

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in Scienze Statistiche



Un'applicazione combinata di modelli CUB e Anchoring Vignettes per l'analisi di dati ordinali

Relatore: Prof. Omar Paccagnella
Dipartimento di Scienze Statistiche

Laureanda: Serena Berretta
Matricola N. 1105181

Anno Accademico 2016/2017

Oreste stammi da sentire ...

Indice

Introduzione	1
1 Modelli per dati ordinali: CUB	5
1.1 Modello CUB	7
1.2 Modello CUB con covariate	11
1.3 Stima della massima verosimiglianza tramite l'algoritmo E-M	13
1.4 Matrice di Informazione osservata $\mathcal{I}(\hat{\theta})$	21
1.5 Aspetti inferenziali	25
1.6 Applicazioni dei modelli CUB	27
2 Anchoring Vignettes	35
2.1 Assunzioni	38
2.2 Metodo non parametrico	39
2.3 Metodo parametrico	40
2.3.1 Modello <i>Ordered Probit</i>	40
2.3.2 Modello <i>Chopit</i>	42
2.3.3 Estensioni del modello <i>Chopit</i>	45
2.4 Test sulle Assunzioni di RC e VE	46
2.5 Applicazioni Anchoring Vignettes	50
2.6 Combinazione Anchoring Vignettes e modelli CUB	51
3 Caso di studio	57
3.1 Dati <i>Online Banking</i>	57
3.2 Analisi preliminari	61
3.3 Modello CUB	68
3.3.1 AVCUB e test RC	76
3.3.2 Test per l'assunzione di VE	77
3.3.3 Genere della vignette	79
3.3.4 Ordine tra vignettes	82
3.3.5 <i>Online banking: seconda wave</i>	84

3.4	Modello <i>Chopit</i>	85
3.5	Ulteriori applicazioni dell'approccio AVCUB	88
	Conclusioni	91
A	Analisi descrittive <i>online banking</i>	99
B	Questionario <i>online banking</i>	109
C	Questionario progetto VECS	117
	Bibliografia	125

Elenco delle figure

1.1	Proiezione degli item nello spazio parametrico mediante la funzione <i>multicub()</i>	28
1.2	Stima di un modello CUB con covariate.	30
1.3	Stima di un modello CUB sulla soddisfazione lavorativa globale.	31
1.4	Stima di un modello CUB sulla soddisfazione lavorativa in funzione del guadagno mensile netto.	31
1.5	Probabilità stimate dal modello CUB per ciascun profilo.	32
1.6	Distribuzione delle frequenze delle valutazioni di tipo <i>ranking</i>	33
1.7	Probabilità stimate tramite un modello CUB con dati di tipo <i>ranking</i>	34
2.1	Analisi sulla <i>political efficacy</i> di Cina e Messico	37
3.1	Distribuzione della variabile professione nel campione.	61
3.2	Distribuzione delle classi d'età nel campione.	62
3.3	Distribuzione di frequenza delle risposte alle autovalutazioni.	63
3.4	Distribuzione di frequenza delle risposte alle vignettes, quando esse sono proposte dopo l'autovalutazione.	64
3.5	Distribuzione di frequenza delle risposte alle vignettes, quando esse sono proposte prima dell'autovalutazione.	64
3.6	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al genere degli individui.	65
3.7	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al genere degli individui.	66
3.8	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al genere degli individui.	66
3.9	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al genere degli individui.	67
3.10	Distribuzione indicatore oggettivo per il blocco di domande Q13-Q15-Q16.	86

3.11	Distribuzione indicatore oggettivo per il blocco di domande Q19-Q21-Q22.	87
A.1	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al titolo di studio raggiunto dagli individui.	99
A.2	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al titolo di studio raggiunto dagli individui.	100
A.3	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al titolo di studio raggiunto dagli individui.	100
A.4	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al titolo di studio raggiunto dagli individui.	100
A.5	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto all'età degli individui.	101
A.6	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto all'età degli individui.	101
A.7	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto all'età degli individui.	101
A.8	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto all'età degli individui.	102
A.9	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla professione degli individui.	102
A.10	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla professione degli individui.	102
A.11	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla professione degli individui.	103
A.12	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla professione degli individui.	103
A.13	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al tipo di dispositivo impiegato nel questionario dagli individui.	103
A.14	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al tipo di dispositivo impiegato nel questionario dagli individui.	104
A.15	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al tipo di dispositivo impiegato nel questionario dagli individui.	104
A.16	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al tipo di dispositivo impiegato nel questionario dagli individui.	104
A.17	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al tipo di browser impiegato nel questionario dagli individui.	105

A.18	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al tipo di browser impiegato nel questionario dagli individui.	105
A.19	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al tipo di browser impiegato nel questionario dagli individui.	105
A.20	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al tipo di browser impiegato nel questionario dagli individui.	106
A.21	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla presenza di problemi durante il servizio di <i>online banking</i>	106
A.22	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla presenza di problemi durante il servizio di <i>online banking</i>	106
A.23	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla presenza di problemi durante il servizio di <i>online banking</i>	107
A.24	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla presenza di problemi durante il servizio di <i>online banking</i>	107
A.25	Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla presenza di figli a carico del rispondente.	107
A.26	Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla presenza di figli a carico del rispondente.	108
A.27	Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla presenza di figli a carico del rispondente.	108
A.28	Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla presenza di figli a carico del rispondente.	108

Elenco delle tabelle

1.1	Notazione standard dei modelli CUB.	13
1.2	Notazione per la costruzione della matrice di Informazione per modelli CUB(1,0) e CUB(0,1).	23
1.3	Confronto tra modelli tramite test di log-rapporto di verosimiglianza e gradi di libertà associati.	26
1.4	Stima di un modello CUB senza covariate per <i>officeho</i>	29
1.5	Profili costruiti in base alle caratteristiche del soggetto introdotte nel modello.	32
3.1	Ordine di somministrazione del questionario sull' <i>Online Banking</i> per ciascun gruppo.	59
3.2	Stime dei parametri con modello CUB per per la domanda Q13 (autovalutazione sulla facilità di navigazione).	70
3.3	Stime dei parametri con modello CUB per per la domanda Q15 (vignette di Carlo sulla facilità di navigazione).	70
3.4	Stime dei parametri con modello CUB per la domanda Q16 (vignette di Marina sulla facilità di navigazione).	71
3.5	Stime dei parametri con modello CUB per la domanda Q19 (autovalutazione sulla semplicità di esecuzione delle operazioni bancarie online).	71
3.6	Stime dei parametri con modello CUB per la domanda Q21 (vignette di Sofia sulla semplicità di esecuzione delle operazioni bancarie online).	72
3.7	Stime dei parametri con modello CUB per la domanda Q22 (vignette di Pietro sulla semplicità di esecuzione delle operazioni bancarie online).	72
3.8	Stime dei parametri con modello CUB per per la domanda Q13bis (autovalutazione sulla facilità di navigazione).	73
3.9	Stime dei parametri con modello CUB per per la domanda Q15bis (vignette di Carlo sulla facilità di navigazione).	73
3.10	Stime dei parametri con modello CUB per per la domanda Q16bis (vignette di Marina sulla facilità di navigazione).	74

3.11	Stime dei parametri con modello CUB per la domanda Q19 (autovalutazione sulla semplicità di esecuzione delle operazioni bancarie online).	74
3.12	Stime dei parametri con modello CUB per la domanda Q21bis (vignette di Sofia sulla semplicità di esecuzione delle operazioni bancarie online).	75
3.13	Stime dei parametri con modello CUB per la domanda Q22bis (vignette di Pietro sulla semplicità di esecuzione delle operazioni bancarie online).	75
3.14	LR test per la convalida dell'assunzione di RC.	76
3.15	Risultati dell'LR test per l'assunzione di VE.	78
3.16	LR test per la convalida della RC in relazione al genere delle vignettes Q15 (Uomo) e Q16 (Donna).	80
3.17	LR test per la convalida della RC in relazione al genere delle vignettes Q21 (Donna) e Q22 (Uomo).	80
3.18	LR test per la convalida della VE in relazione al genere delle vignettes Q15(Uomo), Q16(Donna), Q21(Donna) e Q22(Uomo).	81
3.19	Test LR per la verifica dell'assunzione di RC per il blocco di domande Q13-Q15-Q16 per i 4 gruppi di intervistati.	82
3.20	Test LR per la verifica dell'assunzione di RC per il blocco di domande Q19-Q21-Q22 per i 4 gruppi di intervistati.	83
3.21	Test LR per la verifica dell'assunzione di VE per le vignettes Q15, Q16, Q21 e Q22, disaggregate per i 4 gruppi di intervistati.	84
3.22	LR test per la convalida della RC per Q15 e Q16 rispetto ai clienti vecchi o nuovi della seconda <i>wave</i> .	85
3.23	LR test per la convalida della RC per Q21 e Q22 rispetto ai clienti vecchi o nuovi della seconda <i>wave</i> .	85
3.24	Metodo AVCUB per la validazione della RC sui dati del LISS panel.	90
3.25	Test LR per la validazione della VE sui dati del LISS panel.	91
3.26	Riassunto dei risultati sull' <i>online banking</i> della validazione delle assunzioni di RC e VE.	97

Introduzione

Un contesto molto ampio in cui si sviluppa lo studio dei dati ordinali sono le indagini socio-economiche, nelle quali è comune raccogliere valutazioni soggettive riguardanti la salute degli individui e/o le loro condizioni di vita, la soddisfazione lavorativa (*job satisfaction*) o la soddisfazione dei clienti (*customer satisfaction*). In generale, le domande sulle attitudini degli individui o sulle loro opinioni provano a misurare una variabile latente continua sottostante che, per ragioni pratiche, è spesso espressa attraverso un set di categorie ordinali che l'individuo sceglie come risposta in base al suo gradimento. In particolare, l'usuale scala di riferimento è quella di Likert (Likert, 1932) che fornisce una descrizione verbale dei livelli di risposta ordinali (ad esempio da "fortemente in disaccordo" a "fortemente d'accordo"). Il modo più naturale di vedere questo processo di misura è, dunque, ipotizzare l'esistenza di una variabile latente non osservata associata ad ogni risposta. Tale variabile è spesso assunta essere descritta da una distribuzione continua che varia da individuo a individuo. Vi è una vasta letteratura sulla modellazione di dati ordinali, dai primi articoli di McKelvey e Zavoina (1975) e McCullagh (1980) ai più recenti libri di testo di O'Connell (2006), Agresti (2010) e Greene & Hensher (2010). Nell'ultima decade sono state introdotte in letteratura nuove ed innovative soluzioni, con l'obiettivo di interpretare e modellare i dati ordinali: tra questi i modelli CUB e l'approccio con le Anchoring Vignettes.

L'approccio con i modelli CUB e quello tramite le Anchoring Vignettes mostrano interessanti similitudini, tra cui la mappatura della variabile continua latente in

valori discreti ordinali. Il processo psicologico, che guida i rispondenti a tradurre il loro giudizio in una risposta di tipo ordinale, è caratterizzato dal fatto che la risposta non è legata esclusivamente al contenuto della domanda. Infatti, la presenza di eterogeneità tra gli individui, dovuta a differenze interpersonali e interculturali nell'interpretare, capire o usare le categorie di risposta per la stessa domanda, può creare una distorsione sistematica nella misurazione delle variabili, producendo eventualmente valutazioni fuorvianti del fenomeno di interesse.

I modelli CUB sono una nuova classe di modelli statistici nei quali la risposta è modellata come una mistura di due componenti latenti: una riguardante il sentimento dell'individuo verso l'oggetto (*feeling*) e un'altra riguardante l'incertezza del processo di risposta. Esistono diverse estensioni di questi modelli che permettono di legare i parametri relativi al *feeling* e all'incertezza a determinate covariate. Differentemente dagli approcci tradizionali, quali i modelli di regressione logistica o i modelli *Probit*, i modelli CUB non introducono nessun parametro di soglia (*threshold*).

Le Anchoring Vignettes sono delle domande aggiuntive indirizzate al rispondente che possono aiutare a identificare le differenze sistematiche nell'uso delle scale di risposta. Difatti, queste descrivono un ipotetico scenario dove un'ipotetica persona è posta in una particolare situazione o condizione. I rispondenti sono poi chiamati a valutare questa persona usando la stessa scala adottata per l'autovalutazione. Questa assunzione, detta di *Response consistency* (RC), in base alla quale si ipotizza che le persone valutino se stessi e le descrizioni degli scenari posti attraverso le vignettes applicando gli stessi stili di risposta, fa parte di una delle assunzioni base per la modellazione delle risposte alle Anchoring Vignettes. L'altra assunzione è detta *Vignette Equivalence* (VE) e prevede che ogni rispondente percepisca le domande delle vignettes allo stesso modo. La validità di queste due assunzioni è tuttora ancora oggetto di dibattito, dato che attualmente non esistono delle soluzioni per testarle che non si basino sull'assunzione di altre forti condizioni (Paccagnella, 2013). Per i modelli che sfruttano le risposte alle Anchoring Vignettes è possibile utilizzare sia

un approccio parametrico che uno non parametrico. La soluzione parametrica è chiamata modello *Chopit* e può essere descritta come una modellazione congiunta di risposte di autovalutazione e risposte alle vignettes, tramite l'utilizzo di un modello *Probit* per dati ordinali. Differentemente dalla soluzione dei modelli CUB, il modello *Chopit* prevede la specifica di soglie (*threshold*) individuali.

Questo lavoro di tesi ha come obiettivo quello di integrare le principali caratteristiche e i vantaggi di questi due approcci, al fine di migliorare la qualità delle analisi dei dati ordinali (per esempio nella comparazione di dati provenienti da diversi Paesi) e superare i limiti che caratterizzano ciascun modello. In particolare, si cercherà di risolvere il problema delle assunzioni di RC e VE, proponendo l'utilizzo della metodologia CUB per effettuare dei test che permettano di verificarne la validità. I risultati di questa tesi possono essere utili in due direzioni: da un punto di vista metodologico, può arricchire la letteratura sulla modellazione dei dati ordinali proponendo modelli che possono spiegare meglio il ruolo dell'incertezza nel processo che porta gli intervistati a tradurre il loro giudizio in una risposta ordinale; allo stesso tempo, può fornire soluzioni per testare le assunzioni delle vignettes senza imporre ulteriori restrizioni e dare indicazioni sulla costruzione del questionario.

Perciò la tesi è organizzata come segue. Il Capitolo 1 è dedicato alla specificazione dei modelli CUB, con un particolare approfondimento sulla loro stima tramite l'algoritmo E-M e sul calcolo della matrice di Informazione osservata. Seguono poi gli aspetti inferenziali e degli esempi di applicazione di questi modelli in differenti ambiti.

Il Capitolo 2 tratta dell'approccio delle Anchoring Vignettes, definendo le sue assunzioni e i diversi metodi di stima (non parametrico e parametrico). Vengono poi riportati i test, già introdotti in letteratura, per la validazione delle assunzioni ed, infine, viene specificata la soluzione proposta in questa tesi che prevede la combinazione dei due approcci per dati ordinali con la costruzione di una funzione che testa l'ipotesi di RC e VE.

Il Capitolo 3 è riservato all'applicazione ad una caso di studio reale, riguardante la soddisfazione di clienti italiani per un servizio di *online banking*, delle tecniche approfondite nei capitoli precedenti. In particolare, dopo delle brevi analisi descrittive si sono applicati i modelli CUB alle singole vignettes e autovalutazioni, per poi andare a testare le assunzioni di RC prima, e VE dopo. Si è approfondita l'analisi valutando possibili variazioni nel considerare il genere della vignettes, l'ordine di esposizione delle vignettes ed anche una seconda serie di dati raccolti a distanza di quattro mesi dalla prima *wave*. Successivamente, si è cercato di validare i risultati tramite il modello *Chopit* ed anche attraverso l'utilizzo di un altro set di dati sulla soddisfazione dopo l'acquisto di un computer portatile, di uno smartphone o di una TV LCD da parte di un campione di clienti olandesi.

L'ultimo capitolo riporta le conclusioni di questo lavoro di tesi, sintetizzando i principali risultati ottenuti con il metodo introdotto per la verifica delle assunzioni previste dalle Anchoring Vignettes, tramite l'utilizzo dei modelli CUB.

Capitolo 1

Modelli per dati ordinali: CUB

I modelli per dati ordinali sorgono, ad esempio, in un contesto in cui ad un soggetto viene richiesto di esprimere una valutazione riguardante differenti argomenti (oggetti, servizi, ecc). Tale giudizio viene espresso tramite una risposta di tipo categoriale. Esistono due tipologie di schemi di risposta: quello *ranking* e quello *rating*. Nell'approccio di tipo *ranking* ogni persona assegna una posizione a ciascun oggetto della lista che deve valutare, dandogli così un giudizio indiretto. Difatti, la posizione assegnata ad un singolo oggetto è fortemente condizionata dalle caratteristiche di tutti gli altri. Ciò suggerisce di prendere in considerazione una distribuzione congiunta delle risposte ai singoli oggetti per un'analisi corretta del fenomeno. Le analisi statistiche prevedono l'uso di modelli con permutazioni o con variabili latenti (Fligner e Verducci, 1993; Marden, 1996; Jöreskog e Moustaki, 2001; Moustaki, 2003). Al contrario, nell'approccio di tipo *rating* il rispondente valuta gli oggetti tramite un set limitato di categorie e sceglie quella che più rappresenta il suo pensiero e la sua preferenza riguardo l'oggetto da valutare. Questo insieme di categorie può avere una scala ordinata o semplicemente una scala nominale.

In questo lavoro di tesi ci riferiremo solo al caso in cui il giudizio dei rispondenti è espresso tramite un approccio *rating* con scala ordinata. In particolare l'usuale scala di riferimento è quella di Likert (Likert, 1932); un esempio, con 5 categorie

possibili di scelta, è il seguente:

1. "Fortemente d'accordo"
2. "D'accordo"
3. "Nè d'accordo nè in disaccordo"
4. "In disaccordo"
5. "Fortemente in disaccordo"

La risposta che il soggetto assegna ad un oggetto è un intero appartenente al supporto $\{1, \dots, m\}$, per un dato numero m di categorie. Le tecniche usate per questo tipo di analisi si basano sui Modelli Lineari Generalizzati (GLM) (McCullagh, 1984; McCullagh e Nelder, 1989; Agresti, 2010). Per questi modelli il modo più naturale per capire il processo di misura è quello di assumere l'esistenza di una variabile latente non osservata sottostante associata alla risposta categoriale osservata. Questo tratto latente è generalmente considerato essere espresso da una distribuzione continua diversa da individuo a individuo. La probabilità che una risposta non superi una determinata categoria è modellata come una funzione lineare di un insieme di covariate. Dunque, basandosi sulle stime delle soglie, per discriminare tra una categoria e l'altra, la variabile latente può essere divisa in classi, così da ottenere la risposta del soggetto in modo discreto. Nel seguito si studierà l'approccio con i modelli CUB, dove si analizza direttamente il processo psicologico che guida i rispondenti alla scelta di una alternativa tra le m a disposizione, senza richiedere alcun calcolo delle soglie. In questo caso le covariate saranno legate direttamente ai parametri tramite una funzione legame (solitamente di tipo *logit*) e, diversamente dai modelli GLM, non prevederanno il legame attraverso il valore atteso¹.

¹Infatti, per un dato numero di categorie, m , differenti coppie di parametri generano lo stesso valore atteso. Dunque, è preferibile un link esplicito tra i parametri del modello e le covariate del soggetto, senza alcun riferimento al valore atteso (Piccolo, 2006).

1.1 Modello CUB

I modelli CUB sono stati introdotti da Piccolo(2003) e D'Elia e Piccolo (2005); successivamente sono stati discussi anche da Iannario (2008), Piccolo e D'Elia (2008) e Iannario e Piccolo (2010, 2012). La loro specificazione si è resa utile per descrivere il processo psicologico che guida il rispondente a tradurre il suo giudizio in una risposta ordinale, esso è infatti il risultato di una complessa decisione. Tale scelta può essere sintetizzata considerando due fattori: l'attrattiva del soggetto verso l'oggetto (detta *feeling*) e l'incertezza che caratterizza la scelta finale, dovuta spesso a circostanze esterne. Dunque, in questa nuova classe di modelli statistici la risposta è modellata come una mistura di due componenti latenti. La prima è relativa al sentimento dell'individuo verso l'oggetto, ossia la parte che riguarda la sua opinione verso ciò che deve valutare. La seconda componente è invece legata all'incertezza del processo di risposta e deriva dall'indecisione della scelta finale (Piccolo 2003; D'Elia e Piccolo 2005). Qui, l'incertezza non è intesa come la solita componente stocastica relativa al campionamento, ma è il risultato di possibili fattori esterni, quali ad esempio:

- il tempo necessario per l'elaborazione della domanda
- la limitata gamma di informazioni disponibili
- la comprensione parziale della domanda
- la pigrizia del soggetto e la volontà di voler scherzare o rispondere il falso (Iannario e Piccolo, 2012)

e così via.

In modo da non creare ambiguità, può essere utile distinguere tra il concetto di incertezza e quello di casualità. L'incertezza si riferisce all'indecisione del soggetto ed è legata alla natura stessa della scelta umana. Essa sarà considerata proprio come componente della struttura del modello CUB. Invece, la casualità si riferisce al modo in cui i dati sono raccolti dalla popolazione ed è legata alla selezione del

campione, agli errori di misura e a conoscenze limitate. Essa è tenuta in considerazione attraverso il paradigma della variabile casuale (Iannario e Piccolo, 2012). Nei modelli CUB la parte di *feeling* è modellata con una variabile casuale *Binomiale traslata* che risulta essere la versione discreta di una componente latente unimodale continua; tale scelta risulta la più appropriata come formalmente giustificato da Iannario e Piccolo (2012). Infatti, sia R la variabile casuale generata dalla selezione di una delle categorie ordinali $r \in \{1, \dots, m\}$, dove r aumenta quando aumenta il *feeling* verso l'oggetto. Se un soggetto sceglie la categoria r significa che considera le precedenti $r - 1$ troppo deboli e le successive $m - r$ troppo forti per esprimere correttamente il proprio giudizio verso l'oggetto. Si assuma $\xi \in [0, 1]$ un numero reale tale che $(1 - \xi)$ sia la probabilità che un giudizio possa essere considerato inferiore a quello scelto dal rispondente (probabilità di successo). Allora, considerando gli $m - 1$ possibili confronti, la probabilità di scegliere la categoria r tra le m possibili ha come distribuzione quella della *Binomiale traslata*²:

$$b_r(\xi) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}, \quad r = 1, \dots, m \quad (1.1)$$

La componente di incertezza è modellata da una variabile casuale *Uniforme discreta*³ che assegna la stessa probabilità di scelta ad ognuna delle m categorie. La funzione di probabilità risulta:

$$U_r(m) = \frac{1}{m}, \quad r = 1, \dots, m \quad (1.2)$$

Con riferimento al comportamento dell'individuo rispondente, l'utilizzo della distribuzione Uniforme sarebbe la soluzione estrema nel caso di una scelta completamente casuale, dove nessuna delle categorie è preferita alle altre.

Il momento della scelta è sempre caratterizzato da entrambi gli aspetti appena descritti, ma non necessariamente con la stessa influenza (50% di *feeling* e 50% di

²Si preferisce la *Binomiale traslata* in quanto la scelta più comune del set di valori assunti dalla variabile è $\{1, \dots, m\}$. Diversamente, il supporto della Binomiale partirebbe da 0.

³La variabile casuale Uniforme massimizza l'entropia tra tutte le distribuzioni con supporto limitato $\{1, \dots, m\}$, fissato m .

incertezza). Come conseguenza di ciò si introduce una variabile casuale mistura costituita dalle due parti propriamente pesate. Allora, si indica con r_i la risposta osservata dell'individuo i -esimo che è la realizzazione di una variabile casuale R distribuita come una *Combinazione di una variabile discreta Uniforme e una Binomiale traslata (CUB)*⁴. Il dataset è composto dunque dalla collezione di giudizi (r_1, \dots, r_n) degli n individui, dove ogni risposta è scelta tra le m categorie possibili $r_i \in \{1, \dots, m\}$. Dunque il modello CUB è così definito:

$$P_r(R = r) = \pi \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1} + (1-\pi) \frac{1}{m}, \quad r = 1, \dots, m \quad (1.3)$$

Questo modello è completamente identificato con $m > 3$ ⁵, come provato da Iannario (2010), ed è ben definito per i parametri $\theta = (\pi, \xi)$ appartenenti allo spazio parametrico $\Omega(\theta) = \{(\pi, \xi) : 0 \leq \pi \leq 1, 0 \leq \xi \leq 1\}$. I modelli CUB risultano essere molto flessibili in quanto la loro struttura stocastica varia molto con il variare dei parametri. A riguardo di quanto appena detto, si possono elencare diverse caratteristiche della distribuzione CUB (Piccolo, 2006):

- è una variabile casuale reversibile:

$$R \sim CUB(\pi, \xi) \implies (m+1-R) \sim CUB(\pi, 1-\xi)$$

- è consistente con l'ipotesi che la popolazione sia composta da due sotto-gruppi di valutatori: un gruppo più informato e riflessivo e un altro più disinformato e istintivo. La loro proporzione relativa è $\frac{\pi}{1-\pi}$.

- emula diverse distribuzioni:

⁴Originariamente questa variabile era chiamata MUB (Mistura di una Uniforme e una Binomiale) come in Piccolo (2003), D'Elia e Piccolo (2005).

⁵Questa restrizione evita il caso di una variabile casuale degenera ($m=1$), di una indeterminata ($m=2$) o di un modello saturo ($m=3$).

- una distribuzione Uniforme se $\pi = 0$
- una distribuzione Binomiale traslata se $\pi = 1$
- una distribuzione Ipergeometrica inversa se $\pi \rightarrow 1$ e ξ tende a 0 o 1
- una distribuzione Normale se $\xi \rightarrow \frac{1}{2}$ e $m \rightarrow \infty$

Il valore atteso e la varianza della variabile casuale R sono rispettivamente:

$$E(R) = \pi(m-1)\left(\frac{1}{2} - \xi\right) + \frac{m+1}{2} \quad (1.4)$$

$$Var(R) = (m-1) \left\{ \pi\xi(1-\xi) + (1-\pi) \left[\frac{m+1}{12} + \pi(m-1) \left(\frac{1}{2} - \xi \right)^2 \right] \right\}$$

Il valore atteso nella formula 1.4 dipende dal segno di $\left(\frac{1}{2} - \xi\right)$; infatti, si avranno valori più bassi della media quando $\xi \rightarrow 1$ come confermato dall'inclinazione (parametro di *skewness*) della distribuzione che è appunto regolata dalla stessa quantità. La variabile è quindi simmetrica solo se $\xi = \frac{1}{2}$. Per interpretare correttamente i parametri coinvolti nel modello CUB è sufficiente pensare che ogni soggetto decide con una certa propensione, π , di scegliere una certa alternativa e $(1 - \pi)$ di sceglierne un'altra casualmente. Quando $\pi \rightarrow 0$ aumenta la propensione ad una scelta completamente casuale di risposta, infatti si misura l'incertezza attraverso la quantità $(1 - \pi)$. Per quanto riguarda l'altro parametro, invece, è sbagliato dire che se $\xi \rightarrow 0$ l'individuo esprime un giudizio positivo riguardo all'oggetto da valutare e viceversa se $\xi \rightarrow 1$. Per interpretarlo correttamente è necessario considerare la direzione della scala di giudizio che dipende da come le risposte sono codificate. Più precisamente, se si lavora con valutazioni di tipo *ranking* (dove $r = 1$ rappresenta il massimo gradimento, mentre $r = m$ il minimo) allora il parametro ξ è direttamente una misura del *feeling*. Invece, se si considerano valutazioni di tipo *rating* (dove $r = 1$ rappresenta il minimo gradimento e $r = m$ la massima soddisfazione), allora

il parametro che misura il *feeling* è $1 - \xi$. Comunque, grazie alla corrispondenza uno a uno tra i modelli CUB e i punti nello spazio parametrico $\Omega(\theta)$, una grande quantità di valutazioni possono essere sintetizzate da un ridotto set di punti nello spazio parametrico dove le coordinate sono espresse da (π, ξ) . Quando vengono collocati i parametri stimati dal modello CUB nello spazio parametrico è possibile confrontare più modelli contemporaneamente sintetizzando gli aspetti salienti contenuti nei dati:

- $(1 - \pi)$ quantifica la propensione dei rispondenti a fornire risposte totalmente casuali: più un punto è spostato a destra del quadrato unitario e più le valutazioni riflettono risposte con alta incertezza. Viceversa, più un item è spostato verso sinistra del grafico, più le risposte sono state ben definite e hanno poca incertezza.
- $(1 - \xi)$ misura il livello di *feeling* espresso dal rispondente verso un oggetto: più vicino è $(1 - \xi)$ al bordo superiore del quadrato unitario e più l'item è preferito. D'altra parte, più i punti sono spostati nella parte inferiore del grafico, meno l'oggetto della valutazione è stato apprezzato dai rispondenti.

1.2 Modello CUB con covariate

La classe dei modelli CUB può essere generalizzata assumendo che, dati n soggetti, i parametri di incertezza e di *feeling* siano, separatamente, legati ad alcune covariate attraverso un link logistico (Iannario, 2008). L'introduzione delle covariate nei modelli CUB permette di migliorarne l'adattamento e di discriminare differenti sottogruppi di persone. Le variabili introdotte nel modello possono essere di tipo quantitativo, quali l'età, il reddito familiare e molte altre, oppure di tipo qualitativo, ad esempio il genere o lo stato civile. Inoltre, è possibile considerare anche covariate che descrivono l'oggetto che il rispondente deve valutare, così da poter catturare possibili reazioni differenti nel processo decisionale in base alle caratteristiche dell'oggetto in esame. Sia Y il vettore delle covariate che si riferiscono all'incertezza

e W quello delle covariate che si riferiscono al *feeling*. In termini matriciali, si indicano come covariate $Y = (1, y_{i1}, y_{i2}, \dots, y_{ip})$ e $W = (1, w_{i1}, w_{i2}, \dots, w_{iq})$, dato un campione di dimensione n di dati di tipo ordinale $r = (r_1, \dots, r_n)'$. In generale, con $m > 3$ categorie, la formulazione di un modello CUB con p e q covariate per l'incertezza e il *feeling* rispettivamente (CUB(p,q)), è espresso come:

- una componente stocastica

$$P_r(R_i = r_i | y_i; w_i) = \pi_i \binom{m-1}{r-1} \xi_i^{m-r} (1 - \xi_i)^{r-1} + (1 - \pi_i) \frac{1}{m}$$

- due componenti sistematiche

$$\pi_i = \pi_i(\beta) = \frac{1}{1 + e^{-y_i \beta}}, \quad \xi_i = \xi_i(\gamma) = \frac{1}{1 + e^{-w_i \gamma}} \quad (1.5)$$

dove $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ e $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_q)'$ sono i vettori dei parametri da stimare. Utilizzando la funzione logistica:

$$\text{logit}(p) = \log \left(\frac{p}{1-p} \right)$$

è possibile riscrivere le relazioni tra le covariate e i parametri dell'equazione 1.5 come segue:

$$\text{logit}(\pi_i) = y_i \beta; \quad \text{logit}(\xi_i) = w_i \gamma; \quad i = 1, \dots, n$$

Esplicitando la distribuzione del modello CUB con covariate si ottiene:

$$P_r(R_i = r_i | y_i; w_i) = \frac{1}{1 + e^{-y_i \beta}} \left[\binom{m-1}{r_i-1} \frac{(e^{-w_i \gamma})^{r_i-1}}{(1 + e^{-w_i \gamma})^{m-1}} - \frac{1}{m} \right] + \frac{1}{m} \quad (1.6)$$

Per visualizzare l'effetto delle covariate sul *feeling* e sull'incertezza è possibile utilizzare diverse tipologie di grafico, ma la principale prevede di rappresentare il parametro $(1 - \xi)$ (oppure $(1 - \pi)$) rispetto alla k -esima covariata di interesse w_k (oppure y_k). Difatti, dall'equazione 1.5 si vede che all'aumentare dei valori di w_k (oppure y_k)

vi è un aumento del *feeling* (o dell'incertezza) se $\gamma_k < 0$ (oppure $\beta_k < 0$), viceversa, vi è un decremento se $\gamma_k > 0$ (oppure $\beta_k > 0$) (Iannario e Piccolo, 2012a). La notazione prevista per i modelli CUB che sarà usata nel seguito è descritta in Tabella 1.1.

Modello	Covariate	Vettore dei parametri	Spazio dei parametri	Numero di parametri
$CUB(0,0)$	no covariate	$\theta = (\pi, \xi)'$	$(0, 1] * [0, 1]$	2
$CUB(p,0)$	covariate per π	$\theta = (\beta', \xi)'$	$\mathbb{R}^{p+1} * [0, 1]$	$p + 2$
$CUB(0,q)$	covariate per ξ	$\theta = (\pi, \gamma)'$	$(0, 1] * \mathbb{R}^{q+1}$	$q + 2$
$CUB(p,q)$	covariate per π e ξ	$\theta = (\beta', \gamma)'$	\mathbb{R}^{p+q+2}	$p + q + 2$

Tabella 1.1: Notazione standard dei modelli CUB.

1.3 Stima della massima verosimiglianza tramite l'algoritmo E-M

Come presentato in Piccolo (2006), si mostreranno in questa sezione i passi che portano alla stima di un modello CUB e dei suoi parametri, tramite l'utilizzo dell'algoritmo E-M⁶. Saranno distinte le casistiche in base alla presenza o meno di covariate per la parte di *feeling* e/o di incertezza, utilizzando la stessa notazione vista nella Tabella 1.1. Per primo si presenterà l'algoritmo in assenza di covariate; successivamente, la procedura sarà mostrata in presenza di un solo set di covariate (o per il *feeling* o per l'incertezza). Infine si mostrerà la generalizzazione del modello CUB, dove sono presenti i set di covariate per entrambe le componenti. Per semplificare l'esposizione degli algoritmi si considererà una sola covariata per componente.

⁶Una discussione generale dell'algoritmo E-M in ambito statistico è presentata in McLachlan e Krishnan (1997) e, per modelli mistura, in McLachlan e Peel (2000)

Innanzitutto, si definisce la funzione di *log-verosimiglianza* dei modelli **CUB senza covariate**:

$$l_{00}(\theta) = \sum_{r=1}^m n_r \log \left\{ \pi \left[b(r; \xi) - \frac{1}{m} \right] + \frac{1}{m} \right\}$$

dove n_r sono le frequenze osservate per ciascun valore di $R = r$ e $b(r; \xi)$ rappresenta la distribuzione della Binomiale traslata.

