1

L'ETEROGENEITÀ DELLE PREFERENZE NEL TRASPORTO MERCI: UN CONFRONTO TRA DIVERSI METODI PER CATTURARLA

Luisa Scaccia, Dipartimento di Statistica, Università degli Studi di Perugia.

ABSTRACT

La struttura delle preferenze relative al trasporto merci è generalmente caratterizzata dalla presenza di una forte eterogeneità dovuta alle specifiche caratteristiche della merce trasportata, dell'azienda che commissiona il trasporto, degli operatori che se ne occupano, dell'area in cui il trasporto si sviluppa. La classe dei modelli logit a parametri casuali si prefigge lo scopo di catturare tale eterogeneità e di integrarla nei processi di stima. In questo articolo verranno confrontati diversi modelli all'interno di questa classe. I modelli in esame differiscono per le assunzioni fatte circa la distribuzione dei parametri nella popolazione. In particolare, verranno considerati sia modelli basati su distribuzioni continue che modelli basati su distribuzioni discrete. I modelli saranno confrontati sulla base della loro capacità di rappresentare l'eterogeneità, della semplicità di interpretazione dei risultati e delle difficoltà computazionali. Il confronto verrà effettuato sulla base di dati derivanti dalle preferenze dichiarate di 51 imprese marchigiane, appartenenti ai settori metallurgico (DJ) e mobile (DN), circa le caratteristiche del loro trasporto merci tipico.

INTRODUZIONE

I processi decisionali relativi alla scelta del trasporto sono generalmente caratterizzati da una forte eterogeneità delle preferenze. Tale eterogeneità risulta particolarmente rilevante nell'ambito del trasporto merci. Le caratteristiche specifiche della merce trasportata, dell'azienda che commissiona il trasporto, dell'operatore che lo effettua, dell'area in cui il

trasporto ha luogo, rappresentano solo alcuni esempi dei fattori che possono determinare eterogeneità nelle preferenze espresse dai soggetti. Tenere conto di tale eterogeneità nei procedimenti di stima dei parametri di un modello per l'analisi delle preferenze permette ovviamente di pervenire ad una rappresentazione più realistica del fenomeno in esame.
Nel caso in cui i dati a disposizione sulla scelta del trasporto siano costituiti da un'unica scelta per ciascun individuo, come di norma avviene nelle indagini basate sulle cosiddette "preferenze rivelate", non è possibile discernere se le differenze osservate tra individui diversi siano ascrivibili a fattori casuali o ad eterogeneità nei gusti. Benchè non sia possibile ottenere dei parametri specifici per ciascun individuo, le differenze tra soggetti possono tuttavia essere catturate assumendo, ad esempio, che alcuni dei parametri del modello seguano una certa distribuzione. Nel caso in cui, invece, i dati siano raccolti sulla base delle "preferenze dichiarate", e per ogni individuo si disponga quindi di un certo numero di scelte effettuate tra profili caratterizzati da livelli diversi degli attributi del servizio di trasporto, è possibile determinare se nei dati è presente eterogeneità, caratterizzare tale eterogeneità e ottenere stime dei parametri specifiche di ciascun individuo.
I modelli a parametri casuali (McFadden e Train, 2000) sono stati ampiamente impiegati allo scopo di catturare e spiegare l'eterogeneità nelle scelte espressa da soggetti diversi. Tali modelli sono sostanzialmente suddivisibili in due categorie: i modelli mixed logit (MXL) e i modelli logit a classi latenti (LCM). Nei MXL l'eterogeneità nei parametri viene trattata attraverso l'assunzione che tali parametri non siano fissi ma seguano una determinata distribuzione nella popolazione. L'analista sceglie la forma distributiva e ottiene i parametri di tale distribuzione attraverso procedure di stima. Le distribuzioni solitamente prese in considerazione sono la normale, la lognormale, la triangolare e l'uniforme. Tuttavia, proprio la soggettività della scelta a priori della forma distributiva e le conseguenze che questa ha sui risultati dell'analisi rappresentano un punto debole di questi modelli (Heckman e Singer, 1984). Inoltre, i parametri di tali modelli non risultano stimabili in forma chiusa e il ricorso a metodi di simulazione è inevitabile, con conseguente aumento dei tempi computazionali. Nei LCM si assume, invece, che la popolazione sia suddivisa in classi caratterizzate, al loro interno, da preferenze omogenee. Questo equivale ad assumere che la distribuzione dei parametri sia concentrata su un numero discreto di punti, pari al numero di classi utilizzato nel modello. In questo modo si evita il ricorso a ipotesi a priori sulla forma della distribuzione dei parametri. L'unica scelta necessaria è relativa al numero di classi latenti da considerare ma questa può essere effettuata a posteriori, confrontando modelli con un numero diverso di classi. Inoltre, per questo modello le stime dei parametri sono disponibili in forma chiusa.
In questo lavoro ci si propone di confrontare i modelli MXL e LCM su base empirica, utilizzando dei dati derivanti dalle preferenze dichiarate di 51 imprese marchigiane, appartenenti ai settori metallurgico (DJ) e mobile (DN), circa le caratteristiche del loro trasporto merci tipico. I modelli saranno confrontati sulla base della capacità di rappresentare l'eterogeneità, della semplicità di interpretazione dei risultati e delle difficoltà computazionali incontrate nella stima dei loro parametri.
Nelle seguenti sezioni vengono brevemente illustrati i modelli messi a confronto, successivamente vengono descritti i dati utilizzati e i risultati ottenuti sulla base dei due diversi modelli e, infine, sono tratte le conclusioni e delineati alcuni spunti di ricerca futuri.

