



Istituto per la Ricerca Valutativa sulle Politiche Pubbliche

# **GLI EFFETTI DEI PROGRAMMI DI STUDIO ALL'ESTERO: L'ESPERIENZA DEL PROGETTO MOS-4**

*a cura di*

*Silvia De Poli, Loris Vergolini e Nadir Zanini*

## Sommario

<b>GLI EFFETTI DEI PROGRAMMI DI STUDIO ALL'ESTERO:</b> .....	<b>1</b>
<b>L'ESPERIENZA DEL PROGETTO MOS-4</b> .....	<b>1</b>
<b>1. INTRODUZIONE</b> .....	<b>1</b>
<b>2. IL PROGRAMMA MOS-4: DESCRIZIONE E SUE FINALITÀ</b> .....	<b>3</b>
<b>3. IL DISEGNO DELLA RICERCA: STRATEGIA DI IDENTIFICAZIONE, DATI E VARIABILI</b> .....	<b>4</b>
<b>4. IL PROCESSO DI RANDOMIZZAZIONE</b> .....	<b>8</b>
<b>5. ANALISI DESCRITTIVE</b> .....	<b>11</b>
5.1    LA SODDISFAZIONE DEI PARTECIPANTI AL PROGRAMMA MOS-4 .....	11
5.2    DIFFERENZE TRA BENEFICIARI E NON BENEFICIARI, PRIMA E DOPO IL TRATTAMENTO .....	13
<b>6. GLI EFFETTI DEL PROGRAMMA</b> .....	<b>15</b>
<b>7. QUALCHE CONSIDERAZIONE CONCLUSIVA</b> .....	<b>16</b>
<b>APPENDICE METODOLOGICA</b> .....	<b>18</b>
IL MODELLO DI MISURAZIONE .....	18
IDENTIFICAZIONE E STIMA.....	26
<b>RIFERIMENTI BIBLIOGRAFICI</b> .....	<b>27</b>

## 1. Introduzione

Oltre un decennio fa i leader dell'Unione Europea hanno concordato sulla necessità di delineare una strategia per rendere l'Europa «*the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion*» (Lisbon European Council, 2000). Al fine di perseguire questo scopo, un ruolo chiave è stato assegnato nel percorso formativo e professionale di ogni cittadino europeo alla *mobilità* per favorire lo scambio di conoscenze e competenze tra i diversi paesi dell'unione.

Per garantire la possibilità di accedere a stimolanti opportunità di apprendimento, la Commissione Europea ha investito notevoli risorse a sostegno di programmi formativi, rivolti a studenti delle secondarie superiori e dell'università e svolti in paesi diversi da quello di residenza. Tra i più noti si possono ricordare i programmi comunitari Erasmus e Leonardo Da Vinci, sebbene siano molteplici quelli organizzati con varia sistematicità da amministrazioni locali con finanziamenti comunitari. Tali iniziative sono, però, più spesso indirizzate agli studenti universitari e a neolaureati, mentre minore attenzione è riservata agli studenti nel ciclo superiore degli studi secondario. Per questi le occasioni di mobilità sono spesso lasciate alla libera iniziativa delle famiglie che, spinte in gran parte dal desiderio di migliorare le abilità linguistiche dei propri figli, investono, qualora le condizioni economiche lo consentano, in corsi di lingua che si tengono all'estero.

Proprio per rimediare alle disuguali opportunità di apprendimento che così si vengono a creare in gran parte del territorio, il Servizio Europa (SE) della Provincia autonoma di Trento (PaT) sta promuovendo da tempo iniziative di mobilità a livello europeo per giovani diplomati o laureati. È, tuttavia, evidente che, quanto prima queste esperienze di miglioramento delle competenze di lingua straniera avvengono, tanto migliori saranno i loro esiti cognitivi. Proprio in base a questa considerazione, il SE ha organizzato un programma di studio all'estero rivolto agli studenti del quarto anno delle scuole superiori della provincia. Questa iniziativa, nota come MOS-4<sup>1</sup>, ha offerto a 200 soggetti la possibilità di partecipare gratuitamente ad un'esperienza di *full immersion* linguistica di 4 settimane, da svolgersi durante l'estate 2012 presso enti di formazione ubicati nel Regno Unito e in Irlanda.

In letteratura esistono alcune ricerche che si sono occupate dello studio degli effetti dei viaggi di studio su tratti della personalità e sulle capacità di apprendimento dei partecipanti. Le evidenze empiriche emerse mostrano come tali esperienze influenzino positivamente i risultati accademici

---

<sup>1</sup> Il programma è stato ideato nell'ambito del Programma Operativo FSE 2007/2013, con il cofinanziamento del Fondo Sociale Europeo, dello Stato italiano e della Provincia autonoma di Trento.

degli studenti che hanno preso parte ai programmi durante il periodo di studio all'università (Sutton e Rubin 2010, Houser et al. 2011). Inoltre, argomentazioni teoriche (McLeod e Wainwright 2008) ed evidenze empiriche (Ismail et al. 2006) hanno indicato che potrebbero avere effetti anche su tratti della personalità come l'autostima, l'apertura mentale e la diversità.<sup>2</sup> Effetti di questo tipo sono confermati anche da studi empirici condotti in modo rigoroso: Williams (2005), ad esempio, mostra come l'aver studiato un periodo all'estero comporti effetti positivi sulla comunicazione interculturale.

Proprio per capire se anche MOS-4 abbia avuto un impatto sui partecipanti, il SE della PaT ha inteso sottoporre il programma ad uno studio di valutazione, coinvolgendo Irvapp fin dalla fase di progettazione del programma. In questo modo è stato possibile implementare un cosiddetto "esperimento sociale" (Heckman e Smith 1995), nel quale gli studenti che hanno presentato domanda di partecipazione e ritenuti ammissibili sono stati assegnati casualmente ad un gruppo di beneficiari – composto dai richiedenti estratti per partecipare al programma – e ad un gruppo di controllo – formato dai richiedenti ammissibili che non beneficeranno del programma. Inoltre, nel caso in cui trattati e controlli siano nella stessa proporzione (50/50), il ricorso alla randomizzazione consente di ottimizzare l'informazione a disposizione al fine di ottenere stime robuste di effetti anche relativamente tenui di un programma, anche con numerosità dei partecipanti relativamente basse.

In questo modo è possibile garantire la presenza di due gruppi di soggetti perfettamente confrontabili da cui è possibile ricavare una stima corretta dell'impatto medio causato dalla misura soggetta a valutazione nei confronti di una serie di variabili risultato, anche dette *outcome*.

L'obiettivo di questo documento è dunque quello di fornire una valutazione di impatto degli effetti del programma MOS-4 su dimensioni di carattere psico-attitudinale quali l'autostima e l'apertura mentale<sup>3</sup>, la conoscenza dell'inglese, le performance scolastiche e sulla propensione di iscriversi all'università. In questo modo sarà possibile capire se programmi di mobilità rivolti a studenti nel ciclo di studi secondari inducano i cambiamenti desiderati e quindi se risulti vantaggioso spendere finanziamenti pubblici per questo tipo di iniziative.

Il resto del documento è organizzato nel seguente modo: nel paragrafo 2 viene descritto il programma e le sue finalità. Nel paragrafo 3 vengono presentati in dettaglio il disegno della ricerca e i dati a disposizione. Il paragrafo 4 riporta alcune analisi descrittive sulla soddisfazione dei partecipanti al programma e sui punteggi medi degli outcome di interesse. Il paragrafo 5 illustra le

---

<sup>2</sup> Tali risultati sono da intendersi puramente descrittivi, in quanto basati solo su un gruppo di individui che hanno partecipato al programma e non sul confronto con un adeguato gruppo di controllo.

<sup>3</sup> Per facilitare la lettura nel corso del testo ci riferiremo a questo insieme di *outcome* con il termine generico "percezione di sé".

evidenze empiriche degli effetti del programma. Il paragrafo 6 riguarda la verifica del processo di randomizzazione. Il paragrafo 7, infine, è riservato a qualche considerazione conclusiva.

## **2. Il programma MOS-4: descrizione e sue finalità**

Il programma MOS-4 offre la possibilità di trascorrere un'esperienza all'estero a studenti che hanno appena concluso il quarto anno di studi in una scuola superiore della provincia di Trento, nell'estate che precede l'inizio dell'ultimo anno del ciclo di studi secondari. Il programma prevede di prendere parte a 4 settimane di *full immersion* linguistica presso enti di formazione ubicati nel Regno Unito e in Irlanda. La partecipazione degli studenti è totalmente gratuita, dato che il programma prevede la totale copertura delle spese di partecipazione da parte dell'amministrazione provinciale attraverso il Fondo Sociale Europeo, che si farà carico delle spese relative a viaggio e assicurazione, vitto e alloggio presso apposite strutture, corso di lingua presso un *college* selezionato. A carico delle famiglie resta solamente una cauzione di €200, restituita a tutti coloro che completeranno le 4 settimane previste dal programma.

Per essere ammissibile al programma era sufficiente che lo studente, residente in Trentino, fosse iscritto al quarto anno di una scuola secondaria superiore di durata quinquennale della provincia di Trento. Inoltre, la condizione economica della famiglia deve essere inferiore al valore 1 dell'indicatore Icef<sup>4</sup>, che in termini di reddito annuo equivalente corrisponde a circa €50.000, o a € 102.000 per una famiglia di tre persone. Un tale valore sta ad indicare che solo gli studenti delle famiglie più ricche (e quindi in grado di provvedere autonomamente ad un'esperienza formativa all'estero) non sono ammessi al programma, per ovvi motivi legati all'equità con cui i fondi pubblici vengono spesi.

