		Women			Men		
		$n$	$\bar{y}$	DS	$n$	$\bar{y}$	DS
No title	FI	11	10503.2	5840.3	10	13980.0	5327.7
	SI		11087.1	4524.0		13363.4	6613.1
Elementary school	FI	117	11691.2	6651.6	72	17662.6	11394.2
	SI		10613.4	9670.7		15129.1	5087.9
«Avviamento»	FI	10	16646.9	10084.6	14	20676.4	15287.8
	SI		8673.9	4055.9		17604.7	11038.8
Italian middle school	FI	80	15296.5	12574.7	107	22855.1	17195.9
	SI		12414.6	7748.3		20290.4	21722.1
Vocational school	FI	14	16335.9	7972.4	25	24131.1	13597.6
	SI		12496.9	7323.9		21931.4	7525.3
School-leaving certificate	FI	165	19740.9	13788.4	148	27098.4	21926.7
	SI		17238.3	9450.3		24619.5	18921.4
Undergraduate	FI	12	17844.4	5375.9	3	19087.3	6217.2
	SI		13876.9	9211.4		21387.1	3465.1
Graduate	FI	105	22418.1	14471.4	94	50800.2	47115.2
	SI		19031.6	17752.3		45476.8	31914.7
Total	FI	514	16725.8	12307.2	473	27975.2	27761.4
	SI		14239.2	11228.0		25033.2	22644.8

#### 4. Analisi dei dati e relazioni tra reddito fiscale e reddito rilevato

Nella Figura 1 sono riportate le distribuzioni dei due redditi, FI e SI, per il campione ristretto ai 987 casi abbinati, dove i valori superiori a 100000€ sono stati eliminati per ottenere una maggiore leggibilità e l'istogramma è stato perequato con la stima di Epanechnikov della densità (*Epanechnikov kernel function*). Si può osservare il picco più elevato e a valori più bassi del reddito nella curva (continua) relativa a SI, che deno-

ta la già notata sottodichiarazione ottenuta in fase di rilevazione, rispetto alla denuncia effettuata al fisco, FI.

**Figure 1** – Epanechnikov kernel density estimate of the fiscal (F) and surveyed (S) income: Year 2002



Nella [Tabella 4](#) sono riportati i valori medi, la deviazione standard, e la numerosità campionaria delle variabili FI e SI per condizione professionale, al fine di migliorare la descrizione del campione disponibile. Dalla [Tabella 4](#) si evince che le donne sottodichiarano nell'intervista più degli uomini:  $-14.9\%$  contro il  $-10.5\%$ . In tre categorie si osserva che le dichiarazioni al fisco sono inferiori, in media, alle dichiarazioni ottenute tramite intervista, ma in tutte e tre le categorie il numero di casi è molto basso. Si osserva, in media, una sopradichiarazione per gli insegnanti uomini con  $+83.5\%$ , per i soci gestori di società con un  $+39.3\%$ , e per pensionati non lavoro con un valore sotto l' $1\%$ . In tutte le altre categorie si osserva, invece, che le dichiarazioni al fisco sono superiori, in media, alle dichiarazioni ottenute tramite intervista. Le differenze rilevanti si hanno tra gli imprenditori uomini che presentano la sottodichiarazione più elevata  $-85.7\%$ , ma il numero di casi è molto basso. Seguono le lavoratrici autonome con  $-64.1\%$ , le parsubordinate con  $-53.7\%$ , le disoccupate con  $-47.8\%$  pur con redditi esigui e numerosità bassa, le pensionate non da lavoro con  $-42.4\%$ . Sembra che tali categorie si fidino meno degli intervistatori e tendano a essere più reticenti, ma la tendenza a sottodichiarare i propri redditi nelle interviste sembra assai diffusa. La differenza media complessiva tra SI e FI è di  $-12.2\%$ . Si evince, quindi, che in alcune categorie le medie sono marcatamente diverse dalle altre e non c'è omogeneità nella variabilità tra le diverse condizioni professionali. Per controllare quanto tali differenze siano statisticamente rilevanti occor-

re tenere conto degli altri fattori, che possono influenzare il comportamento alla risposta in fase di intervista.

