

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

**Il rendimento dell'istruzione:
alcuni problemi di stima**

di Luigi Cannari e Giovanni D'Alessio



Numero 253 - Marzo 1995

Temi di discussione

del Servizio Studi

La serie "Temi di discussione" intende promuovere la circolazione, in versione provvisoria, di lavori prodotti all'interno della Banca d'Italia o presentati da economisti esterni nel corso di seminari presso l'Istituto, al fine di suscitare commenti critici e suggerimenti.

I lavori pubblicati nella serie riflettono esclusivamente le opinioni degli autori e non impegnano la responsabilità dell'Istituto.

Comitato di redazione:

GIORGIO GOMEL, EUGENIO GAIOTTI, CURZIO GIANNINI, LUGI GUISO, DANIELE TERLIZZESE,
SILIA MIGLIARUCCI (*segretaria*).

**Il rendimento dell'istruzione:
alcuni problemi di stima**

di Luigi Cannari e Giovanni D'Alessio

Numero 253 - Marzo 1995

**IL RENDIMENTO DELL'ISTRUZIONE:
ALCUNI PROBLEMI DI STIMA**

di Luigi Cannari (*) e Giovanni D'Alessio (**)

Sommario

In questo lavoro si esaminano alcuni problemi che caratterizzano le stime del rendimento dell'istruzione nella funzione generatrice del reddito (earnings function). Per valutare la presenza di eventuali distorsioni nelle stime del coefficiente della scolarità ottenute con i minimi quadrati ordinari, si ricorre a metodi che utilizzano indicatori di background familiare come variabili strumentali per l'istruzione o sfruttano l'ipotesi di invarianza temporale dell'abilità dei lavoratori, trattata come variabile non osservabile. Gli esperimenti, condotti utilizzando l'indagine sui bilanci delle famiglie della Banca d'Italia, portano alla conclusione che il metodo dei minimi quadrati ordinari determina una significativa sottostima del coefficiente della scolarità; il rendimento di un anno di istruzione sulle retribuzioni dei lavoratori dipendenti, che le stime OLS collocano tra il 3 e il 5 per cento secondo le diverse specificazioni, è valutabile tra il 6 e l'8 per cento sulla base degli stimatori alternativi utilizzati.

Indice

1. Introduzione	p.	7
2. I dati utilizzati	p.	12
3. Le stime <u>panel</u>	p.	14
4. Istruzione e <u>background</u> familiare	p.	20
4.1 Il campione dei figli	p.	20
4.2 Il campione dei capifamiglia	p.	24
5. Considerazioni conclusive	p.	29
Appendice	p.	33
Riferimenti bibliografici	p.	36

(*) Banca d'Italia, Sede di Firenze, Nucleo per la ricerca economica

(**) Banca d'Italia, Servizio Studi.

1. Introduzione¹

Alla stima delle funzioni generatrici del reddito (earnings function) è stata dedicata una considerevole attenzione, a partire dai lavori di Mincer (1958) e Becker (1964)². In tale contesto, l'interesse si è spesso focalizzato su questioni concernenti il rendimento di un investimento in istruzione, sui differenziali salariali settoriali o geografici, sulla presenza di discriminazioni nel mercato del lavoro, ad esempio tra uomini e donne o tra individui di razze diverse³.

Nella formulazione che più frequentemente si riscontra in letteratura, la funzione generatrice del reddito è una relazione lineare tra il logaritmo delle retribuzioni e una serie di variabili esplicative, tra cui l'esperienza lavorativa (o l'età) e il livello di istruzione dei soggetti osservati.

Questa relazione, apparentemente semplice, pone tuttavia alcuni problemi per la stima econometrica dei parametri. Una variabile potenzialmente rilevante per la spiegazione dei differenziali di reddito è l'abilità dei lavoratori, intesa

1. Le opinioni espresse non impegnano la Banca d'Italia. Gli autori desiderano ringraziare Andrea Brandolini, Andrea Giommi, Luigi Federico Signorini, Paolo Sestito e un anonimo referee, per i loro preziosi suggerimenti. Una precedente versione del lavoro è stata presentata al convegno "Quantitative Methods for Applied Sciences" tenutosi a Siena, Certosa di Pontignano, 27-29 giugno 1994. La responsabilità di quanto scritto è, ovviamente, degli autori.
2. Per una rassegna sulle funzioni generatrici del reddito e sulla teoria del capitale umano cfr. Willis (1986).
3. Per contributi italiani in questo campo di ricerca, tra gli altri, cfr. Commissione Carniti (1988), Dagum e Lemmi (1988), Lucifora (1988), Cannari, Pellegrini e Sestito (1989), Brunello e Rizzi (1993), Birindelli e altri (1994).

nella sua più ampia accezione (capacità innate, impegno lavorativo, ecc.). Questa variabile non è facilmente osservabile e le proxies talvolta utilizzate, come ad esempio i risultati di test psico-attitudinali o il Q.I., non suppliscono a questa carenza informativa in modo soddisfacente. L'abilità di cui si tratta è infatti più prossima al concetto di "capacità di far soldi" che a quello di intelligenza, o di cultura generale, o di abilità scolastica. Per questi motivi, nelle applicazioni pratiche l'abilità non viene generalmente inserita tra le variabili esplicative del reddito, ed è pertanto considerata tra i residui del modello⁴. Come noto, nel caso di omissione dal modello di una variabile rilevante, le stime dei parametri ottenute con il metodo dei minimi quadrati ordinari (OLS) sono distorte e inconsistenti.

Problemi di stima derivano anche dal fatto che la scolarità è presumibilmente funzione dell'abilità e quindi è una variabile endogena nel modello. In un mondo caratterizzato da informazione imperfetta, le persone più abili possono studiare di più per inviare un "segnale" circa le proprie capacità ai potenziali datori di lavoro, i quali, a loro volta, possono utilizzare il livello d'istruzione come variabile di screening per la selezione dei lavoratori. Inoltre, se l'abilità accresce il rendimento di un investimento in istruzione, gli individui più abili studieranno più di quelli meno abili, per le maggiori chances di trarre profitto dall'accresciuta scolarità⁵. La stima OLS del coefficiente

-
4. Gli studi che utilizzano proxies tendono in genere a concludere che l'abilità non gioca un ruolo di particolare rilievo come variabile esplicativa dei redditi. Come indicato nel testo, tuttavia, la "capacità di far soldi" può essere ben diversa dalla capacità di superare con successo un test attitudinale.
 5. Per una discussione sulla relazione tra abilità, istruzione e retribuzioni nei modelli di signalling e in quelli di capitale umano cfr., per esempio, Blackburn e Neumark (1993).

dell'istruzione dovrebbe essere pertanto distorta verso l'alto (e inconsistente) a causa della correlazione positiva tra l'abilità (non osservabile e inserita tra i residui) e l'istruzione, congiuntamente con il fatto che l'abilità ha un effetto positivo sul reddito.

Va tuttavia osservato che il grado di istruzione è il risultato, almeno in parte, di un comportamento ottimizzante degli individui, che cercano di massimizzare la propria funzione di utilità intertemporale. In termini di reddito non guadagnato, il costo dell'istruzione è più elevato per le persone più abili, che possono dunque essere indotte a studiare di meno. Sulla base di tale argomentazione, la correlazione tra l'abilità e l'ammontare ottimo di scolarità potrebbe anche essere negativa (al riguardo cfr. Griliches, 1977), comportando una distorsione verso il basso della stima OLS del rendimento dell'istruzione.

La circostanza che le scelte siano, almeno parzialmente, effettuate dai genitori rendono ancor più complesso individuare, a priori, il segno della correlazione tra l'abilità e la scolarità. I genitori potrebbero, da un lato, investire di più nei figli meno capaci, per sostenerne le sorti future o perché il costo opportunità della maggiore scolarità è in questo caso più contenuto; dall'altro, potrebbero investire di più nei figli più abili, quelli che hanno più chances di trarre profitto dall'investimento in istruzione.

Ulteriori problemi di stima possono infine derivare dalla presenza di errori di misurazione nelle variabili esplicative delle retribuzioni che, come noto, rendono le stime OLS inconsistenti.

