



UNIVERSITÀ DEGLI STUDI DI PADOVA

Facoltà di Scienze Statistiche

Corso di studi di Scienze Statistiche Economiche Finanziarie e Aziendali

TESI DI LAUREA

Le Risposte del Consumo delle Famiglie Italiane a
Variazioni Inattese del Reddito a Carattere
Permanente o Transitorio

Relatore: *prof. Guglielmo Weber*

Correlatore: *prof. Alessandro Buccioli*

Laureando: *Brigato Lorenzo*

Anno Accademico 2010/2011

Indice

Introduzione	4
Capitolo 1: La decomposizione degli shock di reddito	6
1.a: Il metodo quasi-sperimentale.....	7
1.b: Decomposizione degli shock di reddito	11
Capitolo 2: L'indagine sul reddito e la ricchezza delle famiglie della Banca d'Italia	28
2.a: Il disegno campionario.....	29
2.b: Il Raccolta e trattamento dei dati	30
Capitolo 3: Il modello economico e l'identificazione dei parametri	36
3.a: Il processo del reddito e del consumo	36
3.b: Identificazione del modello.....	43
3.b.1: Identificazione del modello con errori di misura del consumo	48
3.b.2: Identificazione del modello con errori di misura del reddito.....	49
3.b.3: Identificazione del modello con MA(1) come processo dello shock di reddito transitorio	51
3.b.4: Non stazionarietà dei coefficienti di assicurazione	54
3.c: Modelli di regressione per dati di panel.....	55
3.c.1: Modello ad effetti fissi	56
3.c.2: Modello ad effetti casuali.....	61
3.c.3: Il test di Hausman	63

3.d: La stima delle varianze	66
3.e: Lo stimatore a distanze minime	68
Capitolo 4: I risultati dell'identificazione	72
4.a: Analisi descrittive delle varianze.....	72
4.b: Trattamento dei dati	75
4.c: Generazione dei residui	77
4.d: Stima delle varianze	83
4.e: Identificazione del modello	87
4.e.1: L'effetto dell'istruzione e dell'anzianità	91
4.e.2: Un confronto internazionale	95
Conclusione e sviluppi futuri	100
Appendice: L'effetto dell'arrotondamento dei dati sulle stime ...	103
A.a: Imputazione dei dati.....	106
A.b: Diversa identificazione del modello	110
Bibliografia	115

Introduzione

Sapere come e quanto le famiglie italiane reagiscono a variazioni inattese del reddito, come tale reazione varia in base alle caratteristiche della famiglia e come le famiglie modificano i loro consumi in base al ciclo economico è importante al fine di prevedere il comportamento dei consumatori e per determinare gli effetti di riforme che influenzano il reddito, come politiche fiscali o del lavoro.

Nella sua teoria del reddito permanente, Friedman (1957) teorizza l'esistenza di due tipologie di shock del reddito, suddivise in base alla persistenza, e che hanno un diverso effetto sul consumo: permanenti, che influenzano tutto il reddito futuro della persona e che si trasmettono interamente al consumo, e temporanei, che influenzano il reddito, invece, per solo un breve periodo e che non influenzano il consumo.

Negli anni successivi molti ricercatori hanno cercato di applicare tale teoria ai dati, ottenendo risultati discordanti. Bodkin (1959), ad esempio, trova che gli shock transitori vengono trasmessi in parte al reddito. Anche Hall (1982) ottiene un risultato simile, ma entrambi questi lavori prevedono piena trasmissione degli shock permanenti del reddito al consumo. Blundell,

Pistaferri & Preston (2008) evitano di imporre questa condizione e verificano che, per quanto riguarda gli Stati Uniti, gli shock temporanei non si trasmettono al consumo, mentre quelli permanenti sono parzialmente assicurati.

L'obiettivo di questa tesi è valutare come le famiglie italiane reagiscono agli shock di reddito calcolando la propensione marginale del consumo rispetto a variazioni del reddito. Seguendo la teoria di Friedman e l'esempio di Blundell et al., non imporre vincoli alla capacità delle famiglie di proteggersi contro variazioni del reddito (permanenti o transitorie); partendo da valori di reddito e consumo depurati dagli effetti di caratteristiche demografiche misurabili, otterremo una stima dei coefficienti di assicurazione del consumo, che indicano la parte dello shock di reddito (permanente o transitorio) che viene trasmessa al consumo.

Si sono utilizzati dati di panel biennali sul reddito, sul consumo di beni non durevoli e sulle caratteristiche demografiche delle famiglie relativi al periodo 1998-2008 provenienti dall'Indagine sui bilanci delle famiglie italiane curata dalla Banca d'Italia.

La tesi è organizzata nel modo seguente: Il primo capitolo offre una panoramica dell'evoluzione delle tecniche di scomposizione degli shock di reddito; il capitolo 2 descrive la SHIW, l'indagine dalla quale si estrarranno i dati utilizzati in questa tesi; il capitolo 3 descriverà, oltre al modello economico utilizzato e al metodo di identificazione dei parametri, anche la teoria sottostante le procedure utilizzate per ottenere questi risultati; il capitolo 4 espone i risultati ottenuti in questa tesi; il capitolo 5 conclude.

Capitolo 1: La decomposizione degli shock di reddito

Questo capitolo cercherà di descrivere l'evoluzione dei principali metodi utilizzati per analizzare l'effetto di shock inattesi del reddito sul consumo, rappresentando i vantaggi e le problematiche ad essi legati.

Nel tempo, a partire dalla fine degli anni '50, sono stati sviluppati due diversi approcci all'analisi dell'effetto dei predetti shock: il primo, detto "quasi-sperimentale"¹, cerca di identificare episodi nei quali si siano verificate variazioni inattese del reddito e confronta soggetti economici sottoposti a tale variazione con soggetti (aventi caratteristiche il più possibile simili al primo gruppo) che non siano state influenzati da tale evento. Tale sistema, nonostante gli vengano mosse numerose critiche in merito alla non distinzione degli shock, merita un approfondimento (vedi *sub a.*), in quanto fu il primo utilizzato per affrontare la problematica di legare le dinamiche di reddito e consumo anche con il **supporto di dati** e non solo in maniera teorica.

¹ "Quasi-sperimentale" in quanto il ricercatore non provoca l'evento come in un qualsiasi studio sperimentale, ma si limita ad osservare gli effetti sui soggetti in esame di un evento che è già accaduto.

Inoltre, questo approccio non è, per così dire, chiuso in sé stesso, atteso che può essere utilizzato integrato ad altri metodi di stima degli effetti di vari eventi macroeconomici, allo scopo di valutare come tali eventi influenzino il consumatore².

Il secondo metodo, sviluppatosi per la prima volta negli anni '80 e che sarà, come già anticipato nell'introduzione, l'oggetto della presente tesi, partendo dall'assunto che reddito e consumo seguano un dato processo, utilizza **dati di panel** (o longitudinali) al fine di imporre restrizioni sulle covarianze, in modo da identificare i parametri che caratterizzano il legame tra reddito e consumo.

1.a. Il metodo quasi-sperimentale

Uno dei primi studi che si avvalsero del metodo "quasi sperimentale", termine introdotto da Campbell & Stanley nel 1963, fu quello di Bodkin (1959), che osservò il comportamento di alcuni veterani della seconda guerra mondiale, che avevano ricevuto dei pagamenti inattesi a titolo di indennizzo da parte del NSLI (*National Service Life Insurance*). Bodkin considerò questa circostanza uno shock inatteso e transitorio del reddito e stimò, a partire dall'Urban Consumption Survey pubblicata dall'U.S. Bureau of Labour

² Un esempio in tal senso è rappresentato dallo studio di Jappelli e Pistaferri "Financial Integration and Consumption Smoothing" (2010). Essi considerano l'introduzione dell'Euro in Italia una liberalizzazione finanziaria, che dovrebbe migliorare la capacità delle famiglie di assicurarsi contro shock di reddito. Essi stimano la propensione marginale del consumo rispetto a shock di reddito permanente e transitorio assumendo che tali propensioni assumano valori differenti prima e dopo il 1999, anno di introduzione dell'Euro in Italia. Essi trovano che l'introduzione dell'euro provoca un calo della propensione marginale del consumo rispetto a shock transitori del reddito, ma che tuttavia tale diminuzione non è statisticamente significativa.

Statistics, che la propensione marginale del consumo rispetto a questo shock fosse pari a 0.72, in contrasto con la teoria del reddito permanente di Friedman che, al contrario, avrebbe previsto una propensione marginale non significativa³.

Più recentemente, questa tipologia di studi si concentra - in maniera più funzionale, rispetto al passato, alla valutazione di politiche economiche - sull'effetto che la disoccupazione ha sui consumi e su come tali effetti possano venire attenuati da eventuali sussidi di disoccupazione. Il lavoro di Gruber (1997) rappresenta uno dei primi tentativi di valutare questi effetti; infatti, egli prende in considerazione lavoratori che risultavano occupati in un periodo e disoccupati in quello successivo; utilizzando i dati sul consumo di cibo e sul reddito (derivante da lavoro e da sussidi statali) relativi a questi lavoratori, egli regredisce:

$$\Delta c_t = \alpha + \beta_1 X_t + \beta_2 UI_t + \varepsilon_t$$

³ Tale risultato risulta però essere oggetto di alcune critiche. Kreinin, nel 1961, eseguì un test simile, usando dati sugli indennizzi di guerra pagati a veterani Israeliani (disponibili nell'Israel Saving Survey), trovando un valore della propensione marginale al consumo rispetto a shock transitory del reddito pari a 0,17, pari a un quarto rispetto a quelli trovati da Bodkin. Landsberger (1966) afferma che tali stime siano distorte dal fatto che i pagamenti ricevuti dagli americani siano molto inferiori - proporzionalmente al reddito - rispetto a quelli ricevuti dai soldati Israeliani, e che la propensione marginale al consumo vari in base all'entità dello shock. Tale ipotesi è tuttavia sconfessata da Bodkin (1966), che dividendo il campione da lui precedentemente utilizzato in due sottocampioni in base all'entità dell'indennizzo ricevuto, trova che la propensione marginale del consumo non varia significativamente tra i due gruppi. Lo stesso Friedman, nel 1960, afferma che il risultato di Bodkin sia distorto verso l'alto, poiché non tiene conto del fatto che il reddito permanente sia costante, e quello che sembra una risposta eccessiva del consumo rispetto a shock transitori del reddito è in realtà l'effetto del reddito permanente. Bird & Bodkin (1965), rieseguono le analisi di Bodkin del 1959 tenendo conto di questo fatto, e considerando solo i consumi di beni non durevoli (Bodkin, nel 1959, aveva incluso anche i beni durevoli nella misura del consumo) trovarono che le stime della propensione marginale del consumo non differivano significativamente da quelle previste dalla teoria del reddito permanente, confermando così la teoria di Friedman.

Dove Δc_t è la variazione di consumo di cibo, ricavato da dati PSID; X_t è una serie di caratteristiche esogene come età, sesso, razza, stato civile, educazione del capofamiglia e dimensione del nucleo familiare e UI_t è la parte di reddito che viene elargita dai sussidi di disoccupazione.

Si precisa che Gruber non ha utilizzato i sussidi realmente usufruiti, ma quelli *potenziali* (ovvero quelli che un soggetto con determinate caratteristiche avrebbe diritto a ottenere) per tre ragioni: l'endogeneità dei sussidi richiesti, essi infatti correlano con la durata della disoccupazione, che correla con la variazione del consumo; la gran quantità di errori rilevata nei dati sui sussidi effettivamente elargiti⁴; infine, l'interesse da parte del governo ad avere un risultato su grandezze sulla quale può intervenire (i sussidi potenziali) rispetto a grandezze sulla quale non ha alcun controllo (i sussidi effettivamente elargiti). Un valore di $\beta_2 > 0$ indica che i sussidi sono utili nel mitigare gli effetti che la disoccupazione provoca sul consumo.

Gruber conclude che i sussidi di disoccupazione hanno una grande utilità da questo punto di vista: più precisamente, un aumento del 10% del tasso di sostituzione (la percentuale di reddito da lavoro erogato come sussidio) aumenterebbe del 3,25% i consumi. In assenza di sussidi, il calo dei consumi risulterebbe di circa il 20%. Ciò con piena sostituzione, ovvero se il sussidio fosse pari al reddito da lavoro del consumatore, non ci sarebbe nessuna variazione del consumo.

Sebbene questo tipo di studi abbia il vantaggio di essere di facile comprensione, gli sono state mosse diverse critiche: in primo luogo viene

⁴ Quasi un quarto delle unità statistiche affermava di aver usufruito di sussidi sebbene dichiarasse di non aver avuto nemmeno una settimana di disoccupazione.

preso in esame solamente un tipo di shock per volta ed è molto difficile identificare la tipologia di shock provocata da un determinato evento, o trovarne uno che causi uno shock esattamente del tipo cercato (un raro caso è dato dagli effetti sul reddito delle condizioni atmosferiche in economie prettamente agricole, che generano shock inattesi di tipo transitorio⁵); spesso, viceversa, un singolo evento provoca shock di diverso tipo: negli studi sulla disoccupazione la persistenza dello shock sul reddito provocato da tale evento potrebbe dipendere dalla durata della disoccupazione stessa, dall'età del lavoratore e anche dalle cause che hanno determinato la perdita del posto di lavoro.

Peraltro, la risposta del consumo alla disoccupazione (così come alla malattia e alla disabilità) può essere distorta da alcuni fattori (Stephens, 2001): innanzitutto, la disoccupazione potrebbe non essere un evento improvviso (come nel caso di aziende in crisi da lungo tempo) o involontario. Inoltre, secondo la teoria economica per la quale l'agente economico tende a massimizzare l'utilità marginale e non direttamente il consumo, una risposta eccessiva di questa grandezza rispetto a uno shock transitorio (come quello causato dalla perdita del lavoro) non mette necessariamente in discussione la teoria di Friedman, in quanto la disoccupazione ha come conseguenza diretta un aumento del tempo libero a disposizione del consumatore e se, in via ipotetica, il consumo fosse un bene sostituto di questo tempo, l'effetto di sostituzione si andrebbe a sommare a quello provocato dal calo del reddito, generando così uno shock negativo del consumo di dimensioni maggiori di quelle previste.

⁵ Si vedano a questo proposito gli studi di Deaton & Paxson "Saving, Growth, and Aging in Taiwan" e il successivo "Growth, Demographic Structure, and National Saving in Taiwan"

In via meramente ipotetica, potrebbero essere presenti problemi *moral hazard*: si pensi al caso dei c.d. “falsi invalidi”, ovvero individui che accedono agli aiuti destinati ai disabili, pur non possedendo i requisiti necessari, e di come tale fenomeno possa distorcere i risultati di studi che trattano l’effetto delle pensioni di invalidità hanno sugli shock provocati dalla disabilità (nel 2009 in Italia la percentuale di controlli da parte dell’INPS che hanno portato alla sospensione della pensione di invalidità è del 17%⁶).

Infine, anche shock completamente inattesi potrebbero essere assicurati da forme di welfare che non possono essere osservate con precisione su larga scala (come le assicurazioni private), o dal sostegno (non misurabile) fornito da parenti e amici. Questo spiegherebbe una risposta a uno shock permanente non così rilevante come ci si aspetterebbe: come sopra enunciato, i sussidi hanno un forte effetto di attenuazione degli shock di reddito sul consumo e l’aiuto ricevuto da questo tipo di welfare diminuirebbe l’effetto di uno shock negativo del reddito.

1.b. Decomposizione degli shock di reddito

Il secondo metodo di analisi del legame tra consumo e variazioni inattese del reddito si basa sulla struttura delle varianze di queste due grandezze. Una volta fatte delle assunzioni sul processo del reddito e su quello del consumo, si usano delle restrizioni sulle covarianze di questi processi per stimare i parametri, che descrivono il comportamento di reddito e consumo.

⁶ Con punte del 29% in Basilicata. Fonte: INPS

Per identificare gli shock del reddito sono stati seguiti due differenti metodi: nel primo caso, si utilizzano dati di panel (o più spesso pseudo-panel) e si definiscono gli shock come scostamento delle realizzazioni del reddito rispetto alle attese, che vengono derivate a partire da caratteristiche osservabili dei consumatori; nel secondo caso ci si basa anche su aspettative soggettive quantitative. Quest'ultimo metodo, pur permettendo il distinguo tra assicurazione parziale e superiore informazione da parte del consumatore, ha suscitato alcune critiche a causa della reperibilità e della qualità dei dati sulle aspettative individuali e dalle assunzioni usate per ricavare tali aspettative⁷.

Ciò premesso, prima di procedere alla decomposizione degli shock di reddito e consumo, è necessario fare delle assunzioni sui processi di reddito e consumo. Tali assunzioni partono dal problema di massimizzazione dell'utilità attesa del consumo sotto un vincolo di bilancio intertemporale, orizzonte temporale finito e una condizione terminale sul risparmio. Si arriva così all'equazione di Eulero dell'utilità marginale del consumo:

$$u'(c_{t-1}) = \frac{(1+r)}{(1+\delta)} E_{t-1}[u'(c_t)]$$

Dove c_t è il consumo, r il tasso d'interesse, δ il tasso di sconto e E_{t-1} è l'aspettativa basata sulle informazioni disponibili al tempo $t-1$. Tale equazione indica che, in equilibrio, non si può aumentare l'utilità al margine del consumatore riallocando il consumo nel tempo.

⁷ Ad esempio, si assume la perfetta conoscenza da parte del consumatore dei meccanismi che regolano l'eleggibilità ai sussidi di disoccupazione.

Se il tasso d'interesse è pari al tasso di sconto, allora l'equazione diventa:

$$u'(c_{t-1}) = E_{t-1}[u'(c_t)]$$

Cioè una martinagala o *random walk*.

Una proprietà di questa equazione (notata per primo da Hall, 1978) è che l'utilità marginale varia solo se le attese non vengono realizzate e, quindi, le variazioni di tale utilità non sono prevedibili sulla base di informazioni pregresse. Di conseguenza, una variazione attesa del reddito non dovrebbe avere un impatto sull'utilità marginale del consumo, perché tale variazione è già stata incorporata nelle attese nel momento in cui il soggetto è venuto a sapere di tale evento. Invece, variazioni inattese del reddito variano l'utilità marginale in base alla natura e durata delle variazioni e alla struttura del mercato del credito.

I primi tentativi nel testare le implicazioni di un'utilità marginale che segue un processo di tipo *random walk* si basano sul caso particolare di utilità quadratica⁸. Questo caso è conosciuto come soluzione equivalente al caso di certezza. Sostituendo nell'equazione precedente si ha:

$$c_t = c_{t-1} + \varepsilon_t$$

Dove ε_t è l'effetto sul consumo di tutta la nuova informazione riguardante le fonti di incertezza del consumatore. La principale implicazione di questo risultato è che il livello del consumo, e non solo l'utilità marginale, è una martingala: il miglior previsore del consumo è il

⁸ La funzione di utilità quadratica è: $u(c_t) = ac_t + b/2 c_t^2$. Nel caso generico, nella quale $r \neq \delta$, l'equazione di Eulero diventa $c_{t+1} = \frac{a}{b} \frac{r-\delta}{1+r} + \frac{1+\delta}{1+r} c_t + \varepsilon_t$

consumo del periodo precedente. Inoltre, il consumo cambia solo se le attese non vengono rispettate. Considerando la condizione di un lavoratore, la principale fonte di incertezza è proprio il reddito da lavoro. Se tale reddito è l'unica forma di incertezza, si può scrivere l'equazione precedente come:

$$\Delta c_t = \alpha_t \sum_{i=0}^{T-t} \frac{(E_t - E_{t-1})y_{t+i}}{(1+r)^i}$$

Dove α_t è un fattore di annuitizzazione⁹ che collassa a $r/(1+r)$ nel caso di orizzonte temporale infinito; e $(E_t - E_{t-1})y_{t+i}$ rappresenta le modifiche delle attese sul reddito tra il periodo $t - 1$ e t .

L'equazione offre un'interpretazione strutturale dell'innovazione del consumo: il consumo non varia finchè le attese su tutte le fonti di incertezza (in questo caso il reddito da lavoro) non variano. L'impatto del reddito dipende anche dal fattore di annuitizzazione, che dipende dall'orizzonte temporale e dal tasso d'interesse. Tale equazione è utile perchè suggerisce che diverse assunzioni sul processo del reddito causano diverse risposte del consumo rispetto a shock di reddito.

Semplifichiamo le equazioni assumendo un orizzonte temporale infinito, e consideriamo diversi processi del reddito. Ad esempio, un processo spesso usato per descrivere serie macroeconomiche è un processo del tipo ARMA(1,1):

$$y_t = \rho y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

⁹ $\alpha_t = \frac{r}{1+r} \left(1 - \frac{1}{(1+r)^{T-t+1}}\right)^{-1}$

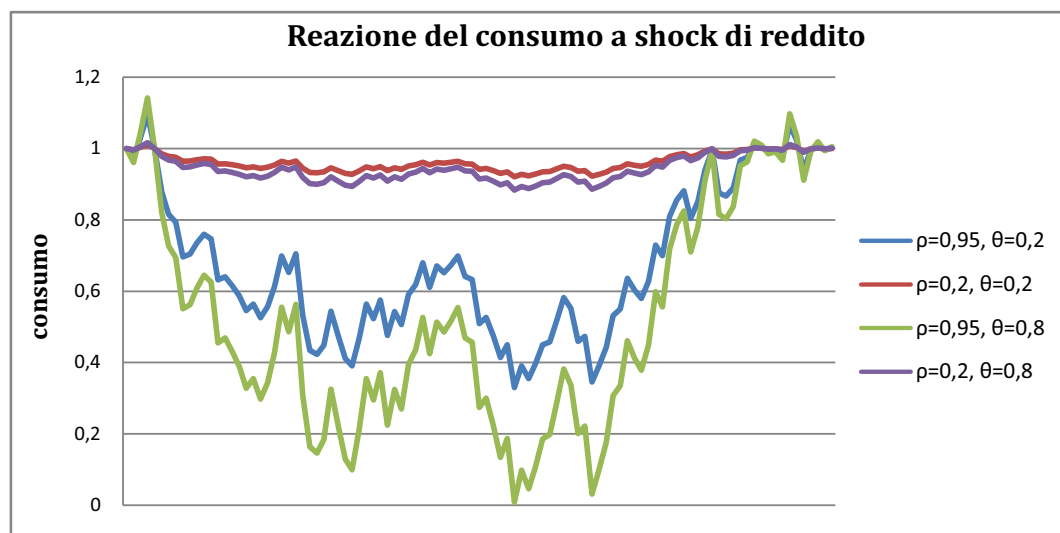
Cosicchè l'equazione della variazione del consumo si può riscrivere come (Jappelli & Pistaferri, 2010):

$$c_t = c_{t-1} + \frac{r}{1+r} \frac{1+r+\theta}{1+r-\rho} \varepsilon_t$$

La variazione del consumo dipende chiaramente dalla persistenza del processo del reddito; più il processo del reddito è persistente, più il consumo è volatile. Considerando l'esempio semplificato di un processo del reddito di tipo AR(1) (e quindi ponendo $\theta = 0$), nel caso che il reddito sia serialmente incorrelato, quindi assumendo $\rho = 0$, la propensione marginale del consumo rispetto a variazioni del reddito è pari a $\frac{r}{1+r}$, cioè solo la parte della revisione delle attese sul reddito relative all'anno appena trascorso.

Al contrario, nel caso che il reddito sia una martingala ($\rho = 1$), la propensione marginale del consumo rispetto a variazioni del reddito è pari a 1, in quanto tutte le variazioni del reddito sono permanenti. Tale diversità di reazione del consumo può essere notata nel grafico in basso, che descrive l'andamento di due processi del consumo. Entrambi i processi hanno un valore iniziale pari a 1, e sono soggetti alle stesse innovazioni (ottenute generando dei valori random da una distribuzione normale con standard error 0,1): le curve corrispondenti a processi più persistenti hanno un andamento molto meno stabile rispetto alle curve che descrivono l'andamento del consumo soggetto ad innovazioni scarsamente persistenti; inoltre, si può notare come l'effetto di shock transitori passati (il valore del coefficiente MA θ) sia scarsamente rilevante nel determinare l'andamento del consumo.

Tuttavia, tale definizione del processo del reddito incorre in una pesante limitazione: essa infatti prevede che il reddito dipenda dal passato solamente attraverso valori precedenti del reddito stesso, eliminando la possibilità dell'esistenza di più processi che influenzano le dinamiche di tale grandezza. Già Friedman (1957) riconobbe che questa possibilità era da prendere in considerazione, in quanto diversi tipi di avvenimenti provocavano diversi effetti sul consumo di chi ne era influenzato. A tal proposito, si pensi alla differenza degli effetti che causano eventi come una variazione degli orari di lavoro, una vincita alla lotteria, o la malattia (che provocano variazioni del reddito transitorie) rispetto agli effetti di promozioni o gravi invalidità (che causano variazioni permanenti del reddito).