**Table 4** – Mean of fiscal income and surveyed income by job position and gender

Job position	W-FI	W-SI	W-n	M-FI	M-SI	M-n
Worker	13370.8	12638.4	51	19937.8	17627.5	77
White-collar worker	18778.5	17015.3	131	27016.6	25884.7	81
Teacher	19916.2	18351.6	39	17256.1	31668.6	7
Executive manager	27521.4	24502.3	12	40434.5	40189.2	36
Manager, high official	52841.5	45465.3	6	68811.8	60247.4	25
Non standard worker	14687.1	6797.5	14	25442.3	16910.2	9
Professional person	29198.0	23982.4	15	50368.1	41243.8	38
Entrepreneur	6692.0	5093.8	1	70303.4	10077.2	3
Self-employed/ artisan	18079.6	6498.4	10	24376.6	14151.6	29
Owner/ co-operator in f.b.	12120.7	11394.3	3	37238.8	33387.7	2
Partner/ managing partner	14648.5	21043.0	8	33283.9	46368.0	10
Unemployed persons	4872.5	2544.0	7			
Retirement pension	16657.1	14485.6	151	21079.8	18971.9	152
Other type of pension	10950.5	6312.3	29	14806.6	14815.5	3
Housewife	12602.2	8707.8	35			
Student	3463.0	464.8	2	4426.0	2353.1	1
<b>Total</b>	<b>16725.8</b>	<b>14239.2</b>	<b>514</b>	<b>27975.2</b>	<b>25033.2</b>	<b>473</b>

La condizione professionale richiede il raggruppamento di alcune categorie, specie quelle con scarsa numerosità, per poterla inserire nei modelli, anche se ciò comporta una perdita di informazione. Nel séguito, la categoria imprenditori comprende gli imprenditori individuali, i titolari o coadiuvanti di imprese familiari, e i soci o gestori di società. La categoria “Altro” include pensionati non da lavoro, casalinghe, disoccupati.

Per descrivere in modo sintetico le differenze del reddito nel campione (finale) dei 987 casi, si è utilizzato un modello di regressione nel quale la variabile dipendente è il reddito complessivo, in un modello è FI e in un altro è SI, e le variabili indipendenti sono le principali disponibili: età, genere, relazione di parentela con il capofamiglia, stato civile, titolo di studio, condizione professionale, settore di attività, tipo di godimento dell’abitazione principale. L’età è stata divisa per 10 al fine di ridurre il suo campo di variazione e introdotta nella forma polinomiale di secondo grado, età e età al quadrato, per tenere conto di una tendenziale non linearità che si osserva nel reddito in funzione dell’età: cresce al crescere dell’età e raggiunto un massimo comincia a decrescere. Le variabili indipendenti sono tutte qualitative, eccetto l’età; sono state introdotte, quindi, come variabili di comodo o indicatrici di categoria (*dummy*) e i coefficienti di regressione indicano di quanto varia il reddito di quella categoria rispetto all’individuo tipo di riferimento: donna, capofamiglia, coniugato, con licenza media, impiegato nel settore dell’industria, e con casa di proprietà. Il modello di regressione presenta eteroschedasticità che è attenuata con la trasformazione del reddito in logaritmi, ma non eliminata per la presenza di alcuni valori anomali (*outliers*), sicché i parametri del modello sono stati stimati con l’opzione robusta di [Stata \(2005\)](#) che produce una stima robusta della varianza e nessuna variazione nella stima dei coefficienti. Per facilitare la leggibilità delle stime si riporta, perciò, quello con le variabili dipendenti originali. La significatività dei

coefficienti rimane immutata anche senza l'opzione robusta e i coefficienti di determinazione sono  $R^2 = 0,339$  per FI e  $R^2 = 0,400$  per SI.

Si può osservare che i due coefficienti dell'età sono statisticamente diversi da zero e con il segno negativo per l'età al quadrato, che indica un andamento parabolico con la concavità rivolta verso il basso, come ci si aspettava. Le donne guadagnano, in media, circa 5000€ in meno degli uomini, a parità di età e di tutti gli altri caratteri inclusi nel modello. Lo stato civile sembra non influenzare il reddito. La relazione di parentela con il capofamiglia sembra influenzare il reddito solo nel campione, ma ciò riflette le differenze di genere combinate con la loro tendenza a dichiarare di meno nell'intervista. Quasi tutti i titoli di studio mostrano differenze significative eccetto la qualifica professionale, che è assimilabile a quello dell'individuo di riferimento che è la licenza media, e il diploma universitario. La condizione professionale mostra differenze significative, in media, per le diverse posizioni con qualche differenza tra i dati rilevati e i dati fiscali. Gli insegnanti, i parasubordinati (nei dati fiscali), i liberi professionisti (nel campione), gli autonomi (nei dati fiscali), e gli imprenditori o soci gestori o proprietari sembrano percepire redditi assimilabili agli impiegati (individuo di riferimento). Nella pubblica amministrazione e nella scuola si guadagna meno degli altri settori. Gli affittuari e i detentori di altre forme di godimento dell'abitazione percepiscono redditi inferiori rispetto ai proprietari e a parità delle altre condizioni.