La funzione di *log-verosimiglianza* del modello **CUB(1,0)**, dove π è espresso da un set di covariate attraverso una funzione, $\pi = f(Y, \beta)$, è:

$$l_{10}(\theta) = - \sum_{i=1}^n \log(1 + e^{-y_i \beta}) + \sum_{i=1}^n \log \left(b(r_i; \xi) + \frac{e^{-y_i \beta}}{m} \right)$$

La funzione di *log-verosimiglianza* del modello **CUB(0,1)**, dove ξ è espresso da un set di covariate attraverso una funzione, $\xi = g(W, \gamma)$, è:

$$l_{01}(\theta) = \sum_{i=1}^n \log \left\{ \pi \left[\binom{m-1}{r_i-1} \frac{(e^{-w_i \gamma})^{r_i-1}}{(1 + e^{-w_i \gamma})^{m-1}} - \frac{1}{m} \right] + \frac{1}{m} \right\}$$

Infine, se entrambi i parametri, π e ξ , sono espressi da un set di covariate tramite delle funzioni, $\pi = f(Y, \beta)$ e $\xi = g(W, \gamma)$, la funzione di *log-verosimiglianza* del modello **CUB(1,1)** risulta:

$$l_{11}(\theta) = \sum_{i=1}^n \log \left\{ \frac{1}{1 + e^{-y_i \beta}} \left[\binom{m-1}{r_i-1} \frac{(e^{-w_i \gamma})^{r_i-1}}{(1 + e^{-w_i \gamma})^{m-1}} - \frac{1}{m} \right] + \frac{1}{m} \right\}$$

Algoritmo E-M per il modello CUB(0,0)

- **Step 0:** si inizializzano i parametri e la *log-verosimiglianza*.

$$\xi^{(0)} = 1 + \frac{0.5 - M_n}{m}; \quad \pi^{(0)} = \min \left(\sqrt{\frac{\sum_{r=1}^m (f_r^2 - \frac{1}{m})}{\sum_{r=1}^m ([b_r(\xi)]^2 - \frac{1}{m})}}, 0.95 \right); \quad l^{(0)} = l_{00}(\pi^{(0)}, \xi^{(0)})$$

dove M_n rappresenta la classe con maggior frequenza osservata:

$$M_n = j \iff f_j > f_k, \quad \forall k \neq j = 1, 2, \dots, m$$

- **Step 1:** si calcola il valore della Binomiale traslata.

$$b(r; \xi^{(k)}) = \binom{m-1}{r-1} (\xi^{(k)})^{m-r} (1 - \xi^{(k)})^{r-1}, \quad r = 1, \dots, m$$

- **Step 2:** si calcolano le quantità utili per l'aggiornamento delle stime dei parametri.

$$\tau(r; \theta^{(k)}) = \left[1 + \frac{1 - \pi^{(k)}}{m\pi^{(k)}b(r; \xi^{(k)})} \right]^{-1}$$

- **Step 3:** si calcola il punteggio medio pesato rispetto alla probabilità a posteriori.

$$\bar{R}_n(\theta^{(k)}) = \frac{\sum_{r=1}^m r n_r \tau(r, \theta^{(k)})}{\sum_{r=1}^m n_r \tau(r, \theta^{(k)})}$$

- **Step 4:** si aggiornano le stime dei parametri π e ξ .

$$\pi^{(k+1)} = \frac{1}{n} \sum_{r=1}^m n_r \tau(r, \theta^{(k)}); \quad \xi^{(k+1)} = \frac{m - \bar{R}_n(\theta^{(k)})}{m - 1}$$

- **Step 5:** le stime aggiornate dei parametri sono $\theta^{(k+1)} = (\pi^{(k+1)}, \xi^{(k+1)})$, con le quali si aggiorna il valore della *log-verosimiglianza*.

$$l^{(k+1)} = l_{00}(\theta^{(k+1)})$$

- **Step 6:** si verifica se effettivamente vi è stato un cambiamento significativo nelle stime rispetto al passo precedente, controllando che la differenza tra le *log-verosimiglianze* di passi successivi sia maggiore o minore di una quantità,

ϵ , fissata a priori (ad esempio, $\epsilon = 10^{-6}$).

$$\text{se } (l^{(k+1)} - l^{(k)}) \geq \epsilon \implies k \rightarrow k + 1 \implies \text{step 1}$$

$$\text{se } (l^{(k+1)} - l^{(k)}) < \epsilon \implies \hat{\theta} = \theta^{(k+1)} \implies \text{stop!}$$

Algoritmo E-M per il modello CUB(1,0)

- **Step 0:** si inizializzano i parametri e la *log-verosimiglianza*.

$$\xi^{(0)} = 1 + \frac{0.5 - M_n}{m}; \quad \beta' = (0.1, 0.1, \dots, 0.1) \in \mathbb{R}^{p+1}; \quad l^{(0)} = l_{10}(\beta'^{(0)}, \xi^{(0)})$$

- **Step 1:** si calcola il valore della Binomiale traslata.

$$b(r_i; \xi^{(k)}) = \binom{m-1}{r_i-1} (\xi^{(k)})^{m-r_i} (1 - \xi^{(k)})^{r_i-1}, \quad i = 1, \dots, n$$

- **Step 2:** si calcolano le quantità utili per l'aggiornamento dei parametri.

$$\pi_i^{(k)} = \frac{1}{1 + e^{-y_i \beta^{(k)}}};$$

$$\tau(r_i; \theta^{(k)}) = \left[1 + \frac{e^{-y_i \beta^{(k)}}}{m b(r_i; \xi^{(k)})} \right]^{-1}$$

per $i = 1, \dots, n$.

- **Step 3:** si calcola il punteggio medio pesato rispetto alla probabilità a posteriori.

$$\bar{R}_n(\theta^{(k)}) = \frac{\sum_{i=1}^n r_i \tau(r_i, \theta^{(k)})}{\sum_{i=1}^n \tau(r_i, \theta^{(k)})}$$

- **Step 4:** si calcola la funzione da massimizzare rispetto al parametro $\beta'^{(k)}$.

$$Q_1(\beta^{(k)}) = - \sum_{i=1}^n \left\{ \log \left(1 + e^{-y_i \beta^{(k)}} \right) + \left(1 - \tau(r_i, \theta^{(k)}) \right) e^{-y_i \beta^{(k)}} \right\}$$

- **Step 5:** si ottiene il valore corrispondente di $\beta^{(k)}$ che massimizza $Q_1(\beta^{(k)})$.

$$\beta^{(k+1)} = \operatorname{argmax}_{\beta} Q_1(\beta^{(k)})$$

- **Step 6:** si aggiorna la stima di $\xi^{(k)}$.

$$\xi^{(k+1)} = \frac{m - \bar{R}_n(\theta^{(k)})}{m - 1}$$

- **Step 7:** le stime aggiornate dei parametri sono $\theta^{(k+1)} = (\beta^{(k+1)}, \xi^{(k+1)})'$, con le quali si aggiorna il valore della *log-verosimiglianza*.

$$l^{(k+1)} = l_{10}(\theta^{(k+1)})$$

- **Step 8:** si verifica se effettivamente vi è stato un cambiamento significativo nelle stime rispetto al passo precedente, controllando che la differenza tra le *log-verosimiglianze* di passi successivi sia maggiore o minore di una quantità, ϵ , fissata a priori (ad esempio, $\epsilon = 10^{-6}$).

$$\text{se } (l^{(k+1)} - l^{(k)}) \geq \epsilon \quad \Longrightarrow \quad k \rightarrow k + 1 \quad \Longrightarrow \quad \text{step1}$$

$$\text{se } (l^{(k+1)} - l^{(k)}) < \epsilon \quad \Longrightarrow \quad \hat{\theta} = \theta^{(k+1)} \quad \Longrightarrow \quad \text{stop!}$$

Algoritmo E-M per il modello CUB(0,1)

- **Step 0:** si inizializzano i parametri e la *log-verosimiglianza*.

$$\pi^{(0)} = \frac{1}{2}; \quad \gamma^{(0)} = (W'W)^{-1}W'n_i \in \mathbb{R}^{q+1} \quad \text{con} \quad n_i = \log \left[\frac{m - r + \frac{1}{2}}{r - \frac{1}{2}} \right];$$

$$l^{(0)} = l_{01}(\pi^{(0)}, \gamma^{(0)})$$

- **Step 1:** si calcola il valore della Binomiale traslata

$$b(r_i; \gamma^{(k)}) = \binom{m-1}{r_i-1} \frac{e^{-(r_i-1)w_i\gamma^{(k)}}}{(1 + e^{-w_i\gamma^{(k)}})^{m-1}}; \quad i = 1, \dots, n$$

con la relazione:

$$\xi_i^{(k)} = \frac{1}{1 + e^{-w_i\gamma^{(k)}}}$$

- **Step 2:** si calcolano le quantità utili per l'aggiornamento dei parametri

$$\tau(r_i; \theta^{(k)}) = \left[1 + \frac{1 - \pi^{(k)}}{m\pi^{(k)}b(r_i; \gamma^{(k)})} \right]^{-1}$$

per $i = 1, \dots, n$.

- **Step 3:** si calcola la funzione da massimizzare rispetto al parametro $\gamma^{(k)}$

$$Q_2(\gamma^{(k)}) = - \sum_{i=1}^n \tau(r_i; \theta^{(k)}) \left\{ (r_i - 1)w_i\gamma^{(k)} + (m - 1) \log \left[1 + e^{-w_i\gamma^{(k)}} \right] \right\}$$

- **Step 4:** si ottiene il valore corrispondente di $\gamma'^{(k)}$ che massimizza $Q_2(\gamma^{(k)})$.

$$\gamma^{(k+1)} = \operatorname{argmax}_{\gamma} Q_2(\gamma^{(k)})$$

- **Step 5:** si aggiorna anche la stima del parametro di incertezza

$$\pi^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \tau(r_i; \theta^{(k)})$$

- **Step 6:** le stime aggiornate dei parametri sono $\theta^{(k+1)} = (\pi^{(k+1)}, \gamma'^{(k+1)})'$, con le quali si aggiorna il valore della *log-verosimiglianza*.

$$l^{(k+1)} = l_{01}(\theta^{(k+1)})$$

- **Step 7:** si verifica se effettivamente vi è stato un cambiamento significativo nelle stime rispetto al passo precedente, controllando che la differenza tra le *log-verosimiglianze* di passi successivi sia maggiore o minore di una quantità, ϵ , fissata a priori (ad esempio, $\epsilon = 10^{-6}$).

$$\text{se } (l^{(k+1)} - l^{(k)}) \geq \epsilon \quad \Longrightarrow \quad k \rightarrow k + 1 \quad \Longrightarrow \quad \text{step1}$$

$$\text{se } (l^{(k+1)} - l^{(k)}) < \epsilon \quad \Longrightarrow \quad \hat{\theta} = \theta^{(k+1)} \quad \Longrightarrow \quad \text{stop!}$$

Algoritmo E-M per il modello CUB(1,1)

- **Step 0:** si inizializzano i parametri e la *log-verosimiglianza*.

$$\beta^{(0)} = (0.1, \dots, 0.1) \in \mathbb{R}^{p+1};$$

$$\gamma^{(0)} = (W'W)^{-1}W'n_i \in \mathbb{R}^{q+1} \quad \text{con} \quad n_i = \log \left[\frac{m - r + \frac{1}{2}}{r - \frac{1}{2}} \right];$$

$$l^{(0)} = l_{11}(\beta^{(0)}, \gamma^{(0)})$$

- **Step 1:** si calcola il valore della Binomiale traslata

$$b(r_i; \gamma^{(k)}) = \binom{m-1}{r_i-1} \frac{e^{-(r_i-1)w_i\gamma^{(k)}}}{(1 + e^{-w_i\gamma^{(k)}})^{m-1}}; \quad i = 1, \dots, n$$

con la relazione:

$$\xi_i^{(k)} = \frac{1}{1 + e^{-w_i\gamma^{(k)}}}$$

- **Step 2:** si calcolano le quantità utili per l'aggiornamento dei parametri.

$$\pi_i^{(k)} = \frac{1}{1 + e^{-y_i\beta^{(k)}}};$$

$$\tau(r_i; \theta^{(k)}) = \left[1 + \frac{e^{-y_i\beta^{(k)}}}{mb(r_i; \xi^{(k)})} \right]^{-1}$$

per $i = 1, \dots, n$.

- **Step 3:** si calcola la funzione da massimizzare rispetto al parametro $\beta^{(k)}$.

$$Q_1(\beta^{(k)}) = - \sum_{i=1}^n \left\{ \log \left(1 + e^{-y_i \beta^{(k)}} \right) + \left(1 - \tau(r_i, \theta^{(k)}) \right) e^{-y_i \beta^{(k)}} \right\}$$

- **Step 4:** si calcola la funzione da massimizzare rispetto al parametro $\gamma^{(k)}$

$$Q_2(\gamma^{(k)}) = - \sum_{i=1}^n \tau(r_i; \theta^{(k)}) \left\{ (r_i - 1) w_i \gamma^{(k)} + (m - 1) \log \left[1 + e^{-w_i \gamma^{(k)}} \right] \right\}$$

- **Step 5:** si aggiornano i valori di $\beta^{(k)}$ e $\gamma^{(k)}$ individuando quelli che massimizzano rispettivamente $Q_1(\beta^{(k)})$ e $Q_2(\gamma^{(k)})$.

$$\beta^{(k+1)} = \operatorname{argmax}_{\beta} Q_1(\beta^{(k)}) \quad \gamma^{(k+1)} = \operatorname{argmax}_{\gamma} Q_2(\gamma^{(k)})$$

- **Step 6:** le stime aggiornate dei parametri sono $\theta^{(k+1)} = (\beta^{(k+1)}, \gamma^{(k+1)})'$, con le quali si aggiorna il valore della *log-verosimiglianza*.

$$l^{(k+1)} = l_{11}(\theta^{(k+1)})$$

- **Step 7:** si verifica se effettivamente vi è stato un cambiamento significativo nelle stime rispetto al passo precedente, controllando che la differenza tra le *log-verosimiglianze* di passi successivi sia maggiore o minore di una quantità, ϵ , fissata a priori (ad esempio, $\epsilon = 10^{-6}$).

$$\text{se } (l^{(k+1)} - l^{(k)}) \geq \epsilon \quad \Longrightarrow \quad k \rightarrow k + 1 \quad \Longrightarrow \quad \text{step1}$$

$$\text{se } (l^{(k+1)} - l^{(k)}) < \epsilon \quad \Longrightarrow \quad \hat{\theta} = \theta^{(k+1)} \quad \Longrightarrow \quad \text{stop!}$$

1.4 Matrice di Informazione osservata $\mathcal{I}(\hat{\theta})$

La matrice di informazione osservata $\mathcal{I}(\hat{\theta})$ viene ottenuta calcolando esplicitamente la derivata seconda della *log-verosimiglianza*. Anche in questo contesto si distinguono i calcoli a seconda del numero di covariate presenti per le componenti di incertezza e di *feeling* (Piccolo, 2006).

Matrice di informazione per il modello CUB(0,0)

Dopo diversi passaggi che coinvolgono tecniche di algebra lineare, si trovano i valori di ciascun elemento della matrice di informazione, nel caso di un modello senza covariate:

$$\begin{aligned}\frac{\partial^2 l(\theta)}{\partial \pi^2} &= -\frac{1}{\pi^2} \sum_{i=1}^n (1 - q_i)^2; \\ \frac{\partial^2 l(\theta)}{\partial \pi \partial \xi} &= \frac{1}{\pi} \sum_{i=1}^n v_i q_i q_i^*; \\ \frac{\partial^2 l(\theta)}{\partial \xi^2} &= -\sum_{i=1}^n q_i^* [u_i - (1 - q_i^*) v_i^2];\end{aligned}$$

dove le quantità v_i , u_i , q_i e q_i^* , per $i = 1, \dots, n$, sono definite:

$$\begin{aligned}v_i &= \frac{m - r_i}{\xi} - \frac{r_i - 1}{1 - \xi}; & u_i &= \frac{m - r_i}{\xi^2} + \frac{r_i - 1}{(1 - \xi)^2}; \\ q_i &= \frac{1}{m p_i(\theta)}; & q_i^* &= 1 - (1 - \pi) q_i.\end{aligned}$$

Con $p_i(\theta)$ si intende la distribuzione di probabilità $p_i(\theta) = P[R = r_i | \pi, \xi]$ definita nel precedente capitolo.

Si ottiene la matrice di varianza e covarianza asintotica $V(\theta)$ delle stime ML di θ , calcolata per $\theta = \hat{\theta} = (\hat{\pi}, \hat{\xi})$:

$$V(\theta) = [\mathcal{I}(\hat{\theta})]^{-1} = \begin{bmatrix} \frac{\partial^2 l(\theta)}{\partial \pi^2} & \frac{\partial^2 l(\theta)}{\partial \pi \partial \xi} \\ \frac{\partial^2 l(\theta)}{\partial \pi \partial \xi} & \frac{\partial^2 l(\theta)}{\partial \xi^2} \end{bmatrix}_{\theta=\hat{\theta}}^{-1}$$

Matrice di informazione per il modello CUB(1,1)

In questo contesto è meglio introdurre delle notazioni per semplificare la stesura dei calcoli per ottenere la derivata seconda della *log-verosimiglianza*. Definiamo:

$$k_i = \binom{m-1}{r-1}; \quad b_i = e^{-y_i\beta}; \quad c_i = e^{-w_i\gamma};$$

$$E_i(\beta) = \frac{1}{1+b_i}; \quad B_i(\gamma) = k_i \frac{(c_i)^{r_i-1}}{(1+c_i)^{m-1}}$$

Dunque, si possono riscrivere sia la distribuzione di probabilità che la funzione di *log-verosimiglianza* utilizzando questa notazione:

$$p_i(\theta) = E_i(\beta) \left\{ B_i(\gamma) - \frac{1}{m} \right\} + \frac{1}{m}$$

$$l(\theta) = \sum_{i=1}^n \log [p_i(\theta)]$$

Inoltre, si definiscono anche le seguenti quantità:

$$F_i(\gamma) = \frac{1}{1+c_i}; \quad a_i(\gamma) = (r_i-1) - (m-1)(1-F_i);$$

$$\tilde{E}_i = E_i(1-E_i); \quad \tilde{F}_i = F_i(1-F_i);$$

$$q_i = \frac{1}{mp_i(\theta)}; \quad q_i^* = 1 - (1-E_i)q_i; \quad \tilde{Q}_i = q_i^*(1-q_i^*)$$

I valori della matrice di informazione osservata sono espressi da:

$$\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \beta_t} = - \sum_{i=1}^n y_{is} y_{it} \left\{ \tilde{E}_i - \tilde{Q}_i \right\};$$

$$\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \gamma_t} = - \sum_{i=1}^n y_{is} w_{it} a_i \tilde{Q}_i;$$

$$\frac{\partial^2 l(\theta)}{\partial \gamma_s \partial \gamma_t} = - \sum_{i=1}^n w_{is} w_{it} \left\{ (m-1) q_i^* \tilde{F}_i - a_i^2 \tilde{Q}_i \right\}.$$

Matrice di informazione per il modello CUB(1,0) e CUB(0,1)

Il risultato generale ottenuto per il modello CUB(1,1) richiede delle modifiche quando solo π o ξ sono funzione di covariate. Nella Tabella 1.2 sono presenti tutte le notazioni utili alla definizione della matrice di osservazione, $\mathcal{I}(\hat{\theta})$, per i due modelli.

Simbolo	CUB(1,0)	CUB(0,1)
$p_i(\theta)$	$E_i(\beta \{B_i(\xi) - \frac{1}{m}\} + \frac{1}{m})$	$\pi \{B_i(\gamma) - \frac{1}{m}\} + \frac{1}{m}$
$B_i(\gamma)$	$B_i(\xi)$	$\binom{m-1}{r-1} \frac{(e^{-w_i \gamma})^{r_i-1}}{(1+e^{-w_i \gamma})^{m-1}}$
$E_i(\beta)$	$\frac{1}{1+e^{-y_i \beta}}$	π
$F_i(\gamma)$	ξ	$\frac{1}{1+e^{-w_i \gamma}}$
q_i	$\frac{1}{m p_i(\beta, \xi)}$	$\frac{1}{m p_i(\pi, \gamma)}$
q_i^*	$1 - q_i [1 - E_i(\beta)]$	$1 - q_i [1 - \pi]$

Tabella 1.2: Notazione per la costruzione della matrice di Informazione per modelli CUB(1,0) e CUB(0,1).

Per il primo modello, CUB(1,0), le componenti della matrice di Informazione sono:

$$\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \beta_t} = \sum_{i=1}^n y_{is} y_{it} \left\{ \tilde{E}_i - \tilde{Q}_i \right\};$$

$$\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \xi} = - \sum_{i=1}^n y_{is} v_i \tilde{Q}_i;$$

$$\frac{\partial^2 l(\theta)}{\partial \xi \partial \beta_t} = - \sum_{i=1}^n y_{it} v_i \tilde{Q}_i;$$

$$\frac{\partial^2 l(\theta)}{\partial \xi^2} = \sum_{i=1}^n \left\{ u_i q_i^* - v_i^2 \tilde{Q}_i \right\}.$$

La matrice risulta:

$$\mathcal{I}(\hat{\theta}) = \begin{bmatrix} \left[\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \beta_t} \right]_{(p+1,p+1)} & \left[\frac{\partial^2 l(\theta)}{\partial \beta_s \partial \xi} \right]_{(p+1,1)} \\ \left[\frac{\partial^2 l(\theta)}{\partial \xi \partial \beta_t} \right]_{(1,p+1)} & \left[\frac{\partial^2 l(\theta)}{\partial \xi^2} \right]_{(1,1)} \end{bmatrix}_{\theta=\hat{\theta}}$$

Per il secondo modello, CUB(0,1), le componenti della matrice di Informazione sono:

$$\frac{\partial^2 l(\theta)}{\partial \pi^2} = \frac{1}{\pi^2} \sum_{i=1}^n (1 - q_i)^2;$$

$$\frac{\partial^2 l(\theta)}{\partial \pi \partial \gamma_t} = \frac{1}{\pi} \sum_{i=1}^n w_{it} a_i q_i q_i^*;$$

$$\frac{\partial^2 l(\theta)}{\partial \gamma_s \partial \pi} = \frac{1}{\pi} \sum_{i=1}^n w_{is} a_i q_i q_i^*;$$

$$\frac{\partial^2 l(\theta)}{\partial \gamma_s \partial \gamma_t} = \sum_{i=1}^n w_{is} w_{it} \left\{ (m-1) q_i^* \tilde{F}_i - a_i^2 \tilde{Q}_i \right\}$$

La matrice risulta:

$$\mathcal{I}(\hat{\theta}) = \begin{bmatrix} \left[\frac{\partial^2 l(\theta)}{\partial \pi^2} \right]_{(1,1)} & \left[\frac{\partial^2 l(\theta)}{\partial \pi \partial \gamma_t} \right]_{(1,q+1)} \\ \left[\frac{\partial^2 l(\theta)}{\partial \gamma_s \partial \pi} \right]_{(q+1,1)} & \left[\frac{\partial^2 l(\theta)}{\partial \gamma_s \partial \gamma_t} \right]_{(q+1,q+1)} \end{bmatrix}_{\theta=\hat{\theta}}$$

1.5 Aspetti inferenziali

Considerando il caso più generale di un modello CUB(p,q) su un campione di dimensione n che misura sia i giudizi degli intervistati attraverso una variabile ordinale sia un set di covariate associate alle componenti di *feeling* e incertezza, la funzione di log-verosimiglianza per il vettore dei parametri $\theta = (\beta', \gamma)'$ è definita come:

$$l(\theta) = \sum_{i=1}^n \log \left[\frac{1}{1 + e^{-y_i \beta}} \left[\binom{m-1}{r_i-1} \frac{(e^{-w_i \gamma})^{r_i-1}}{(1 + e^{-w_i \gamma})^{m-1}} - \frac{1}{m} \right] + \frac{1}{m} \right]$$

Nelle sezioni precedenti si è mostrato come le stime dei parametri θ vengano ottenute attraverso la stima di massima verosimiglianza (ML) tramite l'algoritmo E-M. Inoltre, per ottenere la matrice di informazione, $\mathcal{I}(\hat{\theta})$, e di conseguenza la matrice di varianza e covarianza, $V(\theta)$, deve essere calcolata la derivata di ordine due di $l(\theta)$ attraverso metodi analitici.

Una volta effettuate le stime dei parametri del modello CUB attraverso la verosimiglianza, un altro aspetto cruciale è la validazione di tali stime (Iannario e Piccolo, 2012). Tutti gli aspetti inferenziali sui modelli CUB sono specificati in Piccolo (2006). Questa convalida può essere effettuata attraverso diversi approcci:

- **significatività dei parametri:** attraverso il test di Wald le stime sono confrontate con il loro *standard error* per investigare quali di esse sono realmente significative e apportano informazione utile per la rappresentazione del fenomeno di interesse.
- **confronto delle log-verosimiglianze:** quando i modelli sono annidati, è possibile testare se l'aumento della log-verosimiglianza è significativo. Ciò viene realizzato attraverso la devianza e il corrispondente percentile della distribuzione χ_g^2 (i gradi di libertà g derivano dalla differenza del numero di parametri nei due modelli considerati), verificando se il modello più complesso è una scelta migliore o meno (Agresti 2010). Per stabilire la rilevanza statisti-

ca delle covariate nello spiegare i dati, si confronta la log-verosimiglianza del modello stimato senza covariate con le log-verosimiglianze dei modelli con p e q covariate, rispettivamente, per l'incertezza e il feeling come in Tabella 1.3.

Confronto	Test LR	Gradi di libertà
$CUB(p, 0)$ vs $CUB(0, 0)$	$2(l_{10} - l_{00})$	p
$CUB(0, q)$ vs $CUB(0, 0)$	$2(l_{01} - l_{00})$	q
$CUB(p, q)$ vs $CUB(0, 0)$	$2(l_{11} - l_{00})$	p+q

Tabella 1.3: Confronto tra modelli tramite test di log-rapporto di verosimiglianza e gradi di libertà associati.

- **indici:** l'indice di dissimilarità è utilizzato per valutare l'adeguatezza del modello. Esso confronta le frequenze osservate, f_r , con le probabilità attese $\hat{p}_r = p_r(\hat{\theta})$:

$$\Psi^2 = 1 - \frac{1}{2} \sum_{r=1}^m |f_r - \hat{p}_r|$$

Valori di Ψ^2 maggiori di 0.90 sono compatibili con la validazione del modello utilizzato. Altri due indici per valutare l'adeguamento del modello, e che possono essere usati anche nel caso di modelli non annidati, sono l'indice *BIC* (Schwarz, 1978) e l'indice *AIC* (Akaike, 1974). Essi permettono di misurare il miglioramento in termini di verosimiglianza, considerando anche la penalizzazione data dal numero di parametri ($npar$) introdotti nel modello.

$$BIC = -2l(\hat{\theta}) + (npar) \log n, \quad AIC = -2l(\hat{\theta}) + 2(npar)$$

A valori maggiori di questi indici corrispondono migliori modelli. In questo contesto dei modelli CUB, l'indice BIC è quello preferito, come visto da Iannario e Piccolo (2009).

- **analisi dei residui:** è possibile effettuare una classica analisi dei residui considerando per esempio i residui di Pearson o altre analisi basate sui residui generalizzati come in Di Iorio e Iannario (2012).

1.6 Applicazioni dei modelli CUB

In letteratura sono presenti diverse applicazioni dei modelli CUB nei più svariati ambiti. Tuttavia, anziché riportare una mera lista di lavori che utilizzano questa metodologia, si preferisce mostrare in seguito tre esempi di applicazioni dei modelli CUB che dovrebbero essere esaustivi per evidenziare la potenzialità di questa soluzione: il primo esempio rappresenta una guida per l'utilizzo del pacchetto del software R; il secondo, tratto da Capecchi et al. (2012), e il terzo (Iannario, 2008) rappresentano un esempio di applicazione *rating* e *ranking* rispettivamente. Le tematiche trattate in questi lavori riguardano:

1. La valutazione del servizio d'orientamento universitario ($m = 7$)
2. La valutazione della soddisfazione lavorativa dei laureati ($m = 10$)
3. L'emergenze metropolitane di Napoli ($m = 9$)

Nel primo caso si è valutata l'iniziativa VER che costituisce uno strumento di partecipazione attiva della popolazione studentesca alle attività universitarie di supporto. La valutazione è espressa attraverso un'indagine che la struttura SOF-Tel (Servizio di Ateneo per l'Orientamento, la Formazione e la Teledidattica) effettua su se stessa, per confrontare obiettivi e percezione dei risultati sulla base dei giudizi formulati dagli utenti/studenti, offrendo al governo dell'Ateneo le indicazioni necessarie per la programmazione del cambiamento organizzativo e la trasformazione delle attività di orientamento allo scopo di aumentarne l'efficacia. I dati analizzati, presenti nel pacchetto "CUB" del software R, sono relativi all'anno 2002 ad un'indagine condotta in tutte le Facoltà dell'Università di Napoli Federico II. Ai

partecipanti è stato chiesto di esprimere le proprie valutazioni su una scala di sette punti (1 = "molto insoddisfatto", ... ,7 = "molto soddisfatto") sui seguenti item:

- *informat*: livello di soddisfazione complessivo rispetto alle informazioni ottenute
- *willingn*: livello di soddisfazione per la disponibilità del personale
- *officeho*: giudizio sull'orario di apertura degli sportelli
- *competen*: giudizio sulla competenza del personale
- *global*: soddisfazione globale

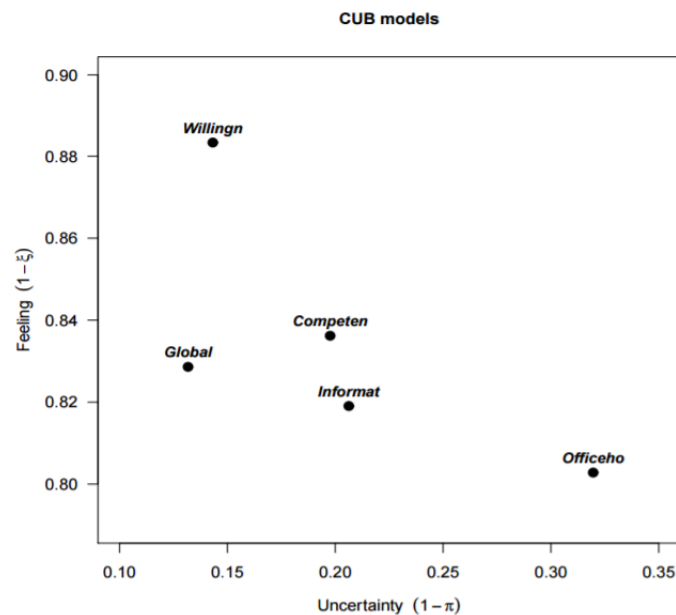


Figura 1.1: Proiezione degli item nello spazio parametrico mediante la funzione *multicub()*.

Nella Figura 1.1 sono proiettati gli item di interesse nello spazio parametrico $\Omega(\theta)$, con coordinate $(1 - \pi, 1 - \xi)$, così da confrontare più modelli contemporaneamente. Per esempio, la valutazione degli studenti verso l'orario d'apertura degli sportelli (*officeho*) è soggetta a maggior incertezza rispetto a tutte le altre. Invece, il livello di soddisfazione per la disponibilità del personale (*willingn*) è il più alto, a rappresentare un ottimo giudizio degli studenti a riguardo.

Consideriamo ora solo la valutazione relativa all'item *officeho* per implementare un modello CUB senza covariate. I risultati delle stime dei parametri di *feeling* e incertezza sono riportati in Tabella 1.4.

Parametro	Stima	Std
π	0.68019	0.01935
ξ	0.19714	0.00588

Tabella 1.4: Stima di un modello CUB senza covariate per *officeho*.

Per migliorare il modello è possibile includere covariate su una o entrambe le componenti. Per la componente di *feeling* si introducono l'età e la frequenza di utilizzo del servizio (0 ="utilizzatori non regolari" , 1 ="utilizzatori regolari"), mentre per l'incertezza si considerano l'età e il genere dello studente. Una volta ottenute le stime dei coefficienti dal modello CUB(p,q), si calcolano i parametri di *feeling* e incertezza attraverso la trasformazione logistica. Nella Figura 1.2 è riportato l'andamento delle stime dei parametri in base alle covariate selezionate. In rosso sono rappresentate le donne utilizzatrici regolari del servizio, che valutano più positivamente, e con un minor livello di incertezza nel rispondere, l'orario degli sportelli, rispetto agli uomini utilizzatori meno frequenti del servizio (curva blu). Inoltre, all'aumentare dell'età degli studenti il livello di incertezza aumenta, mentre diminuisce la componente di *feeling*.

Per quanto riguarda il secondo esempio di applicazione dei modelli CUB, il focus è diretto all'analisi della soddisfazione lavorativa di un sottoinsieme di laureati italiani. I dati sono relativi a due indagini dell'archivio di AlmaLaurea:

- Indagine 2010, sui laureati del 2005 pre-riforma
- Indagine 2012, sui laureati del 2007 post-riforma.

La domanda generale riguarda la soddisfazione globale ed è accompagnata da ulteriori 14 quesiti. Le risposte sono espresse su una scala ordinale da 1 a 10, dove 1 significa "totalmente insoddisfatto" e 10 significa "completamente soddisfatto".

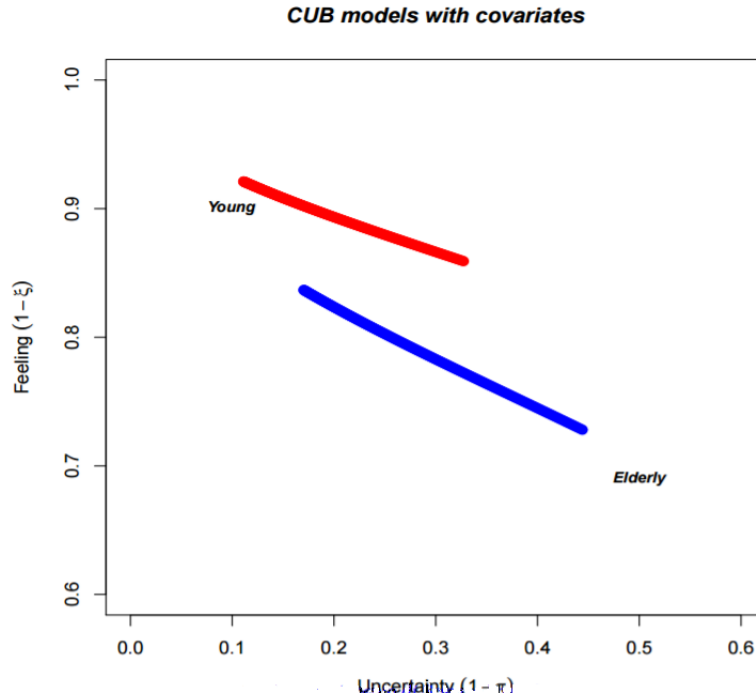


Figura 1.2: Stima di un modello CUB con covariate.

In Figura 1.3 è riportato il risultato della stima di un modello CUB senza covariate applicato alla soddisfazione lavorativa dei laureati pre e post riforma, dal quale si evince che i laureati pre-riforma sono tendenzialmente meno soddisfatti di quelli laureati post-riforma, ma meno incerti nelle valutazioni. Dalla Figura 1.4, invece, si può notare come tale soddisfazione lavorativa sia fortemente influenzata dal guadagno degli studenti nella soluzione lavorativa trovata dopo la Laurea. Difatti, per i laureati con più alti guadagni la valutazione è alta e con un basso livello di incertezza, viceversa per coloro che hanno minori guadagni, la soddisfazione tende a diminuire e aumenta l'incertezza nelle risposte. Inoltre, la distinzione tra i due gruppi di laureati, pre e post riforma, aumenta al diminuire dei guadagni.

Una volta stimato un modello CUB con covariate per entrambe le componenti è possibile costruire dei profili per i quali calcolare le probabilità di osservare le diverse categorie. Alcuni dei possibili profili che si possono costruire in questo contesto sono riportati in Tabella 1.5.