LA RAPPRESENTAZIONE DELL'ETEROGENEITÀ ATTRAVERSO I MODELLI A PARAMETRI CASUALI

Il modello logit multinomiale

Per comprendere il modo in cui i modelli a parametri casuali permettono di rappresentare l'eterogeneità che caratterizza le preferenze, consideriamo, come punto di partenza, il modello logit multinomiale (MNL) che assume invece parametri fissi ed omogeneità nei gusti (Ben-Akiva e Lerman, 1985, Train, 2003). Sotto le assunzioni del MNL, la probabilità che l'individuo n scelga l'alternativa i, all'interno dell'insieme di alternative per lui disponibili C_n, è data da

$$P_n(i) = \frac{\exp(X_i \beta)}{\sum_{j \in C_n} \exp(X_j \beta)} \tag{1}$$

dove X_i è un vettore riga contenente il livello di ciascun attributo osservato per l'alternativa i e β è un vettore colonna di parametri incogniti ma fissi, che assumono cioè lo stesso valore per ogni soggetto.

I modelli logit a parametri casuali possono essere visti come una naturale estensione del modello MNL in cui il vettore parametrico β non è più fisso ma può variare da soggetto a soggetto secondo una certa distribuzione. Indicata con $F(\beta)$ la funzione di ripartizione di β, nel modello logit a parametri casuali la probabilità che l'individuo n scelga l'alternativa i, può esprimersi come

$$P_n(i) = \int \frac{\exp(X_i \beta)}{\sum_{j \in C_n} \exp(X_j \beta)} dF(\beta) \tag{2}$$

Il modelli MXL e LCM sono entrambi ottenuti dalla (2) sostituendo a $F(\beta)$, rispettivamente, una distribuzione continua o una distribuzione discreta.

Il modello mixed logit

Se si assume che $F(\beta)$ sia la funzione di ripartizione di una distribuzione continua, la (2) può essere riscritta nella forma

$$P_n(i) = \int \frac{\exp(X_i \beta)}{\sum_{j \in C_n} \exp(X_j \beta)} f(\beta) d\beta \tag{3}$$

che esprime la probabilità che l'individuo n scelga l'alternativa i nel modello MXL (Green e Hensher, 2003b). In pratica, il modello assume che le preferenze dei soggetti, espresse dal vettore β, varino in maniera continua nella popolazione. La probabilità di ciascuna alternativa viene allora calcolata come media ponderata, su tutto lo spazio parametrico, delle probabilità sotto il modello MNL, utilizzando come pesi le densità di β.

