

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



RELAZIONE FINALE

**SODDISFAZIONE DELLA VITA E DEL REDDITO: EVIDENZE
EMPIRICHE TRA ALCUNI PAESI EUROPEI**

Relatore: Prof. Omar Paccagnella
Dipartimento di Scienze Statistiche

Laureando: Ilaria Zanin
Matricola: N 1103976

Anno Accademico 2016/2017

*A mamma
perchè la miglior arma per vincere una battaglia è l'amore.*

Indice

Introduzione	7
1 Soddisfazione della vita e del reddito	11
1.1 Il benessere	11
1.1.1 Il benessere soggettivo	15
1.2 La soddisfazione della vita	16
1.3 La soddisfazione del reddito	20
1.4 Obiettivo dell'analisi	24
2 Modelli per variabili risposta ordinali	27
2.1 L'approccio delle <i>anchoring vignettes</i>	27
2.1.1 <i>Differential Item Functioning</i>	28
2.1.2 La tecnica dell' <i>anchoring vignettes</i>	30
2.1.3 Le assunzioni	34
2.1.4 I modelli	37
2.1.4.1 Il modello parametrico Chopit	37
2.1.4.2 Estensioni del modello parametrico	44
2.1.4.3 Il metodo non parametrico	49
2.1.4.4 Vantaggi e svantaggi del metodo parametrico e non parametrico	57
2.1.5 Selezione delle <i>vignettes</i>	60
2.1.6 Verifica delle assunzioni	62
2.2 Modello probit ordinato bivariato	70
2.3 Modello di selezione del campione per variabili ordinali	72
2.4 Alcune misure ordinali di associazione	76

3	Applicazioni empiriche	79
3.1	L'indagine di SHARE	79
3.2	Descrizione del dataset	81
3.3	Analisi descrittive	86
3.3.1	Analisi sulle variabili esplicative	86
3.3.2	Analisi sulla soddisfazione	94
3.3.2.1	Soddisfazione del reddito	96
3.3.2.2	Soddisfazione della vita	102
3.3.2.3	Analisi congiunta delle due soddisfazioni . . .	108
3.4	Analisi multivariate	110
3.4.1	Applicazione della tecnica delle <i>vignettes</i>	110
3.4.2	Stima dei modelli	119
	Conclusioni	143
	Elenco delle tabelle	148
	Elenco delle figure	150
A	Descrizione delle variabili in scala europea	151
	Ringraziamenti	155
	Bibliografia	157

Introduzione

I ricercatori sono interessati ad osservare fenomeni con lo scopo di estrarre dei risultati sintetici e significativi: i campi ricercati sono numerosi, da quelli più elementari e a noi vicini, a quelli più complessi. Spesso l'oggetto di indagine è quantitativo e dunque misurabile (ad esempio la temperatura atmosferica, il reddito, ecc), ma altre volte esso è astrattivo e soggettivo (la felicità, la depressione dei soggetti, ecc).

Uno dei temi maggiormente dibattuto in letteratura è il benessere, un concetto che difficilmente può essere racchiuso in un'unica, esemplificativa, globale e al tempo stesso valida definizione. La complessità del fenomeno è tale per cui non risulta così ovvio e banale definire e circoscrivere oggettivamente un concetto così astratto e/o soggettivo. Per molto tempo infatti si è considerato il PIL come lo strumento principale e rappresentativo del benessere; tuttavia ormai molti studiosi convergono in un pensiero comune: un indicatore così "debole" a livello informativo in quanto relativo ad una sola dimensione del benessere costituisce una definizione insufficiente per una realtà che non è e non può essere limitata solo a quella economica. Per tale motivo, molti indicatori oggettivi che misurano tale fenomeno prendono in considerazione diversi fattori, tra cui il benessere soggettivo o alternativamente la soddisfazione della vita, un concetto che per natura non è estraneo ad elementi di soggettività come la personalità, il passato, ecc. Se da una parte tale realtà pone l'individuo come punto focale di studio e dunque comprende tutti quegli elementi estranei all'oggettività, dall'altra aggrega in un'unica definizione diversi suoi domini e valutazioni, come la soddisfazione del lavoro, della salute, sociale, ecc. Ciò rappresenta contemporaneamente un vantaggio

e uno svantaggio perchè consente di avere una valutazione soggettiva e diretta dell'individuo, ma al tempo stesso è legata a tutti i problemi che nascono con valutazioni personali (si veda il Capitolo 1).

Un'ampia parte della letteratura analizza i dati relativi al benessere soggettivo; tuttavia l'errore comune in cui spesso si imbattono gli studiosi è quello di trascurare e dunque ignorare la natura stessa di tale realtà, limitandosi ad analizzare, attraverso appropriati modelli, la relazione tra l'autovalutazione generale della vita e alcune covariate osservabili. Questa tesi tiene conto di tale constatazione e nello specifico prendere in considerazione uno degli aspetti che inevitabilmente è legato concettualmente alla valutazione di un dominio più complesso ovvero la soddisfazione monetaria: infatti un fattore estremamente importante nella vita degli individui, economicamente ed emotivamente, e spesso considerato (erroneamente) come rappresentante del benessere è il reddito. Se è vero che alcune pubblicazioni trovano evidenza del fatto che i soldi non comprano la felicità, attualmente non si può affermare altrettanto della soddisfazione soggettiva del reddito. Dunque, l'obiettivo che si prefigge tale tesi consiste nell'individuare e verificare l'esistenza di una possibile correlazione tra la soddisfazione della vita e quella di un suo dominio economico-monetario, la soddisfazione del reddito. Più precisamente, l'interesse ricade nell'associazione di queste due dimensioni e non in una possibile relazione causale di una sull'altra. Accanto a ciò, si vogliono studiare e determinare i fattori che maggiormente influenzano le due soddisfazioni. Tale argomento in sé non è nuovo in letteratura. Tuttavia, dovendo analizzare valutazioni soggettive, si rischia di imbattersi nel seguente problema: nel momento in cui esprime il suo giudizio, un individuo può usare diverse scale di risposta; in altre parole, persone che hanno un differente background culturale e/o socio-economico, o semplicemente gruppi di individui entro la stessa popolazione o Paese potrebbero comprendere differentemente una stessa domanda. Questa eterogeneità che potrebbe sorgere e che porterebbe ad un'incomparabilità delle risposte tra differenti gruppi è nota in letteratura sotto il nome di DIF (*Differential Item Functioning*), un problema che, se ignorato, potrebbe portare a risultati fuorvianti. Un possibile strumento in grado di correggere tale problematicità è la tecnica delle *anchoring vignettes*: attraverso l'uso di domande aggiuntive basate sulla sfera d'interesse e relative

a soggetti ipotetici è possibile modellare la DIF riscaldando l'autovalutazione degli individui in funzione delle risposte fornite proprio a queste domande aggiuntive. Ciò è possibile grazie al fatto che ciascuna domanda aggiuntiva vede come protagonista un soggetto ipotetico e descrive scenari di situazioni uguali per tutti gli intervistati, riguardo un particolare concetto, fornendo così "un'ancora" di valutazione e confronto. Conseguentemente, l'obiettivo descritto precedentemente assume una connotazione più ampia perchè si vuole analizzare la relazione tra soddisfazione della vita e soddisfazione del reddito correggendo per l'eterogeneità che potrebbe sorgere tra cittadini di diversi Paesi o gruppi, così da non incorrere in conclusioni distorte, un problema di misura di cui la letteratura non ha ancora tenuto conto in maniera esauriente.

Questo lavoro, suddiviso in tre Capitoli, cerca di rispondere alle domande poste attraverso l'analisi dei dati riguardanti alcuni intervistati europei partecipanti all'indagine SHARE tra il 2006 e il 2007.

Il Capitolo 1, in veste discorsiva, fornisce un'introduzione generale alle due sfere d'indagine del benessere, focalizzandosi inoltre sui principali fattori che in letteratura risultano avere un'influenza statisticamente significativa su queste autovalutazioni. Segue una esposizione più accurata dell'obiettivo prefissatosi in tale tesi.

Il Capitolo 2 si pone l'obiettivo di presentare gli strumenti statistici utilizzati in questo lavoro: la tecnica delle *anchoring vignettes*, il modello probit ordinale bivariato e il modello di selezione per variabili ordinali.

Infine nel Capitolo 3 viene fatta una breve panoramica riguardo l'indagine SHARE e la descrizione del dataset analizzato. Segue la presentazione di alcune analisi descrittive associate ai due domini d'interesse e l'esposizione dei risultati emersi dai modelli considerati ed esposti nel Capitolo precedente: lo scopo è quello di rispondere al quesito postosi verificando la validità degli strumenti sfruttati.

La Sezione "Conclusioni" conclude questo lavoro di tesi, evidenziando i principali risultati e fornendo spunti per sviluppi futuri di queste analisi.

Capitolo 1

Soddisfazione della vita e del reddito

1.1 Il benessere

Il benessere costituisce uno dei temi più discussi e dibattuti in letteratura, a cui si dedicano diversi psicologi, statistici, economisti, filosofi, istituzioni, ecc. Al di là del perché di questo interesse, viene naturale chiedersi cosa sia effettivamente il benessere; il dizionario Treccani (2017) riporta le seguenti definizioni:

"1. Stato felice di salute, di forze fisiche e morali. (...) - 2. Condizione prospera di fortuna, agiatezza (...): in particolare, nell'economia moderna, benessere economico, stato di agiatezza collettiva ottenuto attraverso una larga disponibilità dei beni di consumo e un'equa distribuzione della moneta. Con l'espressione economia del benessere si intende anche una corrente di pensiero economico (...) che si è proposta il problema di come influire sull'ordinamento economico per accrescere il benessere di tutti (vedi anche welfare, welfare economics e welfare state). - 3. Sensazione soggettiva di vita materiale piacevole, nell'espressione condizioni di benessere con cui vengono indicati, in termotecnica, i livelli a cui devono trovarsi i vari fattori che influenzano l'abitabilità di un ambiente chiuso (temperatura dell'aria, umidità relativa dell'aria alla temperatura ambiente, temperatura media radiante dalle pareti, ecc.), perché le persone che vi devono soggiornare per un tempo abbastanza

lungo (in genere, superiore a un'ora) si trovino a loro agio, senza avvertire sensazioni spiacevoli di caldo o di freddo."

È evidente come non esista un'unica definizione di benessere. La complessità del fenomeno è tale per cui risulta complicato definire e circoscrivere oggettivamente un concetto così astratto e/o soggettivo e per cui, in senso stretto, non appare ad oggi possibile pervenire ad una sua effettiva e "ultima" misura oggettiva e al tempo stesso valida. Da diversi anni si è sviluppato un acceso dibattito per giungere a una definizione degli indicatori dei livelli di benessere più congrua di quanto non si possa derivare dalla teoria economica. Infatti tradizionalmente gli economisti hanno espresso il benessere (sociale e/o economico) con alcuni indicatori quantitativi, quali ad esempio gli indici di produzione e di consumo di beni e di servizi, il livello di reddito, il tasso di disoccupazione e di crescita industriale. Addirittura, con la diffusione della teoria keynesiana, si è sviluppata una forma di misurazione aggregata (come il PIL -Prodotto Interno Lordo- o in inglese GPD - *Gross Domestic Product*) quale preferibile variabile per la misurazione del benessere economico: in particolare dal periodo successivo alla grande depressione e alla seconda guerra mondiale si è considerato questo tipo di benessere come sinonimo del benessere generale (Giovannini, 2012). Il PIL, considerato da molti come strumento principale per la misurazione dello sviluppo, è tuttavia un indicatore relativo a una sola dimensione del benessere e quindi una definizione insufficiente, sia perché esclude in realtà tutte quelle attività e servizi che sfuggono ad una valutazione di mercato (volontariato, lavoro domestico, ecc), sia perché non considera aspetti sociali (di tipo ambientale, psicologico-sociale, ecc.) che incidono sulla reale struttura di un'economia. Inoltre, limitatamente a questa realtà, assumere la crescita del PIL come indicatore di un benessere maggiore porta a pensare ad un legame monetario, ovvero ad un rapporto diretto tra consumo e benessere, una realtà molto limitata.

Se invece si assume il concetto di benessere in termini più generici, si considera una misurazione composita della qualità della vita e della soddisfazione che non necessariamente è funzione diretta e lineare dei livelli di ricchezza e di consumo. In questo caso, quindi, le variabili da analizzare fanno affidamen-

to, accanto agli aspetti economici, a fattori quali l'istruzione, le condizioni di salute della popolazione, l'ambiente di vita, l'alienazione sociale, ecc, definendo così un concetto ben più ampio di quello economico e sicuramente più efficiente ed appropriato.

Lo sviluppo e il rafforzamento degli indicatori statistici su aspetti sociali ed ambientali e la loro accresciuta diffusione hanno indubbiamente alimentato il dibattito per la costruzione di un quadro informativo che permetta di valutare il livello di benessere e la sua sostenibilità nel tempo (ISTAT, 2016): infatti se gli indicatori utilizzati non sono corretti o non colgono tutte le caratteristiche del benessere, essi possono indurre a prendere decisioni inefficaci o sbagliate (Giovannini, 2012). Un punto di svolta è rappresentato dalla presentazione nel 1990 da parte dell'Onu del primo indice comparativo macroeconomico, l'Indice di Sviluppo Umano (ISU o meglio noto mediante l'acronimo inglese HDI - *Human Development Index*), che affianca l'indicatore sul reddito procapite con quelli sulla speranza di vita e sul livello di istruzione. Nel tempo l'HDI ha inglobato inoltre aspetti legati alla sostenibilità ambientale, un aspetto che non può essere sottovalutato nella nozione di benessere.

Dunque, l'assunzione di un concetto così ampio e di difficile definizione come quello di benessere solleva a sua volta dei problemi che non possono trovare una soluzione unitaria per definirlo, quantificarlo e allo stesso tempo per andar oltre le misure economiche convenzionali come il PIL pro capite; tali problematiche sono diventati cruciali anche nei contesti istituzionali. In questo scenario si inserisce, ad esempio, lo sviluppo dell'indice BES (Benessere Equo Sostenibile), realizzato nel 2011 dall'Istat e dal CNEL, che si pone l'obiettivo di individuare gli indicatori statistici più adeguati al fine di misurare ciascun dominio riconducibile al benessere; in particolare sono stati identificati 12 domini e 130 indicatori che tengono conto sia di aspetti che hanno un diretto impatto sul benessere umano ed ambientale sia di quelli che misurano gli elementi funzionali al miglioramento del benessere della collettività e dell'ambiente che la circonda. Tale indice non valuta quindi solo una dimensione (come l'aspetto economico, come fa il GDP), bensì attraverso delle misure sintetiche ne aggrega diversi (compresa quella del PIL), fra i quali la salute, l'istruzione e la formazione, il lavoro, il benessere economico, le relazioni sociali, il benessere soggettivo, l'ambiente: in questo modo si rende più

agevole il confronto e l'analisi dei fenomeni osservati, rimandando ai singoli indicatori per ulteriori approfondimenti¹.

Una soluzione simile viene proposta dall'Organizzazione per la cooperazione e lo sviluppo economico (OCSE o in inglese OECD), che a partire dal 2001 ha proposto diverse iniziative. A discapito dell'assenza di una singola definizione, OECD (2011) sottolinea che ogni determinante del benessere correla fortemente con le altre; nello specifico l'organizzazione identifica tre pilastri per capire e misurare questo complesso fenomeno, come si può vedere nella Figura 1.1: le condizioni di vita materiale o benessere economico, la qualità della vita (un insieme di attributi non monetari) e infine la sostenibilità dei sistemi socio-economici e naturali dove le persone vivono e lavorano. In particolare, si noti come al centro della valutazione vi siano le circostanze di vita (oggettive e soggettive) ed esperienze del benessere delle persone (individui e famiglie) e come queste siano messe in relazione con fattori temporali del benessere. Conseguentemente le prime due componenti si riferiscono ad un benessere corrente (si può notare la similitudine con i domini del BES), mentre la terza a quello futuro. La selezione degli indicatori che vengono utilizzati per ciascuna dimensione si basa sulla rilevanza politica, sulla qualità dei dati che si stanno considerando, sulla confrontabilità dei concetti e delle domande d'indagine usate e su alcune richieste statistiche (per maggior dettagli si veda OECD (2011)). Dunque la difficoltà e la complessità di individuare o definire un indicatore soddisfacente in grado di considerare tutti i (differenti) fattori non è così banale. Chiaramente sia nella proposta dell'ISTAT (2016) che in quella dell'OECD (2013) vi è l'elevata possibilità che ognuna delle componenti individuate e ciascuna dimensione che le costituisce possano differire tra Paesi e culture.

Altri indici che seguono l'impianto dello sviluppo teorico e metodologico dell'Indice di Benessere Equo Sostenibile e dell'OECD, e dunque che integrano indicatori che tengano conto di fenomeni socio-economici che il PIL stesso non è in grado di cogliere sono ad esempio quello creato da *Barilla Center for Food & Nutrition* (BCFN), il *Genuine Progress Indicator* (GPI), *The Happy Planet Index* (HPI) o il *Canadian Index of Wellbeing* (CIW), solo per citarne

¹Per consultare gli indicatori riassunti di ciascun dominio si consulti la Tavola 1 ISTAT (2016, pg. 14)

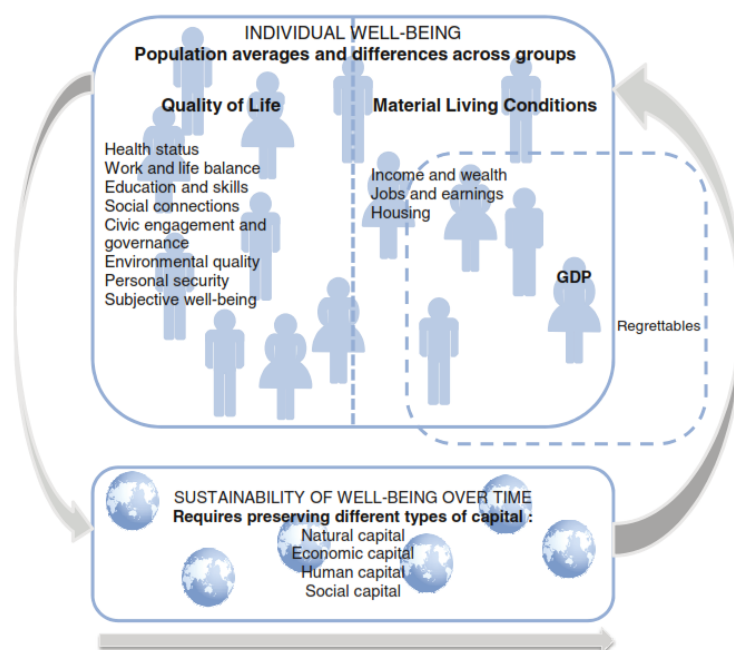


Figura 1.1: Struttura per gli indicatori del benessere secondo OECD. Fonte: OECD (2011).

alcuni².

In conclusione dunque è necessario tenere conto del diverso numero di fattori, rinunciando a definizioni che enfatizzino un elemento o un aspetto particolari a scapito di altri. Il benessere è perciò, in prima battuta, l'esito di una condizione oggettiva legata al verificarsi di una serie molto ampia di fenomeni positivi, o comunque non negativi, che possono cambiare ed evolversi.

1.1.1 Il benessere soggettivo

Scegliere di adottare l'individuo come unico punto focale di studio pone un problema fondamentale (oltre a quelli già presentati nella Sezione precedente), legato alla logica di scelta, misurazione e ponderazione dei diversi fattori soggettivi e oggettivi che concorrono a definire il benessere (Food & Nutrition, 2010). Si è appena visto che alcune soluzioni che misurano il benessere prendono in considerazione il benessere soggettivo come una delle sue

²È possibile consultare una breve descrizione di questi indicatori in Food & Nutrition (2010) per i primi 3 e in <https://uwaterloo.ca/canadian-index-wellbeing/>

dimensioni (OECD, 2011; ISTAT, 2016). Certamente l'inclusione della valutazione personale del benessere rende questa misura più completa e aderente alla reale valutazione della qualità della vita in quanto consente di valutare le eventuali divergenze tra ciò che le persone riferiscono e ciò che viene catturato dagli indicatori oggettivi. Tale processo però ingloba degli elementi di soggettività nella misurazione (ad esempio la cultura, la personalità, ecc), rendendo complessi i confronti fra individui e Paesi diversi. Per questo motivo diversi indici relativi al benessere non considerano tale possibilità, contrariamente ad altri, come l'HDI, che tuttavia rimangono delle misurazioni oggettive.

La mancanza di indicatori che comprendano aspetti soggettivi ed oggettivi, e che allo stesso tempo siano validi indicatori individuali, porta una grande parte dei ricercatori ad indagare separatamente il benessere soggettivo, come una definizione alternativa a quella più generale; così, l'obiettivo diventa quello di identificare la relazione tra le misure soggettive del benessere e le caratteristiche economiche, demografiche, sociali e fisiche di una popolazione. Ciò che però non può essere accettato è che questo corrisponda al "benessere" o che ne costituisca una misura esaustiva (Michalos, 2007): non per niente OECD (2011) e ISTAT (2016) considerano il "benessere soggettivo" come una componente essenziale nella valutazione del "benessere".

La letteratura sul benessere individuale tipicamente distingue tra una componente affettiva e una componente di valutazione del benessere (Michalos (1985), ISTAT (2016), Eurostat (2016)). In particolare la seconda comprende il benessere soggettivo: il dominio di tale concetto riguarda essenzialmente le valutazioni e le percezioni espresse direttamente dagli individui sulla loro vita in generale. Per tale motivo questo aspetto è spesso considerato come la valutazione sulla soddisfazione della vita, un aspetto globale della qualità della vita (Eurostat, 2016) che verrà quindi analizzato con maggior dettaglio nella prossima Sezione.

1.2 La soddisfazione della vita

Di fronte al quadro descritto nelle pagine precedenti e all'impossibilità di fornire concretezza analitica a concetti essenzialmente astratti in un unico indicatore, la soddisfazione della vita sembra essere un criterio alternativo,

coerente e soddisfacente, specialmente quando si fa riferimento alla qualità e al tenore di vita: al centro di tale aspetto vi è la scelta di adottare l'individuo quale punto focale di studio. In letteratura comunque i termini benessere, felicità e soddisfazione della vita sono solitamente interscambiabili tra i ricercatori in quanto tutti contribuiscono al generico concetto di qualità della vita (nell'ottica individuale), nonchè sono strettamente connessi tra loro.

Nel 2016, l'indicatore relativo alla soddisfazione della vita ha mostrato segni di ripresa, dopo il forte calo registrato tra il 2011 e il 2012 e la sostanziale stabilità nel periodo successivo (ISTAT, 2016). In Figura 1.2 viene riportata la valutazione relativa alla soddisfazione della vita delle persone di 16 anni e più residenti nei 28 Paesi dell'Unione Europea, in una scala da 0 a 10. I dati, riferiti al 2013, evidenziano come il livello di soddisfazione dell'Italia sia leggermente inferiore della media europea, mentre superiore di un punto nei Paesi scandinavi, Danimarca e Finlandia in particolare (punteggio medio pari a 8).

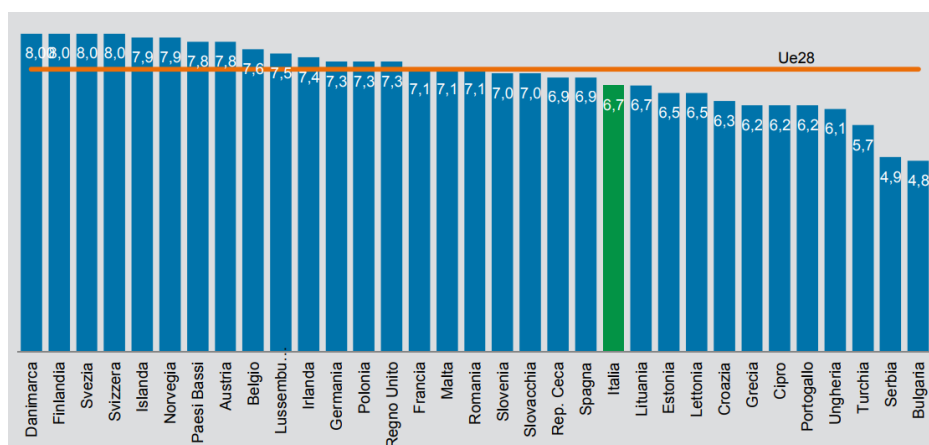


Figura 1.2: Soddisfazione della vita delle persone di 16 anni e più per Paese (Anno 2013). Fonte: ISTAT (2016).

Uno dei dibattiti più discussi in letteratura è quello tra il benessere soggettivo e il reddito. Nella società in cui viviamo il reddito rappresenta una componente importante della vita: un reddito più alto infatti consente, ad esempio, di comprare macchine costose, beni tecnologicamente avanzanti e all'ultima moda o di effettuare viaggi di piacere costosi al di fuori del proprio paese o continente. Spesso la visione generale che si ha è che un reddito più alto

indica un benessere individuale, una soddisfazione, una condizione migliore; si tratta di un pensiero che appoggia la teoria dell'utilità, seconda la quale un reddito più alto consente ad un individuo di raggiungere una curva di indifferenza più alta. Secondo alcuni studi recenti tale fattore rappresenta il "goal" del benessere (inteso come soddisfazione, felicità) soggettivo di un individuo: individui più felici hanno redditi più alti (Carbonell (2005), Clark et al. (2005)); parte di questa correlazione positiva è certamente dovuta a caratteristiche individuali non osservate, come ad esempio la personalità (Dolan et al., 2008), anche se non va dimenticato l'aspetto di causalità inversa (ovvero benessere maggiori determinano redditi futuri più alti). Altre pubblicazioni mostrano invece evidenza del fatto che il benessere correla solo debolmente con il reddito (Easterlin, 2001), che non diminuisce all'aumentare del reddito (Stevenson e Wolfers, 2013) o che i soldi non fanno la felicità (Dolan et al. (2008), Angelini et al. (2014)). Certamente il reddito può rappresentare una componente importante del benessere individuale, ma di base non può essere considerata l'unica.

Fattori come l'età, l'istruzione, il genere, la salute, lo stato occupazionale e le relazioni hanno generalmente una forte associazione con la soddisfazione della vita, indipendentemente dal modello statistico specificato: un approccio che sfrutta classi latenti (Clark et al., 2005), probit ordinale con effetti casuali (Carbonell, 2005), probit ordinale per dati *pooled* (Clark, 2006), logit ordinale (Peiro, 2006), modello Chovit (Angelini et al., 2014). Tuttavia, non vi è sempre concordanza sulla sua positività o negatività: ad esempio Angelini et al. (2014) trovano evidenza del fatto che la soddisfazione decresce se si è di genere maschile, se si ha un'istruzione più bassa, una salute peggiore o non si ha un lavoro, a parità di altre condizioni, mentre aumenta col fatto di avere un partner e con l'avanzare dell'età; Peiro (2006) concorda con l'effetto delle associazioni individuate da questi autori tranne per l'età, che mostra invece un impatto opposto. L'evidenza riguardo l'associazione tra il benessere e l'essere genitori non è univoca (Dolan et al., 2008): in alcune indagini l'effetto non è statisticamente significativo o lo è debolmente per alcuni Paesi (Peiro, 2006), in altre è negativo (Clark (2006), Angelini et al. (2014)).

Generalmente queste sono le variabili più frequentemente considerate in letteratura; è tuttavia possibile consultare una descrizione dettagliata di altri

fattori che possono avere un'influenza sul benessere in Diener (1984) e Dolan et al. (2008), come le relazioni sociali, l'intelligenza, la religione, l'etnia ecc. Non mancano inoltre studi che mostrano differenze anche tra i Paesi: Svezia, Danimarca e Olanda sono quelli solitamente più "felici", Italia e Grecia quelli meno (Clark et al. (2005), Clark (2006), Angelini et al. (2014)).

La soddisfazione della vita può essere vista, nel complesso, come un concetto aggregato che viene spiegato attraverso diversi suoi domini, come la soddisfazione del lavoro, della salute, finanziaria, sociale, amorosa, della propria istruzione, ambientale, del reddito, ecc. Spesso i ricercatori non considerano questa possibilità, ma si limitano ad analizzare la relazione tra l'autovalutazione di questo dominio generale e alcune covariate osservabili attraverso modelli per variabili ordinali. Coloro che invece ritengono importante analizzare anche questi aspetti del benessere lo fanno limitatamente ad un paio di componenti (ad esempio Peiro (2006) ne considera solo tre, la felicità, la soddisfazione della vita e la soddisfazione finanziaria degli individui di 15 Paesi) o escludendo l'analisi della soddisfazione della vita (ad esempio Bonsang e Van Soest (2012) analizzano la soddisfazione con il reddito e quella con il lavoro di individui europei ultracinquantenni). Diversamente, Van Praag et al. (2003) tengono in considerazione sei domini (si veda la Figura 1.3). Vero è che probabilmente alcune determinanti possono essere influenzate dalle altre: ad esempio la soddisfazione sul lavoro ha quasi sicuramente un'associazione con la soddisfazione della salute. Pertanto, la Figura 1.3 deve essere vista come un modello riduttivo nel quale si assume che non vi siano cross-relazioni tra le diverse sfere del benessere (Van Praag et al., 2003). Chiaramente la soddisfazione individuale dipende non soltanto dalle situazioni oggettive del soggetto, ma anche da caratteristiche inosservabili, come la sua personalità (Van Praag et al. (2003) le considerano ad esempio invarianti nel tempo). Analizzando un panel longitudinale relativo a delle famiglie tedesche (distintamente per lavoratori e non della Germania dell'Est e di quella dell'Ovest), gli autori sfruttano un modello ad equazioni simultanee: prima stimano separatamente la relazione tra ciascun dominio con l'insieme X delle covariate scelte (misurabili oggettivamente), successivamente si dedicano a quella del benessere generale con ciascuna singola sua dimensione. In particolare nella

seconda, gli autori inseriscono una variabile Z non osservabile, invariante nel tempo e comune a tutti i domini del benessere, rappresentante tratti inosservabili (come la personalità degli individui o la salute dei figli). Inoltre, per evitare che le singole sfere del benessere siano correlate tra loro e con il termine d'errore della seconda equazione, questa variabile Z viene strumentalizzata³. La conclusione principale a cui giungono gli autori è che si può spiegare il concetto generale di benessere attraverso le sue componenti misurate oggettivamente in tutti e quattro i sottocampioni e che questi presentano circa lo stesso ordine di magnitudo: in particolare la soddisfazione finanziaria, della salute e del lavoro sono le principali determinanti del benessere, mentre l'ambiente e la casa quelli più marginali, specialmente per i cittadini tedeschi dell'Ovest. Per quanto riguarda le covariate osservate, età, reddito ed istruzione sono quelle con associazione più significativa per tutti i domini, anche se gli impatti cambiano da una sfera all'altra e da un sottocampione a un altro; in particolare l'età sembra presentare un comportamento comune parabolico.



Figura 1.3: Modello a due stadi di Van Praag et al. (2003). Fonte: Van Praag et al. (2003).

1.3 La soddisfazione del reddito

Ogni individuo e ogni famiglia dispone di un reddito annuo che può essere differente tra individui che abitano nella stessa città, nella stessa provincia

³Dopo aver stimato i sei domini, si calcolano i residui in modo da stimare la parte Z comune a tutti i residui.

o nello stesso Paese. Il reddito rappresenta un fattore estremamente importante non solo nella vita degli individui, economicamente ed emotivamente, ma anche nell'interesse di enti come, ad esempio, lo Stato: esso consente agli individui l'acquisto di diversi beni e servizi, di vivere in una data località, ecc, mentre allo Stato dà la possibilità di conoscere le entrate del cittadino, ad esempio, attraverso il documento denominato "Dichiarazione dei redditi", e dunque, specificatamente per il livello di reddito di ogni cittadino, permette da una parte di poter fissare i versamenti delle imposte su di essi dovute, dall'altra di offrire delle prestazioni economiche per le famiglie che si trovano in una situazione economica non agevole (si pensi ad esempio al reddito di inclusione, una misura varata in Italia nell'Aprile 2017 e che entrerà in vigore il prossimo anno al fine di assicurare il sostegno economico a tutti i nuclei familiari che hanno una soglia di povertà inferiore a quella assoluta e che soddisfino alcuni requisiti).

La domanda a cui bisogna rispondere è perchè si è interessati al reddito. Molto spesso il reddito viene considerato come l'elemento principale, necessario e determinante di diversi fattori: il tenore di vita degli individui dipende dalla soddisfazione di bisogni essenziali (come le spese di tutti i giorni, l'abbigliamento, l'alloggio, l'assistenza sanitaria, l'istruzione), ma anche da bisogni (come il divertimento e la comunicazione) che vanno al di là dell'essenziale; tali necessità sono garantite o comunque vengono contribuite dal reddito e dal risparmio che da esso si riesce ad accumulare. Il reddito gioca quindi il ruolo di *proxy* per il *welfare* economico e/o di motore delle risorse (Cowell, 2007), anche se non può essere considerato un elemento a se stante: nel modello keynesiano e nel sistema economico raffigura un elemento chiave per garantire la ciclicità dell'insieme di flussi di merci e denaro che circolano tra quattro operatori principali (famiglie, Stato, imprese e sistema bancario). Proprio anche per questo motivo costituisce una delle componenti da considerare negli indicatori del benessere.

Se si pensa al reddito, si potrebbe ritenere semplice delinearne il concetto e i fattori che lo influenzano, ma in realtà in base al tipo di esercizio (lavoro, rendita, attività imprenditoriale, ecc) e a seconda del soggetto che lo percepisce (persone fisiche - si distinguono tra individuo e nucleo familiare - o società, ente o azienda) presenta una definizione e caratteristiche proprie. Comunque,

esso può essere riassunto come la ricchezza (ovvero la quantità di soldi, beni o servizi) che fluisce ad un periodo ben definito (un quinquennio, un anno, un semestre, ecc.) e che un individuo o un'attività commerciale riceve in cambio di un bene o di un servizio o attraverso l'investimento di capitale in asset finanziario; in particolare, trattandosi di un compenso, potrebbe dunque essere soggetto a cambiamenti in un lasso di tempo dato.

Poichè il reddito può rappresentare una delle componenti più importanti per un individuo, è naturale pensare che la sua soddisfazione possa influenzare in qualche modo quella del benessere. Generalmente però i ricercatori sono propensi a disinteressarsi di ciò: la letteratura sulla soddisfazione della vita tende infatti ad enfatizzare il ruolo del reddito (come è già stato detto nella Sezione precedente) e quindi ad analizzare il legame diretto di questo sull'autovalutazione soggettiva, piuttosto che considerare la soddisfazione con questa componente monetaria. Sono pochi gli autori che si sono dedicati proprio a quest'ultima relazione, probabilmente perchè vedono una stretta (e naturale) associazione positiva tra la valutazione soggettiva sul proprio reddito e il reddito percepito; conseguentemente, ciò potrebbe indurre a prediligere una misura oggettiva ad una soggettiva. Eppure, come mostrato da Van Praag et al. (2003), questa è una sfera fondamentale del benessere. Chiaramente il concetto di soddisfazione finanziaria (analizzato da Van Praag et al. (2003)) è ben più ampio rispetto a quello di reddito: infatti il primo è basato anche sui risparmi e sulla disponibilità di assets, e non solo sul reddito.

La soddisfazione del reddito viene solitamente studiata per la popolazione adulta nel contesto delle scale di equivalenza dei nuclei familiari (Bonsang e Van Soest, 2012). Come per la soddisfazione della vita, i fattori come l'età, l'istruzione, il genere, vivere con un compagno hanno generalmente una associazione significativa con questa autovalutazione, indipendentemente dal modello statistico considerato, come ad esempio modelli semiparametrici (Bellemare et al., 2002), modello probit ordinale con l'inclusione di un termine di errore casuale e un effetto fisso nel tempo (Ferrer-i-Carbonell e Van Praag, 2003), modello Chopit (Bonsang e Van Soest (2012); Kapteyn et al. (2013)), sebbene non vi è sempre concordanza sul segno. Per esempio Ferrer-i-Carbonell e Van Praag (2003) (analizzano quattro sottogruppi

di individui tedeschi, lavoratori e non, occidentali e orientali), Bonsang e Van Soest (2012) (esaminano un campione di ultracinquantenni di 11 Paesi Europei) e Bellemare et al. (2002) (considerano un campione di individui tedeschi) trovano che la soddisfazione del reddito diminuisce con l'avanzare dell'età, se si è di genere maschile e col numero di soggetti che vivono in casa, a parità di altre condizioni, mentre aumenta al crescere del reddito e se si ha un partner. Confrontando un campione di soggetti olandesi e americani, Kapteyn et al. (2013) confermano questi effetti per quanto riguarda il fatto di non essere single e per un reddito equivalente maggiore, ma individuano associazioni non significative per le altre variabili. Per quanto riguarda l'istruzione, questa ha un'associazione positiva, anche se in Ferrer-i-Carbonell e Van Praag (2003) questa risulta non significativa per i lavoratori dell'est e negativa per i non-lavoratori dell'est. Il fatto di non avere un'occupazione (Ferrer-i-Carbonell e Van Praag (2003), Bonsang e Van Soest (2012)) o di non godere di una buona salute (Bonsang e Van Soest (2012)) porta gli individui a essere non soddisfatti.

Come per la soddisfazione della vita, vi sono differenze tra i Paesi Europei: in particolare i cittadini danesi si riconfermano i più soddisfatti, mentre quelli olandesi i meno (Bonsang e Van Soest, 2012). Attraverso il modello Chopit, Kapteyn et al. (2013) non trovano differenze di valutazioni tra individui olandesi e americani.

Una particolare analisi sulla soddisfazione del reddito di alcuni studenti viene svolta da Crawford Solberg et al. (2002), le cui analisi conducono a risultati inattesi: in tre studi, gli autori verificano la previsione di questa componente del benessere, esaminando se le discrepanze tra le cose che le persone hanno e quelle che desiderano prevedono la soddisfazione. Il primo studio mostra che questa dipende dalla desiderabilità dei beni materiali che possono essere comprati con il reddito; in particolare coloro che li desiderano meno, sono i più soddisfatti. Il secondo studio dimostra che la soddisfazione del reddito aumenta se la disponibilità economica consente di acquistare beni e servizi che gli individui bramano. Infine il terzo studio trova evidenza del fatto che la soddisfazione passata del reddito è completamente mediata dai suoi desideri.

Le associazioni relative alla soddisfazione del reddito vengono in linea di massima rispecchiate anche in quella finanziaria (si vedano risultati delle analisi di Peiro (2006) e di Van Praag et al. (2003)). La letteratura riguardo questa componente risulta comunque molto residua ed è dunque necessario dedicargli maggior spazio, eventualmente considerando la soddisfazione finanziaria che rappresenta una misura soggettiva più completa.

1.4 Obiettivo dell'analisi

Nelle Sezioni precedenti si è visto come il benessere sia uno dei principali temi di discussione in letteratura; in economia, ad esempio, le misure di questa realtà sono focalizzate principalmente su fattori economici osservabili e oggettivi, come il reddito e la ricchezza. Tuttavia, queste misure monetarie non possono costituire un'accurata proxy del benessere generale di un soggetto. In particolare, qualora l'interesse ricada nella valutazione soggettiva di questa dimensione individuale, la soddisfazione della vita sembra costituire una valida alternativa: questa, tra l'altro, risulta essere influenzata frequentemente e significativamente anche da fattori non economici.

Un'ampia parte della letteratura analizza i dati relativi al benessere soggettivo al fine di individuare le relazioni tra questa misura di autovalutazione e le caratteristiche economiche e non economiche. Eppure economisti, sociologi e psicologici spesso non tengono conto del fatto che la soddisfazione della vita è una dimensione molto ampia che, nel complesso, può essere vista come un concetto aggregato spiegato attraverso i diversi domini che la rappresentano. Tra questi vi è la soddisfazione finanziaria nella quale rientra la soddisfazione del reddito: questa realtà rappresenta infatti la valutazione soggettiva dei fattori economici che, il più delle volte, vengono presi in considerazione come rappresentanti del benessere. Se è vero che in alcune pubblicazioni un reddito più alto contribuisce ad una soddisfazione della vita maggiore, attualmente altrettanto non si può affermare della valutazione soggettiva del reddito. Tale tesi si prefigge proprio questo obiettivo e più precisamente cerca di individuare l'esistenza di una possibile correlazione tra queste due sfere. In secondo luogo, si indagheranno anche i fattori che possono avere un'influenza su que-

ste due. In particolare, poichè un individuo può usare diverse scale di risposta nel momento in cui esprime il giudizio su se stesso (il cosiddetto DIF), senza un'opportuna correzione queste potrebbero risultare distorte; una possibile soluzione è rappresentata dall'idea delle *anchoring vignettes*. Tale problema è stato documentato in precedenti studi, specialmente per la soddisfazione della vita (Angelini et al. (2014)). Per questo motivo, in tale lavoro si è prima di tutto applicata questa tecnica e successivamente si sono analizzate le nuove variabili d'interesse con appropriati modelli, come si vedrà nel Capitolo 3. Nello specifico non ci si focalizzerà quindi sul nesso o relazione causale tra le due sfere del benessere ovvero sulla possibilità che una dei due fenomeni possa contemporaneamente ricoprire il ruolo di variabile esplicativa e variabile risposta. Sebbene la soddisfazione della vita sia rappresentata da un insieme di concetti più che da uno singolo, non si indagano i possibili legami diretti (non mediati) tra questo concetto generale e quello del reddito o quelli indiretti (mediati) su di essa, in quanto prima di studiare questi legami si ritiene importante accettarsi dell'esistenza di un'associazione tra le due soddisfazioni oggetto d'indagine, soprattutto se misurata tenendo in considerazione l'eterogeneità degli individui nell'utilizzo delle scale di risposta per questo tipo di autovalutazioni.

Capitolo 2

Modelli per variabili risposta ordinali

Questo capitolo è dedicato alla presentazione di tre modelli statistici: il primo sfrutta l'informazione delle *anchoring vignettes* al fine di correggere il problema della DIF, il secondo rappresenta un'estensione del modello probit ordinale e infine l'ultimo è un'estensione del modello di Heckman per dati ordinali. Una sezione iniziale è dedicata alla presentazione dell'idea alla base dei modelli che analizzano una variabile dipendente ordinale.