Poiché con questo tipo di processo non si è in grado di analizzare l'evoluzione di questi tipi di shock, risulta essere non adatto per l'utilizzo nelle analisi di cui ai capitoli successivi.

Un modo per risolvere questa limitazione è considerare un processo del reddito che dipenda dal passato attraverso più fattori, in modo da mettere in evidenza che diversi processi di differente persistenza influenzano le dinamiche del reddito. Un processo che possiede tali caratteristiche è il seguente:

$$y_t = P_t + u_t$$

Dove P_t è la componente permanente del reddito, che segue un processo di tipo martingala o *random walk*, mentre ε_t rappresenta la componente transitoria e segue un processo *white noise*. Ora, l'equazione del consumo dipende da entrambi i tipi di shock: in particolare, poiché gli shock sono rispettivamente una martingala e un *white noise*, la propensione marginale del consumo rispetto a questi shock sarà pari a quelle del caso semplificato di un processo del reddito di tipo AR(1) descritto sopra, in quanto il consumo è supposto essere la somma di due processi autoregressivi, uno con ρ pari a 1, e l'altro con coefficiente nullo, indipendenti tra loro. Di conseguenza:

$$\Delta c_t = \varepsilon_t + \frac{r}{1+r} u_t$$

Dove ε_t è l'innovazione del processo della componente permanente del reddito.

Il consumo risponde interamente agli shock permanenti, mentre è quasi insensibile a quelli transitori.

Il primo lavoro che usa questo metodo per scomporre il reddito è di Hall e Mishkin (1982): essi assumono un processo del reddito del tipo

permanente-transitorio come quello descritto precedentemente, utilità quadratica e un processo della componente transitoria del reddito del tipo MA(K). Essi risolvono il problema di ottimo:

$$\max_c \left\{ E_t \left[-\frac{1}{2} \sum_{i=0}^{T-t} (1 + \delta)^{-i} (c_t^* - c_t)^2 \right] \right\}$$

sotto un vincolo di bilancio intertemporale

$$\sum_{i=0}^{T-t} (1 + r)^{-i} (y_{t+i} - c_{t+i}) + A_t = 0$$

dove c_t^* è il livello ottimo di consumo, δ è il tasso di sconto, r il tasso d'interesse e A_t sono gli asset. Si ottiene una condizione del primo ordine pari a:

$$\frac{E_t[(1 + \delta)^{-i} (c_{t+i}^* - c_{t+i})]}{c_t^* - c_t} = (1 + r)^{-i}$$

Semplificando ponendo $\delta = r$, e sostituendo questa equazione al $(1 + r)^{-i}$ presente all'interno del vincolo di budget, si ottiene:

$$\sum_{i=0}^{T-t} (1 + r)^{-i} (E_t[y_{t+i}] - c_{t+i}^* - c_t + c_t^*) + A_t = 0$$

Ponendo rispettivamente il capitale umano, il fattore di annuitizzazione e il consumo atteso futuro pari a:

$$U_t = \sum_{i=0}^{T-t} (1+r)^{-i} (E_t[y_{t+i}]); \quad \alpha_t = 1 / \sum_{i=0}^{T-t} (1+r)^{-i}; \quad C_t = \sum_{i=0}^{T-t} (1+r)^{-i} (c_t^*)$$

Si può riscrivere l'equazione di cui sopra in base al consumo:

$$c_t = c_t^* + \alpha_t (U_t + A_t - C_t)$$

Risultato simile a quello ottenuto in precedenza da Friedman (1957) e Modigliani (1963). In altre parole, una variazione unitaria della ricchezza (umana o non) provoca una variazione del consumo pari a α_t , cioè al valore scontato relativo all'anno preso in considerazione.

Sottraendo all'equazione la parte deterministica (cioè le realizzazioni di queste variabili nel caso di assenza di innovazione) si ottiene:

$$\tilde{c}_t = \alpha_t (\tilde{U}_t + \tilde{A}_t)$$

$$\tilde{A}_t = (1+r)(\tilde{A}_{t-1} + P_{t-1} + u_{t-1} - c_{t-1});$$

$$\tilde{U}_t = \sum_{i=0}^{T-t} (1+r)^{-i} (P_t + u_t)$$

$$\tilde{U}_t = \frac{P_{t-1}}{\alpha_t} + \frac{\varepsilon_t}{\alpha_t} + \sum_{i=0}^{T-t} \left[(1+r)^{-i} \sum_{k=0}^K \phi_k v_{t-k} \right]$$

Dove P_t e u_t sono le componenti permanente e transitoria del processo del reddito.

Citando Hall (1978) *"un consumatore con aspettative razionali che massimizza il valore atteso di una funzione di utilità intertemporale separabile nel*

tempo possiede la seguente condizione: l'utilità marginale attesa del consumo nell'anno successivo dipende solo dal livello attuale del consumo e dai parametri della funzione di utilità." Di conseguenza, l'equazione di innovazione del consumo conterrà solamente quei valori che sono ignoti al tempo $t - 1$.

Consideriamo i fattori dalla quale dipende l'innovazione del consumo. α_t è il fattore di annuitizzazione, di conseguenza esso non viene eliminato. \tilde{A}_t è completamente noto al tempo $t - 1$, poiché r è assunto costante, e quindi non compare nell'equazione. Di conseguenza, l'equazione di innovazione del consumo è pari a α_t volte l'innovazione di \tilde{U}_t .

L'equazione di \tilde{U}_t rappresenta il capitale umano, definito come la somma scontata di tutto il reddito percepito più il valore atteso del reddito futuro. La parte di reddito permanente relativa agli anni precedenti a quello in esame (P_{t-1}) è sicuramente nota, ma così non è per ε_t ; della parte transitoria u_t , invece, non è noto tutto, non conosciamo infatti lo shock transitorio al tempo t . Di conseguenza della sommatoria $\sum_{k=0}^K \phi_k v_{t-k}$ rimarrà solamente $\phi_0 v_t$, ma poiché ϕ_0 è pari a 1, l'ultimo addendo di \tilde{U}_t sarà pari a $\sum_{i=0}^{T-t} [(1+r)^{-i}] v_t$. Moltiplicando per α_t questa equazione, scevra dei componenti già noti, si ottiene la formula dell'innovazione del consumo, pari a:

$$\Delta c_t = \varepsilon_t + \beta_t v_t; \quad \beta_t = \alpha_t \sum_{i=0}^{T-t} (1+r)^{-i}$$

Hall concentra la sua analisi sulla propensione marginale del consumo rispetto a shock inattesi e transitori del reddito, assumendo che un consumatore risponda pienamente agli shock permanenti (il coefficiente di ε_t

è posto pari a 1). Usando dati PSID su consumo di alimenti e reddito negli anni 1969-1975, egli stima, usando il metodo della massima verosimiglianza, un valore di β_t pari a 0.29, troppo alto per essere consistente con la teoria del reddito permanente di Friedman.

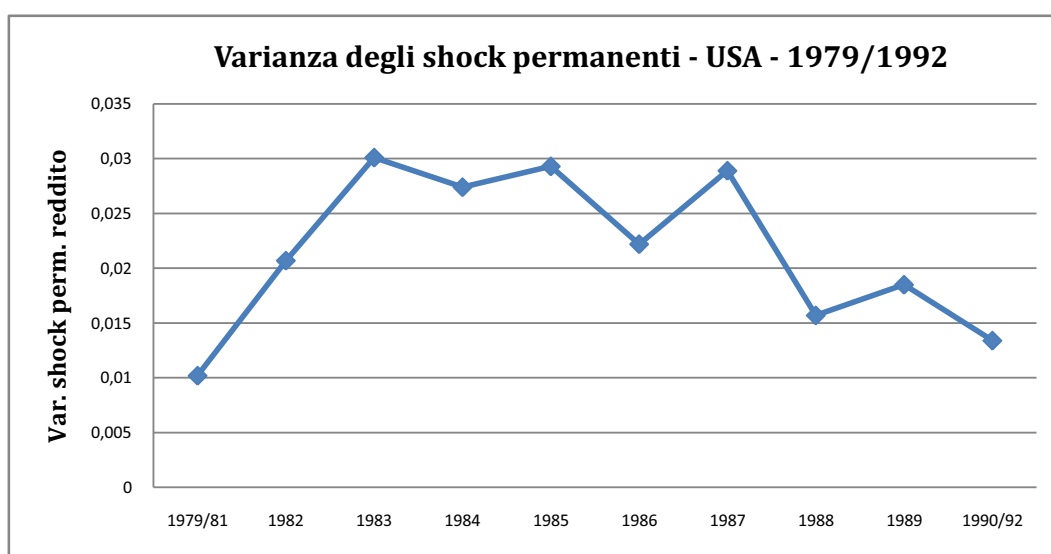
Blundell, Pistaferri e Preston (2008) riprendono il lavoro di Hall, estendendo il suo lavoro ad una funzione di domanda del tipo CRRA (*Constant Relative Risk Adversion*) al fine di approssimare una funzione del consumo a partire da una funzione del processo del reddito del tipo permanente-transitorio.

Ciò porta a sollevare il vincolo di assicurazione nulla sugli shock permanenti. Inoltre, considerano innovazioni in momenti più alti del processo del reddito e la presenza di variazioni nel tempo delle varianze degli shock di reddito, che Hall aveva supposto costanti.

Essi utilizzano, usando un approccio già sperimentato in precedenza da Skinner (1987) e Ziliak (1998), una misura del consumo di beni non durevoli derivante da dati PSID e CEX; essi stimano una funzione di domanda per il consumo di cibo e si regrediscono questi dati sulla spesa di cibo provenienti dal CEX instrumentando la spesa di beni non durevoli, sempre provenienti da questa indagine, con la media, specifica per gruppo, educazione e anno, della paga oraria del marito e della moglie delle famiglie prese in esame. Assumendo monotonia nella domanda di cibo, si può invertire tale equazione per ottenere una misura del consumo di beni non durevoli del PSID a partire dai dati sul consumo di cibo di tale indagine statistica.

Blundell, Pistaferri e Preston usano i risultati ottenuti da una parte, per analizzare le dinamiche di consumo e reddito degli Stati Uniti negli ultimi due decenni, dall'altra, per verificare come il modello di Friedman sul reddito permanente si adatti ai risultati ottenuti.

Le stime delle varianze degli shock transitori e permanenti e la loro evoluzione nel tempo vengono utilizzati per determinare le cause per cui, negli USA, la variabilità del consumo è cresciuta meno rispetto alla variabilità del reddito nei primi anni '80: si afferma che tale fenomeno è causato dalla variazione della persistenza degli shock del reddito in quel periodo; in particolare, durante i primi anni ottanta si può osservare come la varianza degli shock permanenti cresca visibilmente, per poi calare nella fine del decennio.



Inoltre, il coefficiente di assicurazione contro shock temporanei del reddito è molto basso (circa 0.05, e comunque non significativo), in linea con la teoria del reddito permanente; tuttavia, il coefficiente di assicurazione contro shock di reddito permanenti è molto più basso del valore unitario

ipotizzato da Friedman (pari a 0.642), il che indica che i soggetti economici sono in grado, seppure in parte, di assicurarsi contro shock permanenti.

Da notare che tali coefficienti sono più alti (il coefficiente di assicurazione contro shock permanenti è pari a 0.944, non significativamente diverso da 1, mentre quello contro shock temporanei è uguale a 0.076) per le famiglie il cui capofamiglia non ha completato il *college*, mentre sono più bassi (rispettivamente 0.419 e 0.027) per le famiglie con un elevato livello di istruzione. Tale differenza è ancora più accentuata se si considerano le famiglie a più basso reddito, ove i coefficienti assumono i valori di 0.849 e 0.288. Questo significa che sono presenti vincoli di liquidità per le famiglie a più basso reddito e meno istruite, mentre le famiglie più ricche e con un più elevato livello di istruzione hanno un accesso facilitato al mercato del credito, rendendole in grado di assicurarsi meglio contro variazioni inattese del reddito. Tali coefficienti spiegano anche perché la stima di Hall della propensione al consumo rispetto a shock transitori del reddito risultasse così elevata: quello che sembrava un risultato in contrasto alla teoria del reddito permanente era in realtà l'effetto del reddito permanente "mascherato", in quanto egli aveva assunto che uno shock permanente del reddito si trasmette integralmente al consumo.

La scomposizione degli shock di reddito ha diversi vantaggi: innanzitutto, è molto versatile, in quanto permette di stimare la propensione marginale del consumo non solo rispetto a shock di reddito, ma anche di studiare l'effetto di altri tipi di shock come, ad esempio, disoccupazione o mutazioni del tempo atmosferico nel caso di economie prettamente agricole, e di determinare la natura delle variazioni che tali eventi causano sul reddito

e di conseguenza anche sul consumo. Inoltre, c'è una netta separazione tra shock permanenti e transitori del reddito, e di come tali shock si trasmettano al consumo, al contrario dell'approccio quasi-sperimentale, dove si studiava lo shock causato dal singolo avvenimento, senza tener conto dei diversi effetti che tale avvenimento causava.

Anche questo approccio, però, non è esente da difetti: poiché si assume che reddito e consumo seguano determinati processi, è possibile che sia presente distorsione, provocata da errori di specificazione del modello. Inoltre richiede osservazioni di reddito e consumo ripetute negli anni, anche se non necessariamente provenienti dallo stesso dataset.

Una delle principali mancanze di questo approccio è che, nonostante esso riesca a determinare con precisione il collegamento esistente tra reddito e consumo, non riesce a spiegare il motivo per la quale esiste assicurazione contro gli shock; in particolare, non si riesce a separare l'effetto della nuova informazione da quello provocato dalla diversità del set di informazioni a disposizione del consumatore rispetto a quello alla quale ha accesso il ricercatore: il primo set contiene più informazioni rispetto al secondo, e questo potrebbe distorcere i coefficienti di assicurazione contro shock di reddito. Infatti, per includere questa distinzione all'interno dei processi di reddito e consumo, sarebbe necessario introdurre nuovi parametri da stimare che separino gli shock del reddito in attesi ed inattesi, con effetti diversi sul consumo. Tuttavia ciò non è possibile usando solamente dati sulle realizzazioni di reddito e consumo, in quanto tali processi generano un numero di equazioni non sufficiente.

Per ovviare a questo problema si è pensato di utilizzare dati sulle aspettative soggettive sul reddito dei consumatori per ricavare le equazioni necessarie per identificare un siffatto modello. Pistaferri (2001) e Kaufmann & Pistaferri (2009) utilizzano dati sulle aspettative soggettive dei consumatori contenuti all'interno dello SHIW¹⁰, trovando che buona parte degli shock transitori del reddito è in realtà anticipata o errore di misurazione (il 50% è anticipata, e il 30% è errore di misurazione), mentre i due terzi degli shock permanenti del reddito sono rappresentati da nuova informazione. Ciò indica che esiste una distorsione verso l'alto il grado di assicurazione contro shock di reddito, e confermano l'esistenza della differenza di informazioni disponibili al consumatore e al ricercatore.

Sebbene l'uso di aspettative soggettive sul reddito risolva il problema della differenza di informazioni tra consumatore e ricercatore, tale pratica porta con sé problematiche legate alla qualità e all'informatività di questi dati¹¹, e alla loro reperibilità (in Italia, dati di questo tipo sono presenti

¹⁰ Per ricavare i valori delle aspettative di reddito dei consumatori, essi ricorrono ad un lavoro di Guiso, Jappelli e Pistaferri (2002). Essi fanno ricorso ai dati presenti nello SHIW ed assumono che le aspettative di reddito dei consumatori seguano una distribuzione triangolare, con minimi e massimi pari al reddito (minimo e massimo) che il consumatore si aspetta di guadagnare nel caso lavorasse, e con funzione di ripartizione del punto medio della distribuzione pari alla probabilità attesa che ha il consumatore di guadagnare meno della media del reddito atteso. Viene usata la media di questa distribuzione per imputare le aspettative di reddito del consumatore.

Tuttavia questo riguarda solo il caso nella quale il consumatore effettivamente lavori; nelle aspettative di reddito bisogna includere anche il caso della disoccupazione. Si suppone che ogni consumatore abbia perfetta conoscenza dei meccanismi di welfare, e per questo le aspettative puntuali sui sussidi di disoccupazione sono poste pari ai sussidi potenziali.

Si ricava infine l'aspettativa di reddito del consumatore, pari a:

$$E[x] = (1 - p)E[y] + pE[b]$$

Dove p è la probabilità soggettiva di rimanere disoccupato, $E[y]$ è la media del reddito da lavoro atteso (ricavata dall'equazione di cui sopra), e $E[b]$ sono i benefit di cui il consumatore godrebbe nel caso restasse disoccupato.

¹¹ In Kaufmann e Pistaferri (2009), le restrizioni di sovraidentificazione vengono rigettate. Inoltre gli s.e. delle stime sono molto alti; gran parte dei parametri, infatti, risulta essere non

all'interno del SHIW, ma solamente fino al 2001), rendendo difficoltosa l'applicazione di risultati ottenuti in questo modo in altri ambiti.

Per questi motivi, si è deciso di non utilizzare dati sulle aspettative soggettive nelle analisi successive, e di utilizzare solamente dati delle realizzazioni di reddito e consumo.

significativamente diversa da zero. Ad esempio, la stima puntuale del coefficiente di assicurazione del consumo contro shock inattesi transitori del reddito è 0.312 ma ha uno standard error pari a 0.4274.

Capitolo 2: L'indagine sul reddito e la ricchezza delle famiglie della Banca d'Italia

Le indagini della Banca d'Italia sono tra le poche a contenere una dimensione **panel** sia per i dati sul reddito che per quelli sul consumo. Inoltre, sono universalmente accettati come dati di elevata qualità sia per la corrispondenza di tali dati con i principali aggregati nazionali, sia perché i dati sul reddito descrivono il reddito disponibile (cioè al netto dei tributi) e possiedono dati sul consumo di beni non durevole, che ci solleva da molte delle problematiche legate all'utilizzo di dati sul consumo di cibo.

Pertanto, una volta descritto al Cap. 1 come si è giunti alle odierne teorie per la stima della propensione marginale del consumo rispetto a variazioni inattese del reddito e, prima di approfondire i concetti teorici alla base delle analisi di cui ai capp. successivi, è utile descrivere l'indagine della Banca d'Italia per ciò che concerne il disegno campionario, sia per quanto riguarda i problemi affrontati in fase di raccolta e, successivamente, di trattamento dei dati.

2.a. Il disegno campionario

Dopo una serie di indagini condotte congiuntamente nel 1961-62 con Isco, nel 1966 la Banca d'Italia cominciò a pubblicare e commentare i risultati della Survey of Households' Income and Wealth (SHIW). Trentuno indagini sono state pubblicate fino a oggi, annualmente fino al 1987 (a parte per il 1985) e ogni due anni da allora¹². Fino al 1984 il numero delle famiglie intervistate era circa 4.000, e da quell'anno si è deciso di ingrandire il campione, raddoppiando le famiglie intervistate a partire dal 1986.

L'unità statistica di base è la famiglia, definita come gruppo di individui uniti da legami di sangue, matrimonio o affetto, che abitano nello stesso domicilio e condividono parte del reddito. In base a questa definizione, individui che vivono assieme unicamente per ragioni economiche non sono considerati membri della stessa famiglia, mentre solo una unità viene registrata nel caso di più famiglie, legate tra di loro, che vivono insieme.

La procedura di campionamento, rivista nel 1986 sulla base dell'indagine sulla forza lavoro dell'Istat, prevede che le municipalità italiane vengano suddivise in 51 strati, definiti da 17 regioni e 3 classi di popolazione: sopra i 40.000 abitanti, tra 20.000 e 40.000 abitanti, e sotto i 20.000 abitanti. Tutte le municipalità della prima classe vengono incluse, mentre quelle appartenenti alle altre categorie vengono scelte a caso.

Nel 1987 è stata introdotta una componente panel, al fine di migliorare il confronto delle varie grandezze economiche nel tempo. Tale componente è

¹² con l'eccezione del 1998, che vede trascorrere tre anni dall'ultima indagine.

diventata sempre più preponderante nell'indagine, partendo da un 15% nel 1987 (quando vennero contattate nuovamente solo le famiglie che esplicitamente volevano ripetere l'intervista) a un 45% nel 1995 fino al 54,5% nel 2008; l'evoluzione della componente panel nel tempo relativamente agli ultimi anni (assieme al numero di famiglie intervistate) si può vedere nella tabella sottostante:

anno della prima partecipazione all'indagine	anno di rilevazione					
	1998	2000	2002	2004	2006	2008
1998	4478	1993	1224	845	636	538
2000		4128	1014	667	475	398
2002			4406	1082	672	525
2004				4408	1334	995
2006					3811	1143
2008						3632
dimensione campione	7147	8001	8011	8012	7768	7977
quota famiglie panel	37,3	48,4	45,0	45,0	50,9	54,5

Per la parte *panel* del campione, sono state selezionate in primo luogo tutte le famiglie residenti in comuni già oggetto di rilevazione nella precedente raccolta dati; inoltre, altre famiglie sono state estratte casualmente tra quelle rilevate solo nella precedente indagine. Le famiglie non *panel* sono state invece estratte casualmente dalle liste anagrafiche, sia nei comuni *panel* che in quelli non *panel*.

2.b. Raccolta e trattamento dei dati

Con riferimento all'anno 2008, il campione di famiglie effettivamente intervistate è risultato composto da 7.977 famiglie, di cui 4.345 panel e 3.632 non panel. Per ottenere queste interviste, è stato necessario contattare 14.209

famiglie. Le mancate interviste costituiscono un problema nelle indagini statistiche in quanto possono condurre a campioni nei quali i segmenti di popolazione meno disposti a collaborare possono essere sottorappresentati, producendo stime distorte (*selectivity bias*). Gli stimatori utilizzati si basano sulle informazioni di coloro che hanno partecipato alla rilevazione e possono presentare una distorsione che cresce all'aumentare della mancata risposta e del divario tra il valore atteso della variabile oggetto di studio per i rispondenti e per i non rispondenti.

Nel corso degli anni sono state messe a punto delle strategie di rilevazione che hanno consentito di ridurre progressivamente il fenomeno della mancata risposta: prima della rilevazione le famiglie ricevono una lettera volta a sensibilizzarle sull'importanza dell'indagine e a rassicurarle sulla riservatezza delle informazioni fornite; gli intervistatori sono selezionati con sempre maggiore accuratezza; il numero totale di nominativi a disposizione di ciascun intervistatore per raggiungere il proprio obiettivo di famiglie intervistate è stato progressivamente ridotto.

Famiglie contattate e motivi della mancata intervista

Famiglie:	<i>Panel</i>		<i>Non panel</i>		Totale	
	<i>(unità)</i>	<i>(percentuali)</i>	<i>(unità)</i>	<i>(percentuali)</i>	<i>(unità)</i>	<i>(percentuali)</i>
Intervistate	4.345	79,3	3.632	41,6	7.977	56,1
Indisponibili	1.012	18,5	3.589	41,1	4.601	32,4
Irreperibili	120	2,2	1.511	17,3	1.631	11,5
Totale	5.477	100,0	8.732	100,0	14.209	100,0
Ineleggibili (*)	150	2,7	629	6,7	779	5,2

(*) Famiglie non esistenti all'indirizzo anagrafico (indirizzi errati, decessi, trasferimenti).

Per limitare i possibili effetti negativi della mancata partecipazione sono state inoltre adottate alcune misure. In primo luogo, le famiglie che non

è stato possibile intervistare sono state sostituite con altre estratte con criteri casuali negli stessi comuni. La sostituzione avviene attraverso una rigida procedura sotto il controllo della società di rilevazione che stabilisce sia quando un nominativo può essere sostituito, sia il nuovo nominativo da contattare. In secondo luogo, al termine della rilevazione, è stata effettuata una stratificazione a posteriori sulla base di alcune caratteristiche individuali dei soggetti intervistati, che consente di riequilibrare, all'interno del campione, il peso dei diversi segmenti della popolazione. Nel 2008, per intervistare le 7.977 famiglie del campione, è stato necessario porre in atto nei loro confronti 14.839 azioni di contatto. La difficoltà a ottenere l'intervista è crescente al crescere del reddito, della ricchezza e del titolo di studio del capofamiglia: inoltre, se il capofamiglia è imprenditore o dirigente sono necessari un numero superiore di azioni di contatto. Minori difficoltà si incontrano con le famiglie con un ridotto numero di componenti, residenti al Sud e Isole o in comuni di piccole dimensioni, con capofamiglia anziano o non occupato.