L'integrale nella (3) non è tuttavia risolvibile in forma chiusa e richiede l'implementazione di metodi di simulazione. In pratica, si campionano M valori $\beta^{(1)},\ldots,\beta^{(M)}$ dalla distribuzione $f(\beta)$ e sulla base di tali valori si calcola la probabilità simulata

$$\widetilde{P}_n(i) = \frac{1}{M}\sum_{m=1}^{M}\frac{\exp\left(X_i\beta^{(m)}\right)}{\sum_{j\in C_n}\exp\left(X_j\beta^{(m)}\right)} \tag{4}$$

Le probabilità simulate vengono poi inserite nella funzione di verosimiglianza ottenendo la verosimiglianza simulata. I parametri della distribuzione $f(\beta)$ sono quindi stimati massimizzando la verosimiglianza simulata. Per migliorare la copertura del supporto di $f(\beta)$ per un dato M, invece di effettuare estrazioni casuali si possono effettuare estrazioni di Halton (si veda Train, 1999 e Bhat, 2001). Nonostante l'utilizzo delle sequenze di Halton, che permettono di ottenere la stessa copertura con un numero inferiore di simulazioni e di ridurre quindi i tempi computazionali, la stima di un modello MXL presenta una complessità computazionale decisamente più elevata di un modello MNL e tempi di implementazione ampiamente maggiori. Un ulteriore problema è rappresentato dal fatto che le stime, essendo basate sul risultato di simulazioni, sono chiaramente affette da errori di simulazione. Un modo per ridurne l'effetto è quello di aumentare M ma questo determina un ulteriore incremento dei tempi computazionali richiesti dal MXL.

Un altro punto debole del MXL, come accennato in precedenza, è rappresentato dal fatto che la scelta a priori della distribuzione $f(\beta)$ è affidata all'analista e i risultati dell'analisi possono dipendere pesantemente da tale scelta. Spesso viene impiegata una distribuzione unimodale dipendente da un parametro di centralità e un parametro di dispersione, tuttavia, nel caso in cui la distribuzione vera sia, ad esempio, bimodale, l'assunzione fatta determina un'incapacità del modello di rappresentare l'eterogeneità in maniera adeguata. Un modo per ovviare a questo problema consiste nell'assumere per i parametri una distribuzione non-parametrica che concentri probabilità non nulla in un limitato numero di punti dello spazio di definizione del parametro. Questa è la soluzione adottata dal LCM.

Il modello logit a classi latenti

Se si assume che $F(\beta)$ sia la funzione di ripartizione di una distribuzione discreta con probabilità non nulla su un numero limitato di punti, la (2) può essere riscritta nella forma

$$P_n(i) = \sum_{k=1}^{K}\frac{\exp\left(X_i\beta_k\right)}{\sum_{j\in C_n}\exp\left(X_j\beta_k\right)}\omega_k \tag{5}$$

che esprime la probabilità che l'individuo n scelga l'alternativa i nel modello LCM (Boxall e Adamowicz, 2002, Green e Hensher, 2003a). In pratica, il modello assume i soggetti siano classificabili in K classi in modo che i soggetti all'interno di una stessa classe abbiano la stessa struttura di preferenze e che questa sia diversa rispetto a quella dei soggetti nelle altre classi. La frequenza relativa incognita di individui appartenenti alla k-esima classe nella popolazione è ω_k. La probabilità di ciascuna alternativa viene allora calcolata come media

ponderata, su tutte le classi, delle probabilità sotto il modello MNL, utilizzando, come pesi, le frequenze relative delle classi. Quindi la probabilità di ciascuna alternativa è ora calcolabile in forma chiusa, senza dover far ricorso a simulazioni, e ciò permette di contenere i tempi computazionali e di ottenere stime non affette da errori di simulazione. Questo rappresenta un vantaggio notevole del LCM rispetto al MXL. Inoltre, il LCM non richiede assunzioni a priori sulla forma distributiva dei parametri. L'unica cosa che è necessario specificare è il numero di punti in cui tale distribuzione ha probabilità non nulla, ossia il numero di classi latenti. Chiaramente è possibile considerare modelli con un numero diverso di classi latenti e successivamente si può valutare quello che si adatta meglio ai dati, senza dimenticare il criterio della parsimonia. E' ovvio, infatti, che più classi consideriamo, maggiore sarà il valore della log-verosimiglianza nel punto di massimo. Tuttavia, all'aumentare del numero di classi, aumenta notevolmente il numero di parametri, con conseguenti difficoltà di interpretazione dei parametri stessi e perdita di potere esemplificativo da parte del modello.
Si consideri, infine, che il confronto tra modelli con un numero di classi latenti differente non può essere effettuato utilizzando il test del rapporto delle verosimiglianze. Nonostante i modelli risultino annidati, la soluzione di massima verosimiglianza per il modello con un numero inferiore di classi rappresenta una soluzione di frontiera per il modello con un numero maggiore di classi e quindi il test del rapporto delle verosimiglianze non è asintoticamente distribuito come un χ^2. Diversi metodi sono stati proposti in letteratura per ovviare a questo problema. Uno di questi è noto come Bayesian Information Criterion (BIC) (Schwarz, 1978) e consiste nel calcolare, per ciascun modello la quantità

$$BIC = -2 \cdot \log lik + p \cdot \log(N) \tag{6}$$

dove p indica il numero di parametri e N il numero totale di osservazioni. Il modello migliore, secondo questo criterio, è quello per cui il BIC risulta minore.

