INDICE GENERALE

Introduzione

Nel decidere come strutturare il proprio portafoglio, le famiglie italiane si trovano a dover scegliere fra una variegata gamma di strumenti finanziari. Il risparmio può venire depositato per intero in conti correnti oppure può essere investito usando anche strumenti finanziari con un più elevato livello di rischio quali le azioni. Data la sempre crescente disponibilità di strumenti finanziari di quest'ultimo tipo, tuttavia, è sorprendente notare come nemmeno la metà delle famiglie italiane ne faccia ricorso, preferendo allocare le proprie risorse in forme di investimento con un profilo di rischio molto basso. Ciò è ancora più strano se consideriamo il fatto che tali attività finanziarie offrono in media una remunerazione di gran lunga superiore, almeno nel lungo periodo, rispetto alla loro controparte più sicura garantendo quindi un elevato premio per il rischio, il cosiddetto *equity premium*.

Mentre un considerevole livello di ricerca è stato dedicato alle scelte di risparmio delle famiglie, le scelte di allocazione di portafoglio sono state molto meno dibattute. Ciò che ci proponiamo con il presente lavoro è una analisi dal punto di vista econometrico del problema in questione, concentrando l'attenzione su un particolare aspetto di tale problema e cioè alla modellazione della partecipazione al mercato dei titoli rischiosi. Rivolgeremo la nostra attenzione su metodi di modellazione che, avvalendosi dei dati di tipo panel forniti dall'indagine SHIW, condotta in Italia dalla Banca d'Italia, permette di tenere conto di quantità di difficile identificazione nei classici modelli di regressione per dati cross-section. La disponibilità di dati di panel ci permetterà di analizzare inoltre il problema anche da un punto di vista dinamico permettendoci di conoscere l'evoluzione del possesso di particolare categorie di strumenti finanziari nel tempo relativamente ad ogni specifica famiglia.

Il primo capitolo si occuperà di introdurre l'argomento, fornendo le ragioni che ci portano a considerare questa come una delle principali tematiche a riguardo. Si tratterà di definire, dunque, quali strumenti considereremo rischiosi e quali siano le dinamiche attualmente presenti nei mercati italiani. Nel far ciò presenteremo alcune analisi descrittive dei dati SHIW così da fornire un quadro aggiornato di ciò che sta succedendo in Italia sul fronte della partecipazione ai mercati dei titoli rischiosi.

Nel corso del secondo capitolo passeremo quindi alla specificazione dei modelli che applicheremo ai dati, fornendo i metodi di stima che ci permetteranno di identificare i modelli. A questo proposito valuteremo modelli di tipo lineare e non lineare preoccupandoci di analizzare modelli sia di tipo statico che dinamico, con l'obbiettivo di valutare un eventuale effetto di inerzia nel possesso di attività rischiose al netto di elementi di specificità relativi al singolo soggetto. In quest'ottica proporremo, infine, un modello per l'analisi di modelli dinamici lineari che ci permetta di utilizzare per la stima l'intero campione a disposizione che è caratterizzato da una cadenza di rilevazione irregolare.

Il terzo capitolo presenterà infine i risultati delle stime dei modelli presentati nel capitolo precedente fornendo dunque gli effetti sul possesso di titoli rischiosi indotti dalle variabili considerate. Proveremo quindi ad applicare il modello dinamico proposto per l'intero campione di dati a disposizione con l'obbiettivo di valutarne l'efficacia per il caso in questione.

Scelte di portafoglio

Fino a vent'anni fa la teoria economica riguardo alle scelte economiche operate dalle singole famiglie si era focalizzata in modo particolare sulle scelte di consumo o investimento. Si era dunque interessati a "spiegare" le ragioni che spingevano le famiglie a scegliere di consumare un determinato quantitativo di risorse e investirne la restante parte (Deaton 1992). Le scelte di portafoglio, invece, ovvero le scelte riguardanti la ripartizione della quota da investire sui differenti strumenti finanziari a disposizione di ogni singola famiglia sul mercato, erano state relativamente trascurate. La mancanza di interesse in questo ambito della ricerca può essere, almeno in prima battuta, ricondotta alla relativa scarsa varietà di strumenti finanziari fino a poco tempo fa a disposizione dei risparmiatori. Nell'ultimo ventennio, infatti, abbiamo potuto assistere al proliferare di strumenti finanziari che hanno permesso di ampliare in misura considerevole le possibilità di scelta, come testimonia d'altra parte la notevole diffusione dei fondi comuni. Un ulteriore elemento che può aver contribuito a spingere gli studiosi a non rivolgere particolare attenzione a questi argomenti è stata la relativa mancanza di dati ed informazioni sufficientemente dettagliate da permettere lo studio della composizione dei portafogli detenuti dalle famiglie. Negli ultimi anni, tuttavia, in molti Paesi, europei e non, si sono andate sviluppando indagini che fornissero informazioni sufficientemente dettagliate relative a scelte di investimento e consumo. Tali indagini sono inoltre state condotte introducendo (almeno in alcuni casi) una dimensione di tipo panel all'interno del disegno campionario, così da poter sfruttare appieno il recente proliferare di metodologie di stima per questo genere di dati particolarmente adatto agli studi in ambito econometrico. Esempi di indagini di questo tipo sono la SCF (*Survey of Consumer Finances*) negli Stati Uniti, la CSS (*CentER Savings Survey*) in Olanda e la SHIW (*Survey of Household Income and Wealth*) in Italia di si occuperemo più diffusamente in seguito.

Grazie anche a questi elementi si è potuto andare sviluppando, dunque, un filone di ricerca interessato alle scelte di portafoglio operate dalle famiglie (Guiso, Jappelli, Haliassos 2002). Alcuni dei principali ambiti di ricerca in questo campo possono essere : le scelte relative all'accumulazione di ricchezza lungo tutto il ciclo di vita di una persona, su quali basi una famiglia decide o meno di possedere titoli rischiosi e come esse allocano la propria ricchezza fra le diverse forme di investimento disponibili, siano esse rischiose o meno. Queste ultime due scelte possono a prima vista sembrare la stessa, tuttavia, la decisione di investire in un particolare asset si può pensare come svolta in due fasi tra di loro correlate ma distinte: la prima decisione riguarda se investire o meno in tale attività e successivamente, condizionatamente alla scelta di investire, viene definito l'importo da assegnare. Una tale procedura decisionale viene assunta anche dalla maggior parte dei modelli utilizzati per spiegare l'ammontare detenuto dai singoli investitori su una particolare forma di investimento. I motivi che spingono a considerare un processo decisionale così composto sono perlopiù legati alla presenza di vincoli alle vendite allo scoperto e alla presenza di costi di transizione relativi alla partecipazione ai mercati finanziari. Se guardiamo alla distribuzione della ricchezza allocata dall'insieme degli agenti su un particolare asset, notiamo come questi due elementi inducano una forte concentrazione nello zero. La presenza di questi elementi porta tutti coloro che avrebbero preferito detenere quote negative di titoli rischiosi o una quantità positiva ma limitata a decidere di detenere zero, gli uni a causa dell'impossibilità di vendere allo scoperto per il singolo agente, gli altri per la presenza di barriere all'entrata qui rappresentate dai costi di transizione, siano essi monetari, quali i costi di ingresso, informativi o altro .

Nel presente lavoro ci occuperemo unicamente della prima parte di tale processo decisionale, focalizzando la nostra attenzione sulle determinanti della partecipazione al mercato dei titoli rischiosi, tralasciando la seconda parte che come detto, condizionatamente alla prima, si occupa di definirne l'entità.

1.1 La partecipazione al mercato dei titoli rischiosi

Come abbiamo avuto modo di notare, dunque, l'attenzione negli ultimi anni si è rivolta in modo particolare allo studio della composizione dei portafogli detenuti dai singoli risparmiatori ed uno degli argomenti di maggiore interesse a questo riguardo sono le scelte di partecipazione al mercato dei titoli rischiosi.

Infatti uno dei più importanti cambiamenti verificatisi in ambito finanziario negli ultimi vent'anni è stata la diffusione della cosiddetta *equity culture*. In Italia, anche se in misura minore rispetto ad altri paesi industrializzati quali gli Stati Uniti o il Regno Unito, si è assistito ad un notevole incremento nel livello di partecipazione in titoli azionari, dove per livello di partecipazione stiamo ad indicare la proporzione di soggetti che, all'interno della popolazione di riferimento, dichiarano di detenere almeno un titolo azionario. Analizzando tale proporzione, infatti, tenendo in considerazione sia le azioni detenute in maniera diretta sia quelle detenute in forma indiretta attraverso il possesso di altre forme di investimento, è immediatamente evidente un trend crescente come possiamo notare in tabella 1.1.

Tabella 1.1 : Partecipazione in titoli rischiosi

	Stati Uniti	***Regno Unito***	***Olanda***	***Germania***	***Italia***
	Partecipazioni azionarie dirette				
1983	*19.1*	*8.9*	*n.a.*	*9.7*	*n.a.*
1989	*16.8*	*22.6*	*n.a.*	*10.3*	*4.5*
1995	*15.2*	*23.4*	*11.5*	*10.5*	*4.0*
1998	*19.2*	*21.6*	*15.4*	*n.a.*	*7.3*
	Partecipazioni azionarie indirette				
1983	*n.a.*	*n.a.*	*n.a.*	*11.2*	*n.a.*
1989	*31.6*	*n.a.*	*n.a.*	*12.4*	*10.5*
1995	*40.4*	*n.a.*	*29.4*	*15.6*	*14.0*
1998	*48.9*	*31.4*	*35.1*	*n.a.*	*18.7*

Fonte : Guiso, Jappelli, Haliassos "Household portfolios" pag. 9

Tenere in considerazione le quote azionarie detenute indirettamente risulta quanto mai cruciale in quanto più che alla diffusione di quote azionarie detenute in maniera diretta questo incremento nel livello di partecipazione è dovuto ad un forte incremento nella domanda di investimenti attraverso intermediari finanziari, nella domanda cioè di strumenti quali i fondi comuni, le pensioni integrative e le assicurazioni vita.

La crescente ricerca da parte delle famiglie di fondi comuni può essere per lo più ricondotta a tre fattori :

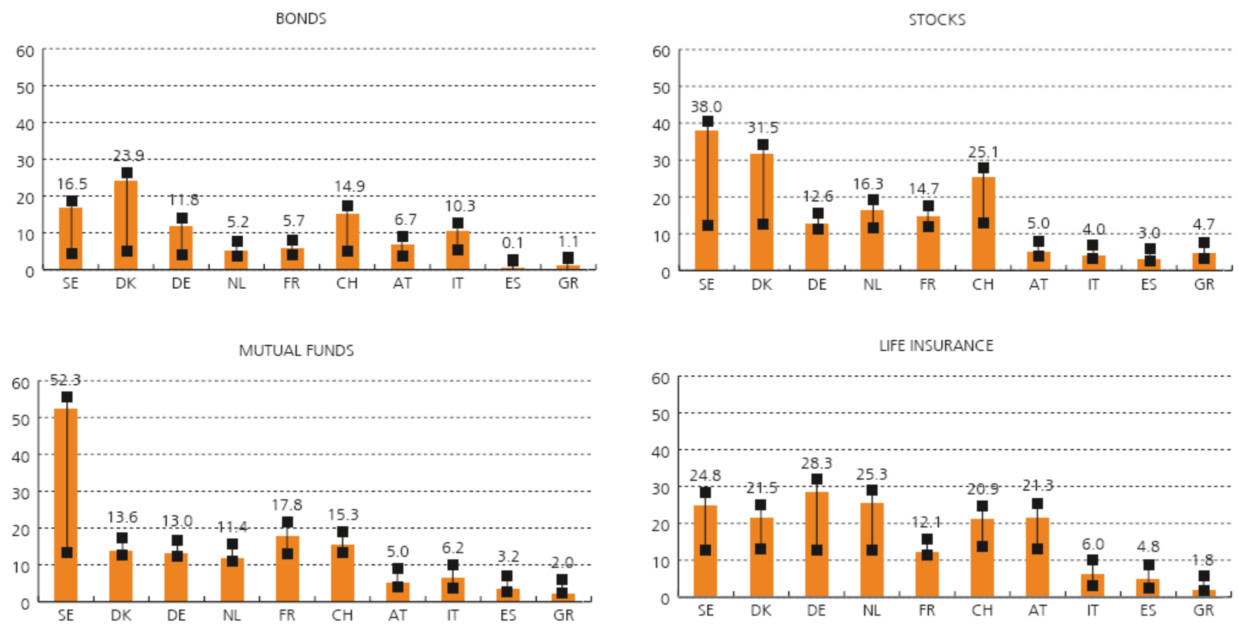
- minor rischio rispetto all'investimento azionario diretto
- minori informazioni e capacità necessarie all'investimento
- minori costi di transazione

In particolare il minore rischio rappresentato dall'investimento in fondi comuni, sostanzialmente dovuto alla possibilità di meglio diversificare il portafoglio di titoli, unito alla possibilità di delegare le scelte a professionisti con la conseguente riduzione del livello minimo di conoscenza, oltre che del tempo, richiesto hanno permesso a questo genere di investimenti di accrescere notevolmente la loro diffusione.

Come avremo modo di notare analizzando i dati a nostra disposizione, tuttavia, questo trend, almeno per quanto riguarda l'Italia, sembra avere subito una battuta d'arresto nel 2000 per motivi che sono oggetto di analisi di questa tesi.

1.2 Lo stockholder puzzle

Nonostante la recente diffusione di titoli rischiosi, però, il numero di famiglie che non partecipano a questo genere di mercato risulta ancora eccezionalmente basso. Tale fenomeno risulta, inoltre, ancor più rilevante in Italia rispetto ad altri paesi europei, come possiamo vedere dalla figura 1.2, dove si riportano le proporzioni di possesso di obbligazioni, azioni, fondi comuni ed assicurazioni vita per alcuni dei maggiori paesi europei nel 2004 per la popolazione degli ultra-cinquantenni. Il livello di possesso di azioni e fondi comuni, in particolare, risulta molto basso in Italia rispetto al livello medio delle altre nazioni.



(figura 1.2)[1]

Questa forte riluttanza da parte degli investitori privati a partecipare al mercato dei titoli rischiosi è noto in letteratura come *stockholder puzzle* ed altro non è che l'analogo micro di ciò che viene definito *equity premium puzzle*.

Quest'ultimo termine venne introdotto per la prima volta da Mehra e Prescott nel 1985 in un articolo nel quale fecero notare come i modelli economici comunemente utilizzati risultassero incapaci di spiegare l'elevata remunerazione storicamente osservata per i titoli azionari sul mercato statunitense. In particolare mostrarono come l'*equity premium* osservato[2] dal 1889 al 1978 era del 6.18% annuo, mentre modelli quali il *Consumption CAPM*, comunemente utilizzato per il prezzaggio degli asset, indicavano un equity premium che si aggirava intorno allo 0.35% annuo. Proprio la

1 Fonte : Health, Ageing and Retirement in Europe - First Results from SHARE (2005). pag. 315

2 Misurato come differenza fra il rendimento medio dei titoli azionari ed il rendimento medio obbligazionario grande differenza fra il valore previsto dai modelli e il rendimento effettivamente verificatosi li portò a coniare questo termine. Notevoli sono stati, in tempi recenti, i tentativi per riuscire a spiegare questa differenza tuttavia ancora non è stato possibile arrivare ad una soluzione univoca.

La presenza di un equity premium per il possesso di titoli azionari così elevato dovrebbero indurre una proporzione molto alta di persone ad utilizzare questo genere di strumenti finanziari a discapito di forme di risparmio meno rischiose e tale opportunità dovrebbe essere avvertita in maniera ancor maggiore dalle persone più giovani che possono, potendo sfruttare un orizzonte temporale più lungo, ridurre il rischio derivante dalla maggiore variabilità di rendimento atteso fornito dai titoli azionari. Numerosi analisi empiriche a riguardo dimostrano, tuttavia, come ciò non accada nei mercati di tutto il mondo e ciò ha indotto a coniare il termine stockholder puzzle. (Haliassos Bertaut, 1995)

Risulta evidente come l'identificazione delle ragioni di una tale riluttanza a detenere azioni risulti di particolare interesse nell'ottica di riuscire a mantenere e rinnovare una cospicua porzione di investitori, così come risulterebbe fondamentale nell'indirizzare decisioni di politica economica, si pensi alla riforma del sistema pensionistico ed in particolare alla previdenza complementare.

Alcune delle motivazioni che inducono una così nutrita porzione di famiglie a decidere di non investire in asset di tipo rischioso possono essere le seguenti:

- Presenza di forti barriere all'entrata. Molte tipologie di contratto prevedono che l'agente corrisponda, al momento della sottoscrizione del contratto, un costo fisso, alcuni, inoltre, necessitano di un investimento minimo al di sotto del quale non è possibile stipulare l'accordo. E' indubbio che la presenza di questo genere di vincoli può scoraggiare alcuni potenziali investitori dall'utilizzare tali strumenti finanziari, così come essi possono alterare in modo significativo le conclusioni in ambito teorico dal momento che, in sostanza, non permettono di costruire portafogli con quantità arbitrarie di risorse investite per singolo asset e ciò va a minare l'ipotesi di completezza del mercato posta come condizione di partenza nella maggior parte dei modelli utilizzati. (Gollier 2002)

- Scarse informazioni riguardo alle possibilità di investimento. Difficilmente l'intera popolazione può essere a conoscenza dell'intera gamma di strumenti finanziari a disposizione, mentre, anche coloro che ne sono a conoscenza, possono non avere sufficiente competenza per potersene avvalere e ciò può contribuire a spiegare la scarsa partecipazione osservata. (King Leape 1998)
- Presenza di rischio individuale non assicurabile. Oltre al rischio connesso all'investimento in sé gli investitori valutano anche la presenza di una componente di rischio che risulta difficile evitare, legata alle possibili variazioni nel livello di reddito della famiglia o alla longevità stessa del singolo individuo che possono aumentare l'aleatorietà dell'ambiente in cui opera l'investitore ed indurlo ad assumersi minori rischi. (Heaton Lucas 2000)

Analogamente a queste spiegazioni sono state proposte altre possibili cause per il fenomeno in questione, tuttavia ancor oggi non è stata data una chiara risposta al problema, lasciando dunque ampio spazio alla futura ricerca.

Nel seguito, una volta analizzata più in dettaglio l'indagine SHIW dalla quale andremo ad utilizzare i dati, ci occuperemo di presentare alcune semplici analisi descrittive dei dati al fine di evidenziare i mutamenti verificatisi in Italia per quanto riguarda il possesso di titoli rischiosi per gli anni che vanno dal 1989 al 2004, nonché proporre alcune relazioni di tipo empirico facilmente individuabili per via grafica.

1.3 L'indagine SHIW

I dati di cui ci serviremo provengono dall'indagine SHIW (*Survey on Household Income and* Wealth) condotta dalla Banca d'Italia. Tale indagine viene svolta dal 1948, mentre dal 1989 è stata introdotta una dimensione panel. Alcune famiglie a partire da quell'anno sono state infatti re-intervistate in periodi successivi sulle stesse quantità, così da ottenere misure ripetute nel tempo per la stessa unità statistica. Il numero di famiglie intervistate è di circa 8000 per ogni indagine, mentre il numero di famiglie panel è andato aumentando negli anni fino ad attestarsi alle circa 3600 famiglie dell'indagine 2004. L'unica eccezione alla cadenza biennale è stata fatta per l'anno 1997 nel corso del quale l'indagine è stata annullata, per problemi in fase di raccolta dei dati, ed è stata ripetuta nel 1998.