2.1 L'approccio delle *anchoring vignettes*

Un questionario è un ottimo mezzo per ottenere direttamente informazioni sugli individui se si analizza una popolazione con caratteristiche simili, ma quando l'indagine coinvolge popolazioni differenti non è più così immediato e banale confrontare i risultati ed affermare che le soglie utilizzate dagli individui per trasformare la propria opinione in una delle categorie proposte siano omogenee. Persone di diversi Paesi o gruppi possono infatti usare scale non eguali per autovalutarsi perchè interpretano o comprendono stesse domande in maniera differente o semplicemente perchè i soggetti possono differire nell'ottimismo, lunaticità, propensione ad usare categorie estreme, o in altre caratteristiche, anche quando sono simili per condizioni economiche e non economiche o quando denotano stesse preferenze. A partire da Terza (1985), molti studiosi hanno questionato l'adeguatezza di scegliere il modello probit ordinale e hanno proposto delle estensioni del modello, come

l'eterogeneità delle soglie, l'eterogeneità dei parametri, ecc (alcuni esempi si possono consultare in Greene e Hensher (2010b)). Le differenze interpersonali e interculturali nell'interpretare, capire o utilizzare categorie di risposta ordinale per una stessa domanda sono noti con il termine DIF (*Differential Item Functioning*). Conseguentemente, confrontare le autovalutazioni tra Paesi o gruppi socio-economici senza considerare l'eterogeneità individuale potrebbe portare a risultati e interpretazioni sistematicamente distorte.

2.1.1 *Differential Item Functioning*

Il termine DIF ha origine nella letteratura psicometrica: individui con la stessa abilità hanno una differente probabilità di dare una certa risposta alle domande o test (Holland e Wainer, 1993). Per capire meglio cos'è il DIF si presentano due esempi.

Il primo è una versione semplificata del DIF (King et al., 2004), utile per focalizzare l'idea. A due individui appartenenti a due Paesi differentemente sviluppati viene chiesto di descriversi come (1) *anziano*, (2) *di media età*, (3) *giovane adulto*, (4) *bambino*, immaginando così di fare inferenza su questa domanda anziché sulla data di nascita (che ovviamente sarebbe preferibile). L'età è infatti una delle principali variabili presenti in un'indagine e ha il vantaggio che il suo valore percepito è tipicamente indistinguibile dalla vera età (come riportato da King et al. (2004)). La Figura 2.1 considera le interpretazioni di due individui e i differenti valori attribuiti ai valori delle soglie τ_1, τ_2, τ_3 tra le 4 possibili modalità di risposta (ordinate dalla più giovane alla più anziana). La scala d'età a sinistra mostra valori delle soglie più bassi e categorizza "anziano" una persona con più di 40 anni; ciò significa che probabilmente questo soggetto vive in un Paese con un'aspettativa di vita bassa. Dall'altra parte, l'individuo a destra fornisce un'autovalutazione migliore in quanto definisce "anziano" un individuo con più di 80 anni e dunque probabilmente è la descrizione di una persona che vive in un Paese più sviluppato. Perciò, se si fosse stati a conoscenza solo della categoria di risposta scelta, non si sarebbe potuto sapere molto sulla vera età della persona e si sarebbe ottenuta una errata classificazione dell'età dei due Paesi, in quanto non si sarebbero conosciute le differenze delle soglie (King et al., 2004).

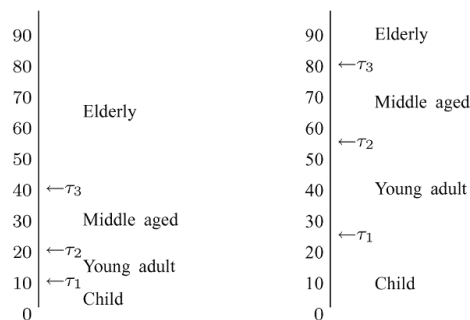


Figura 2.1: Categorizzazione degli anni d'età di due individui. Fonte: King et al. (2004).

In Figura 2.2 viene riportato un secondo esempio, quello di Wand (2013) relativo a due individui (i e j) che interpretano la stessa categoria non ugualmente. Ai due individui è stato chiesto, in un'indagine dell'Organizzazione Mondiale di Sanità del 2002, "*How much say do you have in getting the government to address issues that interest you?*" con le seguenti possibili risposte: (1) *No say at all*, (2) *Little say*, (3) *Some say*, (4) *A lot of say*, (5) *Unlimited say*. Dalla Figura 2.2 emerge l'eterogeneità della scala di risposta dei due soggetti e la differente posizione delle loro soglie τ_{ik} e τ_{jk} (dove $k=1,\dots,4$): dato lo stesso livello latente (inosservato) della domanda di interesse \tilde{y} , il rispondente i dichiara "*Some say*", mentre il risponde j un livello superiore. La stessa situazione si verifica considerando la stessa posizione latente \tilde{y} : infatti, il rispondente i afferma "*Some say*", contrariamente all'intervistato j che sostiene "*Unlimited say*". Più precisamente, il livello "*Some say*" non è lo stesso per i due intervistati: in base all'autovalutazione riportata nel questionario, il primo individuo ha molto meno da dire in merito al governo rispetto al secondo, mentre secondo il vero livello \tilde{y} non vi sono differenze tra i due. L'eterogeneità presente è dunque dovuta alla DIF, che viene espressa attraverso soglie percepite diversamente che separano le categorie di risposta nel giudizio dei soggetti.

È possibile consultare un terzo esempio nel paper di Van Soest et al. (2011), il quale confronta le autovalutazione relativi a problemi con l'alcool.

Dagli esempi emerge dunque come la presenza di eterogeneità individuale

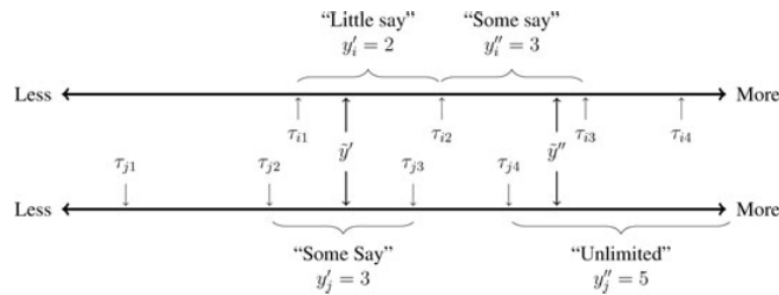


Figura 2.2: Mappaggio dei livelli inosservati continui dell'autovalutazione nelle categorie ordinali riportate dagli intervistati i e j . Fonte: Wand (2013).

possa distorcere e portare a valutazioni fuorvianti; questo problema può nascere quando ai soggetti di Paesi e culture diverse viene domandato di dare la propria opinione a domande soggettive. Nello specifico, viene chiesto di fornire il proprio livello riguardo al costrutto di interesse scegliendo una delle categorie ordinali proposte, ma, nel momento in cui devono dare la propria risposta, esse sono influenzate da fattori esterni come salute, aspettativa di vita, condizioni socio-economiche, passato, ecc. Ciò non consente di avere scale di risposte uguali in quanto non è possibile, utilizzando solo le risposte delle domande soggettive, stabilire quanta parte delle eventuali discrepanze sia dovuta alle differenze oggettive e quanta parte sia dovuta alle soglie (Van Soest et al. (2011), Van Soest e Vonkova (2014)). Ciò causa l'incomparabilità delle loro autovalutazioni. Questa tesi dunque sfrutta un metodo che si focalizza su come superare tale difficoltà.

2.1.2 La tecnica dell'*anchoring vignettes*

Molti ricercatori hanno provato a migliorare il problema dell'incomparabilità interpersonale e cross-culturale ponendo attenzione alle domande e quindi cercando di migliorare le tecniche per porle e crearle, come la loro traduzione, focalizzandosi sul livello di comprensione degli intervistati, ecc (Wand et al., 2011). Tuttavia le risposte riportate nelle domande (soggettive) dell'autovalutazione includono sia il livello reale dell'oggetto di interesse sia la DIF e perciò non è possibile separare le due componenti senza avere a

disposizione ulteriori informazioni. Al fine di poter rendere confrontabili tra individui diversi le autovalutazioni, King et al. (2004) sfruttano lo strumento delle *anchoring vignettes*, sviluppando così un possibile metodo che omogenei le risposte individuali: le *vignettes* hanno una lunga storia nell'analisi di fenomeni sociali ed economici (Alexander e Becker, 1978), ma solo negli anni Duemila grazie a King et al. (2004) vengono introdotte in letteratura come uno strumento addizionale per analizzare variabili ordinali, tenendo conto delle differenze individuali nell'interpretazione delle domande e quindi consentendo di costruire una misurazione del concetto d'interesse libera dalla DIF.

Le *vignettes* sono brevi descrizioni oggettive di individui ipotetici che si trovano in scenari o condizioni ipotetici, pertinenti all'oggetto di interesse (ovvero alla salute, alla soddisfazione della vita, alla disabilità lavorativa,..). Queste domande in cui viene chiesto all'intervistato di valutare certi scenari vengono introdotte nel questionario in aggiunta alla tradizionale domanda di autovalutazione. Se nell'indagine è esposta più di una *anchoring vignette*, ognuna di queste descrizioni presenta un soggetto con differenti caratteristiche o meglio con una diversa situazione associata alla sfera d'interesse, generalmente da una peggiore ad una migliore. Secondo uno studio di Hopkins e King (2010) e contrariamente a King et al. (2004), agli intervistati dovrebbe essere prima chiesto di valutare gli individui ipotetici su una scala che misura un certo fenomeno o atteggiamento in un determinato settore; dopodiché dovrebbero esprimere il proprio giudizio utilizzando la stessa scala con quale hanno valutato le *vignettes*; ciò infatti porta i rispondenti a migliorare la comprensione dell'oggetto di ricerca e la corretta valutazione degli individui ipotetici. Dato che esse descrivono differenti livelli, che sono fissati tra gli intervistati, e che sono scenari di situazioni oggettive uguali per tutti gli intervistati, riguardo a un particolare concetto, le *vignettes* forniscono "un'ancora" (ovvero una scala comune libera dalla DIF) per aggiustare le autovalutazioni; conseguentemente, attraverso questa correzione, si hanno livelli reali uguali e si rendono confrontabili le autovalutazioni fra Paesi o gruppi socio-economici non omogenei. Da qui deriva il termine *anchoring vignettes*.

Un esempio delle domande relative a persone ipotetiche e utilizzate da

King et al. (2004) nell'indagine su *political efficacy*¹ potrebbe essere il seguente:

1. "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative"
2. "[Jane] lacks clean drinking water because the government is pursuing an industrial development plan. In the campaign for an upcoming election, an opposition party has promised to address the issue, but she feels it would be futile to vote for the opposition since the government is certain to win"
3. "[Moses] lacks clean drinking water. He would like to change this, but he can't vote, and feels that no one in the government cares about this issue. So he suffers in silence, hoping something will be done in the future"

Si noti che le tre descrizioni presentano il seguente ordine di scala, dal più libero al meno, e che il nome dei soggetti ipotetici è inserito tra parentesi quadre e questo perchè sarebbe opportuno che i nomi propri siano i corrispondenti tradotti del Paese o del gruppo a cui si sta sottoponendo il questionario, così come per cifre monetarie, grado di istruzione, ecc, in modo da adeguarli alla cultura e al sesso del rispondente. Inoltre, è consigliabile proporre in modo casuale le *vignettes* e non secondo un ordine della scala. La descrizione dovrebbe così essere il più possibile vicino alla realtà e al fenomeno d'interesse, in modo che gli individui possano immedesimarsi efficacemente.

Agli intervistati viene dunque chiesto di valutare questi individui ipotetici, scegliendo una delle k categorie proposte; pertanto, è ragionevole che il numero e il tipo di k modalità sia identico a quello dell'autovalutazione.

¹Per *political efficacy* si intende la fiducia che un cittadino ha verso il proprio governo e la credenza che esso ha nel poter capire ed influenzare le decisioni politiche. In particolare, l'indagine a cui si sta facendo riferimento, è l'Organizzazione Mondiale di Sanità del 2002. Nella Figura 2.2 si erano già valutati due individui della stessa indagine. Le *vignette* che vengono descritte come esempi sono solo 3 delle 6 che sono utilizzate nell'indagine.

Le *vignettes* possono sembrare una tecnica costosa dal momento che sono domande aggiuntive che hanno lo scopo di correggere ogni domanda di autovalutazione. Per ridurre i costi, è possibile sottoporle solo a un piccolo gruppo di soggetti scelti casualmente o ad esempio ogni quarto dei rispondenti oppure includerle in un pretest e non in tutta l'indagine (King et al., 2004; King e Wand, 2007). Nell'utilizzarle, è importante prestare attenzione ad alcuni aspetti: le caratteristiche degli individui ipotetici, il numero di descrizioni ipotetiche da inserire nel questionario e il loro ordine, la possibilità di *inconsistencies* e *tied* (quest'ultimo punto verrà affrontato con maggior dettaglio nella Sezione 2.1.5).

Per quanto riguarda il primo aspetto è già stato detto che gli individui ipotetici dovrebbe avere nomi appropriati; se ciò è troppo costoso una soluzione potrebbe essere quella di usare nomi di genere neutrale (ad esempio Andrea) o riferirsi solo all'iniziale del nome. Allo stesso tempo, le *vignettes* dovrebbero essere scritte così che individui di culture, caratteristiche, Paesi, ..., diversi le interpretino quanto più possibile in modo simile a loro. Ciò è molto difficile e richiede un'accurata attenzione.

Il numero di *vignettes* da chiedere dipende da molti fattori; logicamente potrebbe sembrare che il più sia meglio, ma ciò richiede costi aggiuntivi e risulta spesso non necessario per correggere la DIF. Pertanto la scelta ottimale ricade nel giusto *trade-off* tra riduzione della distorsione e costi dell'indagine, condizionatamente alla natura della DIF e a quanta informazione dispone il ricercatore. Secondo l'esperienza di King et al. (2004) due o tre *vignettes* sono sufficienti, specialmente quando sono rivolte ad una singola cultura o Paese. In generale, è necessaria solo una *vignette* per identificare il modello parametrico e non parametrico, ma è consigliabile appunto usarne più di una per avere più informazione. Altre indicazioni riguardo al numero di queste brevi descrizioni e a quali scegliere nella propria indagine verranno fornite alla Sezione 2.1.5.

La Figura 2.3 rappresenta un esempio della correzione della DIF applicata alle valutazioni di due individui (etichettati con 1, a sinistra, e 2, al centro) sulla *political efficacy* e quindi dell'idea alla base dello strumento che si sta proponendo. In base ai primi due grafici si affermerebbe che l'individuo 1 abbia un livello di valutazione di sé e dei soggetti ipotetici più alto del 2,

seppur entrambi siano d'accordo sulla classificazione ordinale delle *vignettes*. È tuttavia possibile notare che le risposte del primo intervistato variano molto di più rispetto a quelle del secondo (quest'ultime sono molto più concentrate); inoltre, l'autovalutazione del primo si pone tra la valutazione della persona con più bassa voce in capitolo e quella con media, mentre quella del secondo tra la media e la più alta valutazione della descrizione ipotetica. Ciò conferma l'uso di scale di valutazioni differenti. Dal momento che si ipotizza che il livello reale di Alison (o Jane o Moses) sulla *political efficacy* sia lo stesso a prescindere da quale sia l'intervistato, è possibile traslare la scala dell'individuo 2 in una scala confrontabile e uguale a quella del rispondente 1 (grafico a destra della Figura 2.3; si noti infatti che la valutazione dei tre soggetti ipotetici è uguale tra i due partecipanti all'intervista): nel far ciò, emerge una situazione inaspettata, ovvero il secondo individuo ha un livello reale maggiore del primo, persino superiore a quello presentato nel grafico non corretto; in più, Jane viene classificata con una valutazione più alta rispetto al "voto" che il primo intervistato dà a se stesso, ma più bassa rispetto al giudizio che il secondo attribuisce a se stesso.

2.1.3 Le assunzioni

L'approccio delle *anchoring vignettes* si basa su due importanti assunzioni, *response consistency* e *vignette equivalence*.

La prima assunzione afferma che ogni individuo adotta la stessa scala di ri-

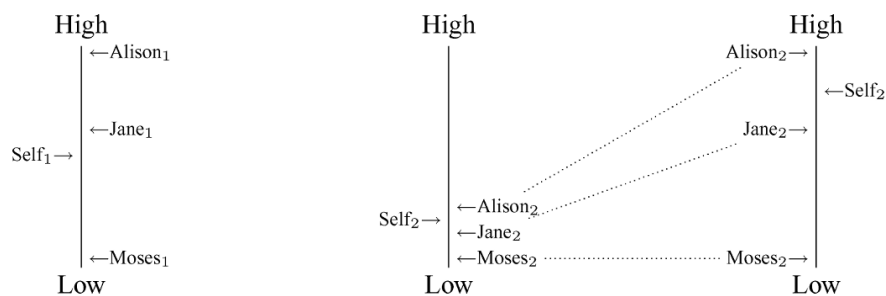


Figura 2.3: Autovalutazioni e valutazione di tre *vignettes* in un contesto di *political efficacy*: confronto tra due individui e correzione della scala di risposta (approccio non parametrico). Fonte: King et al. (2004).

sposta e quindi le stesse soglie sia per l'autovalutazione sia per la valutazione dei soggetti ipotetici. Le risposte potrebbero ovviamente essere influenzate dalla DIF, ma il tipo di DIF è approssimativamente lo stesso tra i due tipi di domanda dell'indagine, per ogni rispondente. In altre parole, il tipo di DIF può variare tra gli individui, ma non fra le risposte sulle autovalutazioni e le *vignettes* che qualsiasi intervistato valuta ad un determinato item. Quindi la DIF di un individuo non viene eliminata, ma attraverso tale assunzione ci si assicura che sia la stessa sia nell'autovalutazione che nelle risposte alle *vignettes*.

La seconda assunzione afferma che le situazioni ipotetiche descritte in ogni *vignette* vengono percepite nello stesso modo e nella stessa scala unidimensionale da parte di tutti gli intervistati, a meno di un errore di misura casuale; conseguentemente si assume che il vero livello della variabile rappresentata da ogni *vignette* venga percepita allo stesso modo da tutti gli intervistati, ad eccezione di un errore casuale. Quindi i rispondenti possono avere delle differenze nella percezione del livello della variabile in ogni scenario, ma queste devono essere casuali e indipendenti dalle caratteristiche della *vignette* stessa (King et al., 2004). Quindi anche nel momento in cui interpretino le domande aggiuntive allo stesso modo, individui diversi possono usare le proprie DIF. Conseguentemente, la *vignette equivalence* consente di ottenere una misura della variabile di interesse libera dalla DIF, poichè la percezione di ogni *vignette* non dipende dalle caratteristiche individuali dei rispondenti.

In Figura 2.4 viene riportato un grafico tratto da una pubblicazione di Bago d'Uva et al. (2011), dove vengono considerate differenze ipotetiche relative alle risposte sulla "gravità" della mobilità fisica tra due stati occupazionali opposti, occupato e non occupato: H_E^* e H_U^* sono rispettivamente le valutazioni del soggetto impiegato e di quello che non ha un lavoro. Il grafico A non è caratterizzato dalla *vignette equivalence* poichè tutte le soglie dell'individuo disoccupato sono più alte di quelle del soggetto che ha un impiego e in particolare $H_U^* > H_E^*$ ovvero l'individuo disoccupato è meno costretto dal problema di mobilità; malgrado ciò, entrambi riportano l'autovalutazione "Moderate". Nel grafico B si considera la valutazione di una *vignette*; sotto l'ipotesi che l'assunzione venga verificata (la situazione ipotetica viene percepita allo stesso modo H_V^*), il disoccupato classifica la situazione dell'in-

dividuo ipotetico Tom come "Severe", la persona occupata come "Moderate". Dal momento che la *vignette* è fissata ed uguale per tutti i rispondenti, la variazione nella classificazione determina un'eterogeneità. L'ultimo grafico si basa sull'assunzione di *response consistency*; conseguentemente, è possibile usare una stessa scala e correggere le risposte rispetto alle soglie, ad esempio, dell'individuo occupato: così facendo, l'autovalutazione del disoccupato è relativa alla categoria "Mild". I tre grafici dunque mostrano che sotto differenti assunzioni, i due intervistati forniscono valutazioni non uguali; dunque l'approccio delle *anchoring vignettes* sfrutta l'idea del grafico B e C per correggere la DIF. Se la *response consistency* non è soddisfatta, le informazioni

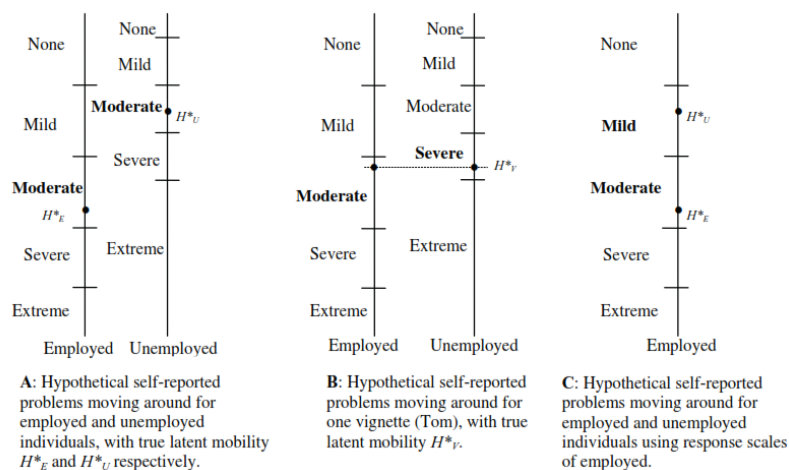


Figura 2.4: Eterogeneità ipotetica dello stato occupazionale nel riportare la mobilità fisica e correzione delle *vignettes*. Fonte: Bago d'Uva et al. (2011).

ottenute dalle risposte alla *vignette* non sono sufficienti a migliorare la comparabilità interpersonale. Questa ipotesi non potrà valere se ci sono influenze sulla propria situazione individuale che sono assenti dalla valutazione della *vignette*. Ad esempio, gli individui non lavoratori possono sperimentare una pressione sociale e/o incentivi finanziari per sottovalutare la propria salute, che però non sono goduti dagli individui ipotetici. In questo caso quindi l'approccio non correggerebbe pienamente la distorsione che ha afflitto le stime dell'impatto della salute sulla partecipazione del mercato del lavoro (Bago d'Uva et al., 2011). Se invece la *vignette equivalence* non è soddisfatta, allora

gli intervistati non stanno rispondendo diversamente in base allo stesso stato, ma stanno rispondendo in base a diversi stati percepiti. Ciò può accadere ad esempio se le descrizioni dello scenario ipotetico sono incomplete (Bago d'Uva et al., 2011).

Negli ultimi anni molti autori si sono dedicati alla ricerca di metodi per testare o rilassare queste due assunzioni, non concordando tuttavia su evidenze univoche. Alcune soluzioni saranno presentate nella Sezione 2.1.6.

2.1.4 I modelli

Attualmente in letteratura vi sono due approcci statistici che sfruttano la tecnica delle *anchoring vignettes*: una soluzione parametrica e una non parametrica. Queste vengono presentate nelle Sezioni 2.1.4.1 e 2.1.4.3. Inoltre si sono sviluppate delle estensioni, essenzialmente per il primo approccio.

2.1.4.1 Il modello parametrico Chopit

Il modello parametrico che utilizza i dati raccolti per mezzo delle *anchoring vignettes* è il *Compound Hierarchical Ordered Probit* (Chopit) (in letteratura si trova anche sotto il nome di Hopit). Questo modello può essere pensato come una generalizzazione del modello Probit ordinato, dove la DIF viene modellata mediante la variazione delle soglie. Entrambi i modelli usano una variabile latente, ma il Chopit consente alle soglie di cambiare tra i rispondenti e ciò avviene in funzione di un insieme di variabili esplicative.

Il modello è costituito da due componenti: una equazione per ognuna delle autovalutazioni e una equazione per ognuna delle risposte alle *vignettes*. In particolare, per ogni domanda e per ogni rispondente si misurano tre livelli della variabile, come si può osservare in Figura 2.5: il reale, il percepito e il riportato (King et al., 2004). Il livello reale è il livello effettivo, attuale della variabile oggetto di studio (libertà, soddisfazione,..) che ogni individuo ha; esso è inosservabile, ma confrontabile tra gli intervistati e viene misurato in una scala continua e unidimensionale. Ovviamente, a causa della sua natura è lecito ritenere che dipenda da caratteristiche individuali. Inoltre viene percepito dagli intervistati correttamente in media, ma con un errore; dunque anche il livello percepito è inosservabile ed è valutato su una scala continua,

eppure è confrontabile solo in media a causa dell'errore casuale. La relazione tra questo livello e quello riportato (osservabile) è la seguente: quando gli individui rispondono ad una domanda, esprimono il loro livello percepito scegliendo una delle (k) categorie ordinali della domanda; tuttavia ogni soggetto usa soglie differenti. Questi due livelli influenzano quello riportato, ovvero la risposta che l'individuo fornisce, che per natura è osservabile, ma incomparabile a causa della DIF (è possibile modellare quest'ultima attraverso la variazione delle soglie).

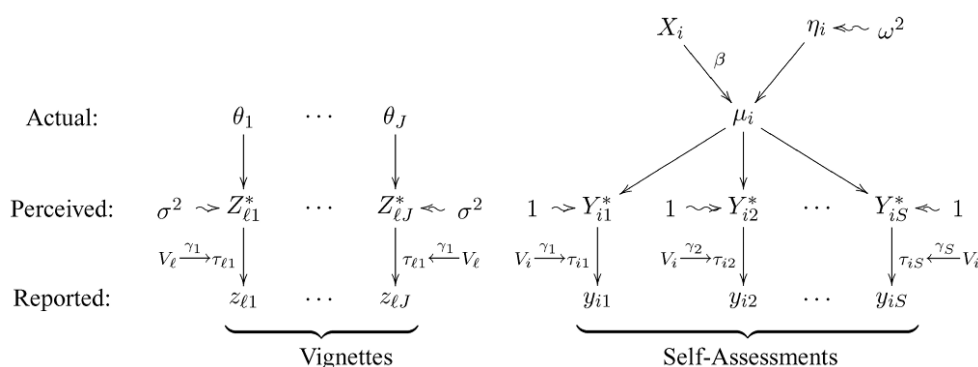


Figura 2.5: Rappresentazione grafica dell'idea alla base del modello parametrico.
Fonte: King et al. (2004).

Componente dell'autovalutazione

Traducendo matematicamente quanto detto a riguardo alla componente dell'autovalutazione ogni intervistato i ha un livello reale μ_i della variabile d'interesse che si assume essere una combinazione lineare di p covariate osservate X_i , le quali rappresentano appunto le caratteristiche individuali. Questo livello viene percepito dagli intervistati con un errore ε_i , come si è detto precedentemente, di distribuzione normale; pertanto $\mu_i = \mu(X_i)$ si distribuisce come un probit ordinato. Y_{is}^* definisce il livello percepito inosservato dell'individuo i all'autovalutazione s e continuo che viene modellato come una normale con media e varianza opportune. Nello specifico si ha:

$$Y_{is}^* = \mu_i + \varepsilon_i = X_i \beta + \varepsilon_i$$

con $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ e $\varepsilon_i \perp X_i$

ove β è un vettore di parametri e $s = 1, \dots, S$ e $i = 1, \dots, N$. Il termine d'errore ε_i può essere dovuto a eterogeneità e/o a un errore nell'aver riporta-

to una risposta (Van Soest e Vonkova, 2012; Paccagnella, 2013); la scelta di porre la media di tale errore pari a 0 è necessaria per poter garantire l'identificazione del modello (problema che verrà affrontato con maggior dettaglio a pagina 42).

Quando un soggetto i deve valutare un'autovalutazione s scegliendo una delle K_s categorie ordinali, non potendo rispondere con una scala continua, egli trasforma il livello percepito continuo Y_{is}^* nella categoria riportata Y_{is} attraverso il seguente meccanismo:

$$Y_{is} = k \quad \text{se} \quad \tau_{is}^{k-1} \leq Y_{is}^* < \tau_{is}^k$$

dove $-\infty = \tau_{is}^0 < \tau_{is}^1 < \dots < \tau_{is}^{K_s-1} < \tau_{is}^{K_s} = +\infty$

con τ_{is} vettore delle soglie e $k = 1, \dots, K_s$. Ciò vuol dire che l'individuo i sceglie la categoria k se e solo se il valore della variabile latente è compreso tra τ_{is}^{k-1} e τ_{is}^k . Un modo naturale per modellare l'eterogeneità riportata consiste nel permettere alle soglie τ_{is} di dipendere da delle caratteristiche osservate V_i (che possono coincidere oppure no con le X_i) adottando, ad esempio, una specificazione lineare (versione originaria di King et al. (2004)), ovvero:

$$\begin{aligned} \tau_s^1 &= \gamma_s^1 V_i \\ \tau_s^k &= \tau_s^{k-1} + e^{\gamma_s^k V_i} \end{aligned}$$

con $k = 2, \dots, K_s - 1$ e γ_s un vettore di parametri ignoti con elementi γ_s^k . L'assunzione esponenziale della seconda equazione è solo una convenienza per assicurare che le soglie siano (monotome) crescenti con k , così da mantenere l'ordine logico della scala.

Dato il modello appena specificato, la distribuzione di probabilità del

livello riportato Y per l'individuo i (con $\sigma_\varepsilon^2=1$) è dunque la seguente:

$$\begin{aligned}
P(Y_i = 1|x_i) &= P(y_i^* \leq \tau_s^1|x_i) = P(\varepsilon_i \leq \tau_s^1 - x_i\beta) = \Phi(\gamma_s^1 V_i - x_i\beta) \\
P(Y_i = 2|x_i) &= P(\tau_s^1 < y_i^* \leq \tau_s^2|x_i) = \\
&= P(\varepsilon_i \leq \tau_s^2 - x_i\beta|x_i) - P(\varepsilon_i \leq \tau_s^1 - x_i\beta|x_i) = \\
&= \Phi(\gamma_s^1 V_i + e^{\gamma_s^2 V_i} - x_i\beta) - \Phi(\gamma_s^1 V_i - x_i\beta) \\
&\dots \\
P(Y_i = k|x_i) &= P(\tau_s^{k-1} < y_i^* \leq \tau_s^k|x_i) = \\
&= P(\varepsilon_i \leq \tau_s^k - x_i\beta|x_i) - P(\varepsilon_i \leq \tau_s^{k-1} - x_i\beta|x_i) = \\
&= \Phi(\gamma_s^1 V_i + \sum_{q=2}^k e^{\gamma_s^q V_i} - x_i\beta) - \Phi(\gamma_s^1 V_i + \sum_{1 < t < k} e^{\gamma_s^t V_i} - x_i\beta) \\
&\dots \\
P(Y_i = K_s|x_i) &= P(\tau_s^{K_s-1} < y_i^* \leq \tau_s^{K_s}|x_i) = P(\varepsilon_i \leq \tau_s^{K_s} - x_i\beta|x_i) \\
&= \Phi(\gamma_s^1 V_i + \sum_{1 < q < K_s} e^{\gamma_s^q V_i} - x_i\beta)
\end{aligned}$$

dove $\Phi(\Delta)$ è la funzione di ripartizione di una variabile normale standard.

Componente delle *vignettes*

Per quanto riguarda la componente delle *anchoring vignettes*, la logica sottostante è la stessa di quella della componente di autovalutazione, come si può vedere anche dalla Figura 2.5. Ovviamente si avranno tante equazioni quante sono le descrizioni ipotetiche per ogni autovalutazione.

Ogni individuo ipotetico i ha un livello reale θ_j (con $j = 1, \dots, J$) che viene misurato su una scala continua, come per il vero livello dell'intervistato. Perciò, il rispondente i percepisce θ_j correttamente in media, ma con un termine d'errore casuale, u_{ij} ; ne consegue:

$$\begin{aligned}
Z_{isj}^* &= \theta_j + u_{isj} \\
\text{con } u_{isj} &\sim N(0, \sigma_{us}^2) \\
u_{isj} &\perp (\varepsilon_i, X_i, V_i)
\end{aligned}$$

dove Z_{isj}^* rappresenta il livello percepito inosservato dell'individuo i alla j -esima *vignette* corrispondente alla s -esima autovalutazione. La varianza

σ_{us}^2 viene assunta uguale non solo tra gli individui ipotetici, ma anche tra gli intervistati; tuttavia è possibile che essa vari tra le *vignettes*, dal momento che le sue stime sono indicatori convenzionali di un aspetto relativo alla bontà di apprendimento corretto di ogni descrizione ipotetica (King et al., 2004). Così come è stato descritto per la componente dell'autovalutazione, l'individuo trasforma la scala continua e inosservata di Z_{isj}^* in una scala discreta scegliendo una delle k modalità (uguali a quelle delle autovalutazioni) attraverso il seguente meccanismo:

$$Z_{isj} = k \quad \text{se} \quad \tau_{isj}^{k-1} \leq Z_{isj}^* < \tau_{isj}^k$$

dove $-\infty = \tau_{isj}^0 < \tau_{isj}^1 < \dots < \tau_{isj}^{K_s-1} < \tau_{isj}^{K_s} = +\infty$

con $k = 2, \dots, K_s - 1$. Le soglie τ_{isj} sono dunque specificate come per la componente di autovalutazione s ; esse son dunque determinate dal vettore di parametri γ uguale a $\gamma_s = (\gamma_s^1, \dots, \gamma_s^{K_s-1})$ e dalle stesse variabili esplicative V_i . Dunque, dal modello emergono chiaramente le due assunzioni su cui si basa il modello:

- la *vignette equivalence* è formalizzata attraverso θ_j che è assunta la stessa per tutti i rispondenti (ovvero non dipende da i); inoltre $f(Z_j^*|X) = f(Z_j^*)$. Quindi il vero livello di ogni *vignette* viene percepito allo stesso modo da tutti gli intervistati;
- la *response consistency* è formalizzata attraverso il vettore di parametri γ che è identico a γ_s ; quindi ogni individuo adotta la stessa scala di risposta (ovvero le stesse soglie) sia per l'autovalutazione sia per la valutazione dei soggetti ipotetici.

Se si volesse ottenere la distribuzione di probabilità della componente riportata sulla descrizione ipotetica condizionatamente a delle caratteristiche V_i per l'autovalutazione s , questa non differirebbe da quella della componente relativa all'autovalutazione; nello specifico:

$$P(Z_i = 1|V_i) = P(z_i^* \leq \tau_s^1|V_i) = P(u_{ij} \leq \tau_s^1 - \theta_j) = \Phi \left(\frac{\gamma_s^1 V_i - \theta_j}{\sigma_u^2} \right)$$

$$\begin{aligned}
P(Z_i = 2|V_i) &= P(\tau_s^1 < z_i^* \leq \tau_s^2|V_i) = \\
&= P(u_{ij} \leq \tau_s^2 - \theta_j|V_i) - P(u_{ij} \leq \tau_s^1 - \theta_j|V_i) = \\
&= \Phi\left(\frac{\gamma_s^1 V_i + e^{\gamma_s^2 V_i} - \theta_j}{\sigma_u^2}\right) - \Phi\left(\frac{\gamma_s^1 V_i - \theta_j}{\sigma_u^2}\right) \\
&\dots \\
P(Z_i = k|V_i) &= P(\tau_s^{k-1} < z_i^* \leq \tau_s^k|V_i) = \\
&= P(u_{ij} \leq \tau_s^k - \theta_j|V_i) - P(u_{ij} \leq \tau_s^{k-1} - \theta_j|V_i) = \\
&= \Phi\left(\frac{\gamma_s^1 V_i + \sum_{q=2}^k e^{\gamma_s^q V_i} - \theta_j}{\sigma_u^2}\right) - \Phi\left(\frac{\gamma_s^1 V_i + \sum_{1 < t < k-1} e^{\gamma_s^t V_i} - \theta_j}{\sigma_u^2}\right) \\
&\dots \\
P(Z_i = K_s|V_i) &= P(\tau_s^{K_s-1} < z_i^* \leq \tau_s^{K_s}|V_i) = P(u_{ij} \leq \tau_s^{K_s} - \theta_j|x_i) \\
&= \Phi\left(\frac{\gamma_s^1 V_i + \sum_{1 < q < K_s-1} e^{\gamma_s^q V_i} - \theta_j}{\sigma_u^2}\right)
\end{aligned}$$

dove $\Phi(\Delta)$ è la funzione di ripartizione di una variabile normale standard.

Identificazione

Senza ulteriori restrizioni, il modello così specificato non risulta identificato; a tale scopo la varianza del termine d'errore del livello percepito, ovvero σ_ε^2 , viene assunta nota e uguale a 1 per ogni autovalutazione s , mentre l'intercetta dell'equazione dell'autovalutazione viene fissata, per comodità, a 0 (da qui deriva che la media dell'errore viene fissata a 0). La scelta di fissare queste soglie nasce con l'idea di poter confrontare il modello Chovit con il Probit ordinato². Un'altra possibile scelta per garantire l'identificazione consiste nel fissare uno dei θ_j e un altro dei θ_j o uno dei β (solitamente $\theta_1 = 0$ e $\theta_{J_s} = 1$ così che la stima di μ possa essere interpretata sulla stessa scala delle *vignettes*).

²Nel modello Probit ordinato si assume solitamente che la varianza del termine d'errore sia pari a 1 e si assegna un valore ad una soglia (per esempio la prima uguale a 0).

Stima del modello

La specificazione che incorpora lo strumento delle *anchoring vignettes* ingloba due modelli di scelta ordinale, la componente di autovalutazione e la componente delle *vignettes*. Conseguentemente la stima consiste nel modellare congiuntamente le risposte delle due componenti mediante un metodo di modellazione probit ordinato. Il metodo standard per stimare i parametri del modello Chopit è quello di massima verosimiglianza. La componente di autovalutazione e quella delle *vignettes* sono caratterizzate dalla propria funzione di verosimiglianza e, essendo indipendenti tra loro ($\varepsilon_i \perp u_{ij}$), è possibile combinarle per ottenere la funzione di verosimiglianza congiunta. Nello specifico la prima è data da:

$$L_s(\beta, \gamma|Y) = \prod_{i=1}^N \prod_{k=1}^{K_s} P(Y_{is} = k) \propto \prod_{i=1}^N \prod_{k=1}^{K_s} [\Phi(\tau_{is}^k | X_i \beta, 1) - \Phi(\tau_{is}^{k-1} | X_i \beta, 1)]^{I(y_i=k)}$$

dove $i = (1, \dots, N)$ è l'indice individuale, $k = (1, \dots, K_s)$ l'indicatore delle categorie della risposta, $\Phi(\Delta) = F(\cdot)$ è la funzione di ripartizione di una variabile normale standard e $I(y_i = k) = \begin{cases} 1 & \text{se } y_i = k \\ 0 & \text{altrimenti} \end{cases}$ è la funzione indicatrice.

La log-verosimiglianza per la componente delle *vignettes* è pari a:

$$L_v(\theta, \gamma|Z) = \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^{K_s} P(Z_{isj} = k) = \\ \propto \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^{K_s} [\Phi(\tau_{is}^k | \theta_j, \sigma_u^2) - \Phi(\tau_{is}^{k-1} | \theta_j, \sigma_u^2)]^{I(z_{isj}=k)}$$

dove i e k coincidono con gli indicatori precedenti, $j = (1, \dots, J)$ è l'indice delle *vignettes* e $I(z_{isj} = k) = \begin{cases} 1 & \text{se } z_{isj} = k \\ 0 & \text{altrimenti} \end{cases}$ è la funzione indicatrice.

Dunque la funzione congiunta è data da:

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

e il vettore delle stime sarà ottenuto dalla massimizzazione della log-verosimiglianza

$$\max_{(\beta, \theta, \gamma, \sigma_u^2)} l(\beta, \theta, \gamma, \sigma_u^2 | Y, Z)$$

Perciò, l'interpretazione della stima dei parametri del vettore β è simile a quella del modello Probit ordinato con la differenza che le soglie stimate τ_i^k sono specifiche per ogni rispondente, anziché esser costanti. Nello specifico, se l'insieme di caratteristiche osservate X_i è coincidente con quello di V_i , allora queste, oltre a poter influenzare il livello riportato, potrebbero aver un'influenza significativa anche negli spostamenti delle soglie tra le diverse k modalità di risposta.

Si tenga comunque presente che la verosimiglianza del modello Chopit potrebbe non risultare semplice nel caso in cui l'ipotesi di normalità e l'assunzione i.i.d (indipendentemente e identicamente distribuite) dei termini d'errore siano incerte. In questi casi quindi sarebbe più ottimale una procedura semi-parametrica o una soluzione parametrica più flessibile, che tuttora non sono state ancora sviluppate.

2.1.4.2 Estensioni del modello parametrico

Si è già detto che il modello Chopit può essere considerato come una generalizzazione del modello Probit ordinato, in quanto ciascuna risposta viene modellata secondo tale approccio: con lo scopo di confrontare i due tipi di approcci, una scelta comune tra gli autori consiste nel fissare l'insieme delle caratteristiche osservate V_i pari a quello di X_i , oltre a fissare i parametri $\sigma_\varepsilon^2 = 1$ e $\mu_\varepsilon = 0$ (pagina 42). Tuttavia il modello Chopit può essere visto anche come una generalizzazione di un modello gerarchico con intercetta casuale, in cui la varianza di secondo livello viene fissata a 1; in questa visione quindi la domanda rappresenta il primo livello, mentre gli individui il secondo livello. L'obiettivo dunque è quello di confrontare gli intervistati non solo in base alle descrizioni ipotetiche, ma anche secondo le autovalutazioni. Nello specifico, supponendo che di avere $s = 1$ autovalutazioni, il modello assumerebbe la seguente forma:

$$\begin{aligned}
Y_{ij}^* &= X_{ij}\beta + \varepsilon_i + u_{ij} \\
\varepsilon_i &\sim N(0, 1) \\
u_{ij} &\sim N(0, \omega^2)
\end{aligned}$$

Diversi ricercatori hanno esteso l'approccio parametrico delle *vignettes* con lo scopo di adattarlo anche ad altri aspetti. Particolare attenzione è stata dedicata alla specificazione delle soglie più che alla varianza condizionata: infatti ci si può aspettare che una caratteristica persistente microeconomica dei dati a livello individuale sia l'eteroschedasticità della varianza (Greene e Hensher, 2010b). Tuttavia ancora non vi sono studi riguardo ad un possibile modello con l'uso delle *anchoring vignettes* che intraprenda questa strada, così come la possibilità di considerare la non-normalità degli errori ε_i e u_{ij} o una specificazione semi-parametrica: si tratterebbe dunque di adattare al contesto di riferimento e dunque generalizzare modelli probit ordinati che già comprendono queste possibili strade come il modello probit ordinato eteroschedastico di Chen e Khan³, la proposta di inserire l'eterogeneità secondo l'idea di Heckman e Singer⁴ o l'approccio semi-non parametrico di Gallant e Nychka⁵ (l'idea e la formulazione dei tre modelli viene riportata rispettivamente da Greene e Hensher (2010b, Cap. 7.4), da Hadri e Mikhail (2014, Cap. 6.5) e da Greene e Hensher (2010b, Cap. 12)). Di seguito vengono riportate le estensioni presenti in letteratura nel caso; si precisa che la formulazione è riferita al caso in cui vi sia una sola autovalutazione ($s = 1$).