Le verifiche empiriche hanno fornito risultati diversi anche se spesso convergenti nell'indicare l'inadeguatezza delle stime OLS.

Affrontando il problema dell'abilità come variabile non osservabile, Griliches (1977) è stato il primo a mettere in evidenza che l'ipotesi, fino ad allora più accreditata, di una distorsione positiva del coefficiente OLS dell'istruzione, dovuta all'omissione dell'abilità, non trova riscontro nei dati. Piuttosto, il coefficiente OLS della scolarità appare sottostimato, indicando la presenza di una correlazione negativa tra abilità e livello di istruzione.

Ulteriori lavori, utilizzando diverse formulazioni del modello e con differenti argomentazioni, hanno confermato la distorsione verso il basso del coefficiente dell'istruzione ottenuto con gli OLS. Griliches, Hall e Hausman (1978), affrontando il problema dell'endogenità dell'istruzione, mostrano che l'uso di variabili strumentali di background familiare produce un incremento del coefficiente della scolarità pari a circa il 50 per cento rispetto a quello ottenuto con gli OLS. A un analogo risultato giungono sia Hausman e Taylor (1981), che utilizzano stimatori con variabili strumentali (IV) basati su restrizioni di esogenità e sull'ipotesi di invarianza temporale dell'abilità, sia Angrist e Newey (1990) che, esaminando i problemi di identificazione e di stima del rendimento della scolarità in modelli a effetti fissi, giungono alla conclusione che il rendimento dell'istruzione è approssimativamente il doppio di quello ottenuto con il metodo OLS.

Altri autori riscontrano distorsioni significative degli stimatori OLS, che sembrano derivare non tanto dall'omissione dell'abilità quanto dalla presenza di errori di misurazione nelle variabili esplicative. In particolare, Ashenfelter e Krueger (1992), utilizzando misurazioni multiple dei livelli di scolarità in un campione di gemelli, concludono che le variabili omesse non determinano distorsioni nel rendimento della scolarità, che invece sarebbe sot-

tostimato dal metodo OLS a causa di errori di misurazione. Angrist e Krueger (1992) ottengono risultati simili: il metodo OLS sottostima il rendimento della scolarità e la distorsione appare dello stesso ordine di grandezza di quella attesa in presenza di errori di misurazione nel livello di istruzione.

Più articolata la conclusione del lavoro di Blackburn e Neumark (1993) secondo cui i metodi basati su variabili strumentali conducono a stime considerevolmente più elevate del rendimento della scolarità solo per le retribuzioni d'ingresso nel mercato del lavoro, mentre per le retribuzioni riferite ad alcuni anni dopo l'entrata nel mercato del lavoro il metodo OLS conduce a una sovrastima del coefficiente della scolarità.

A risultati diversi pervengono invece Angrist e Krueger (1991), i quali mostrano che le stime OLS non differiscono significativamente da quelle ottenute con il metodo delle variabili strumentali, e Ottersten e altri (1994) che concludono che i diversi fattori di distorsione possono operare in direzioni differenti, compensandosi, in una certa misura, l'uno con l'altro.

Lo scopo di questo lavoro è di valutare l'ampiezza e il segno dell'eventuale distorsione delle stime OLS del rendimento dell'istruzione, discutendone le possibili cause. Dopo una breve descrizione dei dati utilizzati (par. 2), si illustrano i metodi di stima e i principali risultati nei paragrafi 3 e 4. Alcune considerazioni conclusive sono infine riportate nel paragrafo 5.

2. I dati utilizzati

I dati utilizzati in questo lavoro sono tratti dall'indagine sui bilanci delle famiglie condotta dalla Banca d'Italia e si riferiscono ai lavoratori maschi che hanno lavorato tutto l'anno a tempo pieno con attività principale di lavoro dipendente⁶.

Una prima analisi è stata condotta utilizzando i dati relativi ai lavoratori dipendenti facenti parte del panel, ovvero del sottoinsieme del campione che dopo essere stato intervistato nel 1989 è stato nuovamente oggetto di intervista nel 1991 (Campione A). Questo sottoinsieme presenta alcuni limiti. Il tasso di risposta è basso e correlato con le caratteristiche familiari (cfr. Cannari e D'Alessio, 1992). Tuttavia, come evidenziato nel paragrafo 3, le stime dei parametri della funzione generatrice del reddito effettuate sul complesso del campione non differiscono significativamente da quelle ottenute sul panel.

Un'analisi alternativa è stata condotta utilizzando variabili di background familiare come strumenti per la scolarità. Per questa analisi si è fatto ricorso alle indagini della Banca d'Italia relative al 1991 e al 1993.

La prima indagine rileva informazioni sulle caratteristiche sociali e demografiche dei soli componenti la famiglia che vivono nella stessa abitazione. L'insieme di individui utilizzato in questo esperimento è dunque necessariamente limitato ai lavoratori dipendenti che convivono con i genitori (Campione B). Questo sottoinsieme del campione è

6. Per una descrizione dettagliata dell'indagine della Banca d'Italia, comprensiva di un confronto tra le stime campionarie e i dati di contabilità nazionale, si rinvia a Brandolini e Cannari (1994). Una descrizione dell'indagine più mirata al mercato del lavoro è riportata in Cannari e Gavosto (1994).

presumibilmente poco rappresentativo della popolazione dei lavoratori dipendenti; nella famiglia di origine potrebbero restare prevalentemente i lavoratori meno abili (che dovrebbero percepire redditi più modesti), ad esempio per beneficiare di economie di scala nella spesa per consumi. Per limitare questo effetto di selezione, si sono presi in considerazione solo i lavoratori più giovani (quelli fino a 30 anni di età). Se le persone all'inizio della vita lavorativa tendono a restare all'interno della famiglia di origine per un certo periodo, ad esempio per risparmiare il necessario per intraprendere la loro vita indipendente, la distorsione di selezione sopra citata dovrebbe essere modesta nelle classi di età più giovani. Valutazioni più precise circa la rappresentatività di questo sottoinsieme sono riportate nel seguito del lavoro. Si ritiene però utile anticipare che le stime OLS relative ai lavoratori nella classe di età 16-30 anni che vivono con i genitori non appaiono significativamente diverse da quelle relative al complesso dei lavoratori della stessa età.

Per contro, l'indagine sul 1993 rileva informazioni sul background familiare del capofamiglia e del coniuge, per tutte le famiglie intervistate; questo insieme (Campione C) non presenta pertanto i problemi di selezione del precedente insieme di individui.

Va rilevato che l'indagine della Banca d'Italia, come ogni indagine statistica di questo genere, è affetta da errori di natura non campionaria⁷. La presenza di errori di misurazione del livello di istruzione (e/o di altre variabili) può far sì che il ricorso a variabili strumentali determini un incremento del coefficiente dell'istruzione nella funzione generatrice del reddito. In questo caso, stime OLS diverse da

7. Una sintetica descrizione del panel della Banca d'Italia e alcune considerazioni sui problemi di qualità dell'indagine sono riportate in appendice.

quelle IV potrebbero ottenersi anche se l'istruzione fosse una variabile esogena. Per questo motivo sono stati esclusi dal campione tutti i records che presentano incoerenze superiori a una soglia prestabilita⁸.

3. Le stime panel

La funzione generatrice del reddito utilizzata nel presente lavoro può scriversi nel modo seguente:

$$(1) \quad Y_{it} = X_{it}\beta + Z_{it}\tau + \alpha_i + \mu_{it}$$

dove Y_{it} è il logaritmo del rapporto tra le retribuzioni nette annuali e le ore mediamente lavorate per settimana, X_{it} è un insieme di variabili esplicative che presentano una dinamica temporale (variabili time-varying) come

8. In particolare:

Campione A: dipendenti maschi presenti nel panel 1989-1991. Si sono escluse le osservazioni per le quali: anni di istruzione + 6 > età; esperienza lavorativa + 14 > età; esperienza con l'attuale datore di lavoro > esperienza lavorativa. Tra i records restanti si sono selezionati quelli con: sesso e livello di istruzione uguali nel 1989 e nel 1991; età nel 1991 = età nel 1989 + 2; differenza tra l'età alla prima occupazione rilevata nel 1989 e l'età alla prima occupazione rilevata nel 1991 inferiore a 1 in valore assoluto (quando la discrepanza era pari a 1 si è utilizzato il valore medio delle due età); esperienza con l'attuale datore di lavoro rilevata nel 1991 uguale a esperienza con il datore di lavoro rilevata nel 1989 + 2 oppure esperienza con il datore di lavoro nel 1991 < 2 (a questa condizione è stata consentita una tolleranza di un 1 anno di differenza in valore assoluto).