L'indagine SHIW è stata largamente utilizzata in passato soprattutto per la qualità delle informazioni messe a disposizione. Le principali fonti di informazione a livello italiano sui redditi e sui consumi delle famiglie italiane sono, infatti, costituite dalle indagini SHIW e SFB (*Survey on Family* Budget). Quest'ultima, condotta dall'ISTAT per calcolare l'indice del costo della vita, contiene informazioni piuttosto dettagliate per quanto riguarda i consumi delle famiglie utilizzando la rilevazione dei dati attraverso i diari di spesa, tuttavia, per le informazioni relative al reddito ed alla ricchezza delle famiglie l'indagine SHIW è ritenuta contenere informazioni di migliore qualità[3].

Dal momento che le quantità di cui ci occuperemo ed in particolare il possesso di titoli rischiosi risulta molto superiore per soggetti caratterizzati da una ricchezza elevata e poiché questo genere di soggetti rappresentano una minima proporzione del totale delle famiglie italiane il semplice campionamento stratificato delle famiglie adottato in questa indagine si traduce in un basso numero di tali famiglie analizzate con una conseguente ripercussione sul livello di dettaglio dell'informazione relativa al possesso di titoli rischiosi. L'indagine SHIW, a differenza di quanto adottato dalla SCF statunitense a dalla CSS olandese, non utilizza un sovracampionamento delle famiglie più ricche. Questo accorgimento permetterebbe di avere a disposizione un maggior numero di dati relativi a queste famiglie, così da descriverne più in dettaglio il comportamento. Tali dati, come ad esempio nel caso statunitense, possono venire riferiti alla popolazione originaria attraverso l'applicazione di determinati pesi alle singole osservazioni, pesi inversamente proporzionali alla probabilità di campionamento della relativa unità statistica. Per l'indagine italiana, invece, i pesi di campionamento forniti per le singole osservazioni si riferiscono al fatto che non tutte le unità statistiche della popolazione hanno uguale probabilità di campionamento a tali pesi permettono, ancora una volta, di riferire i dati alla popolazione originaria.

Ulteriori elementi di difficoltà nell'utilizzare i dati SHIW sono dovuti, in analogia con

---

3 Per una analisi approfondita sulla qualità dei dati resi disponibili dalle indagini proposte rinviamo a Brandolini (1999) la CSS olandese, al fatto che la dimensione temporale dei dati di tipo panel è piuttosto bassa. In altre parole, data la recente nascita della dimensione panel nell'indagine, il numero di periodi per il quale sono state effettuate misurazioni ripetute è ancora basso e tale problema viene amplificato dal fatto che, come vedremo, il panel non risulta equispaziato, introducendo un ulteriore problema per l'utilizzo dei dati nel caso in cui si vogliano adottare modelli di tipo dinamico.

1.4 Alcune analisi descrittive

In questo paragrafo ci occuperemo di dare un'idea più dettagliata dei cambiamenti verificatisi nel livello di partecipazione in attività finanziarie rischiose fra le famiglie italiane negli ultimi anni. A tal proposito è importante definire sin d'ora ciò che nel seguito considereremo titoli rischiosi.

Le informazioni a nostra disposizione ci permettono di suddividere le forme di risparmio adottate da ogni famiglia sostanzialmente in sette categorie : depositi bancari e certificati di deposito, depositi postali, titoli di stato italiani, obbligazioni e quote di fondi comuni, partecipazioni azionarie italiane, gestioni patrimoniali, titoli esteri, prestiti a cooperative. La nostra attenzione sarà rivolta in modo particolare a quella categoria di strumenti finanziari caratterizzata da un discreto livello di rischio ed adotteremo dunque la convenzione di definire titoli rischiosi le quote in partecipazioni azionarie dirette o indirette attraverso il possesso di quote di fondi comuni[4]. Nonostante a partire dal 2000 sia disponibile, per l'indagine SHIW, una

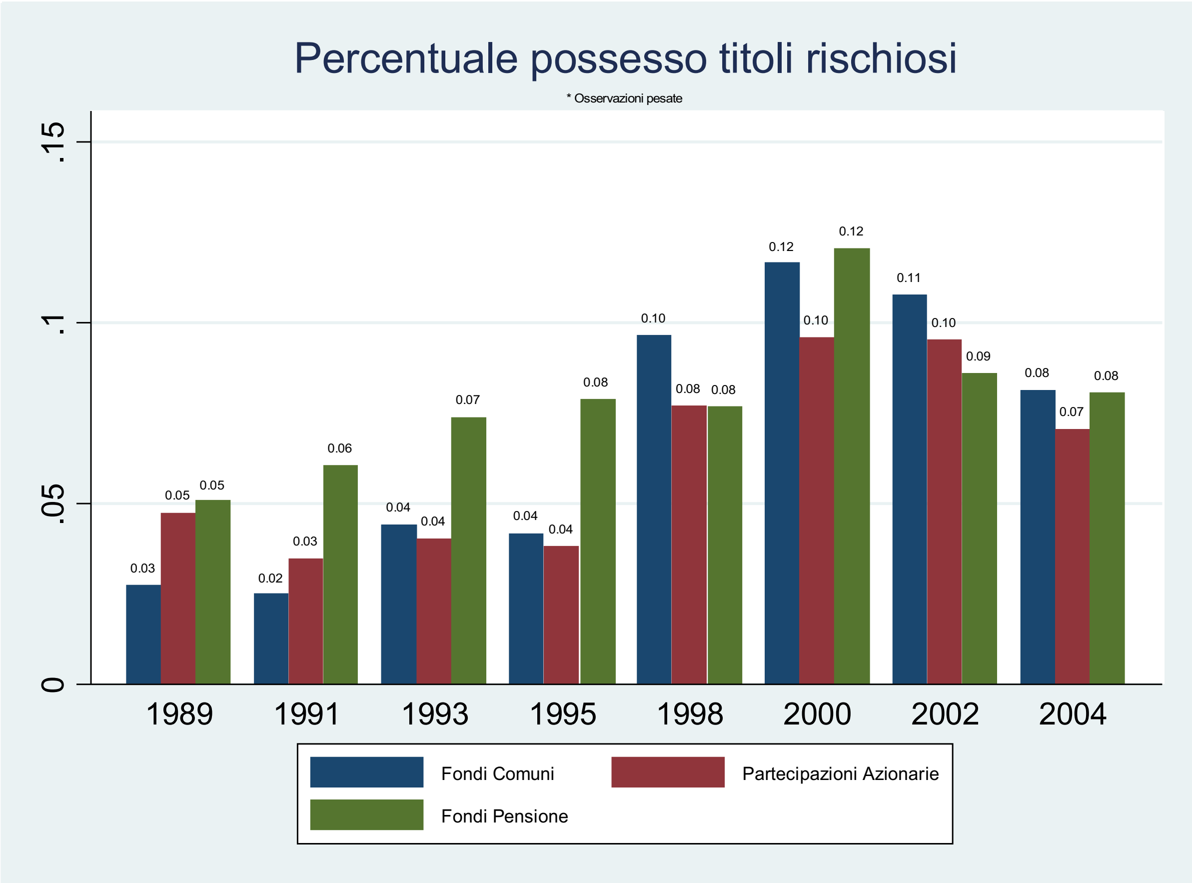
4 Potremmo includere nella categoria di titoli rischiosi anche il possesso di titoli esteri e gestioni patrimoniali, tuttavia per semplicità non terremo conto di questi ultimi strumenti finanziari in quanto il loro livello di partecipazione è molto basso e pressoché nullo se condizionato alla non partecipazione in azioni o fondi comuni.

distinzione fra fondi comuni azionari, obbligazionari, misti e monetari, per gli anni precedenti non è possibile discriminare fra queste tipologie di titoli e questo ci porterà a dover includere nel computo il possesso di fondi comuni indipendentemente dalla loro natura anche se il possesso di fondi comuni obbligazionari o monetari può difficilmente essere considerata un'attività finanziaria rischiosa.

In aggiunta a questo genere di attività sarà ritenuto rischioso anche l'eventuale possesso di pensioni integrative o assicurazioni vita che corrispondano un vitalizio finale al sottoscrittore, dal momento che questo genere di assicurazioni rappresentano più una forma di investimento, equiparabile in qualche misura ad un fondo comune, che un'assicurazione vera e propria.

La figura 1.3 mostra come il possesso di queste tre forme di investimento appena definite siano andate diffondendosi in modo sempre maggiore fino al 2000, in maniera analoga a quanto avvenuto nei mercati stranieri, mentre negli ultimi quattro anni i dati rivelano una tendenza inversa. Le proporzioni sono qui calcolate pesando le osservazioni per tenere conto della procedura di campionamento utilizzata.

Dai dati disaggregati è possibile notare come a guidare la diffusione di tali strumenti finanziari sia stato il notevole aumento, soprattutto a partire dal 1998, del numero di famiglie in possesso di fondi comuni, anche se non è trascurabile l'aumento anche in partecipazioni azionarie dirette. Per tutte e tre le categorie si è tuttavia registrato un aumento, più o meno significativo, nella loro diffusione con una successiva contrazione a partire dall'anno 2002.



(figura 1.3)

Per analizzare più nel dettaglio il possesso di tali titoli rischiosi, con particolare riguardo alla loro diversificazione, riportiamo in tabella 1.4 la diffusione di loro specifiche combinazioni.

Le prime tre colonne della tabella stanno ad indicare il possesso o meno di azioni, fondi comuni o pensioni integrative e per ogni riga viene riportata la percentuale di famiglie con quelle particolari caratteristiche registrata nei diversi anni di indagine.

A supporto di quanto detto in precedenza notiamo che la proporzione di famiglie che detiene unicamente quote di fondi comuni ha subito un forte incremento fra il 1995 ed il 1998 per decrescere leggermente nei tre anni successivi. Un andamento analogo è riscontrabile per tutte quelle combinazioni che prevedano il possesso di fondi comuni, mentre per le altre combinazioni l'andamento risulta un po' più irregolare e di difficile interpretazione, a riprova del fatto che uno dei fattori trainanti nell'aumento della partecipazione in titoli rischiosi sia la crescente domanda di forme di fondi comuni.

Tabella 1.4 : Percentuale di possesso di combinazioni di titoli rischiosi

Azioni	*Fondi Comuni*	*Pensioni Int.*	*1989*	*1991*	*1993*	*1995*	*1998*	*2000*	*2002*	*2004*
0	**0**	**0**	*90.28%*	*90.99%*	*87.69%*	*86.59%*	*79.41%*	*75.63%*	*80.26%*	*81.13%*
1	**0**	**0**	*2.87%*	*1.92%*	*2.54%*	*2.21%*	*3.73%*	*4.14%*	*4.11%*	*3.99%*
0	**1**	**0**	*1.30%*	*1.37%*	*2.33%*	*2.82%*	*5.99%*	*5.72%*	*5.48%*	*4.99%*
0	**0**	**1**	*3.77%*	*4.45%*	*4.99%*	*6.16%*	*4.99%*	*6.25%*	*4.61%*	*5.48%*
1	**1**	**0**	*0.50%*	*0.38%*	*0.94%*	*0.78%*	*2.79%*	*2.79%*	*2.32%*	*1.63%*
1	**0**	**1**	*0.63%*	*0.34%*	*0.49%*	*0.51%*	*0.86%*	*1.75%*	*1.07%*	*0.80%*
0	**1**	**1**	*0.38%*	*0.37%*	*0.70%*	*0.66%*	*1.28%*	*2.12%*	*1.27%*	*1.20%*
1	**1**	**1**	*0.26%*	*0.18%*	*0.32%*	*0.26%*	*0.95%*	*1.60%*	*0.89%*	*0.79%*
Totale			***100%***	***100%***	***100%***	***100%***	***100%***	***100%***	***100%***	***100%***

Nota : i valori 1 e 0 nelle prime tre colonne rappresentano rispettivamente il possesso e non del relativo strumento finanziario.

Suddividendo, più in generale, le famiglie considerate fra possessori di titoli rischiosi e non a seconda che tale famiglia possegga o meno almeno uno fra gli strumenti finanziari appena analizzati, possiamo mettere in evidenza le dinamiche verificatesi negli anni considerati. In tabella 1.5 riportiamo le transizioni verificatesi per gli anni dal 1991 al 2004.

Tabella 1.5 : Transizioni nel possesso di titoli rischiosi

	1989	1991	1993	1995	1998	2000	2002	2004
0 → 0		*84.84%*	*80.86%*	*79.34%*	*70.73%*	*65.07%*	*65.26%*	*68.60%*
0 → 1		*4.24%*	*8.92%*	*6.33%*	*13.12%*	*13.03%*	*6.81%*	*8.77%*
1 → 0		*4.55%*	*4.61%*	*5.76%*	*5.50%*	*7.72%*	*10.86%*	*8.56%*
1 → 1		*6.37%*	*5.61%*	*8.57%*	*10.65%*	*14.18%*	*17.07%*	*14.07%*
0	*90.28%*	*90.99%*	*87.69%*	*86.59%*	*79.41%*	*75.63%*	*80.26%*	*81.13%*
1	*9.72%*	*9.01%*	*12.31%*	*13.41%*	*20.59%*	*24.36%*	*19.73%*	*18.87%*

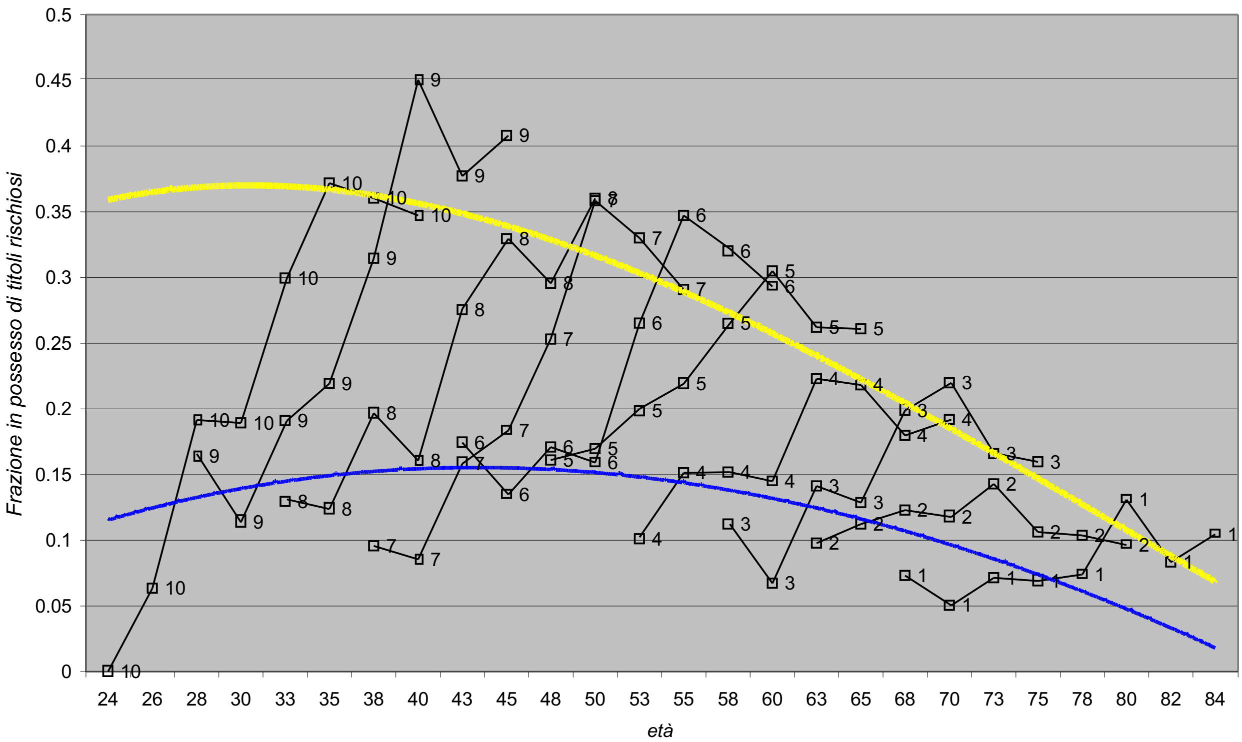
Le percentuali sono calcolate sul totale di osservazioni a disposizione nei due anni considerati, mentre le ultime due righe si riferiscono alle proporzioni di possesso di titoli rischiosi sui dati cross-section.

La percentuale di famiglie in possesso di titoli rischiosi, in analogia con quanto visto per i dati disaggregati, mostra un forte incremento a partire dal 1998 per poi decrescere nel 2002. A riprova di questa tendenza la differenza fra nuovi entrati ed usciti rimane sempre positiva fatta eccezione per l'anno 2002 nel quale la percentuale di persone che smettono di detenere attività rischiose supera di gran lunga quella di coloro che decidono di iniziare a detenerne.

Un elemento di interesse nello spiegare gli andamenti e le dinamiche assunte dal fenomeno della partecipazione in titoli rischiosi può essere rappresentato dalla relazione fra quest'ultima e l'età del capofamiglia. Precedenti studi sui dati italiani per lo più nello scorso decennio hanno evidenziato una forma “a gobba” della relazione fra età e partecipazione. La partecipazione raggiungeva cioè un picco intorno ai 50 anni e tendeva ad affievolirsi sia per età minori che maggiori.

La descrizione di una relazione di questo tipo non può però prescindere da possibili effetti coorte presenti nei dati. Per tenere conto di questo elemento abbiamo suddiviso la popolazione in 10 coorti a seconda dell'anno di nascita a partire dai nati intorno al 1920 per arrivare ai nati intorno al 1965 ed abbiamo quindi rilevato la frazione di possessori di titoli rischiosi nel tempo per ognuna delle coorti definite andando a comporre quelle che vengono chiamate *cohort curves*.

Figura 1.6 : Partecipazione per età e coorte



nota : coorti suddivise in base all'anno di nascita ogni 5 anni, dal 1920 al 1965. Le curve sono spline relative ai primi quattro (blu) e gli ultimi quattro periodi di osservazione (gialla).

Il grafico 1.6 mostra i risultati e ci permette di notare come la presenza di salti fra le cohort curves evidenzi la presenza di effetti coorte o tempo. Esse si sviluppano nel tempo in modo diverso a seconda della coorte a dimostrazione del fatto che soggetti appartenenti a diverse coorti evolvono nel tempo dal punto di vista del possesso di titoli rischiosi in maniera differente. In particolare fra i più giovani si registra un incremento molto più elevato nel tempo di possessori che fra i più anziani e ciò porta a notare come la relazione fra possesso ed età al netto di effetti coorte si sia evoluta nel tempo. Il grafico riporta inoltre due curve *spline* costruite interpolando i dati relativi ai primi quattro (blu) e gli ultimi quattro periodi di osservazione (gialla) e sono interpretabili come l'effetto cercato dell'età sul possesso di titoli rischiosi al netto degli effetti coorte. Tali curve ci permettono di mostrare come mentre per i primi quattro periodi la forma della curva è "a gobba" in linea con i precedenti studi, più recentemente tale relazione indica un possesso maggiore per i soggetti più giovani con un effetto a decrescere all'aumentare dell'età. Le persone più giovani, infatti, dovrebbero tendere a possedere titoli rischiosi in misura maggiore rispetto alla media in quanto possono sfruttare in maniera maggiore l'equity premium garantito da questo tipo di investimenti.

Un ulteriore elemento di forte discriminazione per i possessori di titoli rischiosi è rappresentato dal titolo di studio del capofamiglia. Notiamo infatti che la proporzione di possessori fra i laureati è di 39.45% contro solo il 33.65% dei diplomati ed il 13.69% dei rimanenti.