Van Soest et al. (2011) inseriscono un termine d'errore ξ_i nella prima equazione delle soglie (e quindi è incluso anche nelle altre soglie), soluzione che permette l'introduzione di una qualche caratteristica inosservata dell'intervistato i . Ciò significa che oltre alle covariate osservate V_i , anche l'eterogeneità individuale (casuale) nelle soglie ξ_i può influenzare le soglie. Con questa componente aggiuntiva, le soglie del modello Chopit si modificano nel seguente modo:

³La specificazione della variabile latente è: $y_i^* = X_i\beta + \sigma_i^2\varepsilon_i$ dove $\varepsilon_i \sim N(0, 1)$.

⁴Assumono che la distribuzione di ε_i sia discreta ovvero $Pr(\varepsilon_i = \zeta_m) = \pi_m$ con $m = 1, 2, \dots, M$, $\zeta = (\zeta_1, \dots, \zeta_M)$ è un vettore di M punti di massa.

⁵La specificazione della distribuzione del termine d'errore ε_i proposta dai due autori è

$$f_k(\varepsilon) = \frac{1}{\theta} \left(\sum_{k=0}^{K_s} \gamma_k \varepsilon^k \right)^2 \Phi(\varepsilon)$$

$$\begin{aligned}\tau_i^1 &= \gamma^1 V_i + \xi_i \\ \tau_i^k &= \tau_i^{K-1} + e^{\gamma^k V_i}, \quad k = 2, \dots, K \\ \xi_i &\sim N(0, \sigma_\xi^2) \\ \xi_i &\perp (\varepsilon_i, u_{ij}, X_i, V_i)\end{aligned}$$

Van Soest e Vonkova (2012) dimostrano che questa componente aggiuntiva è utile per ridurre sostanzialmente problemi di errata specificazione, grazie all'inclusione di un solo parametro al modello originario di King et al. (2004). I due modelli sono comunque annidati dal momento che imponendo $\sigma_\xi^2 = 0$ si ottiene il modello standard.

In alcune analisi preliminari, Kapteyn et al. (2007) trovano evidenza del fatto che gli individui potrebbero rispondere diversamente a seconda del genere dei soggetti ipotetici descritti; se ciò si verifica, il livello percepito associato alla *vignette* sarebbe distorto determinando l'uso di soglie differenti nella valutazione di uno scenario con soggetto femminile piuttosto che maschile. Conseguentemente, ciò violerebbe l'assunzione di *vignette equivalence* perché gli individui percepirebbero in modo diverso gli scenari ipotetici. Date queste possibili evidenze, Kapteyn et al. (2007) includono nella componente delle *vignettes* una dummy sul genere della persona ipotetica:

$$Z_{ij}^* = \theta_j + \nu F_{ij} + u_{ij}$$

$$\text{dove } F_{ij} = \begin{cases} 1 & \text{individuo ipotetico : donna} \\ 0 & \text{altrimenti} \end{cases}.$$

Tuttavia tale estensione può essere usata solo quando le descrizioni ipotetiche sono somministrate casualmente agli intervistati. Analizzando il comportamento nell'uso di sostanze alcoliche di alcuni studenti, Van Soest et al. (2011) includono in questa formulazione un'ulteriore termine, specifico per le proprie analisi, ovvero una variabile dummy indicante se all'individuo i è stato mostrato il totale di grammi di alcool che ha una bevanda e se sono stati forniti degli esempi relativi.

Paccagnella (2011) estende il modello standard per tener conto della distorsione di selezione. La sua specificazione è utile quando il questionario è somministrato ad un campione, ma solo alcuni rispondenti sono selezionati (non casualmente) per rispondere alle *vignettes*. Nei confronti di misure di auto-

valutazione tra Paesi questo problema potrebbe portare a risultati inconsistenti (specialmente se il numero di risposte alle *vignettes* è basso) a causa, per esempio, di mancanza di procedure sul campo di lavoro standardizzate (Paccagnella, 2013). Il modello prevede dunque una terza componente alle due già introdotte nel Chopit, detta equazione di selezione:

$$S_i^* = W_i\delta + \eta_i$$

$$S_i = \begin{cases} 1 & \text{se } S_i^* > 0 \\ 0 & \text{altrimenti} \end{cases}$$

dove W_i è un insieme di variabili esplicative (W_i può coincidere anche con X_i , ma è consigliabile che almeno una variabile sia esclusiva di W_i stesso), δ è un vettore di parametri che deve essere stimato e η_i è il termine d'errore. Pertanto, i livelli riportati Y_i e Z_{ij} sono osservati solo se la condizione $S_i^* > 0$ è soddisfatta. Al fine di introdurre una correlazione tra i termini di errore si sfrutta l'approccio di un effetto *shared random* w_i :

$$\varepsilon_i = \phi w_i + \psi_i$$

$$u_{ij} = \lambda_j w_i + \varsigma_{ij}$$

$$\eta_{ij} = w_i + \nu_i$$

dove w_i , ψ_i e ν_i sono distribuiti indipendentemente come una normale standard e ς_{ij} indipendentemente secondo una normale di media zero e varianza σ_ν^2 .

Angelini et al. (2011) introducono il modello Chopit longitudinale, al fine di analizzare fino a quale punto le soglie siano stabili nel tempo. Infatti uno dei problemi che possono insorgere nell'analizzare le scale di risposta è che queste possano differire non solo tra gli individui, ma anche nel tempo, determinando appunto l'incomparabilità delle autovalutazioni. In presenza di dati panel, l'eterogeneità invariante nel tempo può essere considerata attraverso l'uso di modelli ad effetti fissi o causali (Angelini et al., 2011). L'idea alla base di tale modello è dunque quello di combinare le tecniche dei modelli panel con quelle delle *anchoring vignettes*.

Formalmente, dati i periodi temporali $t = 1, 2, \dots, T$ e Y_{it}^* il livello percepito dell'individuo i al tempo t ad una domanda, con $i = 1, 2, \dots, n$ viene modellato come segue:

$$Y_{it}^* = X_{it}\beta_t + \varepsilon_{it}$$

$$\varepsilon_{it} = \eta_i + \omega_{it}$$

ove X_{it} sono variabili esogene varianti nel tempo, β_t è il vettore di parametri che deve essere stimato (senza costante per l'identificazione) e ε_{it} è il termine d'errore che è costituito da una componente specifica individuale η_i e da un errore idiosincratico ω_{it} . Nello specifico, il termine d'errore nell'equazione per l'autovalutazione è definito come:

$$\eta_i \sim N(0, \sigma^2)$$

$$\omega_{it} \sim N(0, 1)$$

$$\eta_i \perp \omega_{it}$$

$$\omega_{it} \perp \omega_{iu}, \quad t \neq u$$

da cui deriva che $V(\varepsilon_{it}) = \sigma^2 + 1$ e $COV(\varepsilon_{it}, \varepsilon_{iu}) = \sigma^2$ per $t \neq u$.

Come nel modello Chopit standard di King et al. (2004) e con la stessa logica, il rispondente i al tempo t esprime il livello percepito Y_{it}^* (inosservato e continuo) nella valutazione discreta Y_{it} scegliendo tra k categorie ordinate ($k = 1, 2, \dots, K$) secondo il seguente meccanismo:

$$Y_{it} = k \quad \text{se} \quad \tau_{it}^{k-1} \leq Y_{it}^* < \tau_{it}^k$$

dove $-\infty = \tau_{it}^0 < \tau_{it}^1 < \dots < \tau_{it}^{k-1} < \tau_{it}^k = +\infty$

Le soglie sono sempre individuali e modellate come una funzione di variabili esogene V_{it} e un vettore di parametri γ_t , ma in più sono tempo-dipendenti:

$$\tau_{it}^1 = \gamma_t^1 V_{it}$$

$$\tau_{it}^k = \tau_{it}^{k-1} + e^{\gamma_t^k V_{it}}$$

con $k = 2, 3, \dots, K - 1$.

Per quanto riguarda la componente delle *anchoring vignettes*, il livello inosservato dell'oggetto d'interesse che è descritto nello scenario j come percepito dall'individuo i al tempo t è modellato come:

$$Z_{ijt}^* = \theta_{jt} + u_{ijt}$$

$$u_{ijt} = \varsigma_{ij} + \nu_{ijt}$$

Quindi ogni intervistato i non percepisce la *vignette* j nello stesso modo in istanti temporali differenti (θ_{jt} costante per i) dove $i = 1, 2, \dots, n$, $t = 1, 2, \dots, T$ e $j = 1, 2, \dots, J$. Così come è stato fatto per la precedente componente, il termine d'errore u_{ijt} è diviso in due termini, una componente specifica individuale e della *vignette* (ς_{ij}), e da un errore idiosincratico (ν_{ijt}). Nello specifico:

$$\begin{aligned}\varsigma_{ij} &\sim N(0, \rho^2) \\ \nu_{ijt} &\sim N(0, \tau_t^2) \\ \varsigma_{ij} &\perp (\nu_{ijt}, \eta_i, \omega_{it}) \\ \nu_{ijt} &\perp \nu_{iju}, \quad t \neq u \\ \nu_{ijt} &\perp (\nu_{ijt}, \omega_{it})\end{aligned}$$

da cui deriva che $V(u_{ijt}) = \rho^2 + \tau_t^2$ e $COV(u_{ijt}, u_{iju}) = \rho^2$ per $t \neq u$.

La valutazione discreta Z_{it} del rispondente i al tempo t è dovuto alla trasformazione del livello percepito Z_{it}^* (inosservato e continuo) scegliendo tra k categorie ordinate ($k = 1, 2, \dots, K$) nel seguente modo:

$$\begin{aligned}Z_{ijt} = k &\quad \text{se} \quad \tau_{it}^{k-1} \leq Z_{ijt}^* < \tau_{it}^k \\ \text{dove} \quad -\infty &= \tau_{it}^0 < \tau_{it}^1 < \dots < \tau_{it}^{k-1} < \tau_{it}^k = +\infty\end{aligned}$$

e in cui le soglie τ_{it}^k sono le stesse dell'equazione dell'autovalutazione, in modo da rispettare l'assunzione di *response consistency*. Per quanto riguarda la *vignette equivalence*, la disponibilità di dati longitudinali permette di rilassare questa assunzione (secondo la formulazione standard essa implicherebbe che $u_{ijt} = \nu_{ijt}$). Ovviamente il Chopit longitudinale può essere combinato con una delle precedenti estensioni (tenendo conto di vari aspetti allo stesso tempo): ad es Angelini et al. (2011) includono nella loro specificazione della *vignette* una variabile dummy indicante il genere dell'individuo ipotetico secondo la formulazione di Kapteyn et al. (2007), dipendente da t oltre che da i e j .

2.1.4.3 Il metodo non parametrico

Diversamente dal modello parametrico, l'approccio non parametrico proposto da King et al. (2004) minimizza le assunzioni statistiche e costituisce quindi un complemento molto utile per la modellazione parametrica. Come

nella soluzione parametrica, si assumono la *response consistency* e la *vignette equivalence*, ma non viene richiesta nessun'altra assunzione aggiuntiva.

L'idea alla base di tale approccio consiste nel ricodificare le autovalutazioni in riferimento all'insieme delle *anchoring vignettes*. In base ad una scala fissata dalla valutazione degli scenari associati agli individui ipotetici in ogni Paese o gruppo socio-economico, si costruisce una nuova variabile libera dalla DIF che può essere analizzata come un'altra qualsiasi variabile ordinale, cioè con un istogramma, delle variabili contingenti o un probit ordinato, ...: quindi è possibile fare confronti diretti tra gli individui oppure modellarla attraverso un metodo parametrico o semi-parametrico.

Per ogni rispondente i con $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, J$ sia Y_i la domanda oggetto d'indagine e sia Z_{ij} la valutazione dello scenario ipotetico j . Come già è detto nella Sezione 2.1.2, le situazioni descritte nelle *vignettes* sono generalmente caratterizzate da un ordine naturale dato dalla severità del problema di interesse descritto (o scelto dal ricercatore); conseguentemente tutti gli intervistati (o la maggioranza) dovrebbero percepire lo stesso ordinamento ovvero $Z_{i,j-1} < Z_{i,j} \forall i, j$. Per questi soggetti è dunque possibile creare una nuova variabile libera da DIF che riscali l'autovalutazione a seconda della loro valutazione alla situazione degli individui ipotetici, chiamata scala C (King et al., 2004):

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

La scala C è funzione delle soglie e dunque la specificazione della precedente variabile è equivalente alla seguente (Wand, 2013):

$$C_i = \begin{cases} 2j & \text{se } \tau_i^{k-1} < Y_i^* \text{ e } Z_{ij}^* < \tau_i^k \\ 2j + 1 & \text{se } Z_{ij}^* < \tau_i^{k'} < Y_i^* < \tau_i^k < Z_{i,j+1}^*, \quad k' < k \end{cases}$$

dove $\tau_i = (\tau_i^1, \tau_i^2, \dots, \tau_i^K)$ è il vettore delle soglie dell'intervistato i -esimo, $k = 1, 2, \dots, K$ indica la modalità optata per la domanda e Y_i^*, Z_{ij}^* sono le

variabili latenti indicanti il livello percepito rispettivamente per l'autovalutazione e per la j -esima *vignette*, come per il metodo parametrico (Sezione 2.1.4.1).

Ovviamente, avere più di una *vignette* potrebbe causare risposte inconsistenti o non così tanto informative quanto si potrebbe pensare (Hopkins e King, 2010) poichè ci possono essere dei soggetti che non capiscono e dunque non rispettano l'ordinamento naturale delle *vignettes*, oppure che percepiscono più scenari ipotetici con lo stesso livello; ciò significa che la variabile risposta associata a queste domande potrebbe indicare un range di valori ordinali più che un singolo valore. Queste divergenze non sono in grado di correggere la DIF e pertanto devono essere raggruppate e trattate come *ties* o inconsistenze; tuttavia, senza ulteriori informazioni l'analista non può distinguere e correggere tali "errori". Perciò la perdita di questa informazione porta ad una inefficienza di tale approccio e dunque ne costituisce un suo punto debole. In Tabella 2.1 viene presentata una lista completa di tutti le classificazioni possibili nel caso di 3 *vignettes* e una autovalutazione. Si supponga che l'ordinamento naturale degli scenari ipotetici sia $Z_1 < Z_2 < Z_3$; in questo caso dunque la variabile C libera da DIF assume sette modalità, così specificate:

$$C_i = \begin{cases} 1 & \text{se } Y_i < Z_{i1} \\ 2 & \text{se } Y_i = Z_{i1} \\ 3 & \text{se } Z_{i1} < Y_i < Z_{i2} \\ 4 & \text{se } Y_i = Z_{i2} \\ 5 & \text{se } Z_{i2} < Y_i < Z_{i3} \\ 6 & \text{se } Y_i = Z_{i3} \\ 7 & \text{se } Y_i > Z_{i3} \end{cases}$$

Nella prima e quarta colonna della Tabella 2.1 è presentato l'ordinamento dei 75 casi possibili, in colonna 2 e 5 i valori di C che sono vettoriali e scalari, mentre nelle rimanenti due colonne viene riportato un valore scelto da una scala da 1 a 7, che ricodifica in valori scalari le inconsistenze. Le prime sette situazioni della parte a sinistra della tabella rispettano l'ordinamento naturale, non hanno *ties* e corrispondono dunque ad una variabile C scalare, mentre tutte le altre sono *ties* o inconsistenze: nello specifico quando

Tabella 2.1: Classificazione dell'autovalutazione Y_s e valutazione di 3 *vignettes* Y_1, Y_2, Y_3 . Fonte: Van Soest e Vonkova (2014).*

Ordinamento	C vettoriale	Celle di C	Ordinamento	C vettoriale	Celle di C
$Y_s < Z_1 < Z_2 < Z_3$	1	1	$Y_s < Z_1 < Z_3 < Z_2$	1	1
$Y_s = Z_1 < Z_2 < Z_3$	2	2	$Y_s = Z_1 < Z_3 < Z_2$	2	2
$Z_1 < Y_s < Z_2 < Z_3$	3	3	$Z_1 < Y_s < Z_3 < Z_2$	3	3
$Z_1 < Y_s = Z_2 < Z_3$	4	4	$Z_1 < Y_s = Z_3 < Z_2$	3-6	16
$Z_1 < Z_2 < Y_s < Z_3$	5	5	$Z_1 < Z_3 < Y_s < Z_2$	3-7	17
$Z_1 < Z_2 < Y_s = Z_3$	6	6	$Z_1 < Z_3 < Y_s = Z_2$	4-7	19
$Z_1 < Z_2 < Z_3 < Y_s$	7	7	$Z_1 < Z_3 < Z_2 < Y_s$	7	7
$Y_s < Z_2 < Z_1 < Z_3$	1	1	$Y_s < Z_2 < Z_3 < Z_1$	1	1
$Y_s = Z_2 < Z_1 < Z_3$	1-4	8	$Y_s = Z_2 < Z_3 < Z_1$	1-4	8
$Z_2 < Y_s < Z_1 < Z_3$	1-5	9	$Z_2 < Y_s < Z_3 < Z_1$	1-5	9
$Z_2 < Y_s = Z_1 < Z_3$	2-5	13	$Z_2 < Y_s = Z_3 < Z_1$	1-6	10
$Z_2 < Z_1 < Y_s < Z_3$	5	5	$Z_2 < Z_3 < Y_s < Z_1$	1-7	11
$Z_2 < Z_1 < Y_s = Z_3$	6	6	$Z_2 < Z_3 < Y_s = Z_1$	2-7	15
$Z_2 < Z_1 < Z_3 < Y_s$	7	7	$Z_2 < Z_3 < Z_1 < Y_s$	7	7
$Y_s < Z_3 < Z_1 < Z_2$	1	1	$Y_s < Z_3 < Z_2 < Z_1$	1	1
$Y_s = Z_3 < Z_1 < Z_2$	1-6	10	$Y_s = Z_3 < Z_2 < Z_1$	1-6	10
$Z_3 < Y_s < Z_1 < Z_2$	1-7	11	$Z_3 < Y_s < Z_2 < Z_1$	1-7	11
$Z_3 < Y_s = Z_1 < Z_2$	2-7	15	$Z_3 < Y_s = Z_2 < Z_1$	1-7	11
$Z_3 < Z_1 < Y_s < Z_2$	3-7	17	$Z_3 < Z_2 < Y_s < Z_1$	1-7	11
$Z_3 < Z_1 < Y_s = Z_2$	4-7	19	$Z_3 < Z_2 < Y_s = Z_1$	2-7	15
$Z_3 < Z_1 < Z_2 < Y_s$	7	7	$Z_3 < Z_2 < Z_1 < Y_s$	7	7
$Y_s < Z_1 = Z_2 < Z_3$	1	1	$Y_s < Z_3 < Z_1 = Z_2$	1	1
$Y_s = Z_1 = Z_2 < Z_3$	2-4	12	$Y_s = Z_3 < Z_1 = Z_2$	1-6	10
$Z_1 = Z_2 < Y_s < Z_3$	5	5	$Z_3 < Y_s < Z_1 = Z_2$	1-7	11
$Z_1 = Z_2 < Y_s = Z_3$	6	6	$Z_3 < Y_s = Z_1 = Z_2$	2-7	15
$Z_1 = Z_2 < Z_3 < Y_s$	7	7	$Z_3 < Z_1 = Z_2 < Y_s$	7	7
$Y_s < Z_1 = Z_3 < Z_2$	1	1	$Y_s < Z_2 < Z_1 = Z_3$	1	1
$Y_s = Z_1 = Z_3 < Z_2$	2-6	14	$Y_s = Z_2 < Z_1 = Z_3$	1-4	8
$Z_1 = Z_3 < Y_s < Z_2$	3-7	17	$Z_2 < Y_s < Z_1 = Z_3$	1-5	9
$Z_1 = Z_3 < Y_s = Z_2$	4-7	19	$Z_2 < Y_s = Z_1 = Z_3$	2-6	14
$Z_1 = Z_3 < Z_2 < Y_s$	7	7	$Z_2 < Z_1 = Z_3 < Y_s$	7	7
$Y_s < Z_1 < Z_2 = Z_3$	1	1	$Y_s < Z_2 = Z_3 < Z_1$	1	1
$Y_s = Z_1 < Z_2 = Z_3$	2	2	$Y_s = Z_2 = Z_3 < Z_1$	1-6	10
$Z_1 < Y_s < Z_2 = Z_3$	3	3	$Z_2 = Z_3 < Y_s < Z_1$	1-7	11
$Z_1 < Y_s = Z_2 = Z_3$	4-6	18	$Z_2 = Z_3 < Y_s = Z_1$	2-7	15
$Z_1 < Z_2 = Z_3 < Y_s$	7	7	$Z_2 = Z_3 < Z_1 < Y_s$	7	7
$Y_s < Z_1 = Z_2 = Z_3$	1	1	$Y_s = Z_1 = Z_2 = Z_3$	2-6	14
$Z_1 = Z_2 = Z_3 < Y_s$	7	7			

* Nella colonna 3 e 6 viene riportata l'etichetta associata ai casi *ties*; in particolare C=1-4, 1-5, 1-6, 1-7, 2-4, 2-5, 2-6, 2-7, 3-6, 3-7, 4-6, 4-7 viene ridefinita rispettivamente con una scala da 8 a 19.

almeno una delle *vignettes* assume la stessa modalità di un'altra si parla di *ties*, negli altri casi caratterizzati da ordinamento incorretto si usa il termine inconsistenza. Conseguentemente, la tabella chiarifica l'impatto di queste incongruenze tra gli scenari ipotetici rispetto all'ordinamento naturale; la colonna 2 e 4 corrispondono ai valori che dovrebbe assumere C se si optasse per una formulazione vettoriale: il range $q - r$ riportato corrisponde al minimo valore della scala C per cui viene soddisfatta una delle relazioni tra l'autovalutazione e una delle *vignettes* e il massimo valore della stessa scala C per cui si verifica la relazione con un'altra *vignette*; ovviamente nell'ipotesi in cui venga soddisfatta solo una delle $2J + 1$ possibilità, si avrà un valore scalare: questo è il caso, ad esempio, di $Y_s < Z_2 < Z_1 < Z_3$, dove C assume la modalità 1 in quanto l'unica relazione che viene verificata è $Y_s < Z_1$. Tale modalità corrisponde alla prima situazione di corretto ordinamento e assenza di *ties* (cioè $Y_s < Z_1 < Z_2 < Z_3$); per poter fare un confronto non parametrico le due situazioni vengono associate in una delle categorie di una nuova variabile unidimensionale che assume il valore 1 nella colonna 3. In alcuni casi però la collocazione di Y_s è ambigua: si prenda in considerazione, ad esempio, $Z_3 < Y_s < Z_1 < Z_2$, avente range $\{q, \dots, r\} = \{1, \dots, 7\}$; se $Z_3 < Y_s$ e Z_1 e Z_2 sono mal riportati, allora la variabile C generalizzata dovrebbe assumere il valore 7, ma se $Y_s < Z_1$ e Z_3 è riportato erroneamente, allora si dovrebbe assegnare la categoria 1. In questo caso dunque non si potrebbe dire nulla sulla vera posizione di Y_s in quanto non si può sapere quale ordinamento non sia stato compreso correttamente dal rispondente; per questo, nell'ultima colonna della tabella viene associato il valore 11. Categorizzando dunque in questo modo ogni osservazione in una specifica cella, si ottiene una variabile C generalizzata con 19 livelli. Idealmente, dunque, dovrebbe verificarsi che le celle non corrette contengano un numero di osservazioni così piccolo che possano essere scartate senza perdita di informazione rilevante (Van Soest e Vonkova, 2014).

Al fine di tentare di superare le inconsistenze della scala C, alcuni ricercatori hanno recentemente proposto delle soluzioni complementari o alternative; l'obiettivo è quello di cercare di collocare queste inconsistenze in una categoria di C nel miglior modo possibile così che simultaneamente si usi l'informazione dei valori scalari e di quelli vettoriali. Per evitare eventuali

fraintendimenti si pensi alla variabile C ricodificata da vettoriale a scalare come a una nuova variabile C^* ; chiaramente C^* avrà le stesse identiche modalità di C per quelle osservazioni a cui è attribuito un valore scalare di C . Nello specifico al momento in letteratura sono presenti tre possibili approcci:

1) *Omissione dei valori vettoriali*:

si crea la variabile C^* ignorando i casi di C con valori associati a un intervallo, che quindi vengono trattati come valori mancanti; i rimanenti valori scalari possono dunque essere rappresentati attraverso un istogramma. Si tratta di un approccio molto semplice che purtroppo comporta una grandissima perdita di informazione, specialmente nel caso in cui a ciascuna modalità di C^* sia associata una frequenza relativamente bassa. Questo approccio non supera le limitazioni della scala C , ma rappresenta un'utile metodo di confronto (*baseline*) con altri approcci che assegnano i valori vettoriali di C a componenti scalari di C^* .

2) *Collocazione uniforme* (King et al., 2004):

a ciascuna modalità scalare di C^* si attribuiscono le frequenze assolute standardizzate per l'ampiezza dell'intervallo, più precisamente le frequenze assolute associate alle osservazioni che hanno come valore minimo dell'intervallo la categoria di C che si sta considerando. Ovviamente più grande è il numero di casi con valori vettoriali, più uniforme e più variabile sarà la nuova variabile C^* .

3) *Collocazione secondo un probit ordinato censurato* (King e Wand, 2007):

il modello probit ordinato censurato ha lo scopo di stimare la proporzione degli individui più "simili" che scelgono risposte con valori vettoriali, ottenendo dunque una stima della proporzione del campione in ogni categoria della nuova variabile osservata codificata C^* . Si tratta di una generalizzazione del modello probit ordinato in quanto tiene conto dei valori vettoriali (detti valori censurati) come variabili dipendenti e considera la probabilità di essere in una specifica singola categoria condizionatamente ad un valore osservato vettoriale o scalare; il termine censurato sta ad indicare che l'individuo i è attualmente nell'intervallo $\{q, \dots, r\}$ e dunque che per valori non compresi in tale intervallo la relativa probabilità di essere osservato è nulla. In sostanza dunque è una combinazione tra il modello probit ordinato e quello tobit. Formalmente, sia C_i la variabile dipendente reale, inosservata e continua con

$i = 1, 2, \dots, n$ che indica il rispondente e sia X_i un vettore di variabili esplicative con β il relativo vettore di parametri associato (per l'identificazione senza termine costante). Si supponga dunque che tale variabile sia combinazione lineare di X_i e dunque che $C_i \sim N(x_i'\beta, 1)$. Essa genera la variabile osservata C_i^* attraverso il seguente meccanismo:

$$C_i^* = \{q, \dots, r\} \quad \text{se } \tau^{q-1} \leq C_i < \tau^r \text{ e } q < r$$

dove $-\infty = \tau^0 < \tau^1 < \dots < \tau^{2J} < \tau^{2J+1} = +\infty$

con $c = \{q, \dots, r\}$ un intervallo avente per estremi uno dei seguenti valori: $(1, 2, \dots, 2J + 1)$. Per ogni individuo i si è dunque interessati a calcolare la probabilità che la variabile latente C_i cada tra due soglie che corrispondono a C_i^* :

$$Pr(C_i^* = \{q, \dots, r\} | x_i) = \int_{\tau^{q-1}}^{\tau^r} N(C_i^* | x_i'\beta) dt.$$

La stima dei parametri, la loro interpretazione, la stima della probabilità di un intervistato a rispondere ad una categoria j , ecc si possono ottenere con le stesse procedure e metodologie dell'approccio probit ordinato. Infatti nel caso in cui $\{q, \dots, r\}$ contenesse un solo valore, la probabilità precedente sarebbe identica a quella del probit ordinato standard e quindi i due modelli coinciderebbero.

Ovviamente si potrebbe soltanto stimare tale modello e interpretare i parametri β esattamente come se si stesse considerando un modello probit ordinato. Tuttavia per poter rappresentare la variabile corretta da DIF si calcola la probabilità condizionata a un insieme di caratteristiche individuali ($x_i = x_0$) e ai valori censurati (c_i); conseguentemente nel modello che si sta analizzando tale probabilità è:

$$Pr(C^* = j | x_0, c_i = \{q, \dots, r\}) = \begin{cases} \frac{Pr(C^*=j|x_0)}{\sum_{a \in c_i} Pr(C^*=a|x_0)} & \text{se } j \in c_i \\ 0 & \text{altrimenti} \end{cases}$$

con $j = 1, 2, \dots, 2J + 1$. Ne consegue che la somma delle probabilità delle modalità all'interno dell'intervallo $\{q, \dots, r\}$ è pari a 1 e al di fuori di questo intervallo è 0. In particolare se c_i è un valore scalare, tale probabilità è pari a 1 per la categoria j , 0 altrimenti. In altre parole, la frequenza relativa di ciascuna modalità scalare di C^* è data dalla media delle previsioni condizionate per una categoria j .

Le probabilità così calcolate possono essere sommate per produrre un istogramma o possono essere interpretate per effetti causali (King e Wand, 2007). Questo approccio risulta quindi robusto per specificazioni errate ed estrae considerevolmente più informazione dai dati relativi alle *anchoring vignettes*. In letteratura la scala C viene utilizzata principalmente per il confronto tra due gruppi o Paesi e dunque comparando i due istogrammi relativi alle nuove variabili (King et al. (2004), Wand et al. (2011), Wand (2013)); conseguentemente, l'interpretazione consiste nel confronto della distribuzione di questa nuova variabile C nei diversi gruppi d'interesse e dunque nel confronto tra le autovalutazioni originarie dei rispondenti di due gruppi e tra quelle corrette secondo la relazione delle valutazioni soggettive e quelle attribuite agli individui ipotetici. Tuttavia, in letteratura non si trovano lavori che sfruttano tale metodo: la facilità di questo strumento consente di fare dei test diagnostici sulle assunzioni del metodo parametrico e di testare la validità delle assunzioni delle *vignettes* (Van Soest e Vonkova, 2014). Molto carente, se non quasi assente, è invece l'utilizzo della nuova variabile corretta dalla DIF come variabile esplicativa (Primi et al. (2016) stimano un modello ad effetti fissi e ad effetti casuali) o come variabile dipendente (Crane et al. (2016) considerano modelli di regressioni ordinali).

Un'alternativa alla scala C è la scala B (Wand, 2013), un approccio che rende credibili i confronti tra gli individui in presenza di DIF, senza richiedere l'assunzione di *vignette equivalence*. Esattamente come la scala di King et al. (2004), questa scala è costruita sulla base delle risposte all'autovalutazione e alle *vignettes*, ma rilassando una delle assunzioni permette ad una modalità di essere rappresentata da un insieme di valori, più che da uno singolo; dunque la soluzione di Wand (2013) richiede meno informazione in presenza di *ties* ovvero quando $Y_i = Z_{ij}$. Se è vero che da una parte tale scala rilassa una delle assunzioni, dall'altra introduce la cosiddetta "*Ordered Preserving Imperfect Anchors*" (OPIA).

La soluzione di Wand (2013) tuttavia non ha ancora avuto particolarmente successo: infatti attualmente in letteratura viene sfruttata solo dallo stesso autore. Non essendo questo metodo d'interesse in questa tesi, si veda Wand (2013) per ulteriori dettagli.

2.1.4.4 Vantaggi e svantaggi del metodo parametrico e non parametrico

Le *anchoring vignettes* hanno il potenziale di ridurre la distorsione, aumentare l'efficienza e rendere le misure più vicine al confronto interpersonale rispetto ad altri metodi esistenti (King et al., 2004). D'altro canto, come tutti i metodi, le due soluzioni proposte da King et al. (2004) presentano vantaggi e svantaggi, riassunti qui di seguito rispettivamente con un [+] e un [-]:

Approccio parametrico:

[+] corregge la DIF permettendo alle soglie di dipendere da delle caratteristiche osservate e considerando l'informazione delle *vignettes*

[+] riduzione dei costi in quanto non è necessario raccogliere le risposte a tutte le *vignettes* per ogni individuo; i ricercatori possono dunque sottoporre gli scenari ipotetici a un sottocampione di quello a cui viene chiesto di dare la propria valutazione alla domanda d'interesse o nel caso di dati panel solo ad alcune coorti.

[+] riconosce che la variabile d'interesse può essere percepita con un errore di misura o in altre parole riconosce che alcuni intervistati possono non individuare l'ordinamento naturale delle *vignettes*

Approccio non parametrico:

[+] corregge la DIF confrontando la valutazione alla domanda di interesse con quella relativa agli individui ipotetici

[-] è necessario disporre delle risposte a tutte le *vignettes* e autovalutazioni per ogni intervistato; ciò richiede costi più alti, nonché una (possibile e) maggior perdita di informazione

[-] deve esistere un ordinamento naturale degli scenari ipotetici; ciò comporta la difficoltà di interpretare possibili divergenze (*ties* e ordinamento incorretto) e conseguentemente possibile perdita di informazione, in assenza di ulteriori assunzioni (nella Sezione 2.1.4.3 si sono viste delle soluzioni alternative)

- [−] oltre alla *response consistency* e alla *vignette equivalence* sono necessarie altre assunzioni, come ad esempio la specificazione lineare della funzione che lega le caratteristiche osservabili con le componenti non osservabili, la forma funzionale delle soglie, la forma della distribuzione dei termini d'errore
- [+] è un approccio molto semplice ed intuitivo, facile da implementare e che non richiede ulteriori assunzioni oltre alla *response consistency* e alla *vignette equivalence*, anche se richiederebbe un numero elevato di dati
- [+] è possibile costruire distribuzioni controfattuali e quindi usare la scala di un gruppo di individui o di Paesi come *benchmark*
- [−] al momento in letteratura non vi sono studi sulla costruzioni di controfattuali con il metodo non parametrico
- [+] consente di stimare direttamente gli effetti delle caratteristiche osservabili degli intervistati, sia nei livelli di autovalutazione sia in quelli delle *vignettes*
- [+] la variabile libera da DIF può essere analizzata attraverso un altro modello, che permette di stimare direttamente gli effetti delle caratteristiche osservabili
- [−] l'assunzione di linearità delle soglie rispetto alle caratteristiche osservabili potrebbe essere troppo restrittiva
- [+] non si assume alcuna specificazione per le soglie usate nella valutazione degli individui ipotetici e di se stessi; in particolare le prime vengono sfruttate per reinterpretare e riscalarle le risposte alle autovalutazioni
- [−] nella costruzione del modello, non si confronta direttamente la valutazione data agli individui ipotetici e a se stessi
- [+] permette di selezionare le *vignettes* che forniscono la migliore informazione estratta sull'uso differente di scale (Sezione 2.1.5)

- [+] considera la variabile originale relativa all'autovalutazione e dunque permette una più facile interpretazione delle k categorie (originarie).
- [±] l'interpretazione dell'oggetto d'interesse non è diretta, ma in relazione all'ordinamento delle *vignettes*; ad ogni modo, il livello di soddisfazione, ad esempio, aumenta al crescere del valore della modalità attribuita alla nuova variabile. Il numero di livelli della variabile C è pari a $2J + 1$, in alcuni casi inferiore al numero k di categorie della variabile risposta determinando quindi una possibile perdita di informazione (recuperabile in parte aggiungendo più *vignettes*)
- [-] se le soglie τ_i sono combinazioni lineari delle stesse caratteristiche osservate del livello reale μ_i (ovvero $V_i \equiv X_i$), il numero di parametri coinvolti nella stima potrebbe essere molto elevato e dunque il costo computazionale aumenterebbe
- [+] il costo computazionale è basso in quanto confronta solo delle valutazioni e non deve stimare alcun modello o parametro. Inoltre non usufruisce dell'informazione delle covariate a disposizione; conseguentemente è possibile usare la variabile C per ulteriori analisi.

In letteratura l'approccio più utilizzato è quello parametrico, probabilmente perchè i suoi vantaggi risultano più convincenti rispetto a quelli del non parametrico. Nonostante la sua forte popolarità, alcuni ricercatori hanno confrontato i due approcci per testarne la validità: se da una parte alcuni studi concludono che i due metodi portano a risultati simili (Jones et al. (2012), Crane et al. (2016)), altri riscontrano la necessità di modelli più flessibili (Van Soest et al. (2011), Van Soest e Vonkova (2014)).

2.1.5 Selezione delle *vignettes*

I risultati dei due metodi dipendono da quali *vignettes* vengono utilizzate nell'analisi, specialmente nel metodo non parametrico dove ha rilevanza anche l'ordine considerato. King e Wand (2007) offrono una possibile soluzione per la selezione delle *vignettes* e il problema del loro ordinamento, successivamente esposta anche da Wand et al. (2011) e basata sul metodo non parametrico. La loro proposta consiste nel considerare tutte le possibili combinazioni di sottoinsiemi di *vignettes* e per ognuna di queste calcolare l'entropia minima e l'entropia basata sui valori stimati dal modello probit ordinale censurato. Dato un numero J di *vignettes*, è possibile scegliere tra

$$J + \binom{J}{2} + \binom{J}{3} + \dots + \binom{J}{J-1} + 1;$$

poichè non sempre il "più" comporta una informazione statisticamente informativa, è necessario avere un *trade-off*: il numero di scenari ipotetici da considerare coincide con un bilanciamento tra l'aver una scala C più precisa (che induca un'uniformità maggiore delle celle) e tra l'aver il minor numero di *ties* ed inconsistenze. La scelta più appropriata ricade dunque nella combinazione di *vignettes* che comportano la minor differenza tra le due misure di entropia, condizionatamente al numero di *vignettes* che si intende utilizzare: una differenza minore indica quindi una distribuzione più uniforme della variabile d'interesse C nei due approcci.

Formalmente, l'entropia è così specificata:

$$H = - \sum_{j=1}^{2J+1} P(\hat{C} = j) \ln P(\hat{C} = j) \in [0, \ln(2J + 1)]$$

dove per convenienza si definisce $-0 \ln(0) \equiv 0$ (dal momento che $\lim_{a \rightarrow 0^+} a \ln(\frac{1}{a}) = 0$). Questa misura è tale che $H = 0$ se e solo se tutti gli intervistati appartengono a una sola categoria di C , mentre è massima se è uguale a $\ln(2J + 1)$ ovvero se $p_j = \frac{1}{2J+1}$.

Nel caso in cui si abbiano valori vettoriali di C si usa l'entropia stimata, quella calcolata in seguito all'allocazione di un probit ordinato censurato con solo l'intercetta, o l'entropia nota (minima) (Wand et al., 2011). Ne derivano le seguenti caratteristiche:

- è funzione delle frequenze p_1, \dots, p_{2J+1} ,
- è una funzione continua e monotona crescente del numero di *vignettes* J ,
- assume il valore minimo quando in una categoria vi sono tutti gli n intervistati e quello massimo quando i soggetti si distribuiscono equamente tra le categorie,
- la quantità di informazione attesa nell'unione degli intervalli più piccoli dovrebbe rimanere la stessa dell'intervallo originale non decomposto, mentre l'informazione attesa negli altri intervalli rimanere la stessa aggiungendo nuove *vignettes*.

Nella Tabella 2.3 vengono riportati i risultati in riferimento a tre dei sei scenari ipotetici (presentati a pagina 32) relativi ai dati utilizzati da King et al. (2004)⁶. La coppia di *vignettes* 13 (cioè 1+3) presenta la differenza di entropia minore e il minor numero di valori intervallari, condizionatamente alle due valutazioni dei soggetti ipotetici; conseguentemente se si vuole ottenere una distribuzione più uniforme possibile e col minor numero di *ties* e inconsistenze attraverso l'informazione di solo due *vignettes*, questa sembra essere la scelta più appropriata. Si noti in particolare che le due misure di entropia relative ai casi in cui si considera una sola valutazione sono identici: l'incertezza dell'entropia stimata è infatti dovuta interamente agli ordinamenti non corretti e questi non sono possibili nel caso singolo scalare.

Se si ha una misura oggettiva è invece possibile selezionare le *vignettes* in base alla correlazione tra l'autovalutazione dell'oggetto di interesse e quella dell'indicatore oggettivo, così come fanno Van Soest et al. (2011).

Per quanto riguarda l'ordine delle *vignettes* da considerare, la definizione della scala C impone che venga fissato a priori. Normalmente esso viene scelto dal ricercatore in base alla gravità o meno della situazione descritta. Wand et al. (2011) propongono una strada alternativa, ovvero quella di considerare l'ordinamento consensuale degli intervistati. Chiaramente, le differenze tra i due tipi di approcci possono essere usati per diagnosticare possibili problemi delle domande associate agli individui ipotetici, come ad esempio la loro

⁶I dati sono scaricabili dalla libreria "*anchors*" del Software RStudio Team, 2015

Tabella 2.3: Sommario dell'entropia e dei casi intervallari di sottoinsiemi di *vignettes*[†]

	<i>Combinazione Vignettes</i>	Entropia Stimata	Entropia minima	N° casi intervallari
1	123	1.710	1.418	328
2	13	1.446	1.352	142
3	23	1.359	1.255	149
4	12	1.498	1.308	253
5	3	0.843	0.843	0
6	2	1.043	1.043	0
7	1	1.084	1.084	0

[†]I numeri 1,2,3 corrispondono rispettivamente alla *vignette* di Moses, Jane e Alison.

traduzione.

L'ordine di scala delle *vignettes* di King et al. (2004) per la *political efficacy* (presentato a pagina 32) è dal più libero al meno; conseguentemente Alison è l'individuo ipotetico più soddisfatto, mentre Moses il meno. Nella Tabella 2.4 vengono riportate le proporzioni dei diversi casi in cui una *vignette* i (in riga) è inferiore a un'altra j (in colonna), ottenute dall'analisi dei dati utilizzati da King et al. (2004). Nello specifico nella matrice triangolare superiore si ha $p_{ij} - p_{ji}$ e nella parte inferiore $1 - p_{ij} - p_{ji}$; valori negativi del primo caso indicano un errato ordinamento degli scenari ipotetici, mentre valori elevati nel secondo caso sono segno della presenza di numerosi *ties*.

L'ordinamento consensuale degli intervistati rispecchia quello della "gravità" descritto nelle tre *vignettes*. Se invece si fosse verificato che una differenza negativa di $p_{23} - p_{32}$, anziché pari a 0.365, ciò sarebbe stato indice del fatto che gli individui avrebbero classificato Jane come più soddisfatta rispetto ad Alison.

2.1.6 Verifica delle assunzioni

Di fianco alla rapida crescita dell'utilizzo di tale strumento, vi è un crescente interesse nel valutare e validare le assunzioni su cui esso si basa. Infatti se individui diversi hanno scale di risposta non uguali e se le assunzioni non

Tabella 2.4: Errato ordinamento e presenza di *ties*

	<1 (Moses)	<2 (Jane)	<3 (Alison)
1	-	0.159	0.438
2	0.388	-	0.365
3	0.290	0.301	-

La proporzione dei casi in cui una *vignette* (in riga) è inferiore a un'altra (in colonna)

risultano essere verificate, le risposte alle autovalutazioni non sarebbero comparabili e dunque i risultati emergenti dalle analisi non sarebbero affidabili, poichè le *vignettes* sarebbero influenzate da eterogeneità, anzichè modellarla. Tuttavia la validità della *response consistency* e della *vignette equivalence* rimane ancora una questione aperta, dal momento che tuttora non esistono test formali. Inoltre le soluzioni attualmente presenti in letteratura verificano le due assunzioni solo attraverso (ulteriori) assunzioni o restrizioni sul modello parametrico Chopit; conseguentemente non si può dire se un'assunzione venga rigettata per la sua effettiva invalidità o perchè le condizioni aggiuntive al modello risultino essere inappropriate. Oltre a questo, le conclusioni dei ricercatori non sono unidirezionali, in quanto le assunzioni sono accettate in alcuni casi e violate in altri. Per questi motivi, tendenzialmente si opta per applicare i metodi, senza accettarsi sulla validità degli assunti su cui si basano. Di seguito si riportano quindi alcune soluzioni proposte per la loro verifica.