Il modello descrive gli effetti (marginali) dei fattori, come in una analisi della covarianza in cui l'età e l'età al quadrato sono covariate e i vari gruppi di indicatori rappresentano i fattori; infatti, mancano le interazioni che potrebbero essere anche all'origine dell'eterogeneità osservata. I risultati mostrano qualche dato poco credibile: gli autonomi guadagnano molto meno della media e gli imprenditori come la media. Ciò sembra corroborare le opinioni degli italiani sull'evasione fiscale: la propensione all'evasione sembra più elevata tra i lavoratori indipendenti che tra quelli dipendenti. Tra questi ultimi, la propensione a evadere risulta minore tra gli operai e risulta maggiore tra i dirigenti e i direttivi, il contrario di quanto osservano [Cannari e D'Alessio \(2007\)](#). Dalla [Tabella 3](#) si evince che vi è anche una differenza di genere per titolo di studio: le donne sottodichiarano sempre, ma l'entità diminuisce al crescere del titolo di studio, in linea con [Cannari e D'Alessio \(2007\)](#) e una differenza media di  $-14,9\%$  su FI. Le donne senza titolo di studio e gli uomini con il diploma universitario dichiarano nell'intervista di più di quanto dichiarino al fisco; mentre gli uomini con altri titoli sottodichiarano in modo circa simile e meno delle donne.

**Table 5 – Coefficients, standard errors (SE), and p-values ( $p>t$ )**

Independent Variables	Dependent Variable: FI			Dependent Variable: SI		
	Coefficient	SE	p>t	Coefficient	SE	p>t
Age/10	8398,8	2511,3	0,00	9357,6	2436,9	0,00
(Age/10)^2	-400,7	242,2	0,10	-642,8	181,0	0,00
Man	5507,8	1499,0	0,00	5010,3	1067,3	0,00
No title	-10013,0	2110,7	0,00	-5031,8	1750,1	0,00
Elementary school	-6327,6	1447,3	0,00	-4274,5	1261,9	0,00
Vocational school	910,7	2033,4	0,65	1013,6	1734,8	0,56
School-leaving certificate	3448,0	1771,8	0,05	4168,4	1851,6	0,02
Undergraduate	3230,6	2338,1	0,17	2218,5	3371,6	0,51
Graduate	12895,1	2956,9	0,00	11288,7	3031,1	0,00
Partner	-2404,3	1754,7	0,17	-3840,7	1231,7	0,00
Children	141,6	2600,0	0,96	-2576,7	1500,1	0,09
Other cohabitant	-4997,7	2850,1	0,08	-3373,1	1989,3	0,09
Unmarried	-1621,5	1965,7	0,41	-2756,0	1530,2	0,07
Separate/ Divorced p.	-2431,1	2642,6	0,36	-959,2	2329,8	0,68
Widower	1039,0	2317,5	0,65	1735,4	1597,2	0,28
Worker	-3027,3	1343,3	0,02	-4774,6	1738,0	0,01
Teacher	498,9	3621,3	0,89	4293,9	5231,0	0,41
Executive manager	6028,4	2334,9	0,01	6817,6	2572,7	0,01
Manager/ high official	30949,8	8227,6	0,00	26722,0	7649,0	0,00
Non-standard worker	-2916,5	3295,3	0,38	-8239,6	3120,1	0,01
Professional person	15685,4	7737,1	0,04	11008,3	7714,5	0,15
Self-employed	-3597,0	3819,5	0,35	-14008,9	2372,4	0,00
Entrepreneur/ M.P./ Owner	-0,3	5356,1	1,00	4683,1	8754,3	0,59
Retired person	-15825,8	3430,9	0,00	-14237,7	2889,6	0,00
Housewife & others	-17319,4	2956,8	0,00	-18795,6	2570,7	0,00
Commerce	-4169,0	2054,2	0,04	-6535,8	2676,8	0,02
Bank & insurance	8306,1	7538,7	0,27	-474,5	5866,0	0,94
School	-13186,9	3656,5	0,00	-14807,9	4788,5	0,00
Health	-5706,1	2812,9	0,04	-6004,6	3452,8	0,08
Public Administration	-11344,9	1815,1	0,00	-10172,2	2315,2	0,00
Services	-7195,9	4960,9	0,15	-9371,4	4617,6	0,04
Other	-5118,3	1894,9	0,01	-1342,6	2948,3	0,65
Rent	-3485,6	1145,6	0,00	-1612,1	1170,2	0,17
Free use	2248,4	2020,2	0,27	-1228,2	1760,3	0,49
Usufruct & other	-7077,0	1566,4	0,00	-4557,8	2220,6	0,04
Constant	-4086,9	8148,4	0,62	-3216,6	8283,9	0,70