Sebbene per un'indagine sul reddito e la ricchezza la mancata risposta costituisca un elemento rilevante nel determinare problemi di qualità delle stime, numerosi altri fattori possono essere indicati come elemento di disturbo nei dati raccolti.

Un aspetto che può influire sulla qualità delle stime riguarda la reticenza delle famiglie a dichiarare le proprie fonti di reddito o le forme di attività finanziarie o reali possedute. Sebbene la partecipazione all'indagine sia volontaria e il contenuto della rilevazione noto all'intervistato sin dall'inizio, è possibile che talvolta questi non risponda con totale sincerità alle domande più personali, quali quelle riguardanti il reddito o la ricchezza.

Per disporre di elementi di giudizio su tale fenomeno, che per sua natura risulta difficilmente investigabile, è stato richiesto agli intervistatori di formulare, immediatamente dopo l'intervista, un giudizio sintetico sulla presunta attendibilità delle risposte su reddito e ricchezza, in base alla rispondenza tra i dati forniti e gli elementi oggettivi a loro disposizione (zona e tipologia dell'immobile di residenza, tenore di vita desumibile dagli arredi, ecc.).

Sebbene il livello di attendibilità sia risultato in media soddisfacente, esso non è omogeneo all'interno del campione. Si riscontrano giudizi migliori per le famiglie con livelli di reddito e ricchezza bassi, residenti al Centro o con capofamiglia con elevato titolo di studio o dirigente. Un'attendibilità leggermente inferiore si riscontra per le famiglie con capofamiglia lavoratore indipendente (imprenditore o lavoratore autonomo). Ulteriori elementi per valutare l'attendibilità delle risposte fornite dagli intervistati possono essere ottenuti comparando le stime desunte dall'indagine con i dati della Contabilità nazionale. Queste comparazioni vanno effettuate con cautela, tenendo conto che, almeno in parte, le differenze riscontrate possono essere dovute alle diverse definizioni adottate. Le stime ricavate dall'indagine sono state, in passato, comparate anche con quelle derivanti dalle dichiarazioni fiscali, segnalando un sostanziale allineamento nei redditi da lavoro dipendente e una sottostima consistente dei redditi da lavoro autonomo dichiarati al fisco rispetto a quelli rilevati nell'indagine.

Altri problemi possono derivare dai processi cognitivi del rispondente; le risposte a determinate domande sono influenzate dalla capacità di elaborazione (ad esempio per le domande ipotetiche) o di memoria (per le domande retrospettive) del rispondente. Vanno inoltre considerati gli

elementi di conoscenza vera e propria, soprattutto quando, per questioni pratiche, è talvolta consentito a un soggetto di rispondere per conto di un altro.

Lo studio degli errori di misura ha permesso di valutare quali siano le variabili presenti nell'indagine che presentino una maggior affidabilità. Con riferimento ai principali aggregati una maggiore affidabilità si può osservare per il reddito, mentre per i consumi e la ricchezza netta tale affidabilità è minore. Tra le componenti di reddito mostrano una migliore qualità i redditi da pensioni e da lavoro dipendente. Il reddito da lavoro autonomo e il reddito da capitale sono nel complesso rilevati con minore affidabilità.

Lo SHIW possiede diverse caratteristiche interessanti: innanzitutto, tale indagine contiene informazioni su consumo, reddito, e ricchezza dei consumatori, variabili non solitamente disponibili all'interno dello stesso dataset; inoltre è presente una grossa componente panel nei dati: tali caratteristiche permettono la stima di processi del reddito, consumo e ricchezza anche molto complessi, e analisi del mercato del lavoro e transizioni di portafoglio. I dati sul consumo non si limitano al solo consumo di cibo, ma sono presenti anche misure del consumo di beni non durevoli e durevoli, superando la principale limitazione che altri dataset usati per testare modelli intertemporali del consumo (ad esempio, lo PSID) possiedono. Inoltre, i dati sul reddito misurano il reddito disponibile, in quanto sono raccolti solo dopo aver tenuto conto delle tasse, cosa che di rado succede negli altri dataset (lo PSID ne è ancora un esempio).

Le misure di reddito, consumo e lavoro di questa indagine ricalcano abbastanza bene gli aggregati nazionali, cosa che indica l'elevata qualità dei

dati; infine, per ognuna delle principali variabili di interesse per i macroeconomisti, lo SHIW ne contiene scomposizioni dettagliate, permettendo ai ricercatori lo sperimentare di diverse misure del reddito (da lavoro, da capitale, trasferimenti), consumo (beni durevoli, non durevoli) e ricchezza (possedimenti reali, finanziari, debiti e altri componenti). Per ulteriori informazioni, e una analisi più dettagliata della qualità dei dati SHIW, si vedano Brandolini e Cannari (1994) e Brandolini (1999).

A causa dell'irregolarità nella rilevazione dei dati che caratterizza questo dataset (i dati più recenti hanno cadenza biennale, a parte l'intervallo 1995-1998) e del fatto che la componente panel è presente solo a partire dal 1987, nella presente tesi si è deciso di utilizzare solamente i dati relativi agli anni dal 1998 al 2008, in quanto dal 1987 al 1995 si disporrebbe di un panel di lunghezza pari a 5 periodi invece dei 6 periodi disponibili usando i dati più recenti.

Se nel caso di un modello statico un panel non equispaziato nel tempo non porta particolari problemi alle stime dei parametri, viceversa, nel caso di modelli dinamici, o in modelli che hanno come variabile dipendente una variabile in differenze prime, questo può portare a distorsioni delle stime, in quanto sono presenti variabili risposta ritardate all'interno del modello; il dato dell'anno 1998 si riferisce a informazioni relative al 1996, tuttavia, tali informazioni non sono disponibili, creando un modello che risulta essere di difficile interpretazione.

Capitolo 3: Il modello economico e l'identificazione dei parametri

Nel presente capitolo andremo a descrivere come si utilizzeranno i dati ricavati dallo SHIW, indagine descritta nel capitolo precedente, per stimare i parametri di un modello economico che lega reddito e consumo. E' necessario tuttavia, per prima cosa, definire tale modello economico, definendo delle equazioni che descrivano il comportamento del reddito e del consumo e il legame tra queste due grandezze.

3.a. Il processo del reddito e del consumo

Il processo del reddito che andremo ad utilizzare è quello conosciuto come processo permanente-transitorio. Questo perchè tale processo permette la presenza di shock al reddito di diversa persistenza, a differenza di altre definizioni del processo del reddito, come quella ARMA, descritta nel capitolo 1.

Assumiamo che il reddito sia influenzato da caratteristiche demografiche come età, titolo di studio, ecc... e che tali caratteristiche siano contenute in un vettore $Z_{i,t}$. L'equazione del reddito sarà quindi pari a:

$$\ln y_{i,t} = \log(Y_{i,t}) = Z'_{i,t} \beta_t + P_{i,t} + v_{i,t}$$

Dove $P_{i,t}$ rappresenta lo shock permanente al reddito, al tempo t , del singolo consumatore i e si assume, seguendo l'esempio di numerosi studi empirici (Abowd, Card, 1989; Meghir, Pistaferri, 2004), che segua un processo *random walk*. Di conseguenza:

$$P_{i,t} = P_{i,t-1} + \delta_{i,t}$$

Dove $\delta_{i,t}$ è un white noise, e quindi serialmente incorrelato. Inoltre:

$$v_{i,t} = \sum_{j=0}^q \theta_j \varepsilon_{t-j}; \theta_0 = 1$$

$v_{i,t}$ rappresenta lo shock transitorio del reddito, e si assume che tale shock, in questa versione del modello, sia di tipo white noise. Nei paragrafi successivi questa restrizione verrà sollevata, assumendo per la parte transitoria del reddito, un processo MA(1). È stato dimostrato (Abowd, Card, 1989; Meghir, Pistaferri, 2004; Hryshko, 2006) che tale frazione del reddito segue un processo di tipo MA(1); disponendo di dati biennali, la prima autocorrelazione che si riesce a calcolare è quella al ritardo 2, che per un processo MA(1) è pari a zero.

Assumere uno shock transitorio del reddito di tipo MA(q) non influisce sull'identificazione del modello: uno shock transitorio di questo tipo

aggiunge q-1 parametri da stimare, ma anche q-1 equazioni; infatti l'autocorrelazione di un processo MA(q) è zero solo per intervalli strettamente maggiori di q.

Si definisce la variazione non spiegata del reddito come la differenza prima del reddito stesso, scevra dell'effetto delle caratteristiche demografiche:

$$y_{i,t} = \Delta(l y_{i,t} - Z'_{i,t} \beta_t) = \delta_{i,t} + \Delta v_{i,t}$$

Si noti che effettuando le differenze prime, si eliminano anche gli effetti delle caratteristiche invarianti nel tempo spesso non misurabili, come intelligenza, produttività, o condizioni iniziali di ricchezza. Se espandiamo l'equazione del reddito in modo da includere questi fattori, si otterrà l'equazione seguente:

$$l y_{i,t} = Z'_{i,t} \beta_t + Q'_i \gamma_i + P_{i,0} + \sum_{j=1}^T \delta_{i,j} + v_{i,t}$$

Dove γ_i sono caratteristiche non misurabili che non variano nel tempo (come la capacità dell'individuo, o la sua intelligenza), Q'_i sono i coefficienti che legano questi fattori al reddito e $P_{i,0}$ sono le condizioni iniziali di ricchezza della famiglia. Se sottraiamo l'effetto delle caratteristiche osservabili e portiamo in differenze prime, otterremo:

$$\begin{aligned} y_{i,t} &= \Delta(l y_{i,t} - Z'_{i,t} \beta_t) = \\ &= Q'_i \gamma_i - Q'_i \gamma_i + P_{i,0} - P_{i,0} + \sum_{j=1}^T \delta_{i,j} - \sum_{k=1}^{T-1} \delta_{i,k} + v_{i,t} - v_{i,t-1} = \end{aligned}$$

$$= \delta_{i,t} + \Delta v_{i,t}$$

Che è esattamente l'equazione di cui sopra.

Una volta esplicitato il processo del reddito, si può modellare come tali shock influenzino il consumo. Consideriamo il problema di massimizzazione dell'utilità del consumo di un soggetto economico con funzione di utilità del tipo CRRA. L'equazione da massimizzare è:

$$\max_c \left\{ E_t \left[\sum_{i=0}^{T-t} (1 + \delta)^{-i} \frac{C_{t+i}^\beta - 1}{\beta} e^{Z_{t+i} \vartheta_{t+i}} \right] \right\}$$

Sotto il vincolo di budget

$$A_{t+j+1} = (1 + r)(A_{t+j} + Y_{t+j} - C_{t+j})$$

Dove δ è il tasso di sconto, r il tasso di interesse, c_t il consumo al tempo t , y_t è il reddito da lavoro al tempo t e A sono i risparmi o asset. β rappresenta l'avversione al rischio del soggetto economico, mentre Z_t rappresentano avvenimenti, comuni a tutti i soggetti economici, che causano variazioni del consumo.

Si ricava la seguente equazione del consumo:

$$C_{t-1}^{\beta-1} = \frac{1 + r}{1 + \delta} e^{\Delta Z'_{t-1} \vartheta_{t-1}} E_{t-1} [C_t^{\beta-1}]$$

Che tuttavia non possiede una soluzione chiusa, e che quindi non è utilizzabile per ricavare una equazione del processo del consumo. Blundell et

al. (2010) utilizzano un approssimazione in serie di Taylor del logaritmo di una sommatoria¹³ per approssimare la funzione del consumo, ottenendo:

$$\Delta \log c_t = \Delta[\log(C_t)] = \Delta Z'_t \vartheta'_t + \eta_t + \Omega_t$$

Dove η_t è uno shock del reddito con media nulla, Ω_t indica tutti gli effetti causati da tassi d'interesse, impazienza e risparmio precauzionale e $\vartheta'_t = \vartheta_t/(1 - \beta)$.

Supponendo che ogni componente idiosincratca presente nell'equazione possa essere adeguatamente approssimata da una componente deterministica Γ_t e una componente stocastica φ_t , e che quindi $\Omega_t = \Gamma_t + \varphi_t$, l'equazione dell'innovazione del consumo si può scrivere nel seguente modo:

$$\Delta c_t = \eta_t + \varphi_{i,t}, \quad c_t = \log(c_t) - \Delta Z'_t \vartheta'_t - \Gamma_t$$

Usando un'equazione del reddito del tipo permanente-transitorio, si ottiene il seguente vincolo di budget intertemporale:

$$\sum_{i=0}^{T-t} q_{t+i} C_{t+i} = \sum_{i=0}^{L-t} q_{t+i} Y_{t+i} + A_t$$

Dove T rappresenta la data della morte, L la data di ritiro dal lavoro, e q_{t+i} un adeguato fattore di sconto.

Approssimando usando la precedente scomposizione in serie di Taylor, e usando un procedimento simile a quello adottato da Hall e Mishkin

¹³ $\log \sum_{k=0}^{S-t} X_{t+k} \approx \sum_{k=0}^{S-t} \alpha_{t+k,S,\delta} [\log(X_{t+k}) - \log(\alpha_{t+k,S,\delta})]$;
 $\alpha_{t+k,S,\delta} = [\exp(\sum_{i=0}^k \delta_{t+i})] / \{1 + \sum_{k=1}^{S-t} [\exp(\sum_{i=0}^k \delta_{t+i})]\}$

(1978) e descritto nel capitolo 1, si considerano le deviazioni dalle componenti deterministiche delle varie grandezze presenti nell'equazione, eliminando tutte quelle variabili già note al consumatore nel periodo $t - 1$. Si ottiene una funzione dello shock del reddito:

$$\eta_t = \pi_t(\delta_{i,t} + \gamma_t \varepsilon_t)$$

Dove $\pi_t = \sum_{i=0}^{L-t} q_{t+i} Y_{t+i} / (\sum_{i=0}^{L-t} q_{t+i} Y_{t+i} + A_t)$ rappresenta la parte del reddito da lavoro futuro della ricchezza della famiglia, e

$$\gamma_t = \frac{r}{1+r} \left[1 + \sum_{j=1}^q \frac{\theta_j}{(1+r)^j} \right],$$

con q ordine del processo MA della componente transitoria del reddito, rappresenta un fattore di annuitizzazione che incrementa con l'età. Di conseguenza, l'equazione dell'innovazione del consumo diventa:

$$c_t = \pi_t \delta_{i,t} + \gamma_t \pi_t \varepsilon_t + \varphi_{i,t}$$

$$c_t = \phi_t \delta_{i,t} + \psi_t \varepsilon_t + \varphi_{i,t}$$

Dove $\phi_{i,t}$ è la propensione marginale del consumo rispetto a shock permanenti del reddito, $\psi_{i,t}$ è la propensione marginale del consumo rispetto a shock transitorio del reddito e $\varphi_{i,t}$ è lo shock transitorio del consumo e che contiene al suo interno anche l'errore di misurazione del consumo (della quale discuteremo nei prossimi paragrafi). Si noti che quanto detto sopra riguardo le componenti invarianti nel tempo presenti all'interno dell'equazione del reddito si può applicare anche a questa equazione.

Tale formulazione del processo del consumo fornisce anche una interpretazione strutturale della stessa: uno shock permanente avrà un effetto sul consumo pari al valore futuro atteso scontato del reddito da lavoro, in quanto tale shock influenzerà tutto il reddito che percepirà da quel momento in poi. Nel caso di uno shock transitorio, l'effetto sul consumo sarà pari allo stesso valore atteso scontato, ma moltiplicato per un fattore di annuitizzazione, rappresentante il fatto che l'effetto dello shock è pari solo alla parte relativa all'anno successivo del reddito futuro atteso.

Questo tipo di modello include al suo interno due opposte teorie del consumo, il modello dei mercati completi con informazione perfetta ($\phi_{i,t} = \psi_{i,t} = 0$) e il modello del reddito permanente di Friedman ($\phi_{i,t} = 1; \psi_{i,t} = 0$). Sotto mercati completi gli agenti economici possono assicurarsi contro tutti gli shock idiosincratici del reddito.

Tuttavia, informazione asimmetrica e *moral hazard* rendono tali assicurazioni difficili da implementare, e infatti non sono osservabili nella realtà. La volatilità del consumo individuale è molto più alta di quella aggregata, fatto che va contro l'ipotesi di mercati completi (Aiyagari, 1994).

Ulteriori test (Jappelli e Pistaferri, 2001) rigettano l'ipotesi di completa assicurazione del consumo, rendendo tale teoria poco adatta a descrivere fenomeni economici. In base a questi risultati, la moderna teoria del consumo si sviluppa a partire dall'assunzione di mercati incompleti (Hall, 1978).

Nella teoria del reddito permanente ($\phi_{i,t} = 1; \psi_{i,t} = 0$), l'unico meccanismo disponibile ai consumatori per stabilizzare il consumo è il risparmio precauzionale; nella versione più estrema della predetta teoria,

non sono previste assicurazioni di alcun tipo e si assumono preferenze quadratiche: uno shock permanente del reddito si trasmette integralmente al consumo.

Questi modelli sono caratterizzazioni estreme del comportamento individuale e del mercato nel quale il consumatore agisce, e il modello che si è ricavato in questo capitolo, nel caso entrambi i coefficienti siano compresi tra 0 e 1, descrive un mercato ad assicurazione parziale, dove gli shock del reddito inattesi sono assicurabili in parte.

Ciò potrebbe essere poichè il consumatore dispone di un set di informazioni maggiore rispetto all'econometrista, o in quanto la persistenza dello shock non è quella prevista (uno shock che viene considerato permanente può venire valutato dal consumatore come avere effetto finito nel tempo, diminuendo così l'effetto di tale shock sul consumo).

3.b. Identificazione del modello

Una volta determinati i processi che descrivono reddito e consumo, si può procedere all'identificazione dei parametri del modello. Seguendo il metodo proposto da Hall e Mishkin (1982) e altri, imponiamo delle restrizioni alle covarianze del processo bivariato di reddito e consumo. Nel caso in esame, si assume che la capacità di assicurare il consumo da parte dei soggetti economici sia invariante nel tempo, e quindi ϕ e ψ sono supposti costanti. Si assume, invece, eteroschedasticità nel tempo per le varianze degli

shock del reddito. L'equazione del reddito si può usare per ottenere le seguenti restrizioni:

$$\begin{aligned} var[\Delta y_{i,t}] &= var[\delta_t] + var[\Delta \varepsilon_t] \\ &= \sigma_{\delta_t}^2 + \sigma_{\varepsilon_t}^2 + \sigma_{\varepsilon_{t-1}}^2 \end{aligned}$$

$$\begin{aligned} cov[\Delta y_{i,t} \Delta y_{i,t-1}] &= cov[\Delta \varepsilon_t \Delta \varepsilon_{t-1}] \\ &= -\sigma_{\varepsilon_{t-1}}^2 \end{aligned}$$

Nel caso di un processo del reddito con shock transitorio di tipo white noise, le covarianze ai ritardi maggiore del primo sono pari a 0

Le restrizioni sulle covarianze del consumo sono le seguenti:

$$\begin{aligned} var[\Delta c_{i,t}] &= \phi^2 var[\delta_t] + \psi^2 var[\varepsilon_t] + var[\varphi] \\ &= \phi^2 \sigma_{\delta_t}^2 + \psi^2 \sigma_{\varepsilon_t}^2 + \sigma_{\varphi}^2 \end{aligned}$$

$$cov[\Delta c_{i,t} \Delta c_{i,t-k}] = 0; \forall k > 0$$

Tale equazione mostra che la volatilità del consumo può aumentare per due ragioni: o per una diminuzione del grado di assicurazione rispetto a shock di reddito (con varianze date), oppure per un aumento delle varianze degli shock di reddito (a parità di grado di assicurazione).

Infine, la covarianza tra reddito e consumo ai vari ritardi è pari a:

$$\begin{aligned} cov[\Delta(C_{i,t}) \Delta(y_{i,t})] &= \phi var[\delta_t] + \psi var[\varepsilon_t] \\ &= \phi \sigma_{\delta_t}^2 + \psi \sigma_{\varepsilon_t}^2 \end{aligned}$$

$$cov[\Delta(C_{i,t}) \Delta(y_{i,t+1})] = \psi cov[\Delta \varepsilon_{t+1} \varepsilon_t]$$

$$= -\psi\sigma_{\varepsilon_t}^2$$

Nel caso di shock temporanei del reddito serialmente incorrelati.

Se si suppone che tali shock seguano un processo del tipo MA(q), allora $cov[\Delta(C_{i,t})\Delta(y_{i,t+k})]$ è pari a 0 per $|k| > q + 1$. Questo aggiunge q parametri extra da stimare (i coefficienti del processo MA), ma anche q momenti extra; in tal modo, l'identificazione del modello non viene inficiata dalla correlazione seriale del processo.

Per identificare i parametri nel modello più semplice servono dati relativi ad almeno 4 anni: $t-2, t-1, t, t+1$. Per ogni singolo periodo vanno identificati 5 parametri: le varianze, di shock del reddito permanenti e transitori, e shock del consumo, indicati rispettivamente da $\sigma_{\delta_t}^2, \sigma_{\varepsilon_t}^2$, che variano nel tempo, e σ_{ϕ}^2 , invariante nel tempo, nonché i coefficienti di assicurazione contro shock permanenti e transitori del reddito ϕ e ψ , assunti stazionari .

Per ottenere i momenti necessari, ci si rifà a un risultato ottenuto da Meghir e Pistaferri (2004). Si può provare, infatti, che:

$$E \left[\Delta y_t \left(\sum_{j=-(1+q)}^{1+q} \Delta y_{t+j} \right) \right] = E[\delta_t^2]$$

Dove $E[\delta_t^2]$ è la varianza dello shock permanente del reddito, e q è l'ordine del processo MA dello shock transitorio del reddito. Applicando tale equazione al nostro caso (shock di reddito transitori *white noise*, e quindi $q = 0$) si ha:

$$E[\Delta y_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})] = \sigma_{\delta_t}^2$$

$$E[\Delta y_t \Delta y_{t-1}] = E[\Delta y_{t+1} \Delta y_t] = -\sigma_{\varepsilon_t}^2$$

L'identificazione della varianza dello shock permanente si basa sul fatto che tale shock è pari alla varianza della variazione del reddito ($E[\Delta y_t \Delta y_t]$) priva degli effetti della componente transitoria ($E[\Delta y_t \Delta y_{t-1}] + E[\Delta y_{t+1} \Delta y_t]$). Di conseguenza, data T la lunghezza del panel, si potranno identificare solo un totale di T-3 varianze dello shock permanente e T-2 varianze dello shock transitorio del reddito. Inoltre, è possibile identificare le varianze degli shock di reddito senza aver bisogno di dati sul consumo, a patto di avere un panel sufficientemente lungo.

I restanti parametri vengono identificati dalle seguenti equazioni; si può provare che:

$$\phi = \frac{E[\Delta C_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]}{E[\Delta y_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]}$$

$$\psi = \frac{E[\Delta C_t \Delta y_{t+1}]}{E[\Delta y_t \Delta y_{t+1}]}$$

$$\begin{aligned} \sigma_{\phi_t}^2 = & E[\Delta C_t(\Delta C_{t-1} + \Delta C_t + \Delta C_{t+1})] - \frac{[E[\Delta C_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]]^2}{E[\Delta y_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]} \\ & + \frac{[E[\Delta C_t \Delta y_{t+1}]]^2}{E[\Delta y_t \Delta y_{t+1}]} \end{aligned}$$

Queste condizioni sui momenti danno la completa identificazione dei parametri d'interesse.

ψ si identifica grazie al fatto che reddito e consumo ritardato correlano attraverso tale coefficiente ($E[\Delta C_t \Delta y_{t+1}] = -\psi \sigma_{\varepsilon_t}^2$). Scalando tale risultato per la varianza dello shock temporaneo ($E[\Delta y_t \Delta y_{t+1}] = -\sigma_{\varepsilon_t}^2$) si ottiene il parametro desiderato. Da notare che esiste anche un'interpretazione IV di tale equazione: ψ è il risultato della regressione di ΔC_t su Δy_t , usando Δy_{t+1} come strumento.

Lo stesso ragionamento si può fare per ϕ , nella quale l'effetto permanente del reddito sul consumo ($E[\Delta C_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})] = \phi \sigma_{\delta_t}^2$) viene scalato per la varianza dello shock permanente; l'interpretazione IV indica ancora una regressione di ΔC_t su Δy_t , ma questa volta lo strumento utilizzato è $(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})$.