Nonostante le risorse stanziare per il programma fossero notevoli, circa €600.000 in grado di sostenere 200 studenti, già dalla fase di disegno del programma risultava chiaro che sarebbero state insufficienti a coprire la totalità delle domande di partecipazione che si sarebbero ricevute. Data l'esigenza di procedere ad una selezione dei beneficiari all'interno dei richiedenti ammissibili, in modo da garantire la fattibilità di un rigoroso studio di valutazione, si è deciso di procedere ad un'estrazione casuale degli studenti che avrebbero potuto beneficiare del programma.

L'idea alla base di questa iniziativa è dunque quella di consentire agli studenti di vivere un'esperienza che possa influenzare il loro modo di pensare o la "percezione di sé", favorendo atteggiamenti come l'apertura verso altre culture e la fiducia in sé stessi, che possono incidere

---

<sup>4</sup> L'acronimo Icef significa "Indicatore della Condizione Economica Familiare". Si tratta di uno strumento che permette di stabilire le concrete condizioni economiche di una famiglia sulla base di informazioni di tipo reddituale, patrimoniale, finanziario e anagrafico. Tale indicatore permette ai nuclei più deboli l'accesso ai benefici sociali e alle agevolazioni tariffarie offerte dalla Provincia di Trento e dagli Enti Locali in Trentino.

sull'interesse verso lo studio. In questo senso MOS-4 si prefigura come un programma indirizzato al miglioramento delle performance scolastiche e delle scelte in ambito di prosecuzione degli studi, prima ancora che allo specifico sviluppo delle abilità linguistiche. Rispetto a precedenti iniziative rivolte a diplomati e laureati, anticipando l'esperienza formativa di un anno, la partecipazione a MOS-4 potrebbe risultare in grado di produrre maggiori effetti, poiché alla conclusione del quinto anno gli studenti potrebbero aver già deciso se proseguire o meno gli studi.

### **3. Il disegno della ricerca: strategia di identificazione, dati e variabili**

Per valutare gli effetti di un intervento è necessario confrontare quello che si è potuto riscontrare tra i beneficiari di un certo programma in seguito alla sua introduzione e quello che sarebbe accaduto se quegli stessi individui non avessero preso parte al programma. Dato che per un individuo è impossibile essere contemporaneamente sia beneficiario che non esserlo, solitamente si ricorre al confronto in termini di condizioni e comportamenti che il programma stesso intende modificare tra un gruppo di beneficiari e un adeguato gruppo di controllo. Il problema principale nella valutazione di impatto diventa dunque quello di ricostruire opportunamente il gruppo di controllo con cui confrontare i partecipanti di un programma. Il rischio è, infatti, quello che le regole per l'ammissibilità al programma selezionino individui con caratteristiche molto specifiche, che li rende non confrontabili a coloro che non hanno beneficiato del programma (Martini e Trivellato 2011).

Quando, come nel caso di MOS-4, tra gli ammissibili richiedenti l'assegnazione degli individui fra beneficiari e non beneficiari avviene in modo causale, ossia quando viene messa in atto una cosiddetta randomizzazione, la comparabilità dei due gruppi è garantita (Bloom 2008). Infatti, tale processo basato sul caso consente di eliminare qualsiasi fonte di distorsione dovuta a fattori che oltre ad influenzare la partecipazione al programma, potrebbero indirettamente causare cambiamenti negli *outcome* di interesse. In altre parole, il processo di randomizzazione, quando attuato e rispettato nel corso dell'intero periodo dell'intervento, assicura che non vi sia il cosiddetto *selection bias*, una distorsione dovuta dalle differenze di partenza tra i due gruppi. Tale distorsione si sarebbe potuta verificare se la selezione dei beneficiari fosse stata attuata su graduatorie ottenute sulla base di caratteristiche degli studenti richiedenti. Tra queste bisogna considerare sia caratteristiche come la conoscenza della lingua inglese o la percezione di sé che il programma intende modificare direttamente, che altre legate indirettamente agli esiti dell'iniziativa, ad esempio la condizione economico-patrimoniale della famiglia misurata dall'indicatore Icef<sup>5</sup> o il titolo di

---

<sup>5</sup> Nel caso di MOS-4, l'iniziativa è rivolta a tutti gli studenti con un indicatore Icef inferiore a 1. Di conseguenza, grazie alla randomizzazione, non sussistono differenze sistematiche rispetto a questa dimensione tra i due gruppi, a differenza

studio dei genitori. In questo caso le eventuali differenze riscontrate a seguito del programma non potrebbero essere imputabili al programma stesso, ma alle differenze di partenza conseguenti al modo in cui sono stati selezionati gli studenti.

Il ricorso alla randomizzazione dei richiedenti fra trattati e controlli è una prassi consolidata soprattutto nei paesi anglosassoni e tipica della sperimentazione di nuovi interventi pubblici. Essa, infatti, consente di identificare l'impatto di programmi rivolti a un numero anche contenuto di beneficiari. D'altro canto, a seconda delle caratteristiche del programma, la sua attuazione può comportare problemi etici, legati all'equità dell'utilizzo del caso come strumento di selezione dei beneficiari, oltre che di fattibilità (Berk 2005). Tali questioni dipendono strettamente dal caso oggetto di studio. Nello specifico caso di MOS-4, una minaccia alla validità del disegno sperimentale attuato è la presenza di soggetti assegnati casualmente al gruppo dei beneficiari che però decidono di non partecipare al programma (e viceversa), ossia i cosiddetti *non compliers*. Questi individui sono infatti particolari, poiché se assegnati al gruppo di trattamento decidono volontariamente di non partecipare (*no show*), oppure se assegnati al gruppo di controllo trovano comunque modo di partecipare al trattamento (*cross-over*). Verosimilmente ciò accade perché essi sono influenzati da alcune caratteristiche personali che oltre ad influire sulla partecipazione al programma potrebbero anche essere determinanti degli outcome del nostro studio.

Nel caso di MOS-4, tra i 281 studenti che hanno fatto richiesta di partecipare all'iniziativa, alcuni studenti che a seguito dell'estrazione casuale erano ammissibili al programma hanno deciso di non beneficiarne, mentre altri studenti che non erano ammissibili hanno comunque avuto delle esperienze di viaggio studio all'estero in Europa.

**Tabella 1** La distribuzioni dei richiedenti tra ammissibili e beneficiari al programma.

	Percezione di sé			Test di inglese		
	Beneficiari	Non beneficiari	Totale	Beneficiari	Non beneficiari	Totale
Ammissibili	186	9	195	186	9	195
Non ammissibili	17	69	86	14	72	86
Totale	203	78	281	200	81	281

In questo caso coloro che hanno partecipato a programmi in paesi anglosassoni sono stati considerati come non ammissibili, ma beneficiari per tutti gli outcome presi in considerazione. Quegli studenti che hanno preso parte a viaggi di studio in paesi non anglosassoni sono stati considerati come non ammissibili, ma beneficiari solo in relazione agli outcome collegati alla

---

delle precedenti edizioni di programmi di formazione all'estero organizzati dal SE, dove venivano generate graduatorie per l'ammissione basate proprio sull'indicatore Icef.

percezione di sé.<sup>6</sup> Come mostrato dalla tabella 1 i *non-compliers* risultano comunque contenuti: circa l'8% del totale nel caso del test di inglese e il 9% nel caso della percezione di sé.

Per poter raccogliere i dati necessari a svolgere l'esercizio di valutazione appena descritto, è stata implementata un'indagine *ad hoc* sulla totalità degli studenti che hanno fatto domanda di partecipazione al programma MOS-4, articolata su due rilevazioni: una al momento della richiesta di partecipazione al programma, l'altra dopo la partecipazione al programma. L'indagine *pre-intervento* condotta ad aprile 2012 ha rilevato una serie di informazioni sugli studenti in termini di caratteristiche socio-demografiche, origini sociali e carriera scolastica pregressa<sup>7</sup>, oltre agli outcome dello studio di valutazione. Ciò consentirà di verificare in modo dettagliato che non vi siano differenze fra beneficiari e non beneficiari prima di prendere parte all'iniziativa. L'indagine *post-intervento*, invece, è stata condotta ad ottobre 2012 e rivolta alla raccolta, per tutti gli studenti richiedenti, delle variabili di interesse su cui l'iniziativa intendeva agire, oltre ad alcune domande rivolte ai soli partecipanti circa la soddisfazione da parte degli studenti nella partecipazione al programma.

Per la raccolta dati si è utilizzata la cosiddetta modalità CAWI (*Computer Assisted Web Interviewing*), in cui gli studenti hanno risposto al questionario nelle aule informatiche degli istituti di provenienza sotto la supervisione degli addetti del SE della PaT, in modo da ridurre la possibilità di *cheating*. La tabella 2 riporta i dati relativi al numero di casi validi<sup>8</sup> intervistati nelle due indagini, da cui emerge come il tasso di risposta all'intervista post-programma sia decisamente buono, attestandosi sul 95%.