## 5. La differenza tra reddito complessivo fiscale e reddito complessivo rilevato

Per individuare i fattori che determinano le differenze tra l'ammontare di SI e l'ammontare FI si è costruita la variabile variazione percentuale tra i due redditi,  $y = 100 \cdot (SI - FI) / FI$ . Si è stimato un modello di regressione con variabile dipendente la  $y$  così definita, e variabili indipendenti tutte quelle incluse nella [Tabella 5](#). Tramite un procedimento di eliminazione progressivo (*backward*) delle variabili non significative si è ottenuto il modello esposto in [Tabella 6](#), dove risultano significative l'età al quadrato con coefficiente negativo e la condizione di imprenditore; tuttavia, il coefficiente di determinazione è molto basso. Se si stimano due modelli separati, per i valori positivi e negativi della  $y$ , allora si ottengono sia coefficienti di determinazione più alti, sia modelli con più variabili esplicative, dove l'età mostra sempre un andamento tendenzialmente non lineare che genera valori negativi per le età avanzate, forse perché gli anziani si fidano meno dei giovani e/o tendono più facilmente a dimenticare. Tra le altre variabili influenti si osservano prevalentemente il genere, alcune variabili indicatrici relative al titolo di studio, alla relazione con il capofamiglia, alla condizione professionale (non standard workers, autonomi, liberi professionisti, altri), e al settore. Le caratteristiche del modello migliorano se si utilizzano i logaritmi.

**Table 6** – Coefficients, standard errors (SE), and p-values ( $p > t$ )

Independent Variables	DV: $\ln(SI/FI)$			DV: $\ln[100 \cdot (SI-FI)/FI]$		
	Coefficient	SE	p>t	Coefficient	SE	p>t
Age/10	0,342	0,088	0,00	0,239	0,089	0,01
(Age/10)^2	-0,030	0,008	0,00	-0,023	0,008	0,00
Man	0,081	0,047	0,09	0,089	0,047	0,06
No title	0,376	0,174	0,03	0,435	0,177	0,01
Elementary school				0,102	0,058	0,08
School-leaving certificate	0,171	0,060	0,00	0,165	0,059	0,01
Children				-0,159	0,069	0,02
Non-standard worker	-0,765	0,180	0,00	-0,730	0,173	0,00
Self-employed	-0,526	0,174	0,00	-0,497	0,173	0,00
Housewife & others	-0,389	0,121	0,00	-0,393	0,120	0,00
Bank & insurance	-0,159	0,096	0,10			
Constant	-1,043	0,236	0,00	3,856	0,244	0,00

L'effetto età agisce secondo un ramo di parabola sicché mostra non linearità e si comprende meglio dai dati esposti nella [Tabella 7](#), dove si può vedere che i giovani tendono a dichiarare molto meno nelle interviste rispetto a quanto dichiarano al fisco: 47.4%. Nelle fasce centrali 40-70 anni c'è sostanzialmente una differenza dell'ordine del 10%. I soggetti sopra i 70 dichiarano meno nelle interviste rispetto al fisco.

**Table 7** – Mean, standard deviation (SD), and number of cases (n) by age classes

Age	≤19	20-29	30-39	40-49	50-59	60-69	70-79	≥80
n	4	72	208	191	175	170	114	53
FI: $\bar{y}$	5234	15409	24181	24471	30932	22442	21048	20660
FI: DS	2716	7130	24779	16423	29378	20948	28483	26854
SI: $\bar{y}$	2755	12377	22055	22044	26861	19027	17085	14057
SI: DS	4233	8044	21702	15207	24619	17746	18226	13964
SI/FI	0,526	0,803	0,912	0,901	0,868	0,848	0,812	0,680
(S-F)/F	-0,474	-0,197	-0,088	-0,099	-0,132	-0,152	-0,188	-0,320

Si è esaminato, infine, anche il divario tra i due dati di SI e FI per decili costruiti su FI; ovvero, si è analizzato l'andamento del divario tra la rilevazione dell'indagine e la dichiarazione al fisco per classi con percentuali di individui pari al 10% a partire dalla classe più povera (per il fisco) a quella più ricca (Tabella 8). La variabile rilevante è  $y = 100 \cdot (SI - FI) / FI$ , che mostra un andamento decrescente a partire da valori positivi nei primi decili fino a valori negativi a partire dal terzo decile, con l'eccezione del quinto che mostra ancora un valore positivo che indica un valore maggiore rilevato nelle interviste rispetto a quello denunciato al Fisco.