L'ultima equazione rappresenta la varianza del consumo privata dell'effetto dello shock permanente ($\frac{[E[\Delta C_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]]^2}{E[\Delta y_t (\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]} = \phi^2 \sigma_{\delta_t}^2$) e di quello transitorio ($\frac{[E[\Delta C_t \Delta y_{t+1}]]^2}{E[\Delta y_t \Delta y_{t+1}]} = -\psi^2 \sigma_{\varepsilon_t}^2$). In questo caso, il più semplice, si può sostituire $E[\Delta C_t (\Delta C_{t-1} + \Delta C_t + \Delta C_{t+1})]$ con $E[\Delta C_t^2]$, tuttavia in versioni più complesse del modello, della quale discuteremo successivamente, è presente un errore di misurazione, che renderebbe la stima distorta. Usando questo risultato, invece, si va a isolare la varianza del consumo, pari a $\phi^2 \sigma_{\delta_t}^2 + \psi^2 \sigma_{\varepsilon_t}^2 + \sigma_{\varphi}^2$, sia con che senza errore di misurazione. Inoltre, non viene influenzata l'identificazione, in quanto i periodi $t-1$ e $t+1$ sono già usati negli addendi successivi.

3.b.1. Identificazione del modello con errori di misura del consumo

Consideriamo il caso realistico nella quale il consumo sia misurato con errore. L'equazione del consumo sarà quindi:

$$c_t^* = c_t + \omega_t^\varepsilon$$

Dove c_t^* è il valore misurato del consumo, formato dal vero valore c_t e da un errore di tipo *white noise* ω_t^ε . La variazione del consumo nel periodo t sarà quindi pari a:

$$\Delta c_t = \phi_t \delta_t + \psi_t \varepsilon_t + \varphi_t + \Delta \omega_t^\varepsilon$$

Similmente all'equazione del reddito, anche in questa equazione la varianza della componente presente solo attraverso la sua differenza prima (in questo caso l'errore di misurazione del consumo) è identificata usando l'autocorrelazione al ritardo uno:

$$E[\Delta c_t \Delta c_{t-1}] = E[\Delta c_{t+1} \Delta c_t] = -\sigma_{\omega_t^\varepsilon}^2$$

Gli altri parametri vengono identificati ancora attraverso le equazioni precedenti, l'unica che richiede una modifica è l'equazione della varianza del consumo, che si riscrive come:

$$\begin{aligned} var[\Delta c_{i,t}] &= \phi^2 var[\delta_t] + \psi^2 var[\varepsilon_t] + var[\varphi] + var[\Delta \omega_t] \\ &= \phi^2 \sigma_{\delta_t}^2 + \psi^2 \sigma_{\varepsilon_t}^2 + \sigma_\varphi^2 + \sigma_{\omega_t^\varepsilon}^2 + \sigma_{\omega_{t-1}^\varepsilon}^2 \end{aligned}$$

Da notare che viene presa in considerazione l'eterogeneità dell'errore di misurazione del consumo, anche se si possono identificare solamente T-1 varianze di tale grandezza.

3.b.2. Identificazione del modello con errori di misura del reddito

Nel caso che il reddito sia misurato con errore, e quindi

$$y_t^* = y_t + \omega_t^y$$

dove y_t^* è il valore misurato del reddito, y_t il vero valore di tale grandezza, e ω_t^y un errore di tipo *white noise*, l'identificazione dei parametri ϕ , $\sigma_{\delta_t}^2$ e $\sigma_{\omega_t^c}^2$ è ancora garantita, anche se vengono usate le grandezze *misurate*, e non quelle reali. Tuttavia, non è più possibile separare $\sigma_{\varepsilon_t}^2$ da $\sigma_{\omega_t^y}^2$ e di conseguenza, la stima di ψ diventa distorta, più precisamente, è possibile solo calcolarne un limite inferiore. Difatti:

$$\psi \geq \frac{E[\Delta C_t^* \Delta y_{t+1}^*]}{E[\Delta y_t^* \Delta y_{t+1}^*]} = \frac{-(\psi \sigma_{\varepsilon_t}^2)}{-(\sigma_{\varepsilon_t}^2 + \sigma_{\omega_t^y}^2)}$$

E' tuttavia possibile ottenenere l'identificazione del parametro anche in questo caso, a patto di porre delle restrizioni. Considerando ε_t come eterogenea, ma assumendo la stazionarietà di ω^y , l'equazione che definisce ψ è:

$$\begin{aligned}\psi &= \frac{E[\Delta C_t^* \Delta y_{t+1}^* - \Delta C_{t-1}^* \Delta y_t^*]}{E[\Delta y_t^* \Delta y_{t+1}^* - \Delta y_{t-1}^* \Delta y_t^*]} = \frac{-(\psi \sigma_{\varepsilon_t}^2) - (\psi \sigma_{\varepsilon_{t-1}}^2)}{-(\sigma_{\varepsilon_t}^2 + \sigma_{\omega_y}^2) + (\sigma_{\varepsilon_{t-1}}^2 + \sigma_{\omega_y}^2)} \\ &= \frac{\psi(\sigma_{\varepsilon_t}^2 + \sigma_{\varepsilon_{t-1}}^2)}{(\sigma_{\varepsilon_t}^2 + \sigma_{\varepsilon_{t-1}}^2) + (-\sigma_{\omega_y}^2 + \sigma_{\omega_y}^2)} = \psi\end{aligned}$$

Usando un ragionamento simile, si può dimostrare che anche della varianza dello shock del consumo non è identificato, e se ne può ricavare solamente un limite superiore.

$$\begin{aligned}\sigma_\varphi^2 &\leq E[\Delta C_t^* (\Delta C_{t-1}^* + \Delta C_t^* + \Delta C_{t+1}^*)] - \frac{[E[\Delta C_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*)]]^2}{E[\Delta y_t^* (\Delta y_{t-1}^* + \Delta y_t^* + \Delta y_{t+1}^*)]} \\ &\quad + \frac{[E[\Delta C_t^* \Delta y_{t+1}^*]]^2}{E[\Delta y_t^* \Delta y_{t+1}^*]}\end{aligned}$$

In quanto l'ultimo addendo è distorto verso l'alto. Infatti:

$$\frac{[E[\Delta C_t^* \Delta y_{t+1}^*]]^2}{E[\Delta y_t^* \Delta y_{t+1}^*]} = \frac{(\psi \sigma_{\varepsilon_t}^2)^2}{-(\sigma_{\varepsilon_t}^2 + \sigma_{\omega_t}^2)}$$

Come indicato precedentemente, l'uso di $[\Delta C_t^* (\Delta C_{t-1}^* + \Delta C_t^* + \Delta C_{t+1}^*)]$ al posto della varianza del consumo all'interno della formula serve a mitigare l'effetto dell'errore di misurazione sulla stima del parametro. Purtroppo, in questo caso, anche assumendo stazionarietà dell'errore non è possibile risalire a una stima corretta dello shock del consumo.

La distorsione che si viene a creare introducendo l'errore di misurazione del consumo è abbastanza marcata: la stima del coefficiente di

assicurazione contro shock transitori del reddito diminuisce di 0.05; le varianze di tali shock aumentano di circa 0.035 per tutti gli anni presi in considerazione (ad esempio, la varianza dello shock transitorio del 2006 passa da 0.0585 a 0.0927), anche gli *standard error* legati a questi parametri aumentano vistosamente.

Per questi motivi, si è deciso di non considerare l'errore di misura del reddito nelle analisi successive.

3.b.3. Identificazione del modello con MA(1) come processo dello shock di reddito transitorio

Numerose analisi sul reddito (Abowd, Card, 1989; Meghir, Pistaferri, 2004; Hryshko, 2006) provano che la componente transitoria del processo del reddito segue un processo di tipo MA(1); di conseguenza sarebbe opportuno estendere il modello finora utilizzato per tenere conto di questo fatto; tuttavia, bisogna ricordare che i dati SHIW presi in esame sono dati raccolti con cadenza biennale, quindi anche estendendo l'analisi a un processo di tipo MA(2), non si sarebbe in grado di determinare il valore del primo dei due coefficienti di un tale processo, a causa della mancanza di dati.

E' tuttavia possibile usare un processo MA(1), supponendo che il gap tra due osservazioni sia pari a due anni.

Le equazioni che descrivono il modello diventano:

$$\Delta(y_{i,t}) = \delta_{i,t} + \Delta v_{i,t}; \quad v_{i,t} = \varepsilon_t + \theta \varepsilon_{t-1}$$

$$\Delta C_t = \phi \delta_{i,t} + \psi \varepsilon_t + \varphi_{i,t} + \Delta \omega_t^c$$

Si aggiunge un'equazione alle restrizioni, in quanto la correlazione del reddito è nulla solo per ritardi superiori a $q + 1$. Le equazioni delle varianze e covarianze del reddito diventano, pertanto:

$$\begin{aligned} var[\Delta(y_{i,t})] &= var[\delta_t] + var[\Delta v_t] \\ &= \sigma_{\delta_t}^2 + \sigma_{\varepsilon_t}^2 + (\theta - 1)^2 \sigma_{\varepsilon_{t-1}}^2 + \theta^2 \sigma_{\varepsilon_{t-2}}^2 \end{aligned}$$

$$\begin{aligned} cov[\Delta(y_{i,t})\Delta(y_{i,t-1})] &= cov[\Delta v_t \Delta v_{t-1}] = \\ &= E[(\varepsilon_t + (\theta - 1)\varepsilon_{t-1} - \theta\varepsilon_{t-2})(\varepsilon_{t-1} + (\theta - 1)\varepsilon_{t-2} - \theta\varepsilon_{t-3})] = \\ &= (\theta - 1)\sigma_{\varepsilon_{t-1}}^2 - \theta(\theta - 1)\sigma_{\varepsilon_{t-2}}^2 \end{aligned}$$

$$\begin{aligned} cov[\Delta(y_{i,t})\Delta(y_{i,t-2})] &= cov[\Delta v_t \Delta v_{t-2}] = \\ &= E[(\varepsilon_t + (\theta - 1)\varepsilon_{t-1} - \theta\varepsilon_{t-2})(\varepsilon_{t-2} + (\theta - 1)\varepsilon_{t-3} - \theta\varepsilon_{t-4})] = \\ &= -\theta\sigma_{\varepsilon_{t-2}}^2 \end{aligned}$$

Le restrizioni su varianza e covarianza del consumo non variano rispetto al caso precedente (il consumo è influenzato solo dagli shock transitori del reddito contemporanei, non da quelli passati), mentre la covarianza tra reddito e consumo ai vari ritardi è pari a:

$$\begin{aligned} cov[\Delta(C_{i,t})\Delta(y_{i,t})] &= \phi var[\delta_t] + \psi var[\varepsilon_t] \\ &= \phi \sigma_{\delta_t}^2 + \psi \sigma_{\varepsilon_t}^2 \end{aligned}$$

$$\begin{aligned} cov[\Delta(C_{i,t})\Delta(y_{i,t+1})] &= \psi cov[\varepsilon_t \Delta v_{t+1}] \\ &= \psi E[\varepsilon_t(\varepsilon_{t+1} + (\theta - 1)\varepsilon_t - \theta\varepsilon_{t-1})] = -\psi(\theta - 1)\sigma_{\varepsilon_t}^2 \end{aligned}$$

$$cov[\Delta(C_{i,t})\Delta(y_{i,t+2})] = \psi cov[\varepsilon_t \Delta v_{t+2}] =$$

$$= \psi E[\varepsilon_t(\varepsilon_{t+2} + (\theta - 1)\varepsilon_{t+1} - \theta\varepsilon_t)] = -\psi\theta\sigma_{\varepsilon_t}^2$$

Come spiegato all'inizio di questo capitolo a proposito del processo del consumo, si può notare come l'aggiunta di un'incognita (il coefficiente MA) non influenza l'identificazione del modello, in quanto vengono aggiunti momenti extra. Tali momenti sono l'autocovarianza al ritardo 2 del reddito e la covarianza allo stesso ritardo tra reddito e consumo.

E' da notare, tuttavia, che si potranno identificare, usando questo modello, meno varianze rispetto a quante se ne potevano trovare con i modelli precedenti. Questo è dovuto al risultato di Meghir e Pistaferri, che qui applicato fornisce la seguente equazione per l'identificazione delle varianze degli shock del reddito:

$$\sigma_{\delta_t}^2 = E[\Delta y_t(\Delta y_{t-2} + \Delta y_{t-1} + \Delta y_t + \Delta y_{t+1} + \Delta y_{t+2})]$$

$$\theta\sigma_{\varepsilon_t}^2 = -E[\Delta y_t\Delta y_{t+2}]$$

Usando un processo MA(1) per descrivere la parte transitoria del processo del reddito, si potranno trovare solamente T-5 (per t che va da 4 a T-2) varianze dello shock permanente del reddito e dello shock del consumo e T-3 (per t che va da 2 a T-2) varianze dello shock transitorio del reddito.

Quando si dispone di panel molto lunghi (come ad esempio quello presente all'interno dello PSID), questo non è un problema, tuttavia, come descritto nel capitolo precedente, i problemi che hanno portato la Banca d'Italia al ritardo di 1 anno nel presentare i dati dello SHIW nel 1997, permette di usare solamente dati a partire dal 1998, disponendo così di un panel con T pari a 6.

Questo porta a ottenere una sola varianza dello shock permanente (e come si vedrà successivamente, anche un solo coefficiente di assicurazione del consumo contro shock permanenti del reddito quando si toglie il vincolo di stazionarietà su questo parametro), rendendo infruttuoso l'utilizzo di questo modello con i dati a nostra disposizione.

3.b.4. Non stazionarietà dei coefficienti di assicurazione

Il modello fin qui descritto è identificato (e anche sovraidentificato) solamente usando varianze e covarianze del reddito e del consumo, grazie al vincolo di stazionarietà imposto ai coefficienti di assicurazione (ϕ e ψ). Quando si solleva tale vincolo, le equazioni presenti non sono più sufficienti a fornire una soluzione per il modello. E' necessario, quindi, trovare altre equazioni, dipendenti sia dai parametri, che da reddito e/o consumo. Il risultato ottenuto da Meghir e Pistaferri (2004), descritto precedentemente, fornisce proprio tale equazione. In tal modo si possono trovare fino a T-3 coefficienti di assicurazione del consumo contro shock permanenti del reddito, e T-2 coefficienti relativi agli shock transitori.

$$\phi_t = \frac{E[\Delta C_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]}{E[\Delta y_t(\Delta y_{t-1} + \Delta y_t + \Delta y_{t+1})]}, 3 < t < T - 1$$

$$\psi_t = \frac{E[\Delta C_t \Delta y_{t+1}]}{E[\Delta y_t \Delta y_{t+1}]}, 2 < t < T - 1$$

Anche in tal caso, assumere un processo di tipo MA(1) per gli shock transitori del reddito riduce, rispettivamente, a T-5 e T-3 i coefficienti di assicurazione contro shock del reddito permanenti e transitori.

Alla luce dei dati a disposizione, e del numero di parametri che si sarebbe in grado di identificare con ogni specificazione del modello finora utilizzata, si è deciso, per le analisi successive, di non utilizzare l'estensione del modello che prevede per lo shock temporaneo del reddito un processo del tipo MA(1), in quanto il panel SHIW che si utilizza ha una lunghezza limitata (6 periodi) e si riuscirebbe a trovare solamente una varianza dello shock permanente del reddito, e un singolo coefficiente di assicurazione contro shock permanenti del reddito, anche assumendo la loro non-stazionarietà.

c. Modelli di regressione per dati di panel

I dati in formato panel combinano le informazioni relative alle caratteristiche di N individui nello stesso istante temporale con quelle rilevate per gli stessi individui in T diversi periodi di tempo. Nei modelli di tipo panel i dati disponibili hanno perciò entrambe le caratteristiche di:

- *Dati Cross Section*: per un dato istante sono osservate le caratteristiche di più individui;
- *Serie storiche*: per un dato collettivo di individui sono rilevate le diverse caratteristiche in diversi istanti.

La convenienza dell'utilizzo dei modelli di tipo panel risiede soprattutto nel guadagno di efficienza della stima perchè il maggior numero di osservazioni che si ha rispetto alla sola dimensione cross section o serie storica genera uno stimatore con varianza più piccola. Quando si parla di dati longitudinali si intende una struttura nella quale generalmente la numerosità degli individui è elevata, mentre quella relativa alla dimensione temporale è piuttosto contenuta.

3.c.1. Modello ad effetti fissi

3.c.1.1. *Il modello a variabili dummy*

Considerando l'*i*-esimo individuo, consideriamo il seguente modello:

$$y_i = \alpha_i + x_i\beta + \varepsilon_i$$

La peculiarità di questa equazione riguarda la costante, che si configura come un vettore di *T* elementi costanti pari ad α_i : questa caratteristica indica innanzi tutto che per ciascun individuo occorre stimare un solo valore della costante e che, per $\alpha_i \neq \alpha_j$, tale costante misura l'effetto individuale, cioè quell'insieme di caratteristiche specifiche proprie di ciascun individuo, e che perciò restano immutate nel tempo. In pratica, nel modello ci sono in tutto $k + N$ parametri da stimare, k contenuti nel vettore β ed N costanti per i diversi individui. Queste costanti rappresentano l'eterogeneità presente tra gli individui nel sistema, caratteristica peculiare dei panel data.

Generallizzando la formula in forma matriciale si ottiene;

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} i_T & 0 & \dots & 0 & X_1 \\ 0 & i_T & \dots & 0 & X_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & X_{N-1} \\ 0 & 0 & \dots & i_T & X_N \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

Dove i_T è un vettore di T elementi pari a 1. In forma compatta si ha perciò:

$$Y = (I_N \otimes i_T) \alpha + X \beta + \varepsilon$$

Poiché in questa formulazione tutte le ipotesi classiche sono rispettate, si può stimare il modello attraverso OLS. Questo modello prende il nome di modello a variabili dummy, in quanto bisogna inserire N dummy (una per unità statistica) all'interno dei regressori. Tale stimatore si può calcolare a partire dalla formula dello stimatore dei minimi quadrati:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} (I_N \otimes i_T)' (I_N \otimes i_T) & (I_N \otimes i_T)' X \\ X' (I_N \otimes i_T) & X' X \end{bmatrix}^{-1} \begin{bmatrix} (I_N \otimes i_T)' Y \\ X' Y \end{bmatrix}$$

Per la proprietà del prodotto di Kronecher vale $(I_N \otimes i_T)' (I_N \otimes i_T) = I_N \otimes i_T' i_T = T I_N$. Sostituendo nell'equazione sopra si ottiene:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} T I_N & (I_N \otimes i_T)' X \\ X' (I_N \otimes i_T) & X' X \end{bmatrix}^{-1} \begin{bmatrix} (I_N \otimes i_T)' Y \\ X' Y \end{bmatrix}$$

Sfruttando un noto risultato sulle matrici partizionate, riferendoci in particolare all'inversione:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1} A_{12} S_2 A_{21} A_{11}^{-1} & -A_{11}^{-1} A_{12} S_2 \\ -S_2 A_{21} A_{11}^{-1} & S_2 \end{bmatrix}$$

$$S_2 = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$$

e dopo alcuni calcoli, si giunge allo stimatore ad effetti fissi, pari a:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} (I_N \otimes i_T)' (Y - X\hat{\beta}) \\ (X'MX)^{-1}X'MY \end{bmatrix}$$

Dove $M = I_N - P$ è la matrice di proiezione che, applicata a una variabile, restituisce lo scostamento dalla media temporale dell'individuo. Tale matrice è, per definizione, quadrata ($NT \times NT$), diagonale a blocchi, simmetrica e idempotente.

3.c.1.2. *Lo stimatore within*

Prendendo in considerazione lo stimatore $\hat{\beta}$ di cui sopra, e tenendo conto che M è simmetrica e idempotente, si può ricavare il seguente risultato:

$$\begin{aligned} \hat{\beta} &= (X'MX)^{-1}X'MY \\ \hat{\beta} &= (\dot{X}\dot{X}')^{-1}\dot{X}\dot{Y} \end{aligned}$$

Tale stimatore è ottenibile dalla regressione OLS di $\dot{Y} = MY$ su $\dot{X} = MX$. In altre parole si usano i minimi quadrati ordinari su dati espressi come deviazioni individuali dalle medie temporali, ottenendo uno stimatore che è numericamente identico allo stimatore a variabili dummy. Tale stimatore è detto anche stimatore *within*, in quanto tiene conto degli effetti individuali grazie alla trasformazione dei dati in deviazioni dalla media, ma allo stesso tempo li elimina dal modello.

Le costanti si possono ricavare a partire dalla formula del modello in forma compatta; si ha:

$$\begin{aligned}
 Y &= (I_N \otimes i_T) \alpha + X \beta + \varepsilon \\
 (I_N \otimes i_T) \alpha &= Y - X \beta \\
 \frac{1}{T} (I_N \otimes i_T)' (I_N \otimes i_T) \alpha &= \frac{1}{T} (I_N \otimes i_T)' (Y - X \beta)
 \end{aligned}$$

Per la proprietà del prodotto di Kronecher di cui sopra si sostituisce $T I_N$ a $(I_N \otimes i_T)' (I_N \otimes i_T)$ alla parte sinistra dell'equazione:

$$\begin{aligned}
 \frac{1}{T} T I_N \alpha &= \frac{1}{T} (I_N \otimes i_T)' (Y - X \beta) \\
 \hat{\alpha} &= \frac{1}{T} (I_N \otimes i_T)' (Y - X \beta)
 \end{aligned}$$

Tale equazione mostra che le costanti sono pari alle medie delle variabili dipendenti meno le medie delle variabili indipendenti, pesate attraverso i regressori *within*. Per ogni individuo, si ha:

$$\hat{\alpha}_i = \bar{y}_i - \bar{x}_i \beta$$

Tali costanti catturano l'effetto di tutte quelle variabili, diverse da individuo a individuo, invarianti nel tempo.

Il limite più evidente di questo stimatore consiste nell'impossibilità di includere nel modello regressori che assumano un valore costante all'interno delle osservazioni relative al singolo individuo: dal punto di vista algebrico calcolare lo scostamento di queste variabili dal loro valore medio individuale produrrebbe colonne di zeri nella matrice dei regressori che quindi non

avrebbe rango pieno. In questo caso il metodo OLS non sarebbe perciò applicabile.

Per la verifica di ipotesi di eterogeneità tra gli individui presi in esame l'ipotesi di nullità dei coefficienti α_i non risulta adatta; piuttosto, si prende in considerazione l'ipotesi nulla di uguaglianza tra tutti i coefficienti (N-1 vincoli), cioè $\alpha_1 = \alpha_2 = \dots = \alpha_N$, e si costruisce un test F; la statistica test è:

$$\frac{\tilde{\varepsilon}' \tilde{\varepsilon} - \hat{\varepsilon}' \hat{\varepsilon}}{\hat{\varepsilon}' \hat{\varepsilon}} \frac{NT - N - K - 1}{N - 1} \sim F_{N-1, NT-N-K-1}$$

Dove $\tilde{\varepsilon}$ sono i residui del modello vincolato (coefficienti uguali tra di loro) e $\hat{\varepsilon}$ sono i residui del modello non vincolato.

la varianza dei coefficienti dello stimatore within è pari a:

$$\text{Var}(\hat{\beta}) = \sigma_{\varepsilon}^2 (X' MX)^{-1}$$

Dove σ_{ε}^2 è la stima corretta e consistente della varianza dei residui:

$$\sigma_{\varepsilon}^2 = \frac{\hat{\varepsilon}' \hat{\varepsilon}}{NT - N - K - 1}$$

Lo stimatore within è:

- BLUE (Best Linear Unbiased Estimator)
- Consistente per $NT \rightarrow \infty$
- Asintoticamente normale: si ha infatti che

$$\sqrt{NT}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \sigma_{\varepsilon}^2 \left(\lim_{NT \rightarrow \infty} \left(\frac{1}{NT} X' MX\right)^{-1}\right)^{-1}\right)$$

3.c.2. Modello ad effetti casuali

Come descritto nel paragrafo precedente, spesso nelle analisi econometriche si ha la necessità di regredire delle grandezze su valori che non variano nel tempo (come la regione di residenza o il sesso dell'individuo), e ciò non è possibile usando lo stimatore *within*. Per questo si è sviluppato uno stimatore che non ha questa limitazione, ma che tuttavia richiede delle condizioni aggiuntive affinché produca stime consistenti. Tale stimatore è detto stimatore ad effetti casuali.