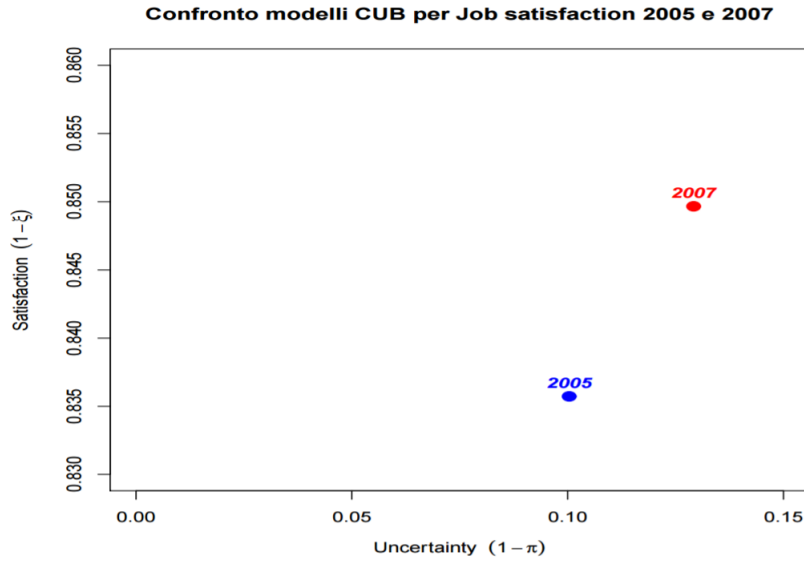


Figura 1.3: Stima di un modello CUB sulla soddisfazione lavorativa globale.

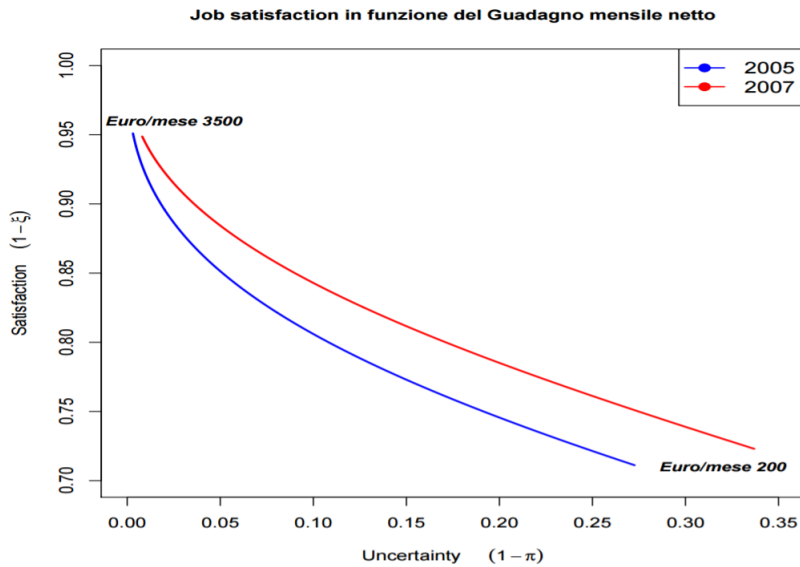


Figura 1.4: Stima di un modello CUB sulla soddisfazione lavorativa in funzione del guadagno mensile netto.

In base ai grafici di Figura 1.5 si conclude che i profili B, E, F e G che hanno problemi riguardo ai guadagni sono quelli più insoddisfatti. Coloro che invece hanno una retribuzione media o alta si ritengono abbastanza soddisfatti, sia che essi siano uomini o donne o più o meno giovani.

	Genere	Cerca lavoro	Reddito	Età	Laurea
A: Donna giovane con lavoro stabile e ben retribuita	1	0	2500	23	
B: Uomo giovane in cerca lavoro, ora mal retribuito	0	1	500	23	
C: Donna di età media in cerca lavoro con medio guadagno	1	1	1500	30	
D: Uomo di età media in cerca lavoro con medio guadagno	0	1	1500	30	
E: Donna di età matura con guadagno basso	1	0	1000	40	
F: Uomo di età elevata con guadagno basso	0	0	1000	50	
G: Donna di età media con lavoro mal retribuito	1	0	500	30	
H: Uomo di età matura in cerca lavoro con guadagno elevato	0	1	2500	40	
I: Uomo in cerca lavoro con età e guadagno medi	0	1	1360	27	

Tabella 1.5: Profili costruiti in base alle caratteristiche del soggetto introdotte nel modello.

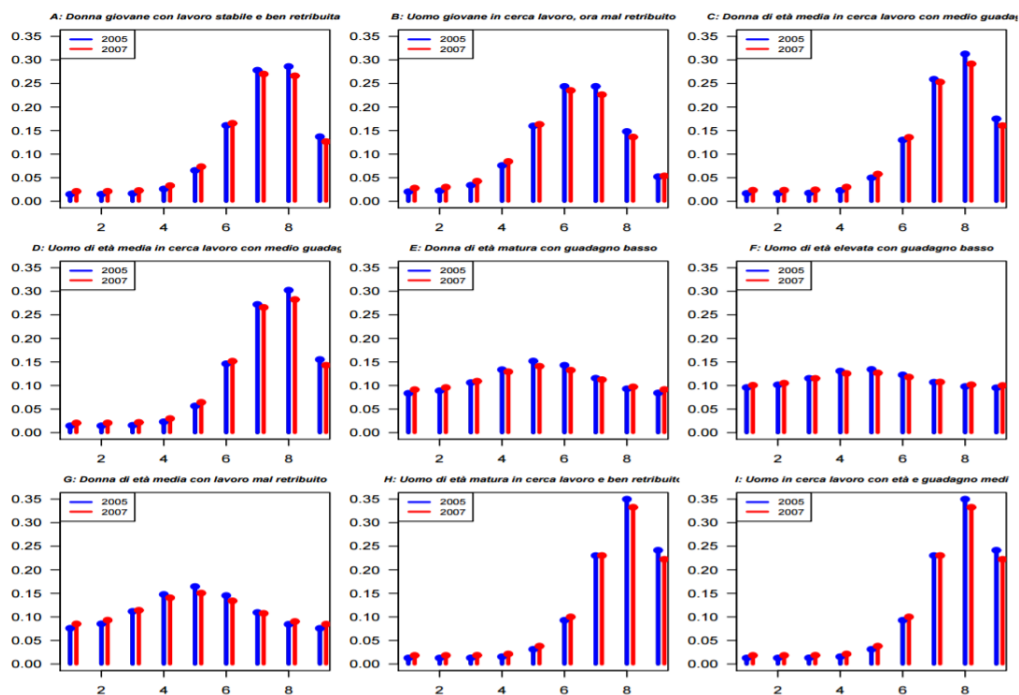


Figura 1.5: Probabilità stimate dal modello CUB per ciascun profilo.

Nel terzo esempio di applicazione dei modelli CUB vengono utilizzate delle risposte di tipo *ranking*. Infatti, è stato chiesto ad un gruppo di $n = 354$ soggetti residenti nell'area metropolitana di Napoli di mettere in ordine le principali emergenze metropolitane della città tra le seguenti:

- Clientelismo e corruzione
- Criminalità organizzata
- Disoccupazione e mancanza di lavoro

- Inquinamento ambientale e carenza di verde
- Mala-sanità
- Micro-criminalità
- Presenza di immigrati
- Pulizia ed igiene del territorio
- Trasporti-traffico-mobilità

È stato chiesto agli intervistati di indicare con il rango 1 il problema ritenuto più grave, e via via crescendo fino ad indicare con rango 9 l'emergenza ritenuta comparativamente meno grave. In questo studio, il parametro relativo al *feeling* esprime il grado di preoccupazione per un problema. Inoltre, sono state raccolte numerose altre informazioni sui soggetti (Genere, Età, Livello di istruzione, Residenza, Condizione occupazionale, etc.)

La distribuzione delle frequenze assolute delle nove emergenze metropolitane è mostrata in Figura 1.6.

Emergenze	Ranghi →								
	1	2	3	4	5	6	7	8	9
<i>Clientelismo e corruzione</i>	21	43	88	78	55	26	19	14	10
<i>Criminalità organizzata</i>	212	94	26	12	5	4	0	0	1
<i>Disoccupazione e mancanza di lavoro</i>	88	84	84	44	26	15	4	6	3
<i>Inquinamento ambientale e carenza di verde</i>	4	5	9	21	49	64	95	70	37
<i>Mala-sanità</i>	2	12	38	78	101	57	31	24	11
<i>Micro-criminalità</i>	20	97	85	72	39	20	12	8	1
<i>Presenza di immigrati</i>	3	4	3	11	11	29	29	70	194
<i>Pulizia ed igiene del territorio</i>	2	7	7	27	40	92	96	62	21
<i>Trasporti-traffico-mobilità</i>	2	8	14	11	28	47	68	100	76

Figura 1.6: Distribuzione delle frequenze delle valutazioni di tipo *ranking*.

Il problema ritenuto più grave dai rispondenti al questionario è stato quello della criminalità organizzata, a seguire la disoccupazione e la micro-criminalità. I grafici della Figura 1.7 mostrano le probabilità assegnate ad ogni categoria tramite la metodologia CUB per ciascun item di valutazione.

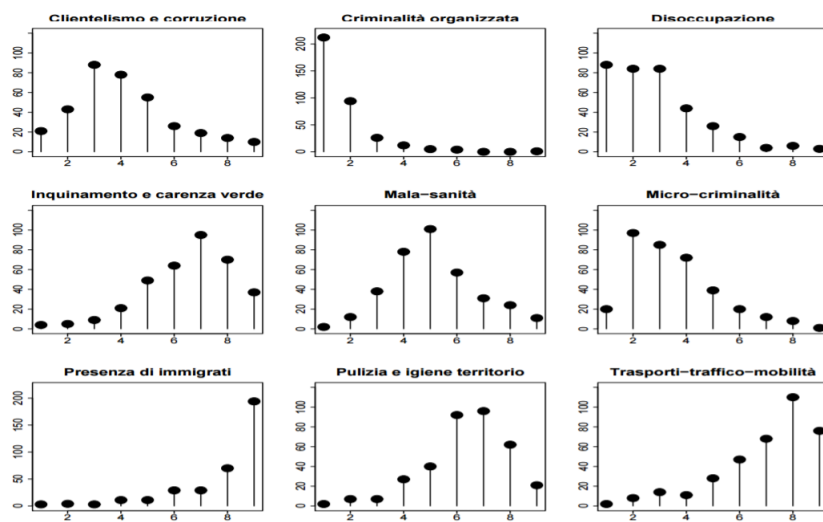


Figura 1.7: Probabilità stimate tramite un modello CUB con dati di tipo *ranking*.

Capitolo 2

Anchoring Vignettes

Nel primo capitolo i modelli CUB hanno mostrato un valido approccio all'analisi dei dati ordinali. Tuttavia, se le indagini coinvolgessero popolazioni con caratteristiche molto diverse tra loro, si renderebbe meno immediato il confronto dei risultati. È in questo ambito che si sviluppano i modelli che fanno uso delle Anchoring Vignettes (King et al., 2004). Le eterogeneità interpersonali e interculturali nell'interpretare, capire o utilizzare categorie di risposta per una stessa domanda, sono noti in psicommetria come DIF (Differential Item Functioning) (Paccagnella, 2011). L'approccio con le Anchoring Vignettes consiste nell'utilizzare una serie di scenari (chiamati appunto vignettes), uguali per tutti gli intervistati, in modo da poter stimare il DIF e correggere i risultati dell'indagine (Green e Hensher, 2010). Le vignettes sono dunque domande supplementari per esaminare gli intervistati, dove si descrive un ipotetico scenario, in cui una ipotetica persona è descritta in particolari situazioni o condizioni. Gli intervistati sono poi invitati a valutare questi ipotetici individui, utilizzando la stessa scala adottata per valutare se stessi.

Quando si usano le vignettes esistono alcune importanti scelte da affrontare: *(i)* quali caratteristiche dell'individuo ipotetico utilizzare; *(ii)* il numero di vignettes da considerare; *(iii)* il loro ordine di esposizione. Per quanto riguarda il primo aspetto, dato che il rispondente dovrebbe considerare la persona descritta come sé stesso, il

nome dell'individuo ipotetico dovrebbe essere appropriato con la sua lingua e la sua cultura. In generale, è preferibile che il nome scelto da usare nella vignettes cambi in base al genere del rispondente, ma essendo questa soluzione troppo costosa, si potrebbe utilizzare un nome di genere neutro o riferirsi solo all'iniziale del nome. In relazione al secondo aspetto da considerare per la costruzione delle vignettes, ossia il numero di esse da proporre al rispondente, la scelta consiste nel *trade-off* tra la riduzione del DIF e i costi di campionamento. Da un punto di vista metodologico, per l'applicazione di questo metodo, una sola vignette sarebbe sufficiente (come sarà più chiaro nel momento in cui verranno introdotte sia la soluzione parametrica sia quella non parametrica che sfruttano le informazioni raccolte mediante le vignettes). D'altra parte, in King et al. (2004), è consigliato il ricorso a due o tre vignettes per ottenere i massimi benefici da questa tecnica. Sempre nello stesso lavoro di King et al. (2004) si trova che è preferibile presentare le vignettes al rispondente in un ordine completamente casuale.

Le indagini campionarie possono rappresentare una valida fonte di informazione sul benessere degli individui sotto forma di valutazioni soggettive riguardanti, ad esempio, la *life satisfaction*, la *job satisfaction*, la salute personale, la *political efficacy*, la qualità dei servizi pubblici, ecc. Mentre questi dati vengono sempre più utilizzati in modo produttivo negli studi sia economici che delle scienze sociali, è nata una persistente preoccupazione che il DIF impedisca la comparabilità interpersonale delle misure soggettive. La soluzione proposta a questo problema di errore di misura è quella di ancorare (Anchoring) la valutazione individuale di una persona alla valutazione di una vignette (King et al, 2004). Dato che quest'ultima è fissa per tutti gli individui, la variazione del suo rating identifica l'eterogeneità osservata purificandola dalla valutazione soggettiva dell'individuo.

Un esempio di un'applicazione delle Anchoring Vignettes è mostrato da un'indagine dell'Organizzazione Mondiale della Sanità che, nel 2002, chiese ad un campione

di cittadini cinesi e messicani di giudicare la propria *political efficacy*¹ in una scala ordinale che andava da “nessuna” a “illimitata”. I primi risultati furono sorprendenti: il 50% dei messicani rispose “nessuna”, contro il 27% dei cinesi come mostrato nel grafico di sinistra di Figura 2.1. Questi risultati porterebbero alla conclusione, dopo un’indagine approssimativa, che gli intervistati messicani abbiano una minore *political efficacy* rispetto ai cittadini cinesi, risultando alquanto strano. Rifacendo l’analisi considerando l’approccio con le Anchoring Vignettes, le conclusioni vengono completamente stravolte: il 40% dei cinesi, infatti, reputa la propria condizione peggiore della peggior vignette, mentre ciò accade solo nell’11% dei messicani come mostrato nel grafico di destra in Figura 2.1 (Chevalier e Fielding, 2011). Questo esempio aiuta a comprendere quanto possano essere importanti le Anchoring Vignettes per confrontare in modo corretto i risultati di indagini su popolazioni estremamente diverse.

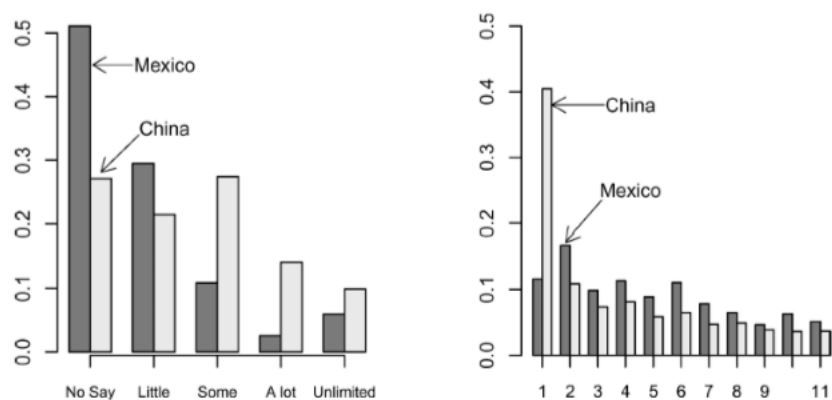


Figura 2.1: Analisi sulla *political efficacy* di Cina e Messico: a sinistra la valutazione originale e a destra la valutazione con l’utilizzo delle Anchoring Vignettes.

¹Per *political efficacy* si intende la fiducia che un cittadino ha nel suo governo e la sua aspettativa di poter capire ed influenzare le decisioni politiche

2.1 Assunzioni

Le Anchoring Vignettes si basano su assunzioni identificative senza le quali l'approccio sarebbe incapace di operare la correzione del DIF. Dunque, per poter utilizzare questa soluzione sono necessarie due assunzioni di base: la prima è chiamata **Response Consistency (RC)**, nella quale si assume che ogni individuo utilizzi le stesse scale di risposta sia valutando la propria situazione, sia valutando quella descritta dalla vignette. Questo significa che può esistere del DIF tra gli intervistati, ma che non ne può esistere tra le risposte di uno stesso individuo (King et al., 2004). Se ciò non fosse verificato, allora le informazioni provenienti dalle risposte della vignette non sarebbero utili per identificare e correggere l'eterogeneità dei dati. Questa assunzione non potrà essere valida se, per esempio, gli individui vogliono riferire in modo errato la propria salute: un esempio tipico di questo ambito è la cosiddetta "*Justification bias*" (Bound, 1991) ossia, fissato il livello della "vera salute", gli individui che non lavorano, in quanto disoccupati, sottostimano la propria salute al fine di giustificare la loro condizione. La seconda assunzione è la **Vignette Equivalence (VE)** che indica che il livello della variabile presentato in ogni vignette viene percepito nella stessa maniera da tutti gli individui. In altri termini, gli intervistati possono avere delle differenze nella percezione del livello della variabile tra i diversi scenari, ma che queste devono essere casuali e indipendenti dalle caratteristiche della vignette stessa (King et al., 2004). L'assunzione di VE può essere violata se le descrizioni delle vignettes sono incomplete e/o equivoche e gruppi di individui interpretano queste descrizioni in modi diversi. Questa assunzione è quella più controversa e rappresenta la critica più forte all'approccio con le vignettes. Difatti, si potranno avere diverse interpretazioni per lo stesso scenario più o meno problematiche a seconda della religione o dello status socio-economico dei Paesi o per altre forti caratterizzazioni della popolazione. Per esempio, si potranno avere violazioni di questa assunzione nel caso in cui si discuta del tema del suicidio nelle popolazioni cattoliche o nelle popolazioni non cattoliche; oppure considerare una

certa quantità di guadagni o di sussidi di disoccupazione nei Paesi con uno stato sociale più sviluppato rispetto a Paesi meno avanzati.

La validità di queste due assunzioni è tuttora oggetto di dibattito. Nel seguito verranno presentati degli studi in proposito, mentre adesso verranno affrontate le tecniche statistiche capaci di implementare l'approccio delle Anchoring Vignettes sia in un ambito non parametrico che in uno parametrico.

2.2 Metodo non parametrico

La soluzione non parametrica non richiede la specificazione di alcuna variabile esplicativa ed assume solo la RC e VE, senza aggiungere ulteriori assunzioni o condizioni. L'idea è quella di ricodificare l'autovalutazione sulla base delle risposte alle vignettes: per ogni rispondente i ($i = 1, \dots, N$), si definisce Y_i la risposta categorica all'autovalutazione e Z_{ij} la valutazione categorica alla vignette j ($j = 1, \dots, J$). In questo contesto è necessario che le situazioni rappresentate nelle vignettes abbiano un ordine naturale dato dalla gravità del problema di interesse e tutti i rispondenti di ciascun Paese/gruppo percepiscano lo stesso ordine : $Z_{i,j-1} < Z_{ij}$ per ogni i e j . Successivamente, le autovalutazioni di ogni rispondente sono riscalate ottenendo la seguente variabile libera dal DIF:

$$C_i = \begin{cases} 1 & \text{se } Y_i < Z_{i1} \\ 2 & \text{se } Y_i = Z_{i1} \\ 3 & \text{se } Z_{i1} < Y_i < Z_{i2} \\ \vdots & \vdots \\ 2J & \text{se } Y_i = Z_{iJ} \\ 2J + 1 & \text{se } Y_i > Z_{iJ} \end{cases}$$

La variabile C_i è stata corretta dal DIF e può essere analizzata come qualsiasi

altra variabile ordinale, ad esempio con un modello *Ordered probit*.

Questo metodo funziona fino a che tutti i rispondenti valutano le vignettes con il loro ordine naturale, ma ciò non sempre accade. Infatti, può succedere che le vignettes siano percepite con un ordine differente o che lo stesso punteggio sia dato a più vignettes; in questo caso queste osservazioni sarebbero trattate separatamente perdendo quindi dell'informazione e rendendo la soluzione non parametrica inefficiente. Inoltre, il metodo non parametrico soffre di un altro svantaggio: richiede infatti che ciascun individuo risponda a tutte le vignettes, comportando quindi alti costi.

2.3 Metodo parametrico

Il modello che permette di utilizzare l'approccio con le Anchoring Vignettes è chiamato *Chopit* (*Compound Hierarchical Ordered Probit*) e rappresenta una generalizzazione del modello per dati ordinali, dove il DIF viene modellato tramite la variazione delle soglie e le vignettes forniscono l'informazione chiave per le stime dei parametri.

2.3.1 Modello *Ordered Probit*

Il modello *Ordered Probit* appartiene alla classe dei modelli con variabile risposta qualitativa ordinale. L'idea centrale per la sua definizione è pensare ad una variabile latente continua, Y^* , che sta alla base del processo decisionale e che viene tradotta dal rispondente in una variabile categoriale Y . Esistono delle soglie che dividono la variabile continua latente in modo da ottenere quella categoriale. Per ogni individuo i ($i = 1, \dots, N$) la variabile latente è una combinazione lineare di alcuni predittori, X_i , più un termine d'errore normalmente distribuito:

$$Y_i^* = \alpha + X_i' \beta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$

dove α è l'intercetta e β il vettore dei parametri associati ad X . Supponendo ad esempio di avere cinque categorie:

$$\left\{ \begin{array}{l} Y_i = 1 \quad \text{se } Y_i^* \leq \tau_1 \\ Y_i = 2 \quad \text{se } \tau_1 < Y_i^* \leq \tau_2 \\ Y_i = 3 \quad \text{se } \tau_2 < Y_i^* \leq \tau_3 \\ Y_i = 4 \quad \text{se } \tau_3 < Y_i^* \leq \tau_4 \\ Y_i = 5 \quad \text{se } Y_i^* > \tau_4 \end{array} \right.$$

dove $\tau = (\tau_1, \tau_2, \tau_3, \tau_4)$ è il vettore dei parametri delle soglie che, essendo uguali per ogni individuo, non dipendono dall'indice i . Il numero di parametri associati alle soglie corrisponde al numero di categorie della variabile ordinale meno uno. La distribuzione di probabilità della variabile osservata è, per esempio, per la modalità due:

$$\begin{aligned} P(Y_i = 2|X_i) &= P(\tau_1 < Y_i^* \leq \tau_2|X_i) = \\ &= P(\epsilon_i \leq \tau_2 - \alpha - X_i'\beta) - P(\epsilon_i \leq \tau_1 - \alpha - X_i'\beta) = \\ &= \phi\left[\frac{\tau_2 - \alpha - X_i'\beta}{\sigma}\right] - \phi\left[\frac{\tau_1 - \alpha - X_i'\beta}{\sigma}\right] \end{aligned}$$

dove $\phi[\cdot]$ è la funzione di ripartizione della Normale standard. Il modello così costruito non è però identificato, in quanto per diversi valori dei parametri si ottengono gli stessi valori della funzione di verosimiglianza. Dunque, è necessario inserire dei vincoli, come ad esempio:

- imporre la varianza unitaria del termine d'errore e la prima soglia uguale a zero ($\sigma = 1, \tau_1 = 0$). Allora la distribuzione di probabilità diventa:

$$P(Y_i = 1|X_i) = \phi[-\alpha - X_i'\beta]$$

$$P(Y_i = 2|X_i) = \phi[\tau_2 - \alpha - X_i'\beta] - \phi[-\alpha - X_i'\beta]$$

$$P(Y_i = k|X_i) = \phi[\tau_k - \alpha - X_i'\beta] - \phi[\tau_{k-1} - \alpha - X_i'\beta] \quad k = 3, \dots, K - 1$$

$$P(Y_i = K|X_i) = 1 - \phi[\tau_{K-1} - \alpha - X_i'\beta]$$

- imporre la varianza unitaria del termine d'errore e l'intercetta pari a zero ($\sigma = 1, \alpha = 0$). I software statistici spesso applicano questo vincolo e la distribuzione di probabilità associata è:

$$P(Y_i = 1|X_i) = \phi[\tau_1 - X_i'\beta]$$

$$P(Y_i = k|X_i) = \phi[\tau_k - X_i'\beta] - \phi[\tau_{k-1} - X_i'\beta] \quad k = 2, \dots, K - 1$$

$$P(Y_i = K|X_i) = 1 - \phi[\tau_{K-1} - X_i'\beta]$$

Il modello *Ordered Probit* è stimato con il metodo di stima della massima verosimiglianza e, nell'ipotesi di $\sigma = 1$ e $\alpha = 0$, la funzione di *log-verosimiglianza* è:

$$\begin{aligned} l(\tau, \beta) = & \sum_{Y_i=1} \log(\phi[\tau_1 - X_i'\beta]) + \sum_{Y_i=k} \log(\phi[\tau_k - X_i'\beta] - \phi[\tau_{k-1} - X_i'\beta]) + \\ & + \sum_{Y_i=K} \log(1 - \phi[\tau_{K-1} - X_i'\beta]) \end{aligned}$$

2.3.2 Modello *Chopit*

Il modello *Chopit* è la soluzione parametrica che sfrutta le informazioni aggiuntive provenienti dalle Anchoing Vignettes, basandosi sul concetto di variabile latente e in cui, differentemente dal modello *Ordinal Probit*, le soglie variano tra i rispondenti in base ad un set di covariate. L'approccio parametrico supera la problematica dell'inefficienza di quello non parametrico, riconoscendo che la variabile di interesse possa essere percepita con un errore di misura, così che le valutazioni sulle vignettes

non siano vincolate al loro ordine naturale. Inoltre, l'applicazione dell'approccio parametrico non necessita della raccolta delle risposte a tutte le vignettes da parte di ciascun intervistato. Tuttavia, anche il modello parametrico soffre di alcuni importanti svantaggi. Infatti, esso richiede ulteriori assunzioni oltre a quelle di RC e VE, tra le quali quella di una specificazione di una funzione lineare che lega le caratteristiche individuali e la variabile di interesse, quella della forma funzionale delle soglie e quella della distribuzione del termine d'errore.

Innanzitutto bisogna specificare che esistono tre diversi livelli in cui la variabile di interesse può essere misurata:

- livello effettivo, che corrisponde al reale livello di interesse ed è misurato in scala continua (μ_i).
- livello percepito, che è quello che il rispondente percepisce, ma che non si può misurare direttamente ed è sempre in scala continua (Y_i^*).
- livello riportato, che è l'unico di cui si ha una vera misura diretta ed è l'espressione del livello percepito in una variabile categoriale ordinata (Y_i).

Considerando l'autovalutazione, il livello effettivo è espresso come combinazione di caratteristiche individuali, X_i , ed il livello percepito è modellato, attraverso questa misura più un suo errore, come segue:

$$Y_i^* = \mu_i + \epsilon_i = X_i' \beta + \epsilon_i \quad \text{con } \epsilon_i \sim N(0, 1)$$

dove è assunta sia l'indipendenza tra gli errori, ϵ_i , e le covariate, X_i , e sia la nullità della costante per avere così un modello identificato. Gli individui sono chiamati a trasformare il proprio giudizio scegliendo una delle K categorie della variabile risposta discreta con il seguente meccanismo:

$$Y_i = k \quad \text{se } \tau_i^{(k-1)} \leq Y_i^* < \tau_i^{(k)} \quad (2.1)$$

dove

$$-\infty = \tau_i^0 < \tau_i^1 < \dots < \tau_i^K = +\infty$$

Come specificato in King et al. (2004) le soglie sono modellate come funzione di covariate individuali V_i (non necessariamente uguali a X_i), nel seguente modo:

$$\tau_i^1 = \gamma^{(1)}V_i$$

$$\tau_i^k = \tau^{(k-1)} + \exp(\gamma^{(k)}V_i), \quad k = 2, \dots, K-1$$

l'utilizzo della funzione esponenziale garantisce un aumento delle soglie con k , comportando un ordine logico nella scala di valutazione.

Per quanto riguarda la parte della modellazione delle vignettes nel *Chopit*, il procedimento è simile al precedente. Innanzitutto si definisce il livello effettivo, θ_j , dell'ipotetico individuo descritto nella vignette j . Il livello percepito dal rispondente corrisponde a:

$$Z_{ij}^* = \theta_j + \eta_{ij} \quad \text{con } \eta_{ij} \sim N(0, \sigma_\eta^2) \quad \text{e} \quad \eta_{ij} \perp (\epsilon_i, X_i, V_i)$$

In questo caso il valore di σ_η^2 può variare tra le diverse vignettes ed, inoltre, la sua stima può essere vista come indicatore di quanto ogni vignette venga capita specificatamente (King et al., 2004). Segue poi il meccanismo di trasformazione del livello percepito in una scala categoriale ordinata, nel seguente modo:

$$Z_{ij} = k \quad \text{se } \tau_i^{(k-1)} \leq Z_{ij}^* < \tau_i^{(k)}$$

dove le soglie $\tau_i^{(k)}$, $k = 1, \dots, K$, sono le stesse specificate nella componente di autovalutazione. In altre parole, quest'ultima uguaglianza corrisponde proprio all'assunzione di RC, mentre il fatto che θ_j , $j = 1, \dots, J$, non dipenda da alcuna covariata, corrisponde effettivamente all'assunzione di VE.

I parametri del modello *Chopit* vengono stimati utilizzando il metodo della massima verosimiglianza. Sia la componente di autovalutazione che le vignettes hanno la propria funzione di verosimiglianza, ed esse vengono combinate per ottenere la verosimiglianza congiunta. Il contributo della parte di autovalutazione alla verosimiglianza corrisponde a:

$$L_s(\beta, \gamma|Y) \propto \prod_{i=1}^N \prod_{k=1}^K \left(\phi[\tau_i^k | X_i \beta, 1] - \phi[\tau_i^{k-1} | X_i \beta, 1] \right)^{I(Y_i=k)}$$

dove $I(Y_i = k)$ è la funzione indicatrice che vale 1 quando $Y_i = k$ e 0 altrimenti.

Il contributo delle vignettes alla verosimiglianza è invece rappresentato da:

$$L_v(\theta, \gamma|Z) \propto \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^K \left(\phi[\tau_i^k | \theta_j, \sigma_\eta^2] - \phi[\tau_i^{k-1} | \theta_j, \sigma_\eta^2] \right)^{I(Z_{ij}=k)}$$

dove $I(Z_{ij} = k)$ è la funzione indicatrice che vale 1 quando $Z_{ij} = k$ e 0 altrimenti.

La funzione di verosimiglianza congiunta è il prodotto di queste due componenti:

$$L(\beta, \theta, \gamma, \sigma_\eta^2|Y, Z) = L_s(\beta, \gamma|Y) L_v(\theta, \gamma|Z)$$

2.3.3 Estensioni del modello *Chopit*

In letteratura sono state introdotte alcune estensioni del modello *Chopit* standard per affrontare altri aspetti delle Anchoring Vignettes.

Kapteyn et al. (2007) hanno incluso un termine d'errore, ψ_i , nella prima equazione delle soglie, per riuscire a tenere conto di caratteristiche dei rispondenti che non vengono osservate (e di conseguenza riuscire ad indebolire un pò l'assunzione di RC). Le equazioni corrispondenti diventano:

$$\tau_i^{(1)} = \gamma^{(1)} V_i + \psi_i$$

$$\tau_i^{(k)} = \tau_i^{(k-1)} + \exp(\gamma^{(k)} V_i), \quad k = 2, \dots, K - 1$$

dove $\psi_i \sim N(0, \sigma_\psi^2)$ e $\psi_i \perp (\epsilon_i, \eta_i, X_i, V_i)$. Van Soest and Vonkova (2014) hanno dimostrato che questa estensione è utile a ridurre problemi di specificazione del modello, con l'aggiunta di un solo parametro da stimare.

Sempre Kapteyn et al. (2007) hanno mostrato che i rispondenti hanno reazioni diverse rispetto al genere degli individui descritti nelle vignettes. Come conseguenza, la loro percezione del valore del livello effettivo dello scenario è distorto ed essi utilizzano diverse soglie quando valutano una vignette che descrive un uomo o una donna. Questa non è altro che un'altra violazione dell'assunzione di VE. E' stato quindi suggerito di costruire delle versioni di vignettes in cui il genere della persona descritta viene specificato ai rispondenti in modo casuale e quindi è stato proposto di includere nell'equazione delle vignettes una variabile che rappresentasse il genere della vignette.

Infine, Angelini et al. (2011) introducono una componente temporale per indagare in che misura gli stili di valutazione rimangano stabili nel tempo. Se la scala di risposta rimanesse la stessa, allora sarebbe possibile sfruttare il fatto che non sia necessario chiedere le vignettes ai rispondenti in tutte le *waves*, ma basterebbe una volta sola. Questa estensione è chiamata modello *Chopit* longitudinale.

2.4 Test sulle Assunzioni di RC e VE

Le due assunzioni delle Anchoring Vignettes sono cruciali per la correzione del DIF, quindi è necessaria una metodologia per testarne la validità. Nonostante ciò non esistono ancora dei test formali e, per verificare la RC e la VE, vengono utilizzate delle soluzioni che impongono ulteriori assunzioni o condizioni. Questo fatto però impedisce di poter concludere i test quando una delle due assunzioni viene rigettata: infatti non si può sapere se l'ipotesi è rigettata per l'assunzione di base o per quella aggiuntiva. Nel seguito saranno riportate alcune delle soluzioni proposte in letteratura.