UN CONFRONTO EMPIRICO TRA I MODELLI

I dati utilizzati

I modelli presentati nella sezione precedente sono confrontati empiricamente sulla base di un insieme di dati concernenti le preferenze dichiarate di 51 imprese marchigiane in materia di trasporto merci. Le imprese intervistate appartengono ai settori ATECO "metallurgia e fabbricazione di prodotti in metallo" (DJ) e "altre industrie manifatturiere" (DN). La popolazione di riferimento è inoltre ristretta alle sole aziende con più di 40 addetti, essendo più probabile rintracciare in esse la figura di un responsabile della logistica da contattare e intervistare. A ciascuna delle imprese intervistate sono stati proposti 15 esercizi di scelta per un totale di 765 osservazioni. Ciascun esercizio di scelta prevedeva tre alternative, o profili di scelta, caratterizzate da livelli diversi degli attributi del servizio. Gli attributi considerati e i relativi livelli sono elencati nella Tabella 1.
Prima di procedere alla somministrazione degli esercizi di scelta ai vari intervistati, si è proceduto all'elicitazione dei loro cut-off chiedendo agli intervistati se vi fossero dei livelli degli attributi da considerare inaccettabili (Swait, 2001). L'inserimento dei cut-off nei

processi di stima permette di adottare curve di utilità "penalizzate", ovvero tali da tenere conto della presenza, nei profili di scelta, di livelli degli attributi che erano stati ex ante dichiarati come inaccettabili.

Tabella 1: Attributi e livelli considerati ai fini dell'analisi congiunta

Attributi	Livelli
Modo:	intermodale stradale
Costo del trasporto:	attuale Inferiore del 5% Superiore del 5% Inferiore del 10% Superiore del 10% Inferiore del 15% Superiore del 15%
Durata del viaggio:	attuale Inferiore di mezza giornata Superiore di mezza giornata Superiore di un giorno Superiore di due giorni
Puntualità:	100% delle spedizioni in orario 85% delle spedizioni in orario 70% delle spedizioni in orario
Danni e ammanchi	Probabilità nulla di danni e ammanchi Probabilità di danni e ammanchi del 5% Probabilità di danni e ammanchi del 10% Probabilità di danni e ammanchi del 20%
Frequenza	Alta Bassa
Flessibilità	Alta Bassa

L'assenza di eterogeneità: il modello logit multinomiale

Per avere un'idea preliminare dei risultati ed escludere dall'analisi successiva quegli attributi il cui effetto sull'utilità non risultasse significativo, si è inizialmente adattato ai dati un semplice modello di tipo MNL. L'utilità è espressa come funzione lineare degli attributi modo, costo, durata, puntualità, danni, frequenza e flessibilità. I cut-off sono stati inseriti sotto forma di variabili dummy: ad esempio, per includere nel modello la presenza di cut-off sul modo, si è creata la variabile $kmodo_{in}$ che assume valore 1 se l'alternativa i viola il cut-off sul modo dell'azienda n. Chiaramente se l'azienda n non esclude a priori la possibilità di utilizzare un trasporto di tipo intermodale, $kmodo_{in}$ sarà pari a 0 per ogni alternativa i. Anche le variabili relative alla presenza di cut-off violati sono state inserite linearmente nella funzione di utilità. In questo modo la violazione del cut-off in un'alternativa viene trattata alla stregua di ogni altro attributo dell'alternativa stessa. Si noti che questo modo di inserire i cut-