Lo stimatore ad effetti casuali assume che gli effetti individuali siano stocastici, cioè fanno parte dell'errore incorrelato con i regressori; questo permette di produrre stime utilizzando variabili invarianti nel tempo. Il modello statistico utilizzato è il seguente:

$$y_i = \alpha + x_i\beta + \gamma_i + \varepsilon_{it}$$

L'effetto delle caratteristiche individuali $\alpha_i = \alpha + \gamma_i$ è scomposto in una costante più una componente che varia da individuo a individuo. Affinchè si ottengano stime consistenti usando questo approccio, bisogna però porre delle condizioni aggiuntive al vettore γ_i :

- $E[\gamma_i] = 0$
- $Var[\gamma_i] = \sigma_\gamma^2 \forall i$ (omoschedasticità)
- $E[\gamma_i\gamma_j] = 0 \forall i \neq j$ (incorrelazione tra gli effetti individuali)
- $E[\gamma_i\varepsilon_{j,t}] = 0 \forall i, j, t$ (incorrelazione tra effetti individuali e residui)

La forma compatta del modello è:

$$Y = \alpha + X\beta + (\gamma \otimes i_T) + \varepsilon$$

Dove γ è il vettore di dimensione N contenente gli effetti individuali. Si può notare che l'errore, pari in questo caso a $(\gamma \otimes i_T) + \varepsilon$, è composto da una componente che varia tra gli individui e il tempo (ε), e una componente che varia solo tra gli individui ($\gamma \otimes i_T$). Date le ipotesi di cui sopra, individuare la matrice delle varianze e covarianze dell'errore Ω diventa fondamentale per la stima dei parametri del modello. Questa matrice è pari a:

$$\begin{aligned}\Omega &= E \left[((\gamma \otimes i_T) + \varepsilon)((\gamma \otimes i_T) + \varepsilon)' \right] \\ \Omega &= E [(\gamma \otimes i_T)(\gamma \otimes i_T)' + \varepsilon\varepsilon'] \\ \Omega &= E[\gamma\gamma' \otimes i_T i_T'] + E[\varepsilon\varepsilon']\end{aligned}$$

Data l'omoschedasticità di γ ed ε , si ha:

$$\Omega = \sigma_\gamma^2 [I_N \otimes i_T i_T'] + \sigma_\varepsilon^2 I_{NT}$$

Questa è una matrice $NT \times NT$ diagonale a blocchi, con ogni blocco, di dimensione $T \times T$, formato da:

$$\Omega_i = \begin{bmatrix} \sigma_\gamma^2 + \sigma_\varepsilon^2 & \sigma_\gamma^2 & \cdots & \sigma_\gamma^2 \\ \sigma_\gamma^2 & \sigma_\gamma^2 + \sigma_\varepsilon^2 & \cdots & \sigma_\gamma^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\gamma^2 & \sigma_\gamma^2 & \cdots & \sigma_\gamma^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

Poiché la matrice di varianze e covarianze ha questa forma, il modello ad effetti casuali deve essere stimato attraverso i minimi quadrati generalizzati, quindi:

$$\hat{\beta} = (X' \Omega^{-1} X)^{-1} (X' \Omega^{-1} Y)$$

Lo stimatore ha le seguenti proprietà:

- Se σ_γ^2 e σ_ε^2 sono noti, lo stimatore GLS è consistente per $N \rightarrow \infty$ e $T \rightarrow \infty$
- Per T dato, lo stimatore GLS è più efficiente dello stimatore within; tale efficienza tende a ridursi all'aumentare di N
- Se $\Omega^{-1} \equiv M$, lo stimatore GLS diventa pari allo stimatore within; questo accade se l'unica fonte di variabilità deriva dagli effetti individuali γ . Ciò accade per
 - $\sigma_\varepsilon^2 = 0$
 - $T \rightarrow \infty$: Considerando il modello per la singola osservazione $y_i - \alpha - x_i \beta = \gamma_i + \varepsilon_{it}$. se $T \rightarrow \infty$, per la legge dei grandi numeri, significa che il valore atteso della componente ε è davvero nullo, quindi l'espressione a sinistra del segno di uguaglianza rappresenta la singola osservazione per γ . In questo caso lo stimatore GLS è consistente.

c.3. Il test di Hausman

Una volta regrediti reddito e consumo su caratteristiche demografiche usando gli stimatori di cui ai paragrafi precedenti, rimane da scegliere quale di questi due stimatori utilizzare per ricavare i residui che verranno utilizzati per le analisi successive; lo stimatore within è costoso in termini di gradi di libertà, in quanto bisogna inserire numerose variabili all'interno del modello;

d'altra parte, lo stimatore ad effetti casuali richiede che gli effetti individuali siano incorrelati con i regressori, altrimenti lo stimatore stesso diventa inconsistente.

Per effettuare questa scelta, si ricorrerà a un test costruito da Hausmann nel 1978 per verificare l'assunzione di ortogonalità di uno stimatore. La costruzione del test richiede che siano disponibili i risultati di due regressioni, uno efficiente e l'altro non efficiente, ma che mantenga la consistenza anche se l'assunzione di ortogonalità viene rigettata. Tali regressori, in questa tesi, sono rispettivamente lo stimatore ad effetti casuali e quello ad effetti fissi.

Il test di Hausman verifica le ipotesi:

$$\begin{cases} H_0: E[X'[(\gamma \otimes i_T) + \varepsilon]] = 0 \\ H_1: E[X'[(\gamma \otimes i_T) + \varepsilon]] \neq 0 \end{cases}$$

Sotto l'ipotesi nulla, la differenza tra i valori dei coefficienti delle due regressioni è non significativa; se verificata, tale ipotesi indica che i due stimatori presi in considerazione restituiscono risultati simili, e quindi è da preferire lo stimatore efficiente. Se invece tale ipotesi viene rigettata, le stime dei coefficienti saranno significativamente differenti, lo stimatore efficiente risulterà essere distorto e inconsistente, e quindi sarà da preferire lo stimatore non efficiente, in quanto questo stimatore mantiene la proprietà di consistenza anche sotto l'ipotesi alternativa.

Il test è basato sulla differenza $\hat{q} = \hat{\beta}_{WG} - \hat{\beta}_{GLS}$; se è verificata H_0 , si userà lo stimatore ad effetti casuali, in quanto efficiente; se è verificata H_1 , si

userà lo stimatore ad effetti fissi, poiché $\hat{\beta}_{GLS}$ sarebbe stimatore distorto, mentre $\hat{\beta}_{WG}$ rimarrebbe consistente.

La statistica test è data da:

$$t = \hat{q}' Var[\hat{q}]^{-1} \hat{q}$$

Sotto H_0 si può dimostrare che la covarianza tra i due stimatori è nulla: consideriamo lo stimatore $\tilde{\beta}$ pari a

$$\tilde{\beta} = \hat{\beta}_{GLS} + k\hat{\beta}_{WG}; \quad k \neq 0$$

Si ha che:

$$\begin{aligned} Var[\tilde{\beta}] &= Var[\hat{\beta}_{GLS}] + k^2 Var[\hat{\beta}_{WG}] + 2k Cov[\hat{\beta}_{WG}, \hat{\beta}_{GLS}] \\ Var[\tilde{\beta}] - Var[\hat{\beta}_{GLS}] &= k^2 Var[\hat{\beta}_{WG}] + 2k Cov[\hat{\beta}_{WG}, \hat{\beta}_{GLS}] \end{aligned}$$

Poiché $Var[\tilde{\beta}] - Var[\hat{\beta}_{WG}] > 0$ per costruzione, deve anche valere

$$k \left[k Var[\hat{\beta}_{WG}] + 2 Cov[\hat{\beta}_{WG}, \hat{\beta}_{GLS}] \right] > 0$$

Ciò vale per $k < 0$ e per $k > -2 \frac{Cov[\hat{\beta}_{WG}, \hat{\beta}_{GLS}]}{Var[\hat{\beta}_{WG}]}$. La condizione di positività di $Var[\tilde{\beta}] - Var[\hat{\beta}_{GLS}]$ è garantita per ogni k se e solo se i due stimatori sono incorrelati.

Alla luce di questo risultato si ha che $Var[\hat{q}] = Var[\hat{\beta}_{WG}] - Var[\hat{\beta}_{GLS}] > 0$.

La distribuzione del test di Hausman è una χ_g^2 con gradi di libertà pari al numero di regressori.

3.d. La stima delle varianze

Una volta stimati gli effetti sul reddito e sul consumo delle caratteristiche demografiche dei consumatori usando i modelli descritti sopra, per avere dati utilizzabili per l'analisi, bisogna ricavare le varianze e covarianze di queste grandezze. Per fare questo consideriamo la forma matriciale dei dati a nostra disposizione. Per semplicità di notazione, definiamo:

$$y_{i,t} = \Delta(\log(Y_{i,t}) - X_i\beta^y) \quad \text{e} \quad c_{i,t} = \Delta(\log(C_{i,t}) - X_i\beta^c)$$

dove β^y e β^c sono i coefficienti risultanti dalle stime degli effetti delle caratteristiche demografiche eseguite rispettivamente su reddito e consumo. Indichiamo con 0 il primo anno di rilevazione (1998) e con T l'ultimo (2008). Per ogni individuo si avrà:

$$y_i = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \vdots \\ y_{i,T} \end{pmatrix}; \quad c_i = \begin{pmatrix} c_{i,1} \\ c_{i,2} \\ \vdots \\ c_{i,T} \end{pmatrix}; \quad i \in [1; N]$$

Tuttavia non tutte le unità statistiche sono state intervistate in tutti gli anni dell'indagine presi in considerazione. Per tener conto di questo, come prima cosa, si porrà pari a zero ogni dato mancante. Inoltre si definiranno i seguenti vettori:

$$d_i^y = \begin{pmatrix} d_{i,1}^y \\ d_{i,2}^y \\ \vdots \\ d_{i,T}^y \end{pmatrix}; \quad d_i^c = \begin{pmatrix} d_{i,1}^c \\ d_{i,2}^c \\ \vdots \\ d_{i,T}^c \end{pmatrix}; \quad i \in [1; N]$$

Dove ogni $d_{i,j}^y = 1$ se è presente un dato per il reddito per l'unità i al tempo j , mentre sarà pari a 0 se quell'unità non è stata intervistata al tempo j ; varrà la stessa cosa per $d_{i,j}^c$ per i dati sul consumo. Sovrapponendo i dati sul reddito e sul consumo dei vari individui si otterrà:

$$x_i = \begin{pmatrix} y_i \\ c_i \end{pmatrix}; \quad d_i = \begin{pmatrix} d_i^y \\ d_i^c \end{pmatrix}$$

Usando questi vettori le stime delle varianze di reddito e consumo e quelle delle covarianze tra reddito e consumo si otterranno attraverso l'equazione:

$$M = \sum_{i=1}^N x_i x_i' \oslash \sum_{i=1}^N d_i d_i'$$

Dove \oslash rappresenta la divisione elemento per elemento. Per ottenere la varianza delle stime appena ottenute, ai fini di inferenza, definiamo i seguenti vettori:

$$m = \text{vech}\{M\}; \quad m_i = \text{vech}\{x_i x_i'\}; \quad D_i = \text{vech}\left\{\sum_{i=1}^N d_i d_i'\right\}$$

La matrice di varianze e covarianze V che può essere usata per l'inferenza si ottiene attraverso l'equazione:

$$V = \left[\sum_{i=1}^N ((m_i - m)(m_i - m)') \otimes D_i D_i' \right] \oslash [D_i D_i']$$

Dove \otimes indica il prodotto elemento per elemento. Le radici quadrate degli elementi sulle diagonali di V sono gli standard error dei corrispondenti elementi di m .

3.e. Lo stimatore a distanze minime

Dopo aver costruito un modello economico che lega reddito e consumo tra di loro, e definito le restrizioni sulle varianze di queste grandezze, scevre degli effetti di caratteristiche demografiche, si sono calcolate le stime delle varianze di reddito e consumo assieme alla matrice delle varianze di tali stime. Ottenuti questi dati, si può ora procedere alla stima dei valori dei coefficienti che descrivono il legame tra reddito e consumo, e le varianze dei vari shock di reddito.

Per stimare questi valori ci si rifà a uno stimatore creato da Chamberlain nel 1982, lo stimatore a distanze minime.

Sia w_i il vettore degli elementi distinti di $x_i x_i'$, definiti nel paragrafo precedente, con varianza non nulla. Le restrizioni vengono esplicitate ponendo m , vettorizzazione della matrice di varianze e covarianze di reddito e consumo (di dimensione $p \times 1$), dipendente da un sistema di equazioni funzione solamente del vettore θ di dimensione $q \times 1$ di parametri ignoti, con $p > q$. Il modello da stimare avrà la forma:

$$m = g(\theta) + Y$$

Con $g(\theta)$ vettore di funzioni dei parametri di interesse (varianza dello shock permanente del reddito, varianza dello shock transitorio del reddito, coefficienti di assicurazione, ecc...) pari a quelle elencate nei capitoli precedenti, e Y errore della stima. Si avrà così, nel modello con errori di misura del consumo:

$$\begin{pmatrix} E[y_1^2] \\ E[y_1 y_2] \\ \vdots \\ E[y_1 y_T] \\ E[y_2^2] \\ \vdots \\ \vdots \\ E[c_1^2] \\ E[c_1 c_2] \\ \vdots \\ E[c_T^2] \end{pmatrix} = \begin{pmatrix} \sigma_{\delta_1}^2 + 2\sigma_{\varepsilon_1}^2 \\ -\sigma_{\varepsilon_1}^2 \\ \vdots \\ 0 \\ \sigma_{\delta_1}^2 + \sigma_{\varepsilon_2}^2 + \sigma_{\varepsilon_1}^2 \\ \vdots \\ \vdots \\ \phi^2 \sigma_{\delta_1}^2 + \psi^2 \sigma_{\varepsilon_1}^2 + \sigma_{\varphi}^2 + 2\sigma_{\omega_1^c}^2 \\ -\sigma_{\omega_1^c}^2 \\ \vdots \\ \phi^2 \sigma_{\delta_T}^2 + \psi^2 \sigma_{\varepsilon_T}^2 + \sigma_{\varphi}^2 + \sigma_{\omega_T^c}^2 + \sigma_{\omega_{T-1}^c}^2 \end{pmatrix} + Y$$

Per ottenere lo stimatore a distanze minime si sceglie il vettore $\hat{\theta}$ che risolve il seguente problema di minimo:

$$\min_{\theta \in \Theta} \left\{ \sum_{i=1}^N (w_i - g(\theta))' K (w_i - g(\theta)) \right\}$$

Tale problema è equivalente al seguente:

$$\min_{\theta \in \Theta} \left\{ (\bar{w} - g(\theta))' K (\bar{w} - g(\theta)) \right\} = \min_{\theta \in \Theta} \left\{ (m - g(\theta))' K (m - g(\theta)) \right\}$$

Per completezza, definiamo dei set di condizioni di regolarità e le proprietà che esse implicano:

Assunzione 1: $m \xrightarrow{q.c.} g(\theta^\circ)$, con θ° veri valori dei parametri;
 T , dominio di θ , è un sottoinsieme compatto di \mathbb{R}^p che contiene θ° ;
 $g(\cdot)$ è continua in T
 K è matrice positiva definita.

Assunzione 2: $\sqrt{N}(m - g(\theta^\circ)) \xrightarrow{d} N(0, V)$;
 T contiene un intorno di θ° nella quale la derivata seconda di $g(\cdot)$ è continua;
 $G = g'(\theta^\circ)$, cioè la derivata prima di $g(\theta)$ calcolata in θ° , ha rango pari a p .

Se le assunzioni 1 sono soddisfatte, allora $\hat{\theta} \xrightarrow{q.c.} \theta^\circ$.

Se le assunzioni 1 e 2 sono soddisfatte, allora $\sqrt{N}(\hat{\theta} - \theta^\circ) \xrightarrow{d} N(0, \Delta)$, con Δ pari a:

$$\Delta = (G'KG)^{-1}G'KVKG(G'KG)^{-1}$$

Una scelta ottimale per K è V^{-1} ; lo stimatore con tale matrice dei pesi viene chiamato stimatore a distanze minime ottime (OMD); se viene scelta come matrice dei pesi la matrice identità, sarà chiamato stimatore a distanze minime a pesi eguali (EWMD); se invece, si sceglie la matrice D , matrice diagonale con gli elementi non nulli pari a $diag(V^{-1})$, si tratterà dello stimatore a distanze minime a pesi diagonali (DWMD). Quest'ultimo caso sarà quello che verrà utilizzato nelle analisi successive.

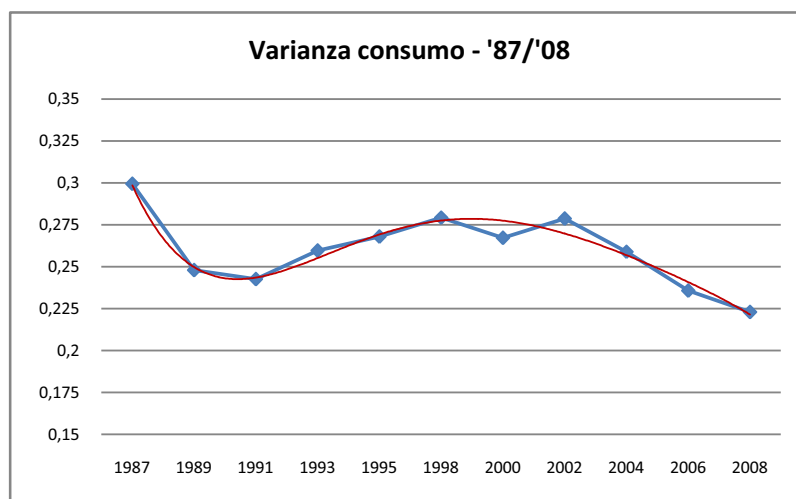
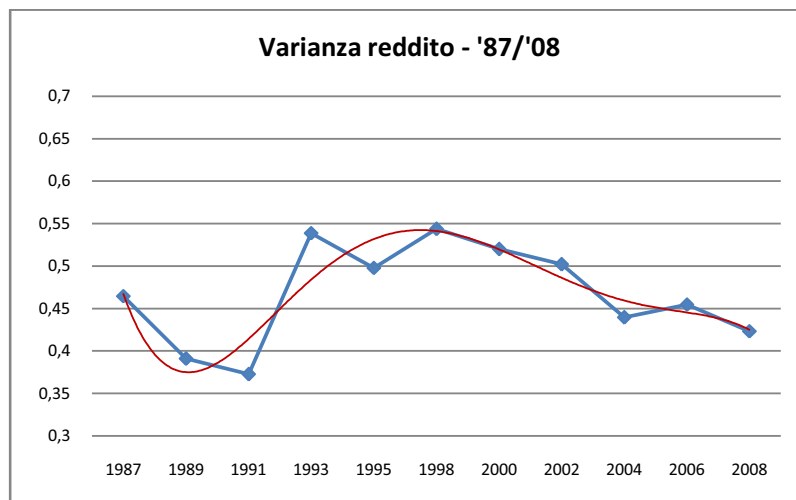
Capitolo 4: I risultati dell'identificazione

In questo capitolo si utilizzeranno i metodi descritti nel capitolo 3 al fine di ottenere delle stime dei parametri che rappresentano il legame tra reddito e consumo, descrivendo sia come tali stime riflettano le caratteristiche dei consumatori italiani, sia come possano essere influenzate dalla natura dei dati utilizzati.

4.a. Analisi descrittive delle varianze

Nel grafico sottostante si può notare l'andamento delle varianze di (log) reddito e consumo a partire dal 1987, anno di revisione del disegno campionario dello SHIW, fino al 2008¹⁴.

¹⁴ Nelle analisi descrittive la limitazione nell'uso dei dati a causa dell'irregolarità della rilevazione può venire ignorata.



Considerando le varianze di reddito e consumo nel tempo si possono osservare due caratteristiche: la varianza del reddito è sistematicamente superiore alla varianza del consumo, e il consumo reagisce agli shock del reddito in maniera limitata.

L'Italia, nel periodo in esame, è andata incontro a profonde modifiche, sia nel mercato del lavoro, sia dal punto di vista economico. Gli anni '90 sono emblematici da questo punto di vista: una serie di riforme del lavoro, tra cui l'abolizione della "scala mobile" nel '92, che legava i salari all'aumento dei prezzi al consumo, la deregolamentazione dei contratti a tempo determinato,

e l'estensione del loro utilizzo da parte delle agenzie di lavoro temporaneo (pacchetto Treu del 1997 e riforma Biagi del 2003) è finalizzata all'aumento sostanziale della flessibilità del lavoro. Tali propositi sono stati raggiunti: l'indice di Protezione Legislativa del Lavoro (Employment Protection Legislation; EPL) diminuisce da un valore di 3,6 della fine degli anni '80, ad un valore di 1,9 nel 2003 e 1,69 nel 2008 (siamo tuttavia ben lontani dal valore di 0,56 degli Stati Uniti)¹⁵

Anche dal punto di vista economico gli anni '90 sono stati un periodo di profondi mutamenti, cominciando dalla riforma bancaria del '93, che incrementò la competitività tra istituti finanziari, la serie di privatizzazioni avvenute in quegli anni, fino all'introduzione dell'Euro nel '99, che rappresenta una liberalizzazione del mercato finanziario, allargando la competizione di questo mercato a livello europeo.

Gli anni 2000, fino al 2008, a parte la recessione degli anni 2001-2004, di entità inferiore rispetto a quella del 1992-1993, rappresentano un periodo di relativa stabilità (o stagnazione) economica, come si può notare dalla graduale diminuzione delle varianze di reddito e consumo; osservando la natura di queste grandezze nel periodo che andremo a esaminare, cioè quello 1998-2008, ci si aspetta che la maggior parte degli shock di reddito in quegli anni sia di natura transitoria, e che esista assicurazione del consumo contro shock di reddito, sebbene non si riesca a determinarne l'entità solamente attraverso analisi descrittive.

¹⁵ Fonte: Organization for Economic Co-operation and Development (OECD)

4.b. Trattamento dei dati

I dati che andremo a utilizzare provengono dalle indagini SHIW della Banca d'Italia dal 1998 al 2008. Si considerano il reddito da lavoro, da pensione e da impresa e si vanno quindi ad escludere i rendimenti derivanti da attività finanziarie. Il consumo di beni non durevoli, invece, viene definito come il totale dei consumi sostenuti in un anno meno quelli derivati dall'acquisto di oggetti preziosi, automobili, i contributi e le donazioni, gli eventuali affitti per l'abitazione o rate di mutui ipotecari e i premi pagati per l'assicurazioni sulla vita o la pensione integrativa. Tale campione è composto in totale da 27.532 famiglie, della quale 8446 formano la componente panel (cioè quelle famiglie che sono state intervistate più di una volta). Dopo che i dati relativi al 1998 e al 2000 sono stati convertiti in Euro (erano espressi in migliaia di lire), si standardizzano a valori 2009 tutti i valori di reddito e consumo, secondo la tabella per la traduzione dei valori monetari fornita dall'ISTAT:

Valori 2009 di un €			
1998	1,2565	2004	1,0974
2000	1,2061	2006	1,0579
2002	1,1467	2008	1,0075

Coefficienti per tradurre valori in Euro in valori Euro 2009; fonte: ISTAT

Prendendo spunto da diversi paper precedenti che hanno affrontato le tematiche della decomposizione degli shock di reddito (Hall, Mishkin, 1982; Blundell, Pistaferri Preston, 2008; Jappelli, Pistaferri, 2010) si decide di considerare solamente alcune unità statistiche, scelte in base a caratteristiche demografiche, al fine di escludere eventi come divorzio, vedovanza o altre cause che pregiudichino la stabilità del nucleo familiare.

Vengono così creati due sottocampioni: il primo campione considera solo le unità statistiche che non hanno cambiato il capofamiglia durante tutto il periodo di permanenza del panel, con capofamiglia di sesso maschile e continuamente sposato in tutte le interviste alla quale ha partecipato. Tale campione è utilizzato per eseguire un confronto coi risultati ottenuti da altri lavori che eseguono analisi simili a quelle presentate in questa tesi su dati italiani (Jappelli & Pistaferri, 2010; Weber & Borella, 2011).

Il secondo campione comprende le unità del primo gruppo di età compresa tra i 30 ed i 60 anni; questo per poter confrontare i risultati ottenuti con lavori che eseguono analisi delle dinamiche del reddito e del consumo su dati di paesi esteri, ed usano un campione dalle caratteristiche simili a questo gruppo (Blundell & Pistaferri & Preston, 2008; Gorodnichenko & Peter & Stolyarov, 2010).