Bago d’Uva et al. (2011) proposero dei test per cercare di validare entrambe le assunzioni utilizzando i dati dell’indagine ELSA (*English Longitudinal Study of Ageing*), in particolare quelli della rilevazione del 2006-2007, che presentavano al loro interno sia le risposte ad alcune vignettes, sia dati su degli indicatori obbiettivi. In questo contesto si presume che quest’ultimi possano catturare le variazioni del costrutto di interesse associato alle caratteristiche individuali. Secondo questa ipotesi, qualsiasi variazione sistematica nelle valutazioni soggettive che rimane dopo il condizionamento agli indicatori obbiettivi può essere attribuita all’eterogeneità non osservata (Kerkhofs e Lindeboom, 1995; Kreider, 1999). Il test dell’assunzione di RC richiede di verificare se le soglie utilizzate dall’individuo per valutare sé stesso, identificate dagli indicatori obbiettivi, sono uguali a quelle utilizzate per valutare la vignette. Anche van Soest ed altri (2007) hanno introdotto un test di RC che si basa sul confronto tra le soglie identificate dalle vignettes e una misura oggettiva (si vedrà dopo nel dettaglio), ma il test di Bago d’Uva et al. (2011) differisce in quanto permette di utilizzare più misure oggettive contemporaneamente (cosa auspicabile quando un singolo indicatore non riesce a catturare tutta l’associazione tra le covariate e il costrutto di interesse). Per la VE, invece, si testa la condizione necessaria di non variazione sistematica con le caratteristiche individuali nella differenza percepita tra due qualsiasi vignettes. Quindi questo test può essere eseguito con qualsiasi dataset contenente almeno due vignettes e non richiede misure oggettive. I ricercatori scelsero una vignette come riferimento (ad esempio $j = 1$) e rilassarono l’ipotesi di VE nelle altre $J - 1$, lasciando dunque che le caratteristiche individuali dei rispondenti influenzassero la valutazione di queste $J - 1$ vignettes:

$$Z_{i1}^* = \theta_1 + \eta_{i1}$$

$$Z_{ij}^* = \theta_j + \lambda_j X_i + \eta_{ij} \quad j \neq 1$$

Il test per la VE è il test di significatività congiunta dei parametri del vettore λ_j :

$$\lambda_j = 0, \quad \forall j$$

Se vale che $\lambda_j \neq 0$, allora significa che le caratteristiche individuali influenzano il livello percepito della vignette j e l'assunzione di VE viene violata. Inoltre, Bago d'Uva et al. (2011) introdussero un test per la RC più debole basato su un'assunzione meno forte riguardo le informazioni contenute nell'indicatore oggettivo: essi testarono se la distanza tra due qualsiasi soglie fosse la stessa nei due approcci (considerando in un caso la RC e nell'altro solo l'indicatore oggettivo). I risultati di questo studio rifiutarono sia l'assunzione di RC che di VE, ma accettarono il test più debole proposto.

Un altro studio condotto da van Soest et al. (2011) ha utilizzato un'intervista via web a 3350 studenti dell'*University College of Dublin* per testare la validità dell'assunzione di RC. Agli studenti è stato chiesto di valutare il loro "rapporto" con l'alcol, chiedendo di descrivere le proprie abitudini, rispetto al consumo di bevande alcoliche, tramite una scala ordinale con cinque modalità. È stato poi chiesto di valutare, utilizzando la stessa scala, il consumo di alcol descritto in alcune vignettes loro presentate. Infine, è stato chiesto di esprimere numericamente l'ammontare del proprio consumo. Proprio quest'ultima informazione, come nello studio di Bago d'Uva et al. (2011), è necessaria per testare la validità dell'assunzione, in quanto rappresenta una misura oggettiva della quantità di alcol assunta dai singoli individui. L'autovalutazione e l'indicatore oggettivo sono generati dallo stesso processo latente, in quanto la prima non è altro che una sorta di valutazione del valore assunto dal secondo. Questa assunzione è chiamata *One Factor (OF)* ed è modellata con un *Ordered Probit* standard, dove le soglie restano costanti tra gli individui:

$$Y_{O_i}^* = X_i' \beta_O + \zeta_{O_i}$$

$$Y_{O_i} = l \quad \text{se } \tau_O^{(l-1)} < Y_{O_i}^* < \tau_O^{(l)}$$

dove $Y_{O_i}^*$ e Y_{O_i} sono rispettivamente il livello percepito e il livello riportato, X_i sono le stesse covariate usate per l'autovalutazione e β è il vettore dei parametri associato. Il numero L di categorie per l'indicatore oggettivo può essere anche diverso dal numero K di categorie usate per l'autovalutazione.. Inoltre, si assume che ζ_{O_i} sia indipendente da X_i , ψ_i e η_{ij} , ma non da ϵ_i ; infatti, $(\zeta_{O_i}, \epsilon_i)$ è distribuito come una Normale bivariata. L'assunzione OF si traduce nell'uguaglianza tra il vettore dei parametri dell'indicatore oggettivo e quello corrispondente all'autovalutazione, ossia $\beta_O = \beta$. Il test proposto da van Soest et al. (2011), per validare la RC, consiste nel confrontare il modello imponendo sia la OF che la RC con il modello in cui si impone solo la OF. I risultati dell'indagine, utilizzando l'indicatore AIC, propongono come modello migliore quello che impone anche l'assunzione di RC.

Peracchi e Rossetti (2013) proposero un test congiunto per verificare la validità sia della RC che della VE, basandosi sul fatto che il modello, sotto entrambe le assunzioni, risultasse sovra-identificato. Questo test non ha la necessità di un'informazione oggettiva e può essere applicato a qualsiasi studio presenti almeno una vignette. Per contro, potrebbe rifiutare l'ipotesi di sovra-identificazione anche per motivi differenti dalle due assunzioni, come un'omissione di alcune variabili rilevanti. Perciò, nel caso in cui venga rifiutata la sovra-identificazione, non si esclude la validità delle due assunzioni. Inoltre, il modello utilizzato per la costruzione di questo test non è esattamente uguale al modello *Chopit* introdotto nell'equazione 2.1 in quanto non viene specificata, nell'equazione delle soglie, la funzione esponenziale.

Può accadere che l'ordine con cui vengono esposte le vignettes rispetto all'autovalutazione sia rilevante per la legittimità delle assunzioni di RC e VE. Analizzando lo stesso dataset considerato in questo lavoro di tesi, Basei (2017) dimostra che la distribuzione della risposta all'autovalutazione cambia a seconda dell'ordine; questo potrebbe dipendere anche dalla validità o meno delle assunzioni. Su questo tema, Hopkins and King (2010) studiarono se l'ordine in cui veniva proposto il questionario

influenzasse le valutazioni dei rispondenti. Inizialmente King et al. (2004) suggerirono di proporre ai rispondenti prima l'autovalutazione e poi le vignettes, affinché la risposta riguardo la propria condizione non sarebbe stata influenzata dagli scenari descritti nelle vignettes. Hopkins and King (2010) analizzarono l'effetto di invertire l'ordine delle domande trovando che effettivamente migliorava la correzione del DIF. Ciò era dovuto al fatto che chi prima rispondeva alle vignettes aveva una concezione più standardizzata di ciò che veniva chiesto, dato che gli scenari sono i principali responsabili per definire la scala di risposta. Infatti, le vignettes fornivano una sorta di chiarimento del significato della domanda e indirizzavano gli individui a fare delle considerazioni sulle proprie risposte.

Riguardo la convalida delle assunzioni delle Anchoring Vignettes in letteratura sono proposti altri approcci come quelli di, ad esempio, Au and Lorgelly (2014) per la *Response Consistency*, Vonkovà and Hullegie (2011) e Rice et al. (2011) per la *Vignette Equivalence*; una review su questo tema è disponibile in Paccagnella (2013).

Questi esempi mostrano come non esista un test formale per accertare la validità delle assunzioni e il dibattito sia ancora molto acceso.

2.5 Applicazioni Anchoring Vignettes

Le Anchoring Vignette sono state incluse in un numero crescente di indagini, tra cui lo studio ELSA (English Longitudinal Study of Ageing), lo studio HRS (Health and Retirement Study), l'indagine sulla salute, l'invecchiamento e il pensionamento in Europa (SHARE) e l'indagine WHS (World Health Surveys). Le applicazioni di questa metodologia stanno aumentando rapidamente e ora includono studi riguardanti la *political efficacy* (King et al., 2004), la *work disability* (Kapteyn et al, 2007), la *job satisfaction* (Kristensen and Johansson, 2008), la *life satisfaction* (Angelini et al., 2014), la salute (Bago d'Uva et al., 2008) e la risposta del sistema sanitario (Rice et al., 2008). Molti di questi studi rivelano una forte eterogeneità tra le per-

sone e quindi importanti impatti dovuti alle correzioni basate sulle vignettes. Tutti i risultati confermano la capacità di questo approccio nel correggere il DIF, ma in mancanza dei test formali per la convalida della metodologia, rimane dell'incertezza circa l'appropriatezza e l'accuratezza di tali "correzioni".

2.6 Combinazione Anchoring Vignettes e modelli CUB

Questo lavoro di tesi vuole affrontare le problematiche legate alle assunzioni di base dell'approccio delle Anchoring Vignettes, al fine di costruire dei test per la RC e per la VE che non utilizzino condizioni o assunzioni aggiuntive. La metodologia impiegata per questo fine è proprio quella dei modelli CUB descritti nel primo capitolo. E' possibile modellare facilmente le valutazioni dei rispondenti, sia nel caso del giudizio su sé stessi sia nella valutazione degli individui ipotetici delle vignettes, dato che entrambe sono espresse tramite una variabile categoriale ordinale. Dunque, l'associazione con i modelli CUB è di fatto immediata.

Per la costruzione di questa nuova metodologia, attua alla verifica della legittimità delle assunzioni di base previste dalle Anchoring Vignettes, è sufficiente identificare quali parametri dei modelli CUB siano coinvolti nella formulazione delle ipotesi di RC e VE e costruire dei test su di essi.

La prima assunzione, quella di RC, prevede che i rispondenti usino le stesse scale di risposta per valutare sé stessi e gli individui ipotetici descritti nelle vignettes; questo equivale a considerare lo stesso livello di incertezza² nel processo che guida i rispondenti ad esprimere le valutazioni su entrambe le domande. In altre parole, si ipotizza che se l'incertezza della vignette e dell'autovalutazione è la stessa allora gli stili di risposta utilizzati dai soggetti sono equivalenti per entrambe le domande

²L'incertezza è il risultato di un insieme di fattori esterni, quali ad esempio la pigritia del soggetto e la volontà di voler scherzare o rispondere il falso, la comprensione parziale della domanda oppure la limitata gamma di informazioni disponibili.

del questionario. Viceversa, se l'incertezza usata dai rispondenti è differente per la vignette e per l'autovalutazione, allora l'ipotesi di RC non può essere validata. Un esempio del parallelismo tra incertezza e stili di risposta è rappresentato dal comportamento degli individui senza un lavoro che tendono a sottostimare la propria salute (rispondere il falso) per giustificare il proprio stato di disoccupazione (la cosiddetta *justification bias*). Oppure, un'altra situazione che lega la componente d'incertezza agli stili di risposta, deriva dalla comprensione parziale della domanda: infatti, può accadere che coloro che non comprendono appieno il significato della vignette non siano in grado di utilizzare la stessa scala di risposta che utilizzerebbero per l'autovalutazione.

Per quanto riguarda la validazione della VE, invece, si è reso necessario un test congiunto dei parametri associati alle covariate della componente di *feeling*. Difatti, questo test equivale a verificare se le caratteristiche individuali dei soggetti influenzino o meno la percezione della vignette e del suo significato. Dunque, per l'applicazione corretta delle Anchoring Vignettes deve valere che le stime dei parametri relativi alla componente che, nei modelli CUB, descrive il sentimento degli individui verso l'oggetto della valutazione, siano tutte contemporaneamente pari a zero.

In generale, per applicare correttamente il modello *Chopit* devono valere contemporaneamente le due ipotesi di RC e VE. Il fatto che con l'approccio proposto in questo lavoro di tesi non siano previste condizioni aggiuntive di applicabilità dei test è legato al non ipotizzare, quando si testa una delle due assunzioni, che necessariamente valga anche l'altra. In altre parole, quando si testa la RC, ad esempio, viene stimato un modello CUB in cui il *feeling* varia da individuo a individuo per mezzo di covariate: questo significa che in questo caso si sta testando la RC, rilassando l'ipotesi di VE. In modo analogo, quando si testa la VE viene stimato un modello CUB dove la componente d'incertezza della vignette dipende da determinate caratteristiche individuali, che non vengono vincolate ad essere uguali a quelle dell'au-

tovalutazione. Questo permette di effettuare i test sulle assunzioni separatamente, senza aggiungere ulteriori condizioni. Si può concludere quindi che, diversamente da altri test presenti in letteratura, il risultato dei test sulla RC e sulla VE così ottenuto non lascerà dubbi sul fatto che questo dipenda da ulteriori condizioni imposte per poterlo applicare, ma permetterà di concludere immediatamente se l'approccio con le Anchoring Vignettes sia corretto o meno per correggere il DIF.

Il procedimento che è stato attuato prevede la combinazione di due modelli CUB: il primo riferito alle risposte delle vignettes e l'altro alle risposte delle autovalutazioni, entrambe con m categorie tra le quali scegliere la valutazione più vicina al livello di soddisfazione percepito. L'equazione relativa al modello CUB per una singola vignette è:

$$P(R_i = r) = \pi_{vi} b_r(\xi_{vi}) + (1 - \pi_{vi}) \frac{1}{m} \quad r = 1, \dots, m \quad (2.2)$$

mentre quella per l'autovalutazione è:

$$P(R_i = r) = \pi_{si} b_r(\xi_{si}) + (1 - \pi_{si}) \frac{1}{m} \quad r = 1, \dots, m \quad (2.3)$$

Come introdotto nel primo capitolo la funzione legame tra le covariate e i parametri usata in un modello CUB è la funzione *logit*:

$$\left\{ \begin{array}{l} \text{logit}(\pi_{vi}) = y_i \beta_v \\ \text{logit}(\pi_{si}) = y_i \beta_s \\ \text{logit}(\xi_{vi}) = w_i \gamma_v \\ \text{logit}(\xi_{si}) = w_i \gamma_s \end{array} \right.$$

dove y e w sono le covariate riferite alla componente di incertezza e di *feeling*, rispettivamente, e β e γ sono i parametri associati alle covariate, stimati tramite l'algoritmo E-M.

Il concetto di RC delle Anchoring Vignettes, nel contesto dei modelli CUB, si traduce nel verificare l'uguaglianza dei parametri delle covariate riferite all'incertezza nei due Modelli 2.2 e 2.3; invece, la verifica dell'assunzione di VE si traduce nel testare se le stime dei parametri associate alle covariate per il *feeling* del Modello 2.2 siano congiuntamente pari a zero. Le due assunzioni possono essere quindi sintetizzate come segue:

$$\begin{cases} H_0 : \beta_v = \beta_s & \rightarrow \text{RC} \\ H_0 : \gamma_v = 0 & \rightarrow \text{VE} \end{cases}$$

Nella pratica, per quanto riguarda la validazione di queste due assunzioni si è proceduto come descritto nei passi seguenti:

1. Si è stimato un modello CUB con covariate (p variabili per la componente d'incertezza e q per la componente di *feeling*) sulle risposte relative ad una vignette.
2. Successivamente, si è stimato un altro modello CUB con covariate (per la componente d'incertezza le stesse utilizzate per la vignette; per il *feeling*, invece, le covariate possono essere diverse) per le risposte relative all'autovalutazione, vincolando i parametri dell'incertezza ad essere gli stessi stimati con il modello sulla vignette precedente (β_v), per poter effettuare così il test sulla RC.
3. A questo punto, sulla stessa risposta all'autovalutazione si è stimato un modello CUB con covariate senza alcun vincolo. Il modello del punto 2 risulta annidato in quest'ultimo, per cui è possibile effettuare un test del rapporto di verosimiglianza (LR test) per verificare l'uguaglianza tra le funzioni di *log-verosimiglianza* dei due modelli annidati. In particolare, la funzione di *log-verosimiglianza* del modello vincolato è stata calcolata utilizzando la risposta all'autovalutazione, le stime dei parametri relativi al *feeling* (γ_s), ottenute tramite algoritmo E-M, e le stime dei parametri relativi all'incertezza della vignette (β_v), ottenute al passo 1. Sulla stessa risposta è stata poi calcolata

la *log-verosimiglianza* utilizzando però le stime dei parametri del feeling (γ_s) e dell'incertezza (β_s), entrambe relative al modello dell'autovalutazione non vincolato. L'LR test verifica se queste verosimiglianze siano uguali utilizzando come distribuzione una *chi-quadro* con p gradi di libertà:

$$LR \sim \chi_p^2$$

dove p corrisponde al numero di covariate considerate per la componente d'incertezza. Nel caso di accettazione dell'ipotesi nulla del test, è quindi possibile affermare la legittimità dell'assunzione di RC.

4. Una volta testata la RC si passa a verificare la VE, appurando che le stime dei parametri della vignette (γ_v) siano tutte congiuntamente pari a 0. Dunque, è stato stimato un modello per la vignette con le stesse covariate per la componente d'incertezza del Modello 2.2, ma che per la parte di *feeling* avesse solo l'intercetta. Questo implica che l'assunzione di RC in questo caso viene rilassata per poter procedere con il test sulla VE. Si procede col verificare l'uguaglianza delle funzioni di *log-verosimiglianza* del modello completo e del modello ridotto, rivelando, in caso affermativo, che i rispondenti giudicano la vignette senza essere influenzati dalle proprie caratteristiche individuali.

L'ordine con cui si stima prima il modello CUB per la vignette e poi quello per l'autovalutazione non è casuale. Infatti, come nel modello *Chopit* le vignettes vengono sfruttate per identificare i parametri delle soglie e solo successivamente le risposte alle autovalutazioni permettono di identificare i parametri dell'equazione di autovalutazione, anche nei modelli CUB si utilizza una procedura simile: le risposte alle vignettes vengono sfruttate per valutare l'incertezza del rispondente e, nel caso di validazione della RC, la stessa incertezza deve valere anche per l'autovalutazione.

Inoltre, con la metodologia appena introdotta, diversamente dai test che si trovano in letteratura, non si lavora con tutte le vignettes contemporaneamente, ma,

viste le caratteristiche del modello CUB e il fatto che una vignette è sufficiente per applicare il metodo delle Anchoring Vignettes, questa nuova soluzione permette di verificare le assunzioni per una vignette alla volta; in questo modo si ha la possibilità di evidenziare la validità o meno delle assunzioni per ogni singola vignette. Negli altri test, invece, si poteva dire solamente se per tutte le vignettes assieme le ipotesi valevano oppure no; dunque, poteva capitare che se l'assunzione non fosse stata convalidata sarebbe stato possibile che il test venisse rigettato perché solo per alcune vignettes le ipotesi non erano valide, ma per altre sì. Questo aspetto consolida ulteriormente l'efficacia di questa strategia che permette di effettuare analisi che in altri contesti non sarebbero possibili.

Il lavoro realizzato con il software statistico R sul codice del programma `CUB.R` (Iannario & Piccolo, 2014) dei modelli CUB, è stato quello di costruire un'ulteriore funzione, chiamata **AVCUB** (combinazione tra **Anchoring Vignettes** e modelli **CUB**), che calcolasse entrambe le equazioni dei Modelli 2.2 e 2.3, fornendo le stime di entrambi e il risultato dei test proposti per la RC e la VE.

Capitolo 3

Caso di studio

Per l'applicazione delle tecniche viste nei precedenti capitoli si è utilizzato un insieme di dati riguardanti le valutazioni sui servizi di *online banking* da parte di un campione di intervistati italiani. Esse costituiscono le risposte ad un questionario sulla soddisfazione dei clienti comprendente sia informazioni sul servizio attivato, sia sulle loro caratteristiche individuali ed anche una serie di vignettes, utili agli scopi di questo lavoro di tesi. Per una ulteriore validazione dei risultati ottenuti si è scelto anche un altro insieme di dati, provenienti in questo caso dal LISS (Longitudinal Internet Studies for the Social sciences) panel e riguardanti la soddisfazione rispetto all'acquisto di alcuni prodotti di telecomunicazioni da parte di un campione di intervistati olandesi. Nel seguito si vedrà nel dettaglio la composizione dei dati e l'applicazione delle tecniche descritte nei capitoli precedenti.

3.1 Dati *Online Banking*

I dati oggetto di questa tesi sono raccolti attraverso un questionario (Appendice B) che analizza l'uso di uno specifico servizio e la corrispondente soddisfazione su di esso riguardo l'*online banking* in Italia. Il progetto di questa raccolta dati è stato definito da un team di ricercatori del Dipartimento di Scienze Statistiche dell'Università di Padova (Proff. Omar Paccagnella, Mariangela Guidolin e Francesca Bassi), mentre

la raccolta vera e propria dei dati è stata realizzata da Doxa (Istituto di Ricerca Statistica e Analisi dell'Opinione Pubblica) in due periodi : Maggio 2015 e Settembre 2015. In questa indagine viene intervistato un solo componente di ogni famiglia, utilizzando il metodo CAWI (Computer Assisted Web Interviewing).

Il questionario è composto da una serie di domande quali il tipo di conto corrente online che è stato aperto, le motivazioni che hanno spinto il cliente a scegliere questa soluzione, da quanto tempo viene sfruttata questa tipologia di servizio, se sono stati riscontrati problemi nell'utilizzo e, nel caso affermativo, viene richiesto un giudizio sui servizi di call center o di assistenza online eventualmente contattati. A seconda dei servizi di *online banking* utilizzati, i due aspetti sui quali si vuole investigare la soddisfazione dei clienti sono la facilità di navigazione e la semplicità di esecuzione delle operazioni del conto corrente online. A questo scopo, sono previste delle domande sulle aspettative che il cliente aveva prima dell'attivazione del servizio e su quanto la sua esperienza ha corrisposto tali aspettative. La differenza tra queste due quantità potrebbe rappresentare una misura oggettiva della soddisfazione del cliente, dato che è la stessa persona che valuta la propria soddisfazione, prima e dopo, con la stessa scala ed inoltre, sarebbe utilizzabile per l'applicazione del test su RC come proposto da van Soest et al. (2011). Successivamente, nel questionario vengono proposte le domande dirette per misurare la soddisfazione dei clienti. A seconda dei servizi utilizzati nel conto corrente online (riassunti dalla domanda Q6) vengono poste ai rispondenti le autovalutazioni o per uno solo dei due aspetti di interesse o per entrambi (facilità di navigazione del conto corrente online - Q13; semplicità di esecuzione delle operazioni del conto corrente online - Q19). Associate a ciascuno di questi aspetti vengono proposte due vignettes per l'eliminazione del DIF. Infine, viene chiesto se il cliente raccomanderebbe il servizio.

Lo studio di questi dati aveva anche come scopo quello di capire se una somministrazione del questionario con un ordine diverso tra vignettes e autovalutazione condizionasse le scelte dei rispondenti. L'insieme dei clienti è dunque stato diviso in

quattro gruppi come riportato in Tabella 3.1. Nei Gruppi 1 e 2 l'autovalutazione è richiesta prima delle vignettes, che sono proposte rispettivamente in ordine invertito, mentre, nei Gruppi 3 e 4 l'autovalutazione segue le due vignettes, sempre in ordine invertito.

Gruppo 1	Autovalutazione - Vignette 1 - Vignette 2
Gruppo 2	Autovalutazione - Vignette 2 - Vignette 1
Gruppo 3	Vignette 1 - Vignette 2 - Autovalutazione
Gruppo 4	Vignette 2 - Vignette 1 - Autovalutazione

Tabella 3.1: Ordine di somministrazione del questionario sull'*Online Banking* per ciascun gruppo.

Le variabili a cui fanno riferimento le autovalutazioni sulla soddisfazione, per entrambi gli aspetti di interesse dello studio, nei Gruppi 1 e 2 sono:

- Q13: livello di soddisfazione per la facilità di navigazione dell'*online banking*
- Q19: livello di soddisfazione per la semplicità dell'esecuzione di operazioni bancarie

mentre, nei Gruppi 3 e 4 le corrispondenti misure saranno rappresentate dalle variabili Q13bis e Q19bis. Le vignettes associate saranno:

- Q15/Q15bis e Q16/Q16bis, per correggere il DIF della Q13/Q13bis
- Q21/Q21bis e Q22/Q22bis, per correggere di DIF della Q19/Q19bis

Tutte le risposte a queste valutazioni sono variabili ordinali con $m = 5$ categorie: da "Molto insoddisfatto/a" a "Molto soddisfatto/a".

Oltre al questionario sulla soddisfazione proposto ai clienti, su di essi si hanno anche informazioni personali quali ad esempio:

- Genere (Maschio, Femmina)

- Fascia d'età (18-34 anni, 35-54 anni, maggiore di 55 anni)
- Regione italiana di residenza
- Professione (Studenti, Lavoratori dipendenti, Lavoratori indipendenti, Altro (pensionato, disoccupato, etc.))
- Titolo di studio (Master/Dottorato, Laurea Specialistica, Laurea Primo Livello-Triennale, Diploma di scuola media inferiore, Licenza di scuola elementare, Nessun titolo di studio). Per semplicità sarà utilizzata anche come variabile dicotomica: alto (dalla Laurea triennale al Dottorato) vs medio-basso (da nessun titolo al diploma) di istruzione.
- Numerosità del nucleo familiare
- Presenza di figli
- Stato civile (Coniugato/a o convivente, Nubile/celibe, Divorziato/a, Separato/a, Vedovo/a).

ed anche informazioni sulle modalità mediante le quali il questionario è stato somministrato:

- Tipo di dispositivo usato per la compilazione (Computer o Tablet)
- Tipo di browser usato per la compilazione (Chrome, Firefox, Internet Explorer, Mozilla, Opera, Safari)
- Versione del browser usato per la compilazione

3.2 Analisi preliminari

Prima di procedere con l'applicazione delle tecniche introdotte nei capitoli precedenti, è necessaria una descrizione delle principali caratteristiche socio-demografiche dei rispondenti del questionario sull'*online banking*. È importante ricordare che solo coloro che hanno completato l'intero questionario, e che quindi hanno almeno un conto corrente che prevede operazioni online, sono considerati nelle analisi descrittive.

In particolare, il campione è composto in prevalenza da uomini (57%), l'età media è di circa 42 anni e quella mediana di 40 anni. Inoltre, solo il 32% degli individui ha conseguito almeno una laurea e più della metà dei rispondenti (55%) non ha a carico alcun figlio. La distribuzione della professione degli individui è riportata in Figura 3.1. La percentuale maggiore si ha in corrispondenza dei lavoratori dipendenti, segue poi la categoria che racchiude i pensionati, i disoccupati, le casalinghe, etc. Infine, le professioni con percentuali più basse nel campione sono i lavoratori indipendenti e gli studenti.

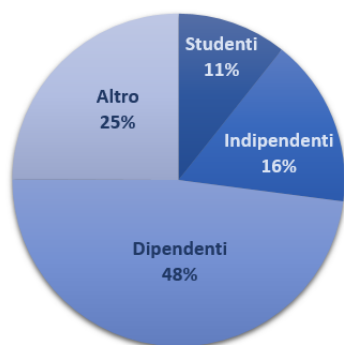


Figura 3.1: Distribuzione della variabile professione nel campione.

Dato che la variabile che rappresenta l'età degli individui verrà utilizzata nelle analisi suddivisa in fasce, si riporta in Figura 3.2 la sua distribuzione nel campione. Si noti come circa l'80% dei rispondenti abbia un'età inferiore ai 55 anni, dunque la clientela che utilizza questa tipologia di servizi può essere ritenuta abbastanza giovane.

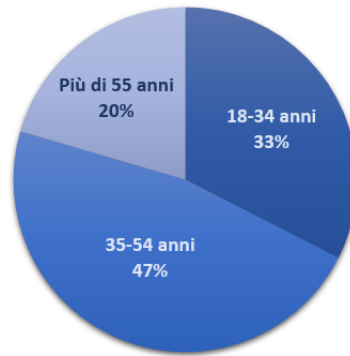


Figura 3.2: Distribuzione delle classi d'età nel campione.

Successivamente, per avere una visione preliminare delle valutazioni dei rispondenti, si riporta la distribuzione delle risposte alle autovalutazioni, per entrambi gli ordini in cui sono proposte, in Figura 3.3. Si può osservare come il servizio di *online banking* sia stato apprezzato dalla maggioranza dei rispondenti, che quindi si sono ritenuti prevalentemente soddisfatti e molto soddisfatti sia per la facilità di navigazione del conto corrente online sia per la semplicità di esecuzione delle operazioni su di esso. Inoltre, si nota che quando l'autovalutazione è posta a seguire delle vignettes associate, la percentuale dei clienti soddisfatti diminuisce: per quanto riguarda la prima autovalutazione, aumenta la percentuale di rispondenti incerti che passa dal 5.29% a circa il 10%; per la seconda autovalutazione, invece, aumenta la percentuale di coloro che sono molto soddisfatti del servizio. Tali cambiamenti possono essere frutto sia di possibili influenze dovute agli scenari descritti nelle vignettes che hanno cambiato l'interpretazione della domanda, sia dell'ordine di presentazione delle vignettes stesse, come evidenziato in Basei (2017).

Per quanto riguarda le distribuzioni delle valutazioni alle vignettes, esse vengono riportate in Figura 3.4 e in Figura 3.5, dove si è distinto l'ordine in cui sono state proposte ai rispondenti rispetto all'autovalutazione. Si nota come le distribuzioni delle vignettes Q16 e Q21 siano nettamente diverse da quelle di Q15 e Q22. Infatti, quest'ultime hanno un comportamento più "standard", nel senso che lo scenario descritto non presenta fatti o circostanze particolarmente negative o problematiche,

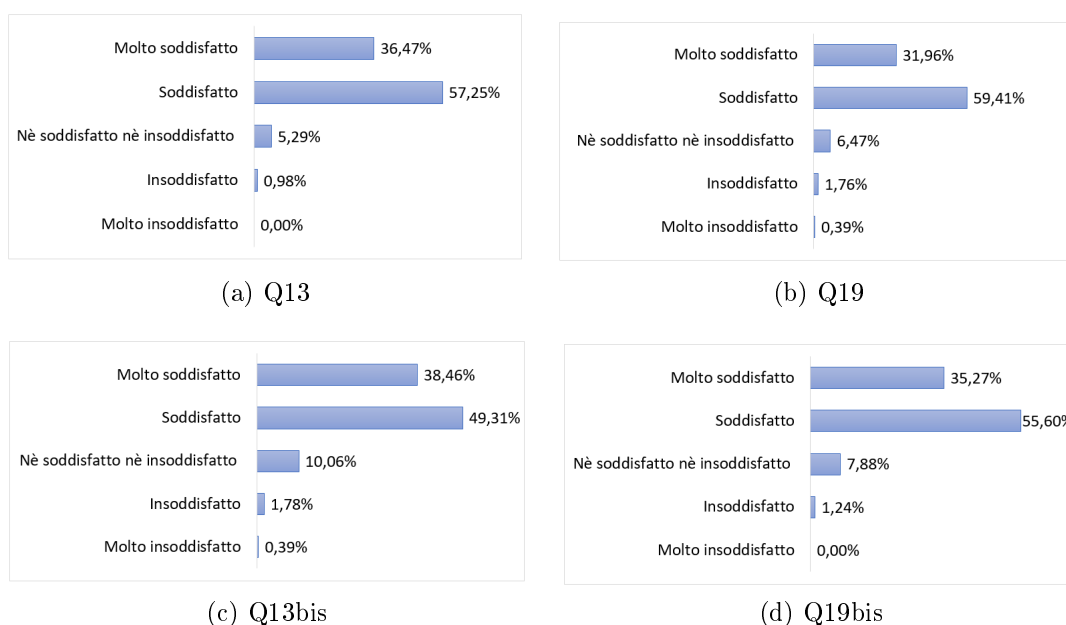


Figura 3.3: Distribuzione di frequenza delle risposte alle autovalutazioni.

quindi non è potenzialmente soggetto a forti interpretazioni individuali. Esse dunque sono valutate positivamente (o negativamente) dalla maggioranza dei rispondenti. Invece, la Q16 e la Q21 hanno una distribuzione delle risposte più simmetrica e non vi è una predominanza di un giudizio nè positivo nè negativo. Questo potrebbe essere dovuto al fatto che la vignette non descrive uno scenario "standard", bensì propone una storia in cui compaiono degli elementi di criticità, lasciando quindi ai rispondenti degli ampi margini di interpretabilità. Ad esempio, la vignette Q16 è la seguente:

"Marina è una casalinga che controlla circa ogni 3 giorni la lista delle spese effettuate dalla sua famiglia con la carta di credito. Un giorno vuole ricontrollare le spese sostenute nel mese precedente, ma non trova il menù a tendina per indicare il mese corretto. Ha bisogno di contattare il call center per risolvere il problema. Grazie all'operatore riesce a far apparire la lista dei movimenti che stava cercando. Quanto è soddisfatta Marina riguardo alla facilità di navigazione del suo conto corrente online principale?"

Tale vignette potrebbe risultare ambigua: il rispondente non sa se giudicare il fatto che Marina non sia riuscita a trovare una soluzione al suo problema autonomamente, anche se poi alla fine, grazie al call center, il problema viene risolto.

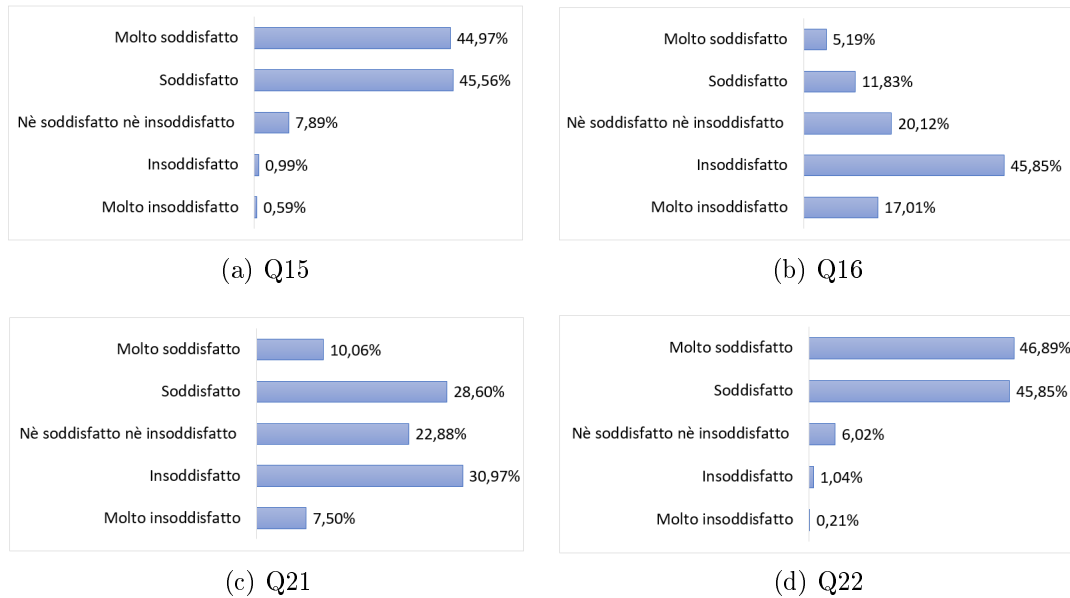


Figura 3.4: Distribuzione di frequenza delle risposte alle vignettes, quando esse sono proposte dopo l'autovalutazione.

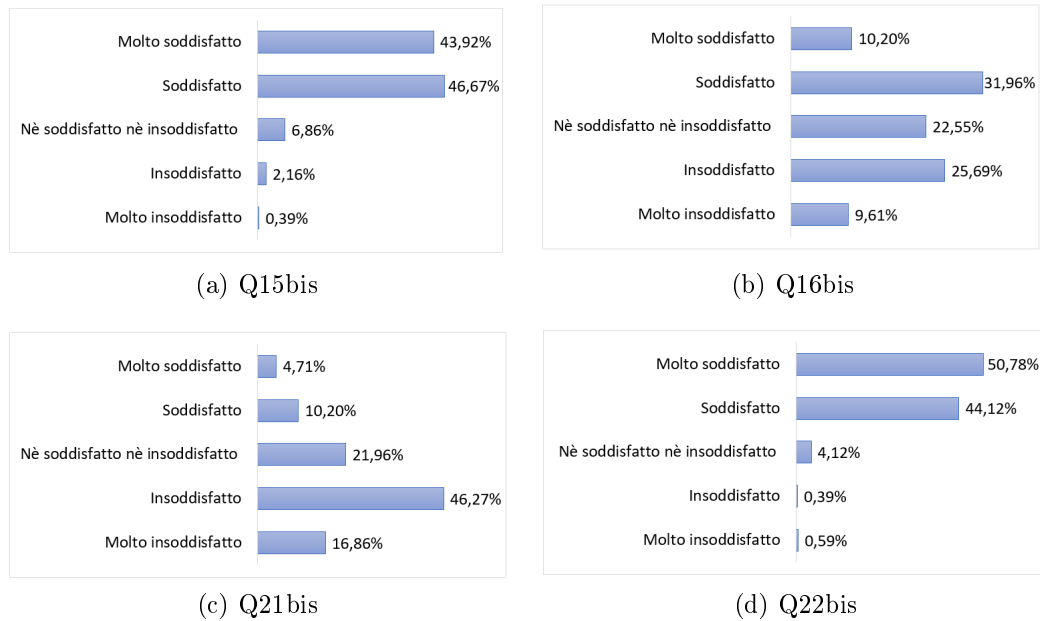


Figura 3.5: Distribuzione di frequenza delle risposte alle vignettes, quando esse sono proposte prima dell'autovalutazione.