Campione B: dipendenti maschi tra 16 e 30 anni conviventi con i genitori. Si sono escluse le osservazioni per le quali: anni di istruzione + 6 > età; esperienza lavorativa + 14 > età; esperienza con l'attuale datore di lavoro > esperienza lavorativa.

Campione C: capifamiglia maschi fino a 60 anni di età. Si sono escluse le osservazioni per le quali: anni di istruzione + 6 > età.

l'esperienza con l'attuale datore di lavoro (tenure), l'età o l'esperienza lavorativa (in termini lineari e quadratici) e Z_i è un insieme di variabili esplicative che a livello individuale sono costanti nel tempo (variabili time-invariant), tra le quali si include la scolarità⁹. La componente α_i rappresenta l'abilità (un effetto fisso individuale non osservabile) e μ_{it} sono i termini di errore. Gli indici i ($i=1, \dots, N$) e t ($t=1, \dots, T$) si riferiscono rispettivamente agli individui e al tempo. Nel campione cross-section T è ovviamente pari a 1; nelle stime panel si dispone invece di due osservazioni per ciascun individuo ($T=2$). I coefficienti β e τ sono i parametri da stimare. Si assume inoltre che μ_{it} siano incorrelati con le colonne (X, Z, α) e abbiano media zero e varianza costante σ^2_{μ} e che gli α_i siano incorrelati tra gli individui e abbiano varianza pari a σ^2_{α} ¹⁰.

Come accennato in precedenza, l'equazione (1) pone alcuni problemi econometrici, a causa della potenziale correlazione tra l'abilità e l'istruzione. Una possibile soluzione al problema è quella di ricorrere a dati panel (Campione A), stimando il sistema trasformato:

9. In realtà il livello di istruzione è variabile nell'arco della vita degli individui, anche dopo che questi hanno iniziato a lavorare. Nel caso dei lavoratori dipendenti a tempo pieno occupati per l'intero anno, i casi di variazione del livello di istruzione sono però assai limitati e in parte dovuti, probabilmente, a errori di misurazione. Nelle stime basate sul panel, pertanto, l'analisi è stata limitata ai soli casi in cui il titolo di studio registrato nel 1989 era uguale a quello del 1991.
10. Funzioni generatrici del reddito come la (1) sono state frequentemente utilizzate in letteratura e mostrano un buon adattamento ai dati anche nel caso italiano. Cannari, Pellegrini e Sestito (1989), replicando l'esperimento di Heckman e Polachek (1974) su dati italiani, affermano che i migliori risultati si ottengono specificando i redditi in forma logaritmica e che l'adattamento del modello non è molto sensibile a trasformazioni di tipo Box-Cox delle variabili esplicative.

$$(2a) \quad Y_{it} - Y_{it-1} = (X_{it} - X_{it-1})\beta + \epsilon_{it}$$

$$(2b) \quad Y_{i.} = X_{i.}\beta + Z_{i.}\tau + \alpha_i + \mu_i.$$

dove:

$$Y_{i.} = (1/T) \sum_t Y_{it}$$

$$X_{i.} = (1/T) \sum_t X_{it}$$

$$\epsilon_{it} = \mu_{it} - \mu_{it-1}$$

$$\mu_{i.} = (1/T) \sum_t \mu_{it}$$

e utilizzando stimatori basati sul metodo dei momenti (GMM), del tipo Hausman e Taylor (1981) o Amemiya e MaCurdy (1986)¹¹. Se si dispone di più osservazioni temporali per ciascun individuo, i coefficienti delle variabili time-varying possono essere stimati sull'equazione differenziata (2a) e le variabili time-varying incorrelate con l'abilità possono essere utilizzate come strumenti per l'istruzione nell'equazione sui livelli (2b).

Prima di commentare i risultati che si ottengono con gli stimatori GMM è però utile comparare il campione panel (Campione A) con il complesso del campione dei lavoratori dipendenti tra 16 e 60 anni di età, al fine di valutare la presenza e la dimensione di eventuali distorsioni dovute a fenomeni di selezione del panel (tav. 1).

Per i maschi tra 16 e 60 anni di età, occupati a tempo pieno per l'intero anno, con attività principale di lavoro

11. Per maggiori dettagli sulle caratteristiche di questi stimatori cfr., per esempio, Arellano e Bover (1990).

dipendente, non si riscontrano differenze marcate tra i dati panel e il campione complessivo del 1991.

L'età media e l'esperienza media sono di quasi due anni superiori nel panel rispetto al campione complessivo a causa dell'"invecchiamento" del primo (si considerano solo gli individui che svolgono da almeno due anni attività di lavoro dipendente); il livello dell'istruzione è mediamente più elevato tra coloro che hanno partecipato a precedenti interviste, probabilmente per la maggiore disponibilità delle persone con una più elevata scolarità a collaborare all'indagine.

Tav. 1

VALORI MEDI E SCARTI QUADRATICI MEDI
DELLE PRINCIPALI VARIABILI, 1991

Variabile	Totale campione (1)		Panel 1989-1991 (2) (Campione A)	
	Media	S.Q.M.	Media	S.Q.M.
LW (3)	6,257	0,334	6,287	0,328
ET (4)	39,216	10,597	40,892	9,947
EX (5)	19,636	11,208	21,400	10,383
TE (6)	13,234	9,758	12,889	9,022
ED (7)	10,054	3,836	10,512	3,739
SI (8)	0,354	0,478	0,335	0,473

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie.

(1) Lavoratori dipendenti maschi nella classe di età 16-60 anni che hanno lavorato a tempo pieno per l'intero 1991. Numero di osservazioni: 3.635. - (2) Lavoratori dipendenti maschi nella classe di età 16-60 anni che hanno lavorato a tempo pieno per gli interi anni 1989 e 1991. Numero di osservazioni: 334 per ciascun anno. I dati riportati nella tavola si riferiscono al 1991. - (3) Logaritmo del rapporto tra le retribuzioni annuali nette e il numero di ore mediamente lavorate alla settimana. - (4) Età in anni. - (5) Esperienza. - (6) Esperienza con l'attuale datore di lavoro. - (7) Numero di anni per conseguire il titolo di studio. - (8) Dummy Sud-Isole.

Le stime OLS condotte sui due insiemi sono abbastanza simili. Nel modello che include l'esperienza lavorativa tra le variabili esplicative, il coefficiente della scolarità risulta pari a 0,0430 nel panel e a 0,0447 in quello del totale dei lavoratori (tav. 2, colonne 1 e 3). Un po' superiori

appaiono invece le differenze che si riscontrano nel modello che include l'età in luogo dell'esperienza (i coefficienti della scolarità sono rispettivamente pari a 0,0394 e a 0,0344; tav. 2, colonne 2 e 4).

Tav. 2

PARAMETRI DELLA FUNZIONE GENERATRICE DEL REDDITO(1)

Parametri (2)	Totale campione 1991		Panel 1989-1991 (Campione A)			
	OLS		OLS		GMM	
Istruzione (ED)	4,47 (0,123)	3,94 (0,114)	4,30 (0,305)	3,44 (0,277)	7,54 (2,315)	10,48 (4,471)
Età (ET)	-	4,86 (0,301)	-	5,21 (0,761)	-	1,48 (1,795)
Età (quadrato) (ET2)	-	-0,048 (0,0038)	-	-0,052 (0,0095)	-	-0,002 (0,0233)
Esperienza (EX)	2,63 (0,156)	-	2,91 (0,427)	-	2,53 (0,512)	-
Esperienza quadrato (EX2)	-0,042 (0,0035)	-	-0,042 (0,0094)	-	-0,019 (0,0148)	-
Tenure (TE)	0,43 (0,064)	0,38 (0,063)	0,42 (0,168)	0,50 (0,164)	0,05 (0,323)	0,29 (0,236)
Sud-Isole (SI)	-4,14 (0,924)	-6,11 (0,904)	-1,53 (2,236)	-2,29 (2,206)	-0,15 (2,840)	-3,80 (3,707)
n.oss.	3635	3635	668	668	668	668
ADJ-R ²	0,38	0,39	0,39	0,40	-	-
r	-	-	-	-	0,68	0,77
Endogene	-	-	-	-	ED	ED
Strumenti (3)	-	-	-	-	EX(4)	ET
					EX2(4)	ET2
					TE	TE

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie.