off nel modello equivale ad assumere dei cut-off "deboli" che l'azienda potrebbe essere disposta a violare nel caso i livelli degli altri attributi risultassero particolarmente vantaggiosi.
La Tabella 2 riporta le stime dei parametri per il MNL. Tale modello è stato ottenuto partendo da un MNL che includesse tutti gli attributi e i rispettivi cut-off e rimuovendo uno ad uno gli attributi o i cut-off con coefficienti non significativamente diversi da 0. La log-verosimiglianza per questo modello è risultata pari a –464.78, significativamente maggiore della log-verosimiglianza per il modello con tutti i parametri uguali a 0, pari a –840.44. L'R^2 aggiustato è 0.44. Per questo modello si ha inoltre BIC = -2·(-464.78)+8·log(765) = 982.68.
I cut-off hanno tutti un effetto negativo: il fatto che essi vengano violati in una determinata alternativa penalizza la probabilità che quella alternativa venga scelta. Gli attributi durata e puntualità in questo modello non risultano avere un effetto sull'utilità se non quello negativo espresso dai rispettivi cut-off. Sembra cioè che le aziende nello scegliere una certa alternativa di trasporto piuttosto che un'altra non prestino attenzione al particolare livello di questi due attributi ma solo al fatto che questo rispetti i loro cut-off. Il costo e i danni hanno invece un effetto negativo sull'utilità solo parzialmente assorbito dall'effetto dei rispettivi cut-off. Il trasporto intermodale, infine, sembra essere visto con interesse da quelle aziende che non lo escludono a priori.

Tabella 2: Stima dei parametri del modello logit multinomiale

Attributo	Coefficiente	Errore standard	Statistica *T*	p-value
Modo	0.838	0.177	4.744	0.0000
Costo	-6.318	1.150	-5.492	0.0000
Danni	-9.054	1.924	-4.706	0.0000
Kmodo	-0.479	0.223	-2.146	0.0319
Kcosto	-1.570	0.329	-4.769	0.0000
Kdurata	-0.835	0.169	-4.936	0.0000
Kpuntualità	-0.720	0.140	-5.153	0.0000
Kdanni	-0.832	0.240	-3.469	0.0005

Il modello mixed logit

La Tabella 3 mostra le stime dei parametri ottenute per un modello di tipo MXL. Inizialmente è stato adattato ai dati un modello in cui tutti i coefficienti fossero casuali e successivamente si sono progressivamente considerati come fissi (utilizzando una procedura di tipo stepwise) quei coefficienti la cui varianza non risultasse significativamente diversa da 0. Si sono inoltre sperimentati diversi tipi di distribuzioni per i parametri casuali. Il modello riportato nella Tabella 3 è quello che sembra spiegare meglio le scelte effettuate in termini di verosimiglianza. Per il costo e i danni si è assunta una distribuzione di tipo normale.
L'incremento nella log-verosimiglianza passando dal modello con parametri fissi a quello con parametri casuali è notevole (da -464.78 a -412.58) evidenziando una forte eterogeneità nei processi di scelta. Per il MXL si ha inoltre BIC = -2·(-412.581)+10·log(765) = 891,56, quindi l'aumento nel numero di parametri, passando dal MNL al MXL, è più che compensato dal miglior adattamento ai dati di quest'ultimo modello. L'eterogeneità colta dal MXL si manifesta in una sensibilità al costo e ai danni piuttosto variabile da azienda ad azienda. Entrambe le distribuzioni hanno una deviazione standard significativamente diversa da 0, pari rispettivamente a 11.6964 e a 15.3906.

Il tempo impiegato per stimare i parametri del MXL, utilizzando 100 estrazioni di tipo Halton, è stato di 165 secondi su un Pentium III, 700 MHz.

Tabella 3: Stima dei parametri del modello mixed logit

Attributo	Coefficiente	Errore standard	Statistica *T*	p-value
Modo	1.0958	0.2161	5.070	0.0000
Costo	-8.5409	2.2593	-3.780	0.0002
Danni	-24.0243	4.8996	-4.903	0.0000
Kmodo	-0.6055	0.2760	-2.193	0.0283
Kcosto	-2.0509	0.4167	-4.922	0.0000
Kdurata	-0.7004	0.1977	-3.543	0.0004
Kpuntualità	-1.0574	0.1715	-6.167	0.0000
Kdanni	-0.5652	0.3256	-1.736	0.0826
sd(Costo)	11.6964	1.8743	6.240	0.0000
sd(Danni)	15.3906	3.0397	5.063	0.0000

Il modello logit a classi latenti

I risultati ottenuti attraverso un modello logit con due classi latenti sono riportati nella Tabella 4. Utilizzando gli stessi attributi si è provato ad adattare ai dati anche un modello con tre e con quattro classi. Tuttavia, nel primo caso si è ottenuta una matrice di covarianze stimata non definita positiva, mentre nel secondo caso si sono ottenute delle stime di alcuni parametri con degli errori standard enormi. Entrambi questi problemi sono generalmente causati dalla scelta di un numero di classi latenti troppo elevato. Quindi, per i dati considerati, senza bisogno di fare ricorso all'uso del BIC, si è scelto un numero di classi pari a due. Il modello finale è stato inoltre ottenuto vincolando quei coefficienti che non erano risultati significativamente diversi nelle due classi, ad essere uguali.