Per una ulteriore descrizione della distribuzione delle risposte nel campione si procede scomponendole in base alle covariate che sono state ritenute d'interesse. Dai grafici successivi è quindi possibile osservare la percentuale di individui che, con quella determinata caratteristica, hanno fornito quella particolare risposta sia all'autovalutazione che alla valutazione delle vignettes.

In Figura 3.6 si mostra la distribuzione del primo blocco di domande sulla soddisfazione (Q13-Q15-Q16), in base al genere dei rispondenti. Non è evidente alcuna forte diversità in base a questa caratteristica, ma bensì si ritrova la netta distinzione tra la distribuzione di Q16 rispetto a quelle di Q13 e Q15. Per quest'ultime, infatti, la percentuale di risposte alle categorie più negative è molto bassa, poiché gli individui, sia maschi che femmine, hanno concentrato le proprie valutazioni su livelli alti di soddisfazione.

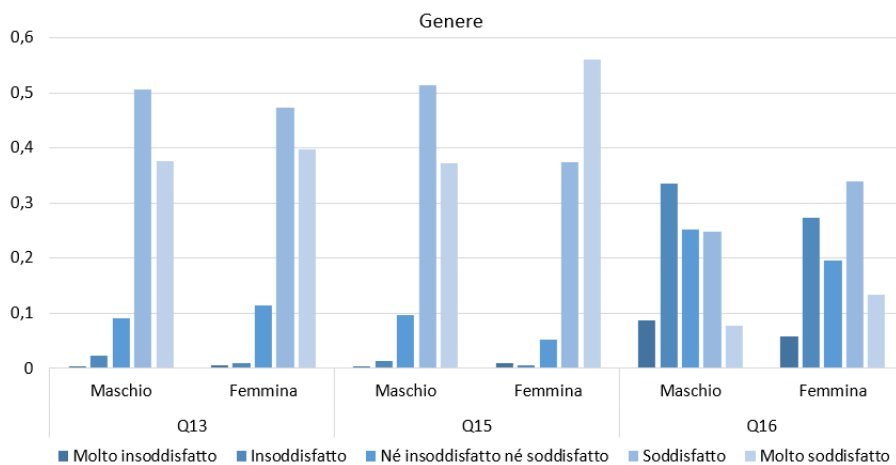


Figura 3.6: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al genere degli individui.

Un risultato simile al precedente si ottiene per il secondo blocco di domande sulla soddisfazione (Q19-Q21-Q22) la cui distribuzione delle valutazioni rispetto al genere è riportata in Figura 3.7. Anche in questo caso la distribuzione delle valutazioni della vignette Q21 ("non standard") si differenzia molto da quelle di Q19 e Q22 e, nuovamente, alcune categorie non vengono addirittura considerate dagli intervistati, indipendentemente dal loro genere.

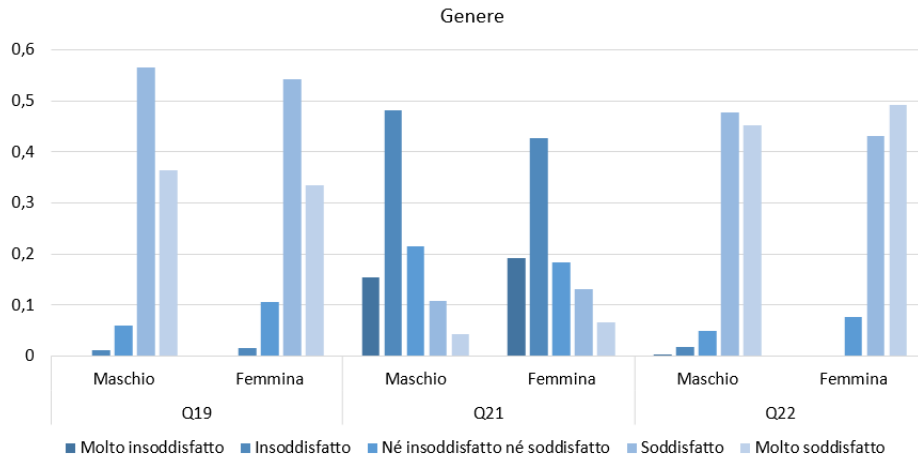


Figura 3.7: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al genere degli individui.

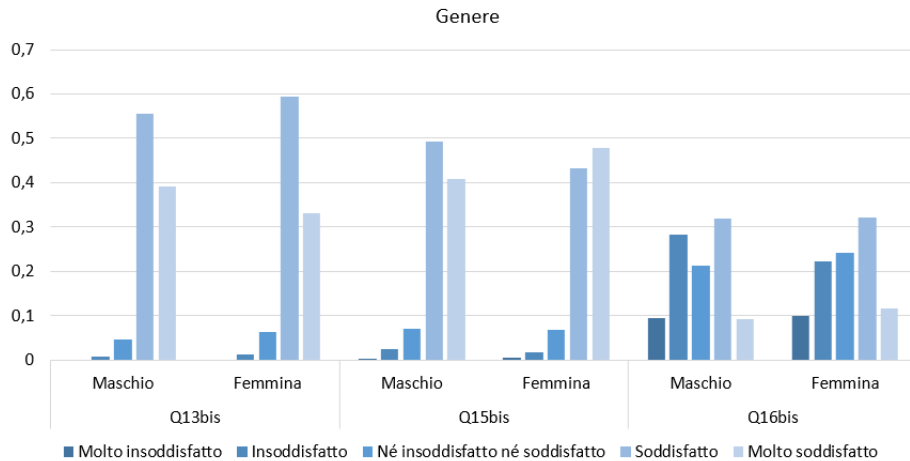


Figura 3.8: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al genere degli individui.

I risultati corrispondenti agli stessi blocchi di domande, con ordine invertito tra vignettes e autovalutazione, vengono riportati in Figura 3.8 e in Figura 3.9, per i quali le conclusioni sono le medesime dell'ordine originale.

La scarsa numerosità di valutazioni per le modalità "Moito insoddisfatto" e "Insoddisfatto" di certe domande del questionario, rispetto a determinate covariate, sarà una problematica molto limitativa sia per la stima dei modelli CUB, ma soprattutto per quella dei modelli *Chopit* e per la loro convergenza.

Per approfondire ulteriormente questo tema si riportano in Appendice A le distribuzioni delle risposte rispetto alle caratteristiche degli individui che, in seguito,

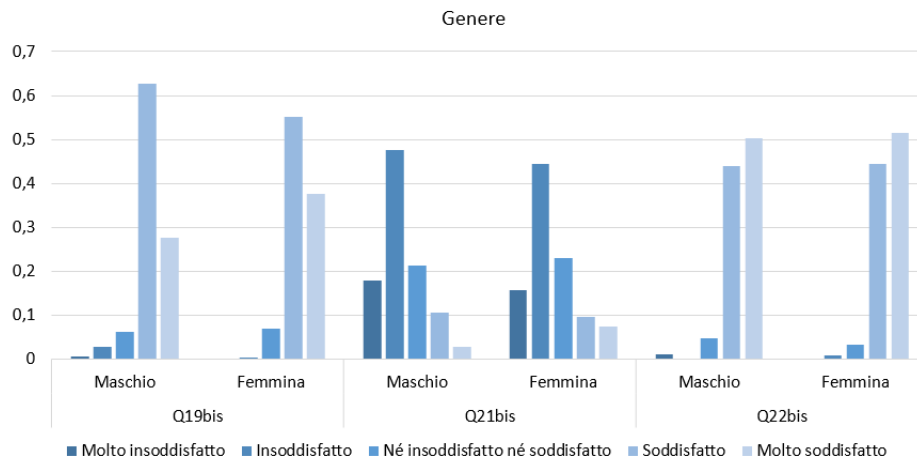


Figura 3.9: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al genere degli individui.

saranno inserite nei modelli. I risultati relativi a queste analisi sono molto simili a quelli ottenuti per la variabile genere e, per questo, i grafici non vengono riportati in questa sezione: *(i)* non si distinguono comportamenti nettamente differenti nelle valutazioni in base alle caratteristiche dell'individuo; *(ii)* Q16 e Q21 (e i rispettivi Q16bis e Q21bis) hanno distribuzioni molto diverse da quelle delle altre vignettes e autovalutazioni; *(iii)* le autovalutazioni e le vignettes "standard" sono state valutate in generale molto positivamente tanto che, quando si analizzano le risposte delle categorie più negative, rispetto alle caratteristiche degli individui, si nota che per molte modalità delle variabili si hanno frequenze molto basse e addirittura nulle.

In conclusione, per tutte le autovalutazioni il giudizio è stato pressoché positivo e spesso le categorie più negative risultano molto poco utilizzate. Per quanto riguarda le vignettes, invece, bisogna distinguere tra quelle "standard" (Q15, Q15bis, Q22 e Q22bis) e quelle "non standard" (Q16, Q16bis, Q21 e Q21bis). Quest'ultime presentano delle risposte ben distribuite su tutte le modalità, proprio perché lo scenario è stato interpretato in modi diversi e non ci sono state categorie poco o per nulla considerate, anche tenendo conto delle caratteristiche individuali dei rispondenti. Viceversa, nelle vignettes "standard" le basse numerosità relative a certe categorie di risposta comporteranno alcune difficoltà nelle stime dei modelli.

3.3 Modello CUB

Come specificato nel Capitolo 1, i modelli CUB non necessitano di alcuna assunzione, se non il fatto che per avere l'identificabilità è necessario un numero di categorie di risposta maggiore di tre; per questo motivo è possibile applicarli facilmente all'insieme di dati dell'*online banking*, dato che per questo dataset il numero di categorie di risposta è pari a cinque.

Si procede inizialmente adattando un singolo modello CUB per ogni risposta ad autovalutazioni e vignettes, ottenendo le stime dei parametri per ciascuna variabile considerata, attraverso lo schema descritto nel primo capitolo di questa tesi. Per la componente di incertezza sono state inserite $p = 12$ covariate, mentre per la parte di *feeling* $q = 9$ covariate, non ritenendo necessario introdurre per quest'ultima componente anche le informazioni riguardanti le modalità di svolgimento del questionario (Touch, Chrome, Explorer) poiché è ragionevole ipotizzare che queste caratteristiche non abbiano un effetto sul *feeling*. A causa dei problemi di numerosità già citati precedentemente, in alcuni modelli non è stato possibile inserire tutte le covariate di interesse.

I modelli CUB sono stati applicati tenendo presente delle suddivisioni previste dalla costruzione del questionario, utilizzando i dati a disposizione in base al gruppo di appartenenza dei rispondenti. I risultati delle stime con modello CUB di coloro che appartengono al primo gruppo, e a cui è stato chiesto di rispondere al blocco di domande Q13-Q15-Q16, sono riportati nelle Tabelle 3.2, 3.3 e 3.4. Invece, le stime per coloro che appartengono sempre al primo gruppo, ma che hanno risposto al blocco di domande Q19-Q21-Q22, sono riportate nelle Tabelle 3.5, 3.6 e 3.7. Infine, per il secondo gruppo, ai quali è chiesto di valutare prima le vignettes e solo dopo dare un giudizio sulla propria soddisfazione, i risultati sono riportati nelle Tabelle 3.8, 3.9, 3.10, per la valutazione della facilità della navigazione e nelle Tabelle 3.11, 3.12, 3.13, per la valutazione della semplicità delle operazioni bancarie. Dai risultati di questi modelli si evince che la componente d'incertezza in questo contesto è pra-

ticamente inesistente, dato che le stime dei parametri associate ad essa sono tutte non significative. L'unica domanda per la quale i valori riportati del p-value sono più bassi rispetto alle altre è la Q21, che raggiunge anche una significatività al 5%, per quanto riguarda la costante della componente d'incertezza quando viene posta dopo l'autovalutazione (Tabella 3.6). In riferimento alla componente di *feeling* e alla significatività delle stime dei parametri ad esse riferite, bisogna invece procedere analizzando caso per caso. Difatti, considerando l'autovalutazione Q13, sia quando è posta prima sia dopo le vignettes, il giudizio su di essa non risulta influenzato da alcuna caratteristica individuale; tuttavia, quando si considerano le vignettes ad essa associate, le valutazioni di queste cambiano in base a certe covariate. Ad esempio, il genere del rispondente, nei modelli per Q15, Q16 e Q15bis, influenza la valutazione della soddisfazione degli individui descritti nelle vignettes. Questo esito può essere dovuto ad un legame con il genere dell'individuo ipotetico rappresentato nello scenario; tenendo conto di questo risultato, in seguito si proverà a considerare anche questa caratteristica nelle analisi. Viceversa, per quanto riguarda il secondo blocco di domande (Q19-Q21-Q22 e Q21bis-Q22bis-Q19bis), l'autovalutazione è condizionata dalle caratteristiche dell'individuo rispondente, specialmente dalla presenza o meno di problemi riscontrati con il servizio di *online banking*. D'altro canto, le vignettes associate, Q21, Q22, Q21bis e Q22bis, non presentano stime di parametri statisticamente significative per le quali le covariate risulterebbero influenzare tali valutazioni. Fanno eccezione però la presenza di un elevato titolo di studio (Laurea, Dottorato o Master) per la vignette Q21, e la presenza di figli nella composizione familiare e di problemi riscontrati con il servizio online, per la vignette Q22bis.

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	360,5810	1315,5870	0,2741	0,7840
Genere	76,7573	305,0110	0,2517	0,8013
Età: 35-54 anni	-69,7360	320,8301	-0,2174	0,8279
Età: 55 anni e più	50,3839	223,0537	0,2259	0,8213
Titolo di studio: alto	101,0092	387,9426	0,2604	0,7946
Lavoratore indipendente	-50,1078	325,2641	-0,1541	0,8776
Lavoratore dipendente	111,9077	542,4812	0,2063	0,8366
Altro (pensionato, disoccupato, etc.)	-25,6372	281,0533	-0,0912	0,9273
Dispositivo Touch	-68,0507	266,0382	-0,2558	0,7981
Riscontro problemi	-240,6499	873,5343	-0,2755	0,7829
Presenza di figli	-162,6569	606,2091	-0,2683	0,7885
Browser: Chrome	51,2564	220,5795	0,2324	0,8163
Browser: Explorer	51,7846	248,5464	0,2084	0,8350
Componente: <i>feeling</i>				
Costante	-0,9530	0,1849	-5,1555	0,0000
Genere	-0,0796	0,1234	-0,6453	0,5187
Età: 35-54 anni	-0,2113	0,1464	-1,4427	0,1491
Età: 55 anni e più	-0,2921	0,1816	-1,6081	0,1078
Titolo di studio: alto	-0,0787	0,1317	-0,5974	0,5503
Lavoratore indipendente	-0,3084	0,2545	-1,2117	0,2256
Lavoratore dipendente	-0,3224	0,2095	-1,5385	0,1239
Altro (pensionato, disoccupato, etc.)	-0,2618	0,2214	-1,1824	0,2371
Presenza di figli	-0,1201	0,1236	-0,9718	0,3311

Tabella 3.2: Stime dei parametri con modello CUB per per la domanda Q13 (autovalutazione sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	2,0486	1,6613	1,2331	0,2175
Genere	-0,2074	1,2730	-0,1630	0,8706
Età: 35-54 anni	-0,1081	1,6387	-0,0660	0,9474
Età: 55 anni e più	-0,9498	1,8532	-0,5125	0,6083
Titolo di studio: alto	-1,5038	1,4073	-1,0686	0,2852
Lavoratore indipendente	0,7214	2,1034	0,3430	0,7316
Lavoratore dipendente	15,5140	720,4108	0,0215	0,9828
Altro (pensionato, disoccupato, etc.)	0,9833	1,9294	0,5096	0,6103
Dispositivo Touch	14,1215	997,5519	0,0142	0,9887
Riscontro problemi	0,0953	1,7906	0,0532	0,9576
Presenza di figli	1,3259	1,3983	0,9482	0,3430
Browser: Chrome	2,0189	1,4895	1,3554	0,1753
Browser: Explorer	14,8252	1104,1120	0,0134	0,9893
Componente: <i>feeling</i>				
Costante	-1,7346	0,2672	-6,4918	0,0000
Genere	-0,4798	0,1396	-3,4360	0,0006
Età: 35-54 anni	0,0242	0,1617	0,1493	0,8813
Età: 55 anni e più	-0,0225	0,1999	-0,1126	0,9104
Titolo di studio: alto	0,0172	0,1418	0,1214	0,9034
Lavoratore indipendente	0,1095	0,3148	0,3478	0,7280
Lavoratore dipendente	0,2579	0,2823	0,9136	0,3609
Altro (pensionato, disoccupato, etc.)	0,1933	0,2999	0,6445	0,5193
Presenza di figli	0,0833	0,1306	0,6375	0,5238

Tabella 3.3: Stime dei parametri con modello CUB per per la domanda Q15 (vignette di Carlo sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	268,3967	456,1704	0,5884	0,5563
Genere	-12,8711	26,7514	-0,4811	0,6304
Età: 35-54 anni	-261,7378	445,2539	-0,5878	0,5566
Età: 55 anni e più	-193,3686	330,4952	-0,5851	0,5585
Titolo di studio: alto	-32,0276	61,4828	-0,5209	0,6024
Lavoratore indipendente	221,3308	383,5790	0,5770	0,5639
Lavoratore dipendente	36,4828	74,6956	0,4884	0,6253
Altro (pensionato, disoccupato, etc.)	249,4034	444,9613	0,5605	0,5751
Dispositivo Touch	-43,9120	81,2196	-0,5407	0,5887
Riscontro problemi	-71,1593	125,5039	-0,5670	0,5707
Presenza di figli	90,8432	155,4544	0,5844	0,5590
Browser: Chrome	-132,4254	228,9840	-0,5783	0,5630
Browser: Explorer	-37,0496	71,1368	-0,5208	0,6025
Componente: <i>feeling</i>				
Costante	-0,1692	0,1697	-0,9975	0,3185
Genere	-0,4221	0,1066	-3,9612	0,0001
Età: 35-54 anni	0,0436	0,1314	0,3315	0,7402
Età: 55 anni e più	-0,0019	0,1486	-0,0127	0,9899
Titolo di studio: alto	0,2728	0,1135	2,4027	0,0163
Lavoratore indipendente	0,0323	0,2182	0,1479	0,8824
Lavoratore dipendente	0,3972	0,1933	2,0550	0,0399
Altro (pensionato, disoccupato, etc.)	0,1225	0,2005	0,6109	0,5413
Presenza di figli	0,0580	0,1066	0,5440	0,5864

Tabella 3.4: Stime dei parametri con modello CUB per la domanda Q16 (vignette di Marina sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	191,3399	14453,6500	0,0132	0,9894
Genere	-18,1166	5004,5550	-0,0036	0,9971
Età: 35-54 anni	25,2056	21178,5600	0,0012	0,9991
Età: 55 anni e più	-65,3341	11304,9200	-0,0058	0,9954
Titolo di studio: alto	-87,5732	7549,1840	-0,0116	0,9907
Lavoratore indipendente	-29,8396	5639,8620	-0,0053	0,9958
Lavoratore dipendente	28,4069	14824,3700	0,0019	0,9985
Altro (pensionato, disoccupato, etc.)	-43,4022	6505,9940	-0,0067	0,9947
Dispositivo Touch	-69,0870	5104,9680	-0,0135	0,9892
Riscontro problemi	-44,9341	4906,0650	-0,0092	0,9927
Presenza di figli	20,6204	10453,1400	0,0020	0,9984
Browser: Chrome	52,2134	11010,9400	0,0047	0,9962
Browser: Explorer	76,0227	11771,4300	0,0065	0,9948
Componente: <i>feeling</i>				
Costante	-1,2749	0,1740	-7,3283	0,0000
Genere	0,0435	0,1240	0,3509	0,7257
Età: 35-54 anni	-0,0626	0,1494	-0,4192	0,6751
Età: 55 anni e più	-0,3318	0,1873	-1,7721	0,0764
Titolo di studio: alto	0,1287	0,1302	0,9892	0,3226
Lavoratore indipendente	-0,3281	0,2349	-1,3969	0,1624
Lavoratore dipendente	-0,4134	0,2013	-2,0535	0,0400
Altro (pensionato, disoccupato, etc.)	-0,1406	0,2123	-0,6623	0,5078
Presenza di figli	-0,1384	0,1251	-1,1063	0,2686
Riscontro problemi	0,4674	0,1258	3,7151	0,0002

Tabella 3.5: Stime dei parametri con modello CUB per la domanda Q19 (autovalutazione sulla semplicità di esecuzione delle operazioni bancarie online).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	2,5812	1,2778	2,0200	0,0434
Genere	-0,9657	0,7043	-1,3711	0,1704
Età: 35-54 anni	-2,9150	2,2049	-1,3220	0,1862
Età: 55 anni e più	-2,4515	2,2966	-1,0674	0,2858
Titolo di studio: alto	-0,1232	0,7974	-0,1545	0,8772
Lavoratore indipendente	3,5493	2,7976	1,2687	0,2045
Lavoratore dipendente	2,1682	2,4263	0,8936	0,3715
Altro (pensionato, disoccupato, etc.)	2,9447	2,6657	1,1047	0,2693
Dispositivo Touch	0,3896	1,2854	0,3031	0,7618
Riscontro problemi	0,1045	0,9110	0,1147	0,9087
Presenza di figli	-1,3038	0,9157	-1,4239	0,1545
Browser: Chrome	-0,1840	0,7473	-0,2463	0,8055
Browser: Explorer	11,2437	141,2069	0,0796	0,9365
Componente: <i>feeling</i>				
Costante	0,4724	0,1843	2,5628	0,0104
Genere	0,0452	0,1341	0,3373	0,7359
Età: 35-54 anni	0,2072	0,1548	1,3383	0,1808
Età: 55 anni e più	0,1027	0,1816	0,5654	0,5718
Titolo di studio: alto	0,4057	0,1517	2,6739	0,0075
Lavoratore indipendente	-0,1186	0,2395	-0,4953	0,6204
Lavoratore dipendente	0,1320	0,2140	0,6166	0,5375
Altro (pensionato, disoccupato, etc.)	-0,0420	0,2237	-0,1876	0,8512
Presenza di figli	0,0504	0,1526	0,3300	0,7414
Riscontro problemi	-0,0446	0,1353	-0,3295	0,7418

Tabella 3.6: Stime dei parametri con modello CUB per la domanda Q21 (vignette di Sofia sulla semplicità di esecuzione delle operazioni bancarie online).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	186,2994	20070,8000	0,0093	0,9926
Genere	90,8904	17147,4700	0,0053	0,9958
Età: 35-54 anni	8,7751	94075,7200	0,0001	0,9999
Età: 55 anni e più	-127,2426	13854,7100	-0,0092	0,9927
Titolo di studio: alto	6,7538	2514,9000	0,0027	0,9979
Lavoratore indipendente	169,3402	19782,3500	0,0086	0,9932
Lavoratore dipendente	47,7956	7046,7850	0,0068	0,9946
Altro (pensionato, disoccupato, etc.)	176,4847	21033,0800	0,0084	0,9933
Dispositivo Touch	-40,2149	5239,3720	-0,0077	0,9939
Riscontro problemi	-129,5160	14930,2700	-0,0087	0,9931
Presenza di figli	39,3898	6149,5770	0,0064	0,9949
Browser: Chrome	-87,8624	9860,0350	-0,0089	0,9929
Browser: Explorer	7,1261	244630	0,0000	1,0000
Componente: <i>feeling</i>				
Costante	-1,5204	0,1981	-7,6758	0,0000
Genere	-0,0753	0,1364	-0,5520	0,5810
Età: 35-54 anni	-0,1217	0,1646	-0,7392	0,4598
Età: 55 anni e più	-0,3719	0,2023	-1,8385	0,0660
Titolo di studio: alto	-0,0826	0,1458	-0,5662	0,5713
Lavoratore indipendente	-0,0938	0,2611	-0,3593	0,7194
Lavoratore dipendente	-0,2233	0,2298	-0,9718	0,3312
Altro (pensionato, disoccupato, etc.)	0,0906	0,2370	0,3822	0,7023
Presenza di figli	0,0414	0,1356	0,3056	0,7599
Riscontro problemi	0,0589	0,1442	0,4088	0,6827

Tabella 3.7: Stime dei parametri con modello CUB per la domanda Q22 (vignette di Pietro sulla semplicità di esecuzione delle operazioni bancarie online).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	488,2095	1008,2450	0,4842	0,6282
Genere	-58,9109	131,1304	-0,4493	0,6533
Età: 35-54 anni	-279,9547	580,0352	-0,4827	0,6293
Età: 55 anni e più	-292,3176	607,9232	-0,4809	0,6306
Titolo di studio: alto	-59,0043	132,2881	-0,4460	0,6556
Lavoratore indipendente	87,5798	200,2813	0,4373	0,6619
Lavoratore dipendente	164,0511	353,7355	0,4638	0,6428
Altro (pensionato, disoccupato, etc.)	132,7752	283,9305	0,4676	0,6400
Dispositivo Touch	-144,0096	303,9316	-0,4738	0,6356
Riscontro problemi	-159,5163	348,9740	-0,4571	0,6476
Presenza di figli	-18,4291	113,2677	-0,1627	0,8708
Browser: Chrome	-57,4884	157,4449	-0,3651	0,7150
Browser: Explorer	9,5855	23220,6000	0,0004	0,9997
Componente: <i>feeling</i>				
Costante	-1,6362	0,1905	-8,5884	0,0000
Genere	0,1373	0,1240	1,1075	0,2681
Età: 35-54 anni	-0,2360	0,1477	-1,5972	0,1102
Età: 55 anni e più	-0,1937	0,1840	-1,0526	0,2925
Lavoratore indipendente	0,2835	0,2537	1,1178	0,2636
Lavoratore dipendente	0,2263	0,2173	1,0412	0,2978
Altro (pensionato, disoccupato, etc.)	0,2967	0,2310	1,2845	0,1990
Presenza di figli	-0,2255	0,1263	-1,7847	0,0743

Tabella 3.8: Stime dei parametri con modello CUB per per la domanda Q13bis (autovalutazione sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	348,3540	997,7986	0,3491	0,7270
Genere	-93,4916	293,1054	-0,3190	0,7497
Età: 35-54 anni	16,5446	403,0215	0,0411	0,9673
Età: 55 anni e più	-149,8896	561,3672	-0,2670	0,7895
Titolo di studio: alto	-31,2424	130,7514	-0,2389	0,8112
Lavoratore indipendente	-61,7902	608,1671	-0,1016	0,9191
Lavoratore dipendente	126,0318	773,2817	0,1630	0,8705
Altro (pensionato, disoccupato, etc.)	44,1350	669,2970	0,0659	0,9474
Dispositivo Touch	-214,6446	620,5367	-0,3459	0,7294
Riscontro problemi	-56,1579	217,3521	-0,2584	0,7961
Presenza di figli	36,2370	222,5379	0,1628	0,8706
Browser: Chrome	-41,2126	151,2352	-0,2725	0,7852
Browser: Explorer	-70,3794	334,2717	-0,2106	0,8332
Componente: <i>feeling</i>				
Costante	-1,8516	0,2174	-8,5192	0,0000
Genere	-0,2627	0,1280	-2,0527	0,0401
Età: 35-54 anni	0,0381	0,1491	0,2554	0,7984
Età: 55 anni e più	-0,2078	0,1958	-1,0615	0,2885
Lavoratore indipendente	0,2392	0,2842	0,8416	0,4000
Lavoratore dipendente	0,5899	0,2436	2,4215	0,0155
Altro (pensionato, disoccupato, etc.)	0,6046	0,2593	2,3318	0,0197
Presenza di figli	-0,3447	0,1284	-2,6848	0,0073

Tabella 3.9: Stime dei parametri con modello CUB per per la domanda Q15bis (vignette di Carlo sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	71,9945	208,5003	0,3453	0,7299
Genere	-7,8332	11,3977	-0,6873	0,4919
Età: 35-54 anni	-77,7494	644,7824	-0,1206	0,9040
Età: 55 anni e più	-85,5359	646,0797	-0,1324	0,8947
Titolo di studio: alto	11,1668	9,5334	1,1713	0,2415
Lavoratore indipendente	19,6154	609,7403	0,0322	0,9743
Lavoratore dipendente	28,4512	610,3069	0,0466	0,9628
Altro (pensionato, disoccupato, etc.)	26,3073	610,3157	0,0431	0,9656
Dispositivo Touch	-34,1445	36,4813	-0,9359	0,3493
Riscontro problemi	21,3382	20,5022	1,0408	0,2980
Presenza di figli	-20,7927	20,6951	-1,0047	0,3150
Browser: Chrome	-3,5354	10,7030	-0,3303	0,7412
Browser: Explorer	-14,3317	20,6966	-0,6925	0,4886
Componente: <i>feeling</i>				
Costante	-0,0595	0,1462	-0,4067	0,6842
Genere	-0,1759	0,1106	-1,5908	0,1117
Età: 35-54 anni	-0,0936	0,1338	-0,6994	0,4843
Età: 55 anni e più	-0,2161	0,1670	-1,2941	0,1956
Lavoratore indipendente	0,1317	0,2025	0,6504	0,5154
Lavoratore dipendente	0,2149	0,1712	1,2548	0,2096
Altro (pensionato, disoccupato, etc.)	0,0834	0,1874	0,4450	0,6563
Presenza di figli	-0,0545	0,1293	-0,4210	0,6738

Tabella 3.10: Stime dei parametri con modello CUB per per la domanda Q16bis (vignette di Marina sulla facilità di navigazione).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	41,7074	4264,9610	0,0098	0,9922
Genere	13,0615	207,4314	0,0630	0,9498
Età: 35-54 anni	-23,0142	517,6218	-0,0445	0,9645
Età: 55 anni e più	6,5468	3641,1700	0,0018	0,9986
Titolo di studio: alto	22,97189	439,1890	0,0523	0,9583
Lavoratore indipendente	17,5862	4280,6520	0,0041	0,9967
Lavoratore dipendente	8,5720	4267,3900	0,0020	0,9984
Altro (pensionato, disoccupato, etc.)	-5,5868	4260,9200	-0,0013	0,9990
Riscontro problemi	-1,0896	1,9534	-0,5578	0,5770
Presenza di figli	-24,4621	456,4513	-0,0536	0,9573
Browser: Chrome	10,8916	207,4209	0,0525	0,9581
Browser: Explorer	-0,2426	1,9001	-0,1277	0,8984
Componente: <i>feeling</i>				
Costante	-1,1411	0,2008	-5,6814	0,0000
Genere	-0,2272	0,1209	-1,8798	0,0601
Età: 35-54 anni	-0,3252	0,1410	-2,3063	0,0211
Età: 55 anni e più	-0,3007	0,1794	-1,6759	0,0938
Titolo di studio: alto	0,0962	0,1222	0,7874	0,4310
Lavoratore indipendente	-0,0703	0,2459	-0,2859	0,7750
Lavoratore dipendente	-0,1688	0,2160	-0,7815	0,4345
Altro (pensionato, disoccupato, etc.)	-0,1925	0,2319	-0,8300	0,4066
Presenza di figli	-0,0300	0,1191	-0,2521	0,8010
Riscontro problemi	0,4617	0,1221	3,7807	0,0002

Tabella 3.11: Stime dei parametri con modello CUB per la domanda Q19 (autovalutazione sulla semplicità di esecuzione delle operazioni bancarie online).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	1,6626	1,8307	0,9082	0,3638
Genere	-2,1322	1,4471	-1,4735	0,1406
Età: 35-54 anni	-1,0996	1,2113	-0,9077	0,3640
Età: 55 anni e più	2,1110	1,9663	1,0736	0,2830
Titolo di studio: alto	17,6734	570,2169	0,0310	0,9753
Lavoratore indipendente	-2,1572	2,2421	-0,9622	0,3360
Lavoratore dipendente	1,9172	1,6510	1,1613	0,2455
Altro (pensionato, disoccupato, etc.)	13,5753	337,1300	0,0403	0,9679
Riscontro problemi	0,7524	1,3282	0,5665	0,5711
Presenza di figli	-0,2667	0,9668	-0,2759	0,7826
Browser: Chrome	0,9703	1,2251	0,7920	0,4284
Browser: Explorer	0,4874	1,3022	0,3743	0,7082
Componente: <i>feeling</i>				
Costante	0,6961	0,2417	2,8796	0,0040
Genere	0,0349	0,1142	0,3056	0,7599
Età: 35-54 anni	-0,0666	0,1364	-0,4881	0,6255
Età: 55 anni e più	0,1532	0,1593	0,9619	0,3361
Titolo di studio: alto	-0,1602	0,1160	-1,3803	0,1675
Lavoratore indipendente	0,3395	0,2932	1,1577	0,2470
Lavoratore dipendente	0,2346	0,2551	0,9199	0,3576
Altro (pensionato, disoccupato, etc.)	-0,2574	0,2593	-0,9928	0,3208
Presenza di figli	-0,0078	0,1103	-0,0708	0,9436
Riscontro problemi	0,0127	0,1210	0,1052	0,9162

Tabella 3.12: Stime dei parametri con modello CUB per la domanda Q21bis (vignette di Sofia sulla semplicità di esecuzione delle operazioni bancarie online).

Parametri	Stime puntuali	S.E.	Wald test	P-value
Componente: <i>incertezza</i>				
Costante	30,5374	2674,8140	0,0114	0,9909
Genere	16,2025	1504,2930	0,0108	0,9914
Età: 35-54 anni	-15,1843	1960,9060	-0,0077	0,9938
Età: 55 anni e più	-15,6305	1960,9060	-0,0080	0,9936
Titolo di studio: alto	11,2968	208,4084	0,0542	0,9568
Lavoratore indipendente	-0,5638	3156,9750	-0,0002	0,9999
Lavoratore dipendente	0,6073	3156,9750	0,0002	0,9998
Altro (pensionato, disoccupato, etc.)	17,8174	5448,7330	0,0033	0,9974
Riscontro problemi	15,9028	1781,0940	0,0089	0,9929
Presenza di figli	-14,4949	736,7853	-0,0197	0,9843
Browser: Chrome	0,2834	1,5084	0,1879	0,8509
Browser: Explorer	15,6104	1784,9470	0,0088	0,9930
Componente: <i>feeling</i>				
Costante	-1,7109	0,2359	-7,2537	0,0000
Genere	0,0663	0,1388	0,4780	0,6327
Età: 35-54 anni	-0,1156	0,1631	-0,7085	0,4786
Età: 55 anni e più	-0,0375	0,2074	-0,1808	0,8565
Titolo di studio: alto	0,0631	0,1422	0,4435	0,6574
Lavoratore indipendente	-0,0042	0,2879	-0,0146	0,9884
Lavoratore dipendente	-0,1524	0,2517	-0,6055	0,5449
Altro (pensionato, disoccupato, etc.)	-0,1010	0,2665	-0,3790	0,7047
Presenza di figli	-0,3795	0,1407	-2,6974	0,0070
Riscontro problemi	0,3618	0,1415	2,5570	0,0106

Tabella 3.13: Stime dei parametri con modello CUB per la domanda Q22bis (vignette di Pietro sulla semplicità di esecuzione delle operazioni bancarie online).