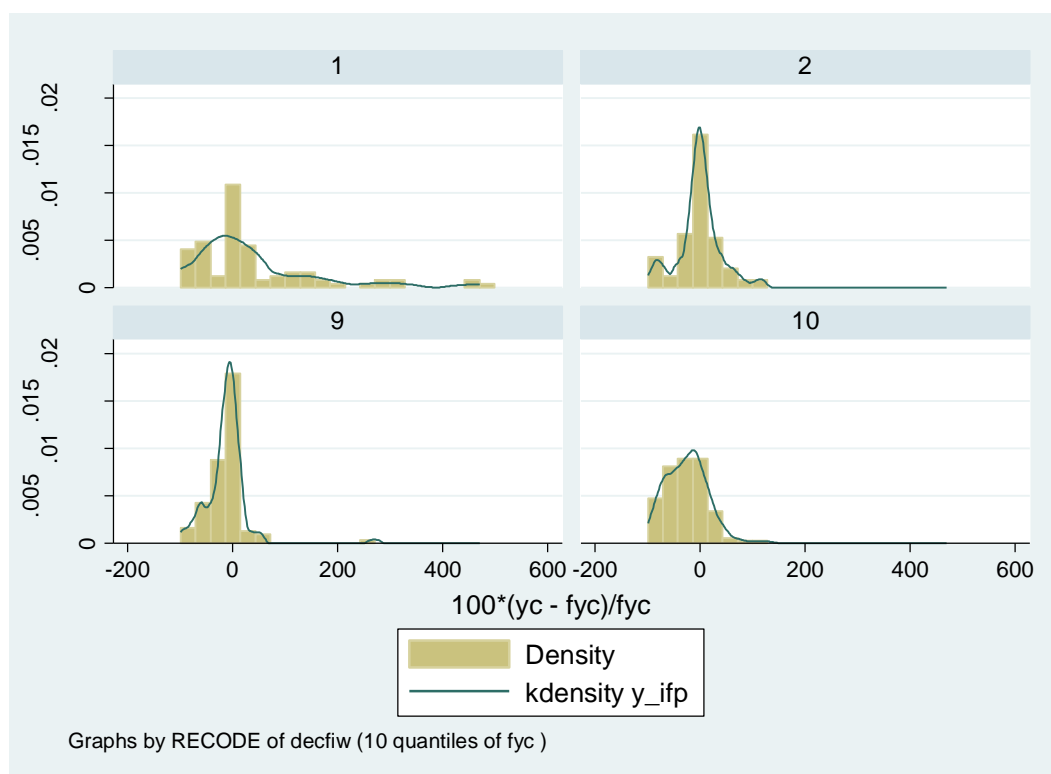
**Table 8** – Minimum, maximum, and mean of fiscal and survey income, mean of percentage variations ( $\bar{y}_\%$ ), and percentage variations of means [ $V\%(\bar{y})$ ]

Decile FI	Min FI	Max FI	Mean FI	n	Min SI	Max SI	Mean SI	$\bar{y}_\%$	$V\%(\bar{y})$
1	138	6127	3691	95	92	64980	6895	490,4	86,8
2	6134	10115	8185	90	352	156040	9599	21,3	17,3
3	10122	12766	11450	89	231	90602	10997	-2,7	-4,0
4	12781	15277	14046	84	1982	25738	13126	-6,5	-6,6
5	15292	17690	16568	92	506	181363	17349	4,6	4,7
6	17848	20259	18984	97	172	100145	17751	-6,6	-6,5
7	20328	22911	21485	103	7092	43709	19747	-8,2	-8,1
8	22966	27110	24883	103	1761	40673	21818	-12,3	-12,3
9	27130	38356	31511	108	1702	112583	29775	-5,1	-5,5
10	38370	304574	67639	126	1585	159116	44278	-29,3	-34,5
<b>Total</b>	138	304574	21807	987	92	181363	19115	44,8	-12,3