**Tabella 2** Numero di casi validi.

	Risposte complete <b>PRE</b>		Risposte complete <b>POST</b>	
	Test di inglese	Questionario	Test di inglese	Questionario
Ammissibili	200	197	195	195
Non ammissibili	96	95	86	86
Totale	296	292	281	281

Per quanto riguarda le informazioni raccolte nel questionario e, più specificatamente, gli *outcome* attinenti alla sfera della percezione di sé<sup>9</sup>, sono state predisposte una serie di batterie di

<sup>6</sup> Per essere considerati come beneficiari questi studenti devono aver partecipato ad un programma della durata di almeno 2 settimane.

<sup>7</sup> Per l'elenco completo si veda tabella 7.

<sup>8</sup> Sono stati esclusi i soggetti con Icef superiore a 1 che non erano ammissibili al programma e che sono stati comunque intervistati.

<sup>9</sup> Queste variabili sono rilevate sulla base di un'autovalutazione soggettiva. L'utilizzo di variabili soggettive come variabili dipendenti resta una questione aperta nelle scienze sociali. Se, da un lato, il ricorso alle variabili soggettive è molto diffuso per rispondere ad alcuni interrogativi circa le attitudini delle persone, dall'altro, alcuni ricercatori



domande.<sup>10</sup> La prima di queste, mutuata dal lavoro di Waxin (2004), ha permesso di individuare cinque dimensioni relative alle opinioni che gli individui hanno di sé stessi:

1. fiducia nelle proprie capacità;
2. capacità di adattamento;
3. orientamento sociale;
4. propensione alla comunicazione;
5. capacità di apertura.

La seconda e la terza batteria sono state mutate dal contributo di Tafarodi e Milne (2002) che hanno permesso di identificare tre ulteriori dimensioni:

6. abilità nel risolvere i problemi;
7. valutazione oggettiva delle proprie competenze (*self-competence*);
8. accettazione delle proprie qualità (*self-liking*).

In particolare, la differenza tra *self-competence* e *self-liking* ricalca la differenza tra due possibili dimensioni dell'autostima (Tafarodi e Swann 2001). La prima riguarda il grado con cui gli individui si percepiscono come capaci ed efficaci nella realtà sociale e si basa su un'autovalutazione dei propri talenti e delle proprie abilità. In questo caso l'individuo assume valore per quello che sa fare. La seconda, invece, riguarda il modo in cui gli individui si sentono di essere persone che valgono e si basa principalmente sull'approvazione sociale. Il valore viene dato per come l'individuo appare ed è centrato su aspetti caratteriali e su caratteristiche morali. Abbiamo quindi quattro dimensioni (2, 3, 4 e 5) che fanno riferimento all'apertura mentale degli studenti e alle loro capacità di adattarsi a diverse situazioni. A queste se ne aggiungono altre quattro (1, 6, 7 e 8) che riguardano la loro autostima e la fiducia che ripongono nelle proprie capacità.

Per quanto riguarda la conoscenza del test di inglese agli studenti interessati dal programma è stato somministrato un test di lingua standardizzato<sup>11</sup> espresso in 90esimi, mentre l'ultimo *outcome* da valutare è rappresentato dalla propensione ad iscriversi all'università. In quest'ultimo caso agli studenti viene chiesto se intendono proseguire gli studi o meno alla fine della scuola secondaria superiore.<sup>12</sup>

---

sollevano alcuni dubbi sull'affidabilità di queste informazioni. In ogni caso, l'utilizzo di variabili soggettive per effettuare confronti tra gruppi di individui, come nel nostro caso, è condiviso (Bertrand e Mullainathan 2001).

<sup>10</sup> Tutte le dimensioni identificate dalle diverse batterie di domande sono state estratte utilizzando l'analisi fattoriale confermativa e si è altresì verificato che i modelli di misurazioni siano equivalenti tra i trattati e i controlli. Le otto dimensioni identificate saranno rilevate tramite altrettante variabili continue. Rimandiamo al paragrafo "il modello di misurazione" dell'appendice metodologica per maggiori dettagli tecnici.

<sup>11</sup> Il test di inglese è stato predisposto da un team di esperti dell'I.S.I.T. (Istituto Universitario per Interpreti e Traduttori) di Trento.

<sup>12</sup> Gli indecisi sono stati trattati alla stregua di chi ha risposto di non voler proseguire gli studi.

#### 4. Il processo di randomizzazione

Per poter attuare il disegno di valutazione proposto è necessario appurare che il processo di randomizzazione messo in atto sia andato a buon fine. Quando infatti si opera un'assegnazione casuale su un numero di richiedenti relativamente basso, come appunto nel caso di MOS-4, il rischio che gli studenti selezionati a partecipare al programma siano diversi dai richiedenti esclusi è comunque da non sottovalutare. Per questo motivo è importante verificare che i due gruppi di studenti coinvolti nel progetto siano effettivamente equivalenti rispetto ad un ampio insieme di variabili di controllo che potrebbero influire sia sulla decisione di partecipare al programma, sia sugli outcome del nostro studio. L'equivalenza dei due gruppi è un requisito necessario per poter comparare beneficiari e non beneficiari nella fase di post-trattamento, ossia al loro ritorno dall'esperienza di studio all'estero.

La tabella 3 riporta i principali risultati emersi dal confronto tra ammissibili e non ammissibili grazie al questionario somministrato prima della partecipazione al programma. Complessivamente non emergono particolari differenze di composizione dei due gruppi in termini di caratteristiche socio-demografiche, origini sociali o carriera scolastica pregressa. Più precisamente, in tabella 3 vengono presentate le medie per trattati e controlli sulle diverse variabili di controllo, ed emerge come non vi siano differenze significative tra i due gruppi. Questo risultato incoraggiante è confermato anche dal modello *probit* presentato nell'ultima colonna. Tale modello stima l'effetto che le diverse variabili hanno sulla probabilità di risultare beneficiario: da tale modello emerge che solo un parametro è statisticamente significativo e lo è in riferimento alla modalità "Imprenditore, libero professionista e dirigente" della classe sociale del padre.<sup>13</sup>

Questi risultati ci portano, quindi, a concludere che il processo di randomizzazione, che rappresenta il primo passo per poter valutare in modo rigoroso il programma MOS-4, è risultato essere molto preciso. A conferma di questo risultato la Figura 1 presenta le probabilità stimate sulla base del modello appena discusso per la probabilità di risultare beneficiari. Si nota chiaramente come vi sia un'ampia sovrapposizione tra le due aree, che conferma come i due gruppi risultino comparabili tra loro per quanto riguarda le caratteristiche osservate prese in esame.

---

<sup>13</sup> Tale risultato, in ogni caso, non sembra inficiare il processo randomizzazione, poiché in direzione opposta a quanto ci aspetteremmo di trovare negli outcome.

**Tabella 3** Differenze medie tra beneficiari e non beneficiari e modello probit per la probabilità di essere trattati

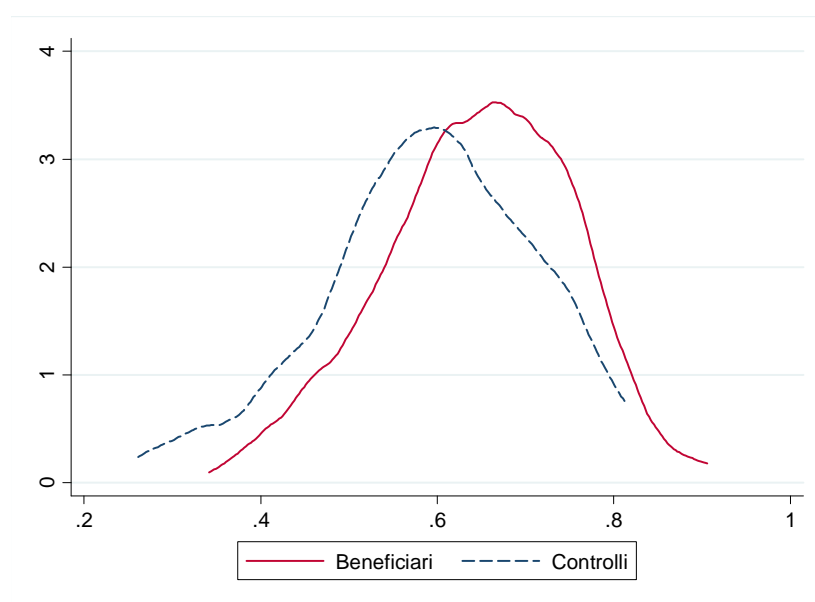
	Beneficiari			Non beneficiari			Modello probit (Trattati=1)	
	Media	Std. error	N	Media	Std. error	N	Stima	Std. error
<b>Variabili di controllo</b>								
Sesso: femmine	0,674	(0,034)	193	0,588	(0,046)	114	0,183	(0,166)
Numero componenti famiglia	4,295	(0,082)	191	4,277	(0,069)	112	0,003	(0,090)
Deprivazioni	3,099	(0,200)	191	2,679	(0,245)	112	0,031	(0,031)
Bocciatura: sì	0,136	(0,025)	191	0,196	(0,038)	112	-0,198	(0,227)
Voto medio (anno in corso)	6,911	(0,056)	191	6,817	(0,075)	112	-0,001	(0,171)
Voto medio (ultimo anno)	7,099	(0,058)	191	6,973	(0,082)	112	0,086	(0,158)
Debiti formativi: sì	0,492	(0,036)	191	0,545	(0,047)	112	0,000	(0,205)
Classe sociale del padre								
Non occupato	0,027	(0,012)	187	0,018	(0,013)	112	-0,217	(0,581)
Impr., lib. prof. e dirigente	0,193	(0,029)	187	0,250	(0,041)	112	-0,487**	(0,247)
Impiegato	0,235	(0,031)	187	0,250	(0,041)	112	-0,300	(0,230)
Autonomo	0,155	(0,027)	187	0,188	(0,037)	112	-0,349	(0,241)
Operaio	0,390	(0,036)	187	0,295	(0,043)	112	0	-
Classe sociale della madre								
Non occupata	0,084	(0,020)	190	0,134	(0,032)	112	-0,253	(0,308)
Impr., lib. prof. e dirigente	0,132	(0,025)	190	0,089	(0,027)	112	0,320	(0,303)
Impiegata	0,463	(0,036)	190	0,464	(0,047)	112	0,032	(0,221)
Autonoma	0,058	(0,017)	190	0,036	(0,018)	112	0,502	(0,404)
Operaia	0,263	(0,032)	190	0,277	(0,043)	112	0	-
Titolo di studio dei genitori								
Fino licenza media	0,305	(0,034)	190	0,339	(0,045)	112	-0,309	(0,276)
Diploma	0,468	(0,036)	190	0,446	(0,047)	112	-0,161	(0,217)
Laurea e oltre	0,226	(0,030)	190	0,214	(0,039)	112	0	-
Condizione madre: occupata	0,758	(0,031)	190	0,714	(0,043)	112	0,007	(0,220)
Condizione padre: occupato	0,897	(0,022)	185	0,856	(0,034)	111	0,206	(0,244)
Straniero	0,187	(0,028)	193	0,193	(0,037)	114	-0,033	(0,224)
Tipo di scuola: liceo	0,513	(0,036)	193	0,465	(0,047)	114	0,019	(0,163)
Costante							-0,287	(1,159)
N							296	
Pseudo-R <sup>2</sup>							0,04	
Test F (21, 274)							0,76	p-value=0,77