Response consistency

La situazione raffigurata negli scenari ipotetici è la stessa per tutti gli intervistati e dunque il livello reale della persona descritta nella *vignette* dovrebbe essere valutato nello stesso modo da tutti gli intervistati; purtroppo ciò non accade in quanto al momento della valutazione ognuno applica la propria scala (si tratta del cosiddetto fenomeno DIF). Al fine di raggiungere l'obiettivo che ogni intervistato usufruisca della stessa scala (o quasi) sia alla domanda sull'autovalutazione sia su quella dello scenario ipotetico, è opportuno che le *vignettes* descrivano il più accuratamente possibile il livello reale dell'individuo ipotetico, adeguato alla sfera di interesse, e attraverso una sintassi

linguistica che sia vicina quanto più immaginabile e possibile a quella dell'intervistato. Nel rispondere poi alla domanda, l'individuo deve convertire il livello percepito dell'autovalutazione e quello dell'individuo ipotetico attraverso una scala di risposta: la *response consistency* assume così che ogni intervistato faccia questa conversione usando la stessa scala di risposta individuale o in altre parole applicando le stesse soglie τ_{is}^k al momento della valutazione $\forall i, k$ con $i = 1, \dots, n$ e $k = 1, \dots, K_s$.

Attualmente il test maggiormente preferito è quello proposto da Van Soest et al. (2011), che richiede però l'utilizzo di un indicatore oggettivo del costrutto di interesse; perciò è necessario che nel questionario sia presente un'ulteriore domanda/misura oggettiva relativa a questo costrutto. Questa misura oggettiva ha le seguenti caratteristiche:

- è in grado di catturare il fenomeno d'interesse in modo oggettivo e quindi anche le differenze dell'autovalutazione e delle caratteristiche individuali osservate;
- non è influenzata da eterogeneità in quanto non è soggettiva e nel rispondere non si applicano scale di risposta diverse;
- è guidata dallo stesso processo latente che genera le autovalutazioni; questa assunzione naturale è denominata assunzione *one-factor* (OF).

Formalmente, alle due equazioni del modello Chopit viene aggiunta una terza equazione, modellata anch'essa come un probit ordinato e così specificata:

$$\begin{aligned}
 Y_{io}^* &= X_i \beta_o + \zeta_{io} \\
 \text{con } \zeta_{io} &\perp (X_i, \xi_i, u_{ij}) \\
 Y_{io} &= l \quad \text{se } \tau_o^{l-1} \leq Y_{io}^* < \tau_o^l \\
 \text{dove } -\infty &= \tau_o^0 < \tau_o^1 < \dots < \tau_o^{L-1} < \tau_o^L = +\infty
 \end{aligned}$$

dove $l = 1, 2, \dots, L$ con L il numero di modalità della domanda "oggettiva" (può essere differente dal numero di modalità K dell'autovalutazione e delle *vignettes*), Y_{io}^* e Y_{io} sono rispettivamente il livello percepito e quello riportato dell'indicatore oggettivo, X_i sono le variabili osservate (identiche a quelle della componente di autovalutazione), β_o il vettore dei parametri, ζ_{io} il termine d'errore (indipendente dalle covariate, dal termine d'errore delle *vignettes* e

dal termine d'errore nella prima soglia (si veda pp. 45)). In particolare ζ_{io} può essere correlato con il termine d'errore della componente dell'autovalutazione ε_i poichè entrambi sono influenzati da un fattore comune inosservato che guida il processo latente d'interesse, dal momento che entrambi sono influenzati da un errore idiosincratico, normalmente distribuito; si assume inoltre che $(\zeta_{io}, \varepsilon_{is})$ si distribuisca come una normale bivariata.

In base al modello così specificato, l'assunzione OF verifica che $\beta_o = \beta$ cioè che il vettore della componente oggettiva sia uguale a quello dei parametri della componente di autovalutazione. Perciò, un test formale si ottiene confrontando il modello in cui si impone la *response consistency* e la *one-factor* con quello in cui si richiede solo la *one-factor*. Il secondo modello risulta essere identificato poichè le *vignettes* vengono sfruttate per stimare i parametri delle soglie attraverso γ^k , mentre la misura oggettiva viene utilizzata per stimare i parametri nell'autovalutazione $\beta_o = \beta$. Anche con queste assunzioni, il modello viene stimato tramite il metodo della massima verosimiglianza.

Van Soest et al. (2011) applicano tale test a un campione di studenti irlandesi, con lo scopo di indagare il loro comportamento nei confronti dell'utilizzo dell'alcool: chiaramente l'autovalutazione chiede di dare la propria valutazione su tale relazione, mentre la misura oggettiva utilizzata fa riferimento a una domanda relativa alla quantità di bevande alcoliche consumate (un indicatore oggettivo e strettamente legato alla sfera d'interesse). I risultati mostrano che il modello che rilassa la *response consistency* fornisce stime peggiori di quello che la include (ovvero ha un AIC più alto) e dunque si è a favore delle *anchoring vignettes*.

Per l'applicazione del test è fondamentale la presenza di un indicatore oggettivo e che tale indicatore sia logicamente legato al costrutto di interesse (cioè deve essere in grado di cogliere tutta la variazione del costrutto d'interesse e delle caratteristiche osservabili che influenzano il comportamento riportato (Bago d'Uva et al., 2011)). Sotto tali assunzioni, qualsiasi variazione sistematica nella valutazione soggettiva che rimane dopo essersi condizionati all'indicatore oggettivo può essere attribuita all'eterogeneità riportata al momento della valutazione. In base a ciò, la scelta adottata da Datta Gupta et al. (2010) risulta essere alquanto discutibile: attraverso l'applicazione dello stesso test di Van Soest et al. (2011) sulla disabilità lavorativa (limi-

tazioni nel tipo e quantità di lavoro a causa di un problema di salute) di alcuni ultracinquantenni di 11 Paesi Europei, questi ricercatori rigettano la validità dell'assunzione. La misura oggettiva da essi presa in esame è la *grip strength*; tale indicatore tuttavia risulta questionabile perchè, ad esempio, possono esservi lavoratori costretti a stare in carrozzina, ma che non presentano disabilità alle mani e dunque che probabilmente hanno un'ottima forza in esse.

Più appropriata invece risulta essere la scelta di Paccagnella et al. (2015), i quali accettano la validità dell'assunzione: analizzando i dati di un'indagine sulla soddisfazione dopo l'acquisto di uno smartphone la misura oggettiva viene costruita per ogni intervistato attraverso il confronto tra le aspettative prima dell'acquisto del prodotto e la sua performance percepita dopo l'acquisto (un'idea simile a quella su cui si fonda un esperimento di Kapteyn et al. (2011), che viene presentata tra qualche riga).

Un'estensione all'approccio di Van Soest et al. (2011) è stato proposto da Bago d'Uva et al. (2011): applicando l'approccio delle *anchoring vignettes* a due distinti domini della salute, la mobilità e la cognizione, questi autori prendono in considerazione più di un indicatore oggettivo. La loro proposta risulta essere particolarmente efficiente nel caso in cui un singolo indicatore non risulti essere in grado di cogliere tutte le associazioni tra le covariate e il dominio oggetto d'indagine: è infatti più credibile che più misure oggettive provvedano informazione sufficiente da supporre che la OF sia verificata, cogliendo così tutti gli aspetti del costrutto d'interesse. Bago d'Uva et al. (2011) introducono inoltre un secondo test (più debole) che è valido anche nel caso in cui non si riesca a cogliere tutta la variazione: verificano se la differenza tra due soglie è la stessa tra il modello che considera la *response consistency* e la *one-factor* e quello che rilassa la prima di queste due assunzioni ovvero $\gamma^k - \gamma^{k-1} = \gamma'^k - \gamma'^{k-1}$. Le misure oggettive utilizzate riguardano capacità motorie (velocità nel camminare e *grip strength*). I risultati che emergono dal loro studio portano a rigettare la *response consistency* per entrambi i due domini analizzati. In letteratura non si sono tuttavia riscontrati altre pubblicazioni riguardo il test debole di Bago d'Uva et al. (2011).

Una differente soluzione è stata proposta da Kapteyn et al. (2011), i quali hanno analizzato cinque sfere della salute (sonno, mobilità, concentrazione,

respirazione, sentirsi giù) relative ad una popolazione di adulti statunitensi. Al fine di assicurare che l'assunzione in esame sia verificata ovvero che non vi sia differenza sistematica tra le scale di risposta usate dagli intervistati per la classificazione dell'autovalutazione e delle *vignettes*, in una prima intervista viene chiesto di dare una valutazione sulla propria salute (per ogni dominio considerato), mentre solo in un'indagine successiva viene domandato di scegliere una delle categorie della scala Likert per la salute dell'individuo ipotetico di una *vignette* opportunamente costruita in base a due domande specifiche della prima indagine (questa viene denominata *replica vignette*). In questo modo gli individui non si rendono conto che le *vignettes* descrivono proprio il loro stato di salute e ciò dovrebbe dunque garantire l'uso della stessa scala di risposta per la valutazione di se stessi e delle persone ipotetiche. Se si verifica tutto ciò e sotto alcune deboli assunzioni⁷, allora la *response consistency* corrisponde all'ipotesi nulla che le due distribuzioni siano le stesse (o quasi); ciò viene verificato attraverso due test non parametrici (test dei ranghi segnati e test dei segni). Oltre a questo viene stimato un modello parametrico che attraverso delle covariate (caratteristiche osservate dei rispondenti e caratteristiche descrittive gli stessi problemi di salute del soggetto in termini del costrutto d'interesse) spiega la relazione tra le autovalutazioni e le valutazioni delle *replica vignettes*, una generalizzazione del modello Chopit e dell'estensione di Kapteyn et al. (2007). La loro analisi porta a non rifiutare l'ipotesi nulla e l'assunzione per il sonno e a non accettarle per le altre 4 sfere della salute analizzate.

Un'ulteriore strada possibile viene data da Angelini et al. (2014): gli intervistati che hanno caratteristiche più vicine agli individui ipotetici dovrebbero dare valutazioni simili alle autovalutazioni e alle *vignettes*. Analizzando i dati relativi alla soddisfazione sulla vita, gli autori individuano correlazioni tra i due tipi di risposta prossime a 1, portando dunque ad accettare l'assunzione. Ovviamente man mano che si selezionano individui più simili a quelli ipotetici si ha una riduzione del campione maggiore.

⁷Le assunzioni su cui si basa sono quattro: 1) le scale di risposta non variano da un'indagine alla successiva; 2) non vi sono errori sistematici nella valutazione; 3) le risposte alla *vignette* e all'autovalutazione usano la stessa scala; 4) la domanda relativa alla misura oggettiva ha un'adeguata descrizione alla sfera d'interesse

Una motivazione che spesso gli autori danno al rifiuto della *response consistency* è l'ordine nel quale vengono presentate le *vignettes*: la valutazione a una *vignette* potrebbe essere influenzata dalla natura della *vignette* precedente (Kapteyn et al., 2011). Questo è un altro motivo per il quale sarebbe opportuno che le *vignettes* fossero mostrate casualmente ai soggetti.

Vignette equivalence

L'assunzione *vignette equivalence* afferma che il vero livello della variabile rappresentata da ogni *vignette* è percepito e interpretato allo stesso modo da tutti gli intervistati; ciò significa che qualsiasi differenza nelle caratteristiche dei rispondenti non influenza la comprensione degli scenari ipotetici.

Si tratta dunque di una forte assunzione poichè potrebbero essere molte le situazioni in cui viene violata, ad esempio a causa di diversi regimi di *welfare* nel caso di confronto fra Paesi diversi ad esempio di salute percepita o di limitazione nello svolgimento di attività lavorative.. Chiaramente anche per questa assunzione l'aspetto fondamentale alla base è il linguaggio e la morfologia usata nel costruire le domande, nonchè i riferimenti al contesto d'interesse. Come già si è detto, altri aspetti come la scelta e l'ordine delle *vignettes* da usare, il genere e il nome dei soggetti ipotetici, ecc sono solo alcuni dei motivi che possono determinare o comunque influire sulla verifica della *vignette equivalence*; riguardo a ciò si son viste delle possibili soluzioni. Bago d'Uva et al. (2011) sono i primi autori a proporre un test sulla validazione di questa condizione, nel caso in cui vi siano almeno due *vignettes*: la loro proposta richiede che non vi sia variazione sistematica nella differenza percepita tra i livelli del costrutto di interesse (nel loro caso la salute) di due qualsiasi *vignettes* e questa rappresenta la condizione necessaria dell'assunzione. Formalmente, il loro test consiste nel rilassare tale costrizione in $J - 1$ *vignettes* e nello specificarne una (ad esempio la $j = 1$) come una funzione indipendente dalle caratteristiche osservate X dell'individuo i ovvero $f(Z_j^*|X) = f(Z_j^*)$; ne consegue:

$$\begin{aligned} Z_{i1}^* &= \theta_1 + u_{ij} \\ Z_{ij}^* &= \theta_j + \varphi_j X_i + u_{ij}, \quad j \neq 1 \end{aligned}$$

dove φ_j è un vettore di parametri. Il modello così specificato risulta essere

identificato; l'identificabilità cadrebbe se anche Z_{i1} fosse funzione delle X_i . In base a questa formulazione, il test di Bago d'Uva et al. (2011) verifica che $\varphi_j = 0, \forall j$. Dunque se statisticamente φ_j è diverso da 0, la *vignette equivalence* risulta essere rifiutata.

In riferimento ad un dominio fisico e a uno mentale della salute (la mobilità e la cognizione) di un campione di individui inglesi di età avanzata, gli autori rigettano questa assunzione.

Peracchi e Rossetti (2013) introducono dei test congiunti, sfruttando il fatto che sotto l'assunzione di *response consistency* e di *vignette equivalence* il modello statistico risulta esser sovraidentificato (si tratta di un modello strettamente legato al Chopit, ma non esattamente lo stesso); in particolare i risultati delle simulazioni di Monte Carlo mostrano che il loro test ha buone proprietà in campioni finiti. Applicando il loro test a diverse sfere del dominio della salute, le restrizioni di sovraidentificazione sono sempre rifiutate; ciò potrebbe però presupporre ulteriori problemi, come ad esempio l'esclusione di alcune variabili importanti.

Angelini et al. (2014) conducono diversi test per testare l'assunzione; prima di tutto verificano che gli intervistati concordino con l'ordinamento globale delle *vignettes* ottenuto dalla media delle risposte categoriale a ciascun individuo ipotetico. Gli autori confrontano questo ordinamento con quello relativo a ciascun Paese preso in analisi nella loro indagine relativa alla soddisfazione sulla vita di ultracinquantenni. Questa verifica viene ripetuta anche dopo aver raggruppato i Paesi in base al loro orientamento religioso (protestanti, cattolici, ex-comunisti) e dopo aver stratificato il campione secondo ciascuna variabile esplicativa osservata. In tutti questi tre casi le conclusioni a cui giungono sono a favore dell'approccio delle *anchoring vignettes*. Un'ulteriore conferma viene riscontrata dalle stime del modello Chopit, separatamente per ognuna delle tre zone "religiose" precedentemente individuate; infatti le stime dei tre modelli risultano essere simili a quelle del campione originario, mostrando dunque che non sono influenzate dalla stratificazione dei Paesi.

2.2 Modello probit ordinato bivariato

In letteratura vi sono diverse estensioni del modello probit ordinato; una di queste segue la logica del modello probit bivariato e dunque associa le tecniche alla base dei modelli per dati ordinali (l'uso di una variabile latente, inosservate) a quelle dei modelli ad equazioni multiple (l'analisi di s variabili d'interesse). Nel caso specifico di due variabili ordinali si parla di probit (o logit) ordinato bivariato, un modello che è stato trovato statisticamente più efficiente e non distorto se (i) i termini d'errore si distribuiscono come una normale bivariata, (ii) se il valore assoluto del coefficiente endogeno ρ è alto e infine (iii) se l'ampiezza del campione è piccola (Sajaia, 2008). Tra i suoi vantaggi, vi è quello di tener conto della correlazione tra le due variabili d'interesse.

Si supponga che ogni intervistato i debba rispondere a due domande di un questionario, $i = 1, 2, \dots, N$; chiaramente ogni intervistato ha un livello reale μ_i per ciascuna domanda d'interesse che si assume essere una combinazione lineare di alcune covariate osservate X_i (ad esempio p) le quali rappresentano le caratteristiche individuali. Questo livello viene percepito dagli intervistati con un errore ε_i di distribuzione normale, specifico per ogni domanda; pertanto ciascun $\mu_i = \mu(X_i)$ si distribuisce come un probit ordinato. Y_{i1}^* e Y_{i2}^* definiscono i livelli percepiti inosservati dell'intervistato i rispettivamente alla prima e seconda autovalutazione; tali livelli sono continui e vengono modellati secondo una normale con media e varianze opportune. Nello specifico dunque si ha (Sajaia (2008), Greene e Hensher (2010b, Cap. 10)):

$$\begin{aligned} Y_{i1}^* &= \mu_{i1} + \varepsilon_{i1} = X_i \beta_1 + \varepsilon_{i1} \\ Y_{i2}^* &= \mu_{i2} + \varepsilon_{i2} = X_i \beta_2 + \varepsilon_{i2} \\ \text{con } \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{bmatrix} &\sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \end{aligned}$$

ove β_1 e β_2 sono due vettori di parametri e $\begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{bmatrix} \perp X_i$. I termini d'errore ε_{i1} e ε_{i2} possono essere dovuti a eterogeneità e/o a un errore nell'aver riportato una risposta; la scelta di porre la media di tale errore pari a 0 è necessaria

per poter garantire l'identificazione del modello, secondo la logica del modello probit ordinato univariato.

Nel momento in cui un soggetto i deve valutare un'autovalutazione scegliendo una delle k (o j) categorie ordinali, non potendo rispondere con una scala continua, egli trasforma ciascun livello percepito continuo Y_{i1}^* e Y_{i2}^* nella categoria riportata Y_{i1} e Y_{i2} attraverso il seguente meccanismo:

$$\begin{aligned} Y_{i1} &= k \quad \text{se} \quad \tau_{i1}^{k-1} \leq Y_{i1}^* < \tau_{i1}^k \\ Y_{i2} &= j \quad \text{se} \quad \tau_{i2}^{j-1} \leq Y_{i2}^* < \tau_{i2}^j \\ \text{dove} \quad &-\infty = \tau_{i1}^0 < \tau_{i1}^1 < \dots < \tau_{i1}^{K-1} < \tau_{i1}^K = +\infty \\ &-\infty = \tau_{i2}^0 < \tau_{i2}^1 < \dots < \tau_{i2}^{J-1} < \tau_{i2}^J = +\infty \end{aligned}$$

con τ_{i1} e τ_{i2} vettori delle soglie rispettivamente per la prima e la seconda domanda, $k = 1, \dots, K$ e $j = 1, \dots, J$. In particolare il numero di modalità K non deve essere necessariamente uguale a quello J . Pertanto l'individuo i sceglie la categoria k se e solo se il valore della variabile latente Y_{i1}^* è compreso tra τ_{i1}^{k-1} e τ_{i1}^k , mentre sceglie il livello j se e solo se il valore della variabile latente Y_{i2}^* è compreso tra τ_{i2}^{j-1} e τ_{i2}^j . Inoltre il modello richiede che queste soglie sian le stesse per tutti i soggetti del campione, indipendentemente dalle loro caratteristiche.

Il metodo standard usato per la stima di tale modello è quello della massima verosimiglianza. Nello specifico, sotto l'assunzione di indipendenza delle osservazioni, la funzione di verosimiglianza è data da

$$L(\theta|Y) = \prod_{i=1}^N \prod_{k=1}^K \prod_{j=1}^J P(Y_{i1} = k, Y_{i2} = j | X_i)^{I(Y_{i1}=k, Y_{i2}=j)}$$

dove θ corrisponde al vettore dei parametri $\tau_1 = (\tau_{11}, \tau_{21}, \dots, \tau_{K1})'$, $\tau_2 = (\tau_{12}, \tau_{22}, \dots, \tau_{J2})'$, $\beta_1 = (\beta_{11}, \beta_{21})'$, $\beta_2 = (\beta_{12}, \beta_{22})'$ e ρ , e

$$I(Y_{i1} = k, Y_{i2} = j) = \begin{cases} 1 & \text{se } y_{i1} = k \text{ e } y_{i2} = j \\ 0 & \text{altrimenti} \end{cases} \text{ è la funzione indicatrice.}$$

Conseguentemente la funzione di log-verosimiglianza è data da

$$l(\theta|Y) = \sum_{i=1}^N \sum_{k=1}^K \sum_{j=1}^J I(Y_{i1} = k, Y_{i2} = j) \ln(P(Y_{i1} = k, Y_{i2} = j | X_i))$$

dove la probabilità congiunta è data da

$$\begin{aligned}
P(Y_{i1} = k, Y_{i2} = j | X_i) &= \\
&= P(\tau_{k-1,1} < Y_{i1}^* \leq \tau_{k,1}, \tau_{j-1,2} < Y_{i2}^* \leq \tau_{j,2}) \\
&= P(\tau_{k-1,1} < x_i' \beta_1 + \varepsilon_{i1} \leq \tau_{k,1}, \tau_{j-1,2} < x_i' \beta_2 + \varepsilon_{i2} \leq \tau_{j,2}) \\
&= P(\tau_{k-1,1} - x_i' \beta_1 < \varepsilon_{i1} \leq \tau_{k,1} - x_i' \beta_1, \tau_{j-1,2} - x_i' \beta_2 < \varepsilon_{i2} \leq \tau_{j,2} - x_i' \beta_2) \\
&= \int_{\tau_{k-1,1} - x_i' \beta_1}^{\tau_{k,1} - x_i' \beta_1} \int_{\tau_{j-1,2} - x_i' \beta_2}^{\tau_{j,2} - x_i' \beta_2} \phi_2(\varepsilon_{i1}, \varepsilon_{i2}, \rho) \partial \varepsilon_{i1} \partial \varepsilon_{i2}
\end{aligned}$$

dove $\phi_2(\cdot)$ è la funzione di densità di una distribuzione normale bivariata di media 0, varianza 1 e correlazione ρ . Attraverso la massimizzazione di questa log-verosimiglianza è dunque possibile ottenere la stima dei coefficienti θ .

Il coefficiente ρ rappresenta una sorta di indice di bontà del modello: infatti un valore statisticamente significativo e differente da 0 indica l'esistenza di un'associazione tra gli errori ε_{i1} e ε_{i2} e dunque tra le variabili Y_{i1} e Y_{i2} . La significatività di tale coefficiente è generalmente verificata attraverso un test di verosimiglianza.

Chiaramente in letteratura esistono ulteriori estensioni di tale modello, alcune che considerano più di due variabili d'interesse, altre che specificano la distribuzione degli errori secondo una forma non parametrica. Per alcuni ulteriori si veda Greene e Hensher (2010b).

2.3 Modello di selezione del campione per variabili ordinali

Un particolare tipo di modello probit è il modello di selezione, utilizzato in letteratura per cercare di risolvere problemi di selezione campionaria: possono esistere dei casi in cui le osservazioni sulla variabile di interesse provengono esclusivamente da un sottoinsieme del suo spazio campionario; in queste situazioni si è di fronte ad una non completa osservabilità del fenomeno, che, per l'appunto, determina una selezione del campione. In generale, tale fenomeno si manifesta qualora si fosse in presenza di un modello per

la popolazione di interesse e si volesse fare inferenza sui parametri del medesimo, ma il processo di campionamento è tale per cui si osserva solo un sottocampione, anzichè lo spazio campionario completo.

Tale modello deve la sua origine a Heckman e al contesto econometrico e per tale motivo è noto come modello di Heckman o tobit II. L'idea sottostante è la capacità di stimare simultaneamente due equazioni: la prima, detta equazione di selezione, implica una regressione probit; la seconda equazione, detta equazione principale, invece è stimata col metodo dei minimi quadrati (OLS). Nel caso più semplice anche la seconda equazione stima la probabilità di una variabile dicotomica differente da quella dell'equazione di selezione, ma legata concettualmente ad essa. L'unica richiesta fondamentale che viene fatta è relativa alla specificazione delle variabili usate nelle due equazioni: l'equazione di selezione deve infatti prevedere almeno una variabile non presente nel modello principale e se possibile non legata concettualmente alla variabile d'interesse dell'equazione principale; ciò è preferibile in quanto questa seconda equazione rappresenta e determina la selezione vera e propria definita nel primo stadio. Se infatti l'insieme delle variabili fosse lo stesso, ci si imbatterebbe nella cosiddetta "identificazione debole": questo metodo di stima consente dunque un'adeguata correzione della distorsione delle stime solo se le variabili esplicative impiegate nel modello di selezione non costituiscono lo stesso insieme di quelle impiegate nel modello strutturale.

In particolare ciò che il modello di Heckman dimostra è che se si stimasse la variabile d'interesse con un semplice modello ad esso adeguato, senza dunque considerare l'equazione di selezione, non si terrebbe in considerazione l'endogeneità tra le due variabili che tale modello di selezione mette in relazione. Conseguentemente, se si ignorasse il problema di selezione, si otterrebbero stimatori distorti e inconsistenti per tutti i parametri del modello (Greene e Hensher, 2010b, Cap. 11.2).

Formalmente, il modello di selezione è costituito da un sistema di equazioni per due variabili risposta, dove l'osservazione della seconda variabile avviene se e solo se è possibile osservare la prima. Il sistema è specificato attraverso due variabili latenti S_i^* e Y_i^* , dove la prima è specifica per la selezione dell'individuo i , mentre la seconda costituisce il livello percepito inosservato di questo stesso soggetto, se, ad esempio, si suppone dover analizzare l'au-

tovalutazione fornita in un questionario. Tali variabili sono assunte continue e vengono modellate secondo una normale di media e varianze opportune. Nello specifico si ha (De Luca e Perotti (2011)):

$$\begin{aligned} S_i^* &= X_{i1}\beta_1 + v_i \\ Y_i^* &= X_{i2}\beta_2 + \varepsilon_i \\ \text{con } \begin{bmatrix} v_i \\ \varepsilon_i \end{bmatrix} &\sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \end{aligned}$$

ove $i = 1, \dots, N$ indica l'intervistato i , β_1 e β_2 sono due vettori di parametri, X_{i1} e X_{i2} sono due insiemi di variabili esplicative differenti e $[v_i, \varepsilon_i]$ è il vettore contenente i termini d'errori di ciascuna equazione tale che $\begin{bmatrix} v_i \\ \varepsilon_i \end{bmatrix} \perp X_i$.

La correlazione dei termini d'errore v_i e ε_i considera dunque l'endogeneità che potrebbe esistere tra la variabili d'interesse dell'equazione di selezione e quella dell'equazione principale.

Le due variabili S_i^* e Y_i^* sono inosservabili e associate alle due variabili d'interesse S_i e Y_i , dove nel caso in esame si suppone che la variabile dipendente dell'equazione principale abbia una natura categoriale ordinale. Pertanto le due variabili latenti si manifestano solo nel caso in cui:

$$S_i = \begin{cases} 1 & \text{se } S_i^* > 0 \\ 0 & \text{altrimenti} \end{cases} \quad Y_i = \begin{cases} k & \text{se } Y_i^* > 0 \\ 0 & \text{altrimenti} \end{cases}$$

In particolare, la variabile Y_i^* è osservata solo se si riscontra la condizione $S_i = 1$; pertanto il sottocampione identificato da tale restrizione rappresenta il campione di selezione. In particolare gli effetti di selezione si manifestano attraverso la correlazione tra i due errori delle variabili latenti. Inoltre, essendo Y_i una variabile ordinale, essa assumerà uno dei K valori secondo il seguente meccanismo:

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

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

con τ_i vettore delle soglie e $k = 1, 2, \dots, K$ l'indice associato al numero di categorie della variabili Y_i .

In questo caso dunque l'equazione di selezione è un modello probit standard univariato (con intercetta diversa da 0), mentre l'equazione principale è un modello probit ordinato univariato. La relazione tra i due modelli viene dunque considerata attraverso la correlazione dei loro errori. Perciò, è ragionevole che l'identificazione dei parametri del modello richieda le seguenti due restrizioni: la prima, impone che l'intercetta dell'equazione principale sia vincolata a 0 secondo la logica e assunzione standard dei modelli probit (logit) ordinali univariati; la seconda condizione assume che l'insieme delle covariate X_{i1} si distingua da quello di X_{i2} per almeno una variabile. In particolare, se si utilizza l'identificazione della specificazione semiparametrica, si richiede una terza assunzione ovvero che X_{i1} e X_{i2} contengano almeno una variabile continua (De Luca e Perotti, 2011).

Sotto l'assunzione della specificazione parametrica e dunque assumendo che il vettore degli errori $[v_i, \varepsilon_i]$ segua una distribuzione gaussiana bivariata di media 0, varianza unitaria e correlazione ρ e sotto l'assunzione che le osservazioni siano tra loro indipendenti, la funzione di log-verosimiglianza per un campione causale di N osservazioni è data da:

$$l(\theta|Y) = \sum_{i=1}^N \left\{ (1 - S_i) \ln \pi_{0i}(\theta) + \sum_{k=1}^K Y_i I(Y_i = k) \ln \pi_{1ki}(\theta) \right\}$$

dove θ corrisponde al vettore dei parametri $\tau = (\tau_1, \tau_2, \dots, \tau_K)'$, $\beta_1 = (\beta_{11}, \beta_{21})'$, $\beta_2 = (\beta_{12}, \beta_{22})'$ e ρ , mentre $I(Y_i = k) = \begin{cases} 1 & \text{se } y_{i1} = k \text{ e } y_{i2} = j \\ 0 & \text{altrimenti} \end{cases}$ è la funzione indicatrice e $\pi = (\pi_0, \pi_{11}, \pi_{12}, \dots, \pi_{1K})$ sono le probabilità condizionate associate con $K + 2$ possibili realizzazioni di S e Y . Più precisamente:

$$\pi_0(\theta) = Pr(S = 0) = 1 - \phi(\beta_1' X_1)$$

$$\pi_{1k}(\theta) = Pr(S = 1, Y = k) =$$

$$= \phi_2(\beta_1' X_1, \tau_{k+1} - \beta_2' X_2; -\rho) - \phi_2(\beta_1' X_1, \tau_k - \beta_2' X_2; -\rho)$$

con ϕ che denota la distribuzione normale standardizzata e ϕ_2 che corrisponde alla distribuzione gaussiana bivariata di medie 0, varianze unitaria e correlazione ρ . Conseguentemente le stime del vettore θ possono essere ottenute

dalla massimizzazione della log-verosimiglianza così specificata.

La validità del modello e dunque la verifica della selezione del campione può essere verificata attraverso il coefficiente ρ : se è nullo o è statisticamente non significativo (secondo il test rapporto di verosimiglianza), è plausibile l'ipotesi che gli individui siano selezionati casualmente e che stime consistenti di β_2 siano ottenute stimando l'equazione di regressione tramite un modello probit ordinale standard; dunque in questa situazione l'ipotesi di selezione da campione viene a mancare.

Anche per questo modelli esistono ulteriori specificazioni come quella di Miranda e Rabe-Hesketh (2005), i quali usano una specificazione differente degli errori delle due equazioni.

2.4 Alcune misure ordinali di associazione

La correlazione è un indice di associazione tra due covariate X e Y e viene espressa con un numero che oscilla fra -1 e 1; il segno del valore della correlazione indica la direzione dell'associazione, mentre il valore numerico assoluto rappresenta l'intensità: zero indica la mancanza di legame, mentre una quantità più prossima a 1 corrisponde ad una maggiore associazione e quindi una sovrapposizione più prossima e perfetta. Qualora le due variabili abbiano una natura ordinale, la correlazione di Pearson non è appropriata per calcolare la loro associazione in quanto non tiene conto del loro ordinamento, ma le considera (erroneamente) come semplici variabili quantitative (se ad ogni livello della variabile ordinale è associato un valore numerico). In letteratura sono presenti alcune misure per soccombere a tale problema, tra le quali Gamma, Tau-b, D di Somers e correlazione policorica.

L'indice di associazione Gamma ha un'affinità con la correlazione in quanto tratta le due variabili simmetricamente ovvero non è necessario identificare una variabile risposta tra esse; inoltre ha un range compreso tra -1 e 1 e in particolare il valore 0 implica l'indipendenza tra le due variabili. A differenza della correlazione, un valore di Gamma prossimo a 1 in valore assoluto non indica una perfetta relazione lineare, bensì una relazione monotoma. In

generale, poichè le modalità delle due covariate sono ordinabili, possono sussistere fra loro due tipi diversi di relazione: relazione diretta (chiamata anche concordanza) e relazione indiretta (chiamata anche discordanza). Nello specifico la prima si manifesta qualora si sia in presenza di una relazione diretta tra due variabili X e Y ovvero quando a modalità basse di X corrispondono più frequentemente modalità basse di Y , mentre a modalità alte di X corrispondono più frequentemente modalità alte di Y . Vi è invece discordanza (ovvero si è in presenza di una relazione inversa) quando a modalità basse di X corrispondono più frequentemente modalità alte di Y , mentre a modalità alte di X corrispondono più frequentemente modalità basse di Y .

Formalmente, sia x_1, \dots, x_n e y_1, \dots, y_n rispettivamente le osservazioni della variabile X e Y tali per cui π_{ij} raffiguri la distribuzione congiunta di x_i e y_j , con i indica la modalità i -esima della prima variabile ($i = 1, \dots, n$) e j quella j -esima della seconda ($j = 1, \dots, n$); la misura Gamma è dunque data da (Agresti, 2002, Cap. 2.4.4):

$$\gamma = \frac{N_c - N_d}{N_c + N_d}$$

$$\text{dove} \quad -1 < \gamma < 1$$

N_c =numero totale di coppie concordanti

N_d =numero totale di coppie discordanti

ovvero coincide con la differenza tra la probabilità di concordanza e quella di discordanza.

Una limitazione di tale indice è dovuta alla mancata considerazione delle coppie per cui le modalità di una delle due variabili rimangono costanti. Per questo motivo un indice preferibile è la seguente misura non parametrica di correlazione, nota come *tau-b* di Kendall, la quale mantiene le proprietà di simmetria e il range di Gamma (Agresti (2002, pp.68), Stata (2017)):

$$\tau_b = \frac{N_c - N_d}{\sqrt{(N - U_X)(N - V_Y)}}$$

dove

$$-1 < \tau_b < 1$$

$$U_X = \frac{\sum_{i=1}^{N_1} u_i(u_i - 1)}{2} \quad \text{e} \quad V_X = \frac{\sum_{i=1}^{N_1} v_i(v_i - 1)}{2}$$

con $N = \frac{n(n-1)}{2}$, N_1 è il numero di insiemi di valori di X con uguale modalità, u_i è il numero di valori di X con uguale modalità nell'insieme i -esimo, N_2 è il numero di insiemi di valori di Y con uguale modalità e v_i è il numero di valori di Y con uguale modalità nell'insieme j -esimo.

Un'estensione unidirezionale di Gamma è la D di Somers, la quale quindi non considera le due variabili come simmetriche, bensì identifica una delle due come variabile risposta; essa indica l'eccesso proporzionato di coppie concordanti su quelle discordanti tra coppie non legate sulla variabile dipendente.

Formalmente, se la covariata Y è dipendente da X, l'indice è pari a (Agresti (2002, pp.68), Newson (2002)):

$$d_Y = \frac{N_c - N_d}{(N - V_Y)}$$

viceversa, ovvero se Y è la variabile indipendente, si avrà:

$$d_X = \frac{N_c - N_d}{(N - U_X)}.$$

Un quarto metodo risulta essere la correlazione policorica (Greene e Hensher, 2010b, Cap. 10.3): in sostanza si ipotizza che i livelli associati alle due variabili ordinali derivino da un processo inosservabile ovvero da due variabili latenti continue di distribuzione normale che approssimano il livello reale inosservabile a valori interi. In sostanza, le due variabili inosservate seguono una distribuzione normale bivariata e conseguentemente, la correlazione policorica coincide con la stima (di massima verosimiglianza) della correlazione ricavata dal modello probit ordinale bivariato che ha solo l'intercetta in entrambe le equazioni.

Questi coefficienti vengono spesso usati in un test statistico unilaterale per verificare se le due variabili sono dipendenti sotto l'ipotesi nulla H_0 . In particolare, per campioni grandi è solito considerare un'approssimazione della distribuzione normale con media 0 e opportuna varianza per le prime tre misure di associazione. Per la correlazione policorica invece si fa affidamento al test χ^2 di Pearson o al test del rapporto di verosimiglianza.

Capitolo 3

Applicazioni empiriche

In questo capitolo vengono presentati innanzitutto l'obiettivo e i dati per le applicazioni empiriche svolte in questa tesi. Successivamente vengono riportate le analisi esplorative e descrittive del campione e infine i risultati dei metodi utilizzati per la stima delle variabili (ordinali) d'interesse.

3.1 L'indagine di SHARE

L'indagine sulla salute, età e pensionamento in Europa (*Survey of Health, Ageing and Retirement in Europe* -SHARE-) è un database cross-nazionale e micro-panel che raccoglie informazioni sullo stato socio-economico, familiare e sulle condizioni di vita di un enorme numero di individui provenienti da 27 Paesi Europei più Israele; nello specifico sono essenzialmente tre le dimensioni che vengono investigate in questa indagine: 1) reddito e ricchezza personale, 2) salute fisica e mentale, disabilità e mortalità, 3) network e rapporti sociali. Il principale obiettivo che si prefigge SHARE è la comprensione del processo di invecchiamento e di come esso influenzi i cittadini dei diversi Paesi: generalmente infatti l'età influisce su molti domini della vita di un individuo, come il pensionamento, la salute, il tempo libero, la ricchezza monetaria, finanziaria e reale, ecc, tutti fortemente legati tra loro. In particolare l'Europa gode di una ricchezza e diversità di cultura, storia e politica molto più ampia di altri continenti e ciò coinvolge particolarmente l'interesse degli economisti e non solo.

L'approccio longitudinale di SHARE rappresenta un punto di forza per i ricercatori, in quanto consente di confrontare ed aggiornare le informazioni relative a diversi anni. La prima indagine SHARE è stata realizzata nel 2004 e attualmente sono state svolte 6 rilevazioni; la settima è iniziata quest'anno e si è appena conclusa. Nello specifico, le informazioni sono state raccolte nei diversi anni attraverso la tecnica CAPI (*Computer Assisted Personal Interview*): viene effettuata un'intervista faccia a faccia da un rilevatore presso il domicilio (regolare) della famiglia attraverso l'ausilio di un *personal computer* che gestisce il questionario elettronico (tradotto nella lingua, o in alcuni casi del Paese del cittadino intervistato). Il rispondente è dunque un componente di questa famiglia, anche se vengono raccolte informazioni sul partner, coniuge o figli presenti e non al momento dell'indagine. Le persone vengono escluse da una rilevazione se sono incarcerate, ricoverate in ospedale o si trovano fuori dal Paese durante l'intero periodo dell'indagine, incapaci di parlare la lingua del Paese in cui risiedono momentaneamente o si sono spostati ad un domicilio ignoto.

Uno dei vantaggi della tecnica CAPI è la possibilità di avere un rapporto diretto con i rispondenti, raccogliendo informazioni difficilmente ottenibili oggettivamente con un questionario compilato a penna: durante lo svolgimento dell'indagine, gli intervistati vengono sottoposti anche a dei test cognitivi e fisici e cognitivi, come ad esempio la misura della forza di presa delle mani. Oltre al questionario elettronico CAPI vengono somministrati due questionari cartacei di autocompilazione. Uno di questi, specifico per gli intervistati di età inferiore ai 65 anni (*vignette* di tipo B) e quelli di età superiore a questa (*vignette* di tipo C), è relativo da un'insieme di domande su diversi domini della soddisfazione (vita, salute, reddito, disabilità lavorativa, ecc) e comprende anche alcune *vignettes* ad essi relative. Nello specifico, tale questionario viene sottoposto, casualmente, solo ad un sottogruppo degli intervistati, riducendo così i costi elevati delle *vignettes*, ma determinando un maggior numero di valori mancanti. Inoltre alcune domande non sono presenti in entrambi i questionari e, soprattutto, non vengono somministrate secondo quanto è consigliato da Hopkins e King (2010): ai soggetti viene prima chiesto di valutare se stessi su una scala che misura uno specifico dominio e successivamente viene chiesto di esprimere il proprio giudizio sulla

stessa scala con la quale hanno classificato se stessi; inoltre le valutazioni dei soggetti ipotetici non sono sottoposte casualmente, ma generalmente secondo un ordine (crescente o decrescente) della scala del dominio d'interesse.

3.2 Descrizione del dataset

Il campione di dati preso in considerazione e analizzato in questo lavoro riguarda la popolazione ultracinquantenne europea della seconda rilevazione SHARE, svolta nel biennio 2006-2007 in 11 Paesi Europei e relativa a 22 moduli. Nello specifico, dato l'obiettivo prefissato (l'analisi dell'associazione tra la soddisfazione della vita e quella del reddito) e in riferimento alla possibilità che gli individui possano usare scale di risposta non eguali al momento della valutazione, si è ristretto il campo di indagine ai soli individui che hanno risposto ad entrambe le domande dei due domini d'interesse e ad almeno una delle due relative *vignettes*. Sulla base dell'enorme mole di dati e di moduli a disposizione, sono state selezionate le covariate soggettive ed oggettive (o ricodificate in una scala europea) che potessero avere una qualche relazione con quelle oggetto d'indagine, anche alla luce dei domini del benessere a cui appartengono e ai risultati emersi ed evidenziati da altri autori (si vedano le Sezioni 1.2 e 1.3). Il campione così ottenuto è composto da 7665 soggetti, di cui lo 0.09% non ha fornito una valutazione ad entrambe le domande d'interesse (70 persone) e un'altrettanta percentuale di individui non valuta la propria soddisfazione su uno dei due domini. Tra questi, il 2.27% non fornisce una risposta ad almeno una delle due *vignettes* per un determinato dominio; conseguentemente il numero di intervistati considerati nelle successive analisi è pari a 7353 ovvero il 4.07% in meno di quello di partenza.