Riportiamo, infine, in tabella 1.7 la relazione fra il possesso delle varie categorie di titoli rischiosi da noi definite ed il livello di reddito. La tabella riporta la percentuale di famiglie in possesso di quote del relativo investimento condizionatamente alla categoria di reddito di appartenenza. L'ultima colonna riporta infine la proporzione di possessori per le famiglie più ricche, quelle cioè il cui reddito si attesta sopra il 95° percentile.

Come ci si poteva attendere il possesso aumenta in maniera piuttosto rilevante con il livello di reddito al punto che i soggetti con reddito più basso il livello di possesso è sostanzialmente nullo.

Tabella 1.7 : Possesso di strumenti finanziari rischiosi per fascia di reddito

	Sotto il I quartile	***Tra il I ed il II quartile***	***Tra il II ed il III quartile***	***Sopra il III quartile***	***Sopra il 95° percentile***
Azioni	0.66%	2.29%	5.42%	16.05%	28.98%
Fondi Comuni	0.99%	3.01%	7.07%	16.14%	24.06%
Pensioni Integrative	1.40%	4.31%	8.60%	15.73%	20.82%
Titoli Rischiosi	2.67%	8.34%	17.59%	34.82%	49.06%

Notiamo, inoltre, come, fra i tre tipi di investimento, le pensioni integrative siano le più diffuse per le famiglie un livello di reddito fino al terzo quartile, mentre fra le più ricche rimane molto diffusa la partecipazione azionaria diretta. Nonostante una correlazione positiva fra possesso e reddito, notiamo che anche nella fascia più ricca della popolazione neppure il 50% delle famiglie detiene almeno una di queste tre forme di investimento ed una così bassa partecipazione, a maggior ragione se ci riferiamo a questa categoria di famiglie, risulta di difficile interpretazione.

I Modelli

I dati a nostra disposizione sono, come detto, di tipo panel ovvero osservazioni ripetute nel tempo delle stesse quantità sulle medesime unità statistiche, che sono nel nostro caso rappresentate dalle famiglie italiane. L'idea sottostante al presente lavoro è la stima di modelli del tipo:

$$\begin{aligned} y_{it}^{*} &= x_{it}'\beta + \mu_{it} \\ y_{it} &= 1[y_{it}^{*} \geqslant 0] \end{aligned} \qquad (1)$$

dove y_{it}^{*} è una variabile latente definita su tutto l'asse reale, y_{it} è la variabile osservata dicotomica che descrive il possesso di titoli rischiosi, x_{it}' è un vettore di dimensione *p* contenente le variabili condizionanti mentre μ_{it} è una componente aleatoria; tali grandezze si intendono inoltre riferite all'unità *i* al tempo *t*. I modelli che considereremo nei successivi paragrafi rappresentano in sostanza differenti specificazioni di questo modello relativamente a :

- diversa natura delle variabili esplicative componenti il vettore x_{it}'
- differenti assunzioni sulla componente erratica μ_{it}

Per quanto riguarda la prima caratterizzazione nel seguito considereremo modelli sia di tipo statico, nei quali il vettore di variabili esplicative sarà composto di quantità che supporremo esogene in senso stretto, sia di tipo dinamico, per i quali includeremo anche variabili per loro natura correlate con il termine di errore quali la variabile dipendente ritardata.

Riguardo alla componente di errore tratteremo differenti specificazioni della stessa, oltre che per il tipo di distribuzione assunta vera anche per la struttura di dipendenza nel tempo fra gli errori stessi, struttura che sarà esplicitata in maniera più dettagliata di volta in volta venga proposto un modello.

Nel seguito presenteremo le specificazioni ed i relativi metodi di stima utilizzando la distinzione a seconda della componente erratica, tratteremo dunque inizialmente la stima di modelli di tipo non lineare applicati al caso di dati di panel per poi passare alla specificazione di modelli di tipo lineare.

2.1 Modelli non-lineari per dati di panel

Il modello che ci proponiamo di stimare può dunque essere formalizzato come in (1) e risulta evidente come, in un contesto di questo tipo, le assunzioni sulla componente di errore siano cruciali sia per quanto riguarda l'interpretazione del modello, sia per la stima del modello stesso. L'errore, dal punto di vista interpretativo, racchiude in sé l'informazione, per qualsivoglia ragione, non inclusa nelle variabili esplicative. In particolare ciò può essere dovuto all'omissione di una qualche variabile esplicativa. Le ragioni che possano indurre ad omettere una variabile possono essere dovute alla non disponibilità pratica di tale misura o alla impossibilità di misurare in modo oggettivo la variabile stessa.

In ambito econometrico l'implementazione di modelli come quello in esame si scontrano con l'impossibilità di misurare un determinato set di variabili implicitamente riferibili ad ogni unità statistica. Nel caso in questione, ad esempio, possiamo identificare un insieme di misure difficilmente quantificabili con oggettività quali il livello di accesso al mercato, i costi di transazione ed alcune altre variabili che fanno riferimento a preferenze ed abilità specifiche del soggetto considerato. Una possibile soluzione può essere ottenuta ricorrendo all'utilizzo di *proxy* di tali variabili, quantità cioè fortemente correlate con la variabili di partenza e che possono essere dunque utilizzate per approssimare il valore assunto dalle variabili stesse. Anche ricorrendo a questa procedura, tuttavia, la misurazione di tali caratteristiche risulta piuttosto difficile da operare in maniera oggettiva.

Una formalizzazione che tenga conto dell'effetto indotto nel modello da questo genere di variabili è quella secondo la quale l'errore è formato da due componenti : una specifica di individuo ed uno shock casuale.

Possiamo dunque definire la componente di errore presente in (1) come

$$\mu_{it} = \eta_i + \epsilon_{it}$$

dove ε_{it} è uno shock casuale assunto serialmente incorrelato, tale cioè che

$$E(\epsilon_{it} \epsilon_{js}) = 0 \qquad \forall \quad i \neq j, t \neq s$$

Resta da notare come l'effetto fisso η_i , per come è stato definito, non dipenda dal tempo.

Abbiamo dunque implicitamente assunto che le preferenze siano costanti nel tempo condizionatamente all'unità statistica considerata. Tale affermazione potrebbe non corrispondere sempre al vero, tuttavia, dal momento che eventuali mutazioni nelle preferenze e nelle abilità di un particolare soggetto, così come variazioni nelle condizioni offerte dal mercato sono generalmente di entità minima e che, almeno per il problema di cui qui ci occuperemo, l'orizzonte temporale nel quale i soggetti vengono osservati è relativamente breve, considerare tali variabili come costanti nel tempo è al più una ragionevole approssimazione della realtà.

Il modello sin qui definito è dunque del tipo :

$$\begin{aligned} y_{it}^{*} &= x_{it}'\beta + \eta_i + \epsilon_{it} \\ y_{it} &= 1[y_{it}^{*} \geqslant 0] \end{aligned} \qquad (2)$$

E' bene precisare, inoltre, come la presenza di una variabile risposta di tipo dicotomico porti a dover modellare non la presenza o assenza di una particolare caratteristica, bensì la probabilità di presenza della caratteristica stessa. Un modo analogo di considerare il problema consiste infatti nel trattare la variabile osservata y_{it} come realizzazione di una variabile casuale di Bernoulli con probabilità di successo pari a π_{it}, dove tale quantità si riferisce all'unità *i* al tempo *t*. Ciò che si vuole modellare, nei modelli con variabile dipendente continua così come in questo tipo di modelli, è la media della variabile dipendente condizionatamente ad un set di variabili esplicative. In questo caso, tuttavia, la media di una variabile casuale di Bernoulli è proprio la probabilità di successo e dunque modellare la media condizionata di y_{it} corrisponde a modellare π_{it}.

2.1.1 Conditional Logit Model

La scelta di impiegare modelli di tipo non lineare in contesti nei quali la variabile risposta è di tipo dicotomico deriva dall'esigenza di prendere in considerazione il fatto che il campo di variazione dei valori previsti dal modello stesso deve essere l'intervallo (0,1) poiché che ciò che ci proponiamo di modellare è la probabilità di successo.

Ciò che vogliamo modellare rimane in ogni caso la media condizionata della variabile y che, come detto, nel nostro caso altro non è che la probabilità di successo. L'approccio che utilizzeremo prende spunto da questa osservazione e. a partire dal modello in (2) otteniamo :

$$Pr(y_{it}=1|x_{it})=Pr(y_{it}^{*}>0|x_{it})=Pr(x_{it}'\beta+\eta_i+\epsilon_{it}>0|x_{it})=Pr(\epsilon_{it}>-x_{it}'\beta-\eta_i)$$

dove l'ultima uguaglianza è dovuta all'assunzione di indipendenza dell'errore ε_{it} da x_{it} e da η_i.

Ora se ε_{it} è una variabile casuale con funzione di ripartizione F abbiamo che

$$Pr(y_{it}=1|x_{it})=1-F(-x_{it}'\beta-\eta_i)$$

Le usuali assunzioni sulla forma della componente di errore ricadono su scelte di tipo Normale o Logistico. Tali variabili casuali possiedono la proprietà di simmetria ed è possibile dunque scrivere :

$$Pr(y_{it}=1|x_{it})=F(x_{it}'\beta+\eta_i)$$

In questo genere di contesti la funzione di ripartizione della variabile casuale adottata assume il nome di *funzione legame* e ha sostanzialmente il compito di mappare i valori assunti da $x_{it}'\beta+\eta_i$ appartenenti all'intero asse reale in valori compresi nell'intervallo (0,1) che risulta dunque compatibile con il campo di variazione di una probabilità.

I modelli ottenuti a partire da assunzioni sulla forma logistica o normale della funzione di ripartizione della componente di errore assumono rispettivamente il nome di *modelli logit* o *modelli probit*.

La stima di questo tipo di modelli presenta comunque ancora il problema dell'identificazione fintanto che all'interno del modello è presente l'effetto fisso[1]. Valutando il caso di una funzione legame tipo logit il modello diventa :

$$Pr(y_{it}=1|x_{it})=F(x_{it}'\beta+\eta_i)=\frac{\exp(x_{it}'\beta+\eta_i)}{1+\exp(x_{it}'\beta+\eta_i)} \qquad (4)$$

Nel caso lineare, come vedremo, il problema relativo alla presenza di η_i può essere risolto trovando un'opportuna trasformazione dei dati che renda il modello indipendente dalla componente non osservabile. In questo caso tuttavia non è possibile reperire una trasformazione dei dati con queste caratteristiche e dunque l'approccio alla stima deve essere diverso.

Una possibile soluzione potrebbe essere quella di fare riferimento ad una stima di massima verosimiglianza. Dal momento che abbiamo specificato il modello come in (4) ciò equivale a considerare il modello come proveniente da (1) dove l'errore è di tipo logistico di media nulla. La stima di massima verosimiglianza può quindi fare leva sulla distribuzione già assunta per y_{it} e possiamo, a questo punto, ottenere la funzione di verosimiglianza. Supponendo, per semplicità di esposizione, T=2 abbiamo che la funzione di verosimiglianza per la generica osservazione $\mathbf{y} = (y_{i1}, y_{i2})$ è la seguente :

---

1 Nel seguito proporremo modelli nei quali la componente η_i sarà trattata come un coefficiente specifico di individuo, tralasciando i modelli definiti in letteratura come modelli ad effetti casuali nei quali tale elemento viene considerato una variabile casuale. Il vantaggio di questo approccio consiste nel non dovere introdurre assunzioni sulla forma della relazione fra η_i e le variabili condizionanti, assunzione invece necessaria nei modelli ad effetti casuali. Bover-Arellano (1997)

$$L(\beta,\eta|y,X)=\prod_{i=1}^{N}\prod_{t=1}^{2}\left[\frac{\exp(x_{it}'\beta+\eta_i)}{1+\exp(x_{it}'\beta+\eta_i)}\right]^{y_{it}}\left[\frac{1}{1+\exp(x_{it}'\beta+\eta_i)}\right]^{1-y_{it}}$$

La funzione di verosimiglianza ottenuta, come si può notare, dipende dalla componente non osservabile η_i e dunque non possiamo ottenere stime per i parametri coinvolti massimizzando direttamente questa funzione di verosimiglianza.

L'identificabilità dei parametri passa infatti attraverso l'individuazione di una statistica sufficiente per η_i.

Come noto, dato un modello statistico parametrico, una statistica sufficiente per un parametro è una statistica che riassuma in sé tutte le informazioni per la stima di quel parametro. Più formalmente essa è una funzione dei dati tale che la verosimiglianza basata su di essa sia equivalente alla verosimiglianza basata sui dati di partenza. Come conseguenza dunque la funzione di verosimiglianza condizionata a tale statistica deve essere funzione unicamente dei dati e non dipendere quindi dal parametro. Più in dettaglio, nel nostro caso, una statistica sufficiente per il parametro η_i sarebbe tale che la funzione di verosimiglianza condizionata a tale statistica sarebbe indipendente dal parametro stesso.

Nel caso di una funzione legame di tipo logit ciò che si può facilmente far vedere è che una statistica sufficiente per il parametro desiderato è $S=\sum_{t=1}^{T} y_{it}$ e dunque possiamo ricavare l'espressione della funzione di verosimiglianza condizionata ad S. A tal proposito notiamo che la statistica S, nel caso in cui T=2, ha come supporto l'insieme composto dai numeri naturali 0, 1 e 2, tuttavia, dal momento che il contributo alla verosimiglianza apportato dall'i-esima unità non è altro che Pr ($Y_{i1}=y_{i1}$,$Y_{i2}=y_{i2}$), se la statistica S assume valori 0 o 2 la suddetta probabilità risulta già definita. Ciò si concretizza nel fatto che il contributo alla verosimiglianza apportato da osservazioni nelle quali non si registrano transizioni da uno stato ad un altro è identicamente 1 e dunque la funzione di verosimiglianza così ottenuta è equivalente alla verosimiglianza calcolata sulle sole osservazioni nelle quali si registrino transizioni. La verosimiglianza cercata è a questo punto la funzione di verosimiglianza sulle suddette osservazioni condizionata alla statistica sufficiente S, ovvero :

$$L(\beta,\eta|y,X,S)=\prod_{i\in C}\left[\frac{\exp(x_{it}'\beta+\eta_i)}{1+\exp(x_{it}'\beta+\eta_i)}\right]^{y_{i1}}\left[\frac{1}{1+\exp(x_{it}'\beta+\eta_i)}\right]^{1-y_{i1}} \qquad (5)$$

dove $C=\{i=1,\ldots,N|\sum_{t=1}^{T}y_{it}=1\}$

Il caso per T=2 evidenzia come la funzione di verosimiglianza ottenuta corrisponda a quella di un modello logit costruito per la sola osservazione y_{i1}, condizionatamente ad S=1, infatti, la seconda osservazione risulta combinazione lineare della prima e non porta dunque contributi alla verosimiglianza.

La stima dei coefficienti di interesse può dunque essere ottenuta dalla massimizzazione della funzione di verosimiglianza o, analogamente, della log-verosimiglianza.

Come si può facilmente vedere da (5) la funzione di verosimiglianza ottenuta corrisponde alla stima di un modello logit classico ristretto alle sole osservazioni per cui si osservi transizione nel quale i possibili valori assunti dalla variabile risposta sono (1,0) o (0,1). Gli standard error relativi a quest'ultima stima si dimostrano però non essere corretti e dunque per ottenere quelli esatti abbiamo bisogno di massimizzare la (5).

Ciò che abbiamo sin qui visto si fondava sull'assunzione che la funzione legame fosse di tipo logit. L'applicazione di questa procedura risulta vincolata a questa assunzione in quanto, supponendo ad esempio un modello probit, lo stimatore ci indurrebbe alla ricerca di una statistica sufficiente per η_i nel nuovo modello, tuttavia la sua individuazione non è immediata e dunque la ricerca della verosimiglianza condizionata da massimizzare ai fini della stima non risulta percorribile. Tale procedimento non può dunque essere generalizzato ad una funzione legame arbitraria a meno che la statistica sufficiente per la componente non osservabile non assuma un'espressione tale da rendere possibile l'implementazione della procedura di condizionamento della funzione di verosimiglianza.

Come abbiamo visto questi modelli, a differenza di quelli di tipo lineare, ammettono solo probabilità incluse fra 0 e 1. Uno dei limiti di questa procedura è però rappresentato dal fatto che ciò non ci permette un'immediata interpretazione dei coefficienti. A differenza di ciò che accade, come vedremo, nel caso di modelli di tipo lineare, i coefficienti ottenuti non rappresentano la variazione in probabilità dovuta ad un aumento unitario della variabile relativa. In questo caso infatti :

$$\frac{\partial Pr(y_{it}=1|x_{it})}{\partial x_{it}'} = F'(x_{it}'\beta+\eta_i)\beta \neq \beta \tag{6}$$

dove F' è la derivata della funzione legame F.

Come possiamo notare il coefficiente è in generale diverso dalla variazione in probabilità dovuta alla variabile esplicativa, tuttavia il segno di β e della derivata in (6) sono dello stesso segno dal momento che F , per come è stata fin qui definita, è una funzione di ripartizione e dunque la sua derivata è una densità che per definizione è sempre non negativa. Possiamo dunque concludere dicendo che il coefficiente β ottenuto a partire da un modello non lineare fornisce un'idea immediata del segno della variazione causata sulla probabilità dalla variabile condizionante relativa, ma l'entità di tale variazione risulta determinata da (6) e non direttamente dal valore assunto dal coefficiente.

2.1.2 Stimatore Honoré-Kyriazidou

Nel precedente paragrafo abbiamo fissato l'attenzione sul comportamento della variabile risposta condizionatamente ad un set di variabili che abbiamo definito esogene al modello e che quindi possono essere assunte incorrelate con la componente di errore presente nel modello.

Un elemento importante può tuttavia venire rappresentato dal valore assunto dalla stessa variabile risposta nel periodo precedente in quanto, almeno nel caso che ci proponiamo di studiare, può essere accettato il fatto che una significativa influenza sulla scelta da parte di un individuo di detenere o meno titoli rischiosi in futuro sia esercitata dall'attuale possesso o meno di tale titolo.

Un approccio di questo tipo può essere formalizzato con l'inclusione, all'interno delle variabili condizionanti, della variabile endogena ritardata :

$$Pr(y_{it}=1|x_{it})=x_{it}'\beta+\rho\, y_{it-1}+\eta_i \qquad (7)$$

Importante risulta a questo punto chiarire ciò che rappresenta in questo modello il parametro ρ. Dal momento che la componente μ_{it} è assunta essere, almeno in prima battuta, un *white noise*, ρ non rappresenta tutta la covarianza fra y_{it-1} e y_{it} , ma la rappresenta al netto dell'effetto di η_i. Dal punto di vista interpretativo dunque, nel modello considerato, ρ indica la probabilità di detenere titoli rischiosi al tempo t dato il loro possesso al tempo t-1, ma questa misura viene depurata dall'effetto, piuttosto cospicuo, dovuto al fatto che rimangono inalterate nel tempo variabili quali l'avversione al rischio e, più in generale, le preferenze dell'individuo, nonché le altre variabili condizionanti presenti in x_{it}. Questo genere di correlazione dovuta al fatto che sia y_{it} che y_{it-1} sono determinate da variabili che assumono lo stesso valore prende il nome in letteratura di correlazione *spuria*.

Modelli di questo tipo presentano nuovi tipi di problemi in fase di stima, fra tutti il fatto che la nuova variabile ritardata genera endogenità all'interno del modello che dovrà essere trattata in maniera specifica di volta in volta.