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Per quanto invece concerne le differenze tra beneficiari e non beneficiari relativamente agli outcome da valutare, è possibile condurre delle cosiddette “regressioni placebo”, in quanto svolte prima della partecipazione al programma e, quindi, in assenza di un trattamento. In tal caso, visto che la selezione dei beneficiari è avvenuta in modo casuale quello che ci si aspetta è di trovare effetti nulli dello status di beneficiario. La Tabella 4 mostra le stime degli “effetti grezzi”, dati dalla stima di una regressione per ciascuno degli outcome considerati sullo status di beneficiario o non beneficiario, interpretabile come differenza algebrica fra gli outcome medi dei due gruppi. Inoltre,

in Tabella 4 vengono riportati gli “effetti netti”, scaturiti dagli stessi modelli di regressione precedenti ma stimati considerando anche tutte le variabili di controllo a disposizione, in modo da guadagnare precisione. Anche in questo caso i dati sono confortanti, infatti emerge chiaramente come non vi siano differenze significative tra beneficiari e non beneficiari prima del trattamento. Questo risultato è in linea con quanto ci aspettavamo nel caso in cui il processo di randomizzazione fosse andato a buon fine, infatti nel caso di equivalenza dei due gruppi non ci dovrebbero essere differenze significative tra trattati e controlli prima della partecipazione al programma.

**Figura 1** Area di supporto comune



**Tabella 4** Modelli di regressione placebo per gli effetti dell’essere non beneficiari sugli outcome di interesse

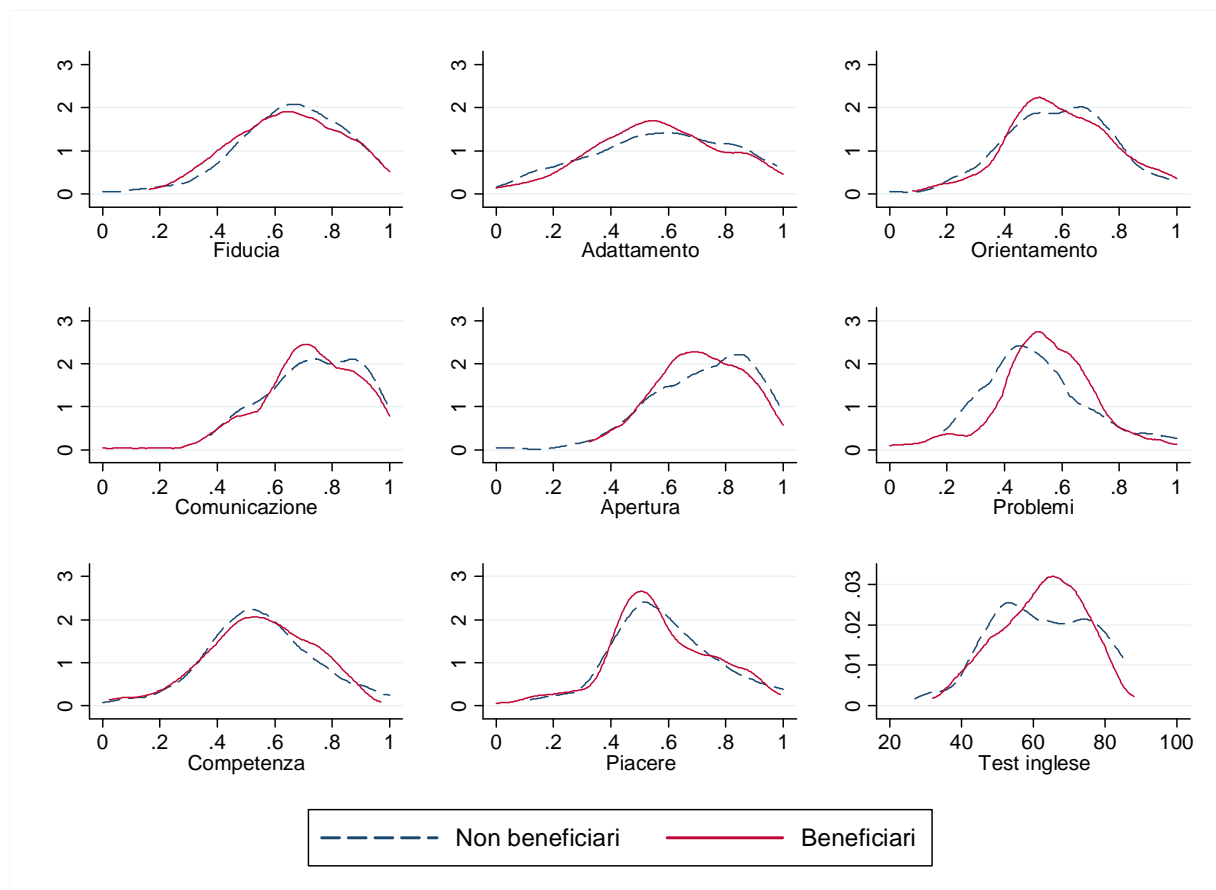
	Beneficiari		Non beneficiari		Effetti grezzi		Effetti netti	
	Media	St. Err.	Media	St. Err.	Stima	St. Err.	Stima	St. Err.
Fiducia nelle proprie capacità	0,649	(0,013)	0,657	(0,017)	-0,008	(0,022)	-0,018	(0,022)
Capacità di adattamento	0,567	(0,016)	0,578	(0,023)	-0,010	(0,028)	-0,020	(0,028)
Orientamento sociale	0,598	(0,013)	0,586	(0,017)	0,011	(0,022)	0,002	(0,022)
Propensione alla comunicazione	0,720	(0,012)	0,739	(0,014)	-0,019	(0,020)	-0,021	(0,020)
Capacità di apertura	0,712	(0,011)	0,724	(0,016)	-0,012	(0,019)	-0,022	(0,019)
Abilità nel risolvere i problemi	0,543	(0,012)	0,515	(0,017)	0,028	(0,021)	0,029	(0,021)
<i>Self-competence</i>	0,546	(0,013)	0,552	(0,018)	-0,006	(0,022)	-0,003	(0,023)
<i>Self-liking</i>	0,574	(0,013)	0,587	(0,017)	-0,014	(0,022)	-0,018	(0,023)
Test di inglese	62,39	(0,848)	62,07	(1,292)	0,324	(1,484)	-0,679	(1,288)
Propensione iscrizione università	0,827	(0,027)	0,813	(0,037)	0,015	(0,046)	0,009	(0,046)

Un ultimo controllo di robustezza concerne il confronto tra le distribuzioni degli outcome che possono essere tratti come variabili continue (Figura 2).<sup>14</sup> In questo caso l’intento è di valutare se anche le distribuzioni complessive degli outcome siano sostanzialmente simili tra i gruppi sotto esame. Dall’ispezione grafica emerge come la maggior parte delle variabili mostri distribuzioni

<sup>14</sup> L’unico outcome escluso dalla Figura 2 è la propensione a proseguire gli studi, poiché rappresentata da una variabile dicotomica che assume solo valori 0-1.

sostanzialmente identiche, abbiamo piccole difformità solo nel caso di tre variabili: “Capacità di apertura”, “Abilità nel risolvere i problemi” e “Test di inglese”.