3.3.1 AVCUB e test RC

Per testare l'assunzione di RC con il metodo descritto nel paragrafo 2.6, si è utilizzata la funzione AVCUB costruita proprio per questo scopo. Viene calcolata la *log-verosimiglianza* di due modelli: quello con i parametri della componente di incertezza vincolati e quello senza alcun tipo di vincolo. Il test del rapporto di verosimiglianza verifica se queste due funzioni possano essere ritenute uguali, cosicché la RC possa essere validata. Dunque, si è proceduto testando a coppie l'autovalutazione con ciascuna delle vignette associate, ottenendo i risultati riportati in Tabella 3.14.

	Valore LR	DF	p-value
Q13 & Q15	27,53	12	0,0065
Q13 & Q16	155,03	12	< 2.2E-16
Q13bis & Q15bis	20,76	12	0,0541
Q13bis & Q16bis	175,38	9	< 2.2E-16
Q19 & Q21	83,10	12	1,05E-09
Q19 & Q22	15,70	12	0,2053
Q19bis & Q21bis	66,73	11	5,10E-07
Q19bis & Q22bis	15,48	11	0,1157

Tabella 3.14: LR test per la convalida dell'assunzione di RC.

Per quanto riguarda i risultati relativi all'autovalutazione rappresentata dalla variabile Q13 è possibile affermare che l'assunzione di RC dipende fortemente dalla vignette associata. Infatti, se si valuta l'LR test tra Q13 e Q15 i risultati sono molto diversi da quelli ottenuti tra Q13 e Q16. Il giudizio dei rispondenti e la relativa scelta della scala di valutazione sono in forte relazione con la chiarezza espositiva della vignette. Nel caso di Q16, come è già stato detto, tale chiarezza non sembra essere sufficiente e gli intervistati non utilizzano la stessa scala adottata per l'autovalutazione, dunque l'assunzione di RC non è valida. In riferimento a Q15, invece, è necessario, affinché la RC valga, che sia proposta al rispondente prima la vignette e poi successivamente l'autovalutazione (Gruppo 3), anche se ad un livello di significatività del 5%, ma non del 10%.

Per il blocco di risposte riferito all'autovalutazione rappresentata da Q19 il di-

scorso è simile al precedente: anche in questo caso si ha una vignette più "standard" (Q22) e una vignette "non standard" (Q21). Pertanto l'assunzione di RC vale solo quando l'autovalutazione è associata alla vignette Q22. Rispetto a Q15, dove l'ordine tra vignette e autovalutazione influiva sull'utilizzo della scala di risposta degli intervistati, con lo scenario descritto nella vignette Q22 questa distinzione non avviene, quindi, essa non fornisce alcun ulteriore chiarimento del significato della domanda che vuole valutare la soddisfazione dell'intervistato (Hopkins & King, 2010).

3.3.2 Test per l'assunzione di VE

La validazione dell'assunzione di VE è basata sul fatto che non dovrebbero esistere differenze nella percezione delle vignettes tra i rispondenti. Nel contesto dei modelli CUB essa si traduce nel testare la significatività delle stime dei parametri associati alle covariate della componente di *feeling* delle vignettes. Se questa assunzione non valesse la percezione delle vignettes dipenderebbe da caratteristiche individuali, impedendo così la correzione del DIF.

Per costruire il test che verifichi la validità della VE si è proceduto sfruttando nuovamente l'LR test. In questo caso si è stimato prima un modello CUB per la vignette, considerando tutte le covariate per la componente di *feeling* e se ne è calcolata la *log-verosimiglianza*. Successivamente, per la stessa vignette si è stimato un modello CUB con solo l'intercetta per la parte di *feeling* e le stesse variabili del modello precedente per la parte di incertezza, ottenendo una seconda *log-verosimiglianza*. A questo punto con l'LR test se ne è verificata l'uguaglianza; in caso affermativo questo confermerebbe l'indipendenza tra le caratteristiche individuali (osservabili) e l'interpretazione della vignette e sarebbe una convalida della corretta assunzione di VE. I risultati relativi ai dati dell'*online banking* sono riportati in Tabella 3.15. Considerando la vignette Q15, gli individui rispondono senza essere influenzati dalle caratteristiche individuali solo nel caso in cui essa sia proposta dopo l'autovalutazione, anche se all'1% di livello di significatività, ma non al 5%. Quando nella

precedente sezione si erano riportate le stime del modello CUB su questa vignette per entrambi gli ordinamenti proposti (Tabelle 3.3 e 3.9), le significatività delle stime dei parametri indicavano già che vi fosse differenza tra quando la vignette era proposta prima dell'autovalutazione o dopo: infatti si aveva un solo parametro significativo per la componente di *feeling* considerando Q15, invece risultavano ben quattro per Q15bis. Viceversa, per Q16 la VE vale (fortemente) solo nel caso del Gruppo 3 (Q16bis); anche in questo caso si ha il parallelismo con quanto ottenuto con i modelli CUB della sezione precedente, per i quali il modello su Q16bis non aveva alcun parametro significativo per la componente di *feeling* (Tabella 3.10), mentre chi rispondeva a Q16 era influenzato sia dal genere che dal titolo di studio (Tabella 3.4). Per quanto riguarda Q21 e Q22 i migliori risultati, in termini di validazione della VE, sono ottenuti quando entrambe le vignettes sono poste dopo l'autovalutazione, cioè quando le covariate hanno meno influenzato i rispondenti ad esprimere la loro valutazione sulla soddisfazione degli individui delle vignettes.

	Valore LR	DF	p-value
Q15	16,33	8	0,0379
Q16	42,45	8	0.0011
Q15bis	39,01	7	0.0020
Q16bis	8,08	8	0.4253
Q21	12,32	9	0,1959
Q22	8,20	9	0,5138
Q21bis	22,98	9	0,0062
Q22bis	17,08	9	0,0475

Tabella 3.15: Risultati dell'LR test per l'assunzione di VE.

3.3.3 Genere della vignette

Facendo riferimento al lavoro di Kapteyn et al. (2007), nel quale si affrontava la tematica della scelta del genere dell'individuo ipotetico della vignette, si valuta la validità della RC e della VE considerando, per ciascuna vignette, solo rispondenti con lo stesso genere della persona descritta in essa e successivamente con genere opposto. L'obbiettivo è quello di riconoscere reazioni diverse dei rispondenti in base al genere della vignette e capire se la scala di risposta utilizzata nel giudicare gli individui ipotetici, sia la stessa usata per sé stessi. Inoltre, si vuole verificare se vi sia un utilizzo di scale differenti quando i rispondenti devono valutare un uomo o una donna; questo non sarebbe altro che una violazione della VE e potrebbe essere legato proprio al genere dell'individuo rispondente; infatti, si è visto nella precedente sezione che il genere influenzava il giudizio dei rispondenti sulla soddisfazione percepita degli individui ipotetici, comportando a volte anche un rifiuto dell'assunzione. Dunque, si vuole indagare se, in termini di accettazione delle assunzioni, le conclusioni siano diverse tra il caso in cui una persona valuta un individuo dello stesso genere o con genere opposto rispetto a quando il genere della vignette non era stato considerato.

E' d'obbligo sottolineare che questa procedura ha praticamente dimezzato il dataset con cui si stimavano i modelli, per cui non è stato possibile introdurre tutte le covariate di interesse nella loro costruzione.

I risultati del test sulla RC relativi al blocco di domande Q13-Q15-Q16 sono riportati in Tabella 3.16, mentre quelli relativi al blocco Q19-Q21-Q22 in Tabella 3.17. Si nota che quando un rispondente si trova a valutare una persona del genere opposto è più propenso a valutarla oggettivamente, mantenendo la stessa scala usata per valutare sé stesso. Viceversa, quando valuta una persona dello stesso genere la sua oggettività diminuisce. I valori del test LR, comunque, forniscono risultati in favore dell'assunzione di RC in entrambi i casi per le vignettes Q15 e Q22. Per le vignettes "non standard", invece, la divisione rispetto al genere dell'individuo

descritto nello scenario migliora i risultati ottenuti in precedenza, in particolare, si ha che nella vignette Q21, quando è proposta prima dell'autovalutazione, l'assunzione di RC è valida se il rispondente è del genere opposto.

In conclusione, quando è introdotto il genere dell'individuo ipotetico della vignette per la convalida dell'assunzione di RC, l'importanza dell'ordine di esposizione tra l'autovalutazione e le vignettes tende a svanire e quando l'assunzione è valida in un gruppo, lo è anche nell'altro.

	Genere	Valore LR	DF	p-value
Q13 & Q15	Uguale	19,16	9	0,0239
	Opposto	16,84	9	0,0513
Q13 & Q16	Uguale	104,22	9	< 2.2E-16
	Opposto	88,49	9	3,27E-12
Q13bis & Q15bis	Uguale	12,15	8	0,1448
	Opposto	13,20	8	0,1052
Q13bis & Q16bis	Uguale	109,46	8	< 2.2E-16
	Opposto	111,51	8	< 2.2E-16

Tabella 3.16: LR test per la convalida della RC in relazione al genere delle vignettes Q15 (Uomo) e Q16 (Donna).

	Genere	Valore LR	DF	p-value
Q19 & Q21	Uguale	84,43	9	2,12E-11
	Opposto	44,63	9	0,0011
Q19 & Q22	Uguale	11,59	8	0,1707
	Opposto	9,23	9	0,4167
Q19bis & Q21bis	Uguale	46,43	8	1,97E-04
	Opposto	15,47	8	0,0507
Q19bis & Q22bis	Uguale	12,26	9	0,1402
	Opposto	16,91	9	0,0501

Tabella 3.17: LR test per la convalida della RC in relazione al genere delle vignettes Q21 (Donna) e Q22 (Uomo).

Per quanto riguarda l'assunzione di VE i risultati del test sono riportati in Tabella 3.18. Si osserva che in certi casi la suddivisione in base al genere ha permesso di convalidare l'assunzione di VE, dove prima veniva rifiutata (Tabella 3.15). Ad esempio, con Q15 l'ipotesi di VE è accettata in entrambi i gruppi di rispondenti, sia nel caso in cui il genere tra rispondente e vignette è lo stesso sia quando è

diverso, mentre prima la stessa assunzione per questa vignette non era stata validata. Invece, Q16 ed anche Q16bis convalidano l'ipotesi di VE solo quando sono poste a rispondenti di genere opposto alla vignette, ossia agli uomini. Infine, per l'ultimo blocco di domande, come si è già riscontrato in precedenza, Q21 è più problematica di Q22 che invece non ha alcun dubbio sulla validazione sia della RC che della VE.

Concludendo, il genere della vignette è sicuramente un fattore molto importante per la costruzione degli scenari da descrivere, quindi è necessario tenerne conto quando si analizzano i dati a disposizione. Infatti, potrebbe non essere un caso che le due vignettes più problematiche, Q16 e Q21, descrivano entrambe un individuo femminile.

	Genere	Valore LR	DF	p-value
Q15	Uguale	5,84	7	0,5585
	Opposto	4,57	7	0,7129
Q16	Uguale	20,15	7	0,0053
	Opposto	12,01	7	0,1003
Q15bis	Uguale	13,14	7	0,0687
	Opposto	10,22	7	0,1765
Q16bis	Uguale	24,68	7	0,0009
	Opposto	10,42	7	0,1661
Q21	Uguale	25,07	7	0,0656
	Opposto	17,33	7	0,0217
Q22	Uguale	7,95	7	0,4843
	Opposto	5,00	7	0,4036
Q21bis	Uguale	13,28	7	0,0007
	Opposto	16,40	7	0,0154
Q22bis	Uguale	6,49	7	0,3367
	Opposto	7,25	7	0,6605

Tabella 3.18: LR test per la convalida della VE in relazione al genere delle vignettes Q15(Uomo), Q16(Donna), Q21(Donna) e Q22(Uomo).

3.3.4 Ordine tra vignettes

Data la rilevanza dell'ordine delle domande nel questionario riscontrata per la validazione delle assunzioni di RC e di VE, si è approfondita l'analisi investigando se anche l'ordine tra le stesse vignettes potesse influire sulle scale di risposta degli individui. Si sono considerate le risposte alle vignettes in base all'ordine in cui erano state collocate nel questionario e per ciascuna sono state verificate le due assunzioni tramite il metodo AVCUB. Dunque, fissato il Gruppo di appartenenza dei rispondenti, si è verificata per prima la RC per ogni coppia di autovalutazione-vignette. I risultati sono riportati nelle Tabelle 3.19 e 3.20. La vignette Q15 ha un diverso comportamento in base all'ordine in cui è posta rispetto a Q16: quando essa segue Q16 la RC cade, indipendentemente dalla posizione dell'autovalutazione. Invece, la vignette Q16 che fino ad ora non aveva mai validato la RC trova un risultato positivo quando è posta come prima domanda del blocco, sia prima di Q15 che dell'autovalutazione. Per il blocco costituito da Q19-Q21-Q22 le conclusioni sono pressoché le stesse in qualsiasi ordinamento venga considerato; infatti, sia nel caso in cui sia proposta prima l'autovalutazione e poi le vignettes oppure il viceversa o ancora sia invertito l'ordine tra Q21 e Q22, la RC per Q21 non è mai validata, mentre quella per Q22 lo è sempre.

	AVCUB	Valore LR	DF	p-value
Gruppo 1 (Q13-Q15-Q16)	Q13-Q15	11,69	10	0,3063
	Q13-Q16	54,55	9	1,49E-05
Gruppo 2 (Q13-Q16-Q15)	Q13-Q15	26,70	10	0,0029
	Q13-Q16	87,31	9	5,62E-12
Gruppo 3 (Q15-Q16-Q13)	Q13bis-Q15bis	8,72	9	0,4640
	Q13bis-Q16bis	151,58	8	< 2.2E-16
Gruppo 4 (Q16-Q15-Q13)	Q13bis-Q15bis	22,70	9	0,0069
	Q13bis-Q16bis	17,14	8	0,0287

Tabella 3.19: Test LR per la verifica dell'assunzione di RC per il blocco di domande Q13-Q15-Q16 per i 4 gruppi di intervistati.

	AVCUB	Valore LR	DF	p-value
Gruppo 1 (Q19-Q21-Q22)	Q19-Q21	74,704	9	1,81E-09
	Q19-Q22	13,308	9	0,1492
Gruppo 2 (Q19-Q22-Q21)	Q19-Q21	93,348	9	3,46E-16
	Q19-Q22	1,5384	9	0,9969
Gruppo 3 (Q21-Q22-Q19)	Q19bis-Q21bis	53,487	8	8.68e-09
	Q19bis-Q22bis	4,9431	8	0,7636
Gruppo 4 (Q22-Q21-Q19)	Q19bis-Q21bis	40,487	8	2,60E-06
	Q19bis-Q22bis	10,173	8	0,2531

Tabella 3.20: Test LR per la verifica dell'assunzione di RC per il blocco di domande Q19-Q21-Q22 per i 4 gruppi di intervistati.

Per quanto concerne la verifica dell'assunzione di VE tramite il test LR i risultati ottenuti sui dati dell'*online banking* sono riportati in Tabella 3.21. L'assunzione di VE per le prime vignettes, Q15 e Q16, sembra non valere sempre, ma dipende fortemente da come è scelto l'ordine di esposizione. Viceversa, nel secondo blocco quest'ipotesi sembra più solida e i test avvalorano il fatto che le caratteristiche degli individui non interferiscano con l'interpretazione delle vignettes. La maggior parte dei rispondenti del questionario sull'*online banking* risponde ad entrambi i blocchi di domande e solo una piccolissima porzione valuta solo la propria soddisfazione a Q19. Quindi, coloro che si trovano a rispondere alle vignettes Q21 e Q22 hanno già risposto a Q15 e Q16 in precedenza e sembra che da queste abbiano capito l'intento della tecnica delle Anchoring Vignettes, comportando una maggiore obbiettività nelle risposte successive.

Da questi risultati potrebbe essere possibile evincere che solo le ultime vignettes di un questionario possano essere utilizzate per l'eliminazione del DIF, dato che esse sono quelle per cui le assunzioni di base delle Anchoring Vignettes sono validate. Le prime vignettes, invece, potrebbero servire all'intervistato per capire meglio questo approccio, ma non potrebbero essere usate nell'eliminazione del DIF, dato che le due assunzioni non sono sempre valide.

	Clients	LR Value	DF	p-value
Q15	Gruppo 1	9,63	8	0,2922
	Gruppo 2	15,77	8	0,0489
	Gruppo 3	17,91	8	0,0219
	Gruppo 4	20,48	8	0,0087
Q16	Gruppo 1	23,07	8	0,0033
	Gruppo 2	20,93	8	0,0073
	Gruppo 3	19,46	8	0,0126
	Gruppo 4	17,92	8	0,0218
Q21	Gruppo 1	5,46	8	0,7071
	Gruppo 2	5,30	8	0,7255
	Gruppo 3	14,88	8	0,0616
	Gruppo 4	12,32	8	0,1375
Q22	Gruppo 1	13,00	8	0,1120
	Gruppo 2	3,94	8	0,8626
	Gruppo 3	6,54	8	0,5870
	Gruppo 4	8,19	8	0,4152

Tabella 3.21: Test LR per la verifica dell'assunzione di VE per le vignettes Q15, Q16, Q21 e Q22, disaggregate per i 4 gruppi di intervistati.

3.3.5 *Online banking: seconda wave*

Dopo quattro mesi, lo stesso questionario è stato riproposto sia ad una parte dei clienti che aveva partecipato già alla prima *wave*, sia a dei nuovi clienti. In questo modo si è tentato di individuare un comportamento diverso tra questi due gruppi, cercando di capire se i clienti che ripetevano il questionario rispondessero diversamente, ricordandosi della loro esperienza pregressa.

In questo lavoro, per sintesi, non sono riportate tutte le analisi a riguardo, ma solo le più significative.

I risultati della seconda *wave* non hanno fornito ulteriori dimostrazioni dei risultati già ottenuti nella prima: per le vignettes "standard" infatti la RC continua a valere, mentre per le vignettes Q16 e Q21 no. I clienti che già avevano risposto alle stesse domande nei mesi precedenti non sembrano avere comportamenti differenti, bensì continuano a rispondere con scale diverse, rispetto a quelle adottate nell'autovalutazione, a queste due vignettes (Tabelle 3.22 e 3.23). Questo risultato è sintomo del fatto che per le vignettes "non standard" ci sia un problema di formulazione

e di costruzione dello scenario rappresentato. Sarebbe opportuno proporre nuove vignettes, in sostituzione a queste, che meglio descrivano il soggetto ipotetico da valutare.

	Clienti	Valore LR	DF	p-value
Q13 & Q15	Vecchi	6,52	8	0,5889
	Nuovi	9,65	8	0,2906
Q13 & Q16	Vecchi	77,89	8	1,30E-10
	Nuovi	101.5	8	< 2.2E-16
Q13bis & Q15bis	Vecchi	8,11	8	0,4225
	Nuovi	10,00	8	0,2652
Q13bis & Q16bis	Vecchi	90,62	8	3,48E-13
	Nuovi	126.82	8	< 2.2E-16

Tabella 3.22: LR test per la convalida della RC per Q15 e Q16 rispetto ai clienti vecchi o nuovi della seconda *wave*.

	Clienti	Valore LR	DF	p-value
Q19 & Q21	Vecchi	37,23	8	1,04E-02
	Nuovi	59,40	8	6,11E-07
Q19 & Q22	Vecchi	12,02	8	0,1501
	Nuovi	9,08	8	0,3358
Q19bis & Q21bis	Vecchi	45,83	7	9,45E-05
	Nuovi	103.88	7	< 2.2e-16
Q19bis & Q22bis	Vecchi	10,02	7	0,1877
	Nuovi	7,46	7	0,3825

Tabella 3.23: LR test per la convalida della RC per Q21 e Q22 rispetto ai clienti vecchi o nuovi della seconda *wave*.

3.4 Modello *Chopit*

L'approccio standard utilizzato quando si è in presenza di dati raccolti mediante vignettes è quello dei modelli per dati ordinali, *Ordinal probit* o *Chopit*. Essi sono stati presentati nel secondo capitolo, dove è stata discussa la possibilità di testare le assunzioni di RC e VE tramite dei test costruiti *ad hoc*. Tra le condizioni per l'applicazione di questi test vi è la presenza di un indicatore oggettivo che fa riferimento all'entità che si sta valutando. Nel caso dell'*online banking*, questo indicatore

oggettivo potrebbe essere riconosciuto nelle due domande poste prima delle autovalutazioni durante il questionario. Esse riguardano il livello di aspettative precedenti all'attivazione del servizio e il livello di aspettative corrisposte in seguito all'utilizzo. La differenza tra queste due quantità fornisce il livello di soddisfazione del cliente sul servizio; infatti, per esempio, se tale differenza fosse negativa significherebbe che le aspettative del cliente sono state deluse e il servizio non è stato soddisfacente tanto quanto si sarebbe aspettato. Viceversa, un valore superiore a zero indicherebbe che l'esperienza con l'*online banking* è stata superiore alle proprie aspettative.

In Figura 3.10 è riportata la distribuzione dell'indicatore oggettivo per il blocco di domande Q13-Q15-Q16, dove è evidente che le aspettative dei clienti siano state prevalentemente poco soddisfatte. Invece, come mostrato in Figura 3.11, per il blocco Q19-Q21-Q22 le aspettative per la semplicità delle operazioni bancarie sono state più soddisfatte.

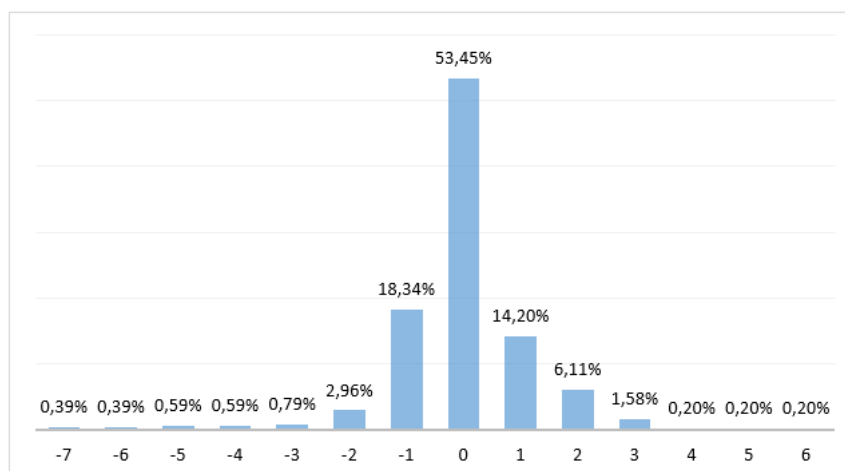


Figura 3.10: Distribuzione indicatore oggettivo per il blocco di domande Q13-Q15-Q16.

L'obiettivo dell'introduzione di questa metodologia era quello di dare conferma dei risultati ottenuti con l'approccio CUB riguardo i test delle assunzioni di RC e VE. Infatti se anche con il modello *Chopit*, attraverso l'ausilio dell'indicatore oggettivo, fosse risultato possibile accettare le assunzioni solo nel caso delle vignettes Q15 e Q22, rifiutandole per Q16 e Q21, allora la RC, e di conseguenza la VE, risulterebbero

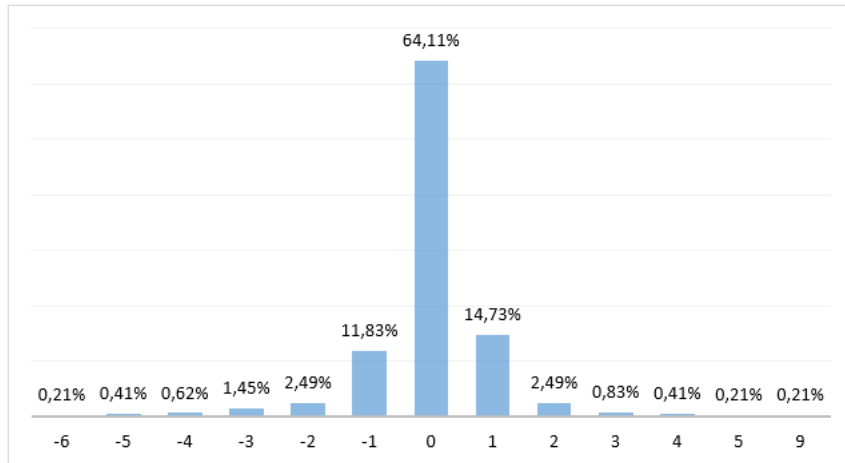


Figura 3.11: Distribuzione indicatore oggettivo per il blocco di domande Q19-Q21-Q22.

legate solo ad un aspetto espositivo delle vignettes. In altre parole, l'utilizzo di scale di risposta diverse nei rispondenti sarebbe vincolato ad una poco chiara esposizione dello scenario presentato nella vignette. Diversamente, il modello *Chopit* poteva essere sfruttato anche per la verifica delle assunzioni nei diversi gruppi di rispondenti, per avere riscontro del fatto che l'ordine influisse sull'utilizzo delle scale di risposta.

Purtroppo con il dataset dell'*online banking*, per via della scarsa numerosità di alcune categorie di risposta (in particolare quelle di tipo negativo), è stato impossibile riuscire a stimare un modello con covariate a causa di problemi di convergenza. Questa problematica ha reso l'introduzione dei modelli CUB ancor più indispensabile per questo caso di studio, dato che con le metodologie standard sarebbe stata impossibile qualsiasi applicazione.

Per validare i risultati ottenuti con i modelli CUB allora si è proceduto con l'analizzare un altro insieme di dati e applicare anche ad esso la nuova metodologia introdotta per testarne le assunzioni di RC e VE.

3.5 Ulteriori applicazioni dell'approccio AVCUB

In questo paragrafo si esplorerà un ulteriore insieme di dati che sono stati estratti dal LISS (*Longitudinal Internet Studies for the Social sciences*) panel, che fa parte di un progetto intitolato MESS (Measurement and Experimentation in the Social Sciences), il quale mira a raccogliere dati su una vasta gamma di argomenti, come la famiglia, la situazione economica, il lavoro, la scuola, la salute, la religione, l'etnia e la politica. In particolare, l'attenzione di questa tesi sarà volta verso il modulo denominato VECS (*Vignette Evaluation of Customer Satisfaction*) che è stato costruito per misurare la soddisfazione dei clienti, sfruttando la tecnica delle vignettes, in relazione all'acquisto di tre differenti prodotti: Computer portatile, Smartphone e TV LCD.

Il questionario, riportato nell'Appendice C, è strutturato nel modo seguente: innanzitutto, alle persone viene chiesto se hanno acquistato uno dei tre prodotti sopra elencati negli ultimi 6 o 9 mesi, e se hanno riscontrato problemi con tale prodotto. A coloro che hanno effettivamente effettuato l'acquisto, vengono chieste le loro aspettative sulla qualità complessiva, sul rendimento del prodotto e se suggerirebbero l'acquisto a parenti o amici. Se nel periodo di riferimento è stato acquistato più di un prodotto, il programma seleziona casualmente solo un bene a cui si riferiscono le domande e le successive vignettes. Seguono quindi le domande su quanto il rispondente sia soddisfatto del proprio acquisto (autovalutazione) e due vignettes legate al prodotto in esame. Alla fine vengono proposte alcune domande di valutazione che riguardano il questionario appena completato.

Oltre al questionario sul prodotto acquistato e sulla relativa soddisfazione, gli intervistati completano un questionario separato per mezzo del quale vengono raccolte informazioni di tipo socio-demografico. Le variabili raccolte relative a ciascun individuo sono: il genere, l'età, lo stato civile, il livello di istruzione, il reddito lordo e netto mensili.

I rispondenti al questionario vengono divisi in 2 gruppi, A e B: quelli assegnati al

gruppo A ricevono la domanda sulla valutazione della propria soddisfazione prima della coppia di domande di vignettes, in cui devono giudicare la soddisfazione di individui ipotetici. Viceversa, per gli intervistati del gruppo B, l'ordine delle domande viene invertito, ossia gli intervistati prima valutano la situazione delle vignettes e poi forniscono la soddisfazione al proprio prodotto. In questo modo è stato possibile valutare il comportamento della validazione delle due assunzioni di RC e VE nelle due diverse somministrazioni.

La popolazione di riferimento per il LISS panel è quella che parla la lingua olandese ed è residente permanentemente nei Paesi Bassi. Lo stesso questionario è stato somministrato allo stesso campione per 3 waves successive. I dati utilizzati in questa tesi fanno riferimento, però, solo alla prima rilevazione, avvenuta nel Novembre del 2011. In questa rilevazione il questionario è stato completato da 5368 individui. Il numero di persone che ha risposto alle domande sulla soddisfazione, su almeno uno qualsiasi dei tre prodotti, è 2273 e questi sono stati divisi equamente tra gruppo A e gruppo B.

Per le analisi descrittive si fa riferimento a Paccagnella (2011); questo elaborato, invece, sarà riservato esclusivamente alle applicazioni approfondite nei capitoli precedenti. In particolare, per poter utilizzare l'approccio con le Anchoring Vignettes e correggere così le autovalutazioni dal DIF, è necessario che valgano le due assunzioni di RC e VE. Dunque, si è reso necessario, anche per questo dataset, verificare in quali condizioni esse valessero.

Il test utilizzato per la convalida della RC è quello tramite la funzione AVCUB costruita appositamente. Quindi, come descritto nei precedenti capitoli, si è proceduto con la stima di due modelli CUB per l'autovalutazione: uno vincolato e uno no, e si è valutata l'uguaglianza tra le funzioni di *log-verosimiglianze*. Questo procedimento è stato ripetuto per entrambi i gruppi di intervistati e per le tre diverse categorie di prodotto. Queste suddivisioni hanno ridotto notevolmente la numerosità sulla quale stimare i modelli per cui non è stato possibile introdurre tutte le covariate del

questionario supplementare riguardante le informazioni socio-demografiche.

I risultati ottenuti con l'approccio CUB sono riportati in Tabella 3.24, dai quali si evidenzia un netto comportamento della validazione dell'assunzione di RC; infatti per tutte e tre le categorie di prodotti, il test viene sempre accettato quando le vignettes sono poste prima dell'autovalutazione (gruppo B), mentre, quando l'ordine è invertito (gruppo A), allora i rispondenti usano scale diverse per giudicare sé stessi e gli individui ipotetici (a parte il caso dei TV LCD). Questo risultato dà un'importante indicazione sulla costruzione del questionario, affinché la metodologia delle Anchoring Vignettes sia applicata correttamente.

	Gruppo	Vignettes	LR	DF	p-value
Computer portatili	A	v1	22,31	9	0,0079
		v2	29,29	9	0,0005
	B	v1	0,90	9	0,9990
		v2	5,45	9	0,7937
Smartphone	A	v1	19,04	6	0,0041
		v2	24,72	6	0,0030
	B	v1	3,05	6	0,8023
		v2	4,71	6	0,5820
TV LCD	A	v1	16,79	6	0,0101
		v2	8,90	6	0,1792
	B	v1	11,12	6	0,0847
		v2	13,56	6	0,0349

Tabella 3.24: Metodo AVCUB per la validazione della RC sui dati del LISS panel.

Per avere la conferma della validità dell'approccio con le vignettes è necessario testare ancora l'assunzione di VE sfruttando, come visto in precedenza, l'LR test e verificando l'uguaglianza tra la *log-verosimiglianza* del modello completo (con tutte le covariate per la componente di *feeling*) e quella del modello con la sola intercetta. I risultati di Tabella 3.25 confermano la validità dell'assunzione per qualsiasi gruppo e per qualsiasi bene acquistato. Questa conclusione è dovuta probabilmente al fatto che la popolazione olandese è una popolazione estremamente omogenea, con alti livelli di istruzione e che riesce ad interpretare in maniera uniforme gli scenari posti nelle vignettes, senza essere influenzato dalle proprie caratteristiche individuali. In

conclusione, l'assunzione di VE è accettata per tutte le vignettes del questionario.

	Gruppo	Vignettes	LR	DF	p-value
Personal Computer	A	v1	8,45	10	0,5852
		v2	5,65	9	0,7743
	B	v1	1,50	9	0,9598
		v2	4,85	6	0,5638
Smartphone	A	v1	9,65	10	0,4719
		v2	11,08	9	0,2701
	B	v1	1,68	9	0,9468
		v2	2,32	6	0,8884
Televisori	A	v1	9,65	10	0,4719
		v2	7,97	9	0,5375
	B	v1	1,70	9	0,9452
		v2	1,46	6	0,9620

Tabella 3.25: Test LR per la validazione della VE sui dati del LISS panel.

Visti i risultati dei test ottenuti, si può affermare che l'approccio con le Anchoring Vignettes con i dati del LISS panel è appropriato quando l'ordine tra vignettes e autovalutazione prevede quest'ultima come ultima domanda. In questo modo si possono vedere le vignettes come delle domande che chiarificano il significato della valutazione della soddisfazione, portando il rispondente a contestualizzare meglio le proprie risposte.

Conclusioni

In questo capitolo si cercherà di sintetizzare ciò che è stato visto in questa tesi, rispetto agli scopi descritti nell'introduzione.

L'obiettivo di questo lavoro era di introdurre una nuova soluzione che testasse le assunzioni previste dall'approccio delle Anchoing Vignettes, senza l'utilizzo di ulteriori condizioni di applicabilità del test stesso. La tecnica che è stata usata a tale scopo ha previsto l'introduzione di un'altra classe di modelli per dati ordinali, i modelli CUB, che hanno il vantaggio di non necessitare di particolari condizioni di utilizzo. Questo ha permesso di costruire due test, uno per la *Responce Consistency* (RC) e uno per la *Vignette Equivalence* (VE), in grado di fornire direttamente i risultati per la convalida del corretto utilizzo delle vignettes per l'eliminazione del DIF.

Il risultato di tutte queste considerazioni ha portato alla definizione di una nuova funzione in R, chiamata AVCUB, in grado di fornire la stima dei modelli CUB sulle domande di vignettes e autovalutazioni ed anche i risultati dei test sulle due assunzioni di base. Difatti, è bastato identificare quali componenti dei modelli CUB fossero coinvolte nelle ipotesi di RC e VE e, con un semplice test (LR test), è stato possibile verificarle efficacemente. Per quanto riguarda la prima assunzione testata, quella di RC, è bastato verificare che la stima dei parametri delle covariate della componente di incertezza tra la vignette e l'autovalutazione fosse la stessa. Invece, per la validazione della VE si è reso necessario un semplice test congiunto sulle stime dei parametri delle covariate della componente di *feeling*.