A questo proposito diventa cruciale la scelta riguardante il metodo di trattazione delle condizioni iniziali. Nel caso i dati a disposizione coprano un periodo relativamente limitato è necessario imporre una distribuzione di probabilità alla variabile risposta nel periodo iniziale. L'approccio proposto da Heckman (1981) consiste nell'includere un'equazione di tipo statico per la condizione iniziale nel modello. In particolare i coefficienti associati alle variabili per il modello al tempo 0 possono differire dai coefficienti del resto del modello. Nel modello che considereremo, inoltre, vedremo come non sia necessario imporre assunzioni sulla forma della relazione fra le variabili esplicative ed y_{i0}. L'utilizzo di questo approccio non renderà necessario, infine, disporre delle osservazioni relative alle variabili esplicative nel periodo iniziale, cosa che può risultare in alcuni casi ininfluente, ma che può portare a preferire questo metodo nei casi in cui tali informazioni siano di difficile reperibilità.

La forma funzione legame utilizzata è di tipo logit e dunque formalizzando tale approccio avremo che :

$$Pr(y_{i0}=1|x_i, \eta_i)= p_0(x_i, \eta_i)$$

$$Pr(y_{it}=1|x_{it}', \eta_i, y_{i0}, \dots, y_{it-1})=\frac{\exp(x_{it}'\beta+\rho y_{it-1}+\eta_i)}{1+\exp(x_{it}'\beta+\rho y_{it-1}+\eta_i)}$$

Nel caso in cui non siano presenti altre variabili condizionanti nel modello ad eccezione della variabile y ritardata, si dimostra che, per T=3, condizionare la verosimiglianza alla statistica sufficiente $S=\sum_{t=1}^{2} y_{it}$ conduce ad una funzione di verosimiglianza che non dipende dalla componente non osservabile. Tale osservazione non è più valida una volta introdotto un insieme di variabili condizionanti X_{it} nel modello. Per T=3 infatti il contributo alla verosimiglianza relativo dall'i-esima unità statistica è il seguente :

$$Pr(y_{i0}=d_0, y_{i1}=d_1, y_{i2}=d_2, y_{i3}=d_3|x_t)= p_0(x_i, \eta_i)^{d_0}[1-p_0(x_i, \eta_i)]^{1-d_0}\times$$

$$\times\prod_{t=1}^{3}\left[\frac{\exp(x_{it}'\beta+\rho d_{t-1})}{1+\exp(x_{it}'\beta+\rho d_{t-1})}\right]^{d_t}\left[\frac{1}{1+\exp(x_{it}'\beta+\rho d_{t-1})}\right]^{1-d_t}$$

condizionando tale funzione di verosimiglianza alla statistica sufficiente S otteniamo che, indicati con A e B gli eventi mutuamente esclusivi

$$A=\{(y_{i0}, \dots, y_{i3})\in\{0,1\}^4 : y_{i0}=d_0, y_{i1}=1, y_{i2}=0, y_{i3}=d_3\}$$

$$B=\{(y_{i0},..,y_{i3})\in\{0,1\}^4 : y_{i0}=d_0, y_{i1}=0, y_{i2}=1, y_{i3}=d_3\}$$

possiamo definire le probabilità di questi due eventi come :

$$Pr(A|x_i)=p_0(x_i,\eta_i)^{d_0}[1-p_0(x_i,\eta_i)]^{1-d_0}\times\frac{\exp(x_{i1}\beta+\rho d_0+\eta_i)}{1+\exp(x_{i1}\beta+\rho d_0+\eta_i)}\times$$
$$\times\frac{1}{1+\exp(x_{i2}\beta+\rho+\eta_i)}\times\left[\frac{\exp(x_{i3}\beta+\eta_i)}{1+\exp(x_{i3}\beta+\eta_i)}\right]^{d_3}\left[\frac{1}{1+\exp(x_{i3}\beta+\eta_i)}\right]^{1-d_3}$$

$$Pr(B|x_i)=p_0(x_i,\eta_i)^{d_0}[1-p_0(x_i,\eta_i)]^{1-d_0}\times\frac{1}{1+\exp(x_{i1}\beta+\rho d_0+\eta_i)}\times$$
$$\times\frac{\exp(x_{i2}\beta+\eta_i)}{1+\exp(x_{i2}\beta+\eta_i)}\times\left[\frac{\exp(x_{i3}\beta+\rho+\eta_i)}{1+\exp(x_{i3}\beta+\rho+\eta_i)}\right]^{d_3}\left[\frac{1}{1+\exp(x_{i3}\beta+\rho+\eta_i)}\right]^{1-d_3}$$

E la verosimiglianza condizionata alla statistica sufficiente diventa il prodotto dei contributi apportati dalle singole osservazioni, in altre parole il contributo dell'i-esima osservazione risulta essere :

$$L(\beta,\rho,\eta_i|X,S=1)=\prod_{i=1}^{N} P(A|A\cup B)^{y_{i1}} P(B|A\cup B)^{1-y_{i1}}$$

dove

$$\begin{aligned}&P(A|A\cup B)=1-P(B|A\cup B)\\&P(B|A\cup B)=\frac{P(B)}{P(A\cup B)}\end{aligned}\qquad(9)$$

e la probabilità a denominatore, date le probabilità appena definite, diventa

$$P(A\cup B)=\frac{p_0(x_i,\eta_i)^{d_0}[1-p_0(x_i,\eta_i)]^{1-d_0}}{1+\exp(x_{i1}\beta+d_0\rho+\eta_i)}\times$$
$$\times\left[\frac{\exp(x_{i1}\beta+d_0\rho+\eta_i)\exp(d_3x_{i3}\beta+d_3\eta_i)}{[1+\exp(x_{i2}\beta+\rho+\eta_i)][1+\exp(x_{i3}\beta+\eta_i)]}+\frac{\exp(x_{i2}\beta+\eta_i)\exp(d_3x_{i3}\beta+d_3\rho+d_3\eta_i)}{[1+\exp(x_{i2}\beta+\eta_i)][1+\exp(x_{i3}\beta+\rho+\eta_i)]}\right]$$

Nel calcolo della verosimiglianza condizionata, inoltre, analogamente a quanto visto per il caso statico, consideriamo solo le osservazioni per le quali si verifica un cambiamento nello stato fra il tempo 1 e il tempo 2 in quanto le rimanenti osservazioni non producono informazioni aggiuntive ai fini della stima, dando luogo a verosimiglianze equivalenti a quella considerata.

Si può, tuttavia, notare che nell'espressione (9) è ancora presente η_i. Ciò è dovuto al fatto che nel modello statistico parametrico considerato, nel quale in particolare sia presente un insieme di variabili esplicative oltre alla variabile risposta ritardata, la statistica S non è più statistica sufficiente per η_i e dunque condizionare la verosimiglianza a tale statistica non risolve il problema[2].

Honoré e Kyriazidou mostrano però che se $x_{i2} = x_{i3}$ allora diventa possibile semplificare l'espressione per la probabilità dell'unione e dunque la probabilità condizionata di B è

$$\begin{aligned}
&P(B|A\cup B, x_{i2}=x_{i3})=\\
&\qquad\frac{\exp(x_{i2}\beta+\eta_i)\exp(d_3 x_{i2}\beta+d_3\rho+\eta_i)}{\exp(x_{i1}\beta+d_0\rho+\eta_i)\exp(d_3 x_{i2}\beta+d_3\eta_i)+\exp(x_{i2}\beta+\eta_i)\exp(d_3 x_{i2}\beta+d_3\rho+d_3\eta_i)}\\
&\qquad=\frac{\exp(x_{i2}\beta+d_3\rho+\eta_i)}{\exp(x_{i1}\beta+d_0\rho+\eta_i)+\exp(x_{i2}\beta+d_3\rho+\eta_i)}\\
&\qquad=\frac{1}{1+\exp[(x_{i1}-x_{i2})\beta+(d_0-d_3)\rho]}
\end{aligned}$$

dove l'ultimo passaggio è stato ottenuto moltiplicando numeratore e denominatore per l'inversa del numeratore stesso.

In quest'ultima espressione non troviamo più la componente non osservabile η_i e dunque la verosimiglianza condizionata a tali eventi risulta calcolabile. Da (9) otteniamo infatti che

2 Chamberlain (1984)

$$P(A|A\cup B, x_{i2}=x_{i3}) = 1 - P(B|A\cup B, x_{i2}=x_{i3}) = \frac{\exp[(x_{i1}-x_{i2})\beta + (d_0 - d_3)\rho]}{1+\exp[(x_{i1}-x_{i2})\beta + (d_0 - d_3)\rho]}$$

e quindi la funzione di verosimiglianza condizionata calcolabile assume la seguente espressione :

$$L(\beta, \rho|x_i, S=1, x_{i2}=x_{i3}) = \prod_{i\in C} \frac{\exp[(x_{i1}-x_{i2})\beta + (d_0 - d_3)\rho]^{y_{i1}}}{1+\exp[(x_{i1}-x_{i2})\beta + (d_0 - d_3)\rho]} \qquad (10)$$

dove C è l'insieme $\{(y_{i0}, \ldots, y_{i3}, x_{i1}, \ldots, x_{i3}) \in \{0,1\}^4 \times \mathbb{R}^3 \mid y_{i1} + y_{i2} = 1, x_{i2} = x_{i3}\}$

Una volta definita l'espressione della verosimiglianza condizionata da massimizzare per ottenere le stime di interesse possiamo fare alcune considerazioni sulle restrizioni per l'identificabilità del modello appena definito. Le osservazioni influenti in fase di calcolo della verosimiglianza sono, infatti, le sole osservazioni appartenenti all'insieme C, ovvero le osservazioni per le quali :

- S = 1 ; deve dunque verificarsi una transizione nella variabile risposta fra il tempo 1 ed il tempo 2;
- $x_{i2} = x_{i3}$; abbiamo bisogno di osservazioni nelle quali le variabili condizionanti rimangano inveriate negli ultimi due periodi
- $x_{i1} \neq x_{i2}$; implicita nell'espressione (10) vi è, infatti, l'assunzione che esista variabilità fra le variabili esplicative nei primi due periodi di osservazione, tale condizione assicura che il parametro β non venga moltiplicato per quantità identicamente nulle e assicura dunque l'identificabilità del parametro stesso.

L'uguaglianza fra l'insieme condizionante ai tempi 2 e 3, assunzione, come abbiamo visto, cruciale per l'identificazione dei parametri, può, soprattutto in ambito econometrico, rappresentare un grosso ostacolo all'applicazione di questo modello a dati reali in quanto risulta spesso poco probabile che variabili tipicamente utilizzate come esplicative, quali reddito e consumo, possano rimanere invariate per due periodi di tempo successivi. Elemento da non dimenticare, inoltre, è il fatto che, come evidenziato nel terzo punto, le variabili in questione devono ammettere variabilità non nulla nei primi due periodi. La soluzione a questo problema, proposta dagli stessi Honoré e Kyriazidou, è quella di pesare le osservazioni in (10) con una funzione Kernel applicata alla differenza fra x_{i2} ed x_{i3}.

Le funzioni di tipo Kernel, tipicamente utilizzate per ottenere stime non-parametriche della funzione di densità di una variabile casuale, sono applicazioni definite in modo tale da assumere valore massimo per valori nulli e decrescono a 0 per valori, in valore assoluto, maggiori di 0. La funzione Kernel maggiormente utilizzata in questi ambiti risulta essere la funzione di Epanechnikov che assume la seguente forma

$$K(u)=\frac{3}{4}(1-u^2)1_{(|u|<1)}$$

Tale funzione assume valore massimo, pari a ¾, per $u = 0$ mentre assume valore 0 per valori di u non appartenenti all'intervallo (-1,1). Rimane da notare come rappreswnti essa stessa una funzione di densità in quanto non-negativa e tale che l'integrale esteso al supporto (-1,1) vale 1. Generalmente non viene utilizzata direttamente sotto questa forma ma si utilizza un fattore di scala h detto *bandwidth* che permette di variare a piacimento il supporto sul quale la funzione risulta definita. Infatti la funzione

$$\frac{1}{h} K\left(\frac{u}{h}\right) = \frac{3}{4h}\left[1-\left(\frac{u}{h}\right)^2\right] 1_{\left(\left|\frac{u}{h}\right|<1\right)}$$

ha come supporto l'intervallo (*-h* , *h*) a fronte delle stesse proprietà della funzione precedente, in particolare il fattore moltiplicativo introdotto consente che la funzione sommi ancora ad 1. Esistono in letteratura diversi tipi di funzioni kernel che riflettono a particolari esigenze in ambito di modellazione non parametrica, citiamo fra queste le funzioni di tipo uniforme, triangolo e quadratica con supporto (-1,1) e la funzione kernel gaussiana che si contraddistingue per essere definita sull'intero asse reale. Quest'ultima assume l'espressione di una funzione di densità Normale di media zero e varianza uno ed in questo caso i diversi valori assunti dalla *bandwidth* si traducono in una diversa varianza della densità ottenuta.

Una funzione Kernel definita sulla differenza fra x_{i2} ed x_{i3} attribuisce peso elevato ad osservazioni per le quali tale differenza risulta prossima a zero e peso minore per le restanti proporzionalmente al valore assoluto di tale differenza. Per quanto riguarda la scelta del parametro di *bandwith* Honorè e Kyriazidou pongono la sola restrizione che tale valore tenda a zero all'aumentare della numerosità campionaria a disposizione.

La funzione da massimizzare nella maggior parte della applicazioni sarà dunque

$$\prod_{i \in D} K\left(\frac{x_{i2}-x_{i3}}{h}\right) \frac{\exp[(x_{i1}-x_{i2})\beta+(d_0-d_3)\rho]^{y_{i1}}}{1+\exp[(x_{i1}-x_{i2})\beta+(d_0-d_3)\rho]}$$

con $D=\left\{(y_{i0},\ldots,y_{i3})\in\{0,1\}^4 \mid y_{i1}+y_{i2}=1\right\}$

Honorè e Kyriazidou forniscono inoltre la distribuzione asintotica, sotto condizioni di regolarità piuttosto deboli eccezion fatta per l'assunzione di continuità delle variabili appartenenti all'insieme condizionante, di tale stimatore affermando che essa risulta essere normale.

L'utilizzazione di un insieme di variabili condizionanti di dimensione superiore ad uno passa per l'individuazione di una funzione Kernel definita in $\mathbb{R}^k$ a valori reali, dove *k* è la dimensione del vettore di variabili esplicative. A questo proposito una generalizzazione della funzione Kernel di Epanechnikov è la seguente :

$$K(u) \propto (1-u^T u) 1(u^T u<1)$$

dove il segno di proporzionalità si deve al fatto che tale funzione non integra ad uno sul suo supporto. Nonostante siano disponibili funzioni Kernel definite su spazi multidimensionali la loro utilizzazione nella pratica si scontra con problemi dal punto di vista computazionale sempre maggiori all'aumentare della dimensione del vettore condizionante dovuti principalmente alla scelta della *bandwidth* più opportuna ed alla identificazione della relativa costante di proporzionalità. Tali difficoltà si traducono in ambito pratica in un'accurata analisi delle variabili incluse nel modello, evitando dunque di includere variabili già ritenute poco significative ed evitando più in generale di includere un numero di variabili molto elevato.

Un ultimo risultato deve infine essere riportato riguardo a questo stimatore e riguarda il rateo di convergenza asintotica. Honoré e Kyriazidou mostrano infatti come lo stimatore risultante converga alla distribuzione asintotica ottenuta con una velocità inferiore rispetto all'usuale $\frac{1}{\sqrt{n}}$.

2.2 Modelli lineari per dati di panel

Come abbiamo avuto modo di vedere nei paragrafi precedenti la stima di modelli a risposta discreta per dati di panel è fortemente condizionata dalle ipotesi necessarie per l'identificazione. Sia per il caso statico che per quello dinamico, infatti, l'identificazione, almeno per i modelli trattati in precedenza, passa per l'assunzione di una componente di errore di tipo logistico e la stima dei parametri avviene per sfruttando il solo gruppo di osservazioni per le quali venga osservata transizione nei periodi intermedi, nel caso dinamico, inoltre, lo stimatore proposto ha un rateo di convergenza inferiore a $\frac{1}{\sqrt{n}}$. Nel tentativo di proporre una più semplice soluzione al problema di modellazione sin qui proposto che permetta di rilassare alcune ipotesi, tratteremo, nei successivi due paragrafi, il problema della stima di modelli di tipo lineare.

Sin qui la probabilità di successo è stata modellata attraverso una funzione di una combinazione lineare delle esplicative e di un effetto fisso di individuo. Un metodo alternativo è tuttavia rappresentato dalla modellazione della probabilità di possesso direttamente con una combinazione lineare delle esplicative. La probabilità cercata sarà dunque:

$$Pr(y_{it}|x_{it},\eta_i)=x_{it}'\beta+\eta_i$$

Ciò corrisponde a non imporre alcuna restrizione sulla componente di errore (eccezion fatta per la media nulla condizionatamente all'insieme informativo) e ad assumere di osservare direttamente y_{it}^* nel modello in (2).

2.2.1 Modello lineare classico

La scelta di un errore di tipo classico, e dunque la scelta di non imporre restrizioni né alla forma né al supporto della funzione di densità relative all'errore, in un contesto nel quale la quantità modellata è una probabilità offre l'indubbio svantaggio che nulla assicura che Pr (y_{it} | x_{it} , η_i) sia compresa fra 0 e 1. D'altro canto, tuttavia, la stima di questo tipo di modelli risulta, almeno dal punto di vista computazionale, notevolmente facilitata e dunque, almeno in prima approssimazione, la loro scelta può risultare opportuna.

Un'ulteriore considerazione in favore di un modello lineare deriva del fatto che il vettore di coefficienti β così stimato offre un'interpretazione immediata. I coefficienti così ottenuti possono infatti essere direttamente interpretati come la variazione nella probabilità di successo dovuta all'aumento di una unità della variabile associata al coefficiente stesso. Infatti se $Pr(y_{it}|x_{it},\eta_i)=x_{it}'\beta+\eta_i$ allora

$$\frac{\partial Pr(y_{it}|x_{it},\eta_i)}{\partial x_{it}'}=\beta$$

La trattazione di modelli di tipo lineare con la presenza di un effetto specifico di individuo rappresenta il classico caso di stima di un modello lineare per dati di panel, la stima del quale dipende in maniera cruciale dalle considerazioni sulla componente non osservabile.

Il dataset a nostra disposizione, che supponiamo bilanciato, è in generale composto da T diverse osservazioni su ciascuno degli N individui oggetto dell'indagine, si tratta dunque di una matrice di N*T righe e p+1 colonne, dove p è il numero di variabili esplicative incluse. Una stima consistente della quantità di interesse β può essere ottenuta dunque con una stima OLS del dataset in esame sotto le seguenti condizioni:

$$E(\eta_i|x_{it}')=0$$
$$E(\mu_{it}|x_{it}')=0$$

Mentre la seconda condizione, corrispondente a supporre l'esogenità dell'insieme condizionante, può, seppure a volte con qualche riserva, essere considerata vera, possiamo difficilmente ritenere giustificata, almeno per il caso in questione, l'assunzione che la componente specifica di individuo risulti incorrelata con le variabili condizionanti. L'avversione al rischio del soggetto, le preferenze in materia di scelte in campo finanziario, l'abilità in generale del soggetto sono tutte variabili impossibili da misurare con oggettività e sono dunque tutte quantità che vanno a finire, all'interno del nostro modello, in η_i e possiamo ragionevolmente ritenere che le variabili condizionanti che andremo ad utilizzare in fase di stima quali il reddito e la ricchezza siano correlate con tale componente. Possiamo infatti ritenere che, ad esempio, un soggetto con un alto livello di reddito possegga un profilo di rischio piuttosto diverso da un soggetto con un livello di reddito più basso.