Le numerosità campionarie dei vari sottoinsiemi sono osservabili nella tabella sottostante.

	n° famiglie	n° famiglie panel
Campione completo	27532	8446
Con stesso capofamiglia	26085	7039
Di sesso maschile	16976	4788
Continuativamente sposato (primo gruppo)	13725	3833
Di età compresa tra i 30 ed i 60 anni (secondo gruppo)	8831	2633

4.c. Generazione dei residui

Una volta ridotto il campione a quelle famiglie dotate di una certa stabilità, si procede ad eliminare dai valori di (log)reddito e consumo delle famiglie tutti gli effetti delle caratteristiche demografiche. Per fare ciò, si regredisce il reddito (e il consumo) su: età (e il suo quadrato), numero di percettori di reddito da lavoro (e il suo quadrato), numero di percettori di pensione (e il suo quadrato), numero di membri della famiglia suddivisi in classi di età e dummy per regione di residenza, titolo di studio e tipo di occupazione (lavoratore autonomo o dipendente) del capofamiglia e disoccupazione o non-occupazione del capofamiglia.

Poiché si dispone di dati di panel sia di reddito che di consumo, al fine di ottenere stime che tengano conto della dimensione temporale dei dati, si utilizzano modelli di regressione ad effetti fissi e ad effetti casuali, e si testerà la non distorsione di quest'ultima regressione attraverso il test di Hausman descritto nel capitolo 3.

I risultati della regressione ad effetti casuali sul reddito delle famiglie del primo gruppo¹⁶ sono i seguenti :

```

Random-effects GLS regression              Number of obs   =   22303
Group variable: nquest                    Number of groups =   13686

R-sq:  within = 0.2399                    Obs per group:  min =    1
        between = 0.5666                  avg   =    1.6
        overall = 0.5489                  max   =    6

Random effects u_i ~ Gaussian             Wald chi2(19)   =  20578.73
corr(u_i, X) = 0 (assumed)                Prob > chi2     =    0.0000

```

ly2	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
eta	.033641	.0027895	12.06	0.000	.0281736 .0391085
eta2	-.0002244	.0000253	-8.87	0.000	-.000274 -.0001748
a_N	.050636	.0094256	5.37	0.000	.0321622 .0691098
a_S	-.2457266	.0100807	-24.38	0.000	-.2654844 -.2259688
s_medie	.2205724	.0094848	23.26	0.000	.2019826 .2391622
s_sup	.4465924	.0098473	45.35	0.000	.427292 .4658928
s_lau	.7748176	.0135889	57.02	0.000	.7481839 .8014512
q_aut	.117687	.0093345	12.61	0.000	.0993917 .1359823
nperl	.5104178	.0106956	47.72	0.000	.4894548 .5313808
nperl2	-.0571141	.0029127	-19.61	0.000	-.0628229 -.0514052
pensionati	.2346142	.0070176	33.43	0.000	.22086 .2483684
h_disocc	-.5537659	.0170996	-32.38	0.000	-.5872805 -.5202513
nonocc	-.1636055	.054252	-3.02	0.003	-.2699376 -.0572735
figli_minori	.0071591	.0047791	1.50	0.134	-.0022077 .016526
nm_18_26	.0054616	.0055112	0.99	0.322	-.0053401 .0162633
nm_27_40	.0254924	.0065317	3.90	0.000	.0126905 .0382942
nm_41_60	.0433749	.0089818	4.83	0.000	.0257709 .0609789
nm_61_70	.0445239	.0109067	4.08	0.000	.0231471 .0659007
nm_71	.0250844	.0126309	1.99	0.047	.0003283 .0498404
_cons	8.290902	.0704912	117.62	0.000	8.152742 8.429062
sigma_u	.3167329				
sigma_e	.29300877				
rho	.53884968	(fraction of variance due to u_i)			

¹⁶ Le regressioni su reddito delle famiglie appartenenti al secondo gruppo non danno risultati significativamente differenti rispetto a quelli esposti.

I risultati della regressione ad effetti fissi sono invece i seguenti (le variabili di area geografica di residenza vengono omesse in quanto invarianti nel tempo):

```

Fixed-effects (within) regression      Number of obs   =   22303
Group variable: nquest                Number of groups =   13686

R-sq:  within = 0.2531                Obs per group:  min =    1
        between = 0.3223                avg   =    1.6
        overall = 0.3214                max   =    6

corr(u_i, Xb) = 0.0784                F(17,8600)     =   171.44
                                         Prob > F       =    0.0000

```

ly2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
eta	.050891	.0062255	8.17	0.000	.0386875	.0630944
eta2	-.0003166	.0000548	-5.77	0.000	-.0004241	-.0002091
s_medie	.0187007	.0352524	0.53	0.596	-.0504025	.0878039
s_sup	.077346	.04056	1.91	0.057	-.0021614	.1568533
s_lau	.0882766	.0690252	1.28	0.201	-.0470294	.2235826
q_aut	.1062624	.0204829	5.19	0.000	.0661111	.1464137
nperl	.4328388	.0161301	26.83	0.000	.4012198	.4644577
nperl2	-.0488973	.0042295	-11.56	0.000	-.0571882	-.0406065
pensionati	.1831373	.011738	15.60	0.000	.160128	.2061466
h_disocc	-.4109586	.0252272	-16.29	0.000	-.4604099	-.3615073
nonocc	-.1586306	.0723667	-2.19	0.028	-.3004866	-.0167746
figli_minori	.0198149	.0119033	1.66	0.096	-.0035185	.0431482
nm_18_26	.0277214	.0105386	2.63	0.009	.0070632	.0483795
nm_27_40	.0415231	.0115173	3.61	0.000	.0189464	.0640997
nm_41_60	.048006	.0153997	3.12	0.002	.0178189	.0781932
nm_61_70	.0499876	.0183789	2.72	0.007	.0139607	.0860146
nm_71	.0508173	.0212696	2.39	0.017	.0091239	.0925107
_cons	7.890279	.1716373	45.97	0.000	7.553828	8.226729
sigma_u	.51877363					
sigma_e	.29300877					
rho	.75814389	(fraction of variance due to u_i)				
F test that all u_i=0:			F(13685, 8600) =	2.85	Prob > F = 0.0000	

Il test di Hausman di differenza dei coefficienti delle due regressioni rifiuta l'ipotesi nulla:

```

----- Coefficients -----
      |          (b)          (B)          (b-B)          sqrt(diag(V_b-V_B))
      | Fixed eff.  Random eff.  Difference          S.E.
-----+-----
      eta |      .050891      .033641      .0172499      .0055655
      eta2 |     -.0003166     -.0002244     -.0000922      .0000487
      s_medie |    .0187007    .2205724     -.2018717      .0339525
      s_sup |    .077346     .4465924     -.3692464      .0393465
      s_lau |    .0882766     .7748176     -.686541      .0676744
      q_aut |    .1062624     .117687      -.0114246      .0182322
      nperl |    .4328388     .5104178     -.077579      .0120742
      nperl2 |   -.0488973     -.0571141      .0082167      .0030667
      pensionati | .1831373     .2346142     -.0514769      .0094093
      h_disocc |  -.4109586     -.5537659      .1428073      .0185476
      nonocc |  -.1586306     -.1636055      .0049749      .0478921
      figli_minori | .0198149     .0071591      .0126558      .0109018
      nm_18_26 | .0277214     .0054616      .0222598      .0089827
      nm_27_40 | .0415231     .0254924      .0160307      .009486
      nm_41_60 | .048006      .0433749      .0046311      .0125091
      nm_61_70 | .0499876     .0445239      .0054637      .0147927
      nm_71 | .0508173     .0250844      .0257329      .017113
-----+-----
      b = consistent under Ho and Ha
      B = inconsistent under Ha, efficient under Ho

Test:  Ho:  difference in coefficients not systematic
      chi2(17) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
              =      331.27
      Prob>chi2 =      0.0000

```

In entrambe le regressioni, la differenza di coefficienti tra le due regressioni è significativa, come dimostrano il test di Hausman eseguito. Questo porterebbe ad escludere i coefficienti ottenuti attraverso lo stimatore ad effetti casuali, in quanto distorti. Tuttavia, alcuni dei regressori assumono valori insoliti rispetto a quello che sarebbe plausibile aspettarsi : ad esempio, nella regressione del reddito, i coefficienti relativi al titolo di studio del capofamiglia assumono valori insolitamente bassi e, soprattutto, non significativi, così come per il sesso.

Osservando l'equazione di stima della regressione ad effetti fissi si nota che le stime sono ottenute attraverso una regressione OLS sulle

deviazioni dalle medie dei dati. Questo significa che la regressione tiene conto solamente della parte panel del campione (le famiglie non panel, infatti, avendo una sola sola osservazione, hanno, ovviamente deviazione dalla media nulla). Poichè tale regressione comprende solo le deviazioni dalle medie, una grossa fetta del campione verrà esclusa, in quanto i parametri relativi ad alcune caratteristiche demografiche non variano nel tempo (un esempio di variazione sono le famiglie in cui è mutato il numero di membri a causa della nascita di un figlio, o perchè un figlio adulto ha abbandonato la casa dei genitori; si ricorda, inoltre, che dal campione sono esclusi tutti i casi di divorzio o vedovanza).

Di conseguenza, la differenza dei coefficienti nelle due regressioni non è dovuta alla distorsione della regressione ad effetti casuali, ma al fatto che le due regressioni usano campioni diversi e misurano cose diverse; ad esempio, la regressione ad effetti fissi mette in evidenza come l'effetto sul reddito di un titolo di studio, conseguito quando si è già inseriti nel mondo del lavoro (i capofamiglia presenti in questo campione hanno un'età minima di 30 anni), sia molto inferiore all'effetto di un titolo conseguito in età scolare.

Per quanto concerne il titolo di studio, bisogna anche rilevare l'esistenza di diversi errori di misurazione in tale variabile. Delle 848 famiglie che negli anni hanno dichiarato di aver "variato" il titolo di studio, quasi il 20% dichiara che il suo grado di istruzione è calato (da diploma di scuola superiore a licenza media, ad esempio), specialmente in famiglie con capofamiglia molto anziano. All'opposto, ci sono casi nei quali il capofamiglia dichiara di aver conseguito in 2 anni (intervallo tra le due interviste) ben 2 o 3 titoli di studio (come passare da non avere nessun titolo a conseguire il diploma di scuola superiore). Tali dichiarazioni sono

evidentemente frutto di incomprensioni della domanda o errori di misurazione; questa problematica risulta essere molto grave nei dati relativi al 2008: come si può notare nel grafico sottostante, di coloro che dichiarano di aver cambiato titolo di studio rispetto al 2006, più del 75% dichiara di aver perso un titolo (*sic!*).



Per questo motivo, si decide di escludere dal campione tutte quelle famiglie che dichiarano di aver perso un titolo di studio o di averne guadagnati 2 o più tra due interviste. Questo elimina dal campione 99 famiglie. Le stime ottenute dal campione che esclude queste famiglie, tuttavia, non subiscono variazioni significative, dato lo scarso numero di queste ultime rispetto al totale del campione.

Alla luce di tali fatti, si decide di utilizzare per la stima delle varianze e l'identificazione del modello i residui provenienti dalla regressione ad effetti casuali.

4.d. Stima delle varianze

Una volta ottenuti i residui a partire dalla stima ad effetti casuali, si calcolano le varianze di reddito e consumo ai vari ritardi. Le varianze e covarianze ai ritardi 1 e 2 del reddito sono visibili nella tabella sottostante (*standard error* tra parentesi):

	$Var[\Delta y_t]$	$Cov[\Delta y_{t+1}\Delta y_t]$	$Cov[\Delta y_{t+2}\Delta y_t]$
2000	0,2065 (0,0223)	-0,0752 (0,0186)	0,0130 (0,0077)
2002	0,1680 (0,0231)	-0,0662 (0,0086)	-0,0130 (0,0065)
2004	0,1462 (0,0101)	-0,0515 (0,0073)	-0,0011 (0,0049)
2006	0,1452 (0,0110)	-0,0562 (0,0095)	*** **
2008	0,1496 (0,0312)	*** **	*** **

L'andamento delle varianze degli shock di reddito indica come queste, da valori abbastanza alti del 2000, cali drasticamente fino al 2004, per poi mantenersi stabile. La covarianza al ritardo 1, nel modello economico preso in considerazione (shock di reddito transitorio *white noise*) rappresenta la varianza dello shock transitorio del reddito ($cov[\Delta y_{i,t+1}\Delta y_{i,t}] = -\sigma_{\varepsilon_t}^2$); dati i valori di questa tabella e, in linea con quanto osservato nell'evoluzione della varianza del reddito e del consumo nel paragrafo precedente, ci si aspetta che tale varianza sia più alta di quella dello shock permanente. L'autocovarianza al ritardo 2 dovrebbe indicare la presenza di autocorrelazione nel processo dello shock transitorio, nel caso in cui si supponga che la parte transitoria del reddito segua un processo MA(1); tuttavia, tali covarianze risultano essere nella maggior parte dei casi non

significative, di conseguenza si assumerà l'incorrelazione della parte transitoria del reddito con il suo passato.

I momenti del reddito sono sufficienti per fare inferenza sulle variazioni del reddito (e sulla natura persistente o transitoria di tali variazioni); tuttavia, essi non sono sufficienti per arrivare a delle conclusioni sulle variazioni del consumo. Per fare ciò, è necessario osservare le varianze e autocovarianze del consumo, nonché le covarianze tra reddito e consumo ai vari ritardi.

	$Var[\Delta c_t]$	$Cov[\Delta c_{t+1}\Delta c_t]$	$Cov[\Delta c_{t+2}\Delta c_t]$
2000	0,1321 (0,0069)	-0,0393 (0,0048)	0,0010 (0,0045)
2002	0,1330 (0,0072)	-0,0717 (0,0069)	-0,0016 (0,0040)
2004	0,1432 (0,0074)	-0,0565 (0,0050)	-0,0081 (0,0042)
2006	0,1239 (0,0068)	-0,0410 (0,0050)	*** **
2008	0,0914 (0,0043)	*** **	*** **

Com'era plausibile aspettarsi dai valori delle varianze del consumo osservati all'inizio del capitolo, anche i valori delle varianze delle differenze prime del consumo sono inferiori rispetto a quelli delle differenze prime del reddito; inoltre, la loro variazione nel tempo è molto più graduale rispetto a quella che si è osservata per il reddito: le varianze aumentano gradualmente fino al 2004 per poi scendere negli anni successivi. La varianza degli shock del consumo rappresenta, oltre al collegamento degli shock del consumo agli shock di reddito, anche il contributo di shock delle preferenze dei consumatori e dell'errore di misurazione. Tale errore è rappresentato dalle

covarianze al ritardo 1; esso risulta essere molto grande¹⁷, a parte per l'anno 2000, circostanza di cui bisogna tener conto nelle analisi successive. Le covarianze ai ritardi 2 e successivi sono statisticamente non significative ed economicamente non rilevanti, questo in accordo con le ipotesi esposte nel capitolo precedente, che impongono la nullità di questi valori.

Le covarianze tra reddito e consumo nei vari periodi a vari ritardi sono le seguenti:

	$Cov[\Delta y_t \Delta c_t]$	$Cov[\Delta y_{t+1} \Delta c_t]$	$Cov[\Delta y_t \Delta c_{t+1}]$
2000	0,0560 (0,0051)	-0,0189 (0,0046)	-0,0210 (0,0053)
2002	0,0535 (0,0062)	-0,0274 (0,0062)	-0,0247 (0,0056)
2004	0,0552 (0,0058)	-0,0230 (0,0047)	-0,0223 (0,0046)
2006	0,0564 (0,0052)	-0,0172 (0,0037)	-0,0204 (0,0042)
2008	0,0413 (0,0038)	*** ***	*** ***

La covarianza tra variazioni del reddito e del consumo contemporanee dà informazioni sull'effetto degli shock di reddito sul consumo, nel caso che gli errori di misura del reddito e del consumo siano indipendenti tra di loro; tale covarianza rimane stabile per tutto il periodo considerato, per poi calare nel 2008.

La covarianza tra la variazione del consumo attuale e la variazione futura del reddito riflette il livello di assicurazione rispetto a shock di reddito transitori. Nel caso del modello del reddito permanente di Friedman, tale covarianza dovrebbe essere nulla, e valori diversi da zero indicherebbero che

¹⁷ Una stima approssimativa della parte della varianza dello shock del consumo priva dell'errore di misurazione è data sottraendo da questa grandezza 2 volte la covarianza al ritardo 1 del consumo nello stesso periodo.

gli shock di reddito futuri sono conosciuti dal consumatore, che di conseguenza modificherebbe il suo consumo prima che le variazioni del reddito avvengano.

La covarianza tra la variazione del reddito corrente e la variazione del consumo futura non ha interpretazione nel modello ad assicurazione parziale con mercato del credito perfetto che consideriamo in questa tesi, tuttavia, può avere importanza in specificazioni alternative di questo modello, nelle quali sono presenti vincoli di liquidità¹⁸.

In conclusione, i valori di varianze e covarianze ai vari ritardi delle differenze prime di reddito e consumo rispecchiano il modello economico teorizzato nei capitoli precedenti, caratterizzato da variazioni del reddito transitorie di tipo *white noise* e dalla presenza di errori di misura del consumo. Tuttavia, la significatività della covarianza tra variazioni del consumo presente e variazioni del reddito future ($Cov[\Delta y_{t+1} \Delta c_t]$) suggerisce che, nel caso della teoria del reddito permanente di Friedman, il consumatore possieda informazioni sulle variazioni di reddito future o che, in alternativa, ci sia una non perfetta assicurazione degli shock transitori di reddito da parte dei consumatori. Inoltre, la non nullità della covarianza tra variazioni del reddito attuale e variazioni del consumo future indica la presenza di vincoli di liquidità: tali vincoli limiterebbero la capacità dei consumatori di assicurarsi contro shock transitori del reddito, che si rifletterebbe in un valore del coefficiente di assicurazione maggiore di zero.

¹⁸ si veda Flavin, "The adjustment of Consumption to Changing Expectations About Future Income", per degli approfondimenti in tale ambito.

4.e. Identificazione del modello

Da principio, si sono confrontate le varie specificazioni del modello tra di loro, per stabilire quale sia la più rappresentativa della realtà. Si confrontano i modelli senza errori di misura del consumo, con errore di misura del consumo sia omoschedastica che eteroschedastica e con errore di misurazione del reddito. Ci si è avvalsi dello stesso campione, più generale rispetto a quelli utilizzati successivamente, che esclude dal solo le famiglie che hanno cambiato capofamiglia nel corso delle interviste. Le stime ottenute utilizzando i modelli spiegati sono le seguenti:

n=26085 n(panel)=7022				
Varianza shock permanente del reddito				
$\sigma^2_{\delta,2002}$	0,0444 (0,0080)	0,0465 (0,0082)	0,0388 (0,0094)	0,0357 (0,0094)
$\sigma^2_{\delta,2004}$	0,0516 (0,0091)	0,0542 (0,0093)	0,0425 (0,0102)	0,0398 (0,0102)
$\sigma^2_{\delta,2006}$	0,0310 (0,0051)	0,0341 (0,0055)	0,0395 (0,0097)	0,0373 (0,0102)
Varianza shock transitorio del reddito				
$\sigma^2_{\varepsilon,2000}$	0,0886 (0,0113)	0,0926 (0,0121)	0,0949 (0,0124)	0,1303 (0,0636)
$\sigma^2_{\varepsilon,2002}$	0,0728 (0,0084)	0,0725 (0,0085)	0,0750 (0,0082)	0,1104 (0,0576)
$\sigma^2_{\varepsilon,2004}$	0,0593 (0,0047)	0,0590 (0,0048)	0,0621 (0,0047)	0,0973 (0,0583)
$\sigma^2_{\varepsilon,2006}$	0,0642 (0,0069)	0,0645 (0,0072)	0,0585 (0,0092)	0,0927 (0,0592)

Coefficiente di assicurazione contro shock permanenti del reddito				
Φ	1,0180	1,0125	0,9998	0,9774
	(0,1626)	(0,1596)	(0,2390)	(0,2510)
Coefficiente di assicurazione contro shock transitori del reddito				
ψ	0,2674	0,2251	0,2628	0,2111
	(0,0301)	(0,0286)	(0,0344)	(0,1173)
Varianza shock delle preferenze dei consumatori				
σ^2_{φ}	0,0812	-0,0250	-0,0185	-0,0364
	(0,0065)	(0,0083)	(0,0107)	(0,0575)
Varianza dell'errore di misura del consumo				
$\sigma^2_{u,2000}$		0,0524	0,0526	0,0507
		(0,0022)	(0,0030)	(0,0032)
$\sigma^2_{u,2002}$			0,0658	0,0643
			(0,0049)	(0,0051)
$\sigma^2_{u,2004}$			0,0624	0,0579
			(0,0039)	(0,0039)
$\sigma^2_{u,2006}$			0,0390	0,0446
			(0,0024)	(0,0027)
Varianza dell'errore di misura del reddito				
σ^2_v				-0,0081
				(0,0117)
test χ^2	2779,30	304,90	236,98	555,70
p-value	(0,0000)	(0,0000)	(0,0000)	(0,0000)

Si può notare che la varianza degli shock delle preferenze dei consumatori diventa negativo, anche se non significativo, introducendo l'errore di misurazione del consumo. Non imponendo ulteriori restrizioni sui parametri, oltre a quelle ricavate dal modello economico, è possibile che

succeda¹⁹; tuttavia, anche imponendo la positività di tale parametro, si otterrebbero risultati non molto diversi da quelli qui presentati²⁰; ad ogni modo, le cause che possono portare a questo risultato sono molteplici, come un'errata specificazione del modello o degli errori di misura del consumo, o come una distorsione dei dati sul consumo utilizzati; la varianza degli shock delle preferenze dei consumatori e le varianze degli errori di misura del consumo appaiono solo nelle equazioni della varianza e dell'autocovarianza al ritardo 1 del consumo, di conseguenza, eventuali errori di misura del reddito non influenzerebbero queste stime. Come vedremo in seguito, i dati sul consumo soffrono di problemi di arrotondamento e accumulazione, che causano i problemi di cui sopra.

Confrontando i modelli che prevedono la presenza di soli errori di misurazione del consumo, si può notare che le varianze di tali errori di misura sono significativamente diverse tra di loro; inoltre, supporre l'omoschedasticità degli stessi modifica le stime, abbassando quella di ψ e rendendo la varianza degli shock di preferenze dei consumatori significativamente minore di 0. Questi fatti ci indicano che tra i due modelli bisogna scegliere quello con errore di misura del consumo eteroschedastico.

L'errore di misurazione del reddito, una volta introdotto all'interno del modello, si rivela essere non significativo in termini di varianza. Peraltro, sono presenti evidenti distorsioni nella stima dei parametri, specialmente per quel che riguarda il coefficiente di assicurazione contro shock transitori e la

¹⁹ Al pari, ad esempio, dei modelli ARCH e GARCH in assenza di restrizioni di positività dei parametri.

²⁰ Blundell et al. (2008) stimano, usando dati americani provenienti dal PSID, una varianza delle variazioni delle preferenze dei consumatori pari a 0.0105

varianza degli stessi shock transitori del reddito, quindi si decide di non utilizzarlo.

Tra i due modelli che restituiscono valori dei parametri simili e che non mostrano segni di distorsione, cioè quello senza errori di misura e quello con errori di misura del consumo eteroschedastici, si sceglie quello con errori di misura, in quanto più informativo rispetto al processo del consumo, anche se in presenza di dati affetti da problemi di arrotondamento e accumulazione.

Tali risultati sembrano essere in linea con altri lavori che hanno affrontato il medesimo problema e che hanno usato dati italiani, i cui valori dei coefficienti di assicurazione sono mostrati nella tabella sottostante²¹.

In tutti e tre i casi, il coefficiente di assicurazione contro shock permanenti non è significativamente diverso da 1; ciò significa che le famiglie italiane non sono in grado di assicurarsi contro shock permanenti del reddito e questi si trasmettono integralmente al consumo. Questo è esattamente quello che prevede la teoria del reddito permanente di Friedman; tuttavia, questa teoria prevede che non esistano meccanismi di assicurazione di alcun tipo, quando, invece nella realtà questi metodi esistono, come le assicurazioni private o il *welfare* statale. Tali stime riflettono gravi inefficienze dello stato del *welfare* nel proteggere i consumatori contro eventi inattesi.