**Figura 2** Distribuzione degli outcome di interesse tra beneficiari e non beneficiari



## 5. Analisi descrittive

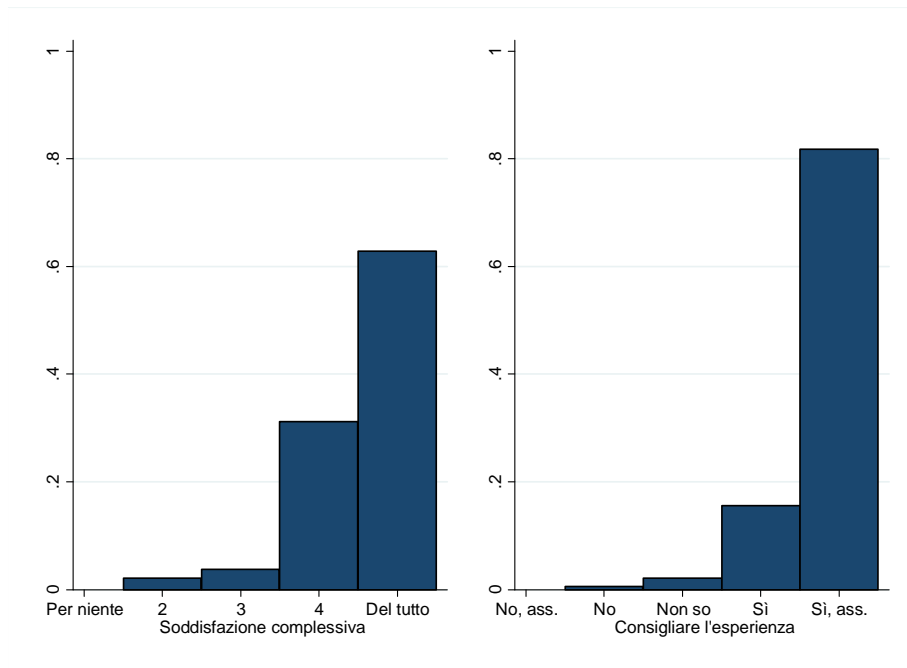
Nel corso di questo paragrafo verranno presentate alcune analisi descrittive tratte dai dati appena descritti circa il livello di soddisfazione espresso dai partecipanti verso il programma MOS-4. In secondo luogo, mostreremo i punteggi medi di beneficiari e non beneficiari, prima e dopo il trattamento, con l’obiettivo di mappare gli andamenti nel tempo sugli outcome di interesse, che non costituiscono una misura dell’impatto del trattamento (che verrà invece presentata nel paragrafo 5).

### 5.1 La soddisfazione dei partecipanti al programma MOS-4

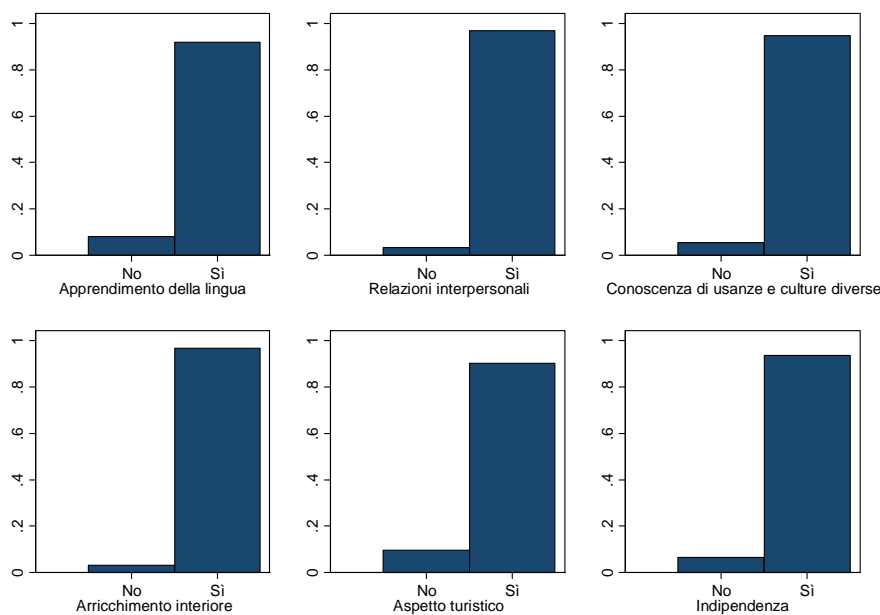
Un primo risultato di interesse riguarda sicuramente il grado di soddisfazione che i beneficiari hanno manifestato nei confronti dell’esperienza. Dalla figura 3 emerge come la soddisfazione complessiva sia molto elevata, tanto che quasi la totalità dei partecipanti si sente di consigliare tale esperienza ad amici e conoscenti. I livelli di soddisfazione sono particolarmente elevati anche in relazione a particolari aspetti quali: l’apprendimento della lingua, le relazioni interpersonali stabilite durante l’esperienza, il venire a contatto con culture diverse, l’arricchimento interiore, l’aspetto

turistico e l'indipendenza (figura 4). Infatti, per tutti questi aspetti si dicono soddisfatti più del 90% dei partecipanti. Per quanto riguarda, invece, gli aspetti problematici collegati all'esperienza all'estero emerge come, in generale, la maggior parte degli studenti che ha partecipato al programma, a parte un diffuso disorientamento iniziale, non ha manifestato grossi problemi (figura 5).

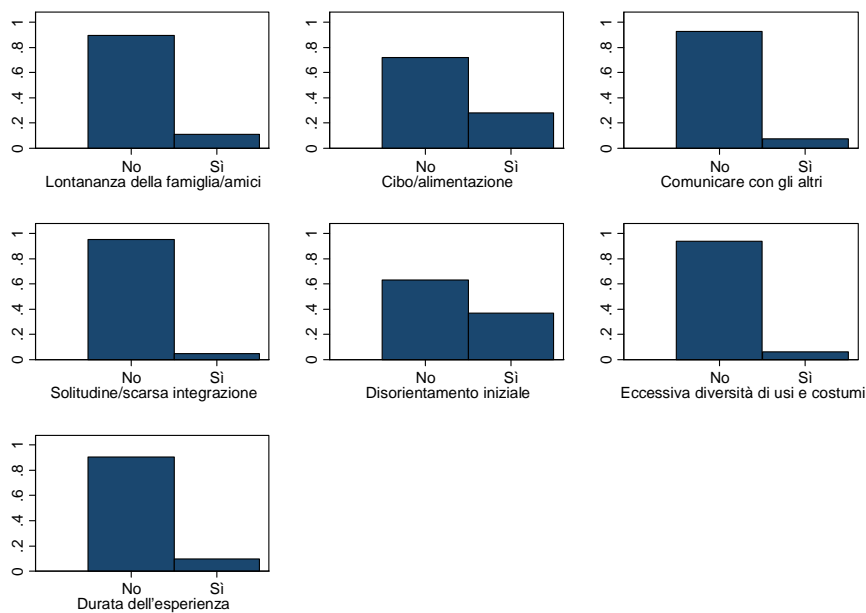
**Figura 3** La soddisfazione generale verso l'esperienza all'estero.



**Figura 4** La soddisfazione verso particolari aspetti dell'esperienza all'estero.



**Figura 5** Gli aspetti problematici collegati all'esperienza all'estero.



### 5.2 Differenze tra beneficiari e non beneficiari, prima e dopo il trattamento

Un risultato di maggior interesse riguarda sicuramente i punteggi ottenuti dagli studenti in relazione al test di inglese (tabella 5).<sup>15</sup> Se ci concentriamo sui dati post-intervento, emerge chiaramente la presenza di un effetto positivo e statisticamente significativo dell'aver partecipato ai corsi all'estero. Più precisamente, i beneficiari hanno in media 6 punti in più rispetto ai non beneficiari. Il quadro, però, cambia se attuiamo un confronto longitudinale prima-dopo. Infatti, se guardiamo i punteggi totali osserviamo come si passi da un punteggio medio di circa 62/90 ad uno di circa 59/90, abbiamo quindi una riduzione media di circa 3 punti, che risulta essere significativa. Se però attuiamo il confronto prima-dopo all'interno dei due gruppi la riduzione che si riscontra non è statisticamente significativa. Infatti, gli intervalli di confidenza relativi ai due gruppi si sovrappongono. In ogni caso un calo nelle performance dei non beneficiari era prevedibile per il fatto che molto probabilmente non hanno avuto occasione di esercitare il proprio inglese durante l'estate. Più sorprendente è, invece, il mancato miglioramento relativamente ai beneficiari.

Una possibile spiegazione può riguardare la scarsa accuratezza che gli studenti potrebbero aver messo nel rispondere al test di inglese post-programma. L'unica variabile che ci permette di valutare questa possibilità riguarda il tempo<sup>16</sup> medio di risposta al test di inglese pre e post intervento (tabella 6). Da tali risultati emerge come in totale il tempo di risposta al test si sia ridotto in modo significativo di circa 6 minuti e che questa riduzione è decisamente più marcata per i non

<sup>15</sup> Le analisi presentate nelle tabelle 3, 4 e 5 sono svolte solo sui compliers.

<sup>16</sup> Tale variabile è stata costruita applicando la seguente formula: tempo=minuti+secondi/60.

beneficiari (quasi 8 minuti in meno). La riduzione dei tempi di risposta relativamente ai beneficiari non è affatto sorprendente, in quanto a seguito del periodo di studio all'estero potrebbero sentirsi più sicuri di sé e rispondere quindi più velocemente al questionario. Più preoccupante è invece il calo riferito ai controlli che potrebbe rappresentare uno scarso impegno allo svolgimento del test a seguito di un programma dal quale sono stati esclusi.