(1) Dipendenti maschi che hanno lavorato tutto l'anno a tempo pieno. Variabile dipendente: log del rapporto tra le retribuzioni nette e le ore lavorate. Come strumenti per le ore lavorate si utilizzano le seguenti variabili: età, esperienza lavorativa, tenure, area geografica, condizione professionale, settore di attività, livello di istruzione. Oltre alle variabili esplicative riportate nella tavola nel modello è inclusa l'intercetta e una dummy temporale (quest'ultima solo nei dati panel). - (2) Le stime dei parametri e gli errori standard (tra parentesi) sono moltiplicati per 100. - (3) Strumenti utilizzati per l'equazione sui livelli. - (4) Al livello di significatività dell'1 per cento il test di Hausman non consente di rifiutare l'ipotesi di esogenità di EX e EX2. L'ipotesi è rifiutata al livello del 5 per cento.

L'ipotesi di uguaglianza del coefficiente del rendimento dell'istruzione per il panel e per gli altri individui è stata sottoposta a verifica, effettuando una stima dei parametri della funzione generatrice del reddito sull'intero campione rilevato nel 1991 e introducendo una dummy per l'istruzione in corrispondenza delle osservazioni panel. Il test non consente di rifiutare l'ipotesi, al livello di significatività del 5 per cento.

Ricorrendo agli stimatori GMM e utilizzando l'esperienza (termine lineare e quadratico) e la tenure come strumenti per l'istruzione nell'equazione sui livelli, si osserva in primo luogo che il coefficiente di correlazione dei residui nel periodo di due anni (1989 e 1991) è piuttosto elevato ($r = 0,680$; tav. 2, colonna 5); ciò indica la rilevanza dell'effetto individuale non osservato. Inoltre il coefficiente della scolarità mostra un considerevole incremento (da 0,0430 a 0,0754; tav. 2, colonne 3 e 5). Nel modello specificato in funzione dell'età l'incremento risulta ancora più ampio (da 0,0344 a 0,1048; tav. 2, colonne 4 e 6).

Queste stime presentano tuttavia alcuni limiti: nel primo modello, l'ipotesi di esogenità dell'esperienza appare troppo restrittiva e viene rifiutata dal test di Hausman al livello di significatività del 5 per cento; se tale variabile è trattata come endogena (tav. 2, colonna 5), il modello è esattamente identificato e non può essere sottoposta a test l'ipotesi di esogenità della tenure. Nel secondo modello, che include l'età come variabile esplicativa, la tenure risulta endogena; se si utilizza solo l'età come variabile strumentale, il coefficiente della scolarità aumenta ulteriormente fino a raggiungere l'implausibile valore di 0,245, a causa della variabilità che caratterizza le stime.

4. Istruzione e background familiare

4.1 Il campione dei figli

Una stima alternativa del rendimento della scolarità può essere effettuata ricorrendo a stimatori IV che utilizzano come strumenti variabili di background familiare. In questo paragrafo, tale analisi viene condotta sul campione dei lavoratori dipendenti che convivono con i genitori (Campione B).

Come nel caso del panel, appare utile valutare preliminarmente le eventuali distorsioni dovute al criterio di selezione e, di conseguenza, in che misura i risultati possono essere ritenuti rappresentativi del totale dei lavoratori dipendenti della classe 16-30 anni.

I valori medi delle principali variabili relative al campione complessivo e a quello dei figli sono riportati nella tavola 3.

Le differenze sono in generale abbastanza contenute, anche se non del tutto trascurabili. I lavoratori dipendenti che convivono con i genitori hanno un'età media inferiore di un anno (24,2 contro 25,2), cui corrisponde una minore esperienza lavorativa (4,9 anni contro 6,1 anni). Il livello di istruzione è solo lievemente superiore a quello del totale dei lavoratori dipendenti della stessa classe di età (10,4 contro 10,2). Il logaritmo delle retribuzioni assume valori assai simili nei due casi (5,977 contro 6,046).

**VALORI MEDI E SCARTI QUADRATICI MEDI
DELLE PRINCIPALI VARIABILI, 1991**

Variabile	Totale classe 16-30 anni (1)		Campione dei figli(2) (Campione B)	
	Media	S.Q.M.	Media	S.Q.M.
LW (3)	6,046	0,303	5,977	0,290
ET (4)	25,189	3,509	24,175	3,503
EX (5)	6,145	3,969	4,933	3,467
TE (6)	4,188	3,239	3,537	2,751
ED (7)	10,230	3,062	10,415	2,978
SI (8)	0,279	0,449	0,277	0,448
ETCF (9)	-	-	55,906	6,947
ETCG (9)	-	-	52,246	7,074
EDCF (9)	-	-	6,504	3,701
EDCG (9)	-	-	5,975	3,403

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie.

(1) Lavoratori dipendenti maschi nella classe di età 16-30 anni che hanno lavorato a tempo pieno per l'intero anno. Numero di osservazioni: 876. - (2) Lavoratori dipendenti maschi nella classe di età 16-30 anni che hanno lavorato a tempo pieno per l'intero anno e convivono con entrambi i genitori. Numero di osservazioni: 520. - (3) Logaritmo del rapporto tra le retribuzioni annuali nette e il numero di ore mediamente lavorate alla settimana. - (4) Età in anni. - (5) Esperienza. - (6) Esperienza con l'attuale datore di lavoro. - (7) Numero di anni per conseguire il titolo di studio. - (8) Dummy Sud-Isole. - (9) Nell'ordine: età del genitore capofamiglia e del coniuge del genitore; istruzione del genitore capofamiglia e del coniuge del genitore. Variabili disponibili solo per i lavoratori che convivono con entrambi i genitori.

Le stime OLS condotte su questi due insiemi si presentano abbastanza simili¹². Nel modello che include l'esperienza lavorativa tra le variabili esplicative, il coefficiente della scolarità risulta pari a 0,0468 nel campione dei figli e a 0,0443 in quello del totale dei lavoratori nella classe di età 16-30 anni (tav. 4, colonne 1 e 3). Anche nel modello che include l'età in luogo dell'esperienza i coefficienti della scolarità risultano simili (0,0308 contro

12. Dato il limitato campo di variazione dell'età (da 16 a 30 anni), non si è ritenuto necessario inserire il termine quadratico dell'età (o dell'esperienza) tra le variabili esplicative del modello.

0,0291; tav. 4, colonne 2 e 4). Come per il campione A, l'ipotesi di uguaglianza del coefficiente del rendimento dell'istruzione per il campione in esame e per l'intero campione della stessa classe di età è stata sottoposta a verifica introducendo una dummy per l'istruzione in corrispondenza delle osservazioni del campione B nella stima dei parametri della funzione generatrice del reddito. Come in precedenza, il test non consente di rifiutare l'ipotesi di uguaglianza al livello di significatività del 5 per cento.

Sulla base di questi risultati, l'estensione al complesso dei lavoratori dipendenti (della classe 16-30 anni) dei risultati che si otterranno utilizzando il metodo delle variabili strumentali al campione dei figli (in cui sono disponibili variabili di background familiare) non appare pertanto irragionevole. A margine, si fa notare che, sostituendo l'esperienza con l'età (tav. 4, colonne 3 e 4), il coefficiente dell'istruzione diminuisce da 0,0468 a 0,0308 in quanto a parità di età un livello di istruzione più elevato implica infatti una minore esperienza lavorativa. Il differenziale retributivo del Sud-Isole rispetto al Nord-Centro passa dal -7,2 per cento al -9,7 per cento, segnalando che, a parità di età e di istruzione, l'ingresso sul mercato del lavoro avviene a un'età superiore nel Sud-Isole rispetto al resto del paese.