Il questionario compilato a mano e distinto per l'età degli intervistati contiene le informazioni relative alla soddisfazione della vita e del reddito e agli appropriati scenari ipotetici; in particolare queste sono precedute da una breve introduzione che specifica il loro dominio di riferimento¹. La misura del benessere soggettivo è ottenuta dalla seguente domanda:

¹La breve introduzione è la seguente: "Vorremmo porLe adesso alcune domande su quanto soddisfatto Lei si sente su diversi aspetti della Sua vita."

"Considerando la vita in generale, lei si sente ..."

mentre quella del reddito soggettivo da questa:

"Considerando il reddito totale della sua famiglia, lei si sente ..."

Nello specifico, la prima segue la seconda di due domande (si precisa che nel questionario B queste domande sono precedute da altre quattro domande che non sono presenti invece nel questionario C, un aspetto di cui si terrà conto nelle successive analisi). Successivamente alla singola autovalutazione, viene infatti chiesto agli intervistati di dare il proprio giudizio in riferimento a degli individui ipotetici; anche queste sono anticipate da una breve introduzione:

"Adesso Le descriveremo alcuni tratti della vita di altre persone. Vorremmo sapere come Lei, secondo i Suoi gusti, valuterebbe certi aspetti della vita di queste persone. Nell'esprimere il Suo giudizio, immagini che queste persone abbiano la Sua stessa età e la Sua stessa formazione (a meno che non sia indicato diversamente). Se non è sicuro/a della risposta, per favore indichi quella che Le sembra più plausibile."

Dunque, il soggetto viene informato indirettamente sulla funzione delle *vignettes* poiché gli/le viene chiesto di valutare individui ipotetici secondo proprie caratteristiche; perciò, le persone dovrebbero rispondere alle stesse domande sostituendo il "tu" al nome dell'individuo ipotetico descritto nello scenario. In particolare vengono sottoposte le seguenti *anchoring vignettes* di cui le prime due per la soddisfazione del reddito totale, le altre per quella della vita in generale:

"[Giorgio] è sposato e ha due figli; il reddito totale della sua famiglia al netto delle tasse è di 1450€ al mese"

"[Anna] è sposata e ha due figli; il reddito totale della sua famiglia al netto delle tasse è di 2900€ al mese. "

"[Carlo] ha 63 anni. Sua moglie è morta 2 anni fa e lui trascorre ancora molto tempo pensando a lei. Ha quattro figli e 10 nipoti che lo visitano regolarmente. Carlo riesce a far quadrare i conti, ma non ha soldi per gli "extra", come regali costosi per i nipoti. Recentemente ha dovuto smettere di lavorare

a causa di problemi di cuore. Si sente stanco facilmente. Per il resto, non ha altri gravi problemi di salute."

"[Franca] ha 72 anni ed è vedova. Il suo reddito totale al netto delle tasse è di circa 950€ al mese. È proprietaria della casa in cui vive e ha una larga cerchia di amici. Gioca a carte due volte alla settimana e va regolarmente in vacanza con degli amici. Ultimamente soffre di artrite, che le rende doloroso lavorare in casa e in giardino."

Chiaramente i nomi delle persone ipotetiche e le cifre monetarie rispettano quanto si è detto nella Sezione 2.1.2; nello specifico i primi sono stati opportunamente tradotti per ogni determinato Paese, mentre i secondi sono stati convertiti in Euro e aggiustati per differenze in potere d'acquisto (usando i tassi di cambio a parità di potere d'acquisto dell'OECD): ad esempio nelle *vignettes* francesi sulla soddisfazione economica si trovano i personaggi ipotetici Jim e Anne i quali hanno rispettivamente un reddito totale pari a 1500€ e 3000€².

Infine si precisa che sia le domande sull'autovalutazione dia quelle sulle *vignettes* hanno le stesse 5 categorie di risposta, dalla più soddisfatta alla meno: "Molto soddisfatto", "Soddisfatto", "Né soddisfatto né insoddisfatto", "Insoddisfatto", "Molto insoddisfatto" (per facilitare i confronti e le interpretazioni delle analisi che si sono svolte tale ordine è stato opportunamente invertito).

Le variabili coinvolte nelle analisi (oltre al Paese di residenza) sono le seguenti:

- caratteristiche demografiche: genere, età, istruzione, relazioni;
- caratteristiche familiari: numero di componenti della famiglia dell'individuo, numero di figli ancora vivi, il soggetto ha almeno un genitore ancora vivo;
- stato di salute fisico, mentale e cognitivo: numero di malattie croniche, di sintomi, di problemi di mobilità, indicatori ADL, IADL depressione,

²Le *vignettes* presentate nelle righe precedenti sono specifiche per il questionario italiano

risultati del *fluency test*³, test di forza di presa delle mani⁴;

- caratteristiche economiche: stato occupazionale, ricchezza finanziaria, ricchezza reale, reddito familiare;
- caratteristiche relative all'indagine CAPI: l'individuo chiede almeno una chiarificazione, tipo di questionario relativo alle *vignettes* fornito all'intervistato.

A riguardo delle variabili monetarie, si sottolinea il fatto che si possono distinguere tre tipi di ricchezza e che questi possono essere dei fattori economici particolarmente influenti sul benessere generale. Queste tre componenti della ricchezza totale sono: la ricchezza finanziaria, la ricchezza reale e la ricchezza umana. La prima coincide con il valore totale di mercato di alcuni assets finanziari sicuri e rischiosi come stock, fondi, derivati, bond, conto bancario e postale, polizze assicurative sulla vita, ecc. La seconda misura è stata ottenuta come la somma del quantitativo monetario corrispondente ai beni reali ovvero le terre, le abitazioni, i veicoli o ciò che è ricavato dal proprio business o da altre proprietà. Entrambe questi tipi di ricchezze sono calcolate al netto dei debiti o mutui. Infine la ricchezza umana in questo lavoro viene intesa come i salari e guadagni che percepisce un individuo e dipende dal tipo di occupazione; solitamente viene rappresentata dal reddito e dunque, specialmente per i più poveri, è la componente più importante della ricchezza totale (ulteriore motivo che spinge l'interesse di molti intellettuali).

Nella Tabella 3.1 si riporta un resoconto delle variabili appena elencate assieme alle loro statistiche riassuntive, in base alla suddivisione qui sopra considerata. In particolare, la misura riferita al reddito è relativa al reddito familiare netto mensile (quindi di tutti gli individui che appartengono al nucleo familiare dell'intervistato); questa, come tutte le altre variabili monetarie sono state corrette a parità di potere di acquisto (si veda l'Appendice A

³Agli intervistato viene chiesto di elencare il nome di animali, reali o mitologici, che conoscono; la variabile *fluency test* indica quindi il numero di animali (non ripetuti) elencati correttamente durante il test

⁴La variabile corrisponde al risultato massimo di quattro misurazioni a cui vengono sottoposti gli intervistati e relative alla forza delle mani (due per ogni mano)

per la formula).

In riferimento alla misura del livello di istruzione e a quella della depressione, al fine di uniformarle in tutti gli Stati in esame, sono state usate, per la prima, la codifica ISCED, ovvero uno standard internazionale creato dall'UNESCO per classificare i corsi di studio e i relativi titoli, e per la seconda la scala EURO-D, ovvero uno standard internazionale sviluppato dall'EURODEP *Concerted Action Programme* per classificare i sintomi della depressione. I livelli di ciascuna variabile possono essere consultati nell'Appendice A, dove si trovano anche le specificazioni degli indicatori ADL e IADL.

Tabella 3.1: Descrizione e statistiche riassuntive delle variabili esplicative considerate

<i>Variabile</i>	<i>Descrizione</i>	<i>Media</i>	<i>Deviazione Standard</i>
<u>Demografiche</u>			
Donna	Dummy=1 se la persona è donna	0.552	0.497
Età	Età intervistato nel 2007 (min=34, max=97)	64.377	9.869
Età_med	$(Et\grave{a}-63)/10$ (min=-2.9, max=3.4)	0.138	0.987
Età_med ²	$[(Et\grave{a}-63)/10]^2$	0.993	1.365
<u>Educazione</u>			
Istruzione bassa	Dummy=1 se il codice ISCED è al massimo 1 (baseline)		
Istruzione media	Dummy=1 se il codice ISCED è 2 o 3	0.494	0.500
Istruzione elevata	Dummy=1 se il codice ISCED è almeno 4	0.235	0.424
<u>Relazioni</u>			
Single	dummy=1 se l'individuo è single (baseline)		
Partner	Dummy=1 se l'individuo è sposato o ha un compagno/a con cui vive	0.778	0.416
<u>Familiari</u>			
N componenti	Numero componenti familiari al momento dell'intervista (min=1, max=10)	2.279	1.067
N figli	Numero figli ancora in vita, adottati o naturali (min=0, max=12)	2.138	1.321
Genitore	Dummy=1 se l'individuo ha almeno un genitore ancora vivo	0.259	0.438
<u>Stato Occupazionale</u>			
Pensionato	Dummy=1 se l'individuo è in pensione	0.499	0.500
Occupato	Dummy=1 se l'individuo è impiegato o imprenditore	0.300	0.458
Non Lavora	Dummy=1 se l'individuo è casalinga, disabile o disoccupato (baseline)		
<u>Stato Salute</u>			
Depressione	Dummy=1 se l'individuo riporta almeno un sintomo di depressione giorni (min=0, max=6)	0.749	0.433

ADL	Numero di limitazioni con le attività di tutti i giorni (min=0, max=6)	0.172	0.677
IADL	Numero di limitazioni con attività strumentali di tutti i giorni (min=0, max=7)	0.272	0.863
Croniche	Numero malattie croniche (min=0, max=10)	1.597	1.508
Mobilità	Numero di limitazioni nella mobilità degli arti o motoria (min=0, max=10)	1.418	2.147
Sintomi	Numero di sintomi di malattia (min=0, max=11)	1.748	1.862
<i>Grip strength</i>	Misura della forza della mano (min=3, max=84)	35.111	11.807
<i>Fluency Test</i>	Numero di animali elencati durante il test verbale (min=0, max=65)	19.355	7.475
<u>Paesi</u>			
Germania	Dummy=1 se l'individuo vive in Germania (baseline)		
Svezia	Dummy=1 se l'individuo vive in Svezia	0.062	0.241
Olanda	Dummy=1 se l'individuo vive in Olanda	0.069	0.253
Spagna	Dummy=1 se l'individuo vive in Spagna	0.068	0.252
Italia	Dummy=1 se l'individuo vive in Italia	0.092	0.289
Francia	Dummy=1 se l'individuo vive in Francia	0.131	0.337
Danimarca	Dummy=1 se l'individuo vive in Danimarca	0.050	0.218
Grecia	Dummy=1 se l'individuo vive in Grecia	0.074	0.262
Belgio	Dummy=1 se l'individuo vive in Belgio	0.113	0.317
Repubblica Ceca	Dummy=1 se l'individuo vive in Repubblica Ceca	0.116	0.321
Polonia	Dummy=1 se l'individuo vive in Polonia	0.073	0.261
<u>Ricchezza totale</u>			
Ricchezza reale	Ricchezza reale (in Euro)	190110.1	253671.8
Ricchezza finanziaria	Ricchezza finanziaria (in Euro)	41226.29	84295.56
Reddito familiare	Reddito annuale netto della famiglia (in Euro)	26915.47	44644.37
<u>Indagine CAPI</u>			
Chiarimenti	Dummy=1 se l'individuo chiede almeno una chiarificazione durante l'intervista	0.600	0.490
Tipo_b	Dummy=1 se l'individuo risponde al questionario di tipo B	0.600	0.490

3.3 Analisi descrittive

3.3.1 Analisi sulle variabili esplicative

Il campione analizzato comprende 7353 rispondenti europei, di cui il 55.23% sono donne. L'età media di questi individui è di 64 anni e la mediana di 63; i valori di questi due indicatori sono molto vicini, ma non coincidenti

cosicché la distribuzione è leggermente asimmetrica a destra. Nello specifico si parla di asimmetria positiva e ciò è confermato dall'indice Skewness, che è positivo e pari a 0.438. Il più giovane intervistato è una donna di 34 anni, mentre i più anziani, o meglio anziane, ne hanno 97. Non deve stupire che l'individuo più piccolo abbia un'età inferiore ai 50; questo infatti è dovuto al fatto che si intervistano famiglie in cui almeno uno dei componenti del nucleo familiare sia come minimo cinquantenne. Altre informazioni relative ai percentili sono riportate nella Tabella 3.2; in particolare si specifica che questa variabile è stata creata in riferimento al secondo e ultimo anno dell'indagine SHARE che si sta considerando, ovvero il 2007, e che nella stima dei modelli verrà considerata una sua trasformata.

Tabella 3.2: Distribuzione dell'età dei rispondenti

	Percentili	Più piccoli		
1%	46	34		
5%	51	35		
10%	53	36	Osservazioni	7353
25%	57	37		
			Media	64.377
50%	63		Deviazione	9.869
		Più grandi	Standard	
75%	71	96		
90%	79	96	Varianza	97.399
95%	82	97	Skewness	0.438
99%	88	97	Curtosi	2.642

Una delle relazioni che spesso si riscontra in letteratura è quella relativa allo stato lavorativo e all'istruzione. Per omogeneizzare la frequenza delle categorie relative allo status occupazionale, non è stato utilizzata la variabile originale, bensì se ne è creata una avente tre categorie: le prime due corrispondono a quelle rilevate dall'indagine, mentre la terza comprende tutti gli individui che al momento dell'intervista non lavoravano perchè disoccupati (3.35%), disabili o malati permanenti (3.94%), casalinghe (11.7%) o altro (0.97%). In questo modo circa metà del campione è in pensione, il 29.91% è occupato (impiegato o imprenditore) e il restante 20.4% non ha un lavoro. Per quanto riguarda l'istruzione, il numero di soggetti che non ha alcun grado

di istruzione o che ha un dottorato di ricerca è quasi nullo (rispettivamente il 3.17% e lo 0.33% del totale): come si può vedere in Figura 3.1 i livelli 1, 3 e 5 sono equamente e maggiormente frequenti tra i pensionati e i non lavoratori, mentre tra gli occupati, come ci si può aspettare, vi è una prevalenza di laureati o soggetti con un'istruzione di grado superiore. In particolare essendo molte le classi, alcune di dimensioni limitate, per il proseguo delle analisi si è deciso di rendere più agevole l'interpretazione del livello di istruzione generando una nuova variabile, così come specificato nella Tabella 3.1.

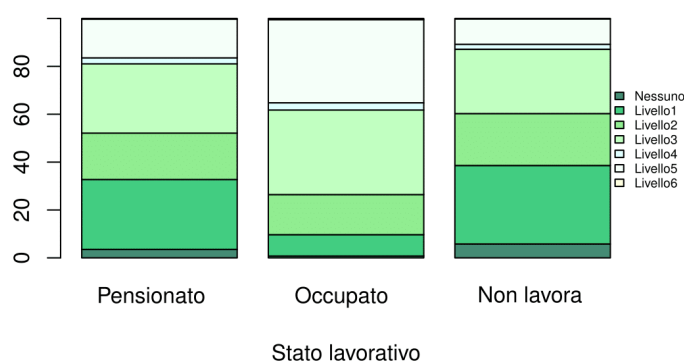


Figura 3.1: Distribuzione del livello di istruzione più status occupazionale

Il questionario raccoglie due tipi di informazioni sulla relazione di un individuo con un partner: lo stato "sentimentale/relazionale" e lo stato civile. In base allo scopo che ci si è prefissati, si è ritenuto più utile analizzare la prima variabile relativa alla relazione con una persona, innanzitutto perchè volendo analizzare la soddisfazione del reddito, un individuo sposato non per forza vive con un partner (e dunque il suo reddito familiare potrebbe non essere influenzato dal suo stato civile) e secondo perchè priva di informazione mancante, al contrario dell'altra. Alla luce di ciò, si sono accorpate in un'unica categoria gli intervistati sposati o che vivono assieme ad un'altra persona, generando così le variabili specificate nella Tabella 3.1: come si vedere in Figura 3.2 questi infatti costituiscono poco più dei $\frac{3}{4}$ del totale degli intervistati, di cui il 94.39% è il rispettivo coniuge; il 22.24% del campione vive da solo: quest'ultime persone sono soprattutto vedove (55.5%), divorziate (24.9%) o non sposate (16.1%), come si può notare dalla Figura 3.3.

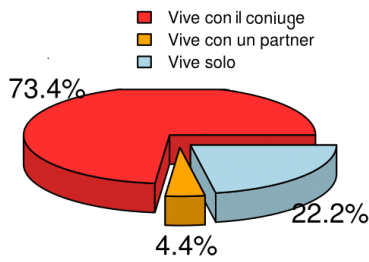


Figura 3.2: Boxplot sulle relazioni

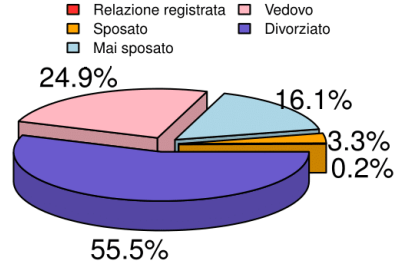


Figura 3.3: Boxplot dello stato civile per chi vive solo

Un'ulteriore informazione legata in parte alla variabile appena analizzata è il numero di componenti della famiglia dell'intervistato: la distribuzione di questa variabile corrisponde a quella attesa; infatti, come si può notare dalla Figura 3.4, il 56.30% vive con un'altra persona (tanto vero che il 95.66% di questi vive con un partner), il 17.18% vive solo e il 14.69% con altre 2 persone. Le successive percentuali diminuiscono velocemente all'aumentare del numero di componenti del nucleo familiare (circa 7.85% per 4 componenti, 2.34% per 5, 1% per 6 e così via).

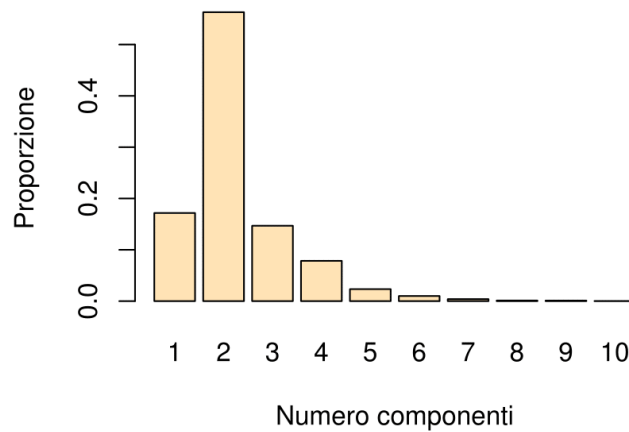


Figura 3.4: Istogramma del numero di componenti del nucleo familiare dell'individuo

Tra le informazioni raccolte ve ne sono due relative alla sfera genitoriale: la prima, comune anche ad altre indagini, è il numero di figli che l'intervistato ha tuttora in vita (solo il 9.37% non è attualmente genitore, mentre tra

chi lo è ha in media, e in mediana, due figli); la seconda, meno frequente, fa riferimento al fatto che almeno uno dei genitori dell'intervistato sia vivo al momento dell'indagine. Per quanto riguarda quest'ultima, il 25.91% fornisce una risposta affermativa: l'età media (e mediana) di questi soggetti infatti è di 56 anni, 10 anni in meno rispetto a coloro che non hanno più alcun genitore ancora in vita.

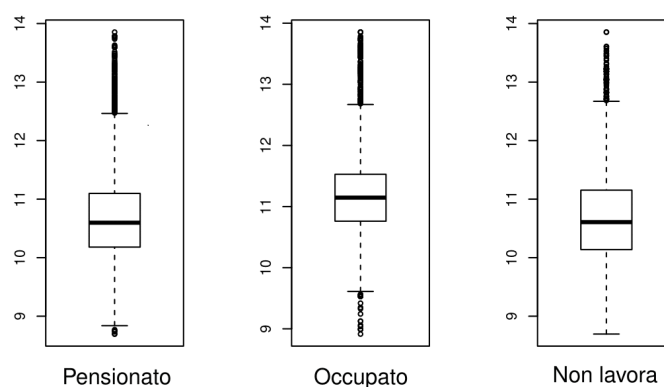
Un ampio numero di domande sono relative allo stato di salute dei cittadini europei: come si può notare dalla Tabella 3.1, in media gli individui presentano almeno un sintomo di malattia e di limitazione nella mobilità, nonché come minimo una malattia cronica. Una certa rilevanza risulta avere anche l'indicatore della depressione secondo la scala EURO-D: in media e in mediana gli individui riportano due sintomi. Nella Tabella 3.3 si possono invece osservare le singole distribuzioni nel dettaglio: tutte le misure relative allo stato di salute hanno percentuali associate al rispettivo numero di malattie/limitazioni/sintomi crescenti al peggioramento della condizione. In particolare nessuno afferma di avere i peggiori sintomi di salute per ognuno dei domini elencati, mentre solo 480 individui non ha nessun problema di salute tra questi; inoltre metà del campione gode di una buona salute in quanto non presenta difficoltà nella mobilità e nelle attività di tutti i giorni, strumentali e non (questi hanno un'età media di 61 anni e solo il 16% non è ne pensionato ne lavoratore).

Fattori come il numero di componenti di un nucleo familiare e il tipo di lavoro dell'individuo e dei suoi familiari possono influire su fattori come ad esempio il reddito familiare. Chiaramente questa quantità monetaria può variare anche in base allo stato lavorativo, specialmente se in famiglia vi è più di un componente che riceve uno stipendio regolarmente o una pensione. Nella Tabella 3.4 si possono osservare le distribuzioni per gli intervistati pensionati, occupati e i non lavoratori: queste sono asimmetriche positivamente e molto leptocurtiche; ciò non stupisce in quanto è noto in letteratura che la distribuzione normale non è appropriata a rappresentare l'aleatorietà dei redditi, in quanto tende a sottovalutare la probabilità attribuita ad eventi estremi (una breve introduzione sulle distribuzioni usate per analizzare questa componente monetaria può essere consultata in Betti (2015)). In questa tesi, al fine di tenere conto dei valori nulli e di mantenerli come tali, si è con-

Tabella 3.3: Distribuzione dei principali indicatori dello stato di salute

N°	Depressione	ADL	IADL	Croniche	Mobilità	Sintomi
0	25.06%	90.94%	85.79%	26.14%	54.15%	28.41%
1	22.55%	5.09%	8.19%	30.26%	14.23%	29.14%
2	15.93%	1.92%	2.84%	20.50%	9.44%	16.66%
3	12.34%	0.87%	1.35%	12.19%	6.90%	10.49%
4	8.73%	0.41%	0.76%	5.93%	4.39%	6.31%
5	5.34%	0.52%	0.44%	2.84%	3.59%	3.55%
6	3.78%	0.24%	0.27%	1.35%	2.73%	2.45%
7	2.56%	-	0.35%	0.41%	2.01%	1.48%
8	1.36%	-	-	0.19%	1.26%	0.84%
9	1.01%	-	-	0.12%	0.94%	0.35%
10	0.49%	-	-	0.04%	0.34%	0.20%
11	0.14%	-	-	-	-	0.07%
NA	0.71%	0.01%	0.01%	0.03%	0.01%	0.05%

siderata una trasformazione arcseno iperbolica; allo stesso tempo tale scelta permette di serbare i valori positivi e negativi e di ottenere una distribuzione più vicina a quella gaussiana, come si può notare dai boxplot in Figura 3.5, sebbene la presenza di numerosi valori anomali.

**Figura 3.5:** Distribuzione arcseno iperbolico del reddito netto familiare

Con la stessa logica, si è operato con le altre due misure della ricchezza, sebbene la gaussianità raggiunta con tale trasformazione non sia così soddisfacente come quella per il reddito, ma più accettabile di quella attribuita alla scala originale (si veda la Figura 3.7): nello specifico, l'indicatore Skewness

Tabella 3.4: Distribuzione reddito netto mensile familiare

	Pensionati	Occupati	Non Lavora
N	3654	2199	1467
%N=0	0	0	0
Min	2986.805	3722.019	1622.043
Max	519730.5	519730.5	563758.4
Media	38894.74	55913.72	39442.19
Deviazione Standard	58657.39	73493.79	58414.98
Skewness	3.733	3.647	3.705
Curtosi	19.153	17.215	19.529
p25	13191.23	23123.89	12636.48
p50	20018.07	37577.96	20211.74
p75	33009.28	56459.58	34918.85
p99	319810.5	531543.6	300118.3

è pari a -1.580 per la ricchezza finanziaria e a -2.174 per quella reale e dunque la distribuzione è asimmetrica negativamente, mentre l'indice di curtosi è pari rispettivamente a 4.484 ed a 7.08 e quindi la prima si avvicina a quello di una distribuzione normale, mentre la seconda indica una distribuzione a code pesanti. Non devono preoccupare invece i valori negativi associati alla ricchezza finanziaria; infatti questa è al netto di debiti e dunque, qualora fosse negativa, tale quantità corrisponderebbe ad un debito del soggetto stesso.

Infine nella Tabella 3.5 si riportano alcune statistiche descrittive per una parte delle variabili esplicative tra gli 11 Paesi Europei: il maggior numero di cittadini è di nazionalità tedesca (15.14%), mentre il minor è di nazionalità francese (5%). La leggera predominanza di donne si mantiene per tutti i Paesi, specialmente per la Repubblica Ceca e la Polonia, mentre l'età media e quella mediana si differenziano da quelle del campione totale, sebbene si conferma l'asimmetria positiva per ogni Stato: i più giovani sono i cittadini polacchi e quelli olandesi, i più anziani gli individui svedesi (tutte e tre rispettivamente di 2 anni in più e in meno in termini di mediana rispetto a quella di tutti i cittadini). Per quanto riguarda l'istruzione, Spagna, Polonia e Italia presentano la percentuale maggiore di soggetti che non hanno un titolo di studio più elevato di quello primario, seppur in mediana solo la Spagna

	Finanziaria	Reale
N	7353	7353
%N=0	11.94%	14.72%
Min	-178690.7	0
Max	994585	1080537
Media	41226.29	196190.4
Deviazione Standard	84295.56	272389.4
Skewness	3.900	4.298
Curtosi	25.130	45.796
p25	480.96	35137.91
p50	10105.15	129599.8
p75	46689.14	260000
p99	415389.6	1236905

Figura 3.6: Distribuzione ricchezza finanziaria e reale

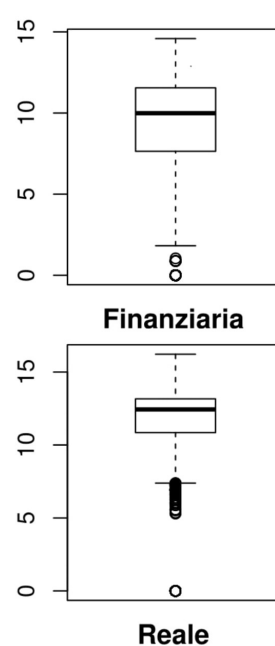


Figura 3.7: Distribuzione arco-iperbolico della ricchezza finanziaria e reale

ha il livello più basso tra tutti i 10 Paesi. Caratteristiche come le relazioni e il numero di componenti rispecchiano la media generale, contrariamente a quella legata al lavoro: la percentuale di pensionati non risulta essere la più alta in Olanda e in Spagna dove, al contrario, il 40% dei cittadini sono rispettivamente occupati e non lavoratori; in Svezia e Repubblica Ceca vi sono invece le proporzioni maggiori di intervistati ex o attuali lavoratori, mentre le inferiori di non occupati. I rispondenti olandesi sono inoltre gli individui con le migliori condizioni di salute, diversamente da quelli dei polacchi che hanno le peggiori. Le differenze maggiori si manifestano nelle quantità monetarie: in termini di mediana la Polonia è il paese più povero per tutte e tre le ricchezze (nel 2007 era lo Stato con il più basso PIL a parità di potere d'acquisto tra gli 11 qui considerati (OECD, 2017)), mentre la Danimarca uno dei più ricchi in mediana. In particolare gli intervistati olandesi si posizionano al primo posto per il più alto reddito mediano, mentre quelli francesi hanno la più alta ricchezza reale mediana, ma la più bassa ricchezza finanziaria dopo Polonia e Grecia.

3.3.2 Analisi sulla soddisfazione

Le analisi descrittive della Sezione precedente consentono di avere un'idea generale delle caratteristiche demografiche, di salute ed economiche dei cittadini europei nel campione in esame. Si procede ora con l'esposizione delle principali evidenze relative alle domande di ricerca, ovvero alla soddisfazione del reddito e della vita, e dunque si vuole indagare se vi siano differenze in termini di soddisfazione riportata nel questionario tra gli individui. In particolare è noto come in Olanda e nei Paesi Nordici in generale sembrano vivere gli individui più felici e più compiaciuti in Europa (Sezione 1.2; <http://www.oecdbetterlifeindex.org/countries/netherlands/>), mentre Svezia e Olanda sono i più ricchi dal punto di vista del reddito (OECD (2017); <http://www.oecdbetterlifeindex.org/topics/income/>). Dal momento che i soldi non comprano la felicità, ma che possono portare ad un benessere migliore, si vuole inoltre verificare se l'evidenzia per questi Stati sia confermata nel campione in esame. Come già detto nella Sezione 3.2, per ciascun dominio vengono formulate tre domande, una relativa alla soddisfazione per-

Tabella 3.5: Media, mediana e composizione di alcune variabili esplicative tra i Paesi Europei

	DE	SW	NL	ES	IT	FR	DK	GR	BE	CZ	PL
N%	15.14	6.19	6.87	6.81	9.21	5.00	13.10	7.40	11.32	11.63	7.34
%Donna	53.73	54.73	52.67	54.49	54.21	55.98	55.45	55.15	54.09	59.65	57.04
Età											
Media	64.92	66.21	61.50	64.22	64.82	64.94	64.06	64.12	65.38	64.49	62.73
Mediana	64	65	60	63	64	63	63	62	64	63	60
Educazione											
%bassa	0.63	30.11	13.06	57.09	49.63	36.41	15.58	33.90	25	20.35	43.89
%media	68.01	36.70	59.21	29.94	41.21	38.86	45.69	32.90	51.44	68.42	38.70
%alta	31.36	33.19	27.73	12.97	9.16	24.73	38.73	20.41	23.56	11.23	17.41
Mediana	media	media	media	bassa	media	media	media	media	media	media	media
Relazioni											
%Single	19.95	21.29	15.45	19.36	16.54	28.80	18.38	28.13	24.76	30.06	23.52
%Partner	80.05	78.71	84.55	80.64	83.46	71.20	81.62	71.87	75.24	69.94	76.48
Componenti											
Media	2.12	2.01	2.29	2.80	2.60	2.12	2.02	2.40	2.12	2.13	2.88
Mediana	2	2	2	2	2	2	2	2	2	2	2
Lavoro											
Pensionato	52.97	58.13	32.67	33.80	52.44	55.43	46.76	37.68	50	65.45	54.26
Occupato	27.84	37.64	41.24	26	18.91	29.08	44.14	34.19	22.88	29.25	19.26
Non Lavora	19.19	4.23	26.10	40.20	28.66	15.49	9.10	28.13	27.12	5.31	26.48
Depressione (mediana)	1	2	1	2	2	2	1	1	2	1	4
Mobilità											
Media	1.36	1.21	0.91	1.80	1.63	1.23	0.82	1.68	1.32	1.48	2.54
Mediana	0	0	0	0	1	0	0	1	0	1	2
Sintomi											
Media	1.62	1.72	1.15	1.65	1.79	1.87	1.49	1.42	1.87	2.09	2.60
Mediana	1	1	1	1	1	1	1	1	1	2	2
Reddito familiare (mediana)	24218.7	25419.5	31712.7	13062.2	18084.1	28016.8	28637.6	14866.8	19904.7	12441.1	93010.8
Ricchezza reale (mediana)	104447.5	96383.3	161999.8	202232.3	219243.16	225612.3	142945.6	152324.7	191655.1	65489	36916.3
Ricchezza finanziaria (mediana)	19249.28	34753.6	24545.42	2948.75	5771.506	17883.33	44026.48	0	33456.01	3349.64	0.000005

sonale e due relative a quella degli individui ipotetici, ovvero alle *anchoring vignettes*.

3.3.2.1 Soddisfazione del reddito

La prima dimensione che si va ad analizzare è la soddisfazione del reddito: questa aumenta quasi proporzionalmente dal livello più basso a quello più alto, tranne per la categoria più elevata che viene valutata da circa un decimo degli intervistati: infatti la percentuale di coloro che sono "*Molto Insoddisfatti*" è pari al 4.88%, mentre per coloro che sono "*Soddisfatti*" corrisponde al 42.65% del campione analizzato (Figura 3.8).

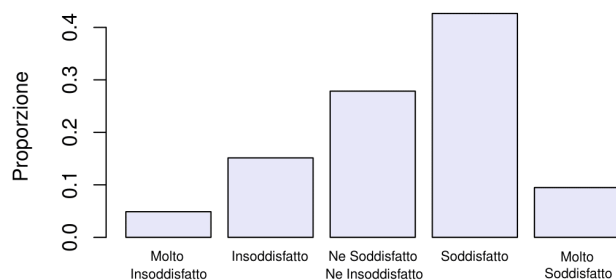


Figura 3.8: Istogramma della soddisfazione del reddito

In riferimento ai singoli Paesi (Figura 3.9), la distribuzione generale non viene rispecchiata da Grecia, Repubblica Ceca e Polonia: il livello di soddisfazione associato alla maggior frequenza non è ben marcato come positivo o negativo; inoltre il 43% degli intervistati polacchi è "*Insoddisfatto*" o "*Molto insoddisfatto*" del proprio reddito. In particolare, in Danimarca e in Olanda vivono gli individui più appagati, rispettivamente con una percentuale pari all'80.89% e al 76.6%: l'evidenza del primo Paese non sorprende poichè è un'ulteriore conferma di quanto emerso nel lavoro di Bonsang e Van Soest (2012) (come evidenziato nella Sezione 1.3), mentre un esito inatteso e opposto risulta per il secondo. Conseguentemente, potrebbe essere che le scale di risposta adottate dagli individui di ciascun Paese siano differenti e dunque che siano caratterizzate da un'eterogeneità tra Paesi: la differenza sostanziale di valutazione sembra emergere tra Polonia e Danimarca (o Germania). I cittadini polacchi, seguiti dai greci, sono mediamente i più pessimisti, mentre i danesi (nuovamente) i più positivi.

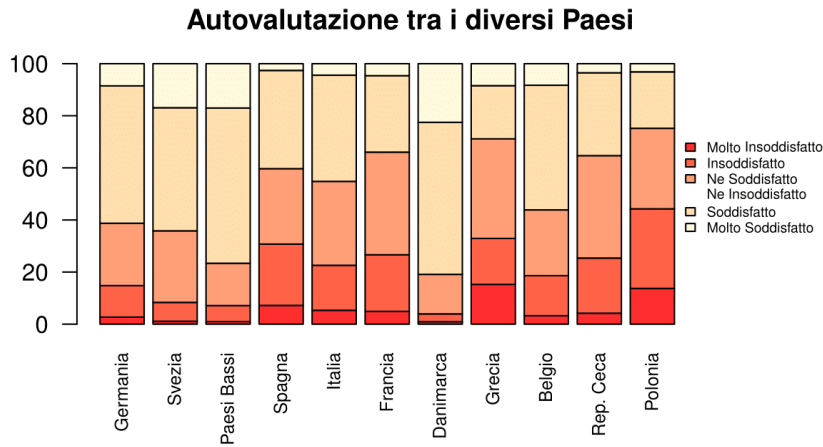


Figura 3.9: Soddisfazione del reddito per Paese

Nella Tabella 3.6 (prime due colonne) si può notare come gli uomini siano leggermente più soddisfatti rispetto alle donne (la differenza è di poco più di 3 punti percentuali), le quali, in opposizione, hanno percentuali più alte nelle altre tre categorie. Perciò, gli intervistati di genere maschile sembrano essere un po' più appagati del proprio reddito familiare rispetto a quelli di genere femminile.

Essendo il reddito una realtà indubbiamente legata alla sfera lavorativa (seppur questa non per forza esclusiva di una ricchezza maggiore), ci si aspetta una differenza di valutazione della sfera d'interesse tra le diverse categorie occupazionali. A dispetto di quanto appena detto, le distribuzioni nelle tre categorie sono molto simili, specialmente per chi ha lavorato o sta lavorando: la modalità "*Soddisfatto*" è caratterizzata dalla più alta frequenza, segue "*Ne soddisfatto né insoddisfatto*". La percentuale di non lavoratori che si ritengono "*Molto insoddisfatti*" è più alta rispetto alle altre due categorie legate al lavoro, le quali sono un po' più "*Molto soddisfatti*", specie gli occupati.

Per quanto riguarda il livello di istruzione si può notare che, indipendentemente dal livello, la soddisfazione aumenta progressivamente dalla categoria "*Molto insoddisfatto*" a "*Soddisfatto*". Ad ogni modo, i cittadini più colti sono più compiaciuti, mentre quelli meno eruditi sono più malcontenti; l'associazione positiva che sembra emergere dalle percentuali trova conferma dalla correlazione policorica, pari a 0.271.

Nelle ultime righe della Tabella 3.6 vengono riportate le percentuali relative alle relazioni dei cittadini: queste confermano quanto ci si potrebbe aspettare, ovvero individui sposati sono più contenti del reddito familiare a loro disposizione; infatti se in una famiglia lavorano più persone (ad esempio il coniuge o il partner), il reddito totale della famiglia dovrebbe essere più alto di quello che potrebbe avere un single, anche se possono esservi altri fattori ad influenzare tale componente soggettiva (ad esempio un imprenditore potrebbe essere più appagato perchè ha un profitto maggiore di un operaio).

Tabella 3.6: Analisi bivariata della soddisfazione del reddito con alcune variabili socio-demografiche ‡

	Molto insoddisfatto	Insoddisfatto	Né soddisfatto Né insoddisfatto	Soddisfatto	Molto soddisfatto
Genere					
Maschi	4.56%	14.28%	27.40%	44.58%	9.17%
Donne	5.15%	15.38%	28.22%	41.07%	9.37%
Lavoro					
Pensionato	4.57%	14.75%	29.06%	42.61%	9.00%
Occupato	2%	12.46%	25.6%	48.16%	11.78%
Non lavora	10.02%	20.18%	28.08%	34.49%	7.23%
Istruzione					
Bassa	8.65%	21.52%	31.22%	32.27%	6.39%
Media	4.02%	14.66%	30.01%	43.21%	8.11%
Alta	2.37%	8.80%	19.44%	53.47%	15.91%
Relazione					
Partner	4.11%	14.39%	26.95%	44.39%	10.16%
Single	7.58%	17.74%	31.01%	36.57%	7.09%

‡ Le percentuali sono calcolate per riga ovvero per il singolo livello della variabile categoriale.

In Figura 3.10 viene riportato un grafico relativo alla percentuale di individui soddisfatti⁵ in ciascun Paese e in riferimento ai seguenti intervalli del reddito familiare netto annuo specifici per ogni Stato:

- <p10: reddito inferiore al percentile 10 del Paese x;
- p10-p25: reddito compreso tra il percentile 10 e 25 del Paese x;

⁵Le percentuali corrispondono al rapporto degli individui che hanno valutato le categorie "Soddisfatto" e "Molto Soddisfatto" sul totale di coloro che possiedono un reddito compreso negli intervalli definiti.

- p25-p50: reddito compreso tra il percentile 25 e 50 del Paese x;
- p50-p75: reddito compreso tra il percentile 50 e 75 del Paese x;
- p75-p90: reddito compreso tra il percentile 75 e 90 del Paese x;
- >p90: reddito superiore o uguale al percentile 90 del Paese x.

In generale la proporzione di soggetti soddisfatti aumenta al crescere della disponibilità economica per quasi tutti i Paesi, evidenziando dunque una relazione positiva tra le due variabili: in particolare questa constatazione è raffigurata perfettamente dalla Svezia. Al contrario, Grecia e Italia sono i due Stati che si distaccano parzialmente da tale realtà: la percentuale di cittadini del primo diminuisce gradualmente se si ha un reddito superiore o pari a quello mediano, mentre quella del secondo presenta un minimo in corrispondenza dell'intervallo p75-p95. Le percentuali del quantile 95-esimo non devono comunque preoccupare se inferiori a quelle dell'intervallo precedente, in quanto rappresentano casi estremi. Il maggior numero di individui soddisfatti sono olandesi e danesi, indipendentemente dagli intervalli di reddito considerati, come ci si poteva aspettare in base a quanto visto nella Figura ?? e alla letteratura documentata nelle precedenti pagine; per quanto riguarda i due intervalli estremi, tra i meno ricchi meno soddisfatti sono gli intervistati polacchi (14.81%), mentre tra i meno poveri quelli greci (26.79%). L'Italia, che risulta avere in mediana il reddito più alto tra gli 11 Paesi del campione (Tabella 3.5), mantiene complessivamente una posizione centrale per quanto riguarda la proporzione di individui appagati indipendentemente dalla ricchezza monetaria posseduta, ad eccezione di coloro che sono economicamente più sfortunati. Si specifica che la variabile relativa al reddito non verrà presa in considerazione nella stima dei modelli poichè potenzialmente endogena con la soddisfazione del reddito, come sembra confermare la relazione monotoma crescente del grafico appena descritto.

Avendo appena analizzato le autovalutazioni sul reddito, viene naturale chiedersi come i cittadini abbiano classificato gli individui ipotetici e, soprattutto, se abbiano fornito delle risposte simili alle due *vignettes*. Da una prima e semplice analisi descrittiva riportata nella Figura 3.11 emerge che le distribuzioni delle valutazioni relative alle *vignettes* sono diametralmente opposte

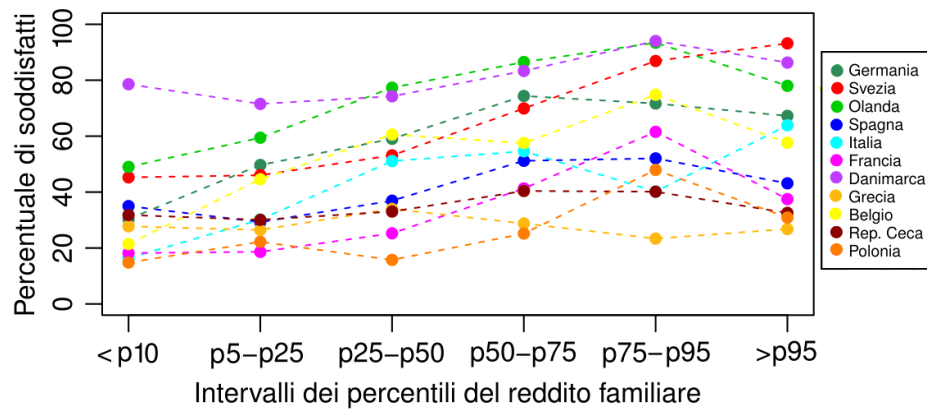


Figura 3.10: Reddito familiare annuale e soddisfazione del reddito totale

tra loro: se quella relativa all'individuo ipotetico Anna segue quella associata all'autovalutazione (seppur con percentuali differenti) e quindi la si ritiene soddisfatta del proprio reddito, non si può affermare altrettanto per Giorgio, al quale quasi metà degli intervistati attribuisce un'insoddisfazione (debole o forte) del dominio del benessere d'interesse, come d'altronde ci si poteva aspettare dal momento che l'unica differenza tra i due individui ipotetici è il reddito totale e che questo risulta essere maggiore proprio per Anna. Ulteriori dettagli sull'ordinamento degli scenari ipotetici verranno approfonditi meglio nella Sezione 3.4.1.