Tabella 4: Stima dei parametri del modello logit a classi latenti

Attributo	Classe	Coefficienti comuni	Errore standard	Statistica *T*	p-value
Modo		0,7449	0,1542	4,830	0,0000
Danni		-10,5852	1,3413	-7,892	0,0000
Kcosto		-1,2541	0,2338	-5,363	0,0000
Attributo	**Classe**	**Coefficiente**	**Errore standard**	**Statistica *T***	**p-value**
Costo	1	-5,5458	1,5716	-3,529	0,0004
	2	-10,2587	0,9731	-10,542	0,0000
Kmodo	1	-1,3892	0,3773	-3,682	0,0002
	2	-0,3668	0,1797	-2,041	0,0413
Kdurata	1	-1,5334	0,2520	-6,085	0,0000
	2	-0,1428	0,1849	-0,772	0,4400
Kpuntualità	1	-1,4448	0,1823	-7,925	0,0000
	2	0,2714	0,0853	3,183	0,0015
Kdanni	1	-1,8725	0,3434	-5,453	0,0000
	2	-0,5201	0,2197	-2,367	0,0180
Probabilità delle classi		**Coefficiente**	**Errore standard**	**Statistica *T***	**p-value**
$\Pr(c=1)$		0,5500	0,0712	7,722	0,0000
$\Pr(c=2)$		0,4500	0,0712	6,318	0,0000

In base a tale modello le aziende sono omogenee per quanto riguarda l'avversione nei confronti dei danni e nei confronti della violazione del cut-off sul costo. Risultano divise in due gruppi distinti per quanto riguarda, invece, l'avversione al costo e alla violazione degli altri cut-off. In particolare, le aziende appartenenti alla prima classe (leggermente più

numerose di quelle appartenenti alla seconda) sembrano essere meno sensibili al costo. Le aziende della seconda classe sembrano, invece, essere meno restie a violare i propri cut-off.
La log-verosimiglianza per il modello presentato nella Tabella 4 è pari a -421.73 e si ha inoltre BIC = -2·(-421.730)+14·log(765) = 936.42. Se confrontiamo il LCM con il MXL, si può notare che la log-verosimiglianza è maggiore nel secondo caso (pari a -412.58), nonostante il numero di parametri considerati sia inferiore (pari a 10 contro i 14 del LCM). Quindi, chiaramente, il criterio del BIC privilegia ampiamente la scelta del MXL per i dati considerati. Confrontando direttamente la Tabella 3 con la Tabella 4 notiamo che mentre nel primo caso le aziende risultano eterogenee rispetto all'avversità ai danni, nel secondo caso tale eterogeneità non emerge. Probabilmente, un'eterogeneità rispetto ai danni sussiste ma questa non può essere colta attraverso una distribuzione discreta. A sua volta, il modello a coefficienti casuali non è invece in grado di cogliere l'eterogeneità delle aziende rispetto ai cut-off.
Il tempo impiegato per stimare i parametri del LCM, utilizzando un Pentium III, 700 MHz, è stato di 7 secondi.