Il ricorso a variabili strumentali (età e istruzione del padre e della madre) modifica sensibilmente il coefficiente della scolarità, che passa da 0,0468 a 0,0786 quando nel modello è inclusa l'esperienza e da 0,0308 a 0,0528 quando si considera l'età come variabile esplicativa (tav. 4, colonne 3 e 5 e colonne 4 e 7). In ambedue i casi, l'ipotesi che l'istruzione sia esogena è rifiutata agli usuali livelli di significatività del 5 per cento.

PARAMETRI DELLA FUNZIONE GENERATRICE DEL REDDITO(1)

Parametri (2)	Totale classe di età 16-30 anni		Individui nella classe di età 16-30 che convivono con i genitori (Campione B)				
	OLS		OLS		IV		
Istruzione (ED)	4,43 (0,305)	2,91 (0,308)	4,68 (0,382)	3,08 (0,409)	7,86 (0,789)	9,17 (1,058)	5,28 (0,946)
Età (ET)	-	3,16 (0,303)	-	2,95 (0,385)	-	-	1,96 (0,550)
Esperienza (EX)	1,34 (0,368)	-	0,76 (0,444)	-	1,31 (0,488)	7,65 (2,323)	-
Tenure (TE)	0,99 (0,368)	0,23 (0,319)	0,81 (0,552)	-0,53 (0,466)	0,88 (0,588)	-4,32 (1,976)	0,42 (0,604)
Sud-Isole (SI)	-4,26 (2,034)	-8,09 (1,941)	-7,16 (2,496)	-9,73 (2,377)	-5,50 (2,682)	-1,82 (3,483)	-8,26 (2,507)
n.oss.	876	876	520	520	520	520	520
ADJ-R ²	0,22	0,29	0,24	0,31	0,18	0,14	0,28
Endogene	-	-	-	-	ED(3)	ED	ED(3)
	-	-	-	-	-	EX(4)	-
Strumenti	-	-	-	-	(5)	(5)	(5)

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie nel 1991.

(1) Dipendenti maschi che hanno lavorato tutto l'anno 1991 a tempo pieno. Variabile dipendente: log del rapporto tra le retribuzioni nette e le ore lavorate. Come strumenti per le ore lavorate si utilizzano le seguenti variabili: età, esperienza lavorativa, tenure, area geografica, condizione professionale, settore di attività, livello di istruzione. Oltre alle variabili esplicative riportate nella tavola nel modello è inclusa l'intercetta. - (2) Le stime dei parametri e gli errori standard (tra parentesi) sono moltiplicati per 100. - (3) Il test di Hausman rifiuta l'ipotesi che ED sia esogena. - (4) Il test di Hausman rifiuta l'ipotesi che EX sia esogena. - (5) Oltre alle esogene del modello si utilizzano i seguenti strumenti: EDCF e EDCG = rispettivamente istruzione del padre e della madre; ETCF e ETCG = rispettivamente età del padre e della madre.

Inoltre anche l'esperienza lavorativa risulta essere correlata con i residui della funzione generatrice del reddito (essendo correlata con la differenza tra l'età e gli anni di istruzione). Se anche quest'ultima variabile viene considerata endogena e quindi trattata come tale, utilizzando anche in questo caso variabili di background familiare come strumenti, il coefficiente della scolarità mostra un ulteriore incremento, divenendo 0,917 (tav. 4, colonna 6). In questo modello, l'ipotesi che la tenure sia esogena non è rifiutata dal test di Hausman; il risultato del test richiede

però una certa cautela, dal momento che gli strumenti utilizzati non sono pienamente soddisfacenti quando le variabili istruzione, esperienza e tenure sono trattate congiuntamente come endogene. Risultati simili si ottengono nel caso del modello che include l'età come variabile esplicativa.

4.2 Il campione dei capifamiglia

Le stime riportate nel precedente paragrafo prestano il fianco ad alcune critiche. In primo luogo il campione dei figli (Campione B) può essere affetto da problemi di autoselezione (come argomentato nel paragrafo 2). Secondariamente, appare dubbia l'esogenità sia dell'esperienza sia della tenure, ma gli strumenti disponibili per trattare tali variabili come endogene, assieme alla scolarità, non sono pienamente soddisfacenti. Per ovviare a questi problemi, si utilizza il campione di capifamiglia rilevati nell'indagine sul 1993 (Campione C) e si ricorre a un modello semplificato in cui le retribuzioni sono espresse in funzione delle sole variabili istruzione, età ed età al quadrato (più il termine di intercetta e la dummy territoriale). I valori medi delle variabili utilizzate sono riportati nella tavola 5.

Anche in questo caso, l'ipotesi di esogenità della scolarità viene rifiutata dai dati. Utilizzando come strumenti il livello d'istruzione dei genitori del capofamiglia, il coefficiente della scolarità aumenta considerevolmente, passando da 0,0546 nel caso di stimatori OLS a 0,0705 nel caso di stimatori IV (tav. 6, colonne 1 e 3).

**VALORI MEDI E SCARTI QUADRATICI MEDI
DELLE PRINCIPALI VARIABILI, 1993**

Variabile	Media	S.Q.M.
LW (2)	6,491	0,372
ET (3)	43,382	8,362
ED (4)	10,183	3,849
SI (5)	0,349	0,477
ETCG (6)	40,154	8,594
EDCG (7)	9,517	4,003
EDP (8)	5,128	3,925
EDM (9)	4,386	3,343

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie.

(1) Capifamiglia dipendenti maschi di età inferiore a 60 anni che hanno lavorato a tempo pieno per l'intero anno. Numero di osservazioni: 2.041. - (2) Logaritmo del rapporto tra le retribuzioni annuali nette e il numero di ore mediamente lavorate alla settimana. - (3) Età in anni. - (4) Istruzione (numero di anni per conseguire il titolo di studio). - (5) Dummy Sud-Isole. - (6) Età del coniuge. - (7) Istruzione del coniuge. - (8) Istruzione del padre. - (9) Istruzione della madre.

Anche queste stime presentano alcuni problemi: per identificare il coefficiente della scolarità sono necessarie alcune restrizioni di esclusione, ovvero le variabili di background familiare non devono influenzare i redditi direttamente. Sarebbe certo difficile sostenere che i datori di lavoro sono disposti a pagare alcuni dipendenti più di altri solo per il fatto che i primi provengono da famiglie, ad esempio, più ricche. Tuttavia si può supporre che i figli delle persone più abbienti siano avvantaggiati, nella ricerca del lavoro, dalla rete di conoscenze della famiglia, riuscendo così a ottenere lavori a più elevato livello retributivo. Questo sarebbe un effetto diretto del background familiare sulle retribuzioni, che si aggiunge a quello, indiretto, attraverso il livello di istruzione (Atkinson, 1983, p. 118). Simili meccanismi non sembrano estranei al mercato

del lavoro italiano, dove le reti informali sono il canale prevalente tra le modalità di ottenimento del lavoro. Casavola e Sestito (1995), per esempio, evidenziano che oltre il 28 per cento degli occupati e delle persone in cerca di occupazione nell'estate del 1992 aveva ottenuto il (più recente) lavoro attraverso segnalazioni di amici, parenti e conoscenti a possibili datori di lavoro.

Tav. 6

**PARAMETRI DELLA FUNZIONE GENERATRICE DEL REDDITO (1)
(CAMPIONE C)**

Parametri (2)	OLS		IV	
Istruzione	5,46	5,03	7,05	6,08
(ED)	(0,172)	(0,190)	(0,364)	(0,321)
Età	2,36	2,59	2,01	2,30
(ET)	(0,747)	(0,744)	(0,766)	(0,753)
Età (quadrato)	-0,015	-0,017	-0,010	-0,013
(ET2)	(0,009)	(0,009)	(0,009)	(0,009)
Sud-Isole	-3,35	-2,99	-3,92	-3,46
(SI)	(1,384)	(1,378)	(1,418)	(1,393)
Istruzione della madre	-	1,13	-	0,60
(EDM)	-	(0,220)	-	(0,256)
ADJ-R ²	0,36	0,37	0,20	0,29
Endogene	-	-	ED	ED
Strumenti (3)	-	-	EDP	EDP
			EDM	EDCG
				ETCG

Fonte: Banca d'Italia, Indagine sui bilanci delle famiglie nel 1993.