**Tabella 5** Punteggi medi sul test di inglese espressi in 90esimi.

Test di inglese <b>PRE</b>				
	Media	Std. err.	Intervallo confidenza (95%)	
Totale	62.32	0.73	60.89	63.75
Non beneficiari	60.57	1.71	57.21	63.93
Beneficiari	62.67	0.86	60.96	64.36
Test di inglese <b>POST</b>				
	Media	Std. err.	Intervallo confidenza (95%)	
Totale	58.68	0.83	57.04	60.32
Non beneficiari	54.26	1.94	50.44	58.08
Beneficiari	60.56	0.91	58.76	62.36

**Tabella 6** Tempo medio di risposta al test di inglese in minuti.

Tempo medio <b>PRE</b>				
	Media	Std. err.	Intervallo confidenza (95%)	
Totale	46.70	0.57	45.58	47.82
Non beneficiari	45.82	1.07	43.72	47.91
Beneficiari	47.04	0.67	45.72	48.37
Tempo medio <b>POST</b>				
	Media	Std. err.	Intervallo confidenza (95%)	
Totale	40.39	0.69	39.03	41.76
Non beneficiari	38.10	1.55	35.05	41.15
Beneficiari	41.26	0.75	39.79	41.73

Quale che sia la motivazione della riduzione dei tempi, vista la strategia di valutazione attuata, i risultati circa gli effetti del programma sulle capacità linguistiche si basano sul solo confronto post. Per questo motivo, i risultati ottenuti in termini di effetto sulle capacità linguistiche restano validi, sotto l'assunzione che beneficiari e controlli si siano approcciati al test con lo stesso impegno. Tale assunto pare realistico, visto che comunque la differenza tra beneficiari e non beneficiari è ridotta (circa 3 minuti) e statisticamente non significativa.



**Tabella 7** Differenze medie tra beneficiari e non beneficiari pre- e post-intervento sugli outcome collegati alla percezione di sé.

	Pre-intervento				Post-intervento			
	Beneficiari		Non beneficiari		Beneficiari		Non beneficiari	
	Media	St. Err.	Media	St. Err.	Stima	St. Err.	Stima	St. Err.
Fiducia nelle proprie capacità	0.651	0.014	0.653	0.020	0.702	0.013	0.605	0.023
Capacità di adattamento	0.570	0.029	0.570	0.029	0.717	0.013	0.657	0.023
Orientamento sociale	0.594	0.013	0.576	0.019	0.607	0.016	0.516	0.024
Propensione alla comunicazione	0.721	0.012	0.729	0.018	0.723	0.015	0.654	0.023
Capacità di apertura	0.710	0.011	0.717	0.019	0.730	0.013	0.658	0.024
Abilità nel risolvere i problemi	0.538	0.012	0.515	0.022	0.661	0.010	0.658	0.016
<i>Self-competence</i>	0.549	0.014	0.559	0.023	0.607	0.015	0.643	0.021
<i>Self-liking</i>	0.574	0.014	0.584	0.023	0.659	0.013	0.697	0.020

Oltre alle competenze linguistiche il programma MOS-4 aveva come obiettivo quello di migliorare l'autostima e di ampliare l'apertura mentale dei ragazzi. Per poter rilevare tali concetti abbiamo fatto ricorso ad una serie di batterie di item le cui informazioni sono state poi riassunte in otto indici. La tabella 7 riporta i valori medi su tali indici divisi tra beneficiari e non beneficiari. Innanzitutto, possiamo notare come, a differenza del test di inglese, si possa osservare un incremento nei livelli medi di autostima e di apertura mentale dei beneficiari dopo la partecipazione al programma. In questo caso le domande poste prima e dopo erano esattamente le stesse. Per questo motivo, diversamente da quanto ottenuto per il test di lingua, l'errore di misura è decisamente più limitato.

## 6. Gli effetti del programma

La tabella 8 riporta i risultati degli effetti della partecipazione al programma, considerando gli outcome rilevati nell'indagine post-intervento. Possiamo limitarci all'indagine post-intervento, in quanto, come esposto nel paragrafo 4, il processo di randomizzazione è andato a buone fine. Più precisamente, la tabella 8 riporta due modelli. Il primo, ottenuto tramite un modello di regressione lineare, considera come variabile causale di interesse l'essere o meno ammissibile al programma fornendo la stima del cosiddetto *Intention-To-Treat* (ITT). Il modello 2, invece, presenta le stime del *Local Average Treatment Effect* (LATE), ottenute mediante ricorso alle variabili strumentali, che permette di valutare l'impatto dell'aver beneficiato del programma limitatamente ai *compliers*, ovvero coloro che hanno rispettato la regola di assegnazione casuale al programma.<sup>17</sup> In generale, emerge chiaramente come la partecipazione al programma MOS-4 abbia avuto un effetto positivo e

<sup>17</sup> Rimandiamo al paragrafo "identificazione e stima" dell'appendice metodologica per una spiegazione dettagliata delle tecniche statistiche utilizzate in questo paragrafo.

significativo sul test di inglese e su cinque degli otto indici collegati all'autostima e alla percezione del sé.<sup>18</sup> Non si evidenzia nessun effetto in relazione alla propensione di iscriversi all'università. Visto che gli outcome sono standardizzati i parametri dei due modelli sono da interpretare in termini di deviazioni standard. In altre parole, il parametro 0.572 relativo alla fiducia nelle proprie capacità significa che i beneficiari hanno avuto un incremento su questo indice, rispetto ai non beneficiari, di circa mezza deviazione standard.

**Tabella 8** Modelli di regressione (ITT) e stime a variabili strumentali (LATE) per gli effetti del trattamento sul test di inglese e sugli outcome collegati alla percezione di sé.

	Modello 1		Modello 2	
	ITT	Std. err.	LATE	Std. err.
Test di inglese	0.218 **	0.105	0.275 **	0.124
Propensione di iscriversi all'università	-0.008	0.047	0.057	0.152
Fiducia nelle proprie capacità	0.429 ***	0.131	0.572 ***	0.163
Capacità di adattamento	0.353 ***	0.129	0.471 ***	0.162
Orientamento sociale	0.421 ***	0.131	0.561 ***	0.166
Propensione alla comunicazione	0.292 **	0.131	0.389 **	0.165
Capacità di apertura	0.309 **	0.128	0.413 **	0.161
Abilità nel risolvere i problemi	0.139	0.130	0.185	0.164
<i>Self-competence</i>	-0.108	0.136	-0.143	0.171
<i>Self-liking</i>	-0.166	0.133	-0.221	0.168

Nota: gli outcome sono standardizzati.

## 7. Qualche considerazione conclusiva

In questo rapporto abbiamo valutato l'efficacia del programma MOS-4 in relazione ad una svariata serie di outcome che vanno dalla conoscenza della lingua inglese all'autostima e all'apertura mentale. Il risultato principale che emerge è sicuramente l'effetto positivo esercitato da tale iniziativa su un'ampia serie di outcome collegati alla sfera psico-attitudinale quali, per esempio, l'apertura mentale, la capacità di adattarsi e la fiducia nelle proprie capacità. Tenendo presente i *caveat* presentati nel paragrafo 5 emerge un effetto positivo anche sulla conoscenza della lingua inglese. In ogni caso, gli effetti più forti si evidenziano in relazione agli outcome collegati alla percezione di sé.

I risultati presentati possono essere ulteriormente arricchiti, in relazione ai risultati scolastici, andando a considerare una serie di informazioni al momento non ancora disponibili quali i voti

<sup>18</sup> I modelli controllano per: caratteristiche socio-demografiche, carriera scolastica, *parental support* e origini sociali. Tali variabili sono inserite con l'obiettivo di guadagnare precisione nelle stime, in quanto è dimostrato come i gruppi di beneficiari e non beneficiari siano equivalenti (si veda paragrafo 6).

ottenuti dagli studenti in inglese alla fine del primo e del secondo quadrimestre, nonché l'iscrizione all'università. Infatti, al momento disponiamo solo di una generica intenzione di iscriversi all'università, entro novembre 2013 saremo, però, in grado di rilevare se gli studenti coinvolti si sono iscritti o meno all'università.