I soggetti nei primi due decili, che sono poveri per il fisco, hanno dichiarato nelle interviste un valore superiore, in media, a quello dichiarato al fisco; mentre nei decili superiori, che sono soggetti ricchi per il fisco, hanno dichiarato nelle interviste un reddito inferiore. I più poveri dichiarano, insomma, di più nelle interviste e i più ricchi dichiarano meno: nel primo decile si ha una media del 490% in più, mentre nell'ultimo decile si ha una media del 29% in meno. La spiegazione del fenomeno non è immediata e la scarsa numerosità del campione limita l'elaborazione dei dati. In prima approssimazione, si può presumere che nelle fasce basse vi siano più facilmente casi particolari e valori anomali; infatti, nel primo decile, il valore di  $y$  nel campione va dal minimo di 92 al massimo di +64980, e nel secondo decile va da 352 a 156040. Può anche essere che

nel primo decile vi sia discreta elusione o erosione perché la frequenza di piccoli introiti non sempre denunciati al fisco potrebbe essere alta o, viceversa, per la scarsa entità della somma potrebbero essere dimenticati in fase di rilevazione o potrebbe essere più diffuso il lavoro in nero. La distribuzione della  $y$  riportata nella [Figura 2](#) per il primo decile mostra una estensione molto ampia a destra che comporta un aumento della media, ma la numerosità dei casi non sembra rilevante. Nel secondo decile la distribuzione della  $y$  si concentra di più intorno allo zero, ma vi è ancora una estensione maggiore a destra. Non si può escludere l'ipotesi che nei decili più bassi di reddito vi sia tra i soggetti scarsa dimestichezza con degli strumenti di indagine e con lo stesso modulo dichiarativo. Negli ultimi due decili si osserva un fenomeno tendenzialmente apposto a quello dei primi due, ma il campo di variazione dei valori negativi rimane pressoché invariato, mentre si riduce marcatamente il campo di variazione dei valori positivi della  $y$ , generando una variazione media negativa e, quindi, una denuncia maggiore al fisco, rispetto a quella dell'intervista.

**Figure 2** – Epanechnikov kernel density estimate of the percentage variations between survey income (SI) and fiscal income (FI) by the two tail's deciles



## 6. Considerazioni conclusive

I risultati, pur con tutte le cautele del caso, sono in linea con quanto trovato da altri autori: l'evasione è relativamente bassa per i lavoratori dipendenti, mentre è alta per i lavoratori indipendenti. In tutti gli studi si assume che i dati rilevati nelle indagini siano più veritieri delle dichiarazioni fiscali, a causa dell'anonimato garantito agli intervistati. Si deve osservare, tuttavia, che ciò non garantisce, di per sé, né la disponibilità a

dichiarare il vero, né la precisione della risposta. Le risposte alle domande sul reddito, senza il controllo dei documenti, possono contenere sia errori di memoria perché si lasciano più frequentemente redditi a componente variabile, sia l'effetto di un atteggiamento tendenziale di sottodichiarazione e di diffidenza. Le rilevazioni dell'indagine ICESMO1 risultano inferiori, in media e in generale, alle dichiarazioni riportate al fisco. L'esito si contrappone a altri risultati, ottenuti su dati aggregati dal confronto delle statistiche dell'indagine della banca d'Italia e le corrispondenti statistiche della base di dati fiscale. Ciò può dipendere sia dalle caratteristiche territoriali, sia dalla maggiore garanzia che un intervistato può imputare alla Banca d'Italia per il prestigio dell'istituzione. Non sfugge a alcuno, tuttavia, che la complessità della determinazione del reddito non permette a un intervistato di essere preciso, senza controllare i documenti, ma in molti casi i documenti non sono stati e non possono essere controllati in fase di intervista. Ci si riserva in futuro di approfondire tale aspetto.

Una spiegazione eccentrica della rilevazione di un reddito inferiore a quello dichiarato al fisco presume che il rumore mediatico sugli studi di settore abbia influenzato i percettori, i quali, temendo che il fine dell'indagine ICESMO fosse il confronto con i dati "ufficiali" in possesso degli enti pubblici, abbiano scientemente abbassato il livello del reddito percepito; ossia, la sottodichiarazione deriva dalla consapevolezza degli individui di essere costantemente monitorati. D'altronde, i cittadini sono sommersi da una miriade di sondaggi che cercano di spiegare le loro scelte e orientamenti politici e sociali.

Un altro elemento di fondamentale importanza è costituito dalle mancate partecipazioni e risposte parziali. Chi sono i non rispondenti? Si assume che siano equamente distribuiti tra le varie categorie e tra gli evasori e non evasori, ma si può ragionevolmente dubitare di tale assunto e, dunque, molti risultati sono distorti. In futuro, ci si propone di valutare anche questo punto in modo indiretto utilizzando i dati amministrativi in modo più esteso e completo. Per verificare chi sono i non rispondenti si può condurre una indagine per accertare le loro caratteristiche rispetto ai dati amministrativi che, seppure dissimulati, potranno dare una indicazione utile sul comportamento dei rispondenti e sulle eventuali possibili distorsioni cui sono soggetti i dati rilevati, ma non sarà facile farli partecipare, visto che hanno già rifiutato di farlo nell'indagine corrente.