I dati dell'*online banking* che sono stati utilizzati per questi fini hanno rivelato diversi risultati interessanti. In particolare, quello che è emerso sono state delle forti indicazioni per la costruzione di un questionario che consenta ai rispondenti di capire e interpretare correttamente le vignettes, senza che vengano violate le assunzioni di base che permettono l'eliminazione delle eterogeneità interpersonali intrinseche nelle valutazioni fornite. La struttura del questionario, con la divisione in quattro gruppi, ha permesso di ricercare l'ordinamento migliore tra vignettes e autovalutazioni, cosicché le assunzioni venissero rispettate. Questo darà la possibilità, nel caso di una nuova *wave*, che l'intero campione sia utilizzato per la stima dei parametri senza alcuna suddivisione in gruppi distinti e senza perdita di informazione; ovvero, a tutti i rispondenti campionati verrà proposto un unico questionario ordinato secondo le indicazioni risultate da queste analisi.

I risultati forniti dai dati dell'*online banking* hanno permesso di effettuare le seguenti conclusioni:

- il genere degli individui della vignette ha un forte impatto sui rispondenti che si trovano a doverla valutare. Le soluzioni che possono essere proposte per non ridurre la numerosità del campione sono di utilizzare nomi di genere neutro o solo le iniziali del nome o, alternativamente, di considerare il genere dell'individuo ipotetico tra i predittori del modello. Altrimenti, i risultati ottenuti hanno dimostrato che persone di genere opposto a quello della vignette, rispetto ai rispondenti dello stesso genere, sembrano più propense ad usare le stesse scale di valutazione che usano per sé stessi (RC).
- A seconda dell'ordine scelto nel proporre vignettes e autovalutazioni, i risultati sono stati abbastanza omogenei, ad eccezione di Q15, per la quale l'assunzione di RC risulta valida in un solo verso. Tuttavia, nel momento in cui viene introdotto il genere della vignette nella costruzione del test, questa differenza scompare. Dunque, anticipare le vignettes prima dell'autovalutazione non aiu-

tava i rispondenti ad avere un chiarimento del significato della domanda sulla valutazione della soddisfazione del servizio.

- Un'altra variazione che portava ad usare diverse scale di valutazione era, invece, l'ordine tra le vignettes stesse. Per la maggior parte dei casi la prima vignette del blocco che veniva proposta era giudicata conformemente con le assunzioni richieste, mentre quelle successive no, soprattutto per il primo blocco Q13-Q15-Q16. Infatti, questo era il primo set di domande che i rispondenti dovevano affrontare, come se dovessero familiarizzare con il questionario; al contrario, nel secondo blocco le problematiche risultavano ridotte.
- il risultato più evidente è quello che per alcune vignettes (Q16 e Q21) le assunzioni non valgono quasi mai, in particolare quella di RC che, se non rispettata, implica diversa incertezza nel valutare vignette e autovalutazione. Ciò è dovuto ad una imprecisa costruzione dello scenario della vignette, che non è risultato ben chiaro e definito ai rispondenti i quali, infatti, hanno risposto con maggiore incertezza alla domanda. Per queste vignettes è necessaria una revisione della descrizione dello scenario per essere riproposte.

Per una miglior comprensione, in Tabella 3.26 vengono sintetizzati i risultati ottenuti sulla validazione delle assunzioni di RC e VE per i dati dell'*online banking*.

Uno dei temi più controversi nel campo delle Anchoring Vignettes è quello che riguarda l'assunzione di VE. Grazie ai modelli CUB è possibile testarla senza alcuna problematica (senza cioè nessuna ipotesi aggiuntiva, anzi senza neppure dover imporre RC) e con un semplice test congiunto sui parametri. Come si può vedere nella tabella riassuntiva è evidente un comportamento legato alle posizioni delle vignettes: infatti, le prime sembrano valutate in base a caratteristiche individuali che pesano sul giudizio del rispondente, ma andando avanti nel questionario, tali caratteristiche sembrano influenzare sempre meno il giudizio dei clienti comportando delle significatività del test sull'assunzione di VE maggiori. Questa conclusione permetterebbe

di costruire un questionario per il quale le prime vignettes siano presenti solo con lo scopo di contestualizzare la tematica d'interesse, mentre le vignettes successive siano usate per la vera e propria correzione del DIF, dato che per esse le assunzioni risultano maggiormente verificate.

Per un'ulteriore conferma dei risultati sulla verifica delle assunzioni, ottenuti attraverso questa metodologia, si è provato a confrontare le tecniche introdotte da van Soest et al. (2007) e Bago d'Uva et al. (2011). Queste utilizzavano una quantità che misurava oggettivamente la soddisfazione del cliente che, nel caso specifico, corrispondeva a quanto le aspettative antecedenti all'attivazione del servizio di *online banking* fossero state soddisfatte dopo il suo utilizzo. Purtroppo, a causa di numerosità esigue, il modello *Chopit* in Stata ha avuto problemi di convergenza ed è stato impossibile confrontare le metodologie dei due diversi approcci, a parità di covariate introdotte nei modelli. Allora, come convalida della tecnica che sfrutta i modelli CUB, si è analizzato un ulteriore dataset, in particolare uno estratto dal LISS panel, appartenente al progetto MESS, che conteneva al suo interno alcune vignettes. Le informazioni estratte dai risultati, riportati nella Sezione 3.5, forniscono spunti interessanti per la costruzione del questionario; infatti in questo caso, è evidente una forte dipendenza dall'ordine tra vignettes e autovalutazione, soprattutto per quanto riguarda l'assunzione di RC. Invece, la VE viene sempre verificata, probabilmente a causa della la bassa eterogeneità della popolazione olandese, a differenza di quella italiana analizzata nel campione dell'*online banking*.

Questo lavoro di tesi potrebbe portare un contributo importante nel campo delle analisi di dati ordinali che sfruttano l'approccio delle vignettes per correggere l'eterogeneità della popolazione del campione, soprattutto nell'ambito della validazione delle sue assunzioni di base che ancora non hanno trovato soluzioni libere da condizioni aggiuntive di applicabilità. La metodologia costruita in questa tesi potrà dare risultati immediati sulla legittimità delle assunzioni di RC e VE per confermare la corretta applicazione delle Anchoring Vignettes.

Vignettes	RC			VE		RC		VE	
	Autovalut.	RC	VE	Genere uguale	Genere opposto	Prima posizione	Seconda posizione	Prima posizione	Seconda posizione
Q15	Q13	0,0065	0,0379	0,0239	0,0513	0,5585	0,7129	0,3063	0,0029
	Q13bis	0,0541	0,0020	0,1448	0,1052	0,0687	0,1765	0,4640	0,0069
Q16	Q13	<0,0001	0,0011	<0,0001	<0,0001	0,0053	0,1003	<0,0001	<0,0001
	Q13bis	<0,0001	0,4253	<0,0001	<0,0001	0,0009	0,1661	<0,0001	0,0287
Q21	Q19	<0,0001	0,1959	<0,0001	0,0011	0,0656	0,0217	<0,0001	<0,0001
	Q19bis	<0,0001	0,0062	<0,0001	0,0507	0,0007	0,0154	<0,0001	<0,0001
Q22	Q19	0,2053	0,5138	0,1707	0,4167	0,4843	0,4036	0,9969	0,1492
	Q19bis	0,1157	0,0475	0,1402	0,0501	0,3367	0,6605	0,2531	0,7636

Tabella 3.26: Riassunto dei risultati sull'*online banking* della validazione delle assunzioni di RC e VE.

Appendice A

Analisi descrittive *online banking*

In questa appendice si riportano le distribuzioni delle valutazioni espresse dai clienti, in possesso di almeno un conto corrente che prevede operazioni online, in base a determinate caratteristiche selezionate. Per ciascuna variabile saranno presenti quattro grafici corrispondenti ai due blocchi di domande per ciascuno dei 2 gruppi considerati. Sarà possibile osservare la percentuale di individui che con quella determinata caratteristica hanno risposto ad una delle domande del blocco con la categoria di valutazione corrispondente.

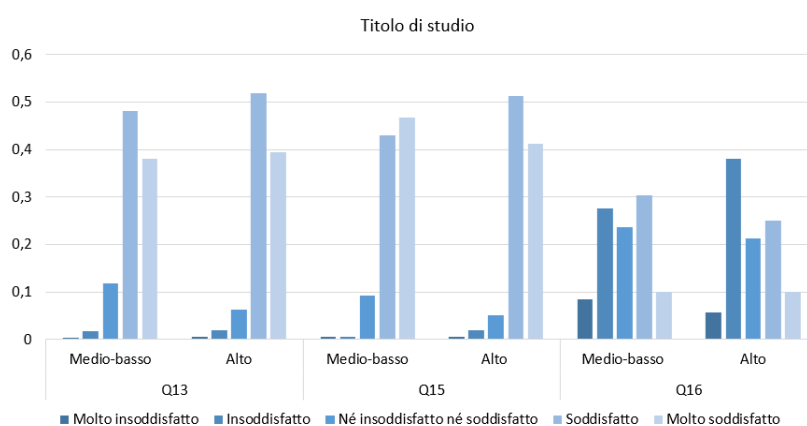


Figura A.1: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al titolo di studio raggiunto dagli individui.

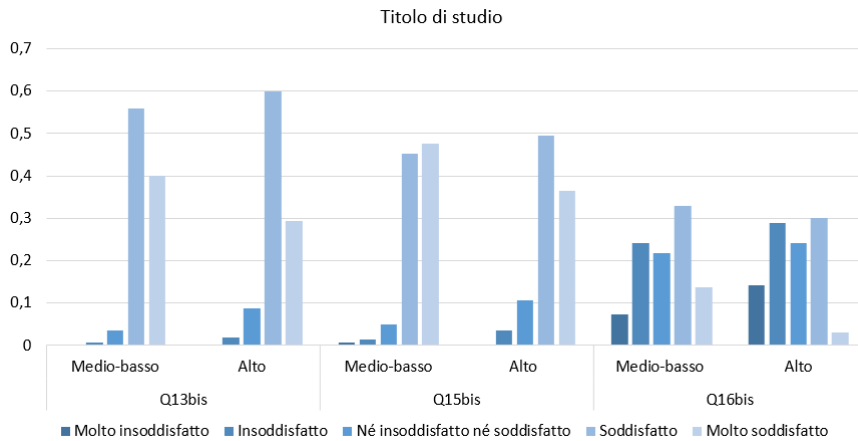


Figura A.2: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al titolo di studio raggiunto dagli individui.

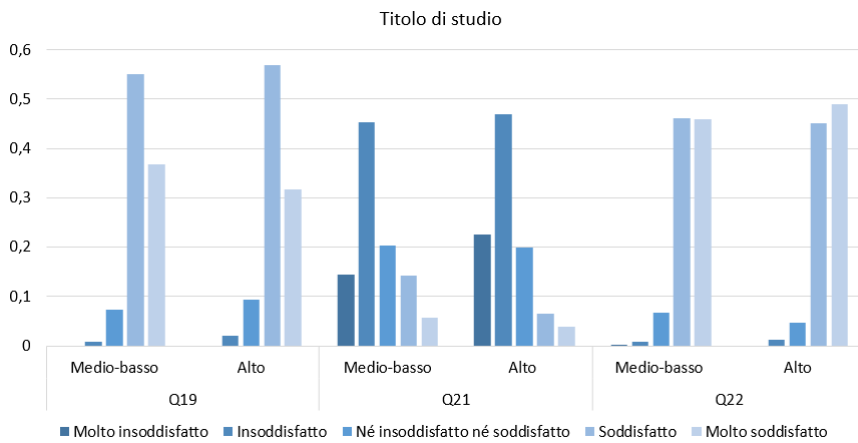


Figura A.3: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al titolo di studio raggiunto dagli individui.

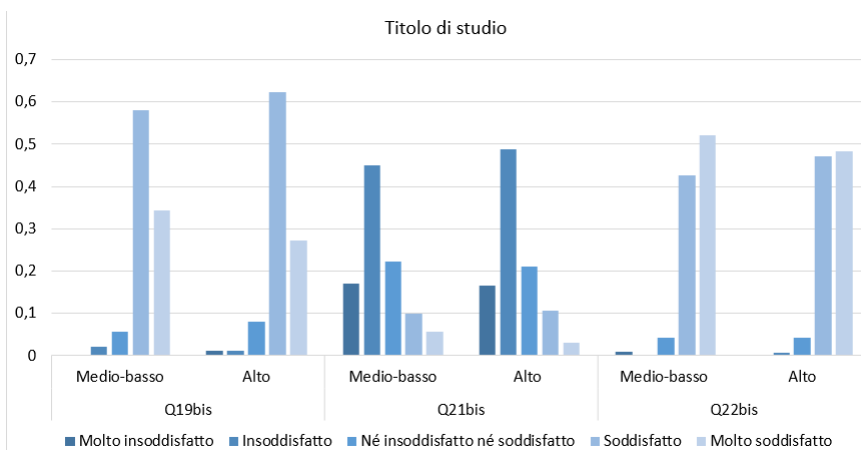


Figura A.4: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al titolo di studio raggiunto dagli individui.

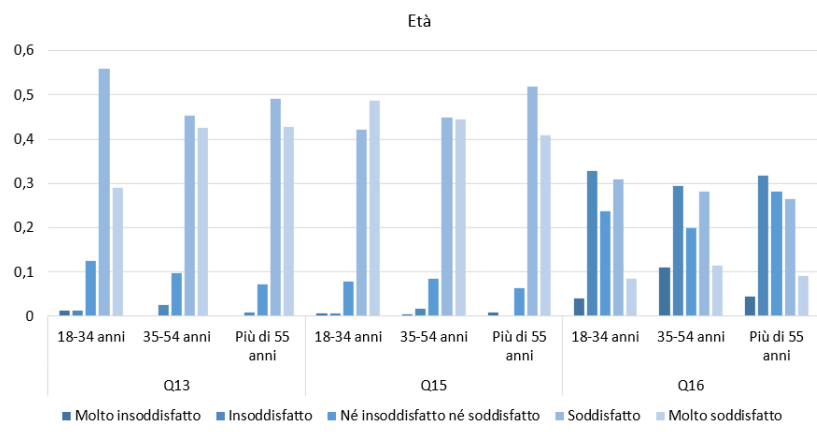


Figura A.5: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto all'età degli individui.

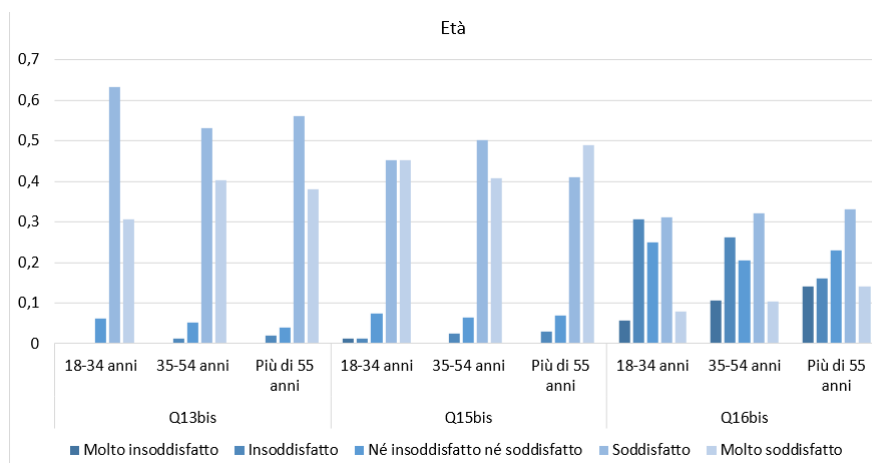


Figura A.6: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto all'età degli individui.

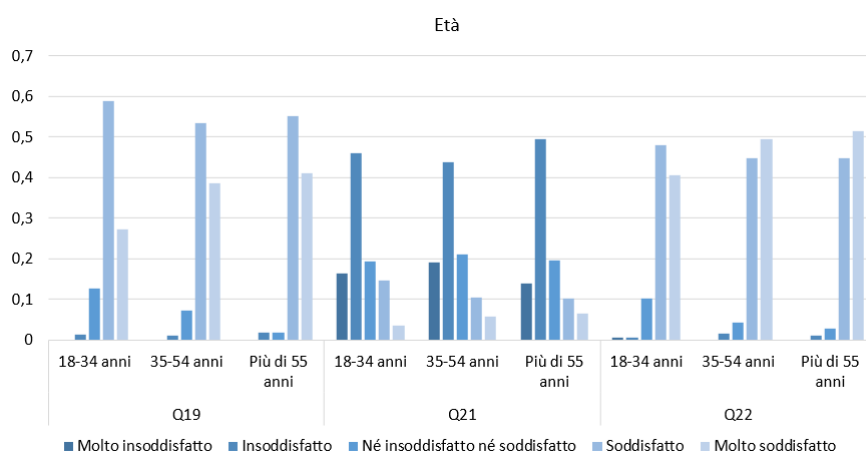


Figura A.7: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto all'età degli individui.

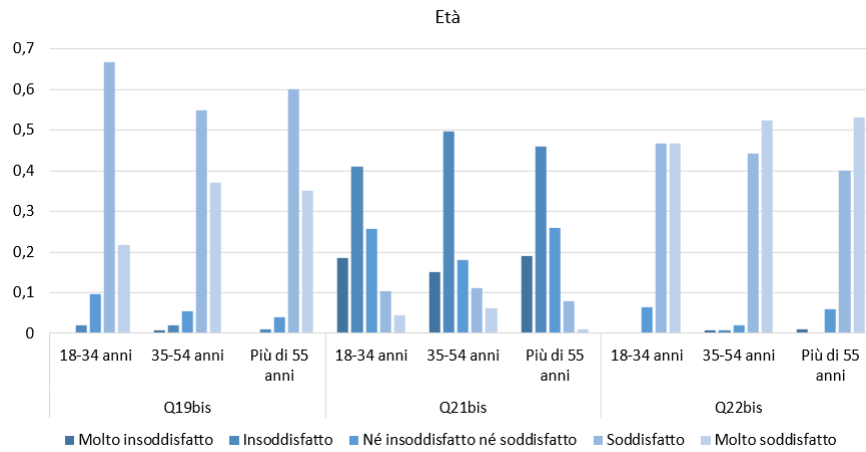


Figura A.8: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto all'età degli individui.

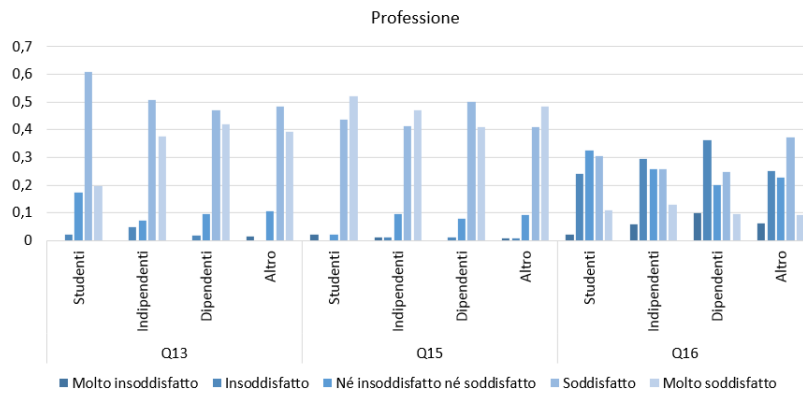


Figura A.9: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla professione degli individui.

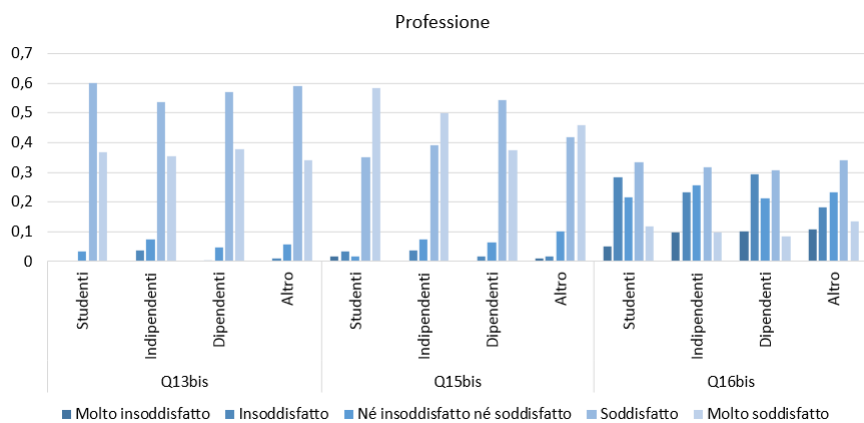


Figura A.10: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla professione degli individui.

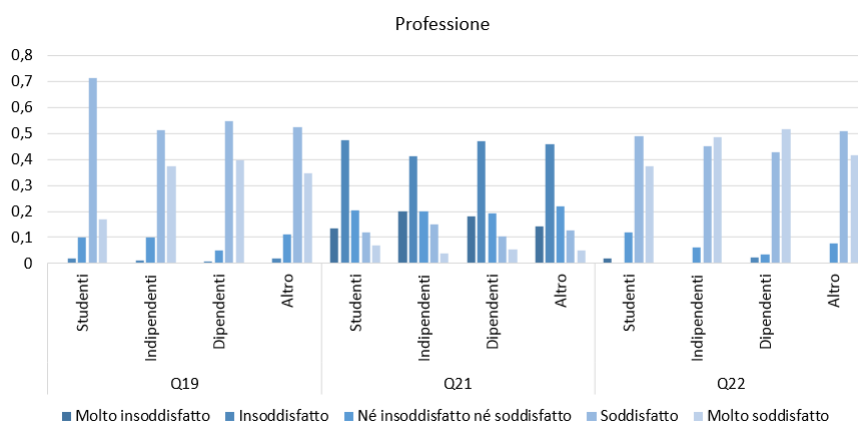


Figura A.11: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla professione degli individui.

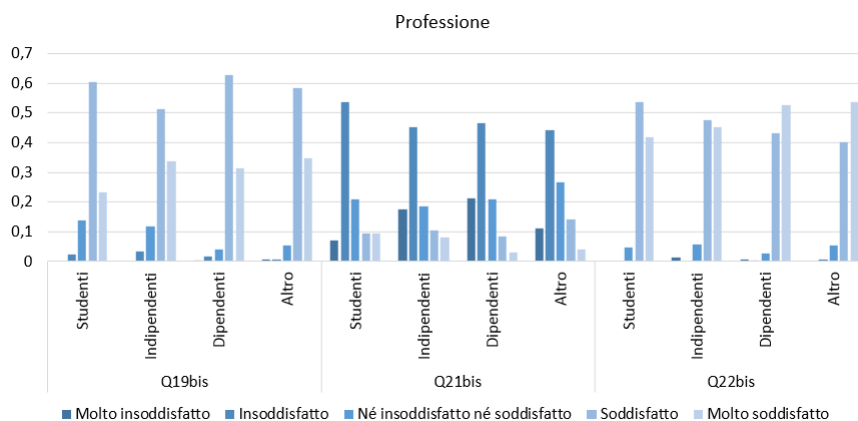


Figura A.12: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla professione degli individui.

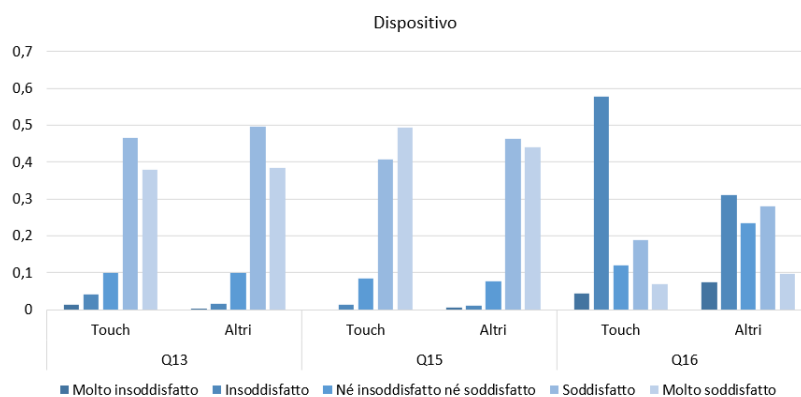


Figura A.13: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al tipo di dispositivo impiegato nel questionario dagli individui.

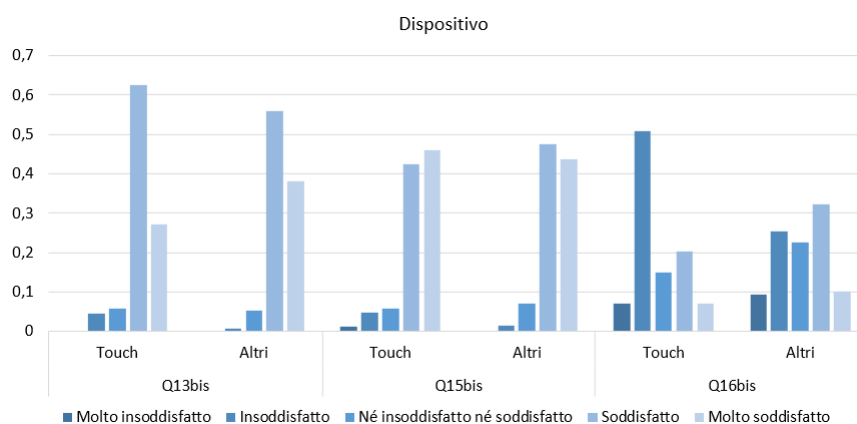


Figura A.14: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al tipo di dispositivo impiegato nel questionario dagli individui.

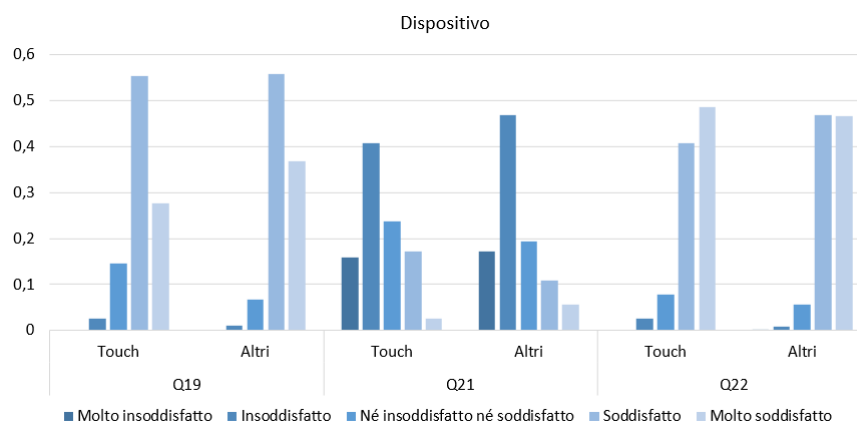


Figura A.15: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al tipo di dispositivo impiegato nel questionario dagli individui.

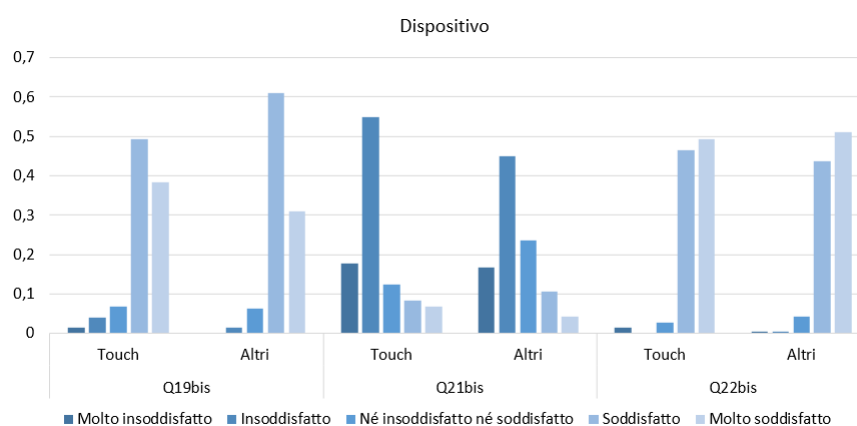


Figura A.16: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al tipo di dispositivo impiegato nel questionario dagli individui.



Figura A.17: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto al tipo di browser impiegato nel questionario dagli individui.

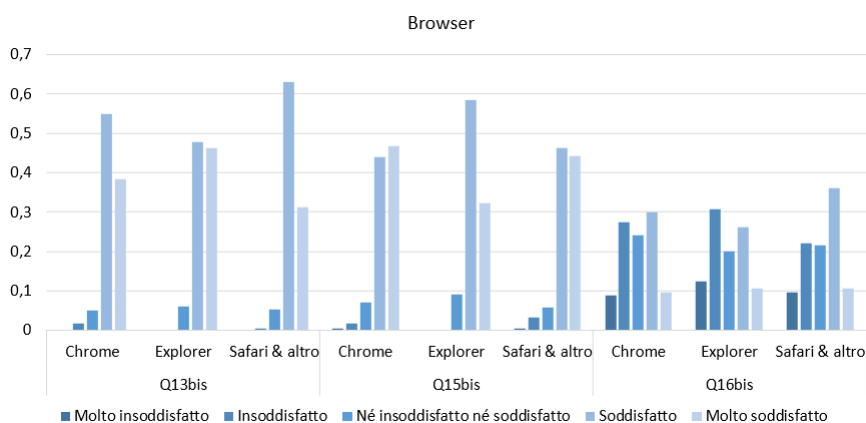


Figura A.18: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto al tipo di browser impiegato nel questionario dagli individui.

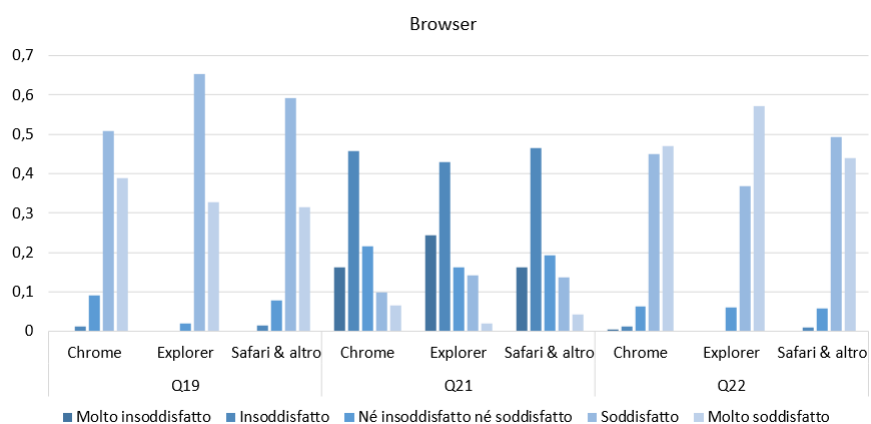


Figura A.19: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto al tipo di browser impiegato nel questionario dagli individui.

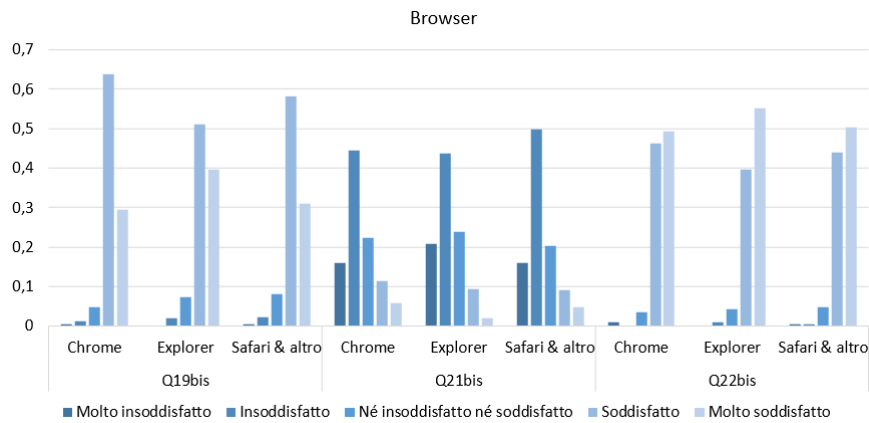


Figura A.20: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto al tipo di browser impiegato nel questionario dagli individui.

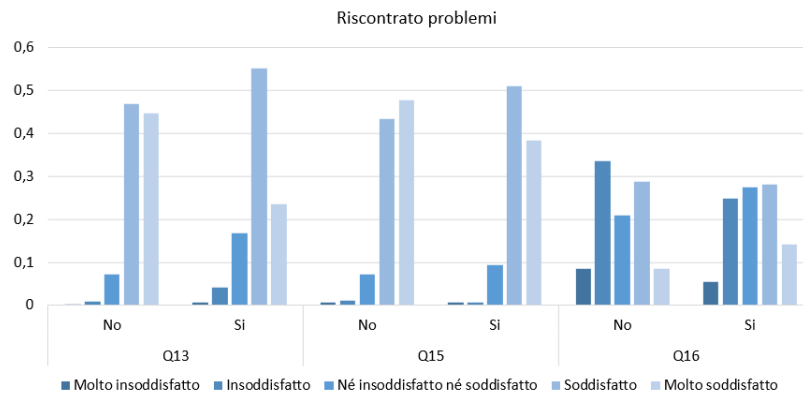


Figura A.21: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla presenza di problemi durante il servizio di *online banking*.

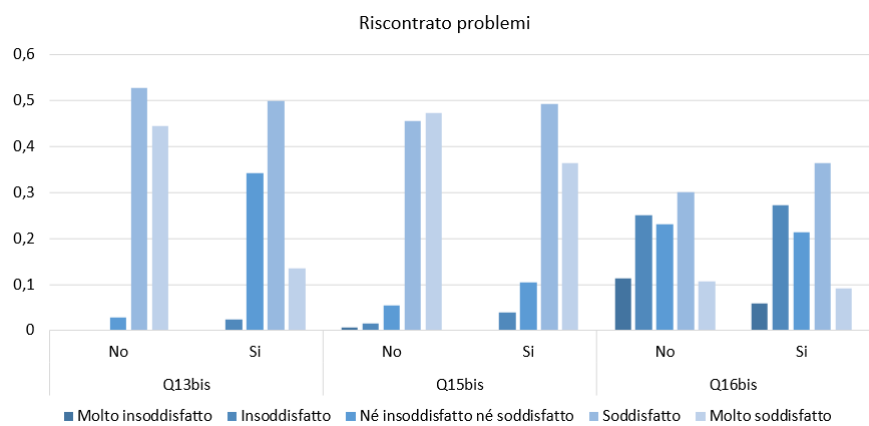


Figura A.22: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla presenza di problemi durante il servizio di *online banking*.

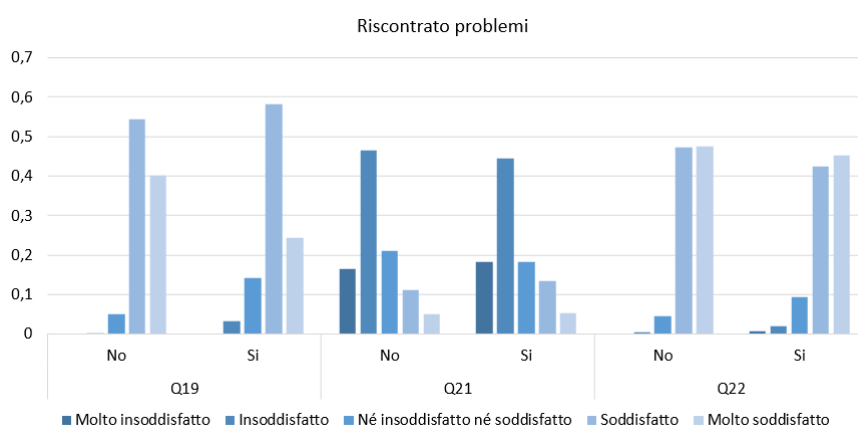


Figura A.23: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla presenza di problemi durante il servizio di *online banking*.