## Appendice metodologica

### *Il modello di misurazione*

Obiettivo di questa appendice è quello di fornire i dettagli tecnici relativi alla rilevazione dell'autostima e dell'apertura mentale. Essendo questi dei concetti complessi si necessita per la loro rilevazione di un elevato numero di domande che sono presentate all'intervistato sotto forma di batteria.<sup>19</sup>

**Tabella A1** Batterie di domande volte a rilevare autostima e apertura mentale

<b>Leggi attentamente le seguenti affermazioni e per ognuna indica quanto descrive te stesso e le tue opinioni su una scala da 1 (per niente) a 5 (del tutto).</b>
Ho fiducia nelle mie capacità per realizzare l'esperienza che mi viene proposta all'estero
Ho già utilizzato capacità simili a quelle che ritengo necessarie per questa esperienza all'estero
Mi sento sufficientemente pronto/a per questa esperienza all'estero
Anche se un paese straniero non mi offre le possibilità che io apprezzo nel mio paese di origine (ad esempio stili di vita, abitudini, cibo, etc.), penso di poter trovare facilmente soluzioni alternative
Se alcune attività che io pratico nel mio paese di origine non sono disponibili nel paese di accoglienza, penso di poterle sostituire facilmente con altre attività disponibili che imparerò ad apprezzare
Faccio facilmente nuove amicizie
Mi sento a mio agio quando incontro delle persone straniere
In generale mi sento a mio agio con gli altri, anche quando non li conosco
Quando parlo con uno straniero nella sua lingua, se non capisco qualche cosa di importante gli chiedo di spiegarmelo di nuovo
Anche se non parlo bene una lingua straniera provo ad utilizzare quello che so
Anche se faccio degli errori nella loro lingua mi piace provare a comunicare con persone straniere
Se uno straniero non comprende quello che dico sono disponibile a ripetere tutte le volte che è necessario
All'estero cerco di comprendere la cultura dei membri del paese d'accoglienza
E' importante avere numerosi contatti con la cultura d'accoglienza all'estero
E' interessante e piacevole conoscere e fare esperienze di nuove culture
<b>Questa domanda si occupa del come riteniate di affrontare i problemi. Ti chiediamo di rispondere indicando il vostro accordo/disaccordo su una scala che va da 1 (per nulla vero) a 4 (totalmente vero), sulle affermazioni seguenti.</b>
Riesco sempre a risolvere problemi difficili se ci provo abbastanza seriamente
Se qualcuno mi contrasta, posso trovare il modo o il sistema di ottenere ciò che voglio
Per me è facile attenermi alle mie intenzioni e raggiungere i miei obiettivi
Ho fiducia di poter affrontare efficacemente eventi inattesi
Grazie alle mie risorse, so come gestire situazioni impreviste
Posso risolvere la maggior parte dei problemi se ci metto il necessario impegno
Rimango calmo nell'affrontare le difficoltà perché posso confidare nelle mie capacità di fronteggiarle
Quando mi trovo di fronte ad un problema, di solito trovo parecchie soluzioni
Se sono in difficoltà di solito posso pensare ad una soluzione
Generalmente posso gestire qualsiasi cosa mi capita

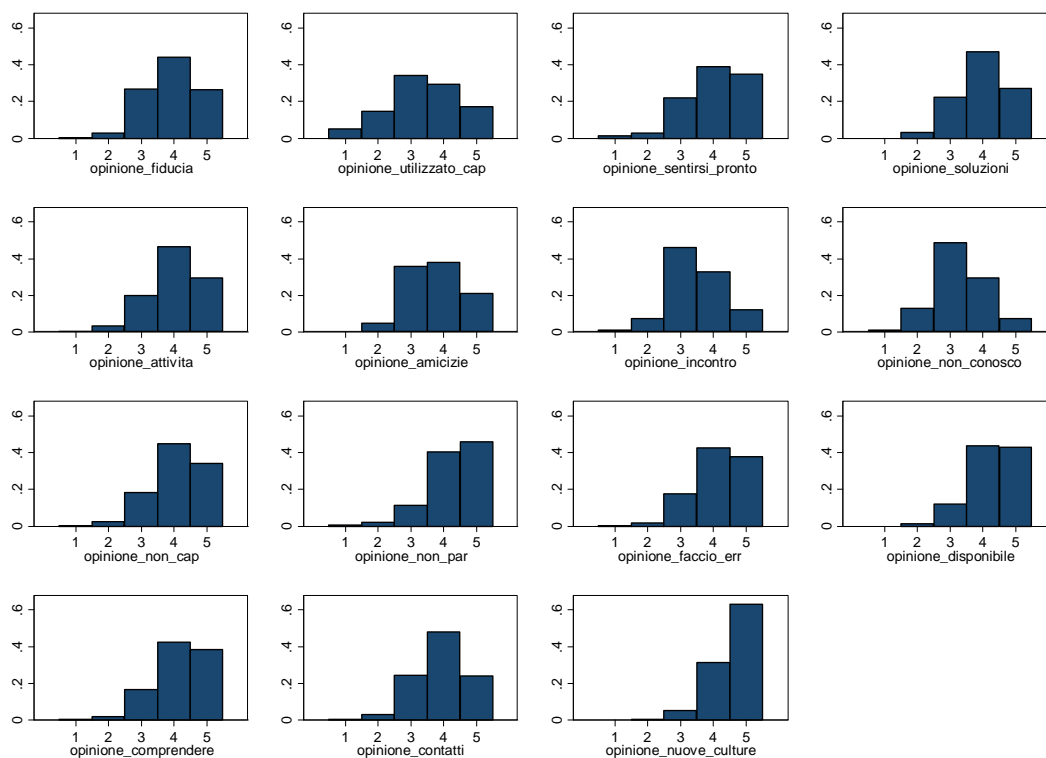
<sup>19</sup> In questo caso si parla, quindi, di batterie di domande o di batterie di item. Infatti, in letteratura, questo tipo domande sono indicate anche con il nome di item.

**Tabella A1** Batterie di domande volte a rilevare autostima e apertura mentale (continua)

<b>Questa domanda vuole invece analizzare “come tu pensi di essere” su una scala che va da 1 (fortemente in accordo) a 4 (fortemente in disaccordo).</b>
Nel complesso, sono soddisfatto di me stesso
A volte penso di non essere all’altezza delle situazioni
Credo di avere molte buone qualità
Sono in grado di fare le cose, così come la maggior parte delle persone
Sento di non avere molto di cui essere orgoglioso
A volte mi sento del tutto inutile
Sento che sono una persona di valore, almeno allo stesso livello degli altri
Vorrei avere più rispetto per me stesso
In generale, tendo a considerarmi un fallimento
Assumo un atteggiamento positivo verso me stesso

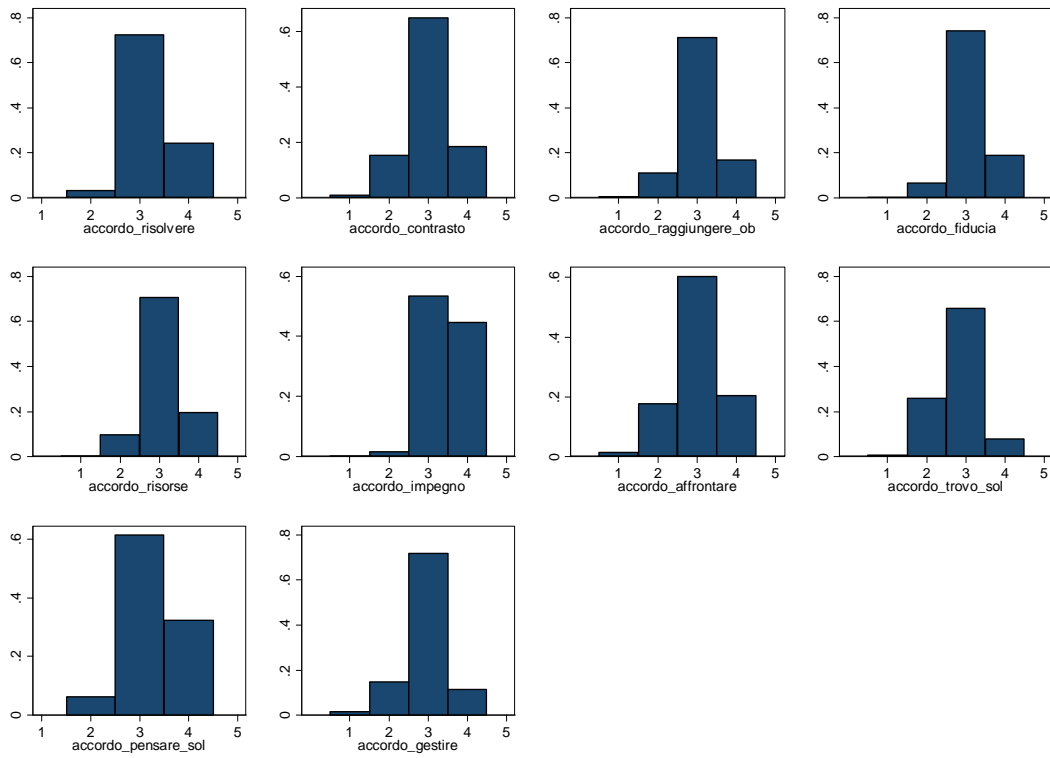
Nel nostro caso le domande utilizzate sono riportate nella tabella A1, mentre le figure A1-A3 mostrano, tramite una serie di diagrammi a barre, la distribuzione di ogni singolo item. Quello che emerge è che la distribuzione dei diversi item non è esattamente simmetrica, ma non mostra nemmeno casi di asimmetrie drammatiche.<sup>20</sup>