(1) Capifamiglia maschi che hanno lavorato tutto l'anno 1993 a tempo pieno come lavoratori dipendenti. Variabile dipendente: log del rapporto tra le retribuzioni nette e le ore lavorate. Oltre alle variabili esplicative riportate nella tavola, nel modello è inclusa l'intercetta. Numero di osservazioni: 2.041. - (2) Le stime dei parametri e gli errori standard (tra parentesi) sono moltiplicati per 100. - (3) Oltre alle esogene del modello. Per la descrizione degli strumenti si veda la tavola 5.

Secondariamente, i genitori più ricchi potrebbero investire maggiormente nell'istruzione "non scolastica" dei loro figli, i quali mostrerebbero un più elevato livello di abilità quando entrano nel mercato del lavoro (Atkinson,

1983, p. 121). Un'istruzione (scolastica) di migliore qualità non sarebbe così l'unico mezzo con il quale i genitori trasmettono ai figli la capacità di "far soldi". Anche questa argomentazione appare potenzialmente rilevante nel caso italiano, date le carenze del sistema educativo pubblico.

In sintesi, vi sono ragionevoli motivi per ritenere che le variabili di background familiare entrino direttamente - oltre che indirettamente, attraverso la scolarità - nell'equazione delle retribuzioni. In questo caso, per risolvere il problema dell'endogenità della scolarità, è però necessario disporre di ulteriori variabili strumentali (oltre alle variabili di background familiare).

Un esperimento in tale direzione è stato condotto utilizzando come strumento aggiuntivo il livello d'istruzione del coniuge del capofamiglia. La specification search conduce a un modello in cui le retribuzioni sono funzione della scolarità, dell'età (termine lineare e quadratico) e del livello d'istruzione della madre del capofamiglia. Ciò evidenzia che, almeno nella precedente generazione, il livello d'istruzione della madre sembra essere più discriminante di quello del padre per identificare lo status della famiglia di origine. È peraltro possibile che venga colto il ruolo giocato dalla madre nell'educazione dei figli.

Utilizzando come strumenti (oltre alle esogene del modello) il livello d'istruzione del padre e quello del coniuge (queste due variabili, singolarmente considerate, superano il test per le restrizioni di sovraidentificazione), il background familiare risulta significativo e ha il segno atteso. L'effetto diretto sulle retribuzioni è limitato, ma non trascurabile: a ogni anno in più di istruzione della madre è associato un incremento del reddito del figlio pari allo 0,6 per cento.

Anche in questo caso l'istruzione risulta endogena; quanto alla stima del rendimento della scolarità l'esperimento, che probabilmente è il più convincente tra quelli effettuati, produce un significativo incremento del coefficiente della scolarità rispetto a quello ottenuto con lo stimatore OLS (da 0,503 a 0,608; tav. 6, colonne 2 e 4), evidenziando un effetto più marcato del livello di istruzione sulle retribuzioni¹³.

Questi risultati richiedono alcune precisazioni; in termini di effetti diretti, le retribuzioni dipendono più dal livello di istruzione che dalle caratteristiche della famiglia di origine; ma occorre ricordare che queste influenzano anche i livelli di scolarità¹⁴. Circa il 24 per cento della varianza riscontrata negli anni di istruzione dei capifamiglia intervistati è spiegata solo dal titolo di studio dei genitori. Pertanto i risultati evidenziano che la trasmissione della "capacità di guadagnare" dai genitori ai figli avviene prevalentemente per via indiretta, attraverso la scolarità¹⁵.

13. Risultati simili si ottengono utilizzando le retribuzioni pro capite in luogo di quelle orarie. In questo caso, utilizzando come strumenti i livelli di istruzione dei genitori, il coefficiente della scolarità incrementa dal 4,73 per cento al 6,60 per cento. Va peraltro segnalato che prendendo in considerazione le retribuzioni pro capite il coefficiente della dummy Sud-Isole diminuisce sensibilmente rispetto a quello stimato per le retribuzioni orarie (-7,7 contro -3,4 per cento). Ciò deriva dal fatto che le ore lavorate per dipendente nel Nord-Centro sono mediamente superiori a quelle del Sud-Isole.

14. Per una rassegna della letteratura empirica sui legami tra i redditi di individui appartenenti a differenti generazioni, cfr. Becker (1993, pp. 263 ss.).

15. A risultati simili, circa il ruolo dell'istruzione nei processi di mobilità sociale in Italia, perviene Schizzerotto (1988, pp. 110-111), dove si legge: "Il destino sociale dei nostri intervistati sembra essere regolato sia dalla classe di provenienza sia dal titolo di studio acquisito ... Con maggiore precisione si può

5. Considerazioni conclusive

In questo lavoro si sono esaminati alcuni problemi che caratterizzano le stime del rendimento dell'istruzione nelle funzioni generatrici del reddito e si è valutato se, e in che misura, le stime OLS del coefficiente della scolarità possono considerarsi corrette.

Nel corso del lavoro si sono ricordati i limiti delle sperimentazioni effettuate. Per quanto riguarda i dati utilizzati, in due dei tre campioni utilizzati, segnatamente il campione panel 1989-1991 (Campione A) e il campione dei figli rilevati nel 1991 (Campione B), sono presenti problemi di selezione. Il fatto che le stime OLS condotte su questi sottoinsiemi del campione siano non significativamente diverse da quelle ottenute operando sul campione complessivo lascia però ritenere che gli effetti determinati dalla selezione non siano tali da inficiare i risultati degli esperimenti. Peraltro il fatto che il campione dei capifamiglia (Campione C), che non è caratterizzato da problemi di selezione, fornisca risultati simili rafforza questa conclusione.

Il lavoro ha messo in evidenza alcuni problemi che riguardano gli stimatori utilizzati, in particolare nella scel-

dire che il grado di scolarità rappresenta una variabile decisiva per entrare o permanere nelle classi ... che traggono dal possesso di risorse culturali la fonte principale della loro situazione di superiorità sociale. L'accesso a queste posizioni è, di fatto, precluso a chi non abbia un titolo di studio sufficientemente elevato. Ed è soprattutto per questo motivo che le classi e i gruppi culturalmente e socialmente privilegiati cercano di riservare ai loro figli chances elevate di ottenere credenziali educative vantaggiose". Per un'analisi di tipo path delle relazioni tra family background, istruzione e retribuzioni, cfr. Antonelli (1987).

ta degli strumenti¹⁶: nelle stime GMM l'esperienza risulta endogena, mentre l'utilizzo dell'età come variabile strumentale è causa di una elevata variabilità dei risultati; nelle stime cross-section le variabili di background familiare potrebbero entrare direttamente nell'equazione delle retribuzioni oltre che essere correlate con l'abilità. Riguardo al primo punto, i risultati ottenuti nel paragrafo 4.1 - utilizzando come ulteriore variabile strumentale il livello d'istruzione del coniuge - evidenziano la significatività dell'effetto diretto del background familiare sulle retribuzioni. Riguardo al secondo aspetto, non si è in grado di fornire alcuna indicazione, non disponendo di variabili che possano essere utilizzate come strumenti per il background familiare. Una ulteriore nota di cautela riguarda l'utilizzo del livello di istruzione del coniuge come variabile strumentale. Una più elevata scolarità di quest'ultimo potrebbe essere correlata con maggiori aspirazioni di carriera e con maggiori redditi; se, come conseguenza, il capofamiglia mostrasse un minore impegno lavorativo (dovendo dedicare più cura ad altri aspetti della vita familiare o a causa del maggiore reddito familiare, che può ridurre l'incentivo a guadagnare) o fosse vincolato nelle scelte di carriera (per esempio, in termini di mobilità sul territorio), vi sarebbe una correlazione negativa tra l'istruzione del coniuge e la componente non osservabile nell'equazione delle retribuzioni del capofamiglia. La scolarità del coniuge non sarebbe pertanto uno strumento appropriato.