CONCLUSIONI

In questo lavoro si sono confrontati su base empirica due modelli comunemente utilizzati in letteratura per tentare di catturare l'eterogeneità presente nei processi di scelta tra un numero discreto di alternative.
Il MXL assume che le preferenze varino in maniera continua nella popolazione di riferimento e che siano quindi rappresentabili attraverso una funzione di densità continua. I punti deboli del modello riguardano la scelta arbitraria di tale densità e il fatto che la probabilità che un individuo effettui una certa scelta sia espressa attraverso un integrale che non è risolvibile in forma chiusa. La stima dei parametri del modello richiede quindi l'impiego di metodi di simulazione. Questo comporta una complessità computazionale elevata e tempi di implementazione che possono diventare considerevoli nel caso in cui le osservazioni siano numerose. Inoltre, le stime sono affette da errori di simulazione. Un ulteriore problema è rappresentato dal fatto che la densità assunta è generalmente unimodale. Se la forma dell'eterogeneità nella popolazione è di tipo diverso, ad esempio bimodale, il MXL potrebbe essere incapace di rappresentare l'eterogeneità in maniera adeguata. Questo è quanto sembra succedere, ad esempio, per il coefficiente del costo. La sua deviazione standard è decisamente grande, probabilmente perchè il modello cerca di adattare all'eterogeneità una distribuzione unimodale (che risulta quindi molto dispersa) mentre nella popolazione esistono in realtà due gruppi distinti con due diverse avversioni al costo, come mostrato dal LCM.
Il LCM assume invece che la popolazione sia divisibile in classi e all'interno di ogni classe i soggetti siano caratterizzati da gusti omogenei. In questo caso, non è necessario fare assunzioni forti sulla forma distributiva dei coefficienti ma è sufficiente stabilire il numero di classi. Questo può essere tuttavia fatto a posteriori, utilizzando criteri quali il BIC che permettano di scegliere il modello che meglio si adatta ai dati. Le stime dei parametri sono ottenute iterativamente, senza bisogno di ricorrere a simulazioni. Esse non sono quindi affette da errori di simulazione e i tempi computazionali per ottenerle sono notevolmente inferiori

rispetto a quelli richiesti da un MXL: nel nostro caso ad esempio la stima dei coefficienti del LCM ha richiesto un tempo 24 volte inferiore rispetto a quello impiegato per stimare i coefficienti del MXL. Tuttavia l'assunzione fatta dal LCM che i soggetti nella stessa classe abbiano tutti esattamente gli stessi gusti non sempre risulta ragionevole. Soprattutto, quando la distribuzione dell'eterogeneità è di fatto continua, il LCM non è assolutamente adeguato e il BIC porterebbe probabilmente a scegliere un modello con un numero enorme di classi. Con riferimento ai dati analizzati, può darsi ad esempio che il LCM fallisca nel cogliere l'eterogeneità delle preferenze rispetto ai danni proprio perchè di fatto queste sono ben rappresentabili attraverso una distribuzione continua, come evidenziato dalla Tabella 3.
Nel complesso, l'eterogeneità presente nei dati considerati sembra essere spiegata meglio da un MXL, sebbene con i limiti visti. Un modo per superare tali limiti è rappresentato dal ricorso ad una modellizzazione più complessa che preveda la possibilità di assumere una distribuzione continua per alcuni dei coefficienti e una distribuzione discreta per altri. Un'ulteriore estensione può essere rappresentata dal rilassamento dell'ipotesi sull'omogeneità all'interno delle classi, ipotizzando ad esempio che, all'interno di ogni classe, le preferenze si distribuiscano secondo una normale. Questo equivale ad assumere un modello MXL in cui la distribuzione dei coefficienti non sia una normale ma una mistura di normali. In questo modo sarebbe possibile rappresentare l'eterogeneità anche nel caso in cui questa avesse distribuzione di tipo multimodale. Chiaramente, un modello del genere richiederebbe il ricorso a metodi di simulazione per poterne stimare i parametri, con le conseguenze viste sull'affidabilità delle stime risultanti e sui tempi computazionali. Tuttavia, in alcune situazioni empiriche, un tale aumento nella complessità computazionale potrebbe essere più che giustificato dalla capacità del modello di adattarsi ai dati e di rappresentare in maniera fedele l'eterogeneità in essi contenuta.

BIBLIOGRAFIA

Boxall, P. C. e W. L. Adamowicz (2002). Understanding heterogeneous preferences in random utility models: a latent class approach. *Environmental and Resource Economics*, **23**, 421-446.

Ben-Akiva, M. e S. R. Lerman (1985). *Discrete choice analysis*, Cambridge Mass, MIT Press.

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research*, **35B**, 677-693.

Green, W. H. e D. A. Hensher (2003a). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research*, **37B**, 681-698.

Green, W. H. e D. A. Hensher (2003b). The mixed logit model: the state of practice. *Transportation*, **30**, 133-176.

Heckman, J. e B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, **52**, 271-320.

Mc Fadden, D. e K. Train (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, **15**, 447-470.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

Train, K. (1999). Halton sequences for mixed logit. *Working paper*, Department of Economics, University of California, Berkeley.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge University Press.