## Riferimenti bibliografici

- Baldini M., Bosi P. e Silvestri P. (a cura di) 2004. *La ricchezza dell'equità. Distribuzione del reddito e condizioni di vita in un'area a elevato benessere*, il Mulino, Bologna.
- Banca d'Italia 2004. *I bilanci delle famiglie italiane nell'anno 2002*, a cura di D'Alessio G., Faiella I., Supplementi al bollettino statistico (nuova serie), anno XIV, n. 12, Banca d'Italia, Roma.
- Bancroft G. 1940. Consistency of Information from Records and Interview, *Journal of the American Statistical Association*, 35, pp. 377–381.
- Bernasconi M., Marenzi A. 1997. Gli effetti redistributivi dell'evasione fiscale in Italia, in *Ricerche quantitative per la politica economica 1997*, Banca d'Italia, Roma, 1999, pp. 493–589.

- Biancotti C., D'Alessio G. e Neri A. 2004. Errori di misura nell'indagine sui bilanci delle famiglie italiane, Temi di discussione, Numero 520, Banca d'Italia, Roma, pp. 1-58.
- Bollinger C.R. 1998. *Measurements Error in the Current Population Survey: a nonparametric look*, Journal of Labor Economics, 16 (3), pp. 576-594.
- Bordignon M. e Zanardi A. 1997. Tax evasion in Italy, *Giornale degli economisti e Annali di Economia*, 56(3-4), pp. 169-210.
- Borus M. 1966. *Response error in survey reports of earnings information*, Journal of the American Statistical Association, 61, pp. 729-738.
- Borus M. 1970. *Response error and question technique in surveys of earnings information*, Journal of the American Statistical Association, 65, pp. 566-575.
- Bound J. e Krueger A. 1991. *The Extent of Measurements Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?*, Journal of Labor Economics, 9 (1), pp. 1-24.
- Bound J., Brown C., Duncan G.J. e Rodgers W.L. 1994. *Evidence on the Validity of Cross-Sectional Data and Longitudinal Labor Market Data*, Journal of Labor Economics, 12 (3), pp. 345-368.
- Bound J., Brown C. e Mathiowetz N. 2001. *Measurement Error in Survey Data*, in Heckman J.J. e Leamer E. (eds.), *Handbook of Econometrics*, vol. 5, Elsevier, Amsterdam, pp. 3705-3843.
- Cannari L., Ceriani V. e D'Alessio G. 1995. Il recupero degli imponibili soggetti a tassazione, in *Ricerche quantitative per la politica economica 1995*, Numero speciale dei Contributi all'analisi economica del Servizio Studi, vol. 2, Banca d'Italia, Roma, pp. 493-589.
- Cannari L. e D'Alessio G. 2007. Le opinioni degli italiani sull'evasione fiscale, Temi di discussione, Numero 618, Banca d'Italia, Roma, pp. 1-51.
- Carstensen L. e Woltman H. 1979. *Comparing Earnings Data from the CPS and Employers Records*, Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C., pp. 168-173.
- Coder J. e Scoon-Rogers L. 1995. *Evaluating the Quality of Income Data Collected in the Annual Supplement to the March CPS and the SIPP*, SIPP Working Paper Series No. 215, U.S. Bureau of the Census, Washington, DC.
- Couper M.P. e Groves R.M. 1992. *The Role of Interviewer in Survey Participation*, Survey Methodology, 18, pp. 263-278.
- Curtin R.T., Juster F.T. e Morgan J.N. 1989. Survey Estimates of Wealth: An assessment of Quality, in Lipsey R.E. e Tice H.S. (eds.), *The Measurement of Saving, Investment, and Wealth*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 52, The University of Chicago Press, Chicago.
- D'Alessio G. e Faiella I. 2002. *Non-response Behaviour in the Bank of Italy's Survey of Household Income and Wealth*, Temi di discussione, Numero 462, Banca d'Italia, Roma, pp. 1-41.
- D'Aurizio L., Faiella I., Iezzi S. e Neri A. 2002. *L'under-reporting della ricchezza finanziaria nell'indagine sui bilanci delle famiglie*, Temi di discussione, Numero 610, Banca d'Italia, Roma, pp. 1-56.
- Duncan G.J. e Hill D.H. 1985. *An Investigation of the Extent and Consequences of Measurements Error in Labor-economic Survey Data*, Journal of Labor Economics, 3 (4), pp. 508-532.