A contrario di quanto esposto nella teoria del reddito permanente, le stime indicano che non esiste completa assicurazione contro shock transitori del

²¹ Le varianze degli shock di reddito non vengono confrontate in quanto vengono utilizzate diversi metodi di identificazione dei parametri: Attanasio et al. identificano gli standard error, e non le varianze; invece Jappelli e Pistaferri stimano queste varianze in maniera non parametrica, come verrà esposto in Appendice.

reddito; poiché tali shock vengono assicurati dai consumatori attraverso il risparmio personale, un valore diverso da zero indica che esistono dei vincoli di liquidità. Questo non è inatteso: l'Italia ha storicamente un mercato del credito molto rigido e una scarsa propensione all'indebitamento.

	Tesi	Attanasio Borella & Nielsen	Jappelli & Pistaferri
Coefficiente di assicurazione contro shock permanenti del reddito			
Φ	1,000 (0,239)	1,123 (0,242)	0,989 (0,056)
Coefficiente di assicurazione contro shock transitori del reddito			
ψ	0,263 (0,034)	0,240 (0,027)	0,282 (0,081)

4.e.1. L'effetto dell'istruzione e dell'anzianità

Per determinare se alcune caratteristiche, come l'età o il titolo di studio, influenzano la capacità di assicurarsi contro shock di reddito delle famiglie, si identificano i parametri del modello scelto nel paragrafo precedente, usando dei sottoinsiemi dei dati del primo gruppo, rispettivamente facendo due separazioni: la prima tra coloro che non hanno cambiato titolo di studio in tutte le interviste effettuate in possesso di diploma di scuola superiore e laurea e coloro con al massimo la licenza media; la seconda tra le famiglie con capofamiglia che al momento di

ingresso nel panel avevano un'età compresa tra i 30 e i 60 anni e quelle con capofamiglia con più di 60 anni²².

I risultati della stima dei parametri del modello con errore di misura del consumo eteroschedastico, sono visibili nella seguente tabella:

Campione di famiglie con capofamiglia invariato, di sesso maschile e sposato					
	c. completo	dipl. o laurea	lic. media	età 30-60	età 61+
n=	13725	5462	8263	8831	4576
n(panel)=	3833	1633	2200	2633	1126
Varianza shock permanente del reddito					
$\sigma^2_{\delta,2002}$	0,0343 (0,0083)	0,0351 (0,0118)	0,0399 (0,0115)	0,0354 (0,0107)	0,0423 (0,0110)
$\sigma^2_{\delta,2004}$	0,0324 (0,0077)	0,0298 (0,0090)	0,0397 (0,0118)	0,0339 (0,0105)	0,0344 (0,0087)
$\sigma^2_{\delta,2006}$	0,0283 (0,0064)	0,0339 (0,0109)	0,0301 (0,0086)	0,0282 (0,0080)	0,0303 (0,0083)
Varianza shock transitorio del reddito					
$\sigma^2_{\varepsilon,2000}$	0,0748 (0,0099)	0,0640 (0,0085)	0,0834 (0,0142)	0,0909 (0,0139)	0,0423 (0,0065)
$\sigma^2_{\varepsilon,2002}$	0,0638 (0,0088)	0,0619 (0,0151)	0,0592 (0,0104)	0,0685 (0,0108)	0,0320 (0,0139)
$\sigma^2_{\varepsilon,2004}$	0,0559 (0,0058)	0,0513 (0,0085)	0,0618 (0,0066)	0,0618 (0,0078)	0,0544 (0,0069)
$\sigma^2_{\varepsilon,2006}$	0,0573 (0,0098)	0,0452 (0,0097)	0,0526 (0,0123)	0,0565 (0,0097)	0,0320 (0,0075)

²² Poiché si stratifica il campione in base a caratteristiche esogene e si stimano parametri diversi per ogni gruppo, si considera la capacità di assicurazione contro shock del reddito all'interno dei gruppi.

Coefficiente di assicurazione contro shock permanenti del reddito					
Φ	1,0106	1,0177	1,0052	1,0012	1,0416
	(0,2350)	(0,3230)	(0,2823)	(0,2811)	(0,2652)
Coefficiente di assicurazione contro shock transitori del reddito					
ψ	0,3178	0,2580	0,2626	0,2590	0,2854
	(0,0439)	(0,0684)	(0,0520)	(0,0440)	(0,0860)
Varianza shock del consumo					
Varcsi	-0,0127	-0,0153	-0,0136	-0,0069	-0,0316
	(0,0096)	(0,0147)	(0,0121)	(0,0110)	(0,0165)
Errore di misura del consumo					
$\sigma^2_{u,2000}$	0,0463	0,0569	0,0371	0,0452	0,0492
	(0,0037)	(0,0063)	(0,0042)	(0,0045)	(0,0063)
$\sigma^2_{u,2002}$	0,0631	0,0641	0,0605	0,0615	0,0673
	(0,0064)	(0,0097)	(0,0084)	(0,0076)	(0,0123)
$\sigma^2_{u,2004}$	0,0584	0,0516	0,0626	0,0538	0,0710
	(0,0047)	(0,0066)	(0,0066)	(0,0051)	(0,0104)
$\sigma^2_{u,2006}$	0,0362	0,0394	0,0333	0,0363	0,0410
	(0,0032)	(0,0047)	(0,0041)	(0,0037)	(0,0063)
test χ^2	164,99	100,20	114,07	110,13	104,90
p-value	0,0000	0,0000	0,0000	0,0000	0,0000

Come si evince da questi risultati, non sembrano esserci differenze significative nella capacità di assicurazione tra coloro con un titolo di studio elevato o meno²³ o tra anziani e meno anziani. La differenza sta soprattutto nel tipo di shock che subiscono le varie categorie. In genere, le famiglie con un capofamiglia più istruito sono meno soggette a shock di reddito, e questa caratteristica è ancora più evidente se si

²³ Stime effettuate suddividendo il campione in laureati e non laureati non producono risultati significativamente diversi.

isolano le famiglie con capofamiglia di età compresa tra i 30 e i 60 anni; sembra, infatti che coloro che hanno un titolo di studio riescano a tutelare il proprio posto di lavoro meglio di chi non ha un pari livello di istruzione. La differenza degli shock di reddito tra chi ha più o meno di 60 anni sta invece nella natura dello shock: chi è più anziano è soggetto a shock permanenti più forti e a shock transitori meno incisivi rispetto a chi ha meno di 60 anni. Questo può essere interpretato come l'effetto del pensionamento: il ritiro dal lavoro è uno shock permanente che avviene principalmente dopo i 60 anni di età e, successivamente alla pensione, il reddito che si percepisce, sebbene minore, è sicuramente meno soggetto a variazioni improvvise rispetto al reddito da lavoro. Al contrario, coloro con meno di 60 anni sono in genere ancora in attività, di conseguenza non sono soggetti allo shock permanente rappresentato dalla pensione, ma il reddito da lavoro è molto più incerto: da qui la differenza che si può notare nella tabella soprastante.

4.e.2. Un confronto internazionale

E' interessante confrontare le reazioni agli shock di reddito di famiglie di paesi diversi; tale confronto si può ottenere stimando le propensioni marginali del consumo di dati italiani e stranieri su un campione di famiglie dalle caratteristiche simili. In un lavoro di Blundell Pistaferri e Preston (2008), vengono stimati questi fattori usando dati sulle famiglie americane.

Si sono, quindi, stimati i valori di tali coefficienti anche su dati italiani, usando un campione il più simile a quello da loro utilizzato, cioè famiglie che non hanno mai cambiato capofamiglia, continuamente sposate per tutta la durata di presenza nel panel, con capofamiglia di sesso maschile di

età compresa tra i 30 ed i 60 anni. La tabella in basso descrive i risultati delle stime su questo campione, e su due sottocampioni divisi in base al titolo di studio come nel precedente paragrafo, raffrontati ai valori ottenuti da Blundell, Pistaferri e Preston su dati americani²⁴:

	Camp. Completo		Alta Istruzione		Bassa Istruzione	
	Italia	USA	Italia	USA	Italia	USA
<hr/>						
Coefficiente di assicurazione contro shock permanenti del reddito						
Φ	1,011	0,642	1,018	0,419	1,005	0,944
	(0,235)	(0,095)	(0,323)	(0,092)	(0,282)	(0,178)
<hr/>						
Coefficiente di assicurazione contro shock transitori del reddito						
ψ	0,318	0,053	0,258	0,077	0,263	0,027
	(0,044)	(0,044)	(0,068)	(0,060)	(0,052)	(0,055)
<hr/>						

Questi risultati indicano che in genere, le famiglie americane hanno una capacità molto superiore rispetto a quelle italiane di assicurarsi contro shock di reddito, ad ogni modo negli Stati Uniti esiste una marcata

²⁴ Le varianze degli shock di reddito non vengono confrontate in quanto relative a periodi differenti; i coefficienti di assicurazione possono essere raffrontati in quanto Blundell et al. accettano l'ipotesi che tali coefficienti siano invariati nel tempo. Il lavoro di Jappelli e Pistaferri (2010) i cui risultati sono visibili nella tabella a pag. yy, sono ricavati a partire da dati che vanno dal 1987 a oggi e non mostrano sostanziali differenze rispetto a quelli ottenuti in questa tesi, indicando che anche in Italia si può fare l'assunzione di invarianza dei coefficienti di assicurazione dei consumatori contro shock di reddito.

Blundell et al. includono nel loro lavoro anche risultati derivati da una suddivisione del campione in base all'anno di nascita separando coloro nati negli anni '30 da quelli nati negli anni '40, tuttavia una tale suddivisione su dati italiani non sarebbe raffrontabile, in quanto si tratta di individui rilevati in periodi diversi, e nati in periodi diversi (Poiché l'età massima di questo campione è 60 anni e i dati utilizzati risalgono al massimo al 1998, l'anno di nascita più "lontano" sarebbe il 1938).

differenza in questa capacità di assicurarsi tra chi possiede una laurea e chi non la possiede, diversamente da quanto si osserva nelle stime ottenute con i dati italiani.

Tale diversità di comportamento delle famiglie potrebbe avere origine nella natura considerata tradizionalmente prudente degli italiani: essi infatti sono molto meno propensi rispetto agli americani (e anche rispetto al resto dell' Europa occidentale) all'utilizzo di strumenti finanziari e in genere a indebitarsi, preferendo ridurre temporaneamente i consumi, piuttosto che intaccare il risparmio. Questo, unito a un mercato del credito molto rigido, causerebbe la scarsa capacità degli italiani di assicurarsi contro variazioni temporanee del reddito.

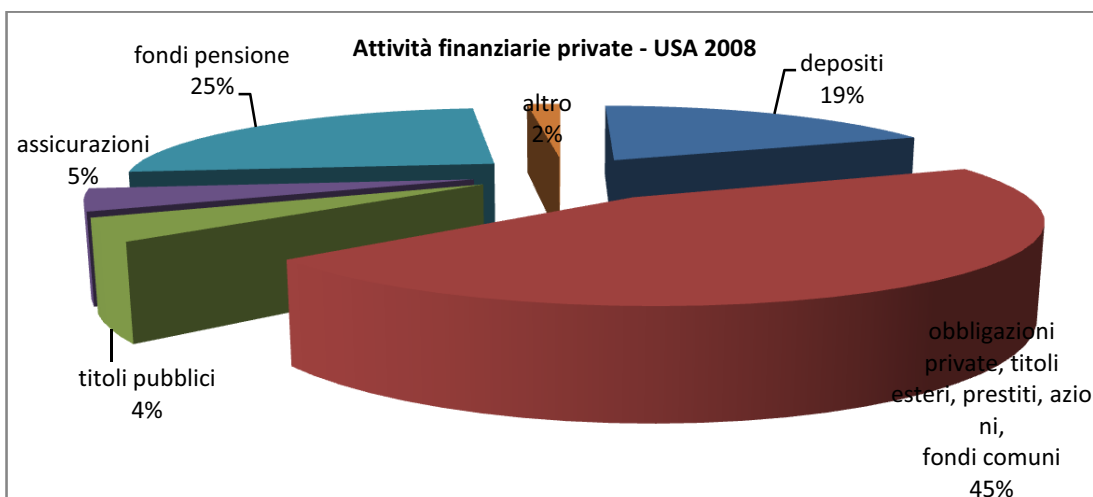
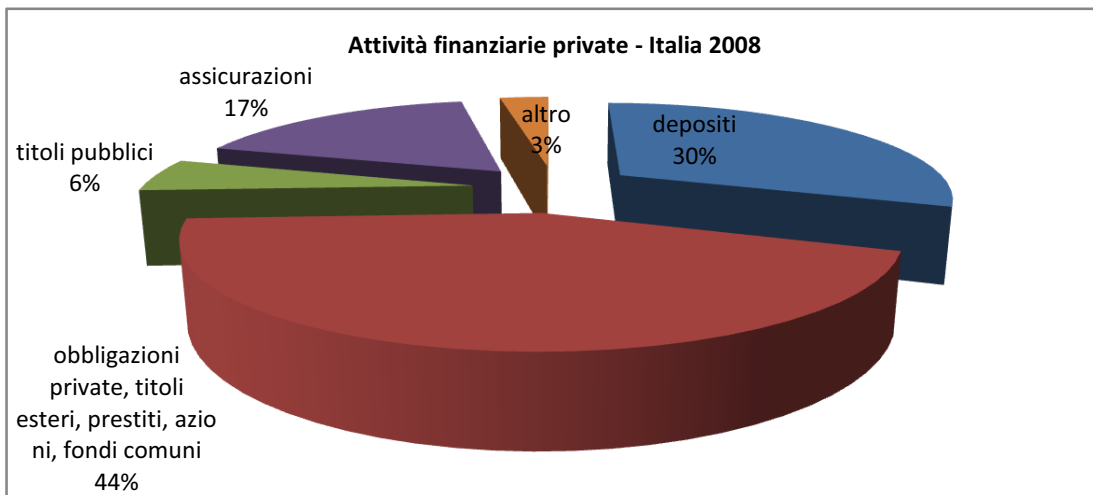
La ricchezza delle famiglie: un confronto internazionale

(valori in rapporto al reddito disponibile)

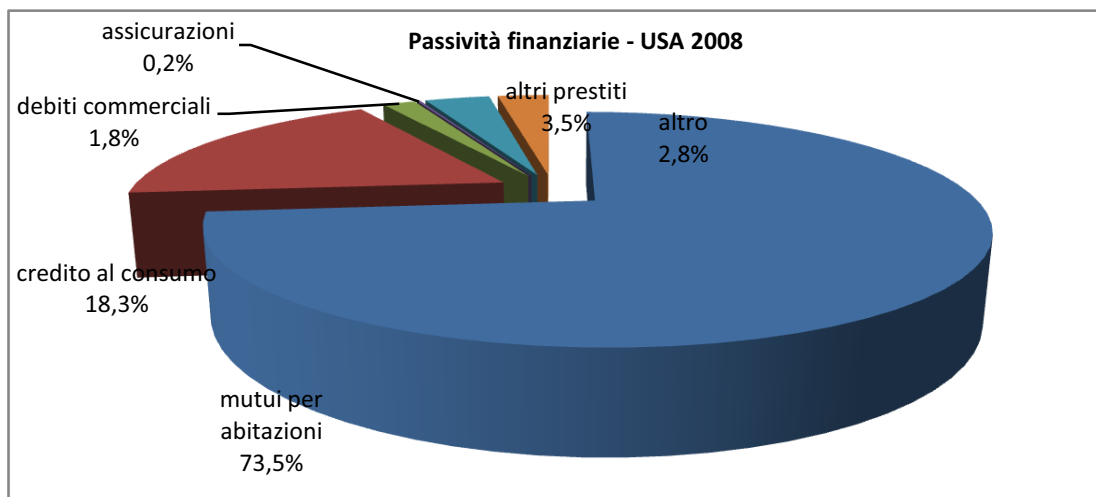
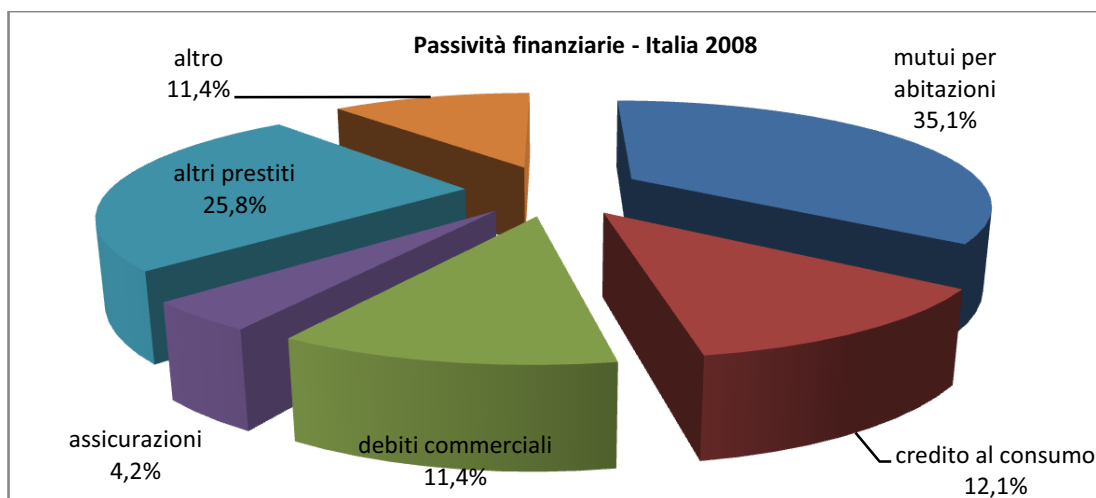
	2000	2001	2002	2003	2004	2005	2006	2007	2008
Attività reali									
Usa	2,24	2,35	2,44	2,56	2,73	2,99	2,90	2,65	2,33
Canada	2,62	2,68	2,81	2,92	3,03	3,18	3,28	3,38	3,34
Giappone	4,12	4,02	3,82	3,70	3,53	3,43	3,44	3,45	..
Germania	3,85	3,81	3,88	3,90	3,94	4,01	4,17	4,30	..
Francia	3,47	3,64	3,88	4,32	4,87	5,48	5,82	5,91	5,65
Regno Unito	3,88	3,91	4,55	4,82	5,27	5,23	5,61	6,01	5,17
Italia	4,03	4,06	4,27	4,51	4,66	4,86	5,08	5,27	5,27
Attività finanziarie									
Usa	4,56	4,21	3,77	4,21	4,41	4,66	4,84	4,87	3,88
Canada	3,53	3,50	3,49	3,45	3,39	3,46	3,50	3,49	3,52
Giappone	4,70	4,78	4,74	4,95	5,01	5,29	5,32	5,11	..
Germania	2,66	2,62	2,58	2,69	2,77	2,87	2,94	3,00	2,83
Francia	2,83	2,66	2,59	2,69	2,79	2,91	3,07	3,11	2,85
Regno Unito	4,97	4,45	3,95	4,11	4,30	4,67	4,91	4,97	4,24
Italia	3,60	3,38	3,31	3,32	3,42	3,53	3,56	3,49	3,11
Passività finanziarie									
Usa	1,01	1,05	1,10	1,17	1,24	1,31	1,35	1,38	1,32
Canada	1,13	1,14	1,17	1,21	1,24	1,29	1,32	1,38	1,42
Giappone	1,35	1,36	1,34	1,34	1,31	1,32	1,30	1,28	..
Germania	1,14	1,12	1,12	1,11	1,10	1,07	1,05	1,02	0,98
Francia	0,77	0,78	0,76	0,80	0,84	0,91	0,96	1,00	1,00
Regno Unito	1,17	1,21	1,34	1,45	1,60	1,62	1,78	1,86	1,81
Italia	0,50	0,49	0,55	0,59	0,63	0,67	0,71	0,74	0,74
Ricchezza netta									
Usa	5,79	5,51	5,11	5,60	5,91	6,34	6,39	6,21	4,90
Canada	5,02	5,03	5,13	5,16	5,18	5,34	5,46	5,50	5,44
Giappone	7,48	7,44	7,22	7,31	7,22	7,40	7,45	7,28	..
Germania	5,37	5,31	5,34	5,48	5,61	5,81	6,06	6,28	..
Francia	5,52	5,52	5,71	6,21	6,82	7,48	7,92	8,03	7,50
Regno Unito	7,68	7,14	7,16	7,48	7,97	8,27	8,75	9,12	7,60
Italia	7,13	6,95	7,02	7,24	7,46	7,72	7,93	8,01	7,64

Tabella 1 - Fonte: La ricchezza delle famiglie italiane 2008, Banca d'Italia

Questa differenza appare evidente se si osservano l'entità e la natura degli asset finanziari privati dei due paesi: le famiglie americane sono molto più indebitate di quelle italiane e, al contempo, possiedono poche forme di investimento a rendimento garantito, come depositi e assicurazioni; da notare, inoltre, che quasi il 26% degli asset finanziari delle famiglie americane sono rappresentati da fondi pensione, funzione che in Italia viene delegata per la maggior parte all'Inps e, quindi, non costituisce risparmio privato.



Inoltre, la natura delle passività finanziarie delle famiglie degli Stati Uniti indica quanto il ricorso al debito sia diffuso tra di loro: sebbene la percentuale delle famiglie che possiedono l'abitazione di residenza nelle due nazioni è simile (68,7% in Italia e 67,9% negli USA)²⁵ il 75% del valore totale dei debiti americani privati è rappresentato da mutui per la casa, quando questa percentuale in Italia è appena del 35%.



Fonte: Banca d'Italia & U.S. Census

²⁵ E' da notare che una buona parte delle famiglie italiane, più del 9%, risiede in una casa in usufrutto gratuito, mentre negli Stati Uniti questo fenomeno è quasi assente (solo lo 0,2% delle famiglie risiede in un'abitazione in usufrutto gratuito)

Conclusione e sviluppi futuri

L'obiettivo di questa tesi è di analizzare come variazioni inattese del reddito delle famiglie italiane influenzino il loro consumo. Tale obiettivo è stato raggiunto attraverso l'uso di un modello economico che prevede l'esistenza di due tipi differenti di variazioni del reddito, suddivise in base alla persistenza, e supponendo che il consumo reagisca diversamente a questi due tipi di variazioni. A partire da dati provenienti dalla Survey of Household Income and Wealth (SHIW) della Banca d'Italia, si sono stimate sia le varianze delle varie tipologie di shock di reddito a cui le famiglie italiane sono state sottoposte nel periodo preso in esame, sia i valori delle propensioni marginali del consumo rispetto alle due tipologie di shock, che rappresentano l'influenza delle variazioni del reddito sul consumo delle famiglie italiane.

I valori delle varianze delle variazioni del reddito ottenute mostrano che nel periodo 2000-2008 le variazioni permanenti del reddito hanno avuto una varianza costante nel tempo, mentre la varianza degli shock transitori ha subito negli anni un calo graduale; ciò è indice di un periodo di relativa stabilità (o stagnazione) economica: gli eventi associati a shock transitori del

reddito hanno un effetto sempre minore sul reddito delle famiglie italiane. Se alcune di queste cause, come malattie o infortuni, sono fisiologiche ed è difficile che la loro incisività vari nel tempo, altre, in particolar modo la disoccupazione, sono legate al periodo economico e quindi osservare una diminuzione della varianza di questi shock può essere interpretato come una maggior stabilità delle condizioni economiche delle famiglie (è da supporre che gli effetti della grande recessione non fossero ancora avvertiti dalle famiglie nel 2008, l'ultimo anno coperto dalla nostra analisi; infatti la raccolta dati dello SHIW viene eseguita nei mesi di marzo e aprile).

Inoltre si osserva come l'età e il titolo di studio influenzino l'entità e la tipologia di shock che si subiscono: famiglie con capofamiglia più istruite subiscono variazioni del reddito meno incisive rispetto a famiglie con un capofamiglia meno istruite. Per quanto riguarda l'età, individui più anziani subiscono shock permanenti più forti e shock transitori meno forti rispetto a chi è più giovane.