Per quanto riguarda gli errori di misurazione, risulta difficile stabilire il ruolo di questa componente nella stima del rendimento dell'istruzione. Indubbiamente, nell'indagine sui bilanci delle famiglie i dati presentano alcuni problemi

16. Va peraltro considerata la fragilità dei metodi che utilizzano variabili strumentali (per esempio cfr. Imbens e Rubin, 1994).

di qualità e gli errori di misurazione non sono trascurabili, soprattutto per certe variabili (esperienza lavorativa e tenure). Nel caso della scolarità, inoltre, non si dispone di informazioni potenzialmente rilevanti per valutare la qualità dell'istruzione, quali il tipo di diploma o di laurea, la votazione conseguita; infine, la partecipazione a certi tipi di corsi (ad esempio scuole di formazione professionale, corsi parauniversitari) non è rilevata. I risultati mostrati, ottenuti su campioni dai quali sono stati eliminati tutti i records con incoerenze superiori a una certa soglia, potrebbero risentire dei residui problemi di qualità dei dati anche se la convergenza dei risultati delle diverse sperimentazioni fa ritenere poco rilevante questo fattore.

Tutti gli esperimenti, sia quello condotto sui dati panel relativi al periodo 1989-1991 (Campione A) che, sfruttando la caratteristica di invarianza temporale della scolarità (almeno nel campione considerato), utilizza alcune variabili time-varying come strumenti, sia quelli che fanno ricorso a metodi di stima che utilizzano variabili di background familiare come strumenti per la scolarità (Campioni B e C), portano a risultati qualitativamente simili: il rendimento dell'istruzione appare significativamente sottostimato dal metodo OLS, indicando che le componenti individuali non osservate risultano negativamente correlate con il livello di istruzione. È possibile che ciò si riscontri a causa di un effetto di sostituzione tra l'abilità e l'ammontare ottimo di anni di studio; per i più abili, il costo dell'istruzione potrebbe essere più elevato che per i meno abili, in termini di redditi non percepiti durante il corso di studi. È forse più plausibile però che gli individui tengano conto, nel massimizzare la propria utilità e quindi nel decidere quanto studiare, anche dei benefici non monetari associati alle posizioni lavorative ottenibili con differenti gradi e tipi di istruzione. In altri termini, si potrebbe studiare di più non tanto, o non solo, per guadagnare di

più, quanto per svolgere un lavoro professionalmente più interessante o di maggiore prestigio sociale; o per accedere a posti di lavoro più sicuri, rinunciando a una maggiore retribuzione in cambio di un minor rischio di disoccupazione.

APPENDICE

Il panel della Banca d'Italia: alcuni problemi di qualità dei dati

Nell'indagine della Banca d'Italia lo schema di campionamento di tipo panel è stato introdotto per la prima volta nel 1989; circa il 15 per cento del campione (1.208 famiglie) fu ottenuto reintervistando famiglie già intervistate nel 1987. Per tenere conto dei cambiamenti demografici nella composizione della famiglia (dovuti a nascite, decessi, matrimoni, divorzi, ecc.) ai capifamiglia fu richiesto di indicare in un'apposita sezione del questionario sesso, età e rapporto di parentela con il capofamiglia di quei componenti che si erano aggiunti alla famiglia, o l'avevano lasciata, nel periodo 1988-1989.

Due principali problemi sorsero. In primo luogo, il tasso di risposta risultò particolarmente basso per le famiglie panel, e correlato con le caratteristiche familiari (cfr. Cannari e D'Alessio, 1992). Secondariamente, errori di misurazione e imprecisioni nella sezione riguardante la variazione di composizione della famiglia ostacolarono la possibilità di collegare i records individuali riferiti ai due anni 1987 e 1989.

Nell'indagine sul 1991 furono contattate solo le famiglie che nel 1989 avevano mostrato la loro disponibilità a essere reintervistate due anni dopo. Ovviamente, questa procedura riduce il costo dell'intervista, ma non rimuove le distorsioni che possono essere determinate dal comportamento di autoselezione dei rispondenti. Il compito di collegare i records individuali di differenti indagini fu assegnato agli intervistatori. Questi avevano a disposizione, al momento dell'intervista relativa al 1991, le informazioni sulla composizione della famiglia e sulle caratteristiche socio-demo-

grafiche dei componenti rilevate con l'indagine sul 1989. In questo modo essi potevano comparare le informazioni relative ai due differenti anni d'indagine e chiedere spiegazioni direttamente alle famiglie in caso di incoerenze nei dati o di variazioni nella composizione familiare.

I problemi di qualità non sono stati tuttavia completamente rimossi. Nella sezione panel sul 1991 si sono riscontrati diversi records incoerenti con le informazioni riportate nel questionario del 1989. I motivi di queste incoerenze sono differenti: gli errori sembrano derivare sia dalle procedure di codifica e data entry, sia dalla regola che permette agli intervistatori di ottenere dal capofamiglia le informazioni sui componenti assenti al momento dell'intervista. Non si può peraltro escludere che alcuni problemi di qualità dei dati siano dovuti agli stessi intervistati che possono talvolta fornire risposte incoerenti. In alcuni casi ciò può dipendere da atteggiamenti di reticenza su particolari aspetti; ad esempio, si può ipotizzare che un intervistato con un basso titolo di studio possa, in particolari circostanze, dichiarare un titolo di studio più elevato di quello effettivamente posseduto; successive interviste, svolte in differenti contesti, possono far emergere l'incoerenza. In altri casi è possibile che la risposta a un quesito del questionario richieda un'elaborazione da parte dell'intervistato; in presenza di situazioni più ambigue di quelle previste nel questionario il soggetto può fornire, in diverse occasioni, risposte involontariamente contrastanti. Nell'ambito della sezione di questionario utilizzata in questo lavoro, si pensi ai problemi di risposta che un quesito del tipo "Da quanti anni lavora con l'attuale datore di lavoro?" può generare nei soggetti che nella fase iniziale della loro attività hanno sperimentato periodi di lavoro non precisamente definiti sotto il profilo contrattuale o previdenziale.

Non va infine trascurato il fatto che la dimensione del questionario è ragguardevole; lo sforzo richiesto ai rispondenti è tutt'altro che trascurabile e l'attenzione tende a decrescere nelle sezioni finali del questionario; d'altra parte, la presenza di molte domande rende difficile per l'intervistatore effettuare controlli di coerenza nel corso dell'intervista.

Riferimenti bibliografici

- Amemiya, T. e T. E. MaCurdy (1986), Instrumental-Variable Estimation of an Error-Components Model, in "Econometrica", vol. 54, n.4, pp. 869-81.
- Angrist, J. D. e A. B. Krueger (1991), Does Compulsory School Attendance Affect Schooling and Earnings?, in "Quarterly Journal of Economics", vol. 106, n. 4, pp. 980-1014.
- _____ (1992), Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery, NBER Working Paper, n. 4067.
- _____ e W. K. Newey (1990), Over-Identification Tests in Earnings Functions with Fixed Effects, Harvard Institute of Economic Research, Discussion Paper, n. 1529.
- Antonelli, G. (1987), Human Resources and Labour Incomes. Demand for Education, Supply of Labour and Comparison between the Private and Public Sector, in "Labour", vol. 1, n. 2, pp. 153-90.
- Arellano, M. e O. Bover (1990), Another Look at the Instrumental-Variable Estimation of Error-Components Models, Centre for Economic Performance, Discussion Paper, n. 7.
- Ashenfelter, O. e A. Krueger (1992), Estimates of the Economic Return to Schooling from a New Sample of Twins, NBER Working Paper, n. 4143.
- Atkinson, A. B. (1983), The Economics of Inequality, Oxford, Clarendon Press.
- Becker, G. S. (1964), Human Capital, New York, NBER.
- _____ (1993), A Treatise on the Family, Cambridge MA, Harvard University Press.
- Birindelli, L., L. Bisaglia, G. Brunello e D. Rizzi (1994), La distribuzione personale e familiare delle risorse, in Secondo rapporto CNEL sulla distribuzione e redistribuzione del reddito in Italia, Roma, CNEL.
- Blackburn, M. e D. Neumark (1993), Are OLS Estimates of the Return to Schooling Biased Downward? Another Look, NBER Working Paper, n. 4259.
- Brandolini, A. e L. Cannari (1994), The Bank of Italy's Survey of Household Income and Wealth, in A.