Questa considerazione ci porta ad escludere, in quanto le assunzioni imposte sono difficilmente applicabili al nostro caso, quella parte di modelli, definiti in letteratura come modelli ad effetti casuali, che si basa sull'incorrelazione di η_i rispetto all'insieme di variabili condizionanti come condizione necessaria per la consistenza della stima ottenuta. Sotto tale ipotesi infatti la stima OLS del modello[3] (2) produce stime consistenti dei parametri di interesse in β, anche se tale stima non è efficiente

3 Lo stimatore prende il nome di *pooled OLS* dal momento che siamo in un contesto di dati di panel. La procedura di stima tuttavia è identica al caso OLS in cui il dataset è di dimensioni NT*p

dal momento che la componente di errore comprende la variabile η_i costante nel tempo e che produce dunque correlazione seriale. Ciò si traduce nel fatto che l'errore non soddisfa le ipotesi del teorema di Gauss-Markov e lo stimatore ottenuto non è quindi BLUE (*Best Linear Unbiased Estimator*). Uno stimatore efficiente può comunque essere ottenuto applicando una trasformata GLS ai dati sfruttando il fatto che la matrice di varianza e covarianze dell'errore ha una struttura, sotto l'assunzione che μ_{it} sia uno shock casuale, nota e dunque non abbiamo bisogno di ricorrere alla procedura FGLS (*Feasible Generalized Least Squares*).

Un approccio che consente, invece, alla componente non osservabile di essere correlata con le variabili condizionanti è il cosiddetto modello ad effetti fissi che consiste nel valutare η_i come un parametro da stimare, un effetto fisso appunto. Esso è tuttavia un parametro cosiddetto *incidentale* nel senso che, essendo specifico di individuo, dovremmo stimare N di questi parametri e il problema non può essere risolto disponendo di osservazioni aggiuntive, in quanto con esse aumenta anche il numero di parametri coinvolti. Ciò che utilizziamo infatti per stimare η_i sono le T differenti osservazioni disponibili per quel particolare soggetto, e dunque la stima di questi parametri con il metodo OLS[4] risulterebbe consistente solo per $T\rightarrow\infty$ e non, come nel caso standard per $N\rightarrow\infty$.

Ciò a cui siamo generalmente interessati è tuttavia β e l'approccio classico passa attraverso il modello in differenze prime, ovvero :

$$\Delta y_{it}=\Delta x_{it}\beta+\Delta \mu_{it} \qquad (3)$$

4 La stima OLS in questo caso sarebbe quella a partire da un modello che include un'intercetta per ogni singolo individuo e che quindi include N+p parametri

dove Δ è l'operatore differenza prima. Il modello che ne risulta può quindi essere stimato in maniera consistente sotto la condizione di stretta esogenità delle variabili condizionanti dal momento che tale trasformazione non dipende dall'effetto fisso dal momento che esso è costante nel tempo. Analogamente a quanto accennato per il modello ad effetti casuali però tale stima non risulta in generale efficiente poiché il nuovo errore trasformato non gode di buone proprietà anche sotto l'assunzione che l'errore di partenza fosse *white noise*. La procedura che ci permette di ottenere una stima efficiente del modello ad effetti fissi, sotto le ipotesi di stretta esogenità delle variabili condizionanti e di errore di partenza di tipo *white noise*, passa attraverso una opportuna trasformata GLS di questo modello. In questo caso, diversamente da prima, la struttura di autocorrelazione dell'errore trasformato non risulta nota e dunque la procedura di stima sarà del tipo FGLS che utilizza una stima della matrice di varianze e covarianze degli errori.

Si dimostra infine che, sotto le precedenti assunzioni, lo stimatore efficiente ottenuto applicando la traformata GLS al modello (3) corrisponde ad una stima OLS del modello in scarti dalla media di gruppo, ovvero alla stima OLS di :

$$y_i^s = X_i^s + \mu_i^s$$

dove y_i^s è il vettore la cui j-esima componente è la differenza fra y_{ij} e la media degli y_{it} e così anche per X_i^s e per μ_i^s.

Fino ad ora abbiamo considerato il caso in cui la componente di errore fosse un *white noise* incorrelato con l'insieme di variabili esplicative, tuttavia, come abbiamo potuto notare in precedenza, i dati a nostra disposizione sono di tipo dicotomico e dunque possono essere considerati realizzazioni di variabili casuali di Bernoulli con una probabilità di successo π_{it} che dipende dai valori assunti dalle variabili esplicative tramite la relazione

$$\pi_{it}=x_{it}'\beta+\eta_i$$

e quindi

$$Var(y_{it})=\pi_{it}(1-\pi_{it})=f(x_{it}'\beta+\eta_i)$$

Gli errori nel modello lineare sono dunque per costruzione eteroschedastici e la varianza dipende dai valori assunti dalle variabili esplicative.

Dal punto di vista applicativo si tratterà dunque di ottenere stime robuste all'ipotesi di eteroschedasticità degli errori dal momento che una tale assunzione è indotta dal disporre di dati a risposta discreta.

Un'ultima considerazione su questo modello riguarda il fatto che le trasformazioni operate sui dati, prima in differenza prime e quindi in scarti dalla media di gruppo, presentano l'inconveniente di rendere identicamente nulla una eventuale variabile che rimanesse costante nel tempo. In fase di stima, infatti, una variabile di questo tipo non potrebbe essere distinta, per come abbiamo definito il modello, dalla componente non osservabile e dunque, con l'obbiettivo di eliminare l'una, eliminiamo anche l'altra. In questo genere di modelli quindi non può essere inclusa una variabile costante nel tempo in quanto nel modello in scarti dalla media di gruppo la variabile non compare più nell'espressione. Il fatto di non poter includere tali variabili non dà luogo tuttavia ad una mancata consistenza delle stime a causa delle variabili omesse in quanto tali variabili sono implicitamente incluse nella componente η_i, dal momento che essa rappresenta tutto ciò che non è incluso nel modello e costante nel tempo. Rimane però il problema associato all'impossibilità di ottenere stime sui coefficienti associati, qualora tali quantità fossero di interesse. Esistono a tal proposito procedure, come quella proposta da Hausman-Taylor, per consentono di reperire stime anche per variabili costanti nel tempo che si basano sulla stima preliminare proposta in precedenza. Una volta reperita tale stima, infatti, è possibile regredire i residui del modello in livelli sulle variabili costanti nel tempo. Tale procedimento, sotto la condizione che queste ultime variabili siano incorrelate con l'effetto fisso η_i, produce stime consistenti per i parametri di interesse e permette dunque di valutare l'effetto indotto da variabili costati nel tempo.

2.2.2 Stimatore di Arellano-Bond

Come abbiamo visto per il caso non lineare, l'aggiunta della variabile endogena ritardata nell'insieme delle variabili condizionanti introduce endogenità all'interno del modello. Per vedere come sia possibile pervenire ad una stima nel caso lineare possiamo provare ad utilizzare lo stesso procedimento che ci ha portato alla stima del modello statico, passando cioè al modello in differenze prime :

$$\Delta y_{it} = \Delta X_{it} \beta + \rho (y_{it-1} - y_{it-2}) + (\mu_{it} - \mu_{it-1}) \qquad (8)$$

Come si può facilmente notare l'errore è ora composto dallo shock al tempo t e al tempo t-1 e dal momento che $\Delta y_{it-1} = y_{it-1} - y_{it-2}$ include y al tempo t-1 abbiamo che Δy_{it-1} risulta *per costruzione* correlata con il nuovo errore generato dalla trasformazione in differenze prime. Lo stimatore OLS applicato a questo modello produce dunque stime inconsistenti dei parametri di interesse.

Il problema nella stima di questo modello, deriva dal fatto che la nuova variabile risulta endogena al modello stesso, tuttavia, a differenza del caso non-lineare, qui disponiamo di una possibilità in più: la stima a variabili strumentali, appositamente ideata per risolvere problemi in cui alcune variabili condizionanti risultino correlate con il termine di errore. Ciò di cui abbiamo bisogno è una variabile correlata con la variabile endogena ed incorrelata con l'errore. Nel caso in questione una possibile soluzione è quella proposta da Anderson-Hsiao e che consiste nell' utilizzare come variabile strumentale y_{it-2} che, come si può facilmente verificare da (8), risulta correlata con Δy_{it-1} e incorrelata con l'errore $\Delta\mu_{it}$.

La stima a variabili strumentali ottenuta utilizzando la variabile y al ritardo 2 come strumento per la stima del modello in (8) produce, a questo punto per costruzione, stime consistenti dei parametri di interesse β e ρ.

Un'ulteriore sguardo al modello permette di notare che anche i valori ritardati per più di 2 periodi della variabile risposta sono incorrelati con l'errore, mentre, dal momento che possiamo scrivere :

$$y_{it}=f(y_{it-1}, y_{it-2}, \ldots)$$

sono correlati con Δy_{it-1} e sono quindi strumenti validi.

Uno stimatore a varianza minore del precedente può dunque essere ottenuto applicando il metodo generalizzato dei momenti dove tra le condizioni di ortogonalità vengano inserite anche le incorrelazioni fra tutti i ritardi di y dopo il secondo. Tale stimatore si dimostra essere lo stimatore efficiente in casi come questo e prende il nome di *stimatore di Arellano-Bond.*

Uno dei principali problemi connessi all'utilizzo della variabile dipendente ritardata come strumento per la stima dei parametri di interesse è rappresentato dalle difficoltà introdotte da strumenti deboli. E' possibile dimostrare, infatti, che la correlazione fra lo strumento utilizzato ed il regressore endogeno si indebolisce se il rapporto fra la varianza dell'effetto fisso e la varianza dell'errore tende all'infinito, oppure se il valore assoluto di ρ tende ai valori di soglia 0 o 1. Gli strumenti utilizzati risultano dunque deboli nel caso in cui l'errore abbia una varianza molto grande rispetto a quella dell'effetto fisso o nel caso in cui il processo autoregressivo sia ad alta o scarsa persistenza. Le motivazioni di quest'ultimo risultato sono, almeno dal punto di vista intuitivo, da ricercare nel fatto che se $\rho = 1$, una volta trasformato il modello in differenze prime, il regressore trasformato diventa un *white noise* e, dal momento che un *white noise* altro non è che uno shock puramente casuale, diventa impossibile riuscire ad identificare una variabile ad esso correlata. In maniera analoga se ρ vale 0 il processo è sostanzialmente un *white noise* e risulta dunque difficile utilizzare come strumenti valori ritardati dal momento che essi sono indipendenti fra loro.

2.3 Modello per dati a rilevazione irregolare

I modelli analizzati fin qui trattano casi standard nei quali la specificazione del modello viene fatta a partire da assunzioni verificate nella maggior parte dei casi e che vanno dunque a coprire un numero di applicazioni molto vasto. Quanto segue, invece, tratta un modello che emerge nel nostro caso specifico a causa della natura particolare dei dati a disposizione. Si tratta dunque di un modello di più limitata applicazione che trova giustificazione nel tentativo di modellare dati che, per loro natura, non soddisfano le usuali assunzioni supposte vere nei modelli trattati sin qui. Per le finalità che ci proponiamo, infatti, questi modelli risultano adatti se, all'interno del nostro dataset, considerassimo solo gli anni di indagine che vanno dal 1998 al 2004 o solo gli anni dal 1989 al 1995, per quanto riguarda periodi di tempo che comprendano anni da entrambi i gruppi, invece, i dati in nostro possesso possono presentare complicazioni nell'applicazione di queste metodologie. Come abbiamo potuto notare in precedenza, infatti, i dati a disposizione vanno dal 1989 al 2004 ad intervalli di due anni eccezion fatta per l'anno 1997, nel quale si è verificato un problema nella raccolta dei dati che ha portato a ritardare di un anno la relativa indagine. Dal punto di vista statistico, dunque, ciò che abbiamo a disposizione è un dataset di tipo panel non equispaziato nel tempo. Le analisi effettuate fino a questo punto si basavano sull'implicita assunzione che il dataset fosse equispaziato ed il venir meno di questa ipotesi basilare, vera nella quasi totalità dei casi, induce la necessità di rivedere la forma degli stimatori proposti nel momento in cui si decida di utilizzare anche osservazioni osservate per periodi pre-1998 e post-1998.

Per quanto riguarda i modelli di tipo statico risulta subito evidente come l'utilizzo di panel non equispaziati non introduca complicazioni a livello di stima del modello e dunque gli stimatori definiti in precedenza sono ugualmente validi se applicati a questi ultimi dati.

Considerare un modello di tipo dinamico introduce, invece, problemi nel momento in cui si vada a definire la struttura del modello per l'anno 1998. Nel caso in questione, infatti, considerando il modello definito in (7) la variabile risposta ritardata si riferisce al valore della variabile y per t=1996, ma noi non disponiamo di tale osservazione e dunque tale modello non può essere utilizzato per trattare i dati a nostra disposizione. Dal punto di vista intuitivo risulta evidente come la presenza di un "salto" temporale maggiore per t = 1998 si rifletta in una diversa dipendenza fra y_{i1998} e la variabile ritardata a nostra disposizione y_{i1995} rispetto alle altre coppie y_{it} e y_{it-2}[5]. Questa diversa dipendenza si riflette in un diverso valore di ρ associato all'anno 1998 e di conseguenza il modello (7) risulta di difficile applicazione.

Nel seguito andremo a definire nel dettaglio un modello opportuno per i nostri dati e proporremo un metodo di stima che ci possa portare a ad ottenere stime delle quantità di interesse anche in casi come questo. Nel farlo prenderemo in esame, per semplicità, solo modelli di tipo lineare, nonostante i già citati limiti di un tale modo di operare, in cui non siamo presenti ulteriori variabili esplicative all'infuori della variabile risposta ritardata.

2.3.1 La struttura del modello

Per meglio valutare i singoli periodi di tempo valutiamo la variabile risposta come generata da un AR(1) con la presenza di un trend differente per periodi antecedenti e posteriori al 1998 :

$$y_{it} = A + By_{it-1} + Ct + Dtd_{t \geqslant 98} + \boldsymbol{x}_{\boldsymbol{it}}{}' \boldsymbol{E} + \eta_i \qquad (10)$$

In cui t rappresenta la variabile anno e $d_{t \geq 1998}$ è una variabile dummy che assume valore 1 per t ≥ 1998 e 0 altrimenti. Inoltre A, B, C e D sono scalari mentre E è un vettore di coefficienti associati alle variabili presenti nel vettore di esplicative x_{it} di

5 D'ora in poi t rappresenterà l'effettivo anno nel quale viene svolta l'indagine e non, come nei paragrafi precedenti, il numero progressivo dell'indagine considerata dal momento che le due notazioni non sono più equivalenti.

dimensione *p* x 1.

La presenza di un trend diverso per valori successivi al 1998 deriva dalla verifica effettuata a questo proposito, specificatamente sui dati in nostro possesso, per le osservazioni osservate dal 1989 al 1995 e dal 1998 al 2004 e che, non includendo la dipendenza fra 1998 e 1995, non risentono del problema in questione.

Sostituendo il valore di y_{it-1} in (10) otteniamo :

$$y_{it}=A+B[A+By_{it-2}+C(t-1)+D(t-1)d_{t-1\geqslant 1998}+\boldsymbol{x_{it-1}}'\boldsymbol{E}+\eta_i]+Ct+Dtd_{t\geqslant 1998}+$$
$$+\boldsymbol{x_{it}}'\boldsymbol{E}+\eta_i$$
$$y_{it}=A(1+B)+B^2 y_{it-2}+C[B(t-1)+t]+D[B(t-1)d_{t-1\geqslant 1998}+td_{t\geqslant 1998}]+$$
$$+\boldsymbol{x_{it}}'(B+1)\boldsymbol{E}+(B+1)\eta_i$$

dove abbiamo assunto $x_{it} = x_{it-1}$ dal momento che nonn possediamo l'informazione relativa alle variabili esplicative al periodo immediatamente precedente, inoltre dal momento che $d_{t\geq 1998} = d_{t-1\geq 1998}$ per $t\neq 1998$ abbiamo che

$$y_{it}=A(B+1)-CB+B^2 y_{it-2}-DBd_{t\geqslant 1998}+C(B+1)t+D(B+1)td_{t\geqslant 1998}+ \qquad (11)$$
$$+\boldsymbol{x_{it}}'(B+1)\boldsymbol{E}+(B+1)\eta_i$$

o analogamente

$$y_{it}=\alpha+\rho\, y_{it-2}+\beta\, d_{t\geqslant 1998}+\phi\, t+\psi\, td_{t\geqslant 1998}+\boldsymbol{x_{it}}'\boldsymbol{\omega}+\tau\,\eta_i \qquad (12)$$

in cui i parametri β e ψ sono legati dalla relazione $\psi=\frac{-\beta(\sqrt{\rho}+1)}{\sqrt{\rho}}$, mentre $\tau=(\sqrt{\rho}+1)$.

Resta da notare come in questo caso, dal momento che abbiamo definito ρ come B^2, sia stata implicitamente assunta la non negatività di ρ, assunzione peraltro ragionevole data l'interpretazione economica del parametro. E' plausibile infatti assumere che la probabilità di possedere determinati titoli rischiosi sia superiore se tali titoli erano già in possesso del soggetto al tempo precedente, mentre risulterebbe sorprendente che il possesso al tempo precedente generasse una minore probabilità di possesso nel periodo successivo.

Il modello fin qui definito risulta appropriato per la modellazione della variabile risposta per t ≠ 1998. Per l'anno 1998 infatti la variabile risposta dipende da se stessa ritardata di tre periodi e, dunque, partendo da (11) abbiamo che :

$$y_{it}=A(B^2+B+1)-CB(1+2\mathrm{B})+B^3\,y_{it-3}-DB(1+2\mathrm{B})\,d_{t\geqslant 1998}+C(B^2+B+1)t+$$
$$+D(B^2+B+1)td_{t\geqslant 1998}+\boldsymbol{x_{it}}'(B^2+B+1)\boldsymbol{E}+(B^2+B+1)\eta_i$$

ovvero

$$y_{it}=\tilde{\alpha}+\tilde{\rho}\,y_{it-3}+\tilde{\beta}\,d_{t\geqslant 1998}+\tilde{\phi}\,t+\tilde{\psi}\,td_{t\geqslant 1998}+\boldsymbol{x_{it}}'\tilde{\boldsymbol{\omega}}+\tilde{\tau}\,\eta_i \qquad (13)$$

Come detto quest'ultimo modello si applica ai nostri dati per t = 1998 e possiamo quindi andare a definire il modello per il generico t. Utilizzando la variabile dummy d_{98} che vale 1 per t = 1998 e 0 altrimenti otteniamo :

$$y_{it}=\alpha+\mu\,d_{98}+\rho\,y_{iLAG}+\gamma\,y_{iLAG}\,d_{98}+\beta\,d_{t\geqslant 1998}+\lambda\,d_{t\geqslant 1998}\,d_{98}+\phi\,t+\kappa\,td_{98}+$$
$$+\psi\,td_{t\geqslant 1998}+\delta\,td_{t\geqslant 1998}\,d_{98}+\boldsymbol{x_{it}}'\boldsymbol{\omega}+\boldsymbol{x_{it}}'\,d_{98}\boldsymbol{\iota}+\tau\,\eta_i+\theta\,\eta_i\,d_{98} \qquad (14)$$

sotto i vincoli :

$$\begin{array}{ll} \alpha+\mu=\tilde{\alpha} & \\ \rho+\gamma=\tilde{\rho} & \phi+\kappa=\tilde{\phi} \\ \beta+\lambda=\tilde{\beta} & \psi+\delta=\tilde{\psi} \\ \boldsymbol{\omega}+\boldsymbol{\iota}=\tilde{\boldsymbol{\omega}} & \tau+\theta=\tilde{\tau} \end{array} \qquad (15)$$

dove y_{iLAG} rappresenta la variabile y ritardata disponibile, essa sarà dunque y_{i1995} per l'anno 1998 e y_{it-2} altrimenti.