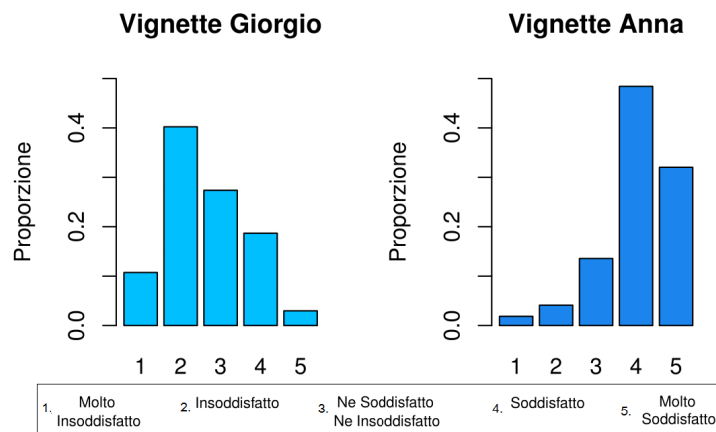


Figura 3.11: Valutazione delle vignettes sulla soddisfazione del reddito

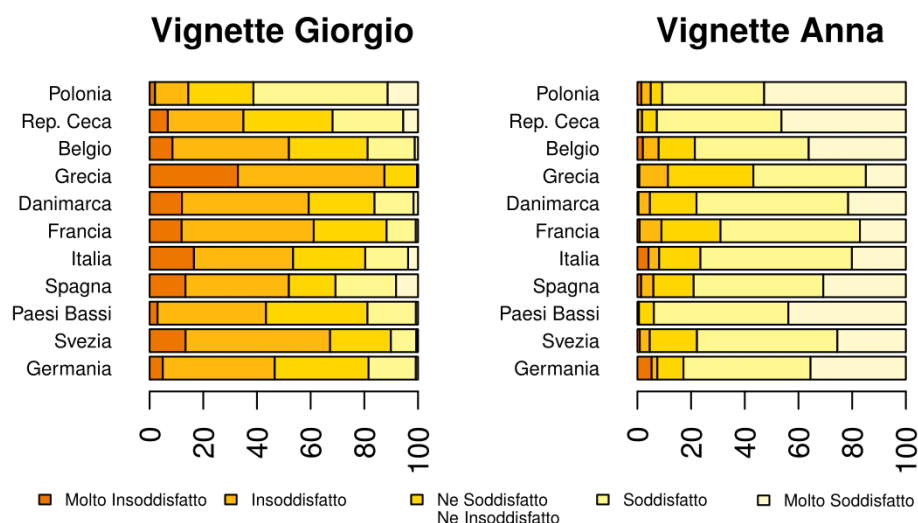


Figura 3.12: Valutazione delle vignettes sulla soddisfazione del reddito per Paese

In particolare, poichè il campione in analisi è relativo a 11 Paesi che hanno culture, religione, ecc diverse, non è detto che le possibili differenze emergenti nel momento in cui classificano la propria soddisfazione si confermino anche nella valutazione degli individui ipotetici. In riferimento alla *vignette* che descrive la situazione di Giorgio, gli intervistati categorizzano essenzialmente l'individuo come "Insoddisfatto" o "Né soddisfatto né insoddisfatto", ad eccezione di Grecia e Polonia: il 54.60% della popolazione greca valuta Giorgio come "Insoddisfatto", mentre il 50% dei cittadini polacchi come "Soddisfatto", un risultato inaspettato in quanto nella Figura 3.9 si autovalutano con il livello negativo opposto, un'evidenza che potrebbe essere legata alla situazione in cui si trova Giorgio. Per quanto riguarda la *vignette* di Anna, i rispondenti sembrano concordare sulla soddisfazione (media-elevata); ciò viene confermato anche dai cittadini greci, anche se vi è una percentuale non trascurabile (e maggiore degli altri Paesi) di individui che non la classificano né come soddisfatta né insoddisfatta. Comunque, la distribuzione differenziata per ciascun Paese rispecchia quella generale in entrambe le descrizioni, seppur non vi sia un riscontro favorevole per la Grecia.

Come ultima analisi, si riportano le correlazioni policoriche tra l'autovalutazione e ciascuna valutazione delle *vignettes*. Ci si aspetta infatti che l'associazione tra le risposte alle due domande aumenti man mano che si individua

un campione con caratteristiche il più possibile affini con quelle attribuite all'individuo ipotetico. I risultati nella Tabella 3.7 confermano debolmente quanto appena supposto; infatti se si considera l'intero campione, non c'è evidenza del fatto che gli individui valutino similmente sia se stessi che i soggetti ipotetici, mentre questa diventa debole (0.14) se ci si focalizza su coloro che hanno la stessa situazione descritta nella *vignette*. Chiaramente, l'ampiezza del campione si restringe man mano che vengono selezionati intervistati che sono "cloni" di quelli ipotetici, specie se queste persone costituiscono un numero molto limitato. In particolare questi risultati potrebbero metter in dubbio l'assunzione *response consistency* viste le deboli associazioni, ipotesi che tuttavia richiederebbe un maggior approfondimento.

Tabella 3.7: Correlazioni tra la soddisfazione del reddito e le rispettive *vignette*

	Selezione del campione attraverso			
	<i>Nessuna</i>	(1) <i>Demografiche</i>	(2) <i>(1)+Familiari</i>	(3) <i>(2)+Reddito</i>
Giorgio	0.029	0.043	0.069	0.145
Anna	0.046	0.035	0.070	0.143

Note: Carlo: Demografiche=è maschio; Familiari=ha due figli ed è sposato; Reddito=ha un reddito compreso nel range (-50,+50) di quello specifico nella *vignette* di ogni Paese descritto;

Anna: Demografiche=è femmina; Familiari=ha due figli ed è sposato; Reddito=ha un reddito compreso nel range (-50,+50) di quello specifico nella *vignette* di ogni Paese descritto

3.3.2.2 Soddisfazione della vita

Si procede ora con la descrizione del secondo dominio d'interesse, la soddisfazione della vita: questa è caratterizzata da una distribuzione molto affine a quella della soddisfazione del reddito e in particolare la frequenza relativa di ciascuna modalità aumenta quasi proporzionalmente dal livello inferiore a quello superiore, tranne per la categoria più alta che viene valutata da circa un diciannovesimo degli intervistati (una proporzione che è quasi il doppio di

quella emersa dalla soddisfazione del reddito): la percentuale di coloro che sono "*Molto Insoddisfatti*" è quasi nulla, mentre coloro che sono "*Soddisfatti*" corrispondono al 58.49% del campione analizzato (Figura 3.13).

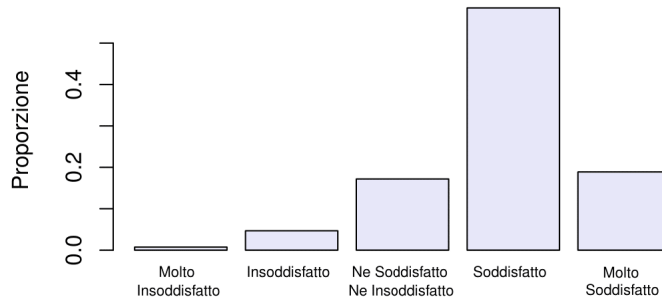


Figura 3.13: Istogramma della soddisfazione della vita

In riferimento ai singoli Paesi, la distribuzione generale è rispecchiata in parte: infatti nella Figura 3.14 si può notare come i cittadini siano sostanzialmente soddisfatti della propria vita. In particolare, Danimarca e Olanda sono i Paesi più appagati anche per questa realtà (rispettivamente con una percentuale pari a 92.42% e al 91.88%), ulteriore conferma di quanto già si è detto e visto nella Sezione 1.2, mentre la Grecia è il Paese con la proporzione più bassa di soddisfatti (relativamente parlando in quanto essa rappresenta circa la metà dei suoi cittadini intervistati). Le frequenze relative ai livelli più bassi di felicità sono abbastanza omogenee tra i diversi Stati e nello specifico inferiori al 7.5%, ad eccezione di Italia (11.82%), Spagna (9.78%) e Grecia (7.54%). In linea generale queste evidenze sono quindi coerenti con quanto è riportato nella Sezione 1.2. Una possibile spiegazione riguardo alla maggior soddisfazione della vita rispetto al reddito potrebbe essere la sua natura, ovvero il fatto che essa sia una componente aggregata di più concetti, come la soddisfazione del lavoro, la soddisfazione della salute, ecc (si vedrà meglio nella prossima Sezione il legame tra le due sfere d'interesse). Comunque, ad un'attenta osservazione, si può notare anche come la distribuzione degli individui polacchi e cechi sia sostanzialmente identica, diversamente da altri Stati: la differenza sostanziale sembra essere tra Polonia e Danimarca o Italia e Svezia.

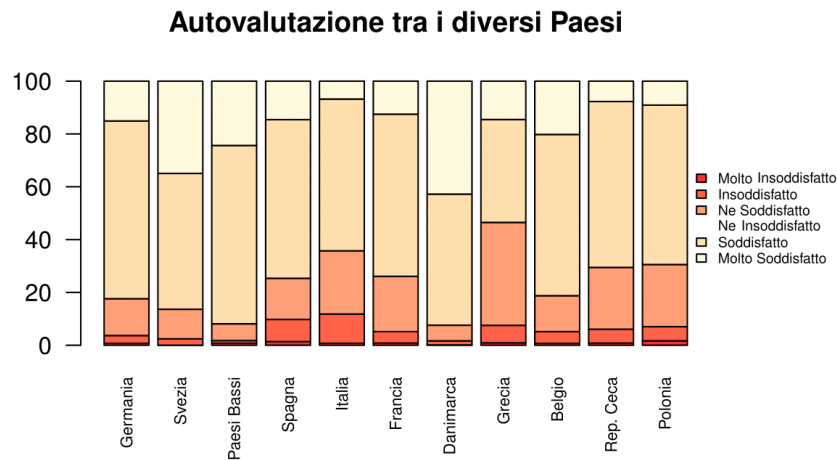


Figura 3.14: Soddisfazione della vita per Paese

Nella Tabella 3.8 viene riportata l'analisi bivariata del dominio d'interesse con alcune variabili socio-demografiche: a primo impatto emerge che il livello di soddisfazione più basso ha una percentuale inferiore all'1% per tutte le modalità, mentre è prossimo all'1.3% per chi rientra nella categoria più sfortunata di istruzione e status occupazionale; diversamente, la proporzione dei "Soddisfatti" è più del 50% dei cittadini con una determinata caratteristica considerata nella tabella, come ci si potrebbe aspettare vista la distribuzione generale della soddisfazione. Gli uomini sono leggermente più soddisfatti rispetto alle donne (la differenza è di poco oltre i 3 punti percentuali), le quali, al contrario, sono più insoddisfatte o indifferenti. Perciò, come per la soddisfazione del reddito familiare, gli intervistati di genere maschile sembrano essere un po' più appagati rispetto a quelli di sesso femminile.

Si è visto nella Sezione 1.2 che un fattore molto rilevante è lo stato occupazionale: in questo campione chi ha lavorato o sta lavorando è più felice, mentre chi non percepisce uno stipendio regolare è indifferente o malsoddisfatto.

La categoria "Soddisfatti" ha la frequenza relativa più elevata per chi ha un'istruzione media, nonostante i più soddisfatti in generale siano quelli con un'educazione più alta, i quali presentano una proporzione più bassa per le modalità inferiori della "felicità". Pertanto, sembra esserci un'influenza positiva tra il fatto di essere più eruditi e quello di essere più "contenti"; ciò pare essere confermato dalla correlazione policorica tra soddisfazione ed edu-

cazione, pari a 0.2.

Nelle ultime righe della Tabella 3.8 si può notare una distribuzione delle relazioni degli individui simile a quella che si è vista per la soddisfazione del reddito: infatti, chi ha un partner è più compiaciuto di circa 13 punti percentuali, mentre chi non ha un compagno tende ad affermare un giudizio non positivo.

Tabella 3.8: Analisi bivariata della soddisfazione della vita con alcune variabili socio-demografiche ‡

	Molto insoddisfatto	Insoddisfatto	Né soddisfatto Né insoddisfatto	Soddisfatto	Molto soddisfatto
Genere					
Maschi	0.73%	3.92%	16.43%	60.05%	18.86%
Donne	0.76%	5.29%	17.80%	57.23%	18.91%
Lavoro					
Pensionato	0.82%	4.32%	17.16%	57.64%	19.56%
Occupato	0.23%	2.27%	13.46%	61.30%	22.74%
Non lavora	1.30%	9.07%	22.70%	52.15%	14.79%
Istruzione					
Bassa	1.26%	7.74%	23.03%	53.14%	14.83%
Media	0.69%	3.88%	16.50%	61.55%	17.38%
Alta	0.29%	2.84%	11.92%	58.22%	26.74%
Relazione					
Partner	0.70%	3.93%	14.85%	59.69%	20.83%
Single	0.92%	7.28%	25.38%	54.31%	12.11%

‡ Le percentuali sono calcolate per riga ovvero per il singolo livello della variabile categoriale.

Avendo appena analizzato le autovalutazioni sulla vita, viene naturale chiedersi come i cittadini abbiano classificato gli individui ipotetici e, soprattutto, se abbiano fornito delle risposte ineguali alle due *vignettes*.

Da una prima e semplice analisi descrittiva, riportata in Figura 3.15, emerge che le distribuzioni delle valutazioni relative alle *vignettes* non son simili tra loro: se quella relativa all'individuo ipotetico Franca segue quella associata all'autovalutazione (seppur con percentuali differenti) e in particolare metà del campione la classifica come "*Soddisfatta*" della propria vita, non si può affermare altrettanto per Carlo, al quale quasi l'80% degli intervistati attribuisce in proporzioni quasi uguali un'insoddisfazione o indifferenza del dominio del benessere generale. Ciò sembra costituire un primo allarme per

quanto riguarda l'ordinamento degli scenari ipotetici, come sarà approfondito meglio nella Sezione 3.4.1.

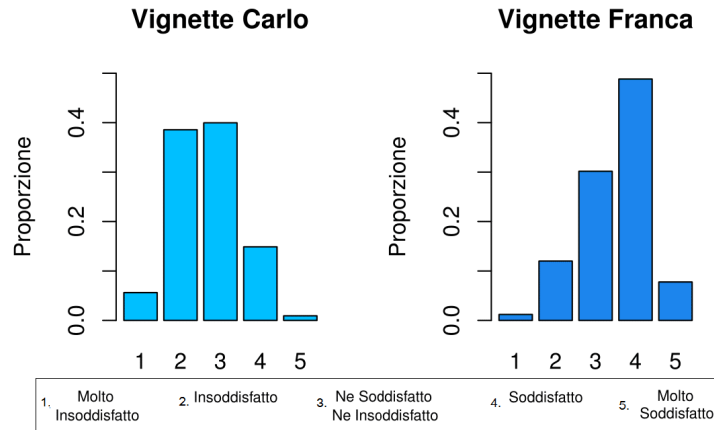


Figura 3.15: Valutazione delle vignettes sulla soddisfazione della vita

In particolare poichè il campione in analisi è relativo a 11 Paesi che hanno culture, religioni, sistemi di *welfare*, ecc diversi, nella Figura 3.16 si riporta la loro singola distribuzione relativa alla valutazione degli individui ipotetici. In riferimento alla *vignette* che descrive la situazione di Carlo, gli intervistati categorizzano essenzialmente l'individuo dello scenario come "*Insoddisfatto*" o "*Né soddisfatto né insoddisfatto*": il 14.18% dei cittadini italiani valuta Carlo come "*Insoddisfatto*", categoria che viene optata solo dallo 0.83% dei danesi. Quindi l'Italia, assieme a Spagna, è il Paese che ha la visione più pessimistica. Al contrario, gli individui danesi e tedeschi hanno un giudizio molto positivo sulla soddisfazione di Carlo, oltre che di se stessi.

Per quanto riguarda la seconda *vignette*, le persone del campione in esame considerano la situazione di Franca più positivamente di quella di Carlo. Ad eccezione di un paio di Paesi (Grecia e Francia), più del 50% dei cittadini di ciascun Stato valuta Franca come soddisfatta: la Danimarca in particolare si riconferma nuovamente come lo Stato con la più alta percentuale di valutazioni "*Soddisfatto*" o "*Molto soddisfatto*" (75%). Le modalità associate all'insoddisfazione sono maggiormente scelte dagli intervistati greci, poco più del 20% contro il 15% circa degli altri cittadini europei. Inoltre, Spagna e Germania sembrano essere i due Paesi con le distribuzioni più simili, diversamente da quanto emerge dalla *vignette* di Carlo.

Entrambe le *vignettes* sembrano comunque essere caratterizzate da distribuzioni eterogenee tra i diversi Stati, specialmente per quelle relative alla valutazione di Carlo.

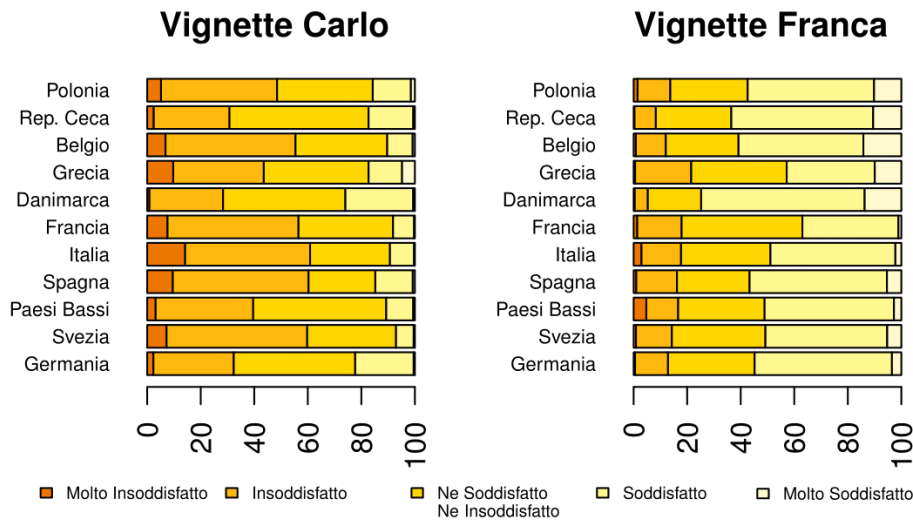


Figura 3.16: Valutazione delle vignettes sulla soddisfazione del reddito per Paese

Come ultima analisi, si riportano alcune correlazioni policoriche tra l'autovalutazione e ciascuna valutazione delle *vignettes*. I risultati nella Tabella 3.9 confermano quanto ci si aspetta, ovvero che l'associazione tra le due domande aumenta man mano che si individua un campione con caratteristiche il più possibile affini con quelle attribuite all'individuo ipotetico; infatti, se si considera l'intero campione, non c'è evidenza del fatto che gli individui valutano similmente sia se stessi che Carlo, mentre questa assume maggior importanza se ci si focalizza sugli intervistati che hanno una situazione simile a quella descritta nella *vignette*. Come già si è detto nelle stesse analisi della soddisfazione del reddito, l'ampiezza del campione si restringe man mano che vengono selezionati individui che sono "cloni" di quelli ipotetici. Si può giungere alle stesse conclusioni se si considera la domanda relativa a Franca. Pertanto queste associazioni discrete sembrano essere a favore della *response consistency*, dal momento che gli individui più simili a quelli ipotetici sembrano dare similari valutazioni a se stessi e a Carlo o Franca.

Tabella 3.9: Correlazioni tra la soddisfazione della vita e le rispettive *vignettes*

	Selezione del campione attraverso					
	<i>Nessuna</i>	(1) <i>Demografiche</i>	(2) <i>(1)+Salute</i>	(3) <i>(2)+Figli</i>	<i>(3)+Lavoro</i>	<i>(2)+Casa</i>
Carlo	0.029	0.091	0.104	0.275	0.452	
Franca	0.156	0.179	0.289			0.331

Note: Carlo: Demografiche=è maschio ed ha un'età compresa tra i 56 e i 65 anni; Salute=ha almeno un sintomo; Figli=ha più di 3 figli; Lavoro=non lavora;

Anna: Demografiche=è femmina ed ha un'età compresa tra i 66 e i 79 anni; Salute=non è depressa e ha almeno una malattia cronica; Casa=possiede una casa

3.3.2.3 Analisi congiunta delle due soddisfazioni

Dalle analisi presentate nella Sezione precedente sembrerebbe esistere un'affinità tra le due componenti d'interesse, in quanto sono emerse evidenze particolarmente simili. In questa Sezione si studia più approfonditamente il legame tra le due sfere e precisamente l'associazione che potrebbe esservi. Nella Tabella 3.10 vengono riportate le percentuali di rispondenti alle due domande: esaminando le valutazioni degli intervistati che rispondono ugualmente ad entrambi i domini d'interesse, la maggior parte dei cittadini afferma di essere soddisfatto, mentre meno risposte vengono fornite nelle valutazioni negative: nello specifico, la proporzione di questi individui aumenta progressivamente man man che il giudizio diventa più positivo, ad eccezione delle modalità più soddisfatte. In particolare si noti come aumenta in generale il numero di individui che tendono ad autovalutarsi più positivamente nei confronti della soddisfazione della vita e più negativamente per quella del reddito: queste percentuali sono infatti più elevate rispetto a quelle associate ad un maggiore grado di compiacimento del proprio reddito ed uno più basso della propria vita. Ciò potrebbe essere dovuto al fatto che la prima realtà (reddito) è un singolo dominio della seconda; tuttavia, non essendo d'interesse in questa tesi il legame causale delle due dimensioni, non si indaga ulteriormente sulla verifica di tale supposizione. Ciò che invece è interessante è l'evidenza di una possibile relazione tra le due autovalutazioni. Infine si noti la nulla (o quasi) proporzione di individui che fornisce giudizi completamente opposti alle due autovalutazioni: anche in questa circostanza comunque, coloro

che affermano di essere "*Molto insoddisfatti*" del proprio reddito, ma "*Molto soddisfatti*" della propria vita sono leggermente superiori a coloro che si valutano "*Molto soddisfatti*" del primo dominio e "*Molto insoddisfatti*" del secondo (0.29% contro 0.03%). Pertanto, gli individui tendono ad avere una stessa (o simile) visione delle due sfere del benessere ed ad essere più positivi per quella generale.

Tabella 3.10: Analisi bivariata della soddisfazione della vita e del reddito

	Soddisfazione della vita					
	Molto insoddisfatto	Insoddisfatto	Né soddisfatto Né insoddisfatto	Soddisfatto	Molto soddisfatto	
Soddisfazione del reddito	Molto					
	Insoddisfatto	0.23%	0.99%	1.69%	1.69%	0.29%
	Insoddisfatto	0.37%	2.03%	4.92%	7.09%	0.73%
	Né soddisfatto	0.10%	0.88%	7.03%	17.31%	2.53%
	Né insoddisfatto	0.03%	0.60%	3.01%	29.17%	9.85%
	Soddisfatto	0.03%	0.18%	0.54%	3.24%	5.49%

Le evidenze appena emerse portano a pensare all'esistenza di una associazione tra le due variabili d'interesse. Data la loro natura ordinale, la correlazione di Pearson non risulta appropriata. Conseguentemente si è ricorso a degli strumenti più adeguati al caso in questione. Nella Tabella 3.11 vengono riportati le misure nella Sezione 2.4: le due domande hanno un'associazione discreta in quanto tutte le misure considerate sono comprese tra lo 0.4 e lo 0.6. In particolare γ assume un valore più alto di τ_b a causa della sua natura ovvero per la mancata considerazione delle coppie per cui le modalità di una delle due variabili rimangono costanti. Le correlazioni D di Somers sono discrete e in particolare mostrano un'associazione più forte se la soddisfazione della vita viene considerata come variabile dipendente. Tuttavia non può essere trascurata neanche l'altra e perciò sarebbe necessaria una maggior indagine nel caso in cui si volesse verificare una relazione di causalità tra le due sfere.

Tabella 3.11: Correlazioni tra la soddisfazione del reddito e la soddisfazione della vita

	γ	τ_b	d_R	d_V	<i>Policorica</i>
Correlazione	0.599	0.409	0.373	0.447	0.524
Standard deviation	0.012	0.009	0.009	0.010	0.011
p-value	0.000	0.00	0.000	0.000	0.000

3.4 Analisi multivariate

Questa Sezione è dedicata alla presentazione dei modelli statistici utilizzati per indagare la correlazione tra la soddisfazione del reddito e quella della vita e al tempo stesso i fattori che maggiormente influenzano tale relazione. La scelta è ricaduta in quei modelli che tengono conto della natura ordinale delle due domande d'interesse e che contemporaneamente correggono le possibili differenze esistenti nelle scale individuali di valutazione. Pertanto si sono prima di tutto aggiustate le autovalutazioni grazie alla tecnica delle *anchoring vignettes*: come si vedrà, ciò si è potuto fare solo per un sottocampione di ciascuna realtà d'interesse. Successivamente, si sono utilizzati dei modelli probit ordinali bivariati al fine di indagare la presenza di un'associazione tra le nuove variabili, in relazione ad alcuni fattori, ma anche tra quelle originarie. Infine, si è stimato un modello che rappresenta un'estensione di quello di selezione di Heckman per testare se la significatività positiva emersa dal modello probit ordinale sia dovuta ad una regola selettiva del campione ovvero se la valutazione riportata per la soddisfazione sia condizionata dall'aver rispettato le condizioni alla base della nuova variabile corretta da DIF.

3.4.1 Applicazione della tecnica delle *vignettes*

Nella Sezione 2.1 si è presentato il problema dell'incomparabilità delle risposte degli individui in quanto affette da eterogeneità (il cosiddetto DIF). Per tener conto di tale aspetto, un possibile strumento è costituito dall'approccio delle *anchoring vignettes*. In letteratura sono disponibili due tecniche

che godono di vantaggi e svantaggi propri, il modello Chopit (soluzione parametrica) e la scala C (soluzione non parametrica). Alla luce dell'obiettivo prefissatosi, la seconda proposta sembra essere più adeguata dal momento che non sfrutta nessun'altra informazione se non quella fornita dall'autovalutazione e dalle valutazioni agli individui ipotetici; in questo modo è possibile creare una nuova variabile libera da DIF e analizzarla attraverso un altro modello: così facendo si stimano direttamente gli effetti delle caratteristiche osservabili correggendo l'eterogeneità delle scale di risposta.

Per ogni rispondente del campione si hanno a disposizione due autovalutazioni e due valutazioni relative a due scenari ipotetici differenti, ma adeguati alla sfera d'indagine. Tuttavia, per poter creare la nuova variabile C è necessario fissare un ordinamento delle *vignettes* a disposizione: la soluzione più appropriata ricade in quello naturale dato dalla gravità della situazione ipotetica descritta, ma è inoltre possibile considerare quello consensuale degli intervistati. Le analisi descrittive relative alla valutazione della soddisfazione dei soggetti ipotetici (si vedano le Sezioni 3.3.2.1 e 3.3.2.2) sembrano fornire valutazioni più positive alle due persone che effettivamente hanno uno scenario più favorevole e migliore, ovvero Anna e Franca; ciò però è una evidenza generale e non assicura che un singolo individuo classifichi Giorgio (o Carlo) come più soddisfatto di Anna (o Franca). Per una maggior sicurezza si indaga meglio e in modo più "matematico" l'ordinamento consensuale degli intervistati: nello specifico, la proporzione dei rispondenti che attribuiscono ad Anna un livello di soddisfazione del reddito più alto rispetto a quello di Giorgio è pari all'84.8%, mentre quella relativa alla classificazione di Franca come persona più "felice" di Carlo è pari al 62.27%. Ciò si può verificare anche con le Tabelle 3.12 corrispondenti a quelle presentate nella Sezione 2.1.5: in entrambe le tabelle la differenza di proporzione $p_{12} - p_{21}$ (valore nella diagonale superiore) è positiva e ciò conferma quanto si è detto in poche righe precedenti. Chiaramente la tabella è molto più utile nel caso in cui il numero delle *anchoring vignettes* sia superiore a due.

Siano Y_i^R e Y_i^V rispettivamente la domanda sulla soddisfazione del reddito e sulla soddisfazione della vita per l'intervistato i e siano Z_{iG}^R , Z_{iA}^R , Z_{iC}^V e Z_{iF}^V le valutazioni di tale soggetto rispettivamente per l'individuo ipotetico

Soddisfazione reddito			Soddisfazione vita		
	<1 (Giorgio)	<2 (Anna)		<1 (Carlo)	<2 (Franca)
1	-	0.822	1	-	0.5364
2	0.127	-	2	0.2909	-

La proporzione dei casi in cui una *vignette* (in riga) è inferiore ad un'altra (in colonna)

Tabella 3.12: Errato ordinamento e presenza di *ties*

Giorgio, Anna, Carlo e Franca. Dal momento che $Z_{iG}^R < Z_{iA}^R$ e che $Z_{iC}^V < Z_{iF}^V \forall i, i = 1, 2, \dots, n$, le nuove variabili libere da DIF sono le seguenti:

$$C_i^R = \begin{cases} 1 & \text{se } Y_i^R < Z_{iG}^R \\ 2 & \text{se } Y_i^R = Z_{iG}^R \\ 3 & \text{se } Z_{iG}^R < Y_i^R < Z_{iA}^R \\ 4 & \text{se } Y_i^R = Z_{iA}^R \\ 5 & \text{se } Y_i^R > Z_{iA}^R \end{cases} \quad C_i^V = \begin{cases} 1 & \text{se } Y_i^V < Z_{iC}^V \\ 2 & \text{se } Y_i^V = Z_{iC}^V \\ 3 & \text{se } Z_{iC}^V < Y_i^V < Z_{iF}^V \\ 4 & \text{se } Y_i^V = Z_{iF}^V \\ 5 & \text{se } Y_i^V > Z_{iF}^V \end{cases}$$

Chiaramente ogni C ha esattamente 5 livelli (nel caso scalare), essendoci $J = 2$ *vignettes* per ciascuna valutazione. Tuttavia col dataset in esame non è possibile creare questa nuova variabile per tutti i 7353 cittadini del campione, senza ulteriori assunzioni o strumenti modellistici, a causa della presenza di risposte inconsistenti. Tale evidenza è la confermata al fatto che gli elementi in diagonale inferiore della Tabella 3.12 mostrano un valore (corrispondente a $1 - p_{12} - p_{21}$) che non è prossimo a 1. Nella Tabella 3.13 si possono osservare nel dettaglio tutte le 13 possibili relazioni di un'autovalutazione Y con due *vignettes*, Z_{i1} e Z_{i2} ove $Z_{i1} < Z_{i2}$: si noti in particolare che il specifico valore numerico associato a Y_i , Z_{i1} e Z_{i2} è irrilevante per definire la nuova variabile. Nelle ultime due colonne della tabella sono inoltre riportate le frequenze assolute delle nuove variabili scalari associate alle due soddisfazioni. Il termine "x" corrisponde alla cella in cui un possibile ordinamento soddisfa le relazioni della variabile C scalare, mentre la colonna "C_i" riporta le modalità associate alla variabile C vettoriale. Conseguentemente i primi 5 casi corrispondono ad un ordinamento corretto ed assenza di inconsistenze, col risultato che C è appunto scalare se tutti gli intervistati rispettano una delle 5 "regole" sul-

la relazione delle tre risposte date. Si hanno *ties* nei successivi casi, i quali producono valori censurati per C solo se la valutazione dell'autovalutazione è uguale a quelle delle *vignettes*. Gli ultimi 5 casi invece fanno riferimento a risposte che attribuiscono un ordinamento incorretto della gravità degli scenari ipotetici rispetto a quello naturale o imposto.

Tabella 3.13: Classificazione delle autovalutazioni e delle relative due *vignettes* di ciascun dominio

Ordinamento	Ordinamenti accettati dalla scala C scalare					C vettoriale		
	$Y_i < Z_{i1}$	$Y_i = Z_{i1}$	$Z_{i1} < Y_i < Z_{i2}$	$Y_i = Z_{i2}$	$Y_i < Z_{i2}$	C_i	C_i^R	C_i^V
$Y_i < Z_{i1} < Z_{i2}$	x					1	822	105
$Y_i = Z_{i1} < Z_{i2}$		x				2	1399	416
$Z_{i1} < Y_i < Z_{i2}$			x			3	1377	507
$Z_{i1} < Y_i = Z_{i2}$				x		4	1872	2143
$Z_{i1} < Z_{i2} < Y_i$					x	5	763	1408
$Y_i < Z_{i1} = Z_{i2}$	x					1	373	150
$Y_i = Z_{i1} = Z_{i2}$		x		x		2-4	315	695
$Z_{i1} = Z_{i2} < Y_i$					x	5	243	1294
$Y_i < Z_{i2} < Z_{i1}$	x				x	1	19	27
$Y_i = Z_{i2} < Z_{i1}$	x			x		1-4	24	65
$Z_{i2} < Y_i < Z_{i1}$	x				x	1-5	17	23
$Z_{i2} < Y_i = Z_{i1}$		x			x	2-5	59	241
$Z_{i2} < Z_{i1} < Y_i$					x	5	70	279

Come si può notare dalla Tabella 3.13, 6233 cittadini (84.77%) rispondono alla soddisfazione del reddito riconoscendo l'ordinamento naturale, mentre 4579 (62.27%) alla soddisfazione della vita. In particolare la distribuzione di queste variabili, riferite quindi solo ai valori scalari, evidenziano la presenza di una maggiore percentuale di individui che mostrano un livello di insoddisfazione (0.05% per l'autovalutazione originaria del reddito e 0.13% per la rispettiva variabile C ; 0.007% per l'autovalutazione originaria della vita e 0.02% per la rispettiva variabile C).

Uno dei vantaggi del metodo non parametrico è quello di poter interpretare la nuova variabile attraverso altri strumenti statistici, come ad esempio un semplice istogramma; tuttavia, uno dei problemi che può sorgere è la presenza e il relativo trattamento delle informazioni "non corrette". Nella Sezione 2.1.4.3 si sono viste alcune possibili soluzioni che vengono riportate nella Figura 3.17 per il dominio del reddito come esempio: le variabili utilizzate per l'allocazione secondo il metodo Probit ordinale censurato sono quelle

relative all'istruzione, all'età e al genere dell'individuo. Come si può notare i metodi mostrano tre distribuzioni sostanzialmente identiche e ciò non sorprende particolarmente dal momento che le frequenze assolute associate ai casi vettoriali non sono particolarmente elevate: pertanto portano alle stesse conclusioni. La differenza che si può riscontrare rispetto all'istogramma sull'autovalutazione originaria (Figura 3.8) è la presenza di un maggior numero di individui che mostrano un alto livello di insoddisfazione.

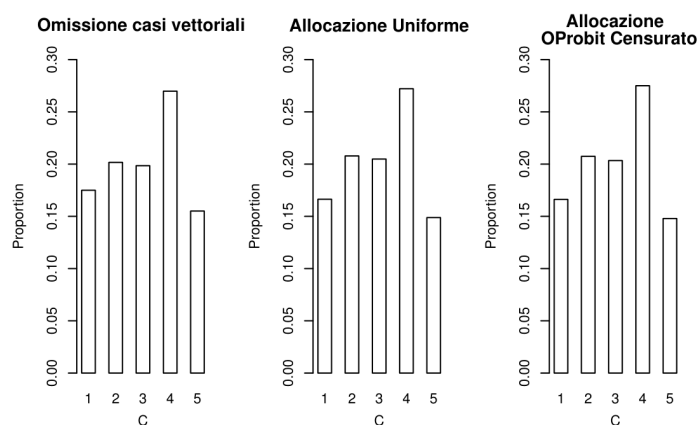


Figura 3.17: Iistogramma con 3 metodi per l'allocazione delle inconsistenze per il reddito

Nello specifico in questa tesi si sono utilizzate le variabili C scalari, secondo la proposta originale di King et al. (2004), seppur con la consapevolezza che lo sfruttamento di tale approccio comporti nel campione in esame una discreta perdita di informazioni, specialmente per quella sulla soddisfazione della vita. Tuttavia in questo modo non è necessario porre nessun'altra assunzione, oltre quelle basilari del metodo non parametrico. Poichè C_i^R e C_i^V sono due variabili che individuano due sottocampioni da quello di partenza, potrebbe sorgere il dubbio che gli intervistati abbiano distinte caratteristiche. A questo scopo, si definiscono con S^R e S^V le due variabili dummy che assegnano il valore 1 se l'individuo ha valutato la propria soddisfazione (rispettivamente del reddito e della vita) e quella delle due *vignettes* associate secondo l'ordinamento precedentemente individuato, ovvero se per questo rispondente la variabile C di una sfera d'interesse assume un valore del range 1-5:

$$S_i^R = \begin{cases} 1 & \text{se } C_i^R \in [1, 5] \\ 0 & \text{altrimenti} \end{cases} \quad e \quad S_i^V = \begin{cases} 1 & \text{se } C_i^V \in [1, 5] \\ 0 & \text{altrimenti} \end{cases}.$$

Nella Figura 3.18 e Figura 3.19 vengono riportate le autovalutazioni relative agli individui che hanno un livello della variabile C scalare ($S^R = 1$ o $S^V = 1$) e per quelli che invece non la presentano ($S^R = 0$ o $S^V = 0$): la distribuzione dell'autovalutazione dei due gruppi di soggetti che rispondono ad uno specifico dominio è sostanzialmente identica; in particolare coloro che rispondono rispettando l'ordinamento naturale sono leggermente più soddisfatti (48.04% contro 52.86% per il reddito e 75.92% contro 78.27% per la vita). Si noti inoltre come la distribuzione dei soggetti per i quali $S^R = 1$ sia sostanzialmente identica a quella del primo grafico della Figura 3.17, se non nelle frequenze.

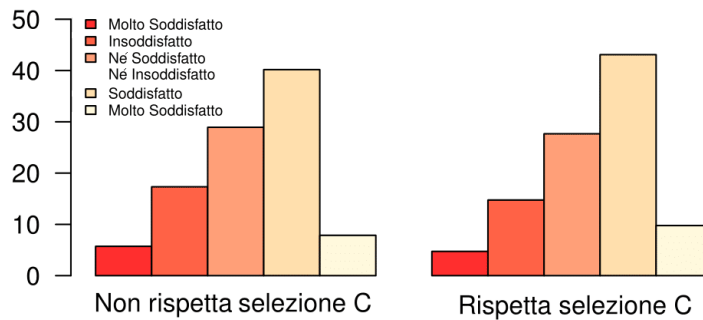


Figura 3.18: Autovalutazione del reddito in base alla selezione di C

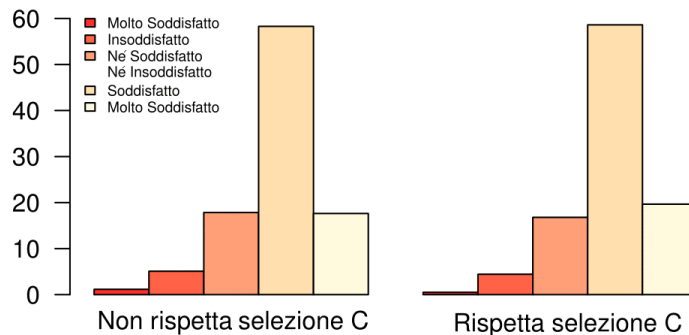


Figura 3.19: Autovalutazione della vita in base alla selezione di C

L'approccio non parametrico delle *anchoring vignettes*, come quello parametrico, si basa su due assunzioni, la *response consistency* e la *vignette equivalence*. In questa tesi si assumo essere verificate; infatti la mancanza di un indicatore oggettivo e di specificazioni non parametriche per saggiarle non consente di applicare una delle procedure teoriche sfruttate da altri autori e presentate nella Sezione 2.1.6. Ad ogni modo per la verifica di ciascuna di queste ci si può valere dei seguenti test "deboli": alcune evidenze riportate nella presentazione delle statistiche descrittive ovvero il fatto che intervistati che hanno caratteristiche più vicine agli individui ipotetici hanno simili valutazioni delle autovalutazioni e delle *vignettes* (Tabella 3.7 e 3.9) supportano la prima assunzione, mentre il fatto che più della metà degli intervistati rispettano l'ordinamento naturale (84.77% per il reddito, 62.27% per la vita) e che tale maggioranza viene accertata anche nei sottogruppi determinati dai singoli Paesi Europei (si veda la Tabella 3.14) sono a favore della *vignette equivalence*. Tuttavia l'eterogeneità tra i vari Paesi che la stessa Tabella 3.14 evidenzia fa capire come il tema della verifica di queste assunzioni non vada sottovalutata.

Nelle Figure 3.20 e 3.21 vengono riportate le distribuzioni dell'autovalutazione originaria e corretta dei due domini d'interesse per i singoli Paesi: una volta che si corregge il DIF si riscontra una maggior proporzione di individui che mostrano un alto livello di insoddisfazione (ad esempio si confrontino le distribuzioni relative alla Polonia e alla Repubblica Ceca). La distribuzione della variabile C mostra una percentuale più alta di cittadini con un livello di soddisfazione più elevato. Inoltre, se una volta che si sfrutta l'informazione delle *anchoring vignettes* emerge che la Grecia è il terzo Paese con una maggior proporzione di individui che mostrano un altro livello di soddisfazione del reddito, mentre la Danimarca è il quarto per quanto riguarda la soddisfazione della vita, diversamente da quanto si riscontra in letteratura. In particolare si noti come i due approcci evidenzino una diversa percentuale di individui francesi con un più alto livello di soddisfazione del benessere generale.