- Fiorio C.V. e d'Amauri F. 2005. Workers' tax evasion in Italy, Working Paper no. 104, Econpubblica: Center for Research on the Public Sector, Università Commerciale Luigi Bocconi, Milano.
- Greenberg D., Moffitt R. e Friedmann J. 1981. *Underreporting and Experimental Effects on Work Effort: Evidence from the Gary Income Maintenance Experiment*, *The Review of Economics and Statistics*, 63, pp. 581-589.
- Greenberg D. e Aalsey H. 1983. *Systematic Misreporting and Effects of Income Maintenance Experiments on Work Effort: Evidence from the Seattle-Denver Experiment*, *Journal of Labor Economics*, 1 (4), pp. 380-407.
- Heise D. 1969. Separating Reliability and Stability in Test-Retest Correlation, *American Sociological Review*, 34(1), pp. 93-101.
- Hox J. e de Leeuw E. 2002. *The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison*, in Groves R.M., Dillman D.A., Eltinge J.L., Little R.J.A., *Survey Nonresponse*, Wiley & Sons, New York, pp. 103-120.
- Hurd M., Juster F.T. e Smith J.P. 2003. *Enhancing the Quality of Data on Income*, *Journal of Human Resources*; 38 (3), pp. 758-772.
- Jäntti M. 2004. *The effect of measurement errors, non-response, and attrition on income inequality, poverty and mobility*, in Ehling M. e Rendtel U. editors, *Harmonisation of panel surveys and data quality*, Statistisches Bundesamt, Wiesbaden (report downloadable from <http://www.abo.fi/~mjantti/>).
- Juster F.T. e Smith J.P. 1997. *Improving the Quality of Economic Data: Lessons from HRS and AHEAD*, *Journal of the American Statistical Association*, 92, pp. 1268-78.
- Kahn R.L. e Cannel C.F. 1957. *The Dynamics of Interviewing*, 11<sup>th</sup> edition, Wiley & Sons, New York. Tr. it. 1968. *La dinamica dell'intervista*, 3<sup>a</sup> edizione, Marsilio Editori, Padova.
- Keating E., Paterson D.G. e Stone C.H. 1950. Validity of Work Histories by Interview, *Journal of Applied Psychology*, 34, pp. 6-11.
- Lessler J.T. e Kalsbeek W. D. 1992. *Nonsampling Errors in Surveys*, Wiley & Sons, New York.
- Mandell L. e Lundsten L.L. 1978. *Some Insight into the Underreporting of Financial Data by Sample Survey Respondents*, *Journal of Marketing Research*, 15 (2), pp. 294-299.
- Marenzi A. 1996. Prime analisi sulla distribuzione dell'IRPEF per categorie di contribuenti e per livelli di reddito, Rossi N. (a cura di), *Competizione e giustizia sociale 1994-1995*, terzo rapporto CNEL sulla distribuzione e redistribuzione del reddito in Italia, il Mulino, Bologna, pp. 305 -341.
- Martini M. 1990. *I dati amministrativi come fonte di informazione statistica sulle imprese*, *Economia & Lavoro*, XXIV (2), pp. 45-58.
- Mellow W. e Sider H. 1983. *Accuracy of Response in Labor Market Surveys: Evidence and Implications*, *Journal of Labor Economics*, 1 (4), pp. 331-342.
- Moore J., Stinson L. e Welniak E.J. 1997. *Income Measurement Error in Surveys: A Review*, U.S. Bureau of the Census, S.M. 97/5, Washington, DC.
- Pedace R. e Bates N. 2000. *Using administrative records to assess earnings reporting error in the survey of income and program participation*, *Journal of Economic & Social Measurement*, 26 (3/4), pp. 173-192.

- Quintano C. e Lucev D. 1990. *Le mancate risposte in esperienze di indagini reddituali*, Quaderni sardi di economia, 20 (3), pp. 253-278.
- Rodgers W.L. Brown C. e Duncan G.J. 1993. *Errors in Survey Reports of Earnings, Hour Worked, and Hourly Wages*, Journal of the American Statistical Association, 88, pp. 1208-1218.
- Stata (2005). *Stata Base Reference Manual*, voll. 1-3, Stata Press, College Station, TX.
- Sudman S. e Bradburn N.M. 1973. *Effects of time and memory factors on response in surveys*, Journal of the American Statistical Association, 68, pp. 805–815.
- Tourangeau R., Rips L.J. e Rasinski K. 2000. *The Psychology of Survey Response*, Cambridge University Press, Cambridge.