Questa espressione incorpora le due precedenti nel senso che per t = 1998 diventa la (13) mentre per i restanti periodi l'espressione risulta analoga alla (12).

Dall'espressione (14) si può tuttavia notare come alcune variabili siano collineari ad altre ed in particolare

$$d_{t\geqslant 1998}\, d_{98} = d_{98}$$
$$td_{98} = 1998\, d_{98}$$

sostituendo nella (14) e raccogliendo otteniamo

$$\begin{aligned} y_{it} = {} & \alpha + (\mu + \lambda + 1998\,\kappa + 1998\,\delta)\, d_{98} + \rho\, y_{iLAG} + \gamma\, y_{iLAG}\, d_{98} + \beta\, d_{t\geqslant 1998} + \phi\, t + \\ & + \psi\, td_{t\geqslant 1998} + \boldsymbol{x}_{it}'\boldsymbol{\omega} + \boldsymbol{x}_{it}'\, d_{98}\,\boldsymbol{\iota} + \tau\,\eta_i + \theta\,\eta_i\, d_{98} \end{aligned} \tag{16}$$

e dunque questo risulta il modello che si adatta alla totalità delle osservazioni a disposizione, indipendentemente dagli anni nei quali sono state effettuate le misurazioni.

Nel seguito ci proponiamo di fornire un metodo di stima per l'espressione ottenuta e saremo interessati, in modo particolare, alla stima del parametro ρ che, data la natura del problema in esame, risulta la quantità di interesse.

Come appare evidente a prima vista l'espressione (16) presenta, in fase di stima, tre problemi fondamentali :

- La presenza di una componente non osservabile per costruzione variabile nel tempo che non può dunque essere trattata come abbiamo visto finora per modelli di tipo lineare.
- La presenza della variabile risposta ritardata che introduce endogenità nel modello.

- L'esistenza di vincoli non lineari sui parametri che influiranno in maniera significativa sulla procedura di stima da utilizzare.

Nei successivi paragrafi proporremo una procedura di stima che tenga conto di queste particolarità e che possa quindi essere utilizzata per ottenere stime consistenti dei parametri di interesse.

2.3.2 La trasformazione in quasi-differenze

L'interazione fra l'effetto fisso e la variabile dummy comporta che l'utilizzo della già menzionata trasformazione in differenze prime non porta all'eliminazione dall'espressione dell'effetto fisso con la conseguente impossibilità di stimare il modello. Non è possibile quindi, in questo caso, far ricorso al classico metodo per la trattazione della componente non osservabile: possiamo tuttavia utilizzare lo stesso principio che ne sta alla base.

L'idea di fondo, nel caso lineare, era quella di individuare una opportuna trasformazione dei dati che porti ad eliminare dal modello la presenza di η_i mantenendo definiti i restanti parametri. Nel caso in questione, un'opportuna trasformazione dei dati potrebbe essere rappresentata da quella che viene definita una trasformazione in quasi-differenze. Si tratta di sottrarre ad ogni variabile se stessa ritardata pre-moltiplicata per una costante *k*, si tratta cioè di definire la trasformazione $y_{it} - k\, y_{it-2}$ attraverso l'identificazione di un opportuno valore di *k*. Questa classe di trasformazioni dei dati racchiude come caso particolare la trasformazione in differenze prime caratterizzata dal fatto che $k = 1$. E' importante tuttavia notare che nulla vieta alla quantità *k* di variare con *t* e, dunque, di sottrarre alla variabile risposta la variabile ritardata moltiplicata per valori differenti nel tempo. Facciamo notare, infine, come tale trasformazione possa essere rappresentata da una matrice che, applicata ad un vettore di osservazioni, produca il vettore di osservazioni trasformate desiderato, infatti siano Π la matrice di trasformazione e υ un vettore di osservazioni definite come

$$\Pi=\begin{bmatrix} 1 & -k_T & 0 & \cdots & & 0 \\ 0 & 1 & -k_{T-1} & 0 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1 & -k_2 \end{bmatrix} \qquad \nu_i=\begin{bmatrix} \nu_{iT} \\ \vdots \\ \nu_{i1} \end{bmatrix} \qquad (17)$$

otteniamo che i dati trasformati non sono altro che $\Pi \nu$.

La trasformazione che cerchiamo deve essere tale che, applicata al vettore di osservazioni $\nu_{it}=\tau \eta_i+\theta \eta_i d_{98}$ restituisca una quantità indipendente da η_i. Il vettore in questione, per t ≠ 1998 vale costantemente $\tau \eta_i$ e dunque per tutte le differenze che non coinvolgano valori al tempo 1998 il valore appropriato di k risulta essere ancora 1. Una diversa trattazione deve, invece, essere riservata alle differenze costruite utilizzando i valori di ν per l'anno 1998, in tal caso infatti

$$\begin{aligned} \nu_{i2000}-\nu_{i1998}&=\tau \eta_i-(\tau+\theta)\eta_i=-\theta \eta_i\neq 0 \\ \nu_{i1998}-\nu_{i1995}&=(\tau+\theta)\eta_i-\tau \eta_i=\theta \eta_i\neq 0 \end{aligned}$$

e dunque per tali differenze un valore di *k* pari a 1 non elimina dall'espressione l'effetto fisso. E' possibile, tuttavia, individuare un valore di *k* che, moltiplicato per la quantità sottratta, produca un risultato pari a zero. Per quanto riguarda la prima differenza, infatti, risolvendo

$$\tau-k_{2000}(\tau+\theta)=0 \quad \Rightarrow \quad k_{2000}=\frac{\tau}{\tau+\theta}=\frac{\tau}{\tilde{\tau}} \qquad (18)$$

otteniamo una valore di k che rende la differenza $\nu_{i2000} - k\nu_{i1998}$ indipendente dall'effetto fisso non osservabile. In modo analogo possiamo ricavare un opportuno valore per la seconda differenza dato da

$$(\tau+\theta)-k_{1998}\tau=0 \quad \Rightarrow \quad k_{1998}=\frac{\tau+\theta}{\tau}=\frac{\tilde{\tau}}{\tau}=k_{2000}{}^{-1} \qquad (19)$$

Applicando, dunque, la trasformazione in (17) in cui i valori di k_t sono identicamente 1 eccetto per i valori di k_{2000} e di k_{1998} definiti secondo la (18) e la (19), otteniamo una matrice Π che trasforma i dati così da eliminare la dipendenza dalla componente non osservabile in maniera analoga a quanto visto per il modello lineare classico. Il problema a questo punto rimane la presenza all'interno dell'espressione di k del coefficiente ρ. Infatti per come sono stati definiti i coefficienti dei modelli

$$\tau=\sqrt{\rho}+1 \qquad \tilde{\tau}=\rho+\sqrt{\rho}+1$$

$$k_{2000}=\frac{\sqrt{\rho}+1}{\rho+\sqrt{\rho}+1} \qquad k_{1998}=\frac{\rho+\sqrt{\rho}+1}{\sqrt{\rho}+1}$$

Questa espressione mette in luce un vincolo, a prima vista inesistente, sul coefficiente ρ da stimare. Il coefficiente da stimare deve infatti essere vincolato ad assumere lo stesso valore utilizzato per ricavare i coefficienti k con i quali si operano le trasformazioni sui dati. Tale vincolo potrà venire implementato, almeno nel nostro caso, attraverso l'individuazione di un metodo di stima che ci permetta di stimare il modello, a partire da un valore iniziale per ρ, e di utilizzare questa stima per differenziare nuovamente i dati a partire dal valore stimato per ρ ed iterare tale procedura per ottenere una convergenza fra il valore utilizzato per ricavare i valori di k e la stima di ρ ottenuta a partire dalla differenziazione operata.

2.3.3 Stima a variabili strumentali vincolata

L'individuazione del modello da utilizzare per la stima ci porta ad analizzare il secondo problema citato in precedenza : l'endogenità introdotta dalla variabile risposta ritardata presente fra le esplicative. Il problema, tuttavia, può venire facilmente aggirato, analogamente a quanto visto in precedenza, stimando il modello utilizzando i ritardi superiori al primo della variabile risposta come variabili strumentali. La stima ottenuta con questo metodo risulta consistente mentre, riguardo all'efficienza, avremmo bisogno di individuare la combinazione di stimatori a variabili strumentali che minimizzi la varianza delle stime in analogia con quanto dimostrato da Arellano-Bond per il modello visto in precedenza.

Per ottenere una stima di tale modello, tuttavia, abbiamo bisogno di passare al modello in quasi-differenze definito in precedenza, giungendo quindi alla seguente espressione :

$$\begin{aligned}\boldsymbol{y}_i = {} & \alpha \Pi \mathbf{1} + (\mu + \lambda + 1998\kappa + 1998\delta) \Pi \boldsymbol{d}_{98} + \rho \Pi \boldsymbol{y}_{iLAG} + \gamma \Pi \boldsymbol{y}_{iLAG} \boldsymbol{d}_{98} + \\ & + \beta \Pi \boldsymbol{d}_{t \geqslant 1998} + \phi \Pi \boldsymbol{t} + \psi \Pi \boldsymbol{t} \boldsymbol{d}_{t \geqslant 1998} + \Pi X_i' \boldsymbol{\omega} + \Pi X_i' \boldsymbol{d}_{98} \boldsymbol{\iota}\end{aligned} \tag{20}$$

dove ora le variabili considerate sono vettori di dimensione T-1 ad eccezione di X_i che rappresenta invece una matrice *p* x (T-1), dove T è il numero di periodi nei quali la relativa osservazione è stata osservata. Le variabili di interazione $\mathbf{y}_{\mathbf{iLAG}}\mathbf{d}_{\mathbf{98}}$ e $\mathbf{td}_{\mathbf{t\geq1998}}$ sono inoltre ottenute moltiplicando i vettori elemento per elemento e non, dunque, attraverso il prodotto vettoriale, essi sono dunque, in generale, così costituiti :

$$\boldsymbol{y}_{iLAG}\boldsymbol{d}_{98}=\begin{bmatrix}0\\0\\0\\y_{i1995}\\0\\0\\0\end{bmatrix} \qquad \boldsymbol{t}\,\boldsymbol{d}_{t\geqslant 1998}=\begin{bmatrix}0\\0\\0\\1998\\0\\0\\0\end{bmatrix}$$

Una volta operata la trasformazione, tuttavia, notiamo come emerga collinearità fra la costante e la variabile dummy per l'anno 1998 trasformate. Nel caso di un soggetto osservato dal 1989 al 2004, per il quale dunque il modello (16) risulta definito per gli anni che vanno dal 1991 al 2004, infatti, esse diventano

$$\Pi\,\mathbf{1}=\Pi\times\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\1-k\\1-k^{-1}\\0\\0\end{bmatrix} \qquad \Pi\,\boldsymbol{d}_{98}=\Pi\times\begin{bmatrix}0\\0\\0\\1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}0\\0\\-k\\1\\0\\0\end{bmatrix}$$

da cui si nota che $\Pi\,\boldsymbol{d}_{98} \;=\; \frac{k+1}{k}\Pi\,\mathbf{1}$.

Da (20) otteniamo quindi

$$\begin{aligned}\boldsymbol{y}_i=&(\alpha+\mu\,c+\lambda\,c+1998\,\kappa\,c+1998\,\delta\,c)\,\Pi\,\mathbf{1}+\rho\,\Pi\,\boldsymbol{y}_{iLAG}+\gamma\,\Pi\,\boldsymbol{y}_{iLAG}\boldsymbol{d}_{98}+\\&+\beta\,\Pi\,\boldsymbol{d}_{t\geqslant 1998}+\phi\,\Pi\,\boldsymbol{t}+\psi\,\Pi\,\boldsymbol{td}_{t\geqslant 1998}+\Pi\,X_i{}'\boldsymbol{\omega}+\Pi\,X_i{}'\boldsymbol{d}_{98}\,\boldsymbol{\iota}\end{aligned} \qquad (21)$$

dove c è la costante di proporzionalità fra la costante e la variabili dummy quasi-differenziate pari a $\frac{k+1}{k}$.

La presenza della collinearità appena evidenziata, così come le uguaglianze individuate in fase di costruzione del modello fra $\mathbf{d}_{t\geq 1998}\mathbf{d}_{98}$ e $\mathbf{d}_{98}$ e fra $\mathbf{td}_{98}$ e 1998$\mathbf{d}_{98}$, non ci permettono l'identificazione della totalità dei parametri presenti all'interno dei modelli definiti da (12) e da (13), per quanto riguarda le ultime due uguaglianze, e nemmeno il modello specificato in (16), a causa di quest'ultima collinearità fra le variabili trasformate. Il fatto che alcune delle variabili del modello risulti combinazione lineare delle rimanenti, tuttavia, non introduce complicazioni per la stima del parametro di interesse e, dal momento che, come affermato in precedenza, l'obbiettivo primario del nostro studio è l'individuazione di una stima del parametro ρ, una procedura di stima che agisca in questo modo può essere considerata soddisfacente.

Abbiamo fino ad ora tralasciato il fatto che alcuni coefficienti in (16) fossero vincolati e ciò vale indubbiamente anche per (21). Un rapido sguardo ai vincoli individuati in (15) ci suggerisce che i coefficienti vincolati ad assumere un particolare valore sono : μ, γ, λ, κ, δ, θ e ι. Il parametro ψ, inoltre, è legato ai valori di β e di ρ attraverso la relazione vista precedentemente, ovvero

$$\psi=\frac{-\beta(\sqrt{\rho}+1)}{\sqrt{\rho}}$$

Di θ ci siamo già occupati in precedenza, nel senso che tale parametro nell'espressione (22) non compare, mentre per quanto riguarda gli altri coefficienti è necessario tenere in considerazione il vincolo definito in precedenza.

Dal modello in quasi-differenze esplicitato in (22) notiamo come rimangano vincolati i parametri γ e ψ, mentre il parametro associato alla variabile dummy per l'anno 1998 trasformata, in quanto somma di quattro parametri vincolati e di un parametro libero, risulta essere libero di variare a piacimento indipendentemente dal valore assunto dagli altri parametri del modello.

Possiamo dunque arrivare ad affermare che uno stimatore appropriato per il modello di partenza può essere ottenuto attraverso una stima a variabili strumentali del seguente modello :

$$\begin{aligned}\boldsymbol{y}_i = {} & \alpha^* \Pi \mathbf{1} + \rho \Pi\, \boldsymbol{y}_{iLAG} + \gamma \Pi\, \boldsymbol{y}_{iLAG} \boldsymbol{d}_{98} + \beta \Pi\, \boldsymbol{d}_{t \geqslant 1998} + \phi \Pi\, \boldsymbol{t} + \psi \Pi\, \boldsymbol{td}_{t \geqslant 1998} + \\ & + \Pi X_i' \boldsymbol{\omega} + \Pi X_i' \boldsymbol{d}_{98} \boldsymbol{\iota}\end{aligned} \tag{22}$$

in cui α^* è un parametro libero diverso dal precedente α. Come nel caso dello stimatore di Arellano-Bond le variabili strumentali possono essere scelte fra i valori della variabile risposta ai ritardi superiori al secondo, mentre i parametri γ , ψ e ι sono vincolati ai restanti attraverso le relazioni

$$\gamma = \rho(\sqrt{\rho} - 1) \qquad \psi = \frac{-\beta(\sqrt{\rho} + 1)}{\sqrt{\rho}} \qquad \iota = \frac{\omega \rho}{\sqrt{\rho} - 1}$$

I vincoli in questione, tuttavia, sono di tipo non lineare: può risultare difficile implementare uno stimatore a variabili strumentali che ammetta la presenza di vincoli non lineari sui parametri. Un approccio alternativo è quello di utilizzare lo stimatore di Arellano-Bond per reperire una stima non vincolata consistente ed efficiente dei parametri da utilizzare per ottenere successivamente stime vincolate dei parametri attraverso il *minimum distance estimator*. Tale approccio permette di arrivare ad una stima dei parametri vincolati a partire da una stima non vincolata dei parametri stessi che abbia le usuali proprietà di consistenza e normalità asintotica. Questo metodo afferma che, sia θ un vettore di parametri di interesse di dimensione S e sia ζ il vettore di parametri in forma ridotta di dimensione P, con P>S, disponendo di una stima $\hat{\zeta}$ dei parametri tale che

$$\sqrt{n}(\hat{\zeta}-\zeta)\sim N(0,\Sigma_0)$$

sia inoltre *h(θ)* la funzione definita in $\mathbb{R}^S \rightarrow \mathbb{R}^P$ che mappa i parametri liberi nei parametri in forma ridotta allora una stima dei parametri vincolati può essere ottenuta attraverso minimizzando la distanza fra i vettori $\hat{\zeta}\, e\, h(\hat{\theta})$ utilizzando la distanza euclidea pesata come misura di prossimità fra i due vettori. La matrice di pesi efficiente in questo caso si dimostra essere $\hat{\Sigma}^{-1}$ dove $\hat{\Sigma}$ è una stima della matrice di varianze e covarianze dei parametri non vincolati tale che $\underset{t\rightarrow\infty}{plim}\,\hat{\Sigma}=\Sigma_0$, si tratta cioè dell'inversa della matrice di varianze e covarianze di un qualsiasi stimatore consistente per ζ. Il problema di minimo così specificato risulta quindi essere

$$\min_{\theta\in\Theta}(\hat{\zeta}-h(\theta))'\hat{\Sigma}^{-1}(\hat{\zeta}-h(\theta)) \qquad (23)$$

Nel nostro caso la stima $\hat{\zeta}$ può essere ottenuta applicando il già citato stimatore a variabili strumentali al modello in quasi-differenze, mentre la funzione *h(θ)* diventa :

$$h(\alpha,\rho,\beta,\phi,\omega)=\begin{bmatrix}\hat{\alpha}^*=\alpha^*\\ \hat{\rho}=\rho\\ \hat{\gamma}=\rho(\sqrt{\rho}-1)\\ \hat{\beta}=\beta\\ \hat{\phi}=\phi\\ \hat{\psi}=\dfrac{-\beta(\sqrt{\rho}+1)}{\sqrt{\rho}}\\ \hat{\omega}=\omega\\ \hat{\iota}=\dfrac{\omega\rho}{\sqrt{\rho}-1}\end{bmatrix}$$

dove θ qui è il vettore di parametri liberi rappresentato da α^*, ρ, β ,φ e ω. Le stime cercate sono a questo punto la soluzione al problema di minimo (23).

Le stime ottenute in questo modo si dimostrano essere efficienti se tale è la matrice di varianze e covarianze delle stime di partenza. Un ulteriore risultato ci fornisce la distribuzione asintotica del vettore θ di parametri liberi stimato, si dimostra infatti che

$$\sqrt{N}(\hat{\theta}-\theta_0)\sim N(0,(H'\hat{\Sigma}^{-1}H)^{-1})$$

con H Jacobiamo della funzione *h(θ)*.