Alcuni risultati interessanti emergono a riguardi del numero di individui e nell'autovalutazione di ciascun Paese. Nella Tabella 3.14 le percentuali sono calcolate sul totale di cittadini di ciascun Stato del campione iniziale di 7353

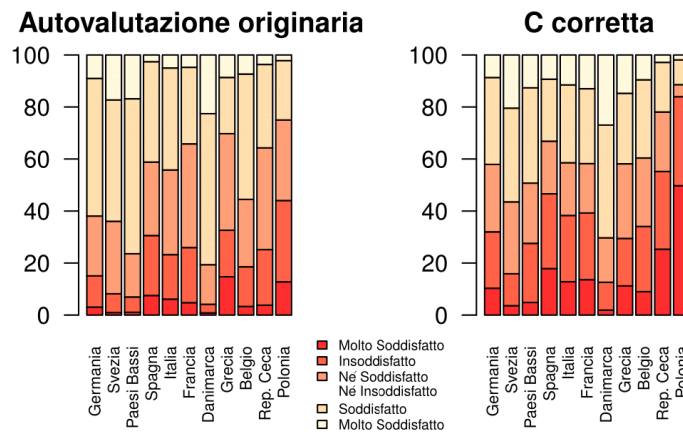


Figura 3.20: Autovalutazione del reddito per il sottocampione $S^R = 1$

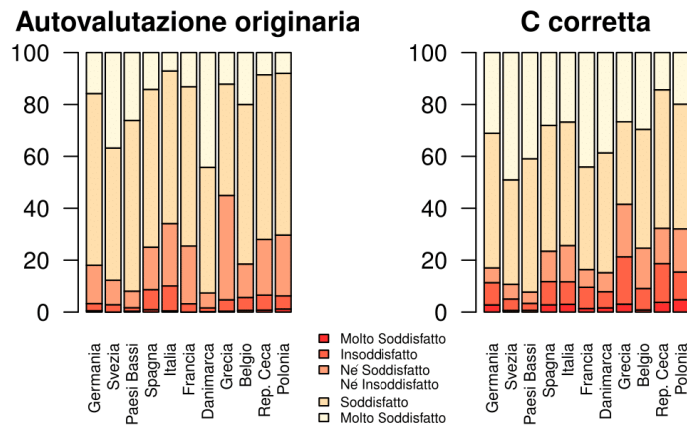


Figura 3.21: Autovalutazione della vita per il sottocampione $S^V = 1$

intervistati. I rispondenti tedeschi, che costituiscono il 15.14% del campione totale, si posizionano al quarto posto per la maggioranza di soggetti che non commettono inconsistenze nella classificazione di se stessi e degli individui ipotetici in riferimento alla soddisfazione del reddito, dopo Polonia, Spagna e Italia, mentre guadagnano il primo posto in riferimento alla soddisfazione della vita. Il Paese più abile nel riconoscimento dell'ordinamento naturale è la Grecia per la soddisfazione monetaria, mentre il Belgio per quella generale. Nella stessa Tabella sono riportate anche le soddisfazioni medie (nell'ipotesi di considerare il compiacimento soggettivo come una variabile numerica): in generale queste risultano essere leggermente maggiori per i Paesi i cui cittadini forniscono valutazioni "ordinabili" secondo una delle condizioni associate alle categorie della variabile C; tuttavia ciò non si verifica per Olanda e Danimarca, per quanto riguarda le risposte relative alla sfera del reddito.

Tabella 3.14: Proporzioni di individui tra i sottogruppi individuati dalla selezione della variabile C in ciascun Paese

	Reddito				Vita			
	$S^R = 0$		$S^R = 1$		$S^V = 0$		$S^V = 1$	
	%	Media	%	Media	%	Media	%	Media
Germania	17.16	3.50	82.84	3.53	47.71	3.92	52.29	3.94
Svezia	8.57	3.67	91.43	3.72	30.11	4.12	69.89	4.22
Olanda	5.94	3.9	94.06	3.85	40.99	4.10	59.01	4.16
Spagna	22.95	3.03	77.05	3.06	35.33	3.75	64.67	3.80
Italia	20.53	3.29	79.47	3.20	35.45	3.51	64.55	3.62
Francia	14.13	3	85.87	3.08	40.22	3.74	59.78	3.85
Danimarca	12.56	4.01	87.44	3.98	34.99	4.30	65.01	4.35
Grecia	6.81	2.54	93.19	2.92	45.59	3.56	54.41	3.62
Belgio	13.22	3.52	86.78	3.41	27.28	3.96	72.72	3.95
Repubblica Ceca	13.57	3.02	86.43	3.10	37.31	3.68	62.69	3.73
Polonia	31.85	2.69	68.15	2.70	37.59	3.68	62.41	3.71

Prima di procedere con le stime dei modelli, nella Tabella 3.15 si riportano alcune analisi descrittive dei due gruppi per ciascun costrutto d'interesse; nel complesso vengono rispettate le evidenze dell'intero campione e non vi sono particolari divergenze, se non di qualche punto percentuale, specialmente per la valutazione soggettiva della vita: infatti i soggetti sono in maggioranza di genere femminile, hanno un'istruzione di grado superiore a quello elementare,

hanno un partner, sono principalmente pensionati e soffrono di problemi di depressione in tutti e quattro i gruppi. Inoltre coloro che rispettano l'ordinamento naturale delle *anchoring vignettes* hanno mediamente 2 anni in meno, citano uno o due animali in più e chiedono meno chiarimenti all'intervistato dell'indagine CAPI.

Alcune differenze si notano nei due gruppi relativi alla sfera soggettiva del reddito: tra coloro le cui risposte rispecchiano l'ordinamento naturale vi è una percentuale più bassa di soggetti occupati (19.89%) e con un'istruzione alta (15.81%), mentre tra coloro per cui è possibile individuare un valore nel range della variabile corretta C vi è una proporzione inferiore di intervistati non occupati e con istruzione bassa; emerge inoltre una differenza di quasi 6 punti percentuali per quanto riguarda la presenza di soggetti depressi. Si precisa infine che non sono presenti particolari divergenze per le variabili esplicative qui non riportate, se non per quelle monetarie.

La Tabella 3.16 mostra le correlazione tra le due variabili corrette dall'eterogeneità delle diverse scale di risposte che possono applicare gli individui; queste misure fanno riferimento a quelle riportate nella Sezione 2.4: le due domande mantengono un'associazione medio-discreta anche una volta che sono corrette, dal momento che tutte le misure considerate hanno un valore compreso tra 0.191 e 0.284, seppur più basse di quelle calcolate in riferimento alle variabili originarie (Tabella 3.11). In particolare le correlazioni D di Somers riconfermano un'associazione più forte se la soddisfazione della vita viene considerata come variabile dipendente, anche se non può essere trascurata nemmeno l'altra. In generale comunque una volta che si corregge per la DIF si conferma l'esistenza di una correlazione tra le due sfere d'indagine, ma di magnitudo inferiore rispetto a quella evidenziata con le valutazioni dirette.

Sembra dunque che l'eterogeneità nelle scale di risposta amplifichi la correlazione esistente fra queste due misure di soddisfazione.

3.4.2 Stima dei modelli

Prima di procedere con la presentazione dei modelli stimati e dei relativi risultati, è bene avere chiara la composizione del campione in esame. Il data-

Tabella 3.15: Media, mediana e composizione di alcune variabili esplicative tra i due gruppi di individui selezionati dalla variabile C

	Reddito		Vita	
	$S^R = 0$	$S^R = 1$	$S^V = 0$	$S^V = 1$
N	1120	6233	2774	4579
<u>Genere</u>				
Donne	56.52%	55%	55.55%	55.03%
<u>Età</u>				
media	66.5	63.99	65.02	63.99
mediana	66	63	64	63
<u>Istruzione</u>				
bassa	39.64%	24.79%	28.05%	26.45%
media	44.55%	50.33%	51.33%	48.31%
alta	15.81%	24.88%	20.62%	25.25%
<u>Relazioni</u>				
Partner	74.24%	76.10%	74.07%	76.76%
<u>Stato occupazionale</u>				
Pensionato	55.45%	48.93%	50.80%	49.39%
Occupato	19.89%	31.86%	27.48%	31.59%
Non Lavora	24.66%	19.21%	21.72%	19.02%
<u>Stato salute</u>				
Depressione	80.27%	73.98%	74.15%	75.41%
<u>Fluency Test</u>				
media	17.27	19.72	18.80	19.68
mediana	17	19	18	19
<u>Indagine CAPI</u>				
Chiarimenti	64.79%	59.18%	61.77%	58.98%

Tabella 3.16: Correlazioni tra la soddisfazione del reddito e la soddisfazione della vita con le variabili corrette

	γ	τ_b	d_R	d_V	<i>Policorica</i>
Correlazione	0.287	0.208	0.191	0.225	0.284
Standard deviation	0.012	0.013	0.018	0.014	0.017
p-value	0.000	0.00	0.000	0.000	0.002

set sul quale si sta lavorando è costituito da 7353 cittadini europei, ognuno dei quali ha fornito una valutazione sulla soddisfazione di se stesso e dei soggetti protagonisti delle due rispettive *vignettes*. Il 7.11% di questi individui non è tuttavia in grado di fornire le tre valutazioni conformemente all'ordinamento generale su entrambi i domini d'indagine, al contrario del 54.15% (pari a 3982 individui). Ad ogni modo, un maggior numero di intervistati riesce ad individuare l'ordinamento naturale delle valutazioni associate al dominio del reddito ed a fornire giudizi coerenti con le assunzioni alla base del metodo non parametrico della tecnica delle *anchoring vignettes*: ciò potrebbe essere dovuto ad una maggior percezione e comprensione delle *vignettes* relative alla soddisfazione del reddito in quanto descrivono essenzialmente poche semplici caratteristiche immediate ovvero il nome, il genere, lo stato civile, il numero di figli e il reddito totale della famiglia. La Figura 3.22 riporta la presentazione grafica di quanto si è appena detto e la relativa frequenza assoluta di ciascun sottoinsieme.

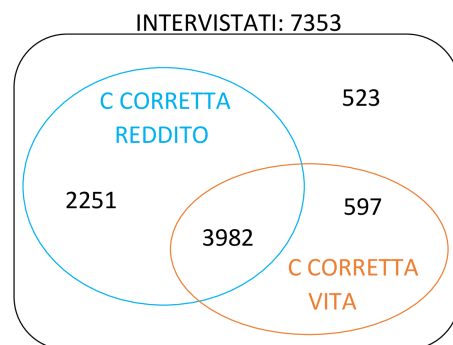


Figura 3.22: Rappresentazione grafica della composizione del campione

Con lo scopo di individuare e quantificare una possibile associazione tra la soddisfazione del reddito e quella della vita, si è stimato il modello probit ordinato bivariato sulle variabili originarie (quelle relative al giudizio diretto degli intervistati) e sulle variabili corrette (quelle "pulite" dell'eterogeneità delle scale di risposta, secondo la proposta di King et al. (2004)): tale modello infatti stima simultaneamente l'equazione specifica per la soddisfazione del reddito e per quella della vita attraverso il metodo della massima verosimiglianza, calcolando la correlazione ρ tra gli errori delle due equazioni. Come detto nella Sezione 2.2, le covariate usate sono le stesse per entrambe le

equazioni. La Tabella 3.17 mostra le stime dei parametri di questi modelli; in particolare, per ogni equazione, il numero di soglie è pari a quattro poiché le categorie delle risposte e quelle delle variabili C sono cinque: nel primo caso la prima soglia divide le categorie "Molto insoddisfatto e "Insoddisfatto", la seconda separa le categorie "Insoddisfatto" e "Né soddisfatto né insoddisfatto", la terza divide "Ne soddisfatto ne Insoddisfatto" e "Soddisfatto" e infine l'ultima distingue le due categorie di soddisfazione; la prima soglia relativa alle variabili corrette C separa il livello 1 dal 2 ovvero $Y_i^R < Z_{iG}^R$ da $Y_i^R = Z_{iG}^R$ se si considera quella relativa al reddito, la seconda divide il livello 2 dal 3 cioè $Y_i^R = Z_{iG}^R$ da $Z_{iG}^R < Y_i^R < Z_{iA}^R$, e così via per le altre due soglie. Inoltre il modello richiede che le soglie del modello probit ordinale bivariato siano le stesse per tutti i soggetti del campione, indipendentemente dalle loro caratteristiche.

Tabella 3.17: Modelli probit ordinato bivariati (campioni distinti)

	AUTOVALUTAZIONI		C CORRETTA	
	Reddito	Vita	Reddito	Vita
Donna	0.123*** (0.042)	0.217*** (0.044)	0.119** (0.056)	0.164*** (0.058)
Età_med	0.137*** (0.025)	0.119*** (0.027)	0.091*** (0.034)	0.170*** (0.035)
Età_med ²	0.045*** (0.013)	-0.011 (0.013)	0.041** (0.017)	0.020 (0.018)
Istruzione media	0.064* (0.035)	-0.026 (0.037)	0.180*** (0.049)	-0.021 (0.050)
Istruzione alta	0.318*** (0.043)	0.032 (0.046)	0.480*** (0.058)	0.057 (0.060)
Partner	0.129*** (0.038)	0.324*** (0.040)	0.093* (0.051)	0.331*** (0.052)
N componenti	-0.010 (0.015)	-0.059*** (0.016)	-0.015 (0.021)	-0.087*** (0.021)
N figli	-0.004 (0.011)	0.023** (0.011)	-0.017 (0.015)	0.014 (0.015)
Pensionato	0.107*** (0.041)	0.150*** (0.043)	0.079 (0.056)	0.137** (0.058)
Occupato	0.206*** (0.043)	0.184*** (0.045)	0.266*** (0.059)	0.229*** (0.060)
Genitore	0.035	0.017	0.049	-0.016

	(0.035)	(0.037)	(0.046)	(0.048)
Ricchezza finanziaria	0.027***	0.013***	0.022***	0.010***
	(0.002)	(0.002)	(0.003)	(0.003)
Ricchezza reale	0.030***	0.016***	0.029***	0.013**
	(0.004)	(0.004)	(0.005)	(0.005)
Depressione	-0.141***	-0.354***	-0.084**	-0.301***
	(0.032)	(0.034)	(0.042)	(0.045)
ADL	-0.035	-0.100***	0.028	-0.100***
	(0.028)	(0.029)	(0.039)	(0.039)
IADL	0.006	-0.006	0.009	0.010
	(0.024)	(0.025)	(0.036)	(0.036)
Croniche	-0.028**	-0.034***	-0.036**	-0.022
	(0.011)	(0.012)	(0.016)	(0.016)
Mobilità	-0.031***	-0.045***	-0.039***	-0.042***
	(0.010)	(0.010)	(0.014)	(0.014)
Sintomi	-0.061***	-0.089***	-0.016	-0.052***
	(0.010)	(0.011)	(0.014)	(0.014)
<i>Grip strength</i>	-0.001	0.001	-0.003	0.004
	(0.002)	(0.002)	(0.003)	(0.003)
<i>Fluency Test</i>	0.003	0.008***	0.006**	0.005*
	(0.002)	(0.002)	(0.003)	(0.003)
Svezia	0.188***	0.393***	0.347***	0.342***
	(0.064)	(0.068)	(0.082)	(0.088)
Olanda	0.395**	0.282***	0.110	0.306***
	(0.061)	(0.064)	(0.082)	(0.088)
Spagna	-0.382***	0.027	-0.136*	0.134*
	(0.067)	(0.070)	(0.092)	(0.096)
Italia	-0.200***	-0.421***	0.073	-0.021
	(0.059)	(0.062)	(0.082)	(0.086)
Francia	-0.501***	-0.170**	-0.140	0.337***
	(0.069)	(0.073)	(0.095)	(0.102)
Danimarca	0.486***	0.626***	0.538***	0.051
	(0.051)	(0.054)	(0.070)	(0.073)
Grecia	-0.533***	-0.428***	0.211**	-0.383***
	(0.065)	(0.067)	(0.091)	(0.094)
Belgio	-0.117**	0.089	-0.062	-0.064
	(0.052)	(0.055)	(0.069)	(0.073)
Repubblica Ceca	-0.347***	-0.231***	-0.643***	-0.382***
	(0.052)	(0.055)	(0.073)	(0.075)
Polonia	-0.496***	0.100***	-1.053***	0.076

	(0.062)	(0.066)	(0.093)	(0.092)
<i>soglia1</i>	-1.371*** (0.123)	-2.528*** (0.140)	-0.607*** (0.167)	-1.929*** (0.178)
<i>soglia2</i>	-0.440*** (0.122)	-1.553*** (0.130)	0.280* (0.166)	-0.968*** (0.172)
<i>soglia3</i>	0.516*** (0.122)	-0.519*** (0.128)	0.996*** (0.166)	-0.467*** (0.172)
<i>soglia4</i>	2.111*** (0.124)	1.411*** (0.129)	2.120*** (0.168)	0.946** (0.172)
<i>Osservazioni</i>	6920		3769	
ρ	0.390*** (0.013)		0.166*** (0.019)	
<i>Test LR di indipendenza</i>	$\chi_1^2=753.33$ p-value=0.000		$\chi_1^2=75.63$ p-value=0.000	

Standard error in parentesi: *** p<0.01, ** p<0.05, *p<0.1

Le stime dei coefficienti riportate permettono di determinare la direzione dell'effetto (inteso come associazione) di ciascuna variabile esplicativa sulla variabile dipendente, mentre non è possibile interpretare la loro magnitudo: un segno positivo di una stima corrisponde a una maggior probabilità di essere soddisfatta, mentre un coefficiente negativo indica una maggior insoddisfazione. La stime degli effetti marginali richiederebbero la specificazione di determinate assunzioni; tuttavia, non essendo obiettivo di questa tesi, si è preferito riportare solamente le stime originarie.

La prima stima da controllare in un modello probit ordinato bivariato è il coefficiente ρ : infatti un valore significativo e differente da 0 indica l'esistenza di un'associazione tra gli errori delle due equazioni e dunque tra le due variabili d'interesse. I test rapporto di verosimiglianza (LR) dei due modelli sono decisamente molto più grandi del valore critico corrispondente al livello di significatività dell'1% e ciò porta dunque a rifiutare l'ipotesi nulla di assenza di correlazione ρ in entrambi i modelli ($\chi_1^2=753.33$, $\chi_1^2=75.63$ rispettivamente per l'autovalutazione originaria e per quella corretta): nello specifico, si individua una discreta correlazione, prossima allo 0.4, per il modello relativo alle variabili originarie della soddisfazione, mentre questa risulta essere decisamente più bassa (pari a 0.166) per il modello riferito alle variabili corrette. Queste evidenze sono in linea con quanto era emerso dalle diverse misure

utilizzate per calcolare l'associazione di queste due realtà.

In generale la significatività e il segno dell'effetto si mantengono uguali per la variabile d'interesse di uno stesso dominio, in tutti e due i modelli; alcune differenze invece si riscontrano nella magnitudo dell'effetto, specialmente per le variabili relative al genere, allo status occupazionale e alla depressione: la prima tende a diminuire, mentre tutte le altre si intensificano, indicando rispettivamente il minor e il maggior impatto che hanno sulle soddisfazioni. In particolare il fatto di essere pensionato rispetto a quello di non essere occupato diventa statisticamente non significativo se si considera la variabile C corretta del reddito, mentre perde di significatività (passa dall'1% al 5%) e diminuisce in termini di magnitudo se si usa quella della vita. Le quantità monetarie, il numero di limitazioni ADL e IADL e i problemi di mobilità hanno un impatto simili o identico in entrambi i modelli e per ambedue le sfere d'interesse, ovvero le stime dei coefficienti sono quasi identici per significatività, segno e magnitudo. Diversamente, le divergenze sostanziali sulla significatività si presentano per le stime specifiche dei singoli Paesi.

Fattori come l'età, l'istruzione, il genere, la salute, lo stato occupazionale e le relazioni confermano la forte associazione con la soddisfazione della vita e quella del reddito, a parità di altre condizioni, come già emerso nelle Sezioni 1.2 e 1.3: ad esempio, c'è una significativa differenza nella soddisfazione tra persone di diversa età, *ceteris paribus*, ovvero al crescere dell'età l'individuo è più soddisfatto; in particolare questa caratteristica demografica ha un legame parabolico e significativo con la soddisfazione del reddito, mentre è lineare per quella della vita: quindi nel primo caso la soddisfazione tende a decresce fino ad una certa età (48 anni) e aumenta con l'avanzare dell'anzianità, mentre nel secondo cresce proporzionalmente con l'età. In particolare, confrontando i risultati del modello che considera la variabile originaria con quelli della variabile corretta, si può osservare come diminuisca e aumenti l'impatto del coefficiente dell'età nel secondo caso rispettivamente per il costrutto d'interesse del reddito e quello della vita. La relazione lineare dell'età con la soddisfazione della vita non trova supporto nel lavoro di Angelini et al. (2014) e di altri studi, i quali trovano evidenza di una relazione parabolica. Se una persona vive con un partner è significativamente più soddisfatta di un intervistato con le stesse caratteristiche, ma che non vive con un com-

pagno/a: in particolare, questa relazione è statisticamente meno significativa per l'equazione relativa alla variabile corretta del reddito (al livello del 10%), mentre all'1% per le altre 3. La soddisfazione è inoltre più elevata per coloro che lavorano o che sono in pensione, rispetto a chi non ha un'occupazione al momento dell'intervista, *ceteris paribus*. In particolare, l'impatto di vivere con un compagno o di essere pensionato perde di significatività nella regressione relativa alla variabile corretta del reddito, mentre il fatto di non vivere solo e di avere un lavoro aumentano lievemente nella regressione specifica per la variabile C della vita, a parità di altre condizioni.

In disaccordo coi risultati presentati da Angelini et al. (2014), la soddisfazione della vita aumenta significativamente se si è di genere femminile (valori più bassi nel modello che sfrutta il metodo non parametrico delle *anchoring vignettes*), mentre non risulta essere statisticamente significativo il fatto di avere almeno uno dei genitori in vita al momento dell'intervista. Queste identiche conclusioni si confermano anche per la soddisfazione della componente monetaria del benessere.

In accordo con quanto emerge dal lavoro di Bonsang e Van Soest (2012), la soddisfazione del reddito diminuisce al crescere del numero di sintomi e di malattie croniche che un individuo ha e se si è di genere femminile, mentre (sorprensamente) non si trova riscontro di un'influenza del numero di persone all'interno del nucleo familiare, diversamente da quanto emerge dal modello Chopit di questi autori. Gli impatti delle due variabili relative allo stato di salute sostanzialmente non variano per il modello che considera le variabili corrette, anche se il numero di sintomi perde la sua significatività. Le stime positive e statisticamente significative relative alla ricchezza finanziaria e reale non sorprendono particolarmente: infatti la soddisfazione aumenta se si dispongono di più risorse monetarie, come evidenziato da altri autori (per queste variabili gli impatti sono praticamente identici nelle due equazioni dello stesso dominio, a parità di altre condizioni). Anche gli effetti delle variabili relative allo stato di salute sono in accordo con quanto ci si potrebbe aspettare: la soddisfazione diminuisce se si gode (o dichiara) una condizione peggiore. In particolare il coefficiente relativo al numero di malattie croniche e quello del numero di sintomi perdono completamente la loro significatività rispettivamente sulla soddisfazione del reddito e della vita

se si considera il modello con le variabili C . Significatività opposte invece si ottengono per le capacità cognitive: la soddisfazione del reddito aumenta e diventa significativa se si considera l'equazione di regressione su C^R , mentre quella della vita cala al 10% per l'equazione che si basa sulla C^V .

I risultati più sorprendenti emergono per le variabili relative ai Paesi: infatti rispetto alla Germania e a parità di altre condizioni, per quanto riguarda il modello relativo alle autovalutazioni essere cittadini di Svezia, Olanda e Danimarca porta ad una maggiore soddisfazione del reddito (come testimoniano Bonsang e Van Soest (2012)), mentre se si abita in tutti gli altri Paesi ad una minore (specie se si è francesi o greci); la direzione del segno dei coefficienti dei Paesi europei si conferma anche per le autovalutazioni della vita, ad eccezione dei cittadini polacchi (sono più soddisfatti di quelle tedeschi, evidenza che non era emersa dalle statistiche descrittive), ma non si può dire altrettanto della significatività: infatti essere cittadini spagnoli o belgi non ha rilevanza rispetto ad essere tedeschi, *ceteris paribus*. Uno scenario quasi completamente differente si presenta quando se si considera il modello con le variabili dipendenti corrette, dove la classifica dei Paesi cambia: infatti una volta che le autovalutazioni vengono corrette dall'eterogeneità di scala si riscontra che i soli intervistati statisticamente significativi in entrambi i domini sono gli svedesi (hanno una soddisfazione positiva e quasi identica sia per il reddito che per la vita, a parità di altre condizioni), i cechi (leggermente più insoddisfatti del reddito che della vita, rispetto alla Germania e a parità di altre condizioni) e i greci (più soddisfatti della propria ricchezza, meno del loro benessere). Il risultato più sorprendente è associato alla soddisfazione del reddito di quest'ultimi cittadini: infatti il modello relativo alle variabili originali evidenziava un'influenza negativa sulla "felicità" economica dei soggetti greci rispetto a quelli tedeschi (-0.533), mentre una volta che si corregge per il DIF emerge un'evidenza opposta, ovvero fortemente positiva (0.211). Comunque, essere spagnoli, italiani o belgi non è statisticamente significativo rispetto ad essere di nazionalità tedesca in entrambe le sfere d'interesse. Ad ogni modo, la Danimarca si riconferma essere il Paese più soddisfatta del proprio reddito, mentre la Svezia è quello più "felice", *ceteris paribus*, a conferma di quanto detto nella Sezione 1.3 e 1.2.

Ricapitolando quindi, se si stima il modello probit ordinale bivariato sulle

autovalutazioni si rischia di interpretare erroneamente alcune evidenze e di distorcere le conclusioni: infatti nella stima relativa alle variabili corrette si trova evidenza di una correlazione più bassa tra le due grandezze d'interesse, impatti più intensificati di alcuni fattori come il fatto di avere un lavoro, il grado di istruzione, il fatto di essere depresso, il numero di sintomi riportati ed effetti non significativi di altre variabili come il numero di figli sulla soddisfazione della vita o le maggiori capacità cognitive sulla soddisfazione del reddito. Vi sono invece effetti identici per alcune variabili come quelle monetarie e quella associata alla dimensione del nucleo familiare. Le maggiori differenze in termini di significatività si riscontrano per le dummy riferite ai Paesi Europei: l'errore in cui ci si imbatterebbe in una regressione sulle variabili dipendenti originali è l'affermazione che, ad esempio, la Danimarca sia il Paese più soddisfatto della propria vita a parità di altre condizioni, constatazione che però viene smentita dal modello con le variabili C e che non rispecchia quanto noto in letteratura. Proprio questi risultati sulle stime delle dummy di Paese evidenziano l'importanza di correggere per il DIF se si vogliono confrontare misure di autovalutazioni di cittadini provenienti da Stati diversi; ci sono differenze in termini di culture, *welfare*, sistemi sanitari, ecc che inevitabilmente si riflettono sugli stili di risposta dei cittadini di ciascun Paese.

Constatato l'esito significativo del coefficiente ρ nei modelli precedenti e soprattutto le differenze in termini di significatività statistica che sono emerse in diverse situazioni, ci si chiede se tale risultati siano in qualche modo influenzati da un problema di osservazioni ovvero se la discreta significatività della correlazione tra la soddisfazione del reddito e quella della vita, derivata dal modello sulle variabili originali, sia dovuta al diverso sottocampione di individui considerati nei due modelli. Infatti, il modello precedente relativo alle autovalutazioni originarie non coinvolge esclusivamente coloro che riconoscono l'ordinamento naturale in entrambi i domini d'indagine (l'intersezione dell'insieme arancione e quello azzurro della Figura 3.22), bensì anche quei intervistati per i quali è stato possibile costruire solo una o nessuna delle due variabili C . Nella Tabella 3.18 vengono quindi riportati i risultati dei modelli probit ordinati bivariati per il campione ristretto di 3982 osservazioni (in

realtà 3769 a causa di alcuni valori mancanti); chiaramente i risultati relativi alla variabile corretta C sono identici a quelli precedenti.

Tabella 3.18: Modelli probit ordinati bivariati (campioni comuni)

	AUTOVALUTAZIONI		C CORRETTA	
	Reddito	Vita	Reddito	Vita
Donna	0.125** (0.057)	0.202*** (0.060)	0.119** (0.056)	0.164*** (0.058)
Età_med	0.163*** (0.034)	0.106*** (0.036)	0.091*** (0.034)	0.170*** (0.035)
Età_med ²	0.056*** (0.017)	0.002 (0.018)	0.041** (0.017)	0.020 (0.018)
Istruzione media	0.080 (0.049)	-0.035 (0.052)	0.180*** (0.049)	-0.021 (0.050)
Istruzione alta	0.331*** (0.059)	-0.015 (0.062)	0.480*** (0.058)	0.057 (0.060)
Partner	0.145** (0.051)	0.365*** (0.054)	0.093* (0.051)	0.331*** (0.052)
N componenti	-0.019 (0.021)	-0.100*** (0.022)	-0.015 (0.021)	-0.087*** (0.021)
N figli	0.005 (0.015)	0.034** (0.016)	-0.017 (0.015)	0.014 (0.015)
Pensionato	0.073 (0.057)	0.147** (0.060)	0.079 (0.056)	0.137** (0.058)
Occupato	0.222*** (0.059)	0.189*** (0.063)	0.266*** (0.059)	0.229*** (0.060)
Genitore	0.019 (0.047)	-0.016 (0.049)	0.049 (0.046)	-0.016 (0.048)
Ricchezza finanziaria	0.026*** (0.003)	0.011*** (0.003)	0.022*** (0.003)	0.010*** (0.003)
Ricchezza reale	0.037*** (0.005)	0.014*** (0.005)	0.029*** (0.005)	0.013** (0.005)
Depressione	-0.136*** (0.043)	-0.423*** (0.047)	-0.084** (0.042)	-0.301*** (0.045)
ADL	0.003*** (0.039)	-0.111*** (0.040)	0.028 (0.039)	-0.100*** (0.039)
IADL	0.002 (0.036)	-0.034 (0.037)	0.009 (0.036)	0.010 (0.036)
Croniche	-0.018**	-0.033***	-0.036**	-0.022

	(0.016)	(0.017)	(0.016)	(0.016)
Mobilità	-0.053***	-0.058***	-0.039***	-0.042***
	(0.014)	(0.015)	(0.014)	(0.014)
Sintomi	-0.058***	-0.086***	-0.016	-0.052***
	(0.014)	(0.015)	(0.014)	(0.014)
<i>Grip strength</i>	-0.002	0.001	-0.003	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
<i>Fluency Test</i>	0.005	0.007**	0.006**	0.005*
	(0.003)	(0.003)	(0.003)	(0.003)
Svezia	0.175**	0.356***	0.347***	0.342***
	(0.085)	(0.090)	(0.082)	(0.088)
Olanda	0.369**	0.270***	0.110	0.306***
	(0.086)	(0.090)	(0.082)	(0.088)
Spagna	-0.446***	-0.262	-0.136	0.134
	(0.094)	(0.100)	(0.092)	(0.096)
Italia	-0.256***	-0.376***	0.073	-0.021
	(0.084)	(0.089)	(0.082)	(0.086)
Francia	-0.650***	-0.185*	-0.140	0.337***
	(0.097)	(0.103)	(0.095)	(0.102)
Danimarca	0.413***	0.599***	0.538***	0.051
	(0.072)	(0.076)	(0.070)	(0.073)
Grecia	-0.544***	-0.498***	0.211**	-0.383***
	(0.093)	(0.097)	(0.091)	(0.094)
Belgio	-0.202***	0.035	-0.062	-0.064
	(0.072)	(0.076)	(0.069)	(0.073)
Repubblica Ceca	-0.426***	-0.238***	-0.643***	-0.382***
	(0.074)	(0.079)	(0.073)	(0.075)
Polonia	-0.556***	0.078	-1.053***	0.076
	(0.090)	(0.095)	(0.093)	(0.092)
<i>soglia1</i>	-1.352***	-2.990***	-0.607***	-1.929***
	(0.170)	(0.205)	(0.167)	(0.178)
<i>soglia2</i>	-0.406***	-1.846***	0.280*	-0.968***
	(0.169)	(0.181)	(0.166)	(0.172)
<i>soglia3</i>	0.536***	-0.784***	0.996***	-0.467***
	(0.169)	(0.178)	(0.166)	(0.172)
<i>soglia4</i>	2.151***	1.184***	2.120***	0.946**
	(0.171)	(0.178)	(0.168)	(0.172)
<i>Osservazioni</i>		3769		3769
ρ		0.379***		0.165***
		(0.018)		(0.019)

<i>Test LR di indipendenza</i>	$\chi^2_1=379.17$ p-value=0.000	$\chi^2_1=75.63$ p-value=0.000
------------------------------------	---------------------------------	--------------------------------

Standard error in parentesi: *** p<0.01, ** p<0.05, *p<0.1

Seppur si sia stimato un modello probit ordinato bivariato sui soli intervistati per i quali è stato possibile creare le variabili corrette, la stima del coefficiente ρ nel caso delle autovalutazioni originarie rimane sostanzialmente uguale (0.379 vs 0.390). Il p-value associato al test rapporto di verosimiglianza, che verifica sotto H_0 l'ipotesi che non vi sia correlazione tra le due variabili d'interesse, porta da affermare al livello dell'1% che non si dovrebbero regredire le due equazioni separatamente come due specificazioni probit ordinati univariati. Pertanto, sebbene si stia analizzando un sottocampione del precedente, l'utilizzo delle variabili originali della soddisfazione mette in luce la stessa discreta correlazione tra i due costrutti del benessere individuale, un'associazione che a quanto pare non dipende dal fatto che alcuni soggetti riconoscono o meno l'ordinamento delle *vignettes*.

Per quanto riguarda invece la stima dei coefficienti non vi sono particolari differenze in termini di significatività e di magnitudo: in generale quelli significativi cambiano dalla seconda cifra decimale. Tuttavia in questo modello avere un'educazione media, anzichè bassa, essere pensionato, piuttosto che non occupato e infine il fatto di soffrire di qualche malattia cronica non risultano essere significativi statisticamente per quanto riguarda la soddisfazione del reddito; queste caratteristiche quindi hanno lo stesso impatto non rilevante di quello individuato dalla stessa equazione in riferimento alla variabile corretta C, ad eccezione di quella relativa allo stato di salute che invece è significativa (più alto è il numero di malattie croniche che un individuo dichiara di avere, più è insoddisfatto del suo reddito). L'unica caratteristica che invece acquisisce significatività al livello dell'1% è il numero di limitazioni ADL: tenendo fissi altri fattori, un soggetto è più compiaciuto del proprio reddito se ha un maggior numero di limitazioni; si tratta di un'evidenza inaspettata, ma che potrebbe trovare un riscontro ragionevole se questi soggetti, ad esempio, percepissero qualche incentivo a causa di questa condizione. Opposto invece è il giudizio sulla soddisfazione della vita: infatti una persona che affronta dei problemi nelle attività di tutti i giorni è meno "felice", *ceteris*

paribus; in particolare si noti come il valore numerico di tale coefficiente sia molto vicino a quello individuato dalla regressione sulla variabile corretta.

Rispetto al precedente, in questo modello aver un lavoro regolare ed avere un partner hanno un'influenza leggermente maggiori sulla soddisfazione del reddito e della vita. Per quanto riguarda le stime dei coefficienti associati all'età questi mantengono la stessa significatività esplorata nelle stime dei modelli della Tabella 3.17: sebbene la stima dei coefficienti dell'età aumentino un po' in magnitudo, l'età minima risulta essere invariata per la soddisfazione del reddito; al contrario, questi soggetti tendono ad essere meno soddisfatti rispetto a individui della stessa età, ma dell'intero campione.

Per quanto riguarda le stime dei coefficienti relativi ai Paesi si ha un incremento in termini assoluti per quelli negativi nell'equazione relativa al reddito, mentre si ha un decremento sempre in termini assoluti di quasi tutti quelli presenti nell'equazione della vita, sia quelli col segno positivo che col segno negativo, evidenziando dunque maggior differenze degli impatti rispetto a quelli individuati correggendo le variabili d'interesse dalla distorsione DIF. Pertanto, anche se si considera il campione ristretto di individui, il fatto di essere cittadini di nazionalità, ad esempio, greca porta ad una minor soddisfazione del reddito rispetto ai tedeschi se si considera la variabile originale, mentre se si utilizza quella corretta, si giunge alla conclusione opposta.

Riassumendo, il modello appena stimato non porta a notevoli divergenze rispetto al precedente, indice del fatto che le differenze emerse tra i due differenti approcci (ovvero utilizzo della variabile corretta e di quella originale delle due soddisfazioni) non sono dovute a un problema di osservazioni, ovvero al diverso sottocampione che si analizza: il coefficiente di correlazione tra le due variabili d'interesse originarie si mantiene discreto e gli impatti delle covariate considerate rimangono sostanzialmente uguali, a conferma di quanto emergeva dalle statistiche descrittive (in media il sottocampione non si distingue significativamente da quello complessivo).

Di fronte all'evidenza che la correlazione individuata dalle variabili libere dal DIF tra la soddisfazione della vita e quella del reddito sia più bassa rispetto a quella ottenuta con l'uso delle originarie e che tale differenza non sia dovuta ad un problema di osservazioni dal momento che la correlazione

permane moderata e sostanzialmente indifferente, ci si pone un ulteriore problema: la ricerca di tale associazione e dei fattori che possono influenzare i due domini soggettivi tra i soggetti per i quali è stato possibile creare le variabili C deve avvenire nella consapevolezza che ciò è possibile solo se la valutazione della soddisfazione (originaria o corretta) è "preceduta" necessariamente dalla capacità di aver riconosciuto correttamente l'ordinamento stesso tra le *anchoring vignettes*. Per stimare in modo esatto l'associazione tra le due soddisfazioni bisogna quindi tenere a mente il fatto che si stanno considerando i soli soggetti per i quali è stato possibile creare le nuove variabili e, di conseguenza, tener conto della selezione che può derivarne nel campione, in modo da non avere stime distorte. In quest'ottica è perciò adeguato l'utilizzo di un modello di selezione, che si basa su due equazioni: quella di selezione che nel caso in esame deve analizzare una variabile dummy che attribuisce il valore 1 se l'individuo i ha riconosciuto l'ordinamento delle *vignettes* delle due sfere di interesse (ovvero $S_i^R = 1$ e/o $S_i^V = 1$) e quella principale che dovrebbe studiare il compiacimento soggettivo del reddito e della vita e quindi della loro associazione. Tuttavia, prima di esaminare congiuntamente le due soddisfazioni si verifica se la selezione esista a livello univariato ovvero considerando solo le singole soddisfazioni. Alla luce di ciò, nelle successive due tabelle si riportano i risultati relativi al modello di Heckman adattato per variabili ordinali: l'osservazione dunque dell'autovalutazione del reddito (vita), o della relativa variabile C, avviene solo se si è risposto alla domanda d'interesse riconoscendo l'ordinamento delle *vignettes* ovvero se $S_i^R = 1$ ($S_i^V = 1$). La Tabella 3.19 riporta le stime di tale modello per il dominio del reddito e la Tabella 3.20 quelle relative alla vita, nei quali si son considerate le variabili "Chiarimenti" e ed età (attraverso "Tipo_b") come covariate aggiuntive dell'equazione di selezione: si tratta di due variabili legate all'indagine CAPI e non direttamente alla capacità di individuare l'ordinamento corretto degli scenari ipotetici. Le variabili esplicative dell'equazione di principale e le restanti covariate dell'equazione di selezione sono le stesse dei modelli precedenti. Si precisa che anche in questo caso non sono stati riportati effetti marginali, per cui è possibile interpretare le stime dei coefficienti solo nella direzione dell'effetto di ciascuna covariata sulla variabile dipendente, mentre non è possibile interpretare la loro magnitudo.

Tabella 3.19: Modelli Heckman per dati ordinali per il dominio del reddito

	Risponde correttamente	Auto- valutazione	Risponde correttamente	C corretta
Donna	0.052 (0.060)	0.147*** (0.046)	0.054 (0.061)	0.118*** (0.045)
Età_med	-0.040 (0.042)	0.130*** (0.027)	-0.030 (0.044)	0.077*** (0.027)
Età_med ²	-0.048*** (0.050)	0.038*** (0.039)	-0.051*** (0.050)	0.033** (0.043)
Istruzione media	0.291*** (0.050)	0.404*** (0.039)	0.294*** (0.050)	0.540*** (0.043)
Istruzione alta	0.291*** (0.063)	0.404*** (0.046)	0.294*** (0.064)	0.540*** (0.053)
Partner	-0.076 (0.054)	0.111*** (0.040)	-0.097* (0.055)	0.099** (0.042)
N componenti	-0.014 (0.022)	-0.016 (0.017)	0.023 (0.022)	-0.012 (0.017)
N figli	-0.017 (0.015)	-0.002 (0.011)	-0.018 (0.015)	-0.016 (0.012)
Pensionato	0.078 (0.058)	0.110** (0.045)	0.105* (0.059)	0.117*** (0.045)
Occupato	0.114* (0.063)	0.215** (0.217)	0.146** (0.064)	0.248*** (0.047)
Genitore	0.042 (0.052)	0.035 (0.036)	0.035 (0.054)	0.030 (0.037)
Ricchezza finanziaria	-0.008** (0.004)	0.023*** (0.002)	-0.007** (0.003)	0.022*** (0.002)
Ricchezza reale	0.012** (0.005)	0.034*** (0.004)	0.013** (0.005)	0.024*** (0.004)
Depressione	-0.051 (0.048)	-0.157*** (0.033)	-0.086* (0.049)	-0.062* (0.034)
ADL	0.054 (0.039)	-0.110 (0.031)	0.062 (0.039)	0.008 (0.032)
IADL	0.043 (0.032)	0.013 (0.026)	0.031 (0.033)	-0.003 (0.029)
Croniche	-0.013 (0.016)	-0.031* (0.012)	-0.012 (0.016)	-0.040*** (0.012)
Mobilità	-0.039*** (0.013)	-0.049*** (0.010)	-0.040*** (0.014)	-0.024** (0.011)
Sintomi	0.047***	-0.032***	0.050***	-0.024**

	(0.015)	(0.011)	(0.015)	(0.012)
<i>Grip strength</i>	0.002	0.001	0.002	-0.002
	(0.003)	(0.002)	(0.003)	(0.002)
<i>Fluency Test</i>	0.013***	0.006***	0.012***	0.006**
	(0.003)	(0.002)	(0.003)	(0.002)
Svezia	0.472**	0.289***	0.494***	0.434***
	(0.101)	(0.067)	(0.103)	(0.074)
Olanda	0.670**	0.513***	0.720***	0.167**
	(0.107)	(0.064)	(0.109)	(0.075)
Spagna	0.058	-0.300***	0.072	0.023
	(0.091)	(0.073)	(0.092)	(0.073)
Italia	0.054	-0.178***	0.053	0.206***
	(0.082)	(0.064)	(0.083)	(0.065)
Francia	0.196*	-0.402***	0.219**	0.056
	(0.100)	(0.076)	(0.101)	(0.075)
Danimarca	0.215***	0.482***	0.231***	0.614***
	(0.073)	(0.054)	(0.074)	(0.058)
Grecia	0.917***	-0.251***	0.784***	0.327***
	(0.110)	(0.069)	(0.110)	(0.084)
Belgio	0.247***	-0.059	0.281***	0.017
	(0.075)	(0.057)	(0.076)	(0.058)
Repubblica Ceca	0.161***	-0.240***	0.225***	-0.443***
	(0.076)	(0.061)	(0.076)	(0.057)
Polonia	-0.342***	-0.538***	-0.337***	-0.869***
	(0.082)	(0.069)	(0.083)	(0.085)
Chiarimenti	-0.037**		-0.054***	
	(0.019)		(0.021)	
Tipo_b	-0.052		-0.337***	-0.036
	(0.060)		(0.065)	
costante	0.426**		0.462**	
	(0.190)		(0.192)	
<i>soglia1</i>		-0.676***		-0.500***
		(0.165)		(0.170)
<i>soglia2</i>		0.127		0.401**
		(0.153)		(0.174)
<i>soglia3</i>		0.966***		1.069***
		(0.149)		(0.177)
<i>soglia4</i>		2.467***		2.185***
		(0.149)		(0.183)
<i>Osservazioni</i>	6920	5873	6920	5873

ρ	0.739*** (0.125)	-0.098 (0.208)
<i>Test LR:</i> $H_0 = \rho = 0$	$\chi_1^2=15.84$ p-value=0.000	$\chi_1^2=0.19$ p-value=0.667

Standard error in parentesi: *** p<0.01, ** p<0.05, *p<0.1

L'elemento fondamentale per capire l'adeguatezza di tale modello è la significatività del coefficiente ρ ovvero del coefficiente di correlazione tra il termine di errore dell'equazione principale e quello di selezione: infatti un valore statisticamente significativo e differente da 0 indica l'esistenza di selezione da campione o in altre parole la necessità di impiegare accanto al modello relativo all'equazione principale (nel caso in esame un modello probit ordinato), un modello che esprima la dipendenza tra questo e la selezione che lo ha in qualche modo generato. Per verificare la significatività di tale dipendenza e dunque l'adeguatezza del modello, è possibile utilizzare un test rapporto di verosimiglianza che saggi l'ipotesi di nullità della correlazione tra gli errori delle due equazioni considerate ($H_0 : \rho = 0$). I test LR dei due modelli mostrano risultati differenti: per il modello che regredisce l'autovalutazione nell'equazione principale si ha un test LR più grande del valore critico corrispondente al livello di significatività dell'1% ($\chi_1^2=15.84$) e ciò comporta il rifiuto dell'ipotesi nulla di assenza di correlazione ρ , mentre per il modello che ha come variabile dipendente la variabile corretta C si accetta l'ipotesi e dunque l'assenza di correlazione: conseguentemente, è sufficiente considerare un semplice modello probit ordinato in quanto non vi è selezione del campione. Ciò vuol dire che la variabile corretta "annulla" la selezione, evidenza che non risulterebbe se si considerasse la variabile originaria e che porterebbe a concludere l'esistenza di una forte selezione del campione ($\rho = 0.739$).