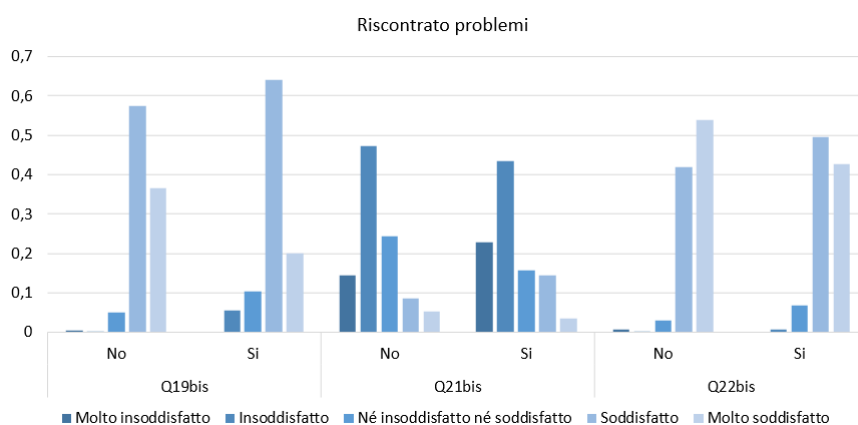


Figura A.24: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla presenza di problemi durante il servizio di *online banking*.

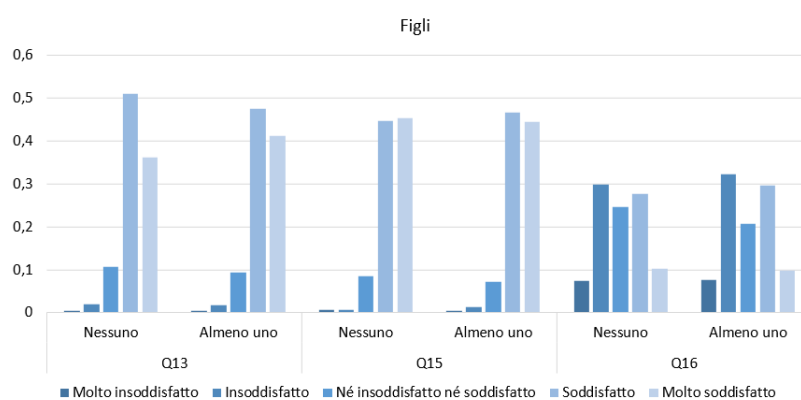


Figura A.25: Distribuzione delle risposte al blocco Q13-Q15-Q16 rispetto alla presenza di figli a carico del rispondente.

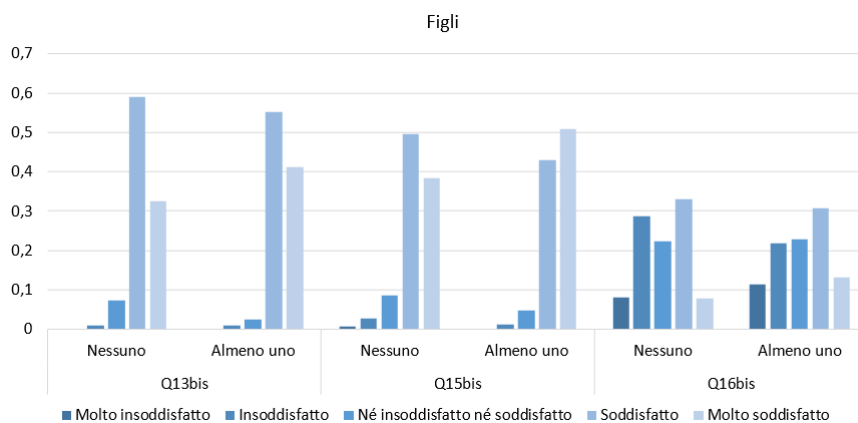


Figura A.26: Distribuzione delle risposte al blocco Q13bis-Q15bis-Q16bis rispetto alla presenza di figli a carico del rispondente.

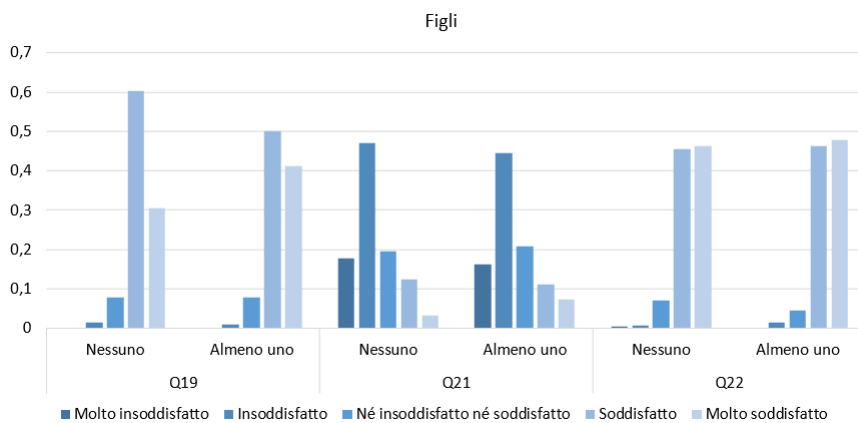


Figura A.27: Distribuzione delle risposte al blocco Q19-Q21-Q22 rispetto alla presenza di figli a carico del rispondente.

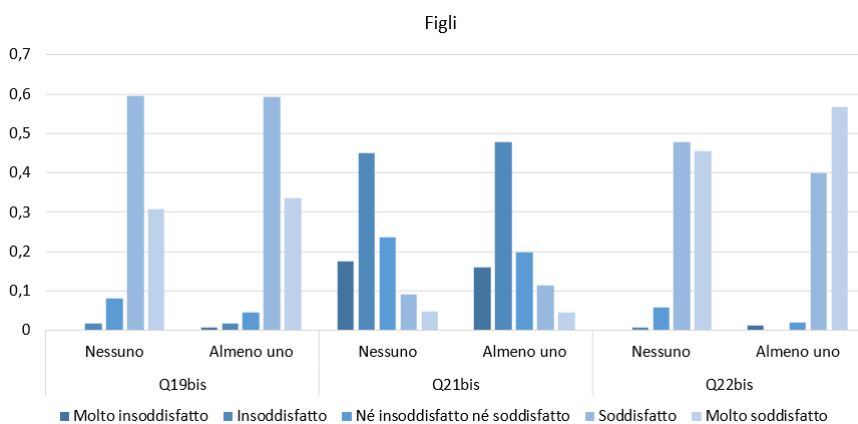


Figura A.28: Distribuzione delle risposte al blocco Q19bis-Q21bis-Q22bis rispetto alla presenza di figli a carico del rispondente.

Appendice B

Questionario *online banking*

- Q1: Lei dispone di almeno un conto corrente bancario o postale?

1. Sì, ne possiedo solo 1.
2. Sì, ne possiedo 2 o più.
3. No, non ne possiedo alcuno.

IF Q1==3

FINE INTERVISTA

ENDIF

IF Q1<3

- Q2: Pensi al conto corrente principale, cioè a quello che utilizza con maggiore frequenza. Di che tipo di conto corrente si tratta?

1. Un conto corrente che consente operazioni solo in agenzia.
2. Un conto corrente che consente operazioni sia in agenzia, sia on-line.
3. Un conto corrente che consente operazioni solo on-line.

- Q3: Sempre facendo riferimento al suo conto corrente principale (cioè a quello che utilizza con maggiore frequenza), perché ha aperto questo tipo di conto?
[RISPOSTA MULTIPLA]

1. Per la comodità di gestione delle operazioni.
2. Per avere interessi alti.
3. Per spese di gestione nulle o estremamente basse.
4. Per il tipo di servizi offerti.
5. Perché ho necessità di un contatto fisico con un operatore.
6. Altro.

- Q4: Da quanto tempo è aperto il suo conto corrente principale?

1. Meno di 1 anno.
2. Da 1 a 2 anni.
3. Da 3 a 5 anni.
4. Da 6 a 10 anni.
5. Oltre 10 anni.

IF Q4==1

- Q5: Si tratta del suo primo conto corrente?

1. Sì, è la prima volta che ho un conto corrente.
2. No, ho solamente cambiato istituto (banca, Posta, etc.) in cui l'ho sottoscritto.
3. No, avevo un conto cointestato con altre persone nello stesso o in un altro istituto (banca, Posta, etc.).

ENDIF

IF Q2==1

FINE INTERVISTA

ENDIF

IF Q2!=1

- Q6: Che tipo di servizi utilizza nel suo conto corrente on-line principale?[RISPOSTA MULTIPLA]

1. Controllo del saldo e della lista movimenti.
2. Controllo dei movimenti della/e carta/e bancomat o postamat.
3. Controllo dei movimenti della/e carta/e di credito.
4. Controllo dei movimenti della/e carta/e prepagata/e.
5. Predisposizione di bonifici o giroconti.
6. Pagamento di utenze, bollettini, bolli o imposte.
7. Ricariche telefoniche.
8. Gestione prestiti o fidi.
9. Trading on-line o investimenti su valute.
10. Altri servizi.
11. Nessun servizio, preferisco relazionarmi solamente con l'agenzia (solamente per coloro il cui conto corrente principale permette operazioni anche in agenzia).

IF Q6==11

FINE INTERVISTA

ENDIF

IF Q6<11

- Q7: Ha mai avuto dei problemi nella visione, navigazione o gestione del suo conto corrente on- line principale?

1. No, mai.
2. Sì, ma non ho contattato nessuno per risolvere il problema.
3. Sì e ho contattato solo il call center per risolvere il problema.
4. Sì e ho contattato solo l'assistenza on-line per risolvere il problema.
5. Sì e ho contattato sia l'assistenza on-line sia il call center per risolvere il problema.

IF Q7==3 o Q7==5

- Q8: Per favore ripensi alle sue esperienze con il call center. Su una scala da 1 a 10 dove 1 indica “Molto insoddisfatto/a” e 10 “Molto soddisfatto/a”, nel complesso in che misura è stato/a soddisfatto/a dalla:

1. Tempestività della risposta [CODIFICARE 1 - 10]
2. Preparazione dell'operatore nel risolvere il problema [CODIFICARE 1 - 10]
3. Gentilezza dell'operatore [CODIFICARE 1 - 10]

ENDIF

IF Q7==4 o Q7==5

- Q9: Per favore ripensi alle sue esperienze con l'assistenza on-line. Su una scala da 1 a 10 dove 1 indica “Molto insoddisfatto/a” e 10 “Molto soddisfatto/a”, nel complesso in che misura è stato/a soddisfatto/a dalla:

1. Tempestività della risposta [CODIFICARE 1 - 10]
2. Qualità della risposta [CODIFICARE 1 - 10]
3. Gentilezza della risposta [CODIFICARE 1 - 10]

ENDIF

- Q10: Per favore ripensi ora alla procedura di apertura del conto principale. Su una scala da 1 a 10 dove 1 indica “Completamente in disaccordo” e 10 “Completamente d'accordo”, in che misura è d'accordo con le seguenti affermazioni:

1. La procedura è stata complessa [CODIFICARE 1 - 10]
2. La procedura ha avuto tempi di attesa molto lunghi [CODIFICARE 1 - 10]
3. La procedura è stata costosa [CODIFICARE 1 - 10]

Adesso vorremmo rivolgerle alcune domande riguardo alla soddisfazione su alcuni aspetti dell'utilizzo del suo conto corrente on-line principale.

IF Q6<5

- Q11: Per favore ripensi alle aspettative che lei aveva riguardo alla facilità di navigazione del suo conto corrente on-line principale. Su una scala da 1 a 10 dove 1 indica “Molto scarse” e 10 “Molto alte”, come valuta queste aspettative? [CODIFICARE 1 - 10]

- Q12: Per favore ripensi alle sue esperienze relative alla facilità di navigazione del suo conto corrente on-line principale. Su una scala da 1 a 10 dove 1 indica “In misura molto limitata” e 10 “In misura molto elevata”, in che misura queste esperienze sono corrisposte alle sue aspettative? [CODIFICARE 1 - 10]

Per le prossime 4 domande verranno creati casualmente 4 gruppi e le domande verranno rivolte nel seguente ordine:

GRUPPO 1: Q13 – Q14 – Q15 – Q16

GRUPPO 2: Q13 – Q14 – Q16 – Q15

GRUPPO 3: Q14bis – Q15 – Q16 – Q13bis

GRUPPO 4: Q14bis – Q16 – Q15 – Q13bis

- Q13: Quanto si ritiene soddisfatto/a dalla facilità di navigazione del suo conto corrente on-line principale?

1. Molto soddisfatto/a
2. Soddisfatto/a
3. Nè soddisfatto/a, nè insoddisfatto/a
4. Insoddisfatto/a
5. Molto insoddisfatto/a

- Q13bis: Tornando alla sua esperienza di utilizzo del suo conto corrente on-line principale, quanto si ritiene soddisfatto/a dalla facilità di navigazione del suo conto corrente on-line principale?

1. Molto soddisfatto/a
2. Soddisfatto/a
3. Nè soddisfatto/a, nè insoddisfatto/a
4. Insoddisfatto/a
5. Molto insoddisfatto/a

- Q14: Le proporremo adesso due esempi di persone che hanno avuto delle esperienze nella navigazione di un conto corrente on-line. Vorremmo sapere come valuta la soddisfazione di queste persone riguardo alla facilità di navigazione del loro conto corrente on-line, utilizzando la stessa scala con la quale in precedenza ha valutato la sua esperienza personale. Nell’esprimere il suo giudizio, immagini che queste persone abbiano la sua stessa età e le sue stesse conoscenze e competenze.

- Q14bis: Le proporremo adesso due esempi di persone che hanno avuto delle esperienze nella navigazione di un conto corrente on-line. Vorremmo sapere

come valuta la soddisfazione di queste persone riguardo alla facilità di navigazione del loro conto corrente on-line. Nell'esprimere il suo giudizio, immagina che queste persone abbiano la sua stessa età e le sue stesse conoscenze e competenze.

- Q15: Carlo è un lavoratore dipendente e da circa 3 anni possiede un conto corrente on-line. Ogni giorno controlla la lista movimenti del suo conto corrente on-line per verificare l'assenza di movimenti anomali: Carlo entra nel sito, cerca la scheda conto corrente e nel menu a tendina seleziona la voce "Movimenti conto". Quindi clicca sull'icona "Ultimi dieci movimenti" e poi controlla la lista. Anche se il caricamento della lista dei movimenti richiede alcuni secondi, Carlo impiega poco meno di un minuto per completare tutta la procedura di controllo. Quanto è soddisfatto Carlo riguardo alla facilità di navigazione del suo conto corrente on-line principale?

1. Molto soddisfatto
2. Soddisfatto
3. Nè soddisfatto, nè insoddisfatto
4. Insoddisfatto
5. Molto insoddisfatto

- Q16: Marina è una casalinga che controlla circa ogni 3 giorni la lista delle spese effettuate dalla sua famiglia con la carta di credito. Un giorno vuole ricontrollare le spese sostenute nel mese precedente ma non trova il menù a tendina per indicare il mese corretto. Ha bisogno di contattare il call center per risolvere il problema. Grazie all'operatore riesce a far apparire la lista dei movimenti che stava cercando. Quanto è soddisfatta Marina riguardo alla facilità di navigazione del suo conto corrente on-line principale?

1. Molto soddisfatta
2. Soddisfatta
3. Nè soddisfatta, nè insoddisfatta
4. Insoddisfatta
5. Molto insoddisfatta

ENDIF

Domande rivolte di nuovo a tutte le persone

IF Q6>4 & Q6<11

- Q17: Per favore ripensi alle aspettative che lei aveva riguardo alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale. Su una scala da 1 a 10 dove 1 indica "Molto scarse" e 10 "Molto alte", come valuta queste aspettative? [CODIFICARE 1 - 10]
- Q18: Per favore ripensi a tutte le sue esperienze relative alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale. Su una scala da 1 a 10 dove 1 indica "Molto limitata" e 10 "Molto elevata", in che misura

queste esperienze hanno corrisposto alle sue aspettative? [CODIFICARE 1 - 10]

Per le prossime 4 domande verranno creati casualmente 4 gruppi (possono essere diversi rispetto a quelli creati in precedenza) e le domande verranno rivolte nel seguente ordine:

GRUPPO 1: Q19 – Q20 – Q21 – Q22

GRUPPO 2: Q19 – Q20 – Q22 – Q21

GRUPPO 3: Q20bis – Q21 – Q22 – Q19bis

GRUPPO 4: Q20bis – Q22 – Q21 – Q19bis

- Q19: Quanto si ritiene soddisfatto/a riguardo alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale?
 1. Molto soddisfatto/a
 2. Soddisfatto/a
 3. Nè soddisfatto/a, nè insoddisfatto/a
 4. Insoddisfatto/a
 5. Molto insoddisfatto/a

- Q19bis: Tornando alla sua esperienza di utilizzo del suo conto corrente on-line principale, quanto si ritiene soddisfatto/a riguardo alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale?
 1. Molto soddisfatto/a
 2. Soddisfatto/a
 3. Nè soddisfatto/a, nè insoddisfatto/a
 4. Insoddisfatto/a
 5. Molto insoddisfatto/a

- Q20: Le proporremo adesso due esempi di persone che hanno avuto delle esperienze nell'utilizzo di un conto corrente on-line. Vorremmo sapere come valuta la soddisfazione di queste persone riguardo alla semplicità di esecuzione delle operazioni del loro conto corrente on-line principale, utilizzando le stesse scale con cui prima ha valutato la sua esperienza personale. Nell'esprimere il suo giudizio, immagini che queste persone abbiano la sua stessa età e le sue stesse conoscenze e competenze.

- Q20bis: Le proporremo adesso due esempi di persone che hanno avuto delle esperienze nell'utilizzo di un conto corrente on-line. Vorremmo sapere come valuta la soddisfazione di queste persone riguardo alla semplicità di esecuzione delle operazioni del loro conto corrente on-line principale. Nell'esprimere il suo giudizio, immagini che queste persone abbiano la sua stessa età e le sue stesse conoscenze e competenze.

- Q21: Sofia è una manager che utilizza quotidianamente internet e da oltre 10 anni possiede conti correnti on-line. Per motivi di tempo, Sofia esegue tutti i suoi pagamenti di bollette, bollettini e altre imposte per mezzo del proprio conto corrente on-line principale. Ogni tanto commette un errore nella digitazione di un codice Iban oppure di un numero di bollettino e non si ferma a controllare subito quanto inserito. Quando l'operazione non va a buon fine, il sistema non segnala l'origine dell'errore e Sofia è costretta a riscrivere tutti i dati precedentemente inseriti. Quanto è soddisfatta Sofia riguardo alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale?

1. Molto soddisfatta
2. Soddisfatta
3. Nè soddisfatta, nè insoddisfatta
4. Insoddisfatta
5. Molto insoddisfatta

- Q22: Pietro è un lavoratore autonomo e nel suo conto corrente on-line dispone anche di un conto titoli sul quale è attivo il servizio di trading on-line. Solitamente accede al servizio ed esegue tutte le sue operazioni senza problemi. La procedura richiede l'inserimento di una serie di codici di sicurezza, ma il tutto si svolge in tempi rapidi e senza particolari difficoltà. Quanto è soddisfatto Pietro riguardo alla semplicità di esecuzione delle operazioni del suo conto corrente on-line principale?

1. Molto soddisfatto
2. Soddisfatto
3. Nè soddisfatto, nè insoddisfatto
4. Insoddisfatto
5. Molto insoddisfatto

ENDIF

Domande rivolte di nuovo a tutte le persone

- Q23: Raccomanderebbe l'utilizzo di servizi banking on-line ad amici e parenti?
1. Sì.
 2. No.

END

Appendice C

Questionario progetto VECS

- fj11a001: Group A or B
 1. A
 2. B
- fj11a003 : in the case of three purchases, on which purchase do we elaborate?
 1. laptop
 2. smartphone
 3. LCD television
- fj11a004: in the case of two purchases, on which purchase do we elaborate?
 1. laptop
 2. smartphone
 3. LCD television

This questionnaire contains some questions about experiences you and others have when purchasing certain products.

- fj11a005-fj11a008: In the last 6 months, have you purchased one or more of the following?
- fj11a005: I bought a laptop [1- No , 2- Yes]
- fj11a006: I bought a smartphone (like Blackberry or iPhone) [1- No , 2- Yes]
- fj11a007: I bought a LCD television [1- No , 2- Yes]
- fj11a008: None of these [1- No , 2- Yes]
if fj11a008=1

- fj11a009: Do you intend to buy a laptop or a smartphone (like Blackberry or iPhone) or a LCD television within the next months? [1- No , 2- Yes]
if fj11a005=1 or fj11a006=1 or fj11a007=1
- fj11a010: In which month did you buy your [if laptop/smartphone/LCD television*]?
 1. May 2011
 2. June 2011
 3. July 2011
 4. August 2011
 5. September 2011
 6. October 2011
 7. November 2011
- fj11a011-fj11a015 Did you experience one of the following after the purchase of your [laptop/smartphone/LCD television*]?
 - fj11a011 A manufacturing defect [1- No , 2- Yes]
 - fj11a012 A delay in the delivery [1- No , 2- Yes]
 - fj11a013 A good with different features with respect to the purchase order [1- No , 2- Yes]
 - fj11a014 A price different with respect to the purchase order [1- No , 2- Yes]
 - fj11a015 None of these [1- No , 2- Yes]
- fj11a016: Before the purchase of your [laptop/smartphone/LCD television*]? you might have known something about the product. Now think back and remember your expectations of the overall quality of the [if laptop/smartphone/LCD television*].
How would you describe your expectations of the overall quality of the [laptop/smartphone/LCD television*]? [CODIFICARE 1 - 10]
- fj11a017: Please consider all your experiences to date with your [laptop/smartphone/LCD television*].
How would you rate your expectations of the overall quality of the [laptop/Apple Iphone/LCD TV]? [CODIFICARE 1 - 10]
- fj11a018: If you think back to all your expectations, to what extent has your [laptop/smartphone/LCD television*] (not) met your expectations? [CODIFICARE 1 - 10]
- fj11a019: Would you suggest the purchase of your [laptop/smartphone/LCD television*] to relatives or friends? [1- Yes , 2- No
if fj11a005=1 or fj11a006=1 or fj11a007=1 and fj11a001=1 and fj11a002=1

- fj11a020: How satisfied are you with your [laptop/smartphone/LCD television*]?
1. Very dissatisfied
 2. Dissatisfied
 3. Neither satisfied, nor dissatisfied
 4. Satisfied
 5. Very satisfied

if fj11a005=1 or if fj11a006=1 or if fj11a007=1 and fj11a001=1 and fj11a002=1
 intro We will now give you some examples of persons who experienced the purchase of some goods. We would like to know how you evaluate the given aspects of their purchases. Please assume that the persons have the same age and background that you have. Please choose one of the five answers for every question.

If laptop and fj11a001=1 and fj11a002=1

- fj11a021: Laura needs a laptop for her work. She went to a specialized shop and bought the laptop version with the largest memory capacity. She had to wait for a couple of days in order to receive it. She was able to learn easily the main features of her laptop. She has never experienced any manufacturing defects.

How satisfied is Laura with her laptop?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if laptop and fj11a001=1 and fj11a002=1

- fj11a022: John makes use of a laptop daily, in particular the internet facilities. He bought a new laptop with the desired features in a shopping centre at a discounted price. Since the first use, the laptop highlighted a slow internet connection. John came back to the shopping centre and the laptop was withdrawn for the assistance. He waited for one week before its delivery. The network interface card was changed. Since then, John has not experienced any other manufacturing defects.

How satisfied is John with his laptop?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied

5. Very satisfied

if smartphone and fj11a001=1 and fj11a002=1

- fj11a023: Mark needs a mobile phone for his work. He went to a specialized shop and bought the last version of a smartphone because its features are suitable for his work. He had to wait for four days in order to receive his smartphone. Reading the user guide, he was able to learn its main features in a couple of days. He has never experienced any manufacturing defects.

How satisfied is Mark with his smartphone?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if smartphone and fj11a001=1 and fj11a002=1

- fj11a024: Anne works part-time, makes various sports and has a lot of friends. She thinks a smartphone can meet her needs. In a shopping centre she immediately bought what she liked. Unfortunately, after two months she experienced a problem in the phone book. She came back to the shopping centre and the smartphone was withdrawn for the assistance. After ten days the phone was delivered to Anne and the problem was solved. She has not experienced any other manufacturing defects.

How satisfied is Anne with her smartphone?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if LCD television and fj11a001=1 and fj11a002=1

- fj11a025: Carry works all day. During the evening she likes watching TV together with her children and husband. She bought a 34 Inch LCD TV with several options, paying a price higher than expected. At home, she needed the assistance to set it up, because the user guide was not clear. Afterwards, she has never experienced any other problems.

How satisfied is Carry with her LCD TV?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied

4. Satisfied
5. Very satisfied

if LCD television and fj11a001=1 and fj11a002=1

- fj11a026: Jim likes watching sports on TV together with his friends. In a specialized shop he bought a 40 Inch LCD TV with a high resolution. He had to wait one week in order to receive it. He has never experienced any manufacturing defects or other problems.

How satisfied is Jim with his LCD TV?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

intro Now we will give you some examples of persons who experienced the purchase of some goods. We would like to know how you evaluate the given aspects of their purchases. Please assume that the persons have the same age and background that you have. Please choose one of the five answers for every question.

if laptop and fj11a001=2 and fj11a002=1

- fj11a034: Laura needs a laptop for her work. She went to a specialized shop and bought the laptop version with the largest memory capacity. She had to wait for a couple of days in order to receive it. She was able to learn easily the main features of her laptop. She has never experienced any manufacturing defects.

How satisfied is Laura with her laptop?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if laptop and fj11a001=2 and fj11a002=1

- fj11a035: John makes use daily of a laptop, in particular the internet facilities. He bought a new laptop with the desired features in a shopping centre at a discounted price. Since the first use, the laptop highlighted a slow internet connection. John came back to the shopping centre and the laptop was withdrawn for the assistance. He waited for one week before its delivery. The network interface card was changed. Since then, John has not experienced any other manufacturing defects.

How satisfied is John with his laptop?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if smartphone and fj11a001=2 and fj11a002=1

- fj11a036 Mark needs a mobile phone for his work. He went to a specialized shop and bought the last version of a smartphone because its features are suitable for his work. He had to wait for four days in order to receive his smartphone. Reading the user guide, he was able to learn its main features in a couple of days. He has never experienced any manufacturing defects.

How satisfied is Mark with his smartphone?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if smartphone and fj11a001=2 and fj11a002=1

- fj11a037: Anne works part-time, makes various sports and has a lot of friends. She thinks a smartphone can meet her needs. In a shopping centre she immediately bought what she liked. Unfortunately, after two months she experienced a problem in the phone book. She came back to the shopping centre and the smartphone was withdrawn for the assistance. After ten days the phone was delivered to Anne and the problem was solved. She has not experienced any other manufacturing defects.

How satisfied is Anne with her smartphone?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if LCD television and fj11a001=2 and fj11a002=1

- fj11a038: Carry works all day. During the evening she likes watching TV together with her children and husband. She bought a 34 Inch LCD TV with several options, paying a price higher than expected. At home, she needed the assistance to set it up, because the user guide was not clear. Afterwards, she has never experienced any other problems.

How satisfied is Carry with her LCD TV?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if LCD television and fj11a001=2 and fj11a002=1

- fj11a039: Jim likes watching sports on TV together with his friends. In a specialized shop he bought a 40 Inch LCD TV with a high resolution. He had to wait one week in order to receive it. He has never experienced any manufacturing defects or other problems.

How satisfied is Jim with his LCD TV?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

if fj11a005=1 or fj11a006=1 or fj11a007=1 and fj11a001=2

intro Now we would like to ask one question about how satisfied you are with some aspects of your purchase.

if fj11a005=1 or fj11a006=1 or fj11a007=1 and fj11a001=2 and fj11a002=1

- fj11a040: How satisfied are you with your [if laptop/smartphone/LCD television*]?

1. Very dissatisfied
2. Dissatisfied
3. Neither satisfied, nor dissatisfied
4. Satisfied
5. Very satisfied

Bibliografia

- Agresti, A. (2010). *Analysis of ordinal categorical data* (Vol. 656). John Wiley & Sons.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Angelini, V., Cavapozzi, D., & Paccagnella, O. (2011). Dynamics of reporting work disability in Europe. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3), 621-638.
- Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2014). Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for Individual-Specific Scale Biases. *Oxford bulletin of Economics and Statistics*, 76(5), 643-666.
- Au, N. and Lorgelly, P.K. (2014). Anchoring vignettes for health comparisons: an analysis of response consistency, *Quality of Life Research*, 23(6), 1721-1731.
- Bago d'Uva, T., Lindeboom, M., O'Donnell, O., & van Doorslaer, E. (2011). Slip-ping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46(4), 875-906.
- Bago d'Uva, T., O'Donnell, O., & van Doorslaer, E. (2008). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, 37(6), 1375-1383.
- Basei, C. (2017). *Anchoring vignettes in customer satisfaction surveys: a longitudinal comparison*. Thesis, University of Padua.
- Bound, J., & Krueger, A. B. (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?. *Journal of Labor Economics*, 9(1), 1-24.
- Capecchi, S., Iannario, M., & Piccolo, D. (2012). *Modelling Job Satisfaction in AlmaLaurea Surveys* (No. 56).
- Chevalier, A., & Fielding, A. (2011). An introduction to anchoring vignettes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3), 569-574.
- Corduas, M., Iannario, M., & Piccolo, D. (2009). A class of statistical models for evaluating services and performances. *Statistical methods for the evaluation of educational services and quality of products*, 99-117.

- D'Elia, A., & Piccolo, D. (2005). A mixture model for preferences data analysis. *Computational Statistics & Data Analysis*, 49(3), 917-934.
- Piccolo, D., & D'Elia, A. (2008). A new approach for modelling consumers' preferences. *Food Quality and Preference*, 19(3), 247-259.
- Di Iorio, F., & Iannario, M. (2012). Residual diagnostics for interpreting CUB models. *Statistica*, 72(2), 163.
- Fligner, M. A. (1993). *Probability models and statistical analyses for ranking data* (Vol. 80). J. S. Verducci (Ed.). New York: Springer-Verlag.
- Greene, W. H., & Hensher, D. A. (2010). *Modeling ordered choices: A primer*. Cambridge University Press.
- Greene, W. H., & Hensher, D. A. (2010). Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. *Transportation*, 37(3), 413-428.
- Hopkins, D. J., & King, G. (2010, March). Improving anchoring vignettes designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly*, 74, 1-22.
- Iannario, M. (2008). A class of models for ordinal variables with covariates effects. *Quaderni di Statistica*, 10, 53-72.
- Iannario, M. (2010). On the identifiability of a mixture model for ordinal data. *Metron*, 68(1), 87-94.
- Iannario, M., & Piccolo, D. (2010). A new statistical model for the analysis of customer satisfaction. *Quality Technology & Quantitative Management*, 7(2), 149-168.
- Iannario, M., & Piccolo, D. (2012). CUB models: Statistical methods and empirical evidence. *Modern Analysis of Customer Surveys: with applications using R*, 231-258.
- Iannario, M., & Piccolo, D. (2014). Inference for CUB models: a program in R. *Stat & Appl*, 12, 177-204.
- Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, 36(3), 347-387.
- Kapteyn, A., Smith, J. P., & van Soest, A. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *The American Economic Review*, 97(1), 461-473.
- Kapteyn, A., Smith, J. P., van Soest, A., & Voňková, H. (2011). Anchoring Vignettes and Response Consistency, WR-840. *RAND Labour and Population Working Paper Series*. RAND Corporation: Santa Monica, CA.

- Kerkhofs, M., & Lindeboom, M. (1995). Subjective health measures and state dependent reporting errors. *Health economics*, 4(3), 221-235.
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American political science review*, 98(1), 191-207.
- Kreider, B. (1999). Latent work disability and reporting bias. *Journal of Human Resources*, 734-769.
- Krishnan, T., & McLachlan, G. (1997). The EM algorithm and extensions. *Wiley*, 1(997), 58-60.
- Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour economics*, 15(1), 96-117.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Marden, J. I. (1996). *Analyzing and modeling rank data*. CRC Press.
- McCullagh, P. (1980). Regression models for ordinal data. *Journal of the royal statistical society. Series B (Methodological)*, 109-142.
- McCullagh, P. (1984). Generalized linear models. *European Journal of Operational Research*, 16(3), 285-292.
- McCullagh, P., & Nelder, J. A. (1989). Generalized Linear Models, no. 37 in Monograph on Statistics and Applied Probability.
- McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1), 103-120.
- Moustaki, I. (2003). A general class of latent variable models for ordinal manifest variables with covariate effects on the manifest and latent variables. *British Journal of Mathematical and Statistical Psychology*, 56(2), 337-357.
- O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables* (No. 146). Sage.
- Paccagnella, O. (2014). A New Tool for Measuring Customer Satisfaction: the Anchoring Vignette Approach. *Italian Journal of Applied Statistics*.
- Paccagnella, O. (2011). Anchoring vignettes with sample selection due to non-response. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3), 665-687.
- Paccagnella, O. (2013). Modeling individual heterogeneity in ordered choice models: Anchoring vignettes and the Chopit model. *Journal of Methodological and Applied Statistics*, 15, 69-89.

- Peel, D., & McLachlan, G. J. (2000). Robust mixture modelling using the t distribution. *Statistics and computing*, 10(4), 339-348.
- Peracchi, F., & Rossetti, C. (2013). The heterogeneous thresholds ordered response model: Identification and inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(3), 703-722.
- Piccolo, D. (2003). On the moments of a mixture of uniform and shifted binomial random variables. *Quaderni di Statistica*, 5(1), 85-104.
- Piccolo, D. (2006). Observed information matrix for MUB models. *Quaderni di Statistica*, 8, 33-78.
- Rice, N., Robone, S., & Smith, P. C. (2008). The measurement and comparison of health system responsiveness. *Centre for Health Economics, University of York*.
- Rice, N., Robone, S. and Smith, P.C. (2011). Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness, *The European Journal of Health Economics*, 12, 141-162.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.
- Van Soest, A., Delaney, L., Harmon, C. P., Kapteyn, A., & Smith, J. P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(3), 575-595
- Van Soest, A., & Vonkova, H. (2014). Testing the specification of parametric models by using anchoring vignettes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(1), 115-133.
- Vonková, H. and Hullege, P. (2011). Is the anchoring vignettes method sensitive to the domain and choice of the vignette?, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174, 597-620.

Ringraziamenti

Vorrei esprimere un grande ringraziamento al mio relatore, il professor Omar Pacagnella, per tutto il suo aiuto, la disponibilità e la pazienza dimostrate negli ultimi mesi. Le Sue conoscenze, il Suo incoraggiamento e la Sua guida hanno permesso la stesura di questa tesi. Non si potrebbe desiderare una miglior supervisione.

Vorrei ringraziare i miei genitori per il supporto che non mi hanno mai fatto mancare, anche se distanti. Non sarei riuscita ad arrivare dove sono oggi senza di loro.

Ringrazio immensamente anche le mie amiche Sara e Sara che, nonostante i chilometri che ci dividevano, sono sempre state un sostegno immancabile. 3S, ma una cosa sola.

Rimane da ringraziare la persona che più di tutte è stata al mio fianco, nei momenti più belli, ma anche in quelli più difficili, che non mi ha mai lasciato sola e che è stata la mia àncora: Federico. Non so cosa il futuro ci riserverà, ma spero di scoprirlo insieme a lui.