L'utilizzo di una procedura di stima di questo tipo ci permette quindi l'individuazione di una stima consistente del parametro di interesse e dei parametri β e φ , fornendone la distribuzione approssimata. E' possibile, inoltre, giungere al valore dei parametri vincolati γ, ψ e ι, dal momento che una loro stima consistente è rappresentata dall'espressione definita in *h* calcolata in $\hat{\theta}$. Gli errori standard relativi possono venire individuati attraverso la procedura nota come *metodo delta*, la quale permette di individuare la distribuzione di una trasformazione di stimatori. Nulla può infine essere affermato, come già visto in precedenza, sui rimanenti parametri che vanno a comporre l'espressione di α^* dal momento che la procedura, analogamente a quanto accade per variabili costanti nel tempo nel modello lineare classico, non riesce a distinguere tali parametri fra loro ma riesce solo a stimare il valore assunto da α^* che rappresenta la trasformazione dei parametri di partenza definita in (21).

Le Stime

Nel precedente capitolo abbiamo introdotto metodologie per la stima di quantità di interesse applicabili allo studio che ci interessa. In questo capitolo ci occuperemo dell'applicazione di tali metodologie al caso in questione definendo di volta in volta le quantità che entreranno in gioco e le variabili che utilizzeremo, interpretando infine ciò che i risultati ci suggeriranno.

Nell'analisi tratteremo dapprima modelli di tipo statico, per i quali utilizzeremo i periodi dal 1989 al 2004 restringendo però l'analisi alle sole famiglie intervistate per almeno quattro periodi, per poi passare a quelli di tipo dinamico, utilizzando il campione panel di dati a nostra disposizione per gli anni dal 1998 al 2004. Un ultimo paragrafo sarà infine dedicato alla stima del modello dinamico per l'intero periodo 1989-2004.

Nello stimare i modelli presentati nel capitolo precedente valuteremo anche le stime relative ad alcuni modelli più semplici e che ignoreranno volutamente alcune assunzioni fatte in sede di individuazione dei modelli con lo scopo di avvalorare la veridicità di tali assunzioni. Proveremo, ad esempio, a stimare modelli che non includano un effetto specifico di individuo così da poter valutare quanto queste ultime stime si discostino dalle stime che ne ammettano la presenza avvalorando così l'ipotesi di partenza.

3.1 Modelli Statici

Nei modelli statici, come abbiamo avuto modo di vedere in dettaglio nel capitolo precedente, le variabili che entrano in gioco sono la variabile risposta, di tipo dicotomico, e un insieme di variabili condizionanti che, almeno per i modelli visti, possono essere assunte strettamente esogene per il modello.

Qui la variabile risposta è ovviamente rappresentata dalla presenza/assenza di almeno uno fra un determinato insieme di titoli rischiosi. Il nostro insieme di titoli rischiosi, che ci permetterà di determinare i valori assunti dalla variabile risposta, risulta essere composto, per le ragioni viste nel primo capitolo, da azioni, fondi comuni di investimento e pensioni integrative.

L'insieme di variabili condizionanti che, almeno in prima battuta considereremo, si compone invece di variabili circa il reddito, la ricchezza, l'età del capofamiglia, la zona di residenza (Nord, Centro, Sud o Isole), una proxy del tenore di vita della famiglia, il sesso, il titolo di studio (laurea, diploma o altro) ed una variabile che indica se il capofamiglia è sposato o meno. La variabile per il tenore di vita è costruita come rapporto fra la superficie abitativa e il numero dei figli. Per rispondere alla necessità di avere a disposizione variabili esogene al modello, inoltre, la variabile reddito e la variabile ricchezza sono valutate rispettivamente come il reddito disponibile netto e la ricchezza reale netta escludendo dal computo le rendite da capitale finanziario e le attività finanziarie. Tali variabili sono inoltre state deflazionate ed espresse in migliaia di euro. Le variabili di tipo anagrafico, infine, si riferiscono al capofamiglia dove capofamiglia è stato considerato essere la persona di sesso maschile all'interno del nucleo famigliare nel caso di coppie sposate e il capofamiglia dichiarato altrimenti. Il campione a nostra disposizione per gli anni che vanno dal 1989 al 2004 si presenta composto da 17648 osservazioni relative a 3417 famiglie osservate per quattro o più periodi di tempo nel corso del periodo di interesse. Le famiglie per le quali venga osservata almeno una transizione nel periodo di osservazione sono infine 609 e dunque solo l'informazione relativa a queste ultime famiglie potrà essere utilizzata per la stima del modello non lineare che andremo a stimare.

Per come risulta definito un modello di tipo statico, dunque, si tratta di stabilire la relazione esistente, lineare o meno, fra questo gruppo di variabili condizionanti X e la variabile indipendente y. Una prima stima di tale effetto potrebbe essere ottenuta regredendo semplicemente y su X nel contesto di un semplice modello di tipo lineare, oppure tenendo in considerazione la natura dicotomica della variabile risposta ed implementando un modello di tipo non-lineare che, a seconda della scelta sulla funzione di legame, che può essere di tipo logit o probit. La tabella 3.1 elenca le stime per i parametri nei modelli lineare e logit.

Per permettere un migliore raffronto fra le stime dei due modelli la terza colonna riporta l'effetto sulla probabilità di successo indotta dall'aumento unitario della relativa variabile condizionante ottenuta dall'espressione riportata in (6) e che, derivando la relativa funzione legame, diventa :

$$\frac{\exp(x_{it}'\beta)}{[1+\exp(x_{it}'\beta)]^2}\beta$$

Tabella 3.1 : Stime modello statico senza eterogeneità individuale non osservata

	Linear Model	***Logit***	***Logit - Effetti Marginali***
Reddito	0.00317 *(0)*	0.02059 *(0)*	0.00306 *(0)*
Ricchezza	0.00006 *(0.003)*	0.00041 *(0)*	0.00006 *(0)*
m^2 / n° figli	0.00032 *(0.039)*	0.0023 *(0.013)*	0.00034 *(0.013)*
Età	-0.00229 *(0)*	-0.01694 *(0)*	-0.00252 *(0)*
Anno 1991	-0.01969 *(0.299)*	-0.20753 *(0.237)*	-0.02929 *(0.211)*
Anno 1993	0.03567 *(0.060)*	0.26027 *(0.116)*	0.04116 *(0.138)*
Anno 1995	0.03639 *(0.051)*	0.26072 *(0.113)*	0.04116 *(0.134)*
Anno 1998	0.12014 *(0)*	0.85174 *(0)*	0.14871 *(0)*
Anno 2000	0.18081 *(0.024)*	1.20807 *(0)*	0.22533 *(0)*
Anno 2002	0.13841 *(0)*	0.95957 *(0)*	0.17458 *(0)*
Anno 2004	0.12866 *(0)*	0.90552 *(0)*	0.16441 *(0)*
Centro	-0.08371 *(0)*	-0.45207 *(0)*	-0.06179 *(0)*
Sud	-0.21297 *(0)*	-1.60719 *(0)*	-0.21074 *(0)*
Superiori	0.10558 *(0)*	0.63869 *(0)*	0.10351 *(0)*
Laurea	0.11306 *(0)*	0.68607 *(0)*	0.1205 *(0)*
Donna	-0.05113 *(0.003)*	-0.48713 *(0)*	-0.07251 *(0)*
Sposato	0.01045 *(0.531)*	0.05592 *(0.620)*	0.00823 *(0.616)*
costante	0.24435 *(0)*	-1.13964 *(0)*	

Nota : p-value in parentesi. La terza colonna riporta gli effetti marginali per il modello logit calcolati nel punto medio delle variabili condizionanti.

I valori riportati nella terza colonna sono dunque calcolati secondo questa espressione, dove le variabili condizionanti sono valutate nel loro valore medio.

Ad un primo sguardo notiamo subito come le stime per i due modelli diano le stesse indicazioni per quanto riguarda la direzione degli effetti indotti dalle variabili, a riprova della buona approssimazione al modello logit fornita da modelli di tipo lineare, approssimazione particolarmente buona in prossimità dei valori medi delle variabili condizionanti.

Le variabili condizionanti risultano quasi tutte significative nel discriminare fra possessori e non ed in particolare notiamo come il modello indichi una probabilità inferiore di circa il 21% per i residenti al Sud ed una probabilità maggiore all'incirca dell'11% per coloro che posseggono un diploma o titoli di studio superiori a parità di altre condizioni. Fortemente significativo risulta anche l'effetto indotto dal reddito per il quale notiamo una probabilità che aumenta del 3% all'aumento di 10.000 euro del livello di reddito *ceteris paribus*.

Le stime ottenute in questo modo, tuttavia, non prendono in considerazione la presenza di una componente non osservabile specifica di individuo che, invece, in sede di definizione dei modelli avevamo ipotizzata particolarmente rilevante. Qualora tale componente risultasse correlata con l'insieme condizionante, infatti, la consistenza delle stime appena ottenute verrebbe meno, rendendo di fatto inutili le considerazioni sui risultati. Queste ci permetteranno tuttavia di valutare in modo più approfondito le stime che andremo ad individuare in seguito, offrendoci la possibilità di confrontare i risultati e concludere sull'adeguatezza delle valutazioni operate sulla componente specifica di individuo.

L'inclusione di questo nuovo elemento porta, dunque, alla specificazione dei modelli statici lineari e non-lineari in cui si tenga conto dell'effetto fisso presentati nel precedente capitolo. Come argomentato in sede di specificazione, tuttavia, questa tipologia di modelli non permette di includere, fra le variabili condizionanti, quantità costanti nel tempo. La presenza di misure di questo tipo, infatti, conduce all'impossibilità di stimare i coefficienti ad esse associati, in quanto l'operazione di eliminazione dell'effetto fisso porta all'esclusione dal modello anche di queste ultime. Nel nostro caso, dunque, non sarà possibile valutare i coefficienti relativi alle variabili relative alla zona di residenza, al titolo di studio, al sesso ed allo stato civile in quanto la variabilità nel tempo di queste quantità risulta piuttosto limitata. La varianza campionaria fra i gruppi calcolata per queste variabili risulta diversa da zero, tuttavia ciò non basta a garantire una adeguata consistenza dello stimatore ottenuto e non potranno dunque essere incluse nell'insieme di variabili condizionanti. L'assenza dal modello di queste variabili, non pregiudica in alcun modo la consistenza dei coefficienti associati alle restanti quantità, dal momento che l'effetto dovuto a sesso, stato civile, titolo di studio e zona di residenza viene semplicemente inglobato nell'effetto fisso ed escluso dal modello lasciando inalterata la consistenza dei restanti coefficienti.

Le stime ottenute relative a quest'ultima tipologia di modelli sono riportate in tabella 3.2 . In essa vengono, inoltre, riportate le stime ottenute considerando la componente non osservabile come effetto casuale.

Tabella 3.2 : Stime modello statico con eterogeneità individuale non osservata

	Linear Model - Random Effects	***Linear Model - Within Group***	***Conditional Logit***
Reddito	0.00261 *(0)*	0.00178 *(0)*	0.0157 *(0)*
Ricchezza	0.00006 *(0.002)*	0.00004 *(0.046)*	0.00035 *(0.086)*
m^2 / n° figli	0.00021 *(0.212)*	-0.00024 *(0.202)*	-0.0015 *(0.532)*
Età	-0.00225 *(0)*	0.0283 *(0.491)*	0.34048 *(0.444)*
Anno 1991	-0.02137 *(0.198)*	-0.81024 *(0.337)*	-0.97154 *(0.293)*
Anno 1993	0.03597 *(0.031)*	-0.08109 *(0.625)*	-0.98659 *(0.584)*
Anno 1995	0.03768 *(0.022)*	-0.13743 *(0.579)*	-1.69378 *(0.528)*
Anno 1998	0.12585 *(0)*	-0.1314 *(0.723)*	-1.85101 *(0.645)*
Anno 2000	0.18816 *(0)*	-0.12611 *(0.781)*	-2.01728 *(0.681)*
Anno 2002	0.14506 *(0)*	0.22967 *(0.668)*	-3.12129 *(0.590)*
Anno 2004	0.13476 *(0)*	0.29807 *(0.629)*	-3.90016 *(0.560)*
Centro	-0.08913 *(0)*		
Sud	-0.22264 *(0)*		
Superiori	0.09392 *(0)*		
Laurea	0.124 *(0)*		
Donna	-0.04976 *(0.024)*		
Sposato	0.00925 *(0.652)*		
costante	0.26447 *(0)*		

Nota : p-value in parentesi

Tale modello troverebbe giustificazione qualora tale componente potesse essere assunta incorrelata con l'insieme delle variabili esplicative considerate, la diversità fra queste stime e quelle ad effetti fissi potrà segnalare, una volta di più, l'inadeguatezza di tale assunzione che avevamo già supposto non verificata. A conferma di ciò troviamo, infatti, che la statistica test di Hausman per verificare l'uguaglianza fra le stime ad effetti fissi e quelle ad effetti casuali prende il valore di 245.79 da confrontare con un chi-quadro con 9 gradi di libertà. Il test in questione, dunque, rifiuta l'ipotesi nulla di uguaglianza delle stime e suggerisce, dal momento che la stima ad effetti fissi risulta consistente indipendentemente dalla relazione esistente fra la componente non osservabile e le variabili condizionanti mentre quella ad effetti fissi è consistente (ed in tal caso anche efficiente) solo se tale componente risultasse incorrelata con le restanti esplicative, la presenza di una componente η_i correlata con le l'insieme di variabili esplicative.

Le stime ad effetti casuali, come è possibile notare, non si discostano molto da quelle ottenute nel modello senza effetto fisso e ciò è dovuto al fatto che il metodo di stima utilizzato nei due casi è piuttosto simile. La stima ad effetti fissi, infatti, non è altro che una stima con il metodo dei minimi quadrati generalizzati del dataset di partenza poiché l'introduzione dell'effetto fisso genera errori correlati fra loro secondo, però, una struttura nota.

Passando a discutere il modello ad effetti fissi notiamo subito come le stime, prima molto significativamente diverse da zero, siano ora, eccezion fatta per il reddito, tutte non significative ad indicazione del fatto che, una volta tenuto conto di tutto ciò che abbiamo detto sta a rappresentare l'effetto fisso, le restanti variabili non riescano più a spiegare in modo significativo il possesso di titoli rischiosi. L'età, in particolare, non riesce a spiegare, condizionatamente all'effetto fisso, il possesso di titoli rischiosi , nonostante la forte relazione notata in sede di analisi descrittiva.

Ciò che precedentemente riusciva meglio a catturare la variabilità della variabile risposta era infatti rappresentato da zona di residenza e titolo di studio, due variabili ora "inglobate" all'interno dell'effetto fisso congiuntamente ad ulteriori variabili specifiche di individuo e di difficile quantificazione ragionevolmente importanti nel discriminare fra possessori e non di titoli rischiosi. L'unica variabile a nostra disposizione che sembra rilevante nel determinare le scelte dei singoli individui è il reddito per il quale comunque questo modello ci indica un aumento di oltre un punto e mezzo percentuale nella probabilità di possesso ogni 10.000 euro di reddito.

Per quanto riguarda il modello non lineare le considerazioni sono analoghe data la significatività dei parametri e i segni degli effetti indotti dalle variabili sulla variabile risposta. Anche in questo caso, dunque, il modello di tipo lineare si dimostra inm grado di approssimare adeguatamente gli effetti stimati attraverso modelli di tipo non lineare.

3.2 Modelli Dinamici

La natura non equispaziata del panel a nostra disposizione ci permette, per quanto riguarda l'analisi di modelli dinamici, di utilizzare solo una parte dei periodi temporali disponibili. Nel seguito utilizzeremo dunque solo i dati disponibili per gli anni di indagine che vanno dal 1998 al 2004. Il campione che tratteremo sarà dunque formato da 7420 osservazioni su 1855 famiglie intervistate in tutti e quattro gli anni in questione. Lo stimatore di Honoré e Kyriazidou che utilizzeremo per la stima del modello non lineare, infine, potrà qui sfruttare l'informazione proveniente dalle 360 famiglie per le quali è stata osservata transizione nello stato fra il 2000 ed il 2002.

Per questo genere di modelli la procedura che adotteremo per verificare le varie assunzioni poste in fase di specificazione del modello sarà analoga a quella proposta per il caso statico. Si tratterà quindi di valutare un modello che non includa una componente specifica di individuo così da analizzare le differenti conclusioni alle quali ci portano questi due modelli.

A tal proposito considereremo inizialmente le stime ottenute da un semplice modello lineare senza effetto fisso di individuo e senza variabili esplicative ad eccezione della variabile risposta ritardata. Il metodo di stima utilizzato in questo caso sarà una semplice stima a variabili strumentali poiché, non imponendo restrizioni sulla componente di errore[1], la variabile ritardata potrebbe essere fonte di endogenità per il modello. La variabile strumentale utilizzata, in modo analogo a quanto visto per il modello lineare dinamico proposto in precedenza, è la stessa variabile risposta ritardata di due periodi.

Le stime fornite da tale procedura indicano un coefficiente associato alla variabile risposta ritardata di 0.947 ed una costante pari a -0.012. Gli standard error stimati

---

1 Assumendo un errore serialmente incorrelato l'esogenità della variabile ritardata sarebbe garantita, tuttavia, non disponendo di informazioni a questo proposito non possiamo escludere un errore autocorrelato. Stimando i modelli utilizzando sia OLS che IV, infatti, notiamo una significativa differenza nelle stime a riprova di un errore autocorrelato.

sono rispettivamente 0.040 e 0.012 indicando una forte significatività della variabile ritardata e un valore non significativamente diverso da zero per la costante. Il modello suggerisce dunque una probabilità di possesso pari al 91% per coloro che già possedevano titoli rischiosi ed un pressoché certo non possesso[2] per coloro che non possedevano titoli rischiosi nel periodo precedente.

Un modello di questi tipo, tuttavia, "spiega" il possesso in un determinato periodo unicamente sulla base del possesso o meno nel periodo precedente. Il valore elevato associato alla variabile ritardata è tuttavia da imputare al fatto che in esso è presente sia ciò che abbiamo definito correlazione pura, cui siamo interessati, sia correlazione spuria fra le due variabili, dettata dal fatto che entrambe, la variabile risposta e se stessa ritardata, sono influenzate da un insieme di variabili costante nel tempo e che genera dunque una correlazione non genuina fra le due variabili. L'obbiettivo della nostra analisi è tuttavia quello di riuscire a distinguere fra questi due effetti.