I risultati che emergono dalla stima del modello di selezione che considera come variabile d'interesse l'autovalutazione mostrano che gli unici fattori in grado di influenzare la capacità di individuare correttamente l'ordinamento delle *vignettes*, a parità di altre condizioni, sono un'istruzione superiore a quella primaria, il possesso di ricchezza reale, il numero di animali conosciuti, l'età, la ricchezza finanziaria, il numero di limitazioni nella mobilità e il fatto

di chiedere almeno un chiarimento⁶. Nello specifico le prime tre caratteristiche inducono una maggior propensione di riconoscere esattamente l'ordinamento degli scenari ipotetici, mentre le restanti una minore probabilità, a parità di altre condizioni. Evidenze interessanti emergono dai coefficienti relativi ai Paesi Europei: il fatto di essere italiano o spagnolo non ha alcun impatto, rispetto a quello di essere tedeschi, mentre quello di essere polacco comporta una minore propensione di individuare l'esatto ordinamento, *ceteris paribus*. Tutti gli altri coefficienti invece risultano essere significativi e con segno positivo.

I fattori che influenzano la soddisfazione del reddito sono gli stessi individuati dai modelli precedenti: il fatto di essere una donna, l'aver un'istruzione media o alta, di essere lavoratore o di esserlo stato, il possesso di ricchezze monetarie e le capacità cognitive portano ad una maggiore soddisfazione. In particolare, con l'avanzare dell'età aumenta la soddisfazione del reddito; questa caratteristica demografica trova nuovamente riscontro della relazione parabolica che instaura con la soddisfazione, anche se l'età minima decade a 46 anni. Come ci si potrebbe aspettare, le variabili associate allo stato di salute ovvero la depressione, il numero di malattie croniche e di quelle della mobilità comportano una minor soddisfazione con l'avanzamento della gravità di salute. Dalle stime dei coefficienti associati ai Paesi europei si ha ulteriore conferma del fatto che Olanda e Danimarca siano gli Stati più soddisfatti rispetto alla Germania, a parità di altre condizioni, la Polonia continua a classificarsi come il più scontento.

Tabella 3.20: Modelli Heckman per dati ordinali per il dominio della vita

	Risponde correttamente	Auto- valutazione	Risponde correttamente	C corretta
Donna	0.061 (0.050)	0.156** (0.062)	0.063 (0.050)	0.167*** (0.056)
Età_med	-0.027	0.133***	-0.032	0.184***

⁶La significatività di molti parametri viene persa con l'inserimento della variabile relativa al *Fluency Test* che un soggetto elenca durante il test; ciò sta ad indicare che l'impatto di tale covariata è talmente forte da "eliminare" altri effetti. Questo risultato porta a pensare che l'aver delle buone capacità cognitive sia talmente tanto importante da essere in grado di contribuire all'individuazione del corretto ordinamento.

	(0.036)	(0.033)	(0.036)	(0.033)
Età_med ²	0.010	-0.006	0.010	0.009
	(0.015)	(0.016)	(0.015)	(0.017)
Istruzione media	-0.022	-0.034	-0.021	-0.003
	(0.043)	(0.046)	(0.043)	(0.046)
Istruzione alta	0.124**	-0.026	0.119**	0.071
	(0.052)	(0.055)	(0.052)	(0.060)
Partner	-0.004	0.297***	-0.003	0.308***
	(0.045)	(0.051)	(0.045)	(0.049)
N componenti	-0.007	-0.073***	-0.007	-0.076***
	(0.018)	(0.020)	(0.018)	(0.020)
N figli	-0.027**	0.031**	-0.026**	-0.019
	(0.013)	(0.014)	(0.013)	(0.015)
Pensionato	0.117**	0.107**	0.114**	0.145**
	(0.049)	(0.056)	(0.049)	(0.058)
Occupato	0.084	0.170***	0.092*	0.244***
	(0.051)	(0.061)	(0.051)	(0.059)
Genitore	0.021	0.003	0.022	0.002
	(0.042)	(0.045)	(0.042)	(0.045)
Ricchezza	-0.001	0.012***	-0.001	0.011***
finanziaria	(0.003)	(0.003)	(0.003)	(0.003)
Ricchezza reale	0.002	0.017***	0.002	0.016***
	(0.004)	(0.004)	(0.004)	(0.005)
Depressione	0.066*	-0.396***	0.063*	-0.304***
	(0.038)	(0.044)	(0.038)	(0.043)
ADL	0.000	-0.100***	0.002	-0.102***
	(0.034)	(0.034)	(0.035)	(0.036)
IADL	0.037	-0.028	0.035	0.001
	(0.029)	(0.030)	(0.029)	(0.033)
Croniche	0.003	-0.036**	0.004	-0.023
	(0.014)	(0.015)	(0.014)	(0.015)
Mobilità	-0.011***	-0.037***	-0.013	-0.028**
	(0.012)	(0.013)	(0.012)	(0.013)
Sintomi	0.007	-0.082***	0.004	-0.063***
	(0.012)	(0.014)	(0.012)	(0.013)
<i>Grip Strength</i>	0.003	-0.001	0.003	0.004*
	(0.002)	(0.003)	(0.002)	(0.003)
<i>Fluency Test</i>	0.005**	0.005*	0.005**	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
Svezia	0.508***	0.228**	0.487***	0.357***
	(0.078)	(0.093)	(0.077)	(0.125)

Olanda	0.173** (0.071)	0.226*** (0.088)	0.167** (0.071)	0.324*** (0.091)
Spagna	0.463*** (0.080)	-0.143*** (0.094)	0.457*** (0.080)	0.118 (0.124)
Italia	0.390 (0.071)	-0.466*** (0.079)	0.387*** (0.071)	-0.012 (0.108)
Francia	0.217*** (0.082)	-0.215** (0.092)	0.215*** (0.082)	0.322*** (0.105)
Danimarca	0.310*** (0.060)	0.493*** (0.083)	0.294*** (0.060)	0.062 (0.089)
Grecia	0.070 (0.076)	-0.463*** (0.084)	0.060 (0.076)	-0.340*** (0.089)
Belgio	0.577*** (0.063)	-0.111 (0.081)	0.572*** (0.063)	-0.055 (0.125)
Repubblica Ceca	0.286*** (0.062)	-0.304*** (0.072)	0.287*** (0.062)	-0.374*** (0.087)
Polonia	0.297*** (0.075)	0.020 (0.092)	0.302*** (0.075)	0.071 (0.102)
Chiarimenti	-0.039** (0.016)		-0.036** (0.017)	
Tipo_b	0.046 (0.051)		0.021 (0.054)	
costante	-0.265* (0.158)		-0.254 (0.158)	
<i>soglia1</i>		-3.080*** (0.208)		-1.768*** (0.354)
<i>soglia2</i>		-2.084*** (0.208)		-0.873** (0.358)
<i>soglia3</i>		-1.113*** (0.228)		-0.361 (0.362)
<i>soglia4</i>		0.689** (0.295)		1.045*** (0.371)
<i>Osservazioni</i>	6920	4326	6920	4326
ρ		-0.527*** (0.141)		-0.040 (0.337)
<i>Test LR:</i> $H_0 = \rho = 0$	$\chi_1^2=2.93$	p-value=0.087	$\chi_1^2=0.01$	p-value=0.911

Standard error in parentesi: *** p<0.01, ** p<0.05, *p<0.1

Se si considera il modello di selezione per il dominio della vita, si giun-

ge alle stesse conclusioni del modello precedente per quanto riguarda la significatività dei coefficienti ρ : l'utilizzo della variabile corretta "annulla" la selezione, ma altrettanto non si può affermare per quella che si basa sulle autovalutazioni originarie: l'associazione che ne deriva mostra un valore significativo al 10% di livello ed un impatto medio alto, pari a -0.527.

Anche in questo dominio non son molti i coefficienti che risultano avere un'influenza sulla probabilità di rispondere secondo le condizioni alla base del metodo non parametrico dell'*anchoring vignettes*: avere un'istruzione elevata piuttosto che quella elementare, essere pensionati anziché non occupati, essere depressi e avere delle buone capacità cognitive porta a una maggior propensione di individuare il corretto ordinamento, a parità di altre caratteristiche; diversamente, il numero figli ancora in vita e il fatto di aver chiesto dei chiarimenti durante l'intervista CAPI mostrano una relazione negativa con tale probabilità: l'effetto della seconda variabile non sorprende dal momento che l'intervista CAPI coinvolge più di 50 domande, pertanto potrebbe essere che il soggetto che chiede parecchie delucidazioni sia poi meno propenso ad individuare correttamente l'ordinamento a causa, ad esempio, della stanchezza derivata dal precedente questionario. Sorprende invece l'influenza della depressione, che tuttavia è al limite della significatività statistica del 10%: potrebbe comunque essere che la loro condizione faciliti la distinzione dei soggetti ipotetici più "contenti" in quanto vivono una condizione più "felice" della loro e quindi distante dalla loro realtà. Caratteristiche come l'età, lo stato di salute e la ricchezza monetaria non hanno invece un impatto statisticamente significativo, a differenza di ciò che accade per il reddito. Evidenze interessanti invece emergono dai coefficienti attribuiti ai Paesi, i quali presentano tutti un'associazione positiva con la probabilità di individuare il corretto ordinamento, persino la Polonia: nello specifico, il fatto di essere intervistati belgi piuttosto che tedeschi porta ad una maggiore propensione, mentre il fatto di essere olandesi ad una minore, rispetto ad essere cittadini di altri Paesi Europei, *ceteris paribus*.

I fattori che invece influenzano una maggior soddisfazione della vita sono gli stessi individuati dai modelli precedenti ovvero essere di genere femminile, essere più anziano, avere un compagno, essere genitore, aver lavorato o essere lavoratore piuttosto che non occupato, possedere una maggior ricchezza

finanziaria e reale e avere delle buone capacità cognitive. Invece, se si ha uno stato di salute peggiore e al crescere del numero dei componenti del nucleo familiare, la soddisfazione della vita diminuisce, come d'altronde ci si poteva aspettare (come testimoniano altri autori e le analisi precedenti). Infine per quanto riguarda gli effetti dovuti al Paese di appartenenza non si notano particolari differenze rispetto a quanto già si era visto nei modelli probit ordinati bivariati, se non nella magnitudo dell'impatto.

I quattro modelli appena riportati mostrano una chiara concordanza: i due modelli che utilizzano le variabili originarie della soddisfazione mostrano l'esistenza di una forte selezione del campione, mentre quelli che considerano le variabili C affermano il contrario. Dunque la correzione delle risposte dalla distorsione della DIF "azzera" la selezione del campione che esiste se si considerano le variabili originarie della soddisfazione: questo porta a concludere l'inappropriatezza di un modello di selezione una volta che si è considerata ed eliminata l'eterogeneità delle scale di risposta, che se non operata, porta ad interpretare erroneamente i risultati.

In altri termini, nelle valutazioni originarie il DIF può manifestarsi introducendo forme di selezione del campione che in realtà non esisterebbero se si analizzassero variabili corrette per questo fenomeno.

Come ultima analisi, si è considerato il modello di selezione per dati ordinali sul solo sottoinsieme di individui che hanno risposto riconoscendo l'ordinamento tra le *anchoring vignettes* in entrambi i domini d'interesse (ovvero i 3982 individui dell'intersezione dell'insieme arancione e azzurro della Figura 3.22). Se è vero quanto si è detto coi risultati precedenti ovvero l'assenza di selezione per il campione che riconosce le assunzioni alla base della scala C in almeno uno dei due domini, non ci potrebbero essere particolari motivi per pensare il contrario per il sottocampione che coinvolge esclusivamente coloro che riconoscono l'ordinamento naturale in entrambe le sfere d'indagine. In quest'ottica, si è stimato il modello di selezione per dati ordinali, separatamente per la variabile corretta del reddito e per quella della vita. I coefficienti ρ confermano i risultati precedenti ovvero la mancanza di selezione ($\rho = -0.162$ con $p\text{-value} = 0.497$ per il primo dominio e $\rho = -0.090$ con $p\text{-value} = 0.759$ per il secondo dominio): in particolare le stime puntuali estremamente simili a

quelle riportate nelle Tabella 3.19 e Tabella 3.20 e per questo non sono nuovamente riportate. Queste evidenze sono un'ulteriore conferma di quanto si è già detto, ovvero che la differente correlazione evidenziata dall'analisi con le variabili corrette e quella con le variabili originarie non è dovuta al fatto che certi individui rispondono riconoscono l'ordinamento naturale, bensì al DIF.

Conclusioni

La soddisfazione e il benessere non sono fenomeni fisici che possono essere misurati facilmente ed oggettivamente, specialmente se si pone l'individuo come punto focale dell'indagine. Di fronte all'impossibilità di racchiudere in un indicatore statistico concetti astratti, la valutazione soggettiva della soddisfazione della vita sembra essere un criterio soddisfacente per cercare di quantificare il benessere. Nonostante ciò, non si deve sottovalutare la natura stessa di tale realtà: se è vero che vi è incertezza nell'individuare una valida definizione, dall'altra è possibile considerare il benessere, complessivamente, come un concetto aggregato che viene spiegato attraverso diversi suoi domini, un aspetto sorvolato molto spesso dai ricercatori. Di fronte a tale evidenza, questa tesi si è posta l'obiettivo di individuare l'esistenza di una possibile correlazione tra la soddisfazione della vita e quella di un suo dominio, la soddisfazione del reddito. Allo stesso tempo si vogliono individuare i fattori che risultano avere una certa influenza su queste due componenti, verificando se le evidenze riscontrate in letteratura siano confermate o meno. Più precisamente, non ci si è focalizzati sul nesso causale di queste due realtà, ma esclusivamente sulla possibile associazione che può esserci tra le due sfere.

Tuttavia, il confronto di autovalutazioni relative ad indagini sul benessere, salute o altri aspetti soggettivi della qualità della vita o della società spesso soffre del fatto che gruppi diversi usano scale di risposta differenti (DIF). Le *anchoring vignettes* sono uno strumento recente che permette di identificare e correggere questa eterogeneità che ne deriva; l'idea sottostante è quella di confrontare le autovalutazioni degli intervistati con un *gold standard* (un'ancora) che è creato all'interno del contesto d'indagine. Certamente tale approccio non risolve tutti i problemi, ma dovrebbe aver il potenziale di

ridurre la distorsione, aumentare l'efficacia e rendere le misurazioni più vicine al confronto interpersonale rispetto ai metodi esistenti. Attualmente in letteratura vi sono due modi di implementare le *vignettes* nelle analisi: l'approccio parametrico (il modello Chopit e le relative estensioni) e quello non parametrico (la scala C). In questa tesi si è considerata la seconda soluzione, in quanto più appropriata con l'obiettivo che ci si è posto e data l'inadeguatezza del primo metodo mostrata in alcuni studi: tale approccio consiste nella costruzione di una nuova variabile corretta dalla DIF, denominata scala C, che sfrutta esclusivamente l'informazione riguardante l'autovalutazione e le *vignettes* di un dominio.

I dati utilizzati in questa tesi riguardano la seconda rilevazione dell'indagine SHARE, svolta nel 2006-2007 in 11 Paesi Europei: si tratta di un'indagine sulla salute, età e pensionamento in Europa; nello specifico, i dati che si sono considerati riguardano gli intervistati che hanno risposto sia alle autovalutazioni sulla soddisfazione della vita e su quella del reddito sia ad entrambe le domande specifiche per ciascun dominio e aventi per protagonisti gli individui ipotetici.

L'applicazione del metodo non parametrico basato sulla tecnica delle *anchoring vignettes* ha portato all'identificazione di due sottogruppi di individui per ciascuna sfera d'indagine: coloro per i quali è stato possibile creare la variabile C e coloro che invece hanno fornito valutazioni degli individui ipotetici con errato ordinamento. In particolare poco più della metà degli intervistati riconoscono l'ordinamento naturale delle *anchoring vignettes* in entrambi i domini d'interesse. Le analisi descrittive rilevano una discreta correlazione tra le due soddisfazioni d'interesse se si considerano le autovalutazioni originarie, mentre questa associazione risulta più bassa se si analizzano le due rispettive variabili libere da DIF (ovvero corrette dall'eterogeneità dell'uso di scale di risposte). In particolare la distribuzione dell'autovalutazione di ciascun dominio è quasi la stessa tra i due gruppi di individui, coloro che non rispondono rispettando l'ordinamento tra le *vignettes* e l'autovalutazione e coloro che invece soddisfano le assunzioni alla base della scala C; emergono invece diverse evidenze se si analizza la distribuzione dell'autovalutazione originaria e di quella corretta. Tuttavia, essendo d'interesse analizzare l'asso-

ciazione tra i due domini e allo stesso tempo individuare quali fattori abbiano maggior influenza su essi, si è considerato un modello probit ordinato bivariato sulle variabili originarie e su quelle corrette: questo risulta essere adeguato per lo scopo in questione, dal momento che le correlazioni tra i due errori risultano essere statisticamente significative al livello del 5%; in particolare si trova conferma di quanto emerso nelle analisi descrittive ovvero l'evidenza di una correlazione più elevata quando non si correggono le autovalutazioni per il DIF.

Il dubbio che potrebbe sorgere è che le diverse associazioni siano dovute ad un problema di numerosità campionaria ovvero al fatto che i soggetti considerati nel modello con la valutazione corretta da DIF siano un sottocampione di quelli analizzati nel modello con l'autovalutazione originaria. Conseguentemente si è ristimato il modello precedente (per l'autovalutazione originaria) esclusivamente per coloro che forniscono risposte secondo l'esatto ordinamento, ottenendo lo stesso risultato: ciò porta a supporre che la differente correlazione ottenuta con le variabili originarie e quella con le variabili C non sia dovuta ad un problema di osservazioni, ma del DIF.

Di fronte a queste evidenze sorge tuttavia la seguente perplessità: la ricerca dell'associazione d'interesse tra le due variabili C deve avvenire nella consapevolezza che tale correlazione deve essere "preceduta" necessariamente dalla capacità di aver riconosciuto correttamente l'ordinamento stesso tra le *anchoring vignettes* o in altre parole che tale correlazione possa essere calcolata solo per coloro che rispondono correttamente ad entrambe le autovalutazioni. Tuttavia, prima di valutare congiuntamente le due soddisfazioni, si è verificata la "distorsione da selezione" separatamente per i singoli domini: il modello stimato (un'estensione del modello di Heckman per dati ordinali) rileva l'assenza di una correlazione forte e significativa se si considerano le variabili corrette e dunque l'assenza di distorsione da selezione, evidenza che invece emerge se si considerano le variabili originarie relative all'autovalutazione.

Conseguentemente i risultati ottenuti con l'esclusiva analisi delle autovalutazioni originali portano a conclusioni divergenti e forvianti rispetto a quelle relative alle variabili corrette dal DIF: dunque se non si riportassero sulla stessa scala le valutazioni dei due domini, si affermerebbe che la soddisfazione della vita e quella del reddito siano fortemente associate tra loro e si

mostrerebbe l'esistenza di una selezione da campione. In altri termini, nelle valutazioni originarie il DIF può manifestarsi introducendo forme di selezione del campione che in realtà non esisterebbero se si analizzassero variabili corrette per questo fenomeno. Dunque l'unico motivo che porta ad avere diverse associazioni è il DIF.

Inoltre non sono da sottovalutare le differenze tra Paesi Europei: secondo la letteratura la Danimarca è il Paese più soddisfatto della propria vita; tuttavia i risultati emersi dal modello che corregge la DIF portano a smentire tale constatazione. Le stime sulle dummy dei Paesi evidenziano dunque l'impatto di correggere il DIF se si vogliono confrontare misure di autovalutazioni di cittadini provenienti da Stati diversi: ci sono differenze in termini di culture, *welfare*, sistemi sanitari, ecc che inevitabilmente si riflettono sugli sili di risposta dei cittadini di ciascun Paese, un problema di cui è necessario tener conto in un'analisi.

Chiaramente il problema dell'eterogeneità delle scale di risposte è talmente ampio che non può essere limitato a questa tesi. Alcuni studi mostrano l'inadeguatezza del modello parametrico Chopit e suggeriscono l'implementazione di soluzioni semiparametriche, più flessibili di quelle parametriche, seppur più restrittive delle non parametriche. Al fine di verificare se anche in altre analisi viene confermata l'evidenza riscontrata, si potrebbe stimare il modello Chopit e il modello probit ordinale bivariato per le due autovalutazioni originarie e verificare se i risultati trovati vengono confutati o meno. Inoltre in questa tesi si è considerato un solo dominio e non si è indagata una relazione causale: conseguentemente, si potrebbero sviluppare anche queste analisi e capire meglio la relazione tra la soddisfazione della vita e i suoi singoli domini.

Elenco delle tabelle

2.1	Classificazione dell'autovalutazione Y_s e valutazione di 3 <i>vignettes</i> Y_1, Y_2, Y_3 . Fonte: Van Soest e Vonkova (2014).*	52
2.3	Sommario dell'entropia e dei casi intervallari di sottoinsiemi di <i>vignettes</i> †	62
2.4	Errato ordinamento e presenza di <i>ties</i>	63
3.1	Descrizione e statistiche riassuntive delle variabili esplicative considerate	85
3.2	Distribuzione dell'età dei rispondenti	87
3.3	Distribuzione dei principali indicatori dello stato di salute	91
3.4	Distribuzione reddito netto mensile familiare	92
3.5	Media, mediana e composizione di alcune variabili esplicative tra i Paesi Europei	95
3.6	Analisi bivariata della soddisfazione del reddito con alcune variabili socio-demografiche ‡	98
3.7	Correlazioni tra la soddisfazione del reddito e le rispettive <i>vignette</i>	102
3.8	Analisi bivariata della soddisfazione della vita con alcune variabili socio-demografiche ‡	105
3.9	Correlazioni tra la soddisfazione della vita e le rispettive <i>vignettes</i>	108
3.10	Analisi bivariata della soddisfazione della vita e del reddito	109
3.11	Correlazioni tra la soddisfazione del reddito e la soddisfazione della vita	110
3.12	Errato ordinamento e presenza di <i>ties</i>	112

3.13	Classificazione delle autovalutazioni e delle relative due <i>vignettes</i> di ciascun dominio	113
3.14	Proporzione di individui tra i sottogruppi individuati dalla selezione della variabile C in ciascun Paese	118
3.15	Media, mediana e composizione di alcune variabili esplicative tra i due gruppi di individui selezionati dalla variabile C . . .	120
3.16	Correlazioni tra la soddisfazione del reddito e la soddisfazione della vita con le variabili corrette	120
3.17	Modelli probit ordinato bivariati (campioni distinti)	122
3.18	Modelli probit ordinati bivariati (campioni comuni)	129
3.19	Modelli Heckman per dati ordinali per il dominio del reddito	134
3.20	Modelli Heckman per dati ordinali per il dominio della vita	137

Elenco delle figure

1.1	Struttura per gli indicatori del benessere secondo OECD. Fonte: OECD (2011).	15
1.2	Soddisfazione della vita delle persone di 16 anni e più per Paese (Anno 2013). Fonte: ISTAT (2016).	17
1.3	Modello a due stadi di Van Praag et al. (2003). Fonte: Van Praag et al. (2003).	20
2.1	Categorizzazione degli anni d'età di due individui. Fonte: King et al. (2004).	29
2.2	Mappaggio dei livelli inosservati continui dell'autovalutazione nelle categorie ordinali riportate dagli intervistati i e j . Fonte: Wand (2013).	30
2.3	Autovalutazioni e valutazione di tre <i>vignettes</i> in un contesto di <i>political efficacy</i> : confronto tra due individui e correzione della scala di risposta (approccio non parametrico). Fonte: King et al. (2004).	34
2.4	Eterogeneità ipotetica dello stato occupazionale nel riportare la mobilità fisica e correzione delle <i>vignettes</i> . Fonte: Bago d'Uva et al. (2011).	36
2.5	Rappresentazione grafica dell'idea alla base del modello parametrico. Fonte: King et al. (2004).	38
3.1	Distribuzione del livello di istruzione più status occupazionale	88
3.2	Boxplot sulle relazioni	89
3.3	Boxplot dello stato civile per chi vive solo	89

3.4	Istogramma del numero di componenti del nucleo familiare dell'individuo	89
3.5	Distribuzione arcseno iperbolico del reddito netto familiare . . .	91
3.6	Distribuzione ricchezza finanziaria e reale	93
3.7	Distribuzione arcseno iperbolico della ricchezza finanziaria e reale	93
3.8	Istogramma della soddisfazione del reddito	96
3.9	Soddisfazione del reddito per Paese	97
3.10	Reddito familiare annuale e soddisfazione del reddito totale . .	100
3.11	Valutazione delle vignettes sulla soddisfazione del reddito . . .	100
3.12	Valutazione delle vignettes sulla soddisfazione del reddito per Paese	101
3.13	Istogramma della soddisfazione della vita	103
3.14	Soddisfazione della vita per Paese	104
3.15	Valutazione delle vignettes sulla soddisfazione della vita	106
3.16	Valutazione delle vignettes sulla soddisfazione del reddito per Paese	107
3.17	Istogramma con 3 metodi per l'allocazione delle inconsistenze per il reddito	114
3.18	Autovalutazione del reddito in base alla selezione di C	115
3.19	Autovalutazione della vita in base alla selezione di C	115
3.20	Autovalutazione del reddito per il sottocampione $S^R = 1$. . .	117
3.21	Autovalutazione della vita per il sottocampione $S^V = 1$. . .	117
3.22	Rappresentazione grafica della composizione del campione . .	121

Appendice A

Descrizione delle variabili in scala europea

Scala ISCED: i livelli sono classificati come segue:

- Livello 0: nessun titolo di istruzione;
- Livello 1: istruzione elementare/primo stadio di istruzione di base;
- Livello 2: istruzione secondaria inferiore (fine dell'istruzione obbligatoria per quasi tutti i paesi, in Italia corrisponde alla scuola secondaria di primo grado);
- Livello 3: istruzione secondaria superiore;
- Livello 4: istruzione post-secondaria non terziaria (corsi pre-universitari, brevi corsi professionalizzanti);
- Livello 5: istruzione terziaria, primo stadio (laurea o laurea magistrale);
- Livello 6: istruzione terziaria, secondo stato (dottorato di ricerca).

Scala EURO-D: Item più alti della scala indicano una maggiore depressione; nello specifico essi sono:

- Item 0 (definito da SHARE): non è depresso;
- Item 1: depressione;

- Item 2: pessimismo per il futuro o aspettative nulle;
- Item 3: desiderio di morte o suicidio;
- Item 4: l'individuo si sente in colpa;
- Item 5: problemi di sonno;
- item 6: poco interesse;
- Item 7: irritabilità;
- Item 8: mancanza o diminuzione del senso di appetito;
- Item 9: ristrette energie, l'individuo si affatica;
- Item 10: difficoltà nel concentrarsi in uno spettacolo o nella lettura;
- Item 11: mancanza di divertimento;
- Item 12: pianto

Indicatore ADL:

- Item 1: fare il bagno;
- Item 2: vestirsi bene;
- Item 3: toilette;
- Item 4: spostarsi;
- Item 5: continenza di feci e urine;
- Item 6: alimentazione

Indicatore IADL:

- Item 1: usare il telefono;
- Item 2: fare la spesa;
- Item 3: preparare il cibo;

- Item 4: governo della casa;
- Item 5: fare il bucato;
- Item 6: mezzi di trasporto;
- Item 7: assunzione farmaci;
- Item 8: uso del denaro

Formula per aggiustamenti PPP: $e \frac{nom}{ppp}$ dove

- e = Valore monetario variabile d'interesse
- nom = Tasso di cambio nominale
- ppp = Tassi di cambio aggiustati a parità di potere d'acquisto

Ringraziamenti

Si ringrazia SHARE per aver concesso l'accesso e l'uso del dataset analizzato in questa tesi.

Un ringraziamento speciale al professor Omar Paccagnella, relatore di questa tesi, per il suo aiuto e pazienza che ha mostrato in questi mesi. Con disponibilità mi ha affiancata in questo lavoro, rendendolo un'esperienza interessante.

Un ringraziamento particolare ai miei genitori, che mi hanno permesso moralmente e materialmente di continuare gli studi universitari, ma soprattutto per il supporto e l'incoraggiamento che mi hanno sempre mostrato. Spero siate orgogliosi di me e che questo giorno possa far dimenticare, in parte, altri "problemi" che si risolveranno, sicuramente. Un ringraziamento altrettanto importante a mia sorella Martina, il "dono" più bello che i miei genitori mi abbiano mai regalato: si dice "Se non capisci come una donna possa amare teneramente sua sorella e al tempo stesso aver voglia di torcerle il collo, allora probabilmente sei figlio unico." e non c'è affermazione più vera, sebbene siano di più i momenti di amore e di complicità. Tu hai sempre creduto in me e continui a farlo.

Grazie a Federico, che ha condiviso con me tutte le gioie e le difficoltà di questo cammino universitario e non. Adesso ci aspettano nuove avventure assieme.

Un ringraziamento ai nonni che so che mi sono vicini anche se non fisica-

mente.

Grazie ai miei amici: oggi condividete con me questa ulteriore pezzetto di vita. Ora non potrò più giocarmi la carta "Devo studiare". Sicuramente vi sarete divertiti e sbizzarriti nuovamente molto per il papiro, non ho dubbi.

Infine, ma non per ultimi, un grazie speciale a tutti i miei compagni universitari, i nuovi e i vecchi Soliti: senza di voi questa esperienza universitaria non sarebbe stata uguale. Siete come una seconda famiglia, un piedistallo su cui poter contare. Strade diverse ci hanno condotto e ci condurranno lontani, ma spero pur sempre vicini.

Bibliografia

URL: <https://uwaterloo.ca/canadian-index-wellbeing/>.

URL: http://dssm.unipa.it/augugliaro/TeachingFiles/Stat1/parte4_3.pdf.

Agresti, A. *Categorical Data Analysis*. 2nd ed. New York: Wiley-Interscience, 2002.

Alexander, C. S. e H. J. Becker. «The use of vignettes in survey research». In: *Public opinion quarterly* 42.1 (1978), pp. 93–104.

An, X. e Y. F. Yung. «Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It». In: (2003). A cura di SAS Institute.

Angelini, V., D. Cavapozzi, L. Corazzini e O. Paccagnella. «Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for Individual-Specific Scale Biases». In: *Oxford bulletin of Economics and Statistics* Vol. 76.5 (2014), pp. 643–666.

Angelini, V., D. Cavapozzi e O. Paccagnella. «Dynamics of reporting work disability in Europe». In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* Vol. 174 (2011), 621–638.

Bago d’Uva, T., M. Lindeboom, O. O’Donnell e E. van Doorslaer. «Slipping Anchors? Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity». In: *The Journal of Human Resources* Vol. 46.4 (2011), pp. 875–906.

Bellemare, C., B. Melenberg e A. Van Soest. «Semi-parametric models for satisfaction with income». In: *Portuguese Economic Journal* 1.2 (2002), pp. 181–203.

Betti, G. *Analisi Statistica del Reddito e delle Condizioni di Vita*. Dipartimento di Economia Politica e Statistica, Università di Siena, inedito, 2015. Cap. 1-2.

- Bonsang, E. e A. Van Soest. «Satisfaction with job and income among older individuals across European countries». In: *Social Indicators Research* 105.2 (2012), pp. 227–254.
- Carbonell, A. Ferrer-i. «Income and well-being: an empirical analysis of the comparison income effect». In: *Journal of Public Economics* Vol. 89.5 (2005), pp. 997–1019.
- Carbonell, A. Ferrer-i e B. MS Van Praag. «Income satisfaction inequality and its causes». In: *Journal of Economic Inequality* 1.2 (2003), pp. 107–127.
- Clark, A., F. Etilé, F. Postel-Vinay, C. Senik e K. Van der Straeten. «Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries». In: *The Economic Journal* 115.502 (2005).
- Clark, Andrew. «A note on unhappiness and unemployment duration». In: (2006).
- Cowell, F. A. «Income Distribution and Inequality». In: *London: STICERD, London School of Economics* (2007).
- Crane, M., C. Risel, S. Greaves e Gebel. K. «Correcting bias in self-rated quality of life: an application of anchoring vignettes and ordinal regression models to better understand QoL differences across commuting modes». In: *Quality of Life Research* Vol. 25(2) (2016), 257–266.
- Crawford Solberg, E., Ed Diener, D. Wirtz, R. E. Lucas e S. Oishi. «Wanting, having, and satisfaction: examining the role of desire discrepancies in satisfaction with income.» In: *Journal of personality and social psychology* 83.3 (2002), p. 725.
- Datta Gupta, N., N. Kristensen e D. Pozzoli. «External validation of the use of vignettes in cross-country health studies». In: *Economic Modelling* Vol. 27.4 (2010), pp. 854–865.
- De Battisti, F., G. Nicolini e S. Salini. «The Rasch model to measure service quality». In: 27 (2003). A cura di Working Paper del Dipartimento di Economia Politica e Aziendale.
- De Luca, G. e V. Perotti. «Estimation of ordered response models with sample selection». In: *The Stata Journal* Vol. 11.2 (2011), pp. 213–239.
- Diener, Ed. «Subjective well-being.» In: *Psychological bulletin* 95.3 (1984), p. 542.

- Dolan, P., T. Peasgood e M. White. «Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being». In: *Journal of economic psychology* 29.1 (2008), pp. 94–122.
- Easterlin, Richard A. «Income and happiness: Towards a unified theory». In: *The economic journal* 111.473 (2001), pp. 465–484.
- Eurostat. «Analytical report on subjective well-being». In: *STATISTICAL WORKING PAPERS* (2016).
- Food & Nutrition, Barilla Center for. *La misurazione del benessere delle persone: il BCFN Index*. 2010. URL: <https://www.barillacfn.com/m/publications/pp-misurazione-benessere-persone.pdf>.
- Giovannini, E. «Indagine conoscitiva sull'individuazione di indicatori di misurazione del benessere ulteriori rispetto al PIL. Audizione del Presidente dell'Istituto nazionale di statistica». In: (2012). URL: http://www.istat.it/it/files/2012/02/Audizione-BES_22_02_2012-DEF_23-FEBBRAIO-2012.pdf.
- Greene, W. e D. Hensher. «Ordered Choices and Heterogeneity in Attribute Processing». In: *Journal of Transport Economics and Policy (JTEP)* Vol. 44.3 (2010), pp. 331–364.
- Greene, W. H. e D. A. Hensher. *Modelling Ordered Choices: A primer*. Cambridge University Press, 2010.
- Hadri, K. e W. Mikhail, cur. *Econometric Methods and Their Applications in Finance, Macro and Related Fields*. World Scientific Publishing Co. Pte. Ltd., 2014.
- Holland, P. W. e H. Wainer. *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum, 1993.
- Hopkins, D. e G. King. «Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability». In: *Public Opinion Quarterly* (2010), pp. 1–22.
- ISTAT. *Rapporto Bes 2016: il benessere equo e sostenibile in Italia*. 2016. URL: <http://www.istat.it/it/files/2016/12/BES-2016.pdf>.
- Jones, A., N. Rice e S. Robone. *A comparison of parametric and non-parametric adjustments using vignettes for self-reported data*. Health, Econometrics and Data Group (HEDG) Working Papers. HEDG, c/o Department of

- Economics, University of York, 2012. URL: <http://EconPapers.repec.org/RePEc:yor:hectdg:12/10>.
- Kapteyn, A., J. P. Smith, A. Van Soest e H. Vonkova. «Anchoring vignettes and response consistency». In: *RAND Labour and Population Working, WR840* (2011).
- Kapteyn, A., J. P. Smith e A. Van Soest. «Are Americans really less happy with their incomes?» In: *Review of Income and Wealth* 59.1 (2013), pp. 44–65.
- Kapteyn, A., J. Smith e A. Van Soest. «Vignettes and self-reports of work disability in the US and the Netherlands». In: *American Economic Review* Vol. 97(1) (2007), pp. 461–473.
- King, G. e J. Wand. «Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes». In: *Political Analysis* 15 (2007), pp. 46–66.
- King, G., C. J. L. Murray, J. A. Salomon e A. Tandon. «Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research». In: *American Political Science Review* Vol. 98 (2004), pp. 191–207.
- Michalos, A. C. *Education, Happiness and Wellbeing*. 2007. URL: <https://www.oecd.org/site/worldforum06/38303200.pdf>.
- «Multiple discrepancies theory (MDT)». In: *Social indicators research* Vol. 16.4 (1985), pp. 347–413.
- Miranda, A., S. Rabe-Hesketh et al. *Maximum likelihood estimation of endogenous switching and sample selection models for binary, count, and ordinal variables*. Rapp. tecn. Centre for Economic Research, Keele University, 2005.
- Newson, Roger. «Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences». In: *The Stata Journal* (2002).
- OECD. *GDP per capita in PPS*. 2017. URL: <http://ec.europa.eu/eurostat/tgm/graph.do?tab=graph&plugin=1&pcode=tec00114&language=en&toolbox=data>.
- «How's life?: Measuring well-being». In: (2011). URL: <http://dx.doi.org/10.1787/9789264121164-en>.

- «OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth». In: (2013). URL: <http://dx.doi.org/10.1787/9789264194830-en>.
- Paccagnella, O. «Anchoring vignettes with sample selection due to nonresponse». In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* Vol. 174 (2011), 665–687.
- «Modelling individual heterogeneity in ordered choice models: Anchoring Vignettes and the Chopit model». In: *Quaderni di Statistica* Vol. 15 (2013), pp. 69–94.
- Paccagnella, O., M. Guidolin, G. Derboni e T. Bago d’Uva. «The anchoring vignette approach to measure customer satisfaction». In: (2015). URL: <http://meetings.sis-statistica.org/index.php/ginilegacy/SIS2015/paper/viewFile/3444/752>.
- Peiro, A. «Happiness, satisfaction and socio-economic conditions: Some international evidence». In: *The Journal of Socio-Economics* 35.2 (2006), pp. 348–365.
- Peracchi, F. e C. Rossetti. «The heterogeneous thresholds ordered response model: Identification and inference». In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* Vol. 176.3 (2013), pp. 703–722.
- Primi, R., C. Zanon, D. Santos, F. De Fruyt e O. P. Jonh. «Can They Make Adolescent Self-Reports of Social-Emotional Skills More Reliable, Discriminant, and Criterion-Valid?» In: *European Journal of Psychological Assessment* Vol. 32(1) (2016), pp. 39–51.
- RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc. Boston, MA, 2015. URL: <http://www.rstudio.com/>.
- Sajaia, Z. «BIOPROBIT: Stata module for bivariate ordered probit regression». In: (2008).
- SHARE. *Release Guide 6.0.0*. 2017. URL: http://www.share-project.org/fileadmin/pdf_documentation/SHARE_release_guide_6-0-0.pdf.
- Stata. *spearman* — *Spearman’s and Kendall’s correlations*. 2017. URL: <https://www.stata.com/manuals13/rspearman.pdf>.
- Stevenson, B. e J. Wolfers. «Subjective well-being and income: Is there any evidence of satiation?» In: *The American Economic Review* 103.3 (2013), pp. 598–604.

- Terza, J. «Ordered Probit: A Generalization». In: *Communications in Statistics - Theory and Methods* Vol. 4 (1985), pp. 1–11.
- Treccani. *Benessere*. 2017. URL: <http://www.treccani.it/vocabolario/benessere/>.
- Van Praag, B. MS, P. Frijters e A. Ferrer-i Carbonell. «The anatomy of subjective well-being». In: *Journal of Economic Behavior & Organization* 51.1 (2003), pp. 29–49.
- Van Soest, A. e H. Vonkova. «Testing Parametric Models Using Anchoring Vignettes against Nonparametric Alternatives». In: (2012). URL: <http://arno.uvt.nl/show.cgi?fid=122505>.
- «Testing the specification of parametric models by using anchoring vignettes». In: *Journal of the Royal Statistical Society: Series A, (Statistics in Society)* Vol. 177(1) (2014), pp. 115–133.
- Van Soest, A., L. Delaney, C. Harmon, A. Kapteyn e J. P. Smith. «Validating the Use of Anchoring Vignettes for the Correction of Response Scale Differences in Subjective Questions». In: *Journal of the Royal Statistical Society Series A, (Statistics in Society)* Vol. 174(3) (2011), pp. 575–595.
- Wand, J. «Credible Comparisons Using Interpersonally Incomparable Data: Nonparametric Scales with Anchoring Vignettes». In: *American Journal of Political Science* Vol. 57, No. 1 (2013), pp. 249–262.
- Wand, J., G. King e O. Lau. «Anchors: Software for Anchoring Vignette Data». In: *Journal of Statistical Software* Vol. 42.3 (2011), pp. 115–133.