- Ando, L. Guiso e I. Visco, Saving and the Accumulation of Wealth, Cambridge, Cambridge University Press.
- Brunello, G. e D. Rizzi (1993), I differenziali retributivi nei settori pubblico e privato in Italia: un'analisi cross-section, in "Politica Economica", n. 3, pp. 339-66.
- Cannari, L. e G. D'Alessio (1992), Mancate interviste e distorsione degli stimatori, Banca d'Italia, Temi di discussione, n. 172.
- _____, e A. Gavosto (1994), L'indagine della Banca d'Italia sui bilanci delle famiglie: una descrizione dei dati sul mercato del lavoro, in "Economia e Lavoro", n. 1, pp. 63-79.
- _____, G. Pellegrini e P. Sestito (1989), Età, istruzione, retribuzioni, in Banca d'Italia, Contributi all'analisi economica del Servizio Studi, n. 5, pp. 45-79.
- Casavola, P. e P. Sestito (1995), Come si cerca e come si ottiene lavoro? Un quadro sintetico sull'Italia e alcune implicazioni macroeconomiche, in A. Amendola (a cura di), Disoccupazione e mercato del lavoro, Napoli, Esi.
- Commissione Carniti (1988), I salari in Italia negli anni ottanta, Venezia, Marsilio.
- Dagum, C. e A. Lemmi (1988), Struttura causale e specificazione della funzione generatrice del reddito nell'esperienza italiana, in "Studi e Informazioni", vol. 11, n. 2, pp. 45-55.
- Griliches, Z. (1977), Estimating the Returns to Schooling: Some Econometric Problems, in "Econometrica", vol. 45, n.1, pp. 1-22.
- _____, B. Hall e J. A. Hausman (1978), Missing Data and Self-Selection in Large Panels, in "Annales de l'INSEE", vol. 30-31, pp. 137-76.
- Hausman, J. A. e W. E. Taylor (1981), Panel Data and Unobservable Individual Effects, in "Econometrica", vol. 49, n.6, pp. 1377-98.
- Heckman, J. J. e S. Polacheck (1974), Empirical Evidence on the Functional Form of the Earning-Schooling Relationship, in "Journal of the American Statistical Association", n. 346, pp. 350-54.

- Imbens, G. W. e D. B. Rubin (1994), On the Fragility of Instrumental Variables Estimators, Harvard Institute of Economic Research, Discussion Paper, n. 1675.
- Lucifora, C. (1988), Struttura di mercato, vincoli finanziari e determinazione del salario: un'analisi empirica dei differenziali interaziendali nell'industria italiana, in ASAP, Rapporto sui salari, Milano, Angeli.
- _____ e P. Sestito (1993), Determinazione del salario in Italia: una rassegna della letteratura empirica, Università Cattolica del Sacro Cuore, Istituto di Economia dell'Impresa e del Lavoro, Discussion Paper, n. 5.
- Mincer, J. (1958), Investment in Human Capital and Personal Income Distribution, in "Journal of Political Economy", vol. 66, n.4, pp. 281-302.
- Ottersen, E. K., E. Mellander, E. M. Meyerson e J. Nilson (1994), Pitfall in the Measurement of the Return to Education: An Assessment Using Swedish Data, Industriens Utredningsinstitut, Working Paper, n. 414.
- Schizzerotto, A. (1988), Il ruolo dell'istruzione nei processi di mobilità sociale, in "Polis", n. 1, pp. 83-123.
- Willis, R. S. (1986), Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions, in O. Ashenfelter e R. Layard, Handbook of Labour Economics, Amsterdam, North-Holland.

ELENCO DEI PIÙ RECENTI "TEMI DI DISCUSSIONE" (*)

- n. 229 — *About the Level of Daylight Credit, Speed of Settlement and Reserves in Electronic Payment Systems*, di P. ANGELINI (agosto 1994).
- n. 230 — *Asymmetries and Nonlinearities in Economic Activity*, di F. FORNARI e A. MELE (agosto 1994).
- n. 231 — *L'attività cross-border delle banche italiane: una verifica empirica*, di C. BENTIVOGLI e A. GENERALE (settembre 1994).
- n. 232 — *La curva dei rendimenti dei BOT come misura dei tassi futuri attesi*, di G. GRANDE (settembre 1994).
- n. 233 — *Assicurazione dei depositi, coefficienti patrimoniali e copertura dei rischi bancari*, di F. DRUDI e R. TEDESCHI (settembre 1994).
- n. 234 — *Predicting Consumption of Italian Households by means of Leading Indicators*, di G. PARIGI e G. SCHLITZER (settembre 1994).
- n. 235 — *L'apertura di sportelli bancari dopo la liberalizzazione: andamento e determinanti*, di R. DE BONIS, F. FARABULLINI e F. FORNARI (novembre 1994).
- n. 236 — *Alternative Estimators of the Cox, Ingersoll and Ross Model of the Term Structure of Interest Rates: A Monte Carlo Comparison*, di C. BIANCHI, R. CESARI e L. PANATTONI (novembre 1994).
- n. 237 — *Capital Structure Decisions of a Public Company*, di O. HART (dicembre 1994).
- n. 238 — *La proprietà cooperativa: teoria, storia e il caso delle banche popolari*, di R. DE BONIS, B. MANZONE e S. TRENTO (dicembre 1994).
- n. 239 — *Controllo di gruppo e incentivi degli azionisti di minoranza: una verifica empirica*, di M. BIANCO, P. CASAVOLA e A. FERRANDO (dicembre 1994).
- n. 240 — *I meccanismi di funzionamento dei circuiti interni dei capitali: un'indagine empirica del caso italiano*, di L. BUZZACCHI e M. PAGNINI (dicembre 1994).
- n. 241 — *Diffusione della proprietà delle imprese fra le famiglie italiane e trasferimento intergenerazionale: alcune evidenze*, di G. D'ALESSIO (dicembre 1994).
- n. 242 — *Origini e natura speciale dell'attività di banca d'affari in Italia*, di M. DE CECCO e G. FERRI (dicembre 1994).
- n. 243 — *Evoluzione degli assetti di controllo: gli investitori istituzionali*, di M. BIANCO e P. E. SIGNORINI (dicembre 1994).
- n. 244 — *Linee di riforma dell'ordinamento societario nella prospettiva di un nuovo ruolo degli investitori istituzionali*, di D. PREITE e M. MAGNANI (dicembre 1994).
- n. 245 — *Efficiency of Bankruptcy Procedures*, di F. CORNELLI e L. FELLI (dicembre 1994).
- n. 246 — *Change of Ownership: Incentives and Rules*, di L. ZINGALES (dicembre 1994).
- n. 247 — *Circolazione della ricchezza e informazioni significative: il problema delle offerte pubbliche di acquisto*, di G. CARRIERO e V. GIGLIO (dicembre 1994).
- n. 248 — *Innovazioni strutturali nel mercato azionario: gli effetti della contrattazione continua*, di C. IMPENNA, P. MAGGIO e F. PANETTA (gennaio 1995).
- n. 249 — *Computable General Equilibrium Models as Tools for Policy Analysis in Developing Countries: Some Basic Principles and an Empirical Application*, di T. BUEHRER e F. DI MAURO (febbraio 1995).
- n. 250 — *The 1992-93 EMS Crisis: Assessing the Macroeconomic Costs*, di L. BINI SMAGHI e O. TRISTANI (febbraio 1995).
- n. 251 — *Sign- and Volatility-Switching ARCH Models: Theory and Applications to International Stock Markets*, di F. FORNARI e A. MELE (febbraio 1995).
- n. 252 — *The Effect of Liquidity Constraints on Consumption and Labor Supply: Evidence from Italian Households*, di S. NICOLETTI-ALTIMARI e M. D. THOMSON (febbraio 1995).

(*) I "Temi" possono essere richiesti a:

Banca d'Italia - Servizio Studi - Divisione Biblioteca e pubblicazioni - Via Nazionale, 91 - 00184 Roma.