Un modo per disaggregare queste due componenti consiste nell'includere un effetto fisso di individuo nell'espressione del modello. In tal modo riusciamo a controllare per un insieme di variabili costanti nel tempo che influenzano sia la variabile risposta che la variabile ritardata. Una prima stima del modello lineare, utilizzando lo stimatore di Arellano-Bond, con effetto fisso di individuo e senza ulteriori variabili esplicative indica un coefficiente per la variabile risposta ritardata di 0.06 con un errore standard di 0.03 ed una costante di -0.027 con errore standard 0.006. L'effetto indotto dal possesso nel periodo precedente sulla probabilità di possesso stimata è

---

2 Il modello indica una probabilità di possesso per i non possessori negativa, ma ciò è dovuto alla mancata restrizione, nel caso lineare, delle probabilità stimate dal modello all'intervallo (0,1).

dunque nell'ordine del 6%. Dal momento che questo coefficiente rappresenta proprio la correlazione pura cui siamo interessati, notiamo come la grande correlazione evidenziata in precedenza sia per lo più spuria poiché poiché subisce un forte ridimensionamento nel caso in cui venga presa in considerazione la presenza di effetti fissi. Gran parte della correlazione fra il possesso di titoli rischiosi nei diversi periodi è dunque attribuibile alla presenza di effetti fissi o, in altre parole, al fatto che le decisioni vengono prese dallo stesso soggetto in diversi periodi di tempo e sulla base delle stesse condizioni di mercato e poiché ogni soggetto possiede caratteristiche che non variano in maniera significativa nel tempo egli tenderà a prendere simili decisioni. Rimane tuttavia presente un effetto dovuto al solo fatto di essere o meno già in possesso di titoli rischiosi quali i costi di avvio/dismissione delle pratiche relative all'acquisto di tali titoli o a tutto ciò che non possa essere altrimenti attribuito al fatto che la scelta è stata fatta dal medesimo individuo.Il p-value associato è tuttavia pari a 0.101 e ciò segnala che i dati a nostra disposizione non ci permettono di rifiutare l'ipotesi di nullità di tale componente agli usuali livelli di confidenza.

L'ultimo passo nella nostra analisi consiste nell'includere un insieme di variabili condizionanti aggiuntive per il modello e ciò che otteniamo è un modello che tiene in considerazione sia l'aspetto dinamico che statico del fenomeno in questione e ciò ci consentirà di valutare eventuali cambiamenti nei valori stimati dei parametri rispetto ai corrispondenti valori già individuati per i due modelli separatamente. Le stime per il modello lineare e logit ottenute sono riportate in tabella 3.4 e si riferiscono alle stime ottenute dalle procedura proposte da Arellano-Bond e da Honoré-Kyriazidou rispettivamente.

Tabella 3.3 : Stime modello dinamico con eterogeneità individuale non osservata

	Dynamic Linear Model	***Honorè Kyriazidou***
Costante	-0.29205 *(0.302)*	
Possesso (t-1)	0.06576 *(0.061)*	0.13385 *(0.827)*
Reddito	0.00136 *(0.023)*	0.0192 *(0.609)*
Ricchezza	-0.00001 *(0.771)*	
m^2 / n° figli	-0.00043 *(0.333)*	
Età	0.11907 *(0.400)*	
Anno 2004	0.04584 *(0.055)*	

Nota : p-value in parentesi

Analogamente a quanto osservato per il modello statico l'unica variabile esplicativa significativamente diversa da zero nel modello lineare è il reddito, mentre per le restanti i dati suggeriscono la nullità dei parametri ad esse associate. Sulla base di questa osservazione, e dal momento che l'inclusione di un elevato numero di variabili esplicative nel modello non lineare risulta difficoltoso a causa della funzione kernel da associare alle osservazioni, il modello logit è stato stimato includendo fra le esplicative il solo reddito. I modello così definiti ci permette di identificare un effetto associato alla variabile ritardata sulla variabile risposta nell'ordine del 6% per il modello lineare, in linea con quanto visto in precedenza per il modello con la sola variabile ritardata fra le esplicative, mentre per il modello non lineare l'effetto, calcolato nel valore medio assunto dalla variabile reddito, è di circa il 3%. Queste ultime stime riportano inoltre degli errori standard per i coefficienti molto elevati che portano ad una forte non significatività dei parametri. Un tale fenomeno può essere tuttavia dovuto alla scarsità di osservazioni a disposizione (solo 360 famiglie) ed al rateo di convergenza inferiore all'usuale $\frac{1}{\sqrt{n}}$ per questo stimatore che, data l'impossibilità di utilizzare l'intero campione di dati a nostra disposizione, può inficiare la precisione delle stime. La stima puntuale indicata dal modello segnala tuttavia un effetto non molto diverso da quello identificato nel modello lineare.

3.3 Modello dinamico per dati a rilevazione irregolare

Presenteremo ora le stime ottenute dall'applicazione del modello dinamico per i dati a rilevazione irregolare specificato nel capitolo precedente. L'utilizzo di un modello di questo tipo è giustificato dall'interesse per la valutazione della correlazione pura esistente qualora l'orizzonte temporale preso in considerazione sia più ampio. Chiaramente l'implementazione di questo modello si basa sull'assunzione che il coefficiente associato alla variabile dipendente ritardata sia costante nel tempo, tuttavia tale ipotesi, che abbiamo considerato verificata nel caso in cui l'orizzonte temporale era più ristretto, è sempre più difficilmente assumibile all'aumentare del numero di periodi considerati.

La stima si basa, inoltre, sull'individuazione di una stima iniziale di ρ che qui reperiremo dalla stima di un modello dinamico lineare sui due gruppi di osservazioni composte dai periodi dal 1898 al 1995 e dal 1998 al 2004. Escluderemo dunque, almeno inizialmente, dalla stima il periodo di transizione fra l'anno 1995 e l'anno 1998. Nel far ciò notiamo come, mentre le osservazioni per il secondo gruppo sono quelle relative alla stima dinamica ottenuta al paragrafo precedente, al primo gruppo appartengono le famiglie intervistate in tutti e quattro gli anni che vanno dal 1989 al 1995 e che risultano essere 827, un numero dunque molto minore rispetto a quelle appartenenti al secondo gruppo. La stima ancora con Arellano-Bond di questo modello porta ad un coefficiente associato alla variabile ritardata di 0.0769 che utilizzeremo come stima preliminare da utilizzare per quasi-differenziare le variabili condizionanti. La stima precedentemente descritta indica inoltre un trend diverso da zero e diverso per i due gruppi di osservazioni per il possesso di titoli rischiosi, elemento che utilizzeremo in fase di stima finale del modello in linea con quanto specificato nel precedentemente capitolo. Una variabile di interazione fra il coefficiente ρ e l'appartenenza al primo gruppo permette di valutare la veridicità dell'ipotesi fatta in precedenza sull'uguaglianza dei coefficienti ρ associati ai due gruppi di osservazioni dal momento che essa è non significativa con un p-value pari a 0.328.

I valori da utilizzare per l'operazione di quasi-differenziazione sono dunque funzione di ρ secondo la relazione :

$$k_{2000}=\frac{\sqrt{\rho}+1}{\rho+\sqrt{\rho}+1} \qquad k_{1998}=\frac{\rho+\sqrt{\rho}+1}{\sqrt{\rho}+1}$$

e dunque assumono qui i valori di 0.94 e di 1.06 .

L'insieme condizionante aggiuntivo sarà composto inoltre, in linea con l'ultimo modello di tipo dinamico stimato al paragrafo precedente, dalla sola variabile di reddito. La stima non vincolata del modello in quasi differenze porta però ad una stima del coefficiente di interesse ρ molto vicina allo zero, al punto che risulta essere negativa. La stima ottenuta è infatti pari a -0.00533 con un errore standard di 0.0395. La statistica test t è pari a -0.13 e dunque a qualunque livello di confidenza non rifiuteremmo l'ipotesi di nullità del coefficiente.

L'identificazione di un valore negativo preclude inoltre la possibilità di reperire, almeno con il metodo da noi indicato, una stima vincolata del modello dal momento che il coefficiente stimato non cade all'interno dell'intervallo di esistenza del coefficiente, condizione posta in sede di specificazione del modello. Tale valore, in particolare, non ci permette di poterne calcolare la radice quadrata presente nell'espressione per $h(\theta)$ e ciò rende impossibile l'applicazione del minimum distance estimator.

Un tale risultato può essere per lo più ricondotto a due fattori principali:

- l'arbitrarietà della scelta del valore iniziale da attribuire a ρ per dare inizio alla stima.
- il valore vicino allo zero assunto dal coefficiente ρ.

Per quanto riguarda il primo punto la nostra scelta è stata quella di inizializzare la stima attribuendo a ρ il valore stimato utilizzando due gruppi di osservazioni separatamente, tuttavia, il numero di osservazioni stiamo tralasciando è piuttosto elevato. Così facendo utilizziamo infatti l'informazione proveniente da 2452 famiglie dove, inoltre, nel valutare coloro che vengono osservati per tutti gli otto periodi in due gruppi separatamente perdiamo la specificità propria della famiglia nel senso che andiamo a tenere conto di due differenti effetti fissi nei due periodi. L'intero campione che utilizziamo invece per la successiva stima è formato da 3417 famiglie e per le famiglie osservate dal 1989 al 2004 andiamo a stimare un solo effetto fisso. La differenza fra queste due stime può risultare dunque anche evidente, andando ad influenzare i risultati finali della stima che ci interessa.

Un ulteriore elemento da tenere in considerazione è il valore del coefficiente ρ che intendiamo stimare. Come affermato a proposito dello stimatore di Arellano-Bond, utilizzare una stima a variabili strumentali in cui gli strumenti siano valori ritardati per più di due periodi della variabile risposta in contesti dinamici può portare ad utilizzare strumenti deboli qualora il vero valore del coefficiente da stimare sia vicino ai valori di soglia. Qui è ragionevole assumere che il coefficiente sia molto prossimo allo zero e dunque la relazione i valori ritardati della variabile risposta può risultare molto debole, con la conseguenza che le stime ottenute tendono a diventare molto imprecise.

Conclusioni

Il lavoro svolto si proponeva come obbiettivo l'identificazione, attraverso l'analisi empirica, delle determinanti nel possesso di titoli di tipo rischioso, tenendo in considerazione la presenza di eterogeneità individuale non osservabile. Le evidenze empiriche stanno ad indicare come, una volta controllato per quest'ultimo elemento, le variabili esplicative, che prima contribuivano in maniera sostanziale a spiegare il possesso, risultano poco influenti. L'unica variabile ancora utile nel discriminare fra possessori e non rimane il reddito, per il quale i modelli stimati indicano un effetto, come ci si poteva attendere positivo, fra il punto e mezzo ed i tre punti percentuali sulla probabilità di possesso conseguenti all'aumento di 10.000 euro del livello di reddito. La significatività riscontrata in tutti i modelli proposti della variabile reddito indica, inoltre, che i portafogli detenuti dalle persone più ricche non possono essere considerati come la versione riscalata di quelli detenuti dai meno ricchi in quanto i primi tendono a possedere una maggiore varietà di strumenti finanziari e non gli stessi con livelli di investimento maggiori. Tale elemento risulta importante in sede di studio della composizione di portafoglio per queste due categorie di famiglie dal momento che il numero di famiglie ricche osservate è sempre piuttosto basso, per ragioni dovute principalmente alla bassa numerosità e dal tasso di non risposta piuttosto alto riscontrato in queste famiglie, e dunque riuscire a caratterizzare in maniera adeguata la composizione del loro portafoglio può risultare difficile.

I modelli dinamici stimati hanno riscontrato infine come la correlazione fra il possesso in periodi differenti, molto elevata inizialmente, sia quasi interamente di tipo spurio. La componente pura può venire identificata nell'ordine del 3% o del 6% a seconda del modello considerato, con coefficienti che tuttavia risultano non particolarmente significativi. Tale effetto risulta ancor più attenuato se facciamo riferimento all'intero orizzonte temporale a disposizione, in quanto il modello specificato porta all'identificazione di un coefficiente stimato praticamente nullo. Le evidenza sono dunque per un effetto di dipendenza dallo stato quantomeno ridotto rispetto a quanto riscontrato in altri paesi quali l'Olanda, per la quale studi precedenti indicavano una dipendenza dallo stato nel possesso di azioni nell'ordine del 17%. (Alessie, Hochguertel, van Soest 2004).

Appendice

Si riporta il listato del programma utilizzato per l'implementazione dello stimatore di Honoré e Kyriazidou in STATA :

```
program define honorekyria, eclass
	version 9.1
	if replay() {
		if "`e(cmd)'" != "honorekyria" {
			error 301
		}
		Replay `0'
	}
	else  Estimate `0'
end

program define Estimate, eclass
	syntax varlist(max=2) [if] [in],[I(varname) T(varname) Level(cilevel) noCONStant
SEMIParametric BANDwidth(real 1)]
	marksample touse
	tokenize `varlist'
	local yvar `1'
	local xvar `2'
	local k: word count `xvar'

	xt_iis `i'
	local ivar "`s(ivar)'"

	cap xt_tis `t'
	local tvar "`s(timevar)'"
	sort `ivar' `tvar'
	tempvar time
```

```
	by `ivar' : gen `time' = _n

	if `k'==1 {
		tempvar x23 x32 x34
		by `ivar' : gen `x23' = `xvar'[_N-2] - `xvar'[_N-1]
		by `ivar' : gen `x32' = `xvar'[_N-1] - `xvar'[_N-2]
		by `ivar' : gen `x34' = `xvar'[_N-1] - `xvar'[_N]
	}

	tempvar y14 y41 y23 y32 y2
	by `ivar' : gen `y14' = `yvar'[_N-3] - `yvar'[_N]
	by `ivar' : gen `y41' = `yvar'[_N] - `yvar'[_N-3]
	by `ivar' : gen `y23' = `yvar'[_N-2] + `yvar'[_N-1]
	by `ivar' : gen `y32' = `yvar'[_N-1] - `yvar'[_N-2]
	by `ivar' : gen `y2' = `yvar'[_N-2]

*******GENERAZIONE NUCLEO*********

	if `k'==1 {
		tempvar nucleos p
		gen `p' = `x34'/`bandwidth'
		gen `nucleos' = 0
		replace `nucleos' = 0.75 * (1/`bandwidth') * (1 - `p'^2) if abs(`p')<=1
	}
	else {
		tempvar nucleos
		gen `nucleos'=1
	}

*****STIMA*****

	if `k'==0 {
		ml model d0 hkyri_ll (`nucleos' `y2' = `y14', `constant') if `touse' & `time'==4 &
`y23'==1, max
		ereturn local method "no-covariates"
```

```
	}
	else {
		if "`semiparametric'" != "" {
			ml model d0 hkyri_sp (`nucleos' `y32' = `x32' `y41', `constant') if `touse' &
`time'==4 & `y23'==1, max difficult
			ereturn local method "semi-parametric"
		}
		else {
			ml model d0 hkyri_ll (`nucleos' `y2' = `x23' `y14', `constant') if `touse' &
`time'==4 & `y23'==1, max
			ereturn local method "parametric"
		}
	}

	tempname b v
	mat `b' = e(b)
	mat `v' = e(V)
	tempvar prob h1_1 lin h2_1 uno
	predict `lin'
	gen double `prob' = exp(`lin') / ( 1 + exp(`lin') )

******CALCOLO VARIANZE*************

	gen double `h1_1' = `y2' - `prob'
	gen double `h2_1' = - ( exp(`lin') / ((1+exp(`lin'))^2) )
	gen `uno' = 1
		if "`constant'" != "" {
			local z "`x23' `y14'"
			local variab "`xvar' `yvar'-1"
		}
		else {
			local z "`x23' `y14' `uno'"
			local variab "`xvar' `yvar'-1 _cons"
		}
	tokenize `z'
```

```
	local l: word count `z'
	tempname h1 h2
	mat `h2'=J(`l',`l',.)
	mat `h1'=J(`l',`l',.)

	forvalues x = 1(1)`l' {
		forvalues j = 1(1)`x' {
			tempvar h2`x'`j' somma`x'`j'
			qui gen double `h2`x'`j'' = `nucleos' * `h2_1' * ``x'' * ``j'' if (`y23'==1 & 
`time'==1 & `touse')
			qui gen double `somma`x'`j'' = sum(`h2`x'`j'')
			mat `h2'[`x',`j'] = `somma`x'`j''[_N]
			mat `h2'[`j',`x'] = `somma`x'`j''[_N]
		}
	}

	forvalues x = 1(1)`l' {
		forvalues j = 1(1)`x' {
			tempvar h1`x'`j' somma2`x'`j'
			qui gen double `h1`x'`j'' = (`nucleos'^2) * (`h1_1'^2) * ``x'' * ``j'' if (`y23'==1 
& `time'==1 & `touse')
			qui gen double `somma2`x'`j'' = sum(`h1`x'`j'')
			mat `h1'[`x',`j'] = `somma2`x'`j''[_N]
			mat `h1'[`j',`x'] = `somma2`x'`j''[_N]
		}
	}

	tempname AVAR
	mat define `AVAR' = (inv(`h2') * `h1' * inv(`h2'))

******RISULTATI*********

	local met = e(method)
	mat `b'=e(b)
	mat `v'=e(V)
```

```
        mat `v'[1,1]=`AVAR'
        mat coln `b'=`variab'
        mat rown `b'=`yvar'
        mat coln `v'=`variab'
        mat rown `v'=`variab'
        qui count if `y23'==1 & `time'==1 & `nucleos'>0
        sca nobs=r(N)
        eret post `b' `v', dep(`yvar')
        eret local cmd "hkyria"
        eret local method = "`met'"
        eret sca N = scalar(nobs)
        Replay, `level'
end

program define Replay
        eret di, level(`level')
end

program define hkyri_ll
        version 9.1
        args todo b lnf

        tempvar theta
        mleval `theta' = `b'

        tempvar lnfj
        qui gen double `lnfj' = $ML_y1 * ln( (exp(`theta')^($ML_y2))/(1+exp(`theta')) )
        mlsum `lnf' = `lnfj'
end

program define hkyri_sp
        version 9.1
        args todo b lnf
```

```
	tempvar theta
	mleval `theta' = `b'

	tempvar lnfj
	qui gen double `lnfj' = $ML_y1 * $ML_y2 * sign(`theta')
	mlsum `lnf' = `lnfj'
end
```

Bibliografia

Abel (1991). “The Equity Premium Puzzle” *Business Review*, Federal Reserve Bank of Philadelphia, 3-14

Alessie, Hochguertel, van Soest (2004) “Ownership of Stocks and Mutual Funds: a Panel Data Analysis” *Review of Economics and Statistics* 86 783-796

Allen, Gale (1994). “Limited Market Participation and Volatility Asset Prices” *American Economic Review* 84, 933-955

Bover, Arellano (1997). “Estimating Dynamic Limited Dependent Variable Models from Panel Data” *Investigaciones Economicas* 21, 141-165

Chamberlain (1980). “Analisis of Covariance with Qualitative Data” *Review of Economic Studies* 47, 225-238

Deaton (1992). “Understanding Consumption” *Clarendon Press*

Guiso, Haliassos, Jappelli et al. (2002) “Household Portfolios”, *The MIT Press*

Haliassos, Bertaut (1995). “Why do So Few Hold Stocks?” *Economic Journal* 105, 1110-1129

Haliassos, Hassapis (2001). “Non-Expected Utility, Saving and Portfolio” *The Economic Journal* 111, 69-102

Haliassos, Michaelides (2003). “Portfolio Choice and Liquidity Constraints” *International Economic Review* 44, 143-177

Heaton Lucas (2000). “Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk” *The Journal of Finance* 3, 1163-1198

Heckman (1981). “The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process” *Structural Analysis of Discrete Data with Econometric Applications*, 179-185

Honoré, Kyriazidou (2000). “Panel Data Discrete Choice Models with Lagged Dipendent Variables” *Econometrica* 68, 839-874

King, Leape (1998). “Wealth and Portfolio Composition: Theory and Evidence” *Journal of Public Economics* 69, 155-193

Kocherlakota (1996). “The Equity Premium: it's Still a Puzzle” *Journal of Economic Literature* 34, 42-71

Manski (1987). “Semiparametric Analysis of Random Effects Linear Models for Binary Panel Data” *Econometrica* 55, 357-362

Mehra, Prescott (1985). “The Equity Premium: a Puzzle” *Journal of Monetary Economics* 15, 145-161

Wooldridge (2002). “Econometric Analysis of Cross-Section and Panel Data” *The MIT Press*