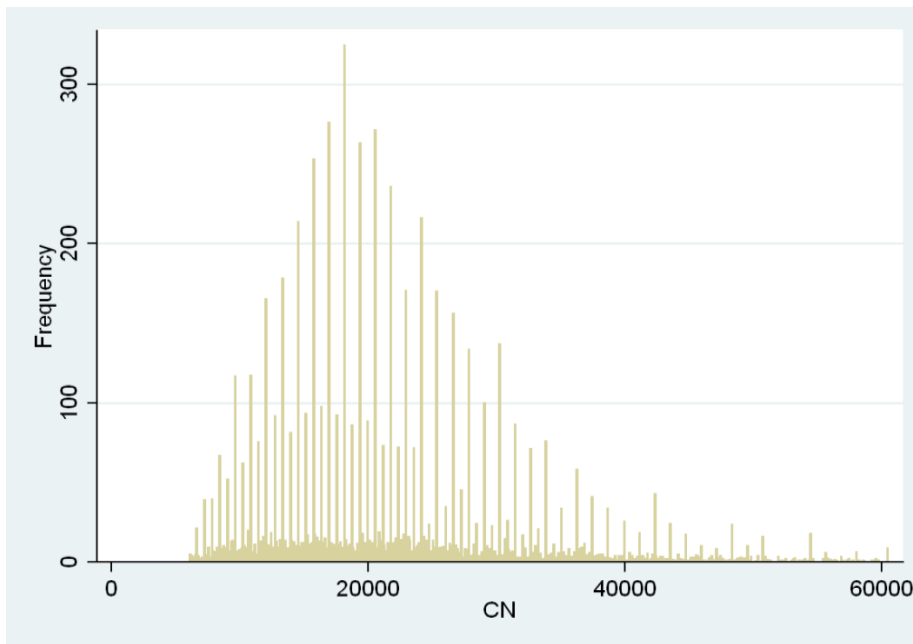
Il modello utilizzato per questa tesi contiene al suo interno sia il modello del reddito permanente di Friedman, dove esiste solo assicurazione completa contro shock transitori del reddito, sia il modello a mercati completi, dove esiste piena assicurazione contro shock di reddito. Le stime dei coefficienti di assicurazione del consumo rifiutano entrambe queste teorie. Tali stime indicano che le famiglie italiane non sono in grado di proteggersi contro variazioni permanenti del reddito, e che, inoltre, non esiste piena assicurazione neanche contro shock transitori, dato che circa il 30% di uno shock transitorio del reddito viene trasmesso al consumo. I valori delle covarianze di reddito e consumo non contemporanee, in particolare la covarianza tra variazioni del reddito presente e variazioni del

consumo future, suggeriscono che ciò può essere causato dalla presenza di vincoli di liquidità, fatto che rifletterebbe la rigidità del mercato del credito italiano e la natura prudente e non propensa all'indebitamento delle famiglie italiane, osservato come si nota dai dati sulle attività e passività finanziarie private e sul possesso dell'abitazione di residenza.

E' possibile che tali risultati siano distorti per la natura dei dati sul consumo utilizzati, di tipo *recall* e quindi soggetti ad arrotondamento. In un recente lavoro Ahmed, Brzozowski & Crossley (2010) sfruttano il particolare disegno campionario del Canadian Food Expenditure Survey (FoodEx), che contiene dati sia *recall* che *diary* sul consumo delle famiglie, per confrontare questi due metodi di raccolta dati, trovando che l'errore di arrotondamento è di natura molto diversa da quella usualmente assunta per le analisi econometriche; sarebbe interessante venire a conoscenza di come tale errore influenzi le particolari stime che abbiamo ottenuto in questa tesi. Il punto principale di questa tesi rimane, tuttavia, che le famiglie italiane hanno una capacità ridotta, rispetto a famiglie di altri paesi, di far fronte a variazioni inattese del reddito, indipendentemente da fattori quali l'età o il titolo di studio.

Appendice: L'effetto dell'arrotondamento sulle stime

Come accennato precedentemente, abbiamo imputato la causa della negatività della stima della varianza delle variazioni delle preferenze dei consumatori a una distorsione della stima dell'errore di misura del consumo. Questo perché i dati sul consumo dello SHIW soffrono di gravi problemi di arrotondamento e accumulazione; già Bound et al. (2001) enfatizzano il fatto che i metodi econometrici usati abitualmente assumono errori di misurazione di tipo *white noise* i.i.d. con distribuzione normale, mentre l'evidenza empirica contraddice questa assunzione. L'istogramma sottostante è relativo al consumo annuale di beni non durevoli delle famiglie italiane relativo all'anno 2008, tolti i valori estremi (sono comprese le famiglie con un valore del consumo compreso tra il 2° e il 98° percentile). È evidente come le osservazioni siano concentrate su pochi valori "tondi", mentre si osservano pochissime osservazioni nei valori intermedi.



Tale problema viene affrontato da Ahmed, Brzozowski & Crossley (2010), che sfruttano il particolare disegno campionario del Canadian Food Expenditure Survey (FoodEx), che prevede che i rispondenti prima stimino qual è il loro consumo di cibo nelle precedenti quattro settimane e, successivamente, registrino i loro consumi di cibo in un diario per le successive due settimane. Questo permette loro di confrontare direttamente dati raccolti attraverso questi due metodi di raccolta dei dati.

I risultati ottenuti da questi ricercatori indicano che l'errore di misura del consumo nei dati *recall* risulta essere molto grande, di distribuzione non normale e di natura molto diversa da quella che viene usualmente assunta; in particolare, si rileva eteroschedasticità, dipendenza dell'errore dal valore del consumo stesso e dipendenza sia della media che della varianza di tale errore dal reddito e da caratteristiche demografiche dell'unità statistica presa in considerazione. Successive analisi indicano che, se si assume che i dati di tipo

diary siano una misura accurata del consumo, quasi il 70% della variazione del consumo tra le unità statistiche è a causa di errori di misurazione.

Nel modello utilizzato in questa tesi, un tale errore di misura porterebbe a stime distorte dei parametri, specialmente quelli che riguardano l'errore di misura del consumo; ad esempio, se tale errore ha media non nulla, la stima della varianza di un errore di misurazione con media nulla risulterebbe distorta verso l'alto. Questo spiegherebbe perché la varianza dello shock di preferenze dei consumatori risulta negativo quando si inserisce l'errore di misurazione del consumo tra le variabili. Tuttavia tale distorsione non è misurabile senza una stima dei parametri ottenuta da dati non affetti da arrotondamento.

Per ovviare ai problemi causati dall'utilizzo di dati di tipo *recall*, è necessaria la presenza di una misura del consumo di tipo *diary* su un campione simile a quello intervistato nell'indagine che usa dati *recall*. In Italia questa misura è data dall'Indagine dei Bilanci delle Famiglie Italiane dell'Istat, che fornisce dati molto accurati sulle abitudini di spesa degli italiani; tuttavia, a questi dati manca la dimensione panel presente nei dati della Banca d'Italia, di conseguenza bisogna trovare dei metodi che permettano l'identificazione del modello, disponendo solamente di dati *cross-section* sul consumo.

Si sono individuati due metodi: uno consiste nell'imputare i dati sul consumo delle famiglie di un'indagine a partire da dati sulla stessa grandezza presi da un campione simile al primo; l'altro nell'avvalersi di un diverso metodo di identificazione dei parametri, non modificando il modello

economico utilizzato, e usare direttamente dati provenienti da una diversa indagine.

A.a. Imputazione dei dati

Il primo metodo, già sperimentato in precedenza da Skinner (1987) e Ziliak (1998), viene utilizzato da Blundell, Pistaferri & Preston (2008) (i cui risultati sono stati mostrati nei paragrafi precedenti) per utilizzare dati PSID sul consumo di beni non durevoli²⁶. Tale indagine possiede una componente panel, tuttavia, sono presenti solamente dati su reddito e consumo di cibo. Per calcolare il consumo di beni non durevoli del PSID si utilizzano i dati del CEX, che contengono informazioni sia sul consumo di cibo che di beni non durevoli, ma manca della dimensione panel. Essi stimano una funzione di domanda per il consumo di cibo a partire da dati CEX:

$$f_{i,t} = Z'_{i,t}\mu + d'_t\delta + \beta(Z_{i,t})c_{i,t} + \varepsilon_{i,t}$$

Dove $f_{i,t}$ è il logaritmo del consumo di cibo, e $c_{i,t}$ è il logaritmo del consumo di beni non durevoli (disponibili solo nel CEX). $Z'_{i,t}$ sono caratteristiche demografiche, d'_t sono dummy per l'anno, $\beta(Z_{i,t})$ è l'elasticità,

²⁶ Blundell Pistaferri e Preston preferiscono tale misura del consumo rispetto a dati sul consumo di cibo, più facilmente reperibili (come fatto da Hall), in quanto tali dati rappresentano una frazione sempre più piccola del consumo totale delle famiglie. Inoltre, la stima dei coefficienti di assicurazione non sarebbe in grado di distinguere una variazione del grado di assicurazione contro gli shock dei consumatori da un mutamento dell'elasticità del consumo di cibo, rischiando di confondere per una variazione della capacità delle famiglie di far fronte agli imprevisti, quello che in realtà è una variazione delle abitudini di consumo dei soggetti in esame.

che si suppone vari nel tempo e attraverso caratteristiche osservabili della famiglia ($Z_{i,t}$); $\varepsilon_{i,t}$ è l'errore, e cattura l'eterogeneità e l'errore di misurazione nella domanda di cibo. Si regrediscono i dati sulla spesa di cibo provenienti dal CEX su caratteristiche demografiche e consumo di beni non durevoli della stessa indagine²⁷. Assumendo monotonia nella domanda di cibo, si può invertire tale equazione, usando i valori dei parametri ottenuti attraverso la regressione su dati CEX, per ottenere una misura del consumo di beni non durevoli del PSID a partire dai dati sul consumo di cibo di tale indagine statistica.

Un metodo simile è utilizzato da Battistin, Miniaci & Weber (2002) per ricavare dati sul consumo di beni non durevoli dello SHIW che non soffrano di problemi di arrotondamento a partire da dati del SFB (Survey on Family Budgets) dell'Istat. Essi stimano, al pari di Blundell et al., una funzione che lega consumo di beni non durevoli, consumo di cibo e caratteristiche demografiche della famiglia, per poi ricavare dati sul consumo di beni non durevoli a partire da dati sul consumo di cibo. Tuttavia, la qualità dei dati ottenuti attraverso tale procedura dipende dai dati sul consumo di cibo usati nell'equazione di imputazione e, poiché, la misura del consumo di cibo osservata nello SHIW può essere diversa dal vero valore non osservato, a causa dell'arrotondamento dei dati che caratterizza questa indagine, questo può portare a distorsioni nella procedura di imputazione.

²⁷ Poiché il consumo di beni non durevoli correla con il consumo di cibo, allora il consumo di beni non durevoli correla anche con l'errore di stima, rendendo la stima attraverso OLS distorta. Si usa perciò una regressione a variabili strumentali, instrumentando la spesa per beni non durevoli (e la sua interazione con dummy per anno, educazione e numero di figli) con la media, specifica per gruppo, educazione e anno, della paga oraria del marito e della moglie (e la loro interazione con dummy per anno, educazione e numero di figli).

Battistin, Miniaci & Weber tengono conto di questo fatto assumendo che:

$$\begin{aligned} W &\sim f(w, \vartheta) \\ W^* &\sim f(W, G) \\ f(G|W) &\sim f(\gamma) \end{aligned}$$

Cioè che il vero valore della variabile di interesse W sia dipendente da un parametro ϑ , ma che possa essere osservata solo una misura di tale variabile, W^* , che può essere espressa come funzione del vero valore w e di una variabile G non osservabile, la cui distribuzione condizionale dato W dipende da γ , parametro del meccanismo di incompletezza. Assumendo che la distribuzione condizionata di W^* sia:

$$f(w^*|w, g, z) = \begin{cases} 1 & \text{se } w^* = W^*(w, g) \\ 0 & \text{se } w^* \neq W^*(w, g) \end{cases}$$

e chiamando $H(w^*)$ la sua inversa, cioè i valori della coppia (W, G) consistenti con il valore w^* , la funzione di massima verosimiglianza dei parametri γ e ϑ è:

$$\prod_{i \in SHIW} \int_{H(w_i^*)} f(g_i|w_i, z; \gamma) f(w_i|z; \vartheta) dg_i dw_i$$

La distribuzione empirica dei dati SHIW suggerisce che le famiglie arrotondano i loro consumi al multiplo più vicino di 100, 500 o 1000. Definendo i tre tipi di errori come:

$$\begin{aligned} w^* &\in R_1 \text{ se } |w - w^*| \leq 50 \\ w^* &\in R_2 \text{ se } |w - w^*| \leq 250 \\ w^* &\in R_3 \text{ se } |w - w^*| \leq 500 \end{aligned}$$

e assumendo che diversi arrotondamenti corrispondano a diversi sottoinsiemi separati del dominio di G , e che quindi esistano due valori di soglia τ_1 e τ_2 che separino questi tre sottoinsiemi essi definiscono:

$$\begin{aligned} H_1 &= [w^* - 50; w^* + 50) \times [\tau_2; +\infty) \\ H_2 &= [w^* - 250; w^* + 250) \times [\tau_1; \tau_2) \\ H_3 &= [w^* - 500; w^* + 500) \times [-\infty; \tau_1) \end{aligned}$$

Ne segue che:

$$H(w^*) = \begin{cases} H_1 \cup H_2 \cup H_3 & \text{se } w^* \in R_3 \Leftrightarrow |w - w^*| \leq 500 \\ H_1 \cup H_2 & \text{se } w^* \in R_2 \Leftrightarrow |w - w^*| \leq 250 \\ H_1 & \text{se } w^* \in R_1 \Leftrightarrow |w - w^*| \leq 50 \end{cases}$$

Si specifica una regressione probit a tre categorie per la distribuzione condizionata di G dato W e Z , ottenendo la distribuzione condizionata di G :

$$f(g|w, z; \gamma) \sim N(\gamma_1 w + \gamma_2 z; 1)$$

Assumendo che la funzione di densità del vero valore del consumo si distribuisca come una normale, e che quindi:

$$f(w|z, \vartheta) = N(\vartheta_1 z; \vartheta_2)$$

Si hanno tutti i fattori dell'equazione di massima verosimiglianza di cui sopra e, quindi, si può procedere alla stima dei parametri ϑ e γ . Una volta stimati i valori di questi due parametri, si può procedere all'imputazione dei dati. Poiché per il teorema di Bayes si ha:

$$f(w, g | w^*, z; \vartheta, \gamma) = \begin{cases} f(w, g | z; \vartheta, \gamma) & \text{se } w^* = W^*(w, g) \\ 0 & \text{se } w^* \neq W^*(w, g) \end{cases}$$

per ogni unità statistica, si imputerà una coppia (W, G) dalla distribuzione $f(w, g | z; \vartheta, \gamma)$ finchè non sarà generata una coppia consistente con il valore osservato w^* , cioè $(w, g) \in H(w^*)$.

I dati così ottenuti fanno capire che i dati provenienti dalle due indagini sono di qualità comparabile, una volta tenuto conto dell'arrotondamento dei dati dello SHIW. Se ciò è particolarmente vero per i dati sul consumo di cibo, lo stesso non si può dire per i dati sul consumo di beni non durevoli: viene rilevata per questi dati una distorsione verso il basso per i dati dello SHIW che non è completamente imputabile all'arrotondamento e che, pertanto, non svanisce nei dati imputati.

A.b. Diversa Identificazione del modello

Il secondo metodo per affrontare il problema dell'arrotondamento dei dati è sviluppare il modello economico in modo che non sia necessario l'uso né di dati di reddito e consumo che provengano dalla stessa indagine, né di dati sul consumo che possiedano una componente panel.

Tale metodo è sviluppato in un lavoro di Jappelli & Pistaferri (2010). Essi si avvalgono dello stesso modello economico utilizzato sia in questa tesi, sia in altri lavori (Blundell, Pistaferri & Preston, 2008; Gorodnichenko, Sabirianova & Stolyarov, 2010; Attanasio, Borella & Nielsen, 2011), ma identificano in maniera diversa i diversi parametri dell'equazione.

In particolare, essi stimano non parametricamente le varianze degli shock di reddito permanenti e transitori, attraverso un risultato di Meghir & Pistaferri (2004): omettendo per semplicità di notazione l'effetto delle caratteristiche demografiche si definiscono:

$$E[(y_t - y_{t-1})(y_{t+1} - y_{t-2})] = \sigma_{\delta_t}^2$$

$$E[(y_t - y_{t-1})(y_{t+1} - y_t)] = \sigma_{\varepsilon_t}^2$$

Dove $\sigma_{\delta_t}^2$ è la varianza dello shock permanente del reddito e $\sigma_{\varepsilon_t}^2$ è la varianza dello shock transitorio. Si può notare come tale identificazione usi gli stessi periodi di dati, $t-2$, $t-1$, t , $t+1$ utilizzati per l'identificazione descritta in questa tesi.

Una volta stimate le varianze degli shock del reddito, l'identificazione dei restanti parametri del modello non avviene, come in questa tesi, attraverso restrizioni poste sulle varianze delle differenze prime del reddito e del consumo, ma attraverso restrizioni sulle differenze prime delle varianze di reddito e consumo:

$$\Delta Var[y_t] = \sigma_{\delta_t}^2 + \sigma_{\varepsilon_t}^2 - \sigma_{\varepsilon_{t-1}}^2$$

$$\Delta Var[c_t] = \phi^2 \sigma_{\delta_t}^2 + \psi^2 \sigma_{\varepsilon_t}^2$$

Dove ϕ e ψ sono le propensioni marginali del consumo verso rispettivamente shock permanenti e transitori del reddito. E' da notare come il considerare la differenza prima delle varianze sollevi il ricercatore dallo stimare la varianza degli shock di preferenze dei consumatori, in quanto considerato omoschedastico.

Per l'identificazione, più precisamente, viene utilizzata la differenza tra queste due grandezze, cioè:

$$\Delta Var[y_t] - \Delta Var[c_t] = (1 - \phi^2)\sigma_{\delta_t}^2 + (1 - \psi^2)\sigma_{\varepsilon_t}^2 - \sigma_{\varepsilon_{t-1}}^2$$

Ovviamente una singola equazione non permette l'identificazione dei due parametri di interesse rimasti; Jappelli e Pistaferri suddividono il campione in base all'anno di nascita e stimano diverse varianze degli shock di reddito per ogni sottoinsieme, in modo da ottenere i momenti necessari per garantire l'identificazione del modello; in altre parole, l'equazione sopra viene riscritta come:

$$\Delta Var[y_{a,t}] - \Delta Var[c_{a,t}] = (1 - \phi^2)\sigma_{\delta_{a,t}}^2 + (1 - \psi^2)\sigma_{\varepsilon_{a,t}}^2 - \sigma_{\varepsilon_{a,t-1}}^2$$

dove a rappresenta l'anno di nascita.

Essi eseguono l'identificazione dei parametri a partire da dati dello SHIW sia del reddito che del consumo, ottenendo risultati, visibili nella tabella a pag. 92, confrontabili con i risultati di questa tesi e di altri lavori eseguiti sugli stessi dati. Essi inoltre testano questo modello nel caso vengano utilizzati, per il consumo, dati provenienti da una diversa indagine (come ad esempio le indagini inglesi Family Expenditure Survey e British Household Panel Survey), ottenendo delle stime accettabili anche in questo caso.

Tale procedura di identificazione esenta il ricercatore dall'utilizzare indagini aventi una componente panel sia del reddito che del consumo, anche se richiede ancora la presenza di dati di panel per il reddito delle famiglie. Tuttavia, per la completa identificazione dei parametri, sono sufficienti dati *cross-section* ripetuti sul consumo, anche non provenienti dalla

stessa indagine statistica dei dati sul reddito. Tale caratteristica, che Jappelli e Pistaferri giudicano utile per le analisi di dati di paesi che non possiedono dati di panel congiunti su reddito e consumo (come la sopracitata Inghilterra), può essere utilizzata anche in Italia, utilizzando i dati di panel sul reddito dello SHIW e i dati *cross-section* sul consumo del SFB, risolvendo in questo modo il problema di arrotondamento dei dati causato dalla natura dei dati sul consumo dello SHIW.

I metodi descritti in questa appendice dovrebbero essere in grado di fornire delle stime dei parametri che descrivono il legame tra reddito e consumo che non siano affetti da distorsione provocata da un arrotondamento dei dati, permettendo una misura della distorsione provocata da tale fenomeno; tuttavia, si delega l'approfondimento di questa tematica a lavori successivi.

Bibliografia

Abowd, & Card. (1989). On the Covariance Structure of Earnings and Hours Changes. *Econometrica* 57(2) , 411-455.

Ahmed, N., Brzozowski, M., & Crossley, T. (2006). Measurement Errors in Recall Food Consumption Data. *IFS Working Papers W06/21, Institute for Fiscal Studies* .

Aiyagari. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics* 109(3) , 659-684.

Ando, & Modigliani. (1963). The 'Life-Cycle' Hypothesis of Saving: Aggregate Implications and Tests. *American Economic Review* 53 , 55-84.

Attanasio, Borella, & Nielsen. (2011). *Characterizing income shocks and their transmission to household consumption*. Palermo: Fondazione "Rodolfo De Benedetti".

Banca D'Italia. (2010). I bilanci delle famiglie Italiane 2008. *Supplemento al Bollettino Statistico* 8.

Battistin, Miniaci, & Weber. (2002). What do We Learn from Recall Consumption Data? *The Journal of Human Resources* 38(2) , 354-385.

Blundell, Pistaferri, & Preston. (2008). Consumption Inequality and Partial Insurance. *American Economic Review* 98(5) , 1887-1920.

- Bodkin. (1959). Windfall Income and Consumption. *The American Economic Review*, 49(4) , 602-614.
- Bodkin. (1966). Windfall Income and Consumption: Reply. *The American Economic Review* 56(3) , 540-546.
- Brandolini. (1999). The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality. *Temi di discussione del Centro Studi della Banca d'Italia* .
- Campbell, & Stanley. (1963). Experimental and Quasi-Experimental Designs for Research on Teaching. In *Handbook of Research on Teaching*. Chicago: Rand McNally.
- Costas, Meghir, & Pistaferri. (2004). Income Variance Dynamics and Heterogeneity. *Econometrica* 72(1) , 1-32.
- Cutler, & Katz. (1992). Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980s. *American Economic Review*, 82(2) , 546-551.
- Deaton, & Paxon. (1994). Intertemporal Choice and Inequality. *Journal of Political Economy* 102(3) , 437-467.
- Deaton, & Paxon. (1993). Saving, Growth and Aging in Taiwan. *NBER Working papers* 4330 .
- Deaton, & Paxson. (1998). Growth, Demographic Structure, and National Saving in Taiwan. *Papers* 183, Princeton, Woodrow Wilson School .
- Friedman. (1957). A Theory of the Consumption Function. *Princeton University Press* .
- Gorodnichenko, Sabirianova, & Stolyarov. (2010). Inequality and Volatility Moderation in Russia: Evidence from Micro-Level Panel Data on Consumption and Income. *Review of Economic Dynamics* 13 , 209-237.
- Gruber. (1997). The consumption smoothing benefits of unemployment insurance. *American Economic Review* 87(1) , 192-205.

- Guiso, Jappelli, & Pistaferri. (2002). An Empirical Analysis of Earnings and Employment Risk. *Journal of Business and Economic Statistics* 20(2) , 241-253.
- Guvenen, & Smith. (2010). *Inferring Labor Income Risk From Economic Choices: an Indirect Approach*. Federal Reserve Bank of Minneapolis.
- Hall. (1978). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy* 86 , 971-987.
- Hall, & Mishkin. (1982). The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households. *Econometrica* 50 , 461-481.
- Hausman. (1978). Specification Tests in Econometrics. *Econometrica* 46(6) , 1251-1271.
- Hryshko. (2006). Identifying Household Income Processes Using a Life Cycle Model. *University of Alberta* .
- Jappelli, & Pistaferri. (2010). Does Consumption Inequality Track Income Inequality in Italy? *Review of Economic Dynamics* 13(1) , 133-153.
- Jappelli, & Pistaferri. (2010). Financial Integration and Consumption Smoothing. *The Economic Journal* 121(553) , 678-706.
- Jappelli, & Pistaferri. (2001). Tax Incentives and the Demand for Life Insurance: Evidence from Italy. *Centre for Studies in Economics and Finance (CSEF) Working Papers 52, Università di Napoli* .
- Jappelli, & Pistaferri. (2010). The consumption response to income changes. *NBER Working paper Series* .
- Johnson, & Smeeding. (1998). Measuring the Trends in Inequality of Individuals and Families: Income and Consumption. *Unpublished* .
- Landsberger. (1966). Windfall Income and Consumption: Comment. *The American Economic Review* 56(3) , 534-540.
- Lee, T. H. (1975). More on Windfall Income and Consumption. *Journal of Political Economy*, 83(2) , 407-417.

Pistaferri. (2001). Superior Information, Income Shocks, and the Permanent Income Hypothesis. *The Review of Economics and Statistics* 83(3) , 465-476.

Skinner. (1987). A Superior Measure of Consumption from the Panel Study of Income Dynamics. *Economic Letters* 23(2) , 213-216.

Stephens. (2001). The Long-run Consumption Effects of Earning Shocks. *The Review of Economical Statistics* 83(1) , 28-36.

Ziliak. (1998). Does the Choice of Consumption Measure Matter? An Application to the Permanent-Income Hypothesis. *Journal of Monetary Economics* 41(1) , 201-216.