**Figura A1** Rappresentazione grafica della distribuzione degli item relativi all’opinione su stessi

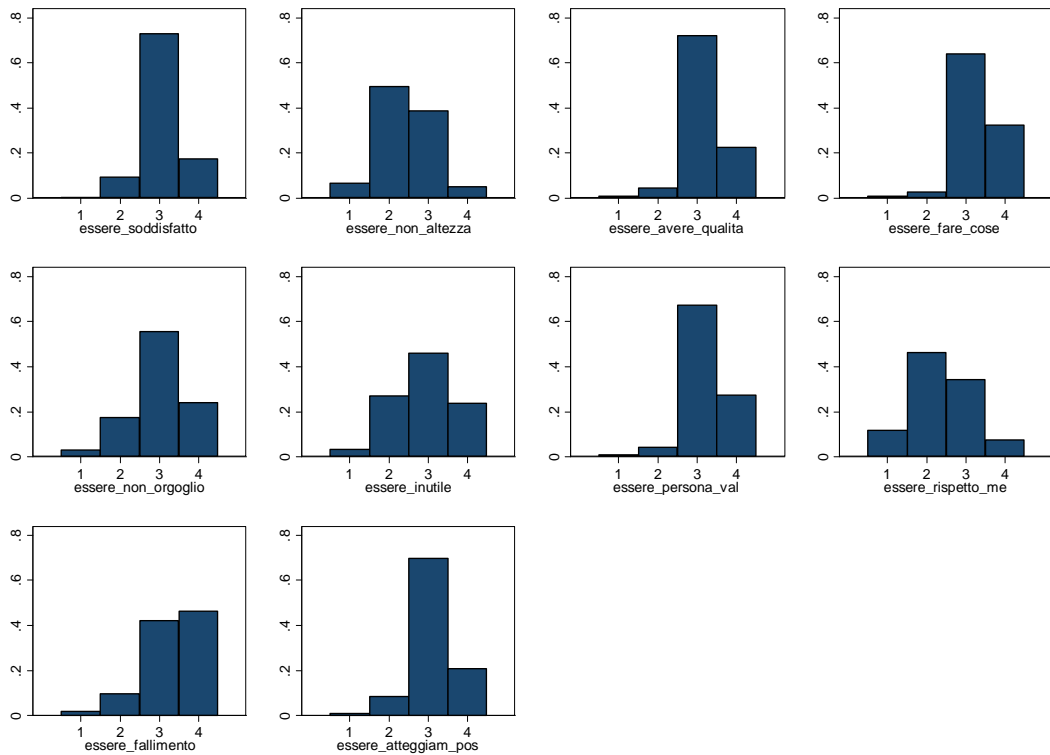


<sup>20</sup> Questa considerazione è fondamentale per il discorso relativo al metodo di stima da applicare nell’analisi fattoriale confermativa. Dati i risultati emersi siamo autorizzati a ricorrere ai metodi classici basati sulle matrici di varianza e covarianza e sulla massima verosimiglianza.

**Figura A2** Rappresentazione grafica della distribuzione degli item relativi al modo in cui si affrontano i problemi



**Figura A3** Rappresentazione grafica della distribuzione degli item relativi al modo in cui si pensa di essere



Nella teoria della misurazione si suppone che la correlazione osservata tra le variabili manifeste sia dovuta alla loro dipendenza comune da una o più variabili latenti,<sup>21</sup> delle quali gli indicatori siano indicatori validi, nel senso che ogni indicatore dipende da una sola variabile latente. Tale dipendenza può essere rappresentata in modi molteplici dando vita a un insieme di strutture differenti. Infatti, possiamo ipotizzare differenti configurazioni relative ai concetti di apertura mentale e di autostima. Per esempio, con riferimento all'autostima possiamo ipotizzare che essa sia identificata da una sola dimensione latente oppure, seguendo Tafarodi e Swann (2001) è possibile pensare ad una struttura caratterizzata da due dimensioni.<sup>22</sup>

A questo punto non resta che comparare tali configurazioni alternative al fine di valutare quella che mostra il miglior adattamento ai dati. Per poter svolgere tale operazione faremo ricorso all'analisi fattoriale confermativa e sceglieremo il modello migliore basandoci su una serie di indici ideati per valutare la bontà di un modello.<sup>23</sup> In generale, esistono indici di natura assoluta e di natura comparativa. I primi prendono in considerazione la capacità del modello di riprodurre la matrice di correlazione osservata, i secondi invece riguardano la scelta tra due o più modelli antagonisti.

Una prima misura che andremo a considerare è la statistica del chi-quadrato, essa può essere utilizzata sia per corroborare o falsificare un modello sia per confrontare modelli annidati (*nested*). Nel primo caso si suggerirà l'ipotesi nulla che il modello stimato è in grado di riprodurre adeguatamente la matrice di correlazione, con il secondo test si andrà a calcolare la differenza dei valori del test chi-quadrato nei due modelli considerati. Tale differenza si distribuisce a sua volta come un chi-quadrato con un numero di gradi di libertà pari alla differenza dei gradi di libertà dei due modelli analizzati. In tal caso l'ipotesi nulla afferma che il modello più parsimonioso (quello con il numero di gradi di libertà maggiore) si adatti ai dati egualmente bene rispetto al modello meno parsimonioso. Inoltre, prenderemo in considerazione l'RMSEA (*Root Mean Square Error of Approximation*) che ha il vantaggio di essere insensibile alla numerosità campionaria, mentre tra quelli comparativi utilizzeremo l'AIC (*Akaike Information Criterion*), il quale non ha un campo di variazione ben definito, ma è in grado di tenere in considerazione sia la bontà di adattamento del modello che il numero di gradi libertà privilegiando così modelli più parsimoniosi.<sup>24</sup>

La tabella A2 riporta la struttura dei diversi modelli con riferimento alle batterie degli item, mentre la tabella A3 mostra il grado di adattamento di tali modelli ai dati. Analizzando la prima

---

<sup>21</sup> I termini variabile latente, variabile non osservata, costrutto latente, dimensione e fattore verranno utilizzati come sinonimi.

<sup>22</sup> Va sottolineato che usiamo uno o più insiemi di item congenerici: ogni item è indicatore di un solo fattore. Se vi sono due o più fattori, la matrice dei pesi fattoriali ha un solo elemento non nullo per ogni riga e i fattori correlano fra di loro.

<sup>23</sup> Si veda Corbetta (2002) per un approfondimento sia sugli indici di adattamento ai dati che sull'analisi fattoriale confermativa.

<sup>24</sup> Come regola si prediligeranno modelli con il valore dell'AIC minori e modelli con un valore dell'RMSEA inferiore a 0,8.

batteria di domande troviamo che la struttura che meglio si adatta ai dati è quella che propone una struttura a cinque dimensioni, infatti troviamo un accordo sia se guardiamo all'RMSEA che se confrontiamo i valori degli AIC.

**Tabella A2** Struttura dei diversi modelli stimati (in grassetto i modelli preferiti)

<b>Opinione su se stessi</b>	<b>Modello 1</b> (una dimensione)	<b>Modello 2</b> (quattro dimensioni)	<b>Modello 3</b> (cinque dimensioni)
Opinione_fiducia	1	1	1
Opinione_utilizzato_cap	1	1	1
Opinione_sentirsi_pronto	1	1	1
Opinione_soluzioni	1	1	2
Opinione_attivita	1	1	2
Opinione_amicizie	1	2	3
Opinione_incontro	1	2	3
Opinione_non_conosco	1	2	3
Opinione_non_cap	1	3	4
Opinione_non_par	1	3	4
Opinione_faccio_err	1	3	4
Opinione_disponibile	1	4	4
Opinione_comprendere	1	4	5
Opinione_contatti	1	4	5
Opinione_nuove_culture	1	4	5

<b>Affrontare i problemi</b>	<b>Modello 1</b> (una dimensione)	<b>Modello 2</b> (due dimensioni)
Accordo_risolvere	1	1
Accordo_contrasto	1	2
Accordo_raggiungere_ob	1	2
Accordo_fiducia	1	1
Accordo_risorse	1	1
Accordo_impegno	1	1
Accordo_affrontare	1	1
Accordo_trovo_sol	1	1
Accordo_pensare_sol	1	2
Accordo_gestire	1	1

<b>Pensare d'essere</b>	<b>Modello 1</b> (una dimensione)	<b>Modello 2</b> (due dimensioni)	<b>Modello 3</b> (due dimensioni mod.)
Essere_soddisfatto	1	1	1
Essere_non_altezza	1	1	1
Essere_aver_qualita	1	2	2
Essere_fare_cose	1	2	2
Essere_non_orgoglio	1	2	1
Essere_inutile	1	1	1
Essere_persona_val	1	2	2
Essere_rispetto_me	1	1	1
Essere_fallimento	1	1	1
Essere_atteggiam_pos	1	2	1

Anche per quanto riguarda la terza batteria non emergono dubbi su quale sia la struttura migliore che è una struttura basata su due dimensioni così come suggerita dai lavori di Tafarodi e Swann (2001) e di Tafarodi e Milne (2002) ma leggermente modificata al fine di ottenere un adattamento migliore.<sup>25</sup> Più complicata è la situazione relativamente alla seconda batteria, infatti i

<sup>25</sup> Tali differenze sono marginali e non inficiano la struttura complessiva del lavoro di Tafarodi, infatti le discrepanze osservate possono essere imputabili ai diversi contesti culturali in cui le batterie di item sono state somministrate.



due modelli mostrano sostanzialmente lo stesso grado di adattamento ai dati. In tal caso, decidiamo di prendere come buono il modello più parsimonioso e cioè quello con una sola dimensione.

**Tabella A3** Adattamento ai dati dei diversi modelli di misurazione

	RMSEA	AIC	CHI <sup>2</sup>	g.l.
<b>Batteria item: opinione su se stessi</b>				
Una dimensione	0,093	9845,340	308,466	85
Quattro dimensioni	0,082	9792,320	253,446	84
Cinque dimensioni	0,079	9777,405	230,531	80
<b>Batteria item: come si affrontano i problemi</b>				
Una dimensione	0,021	4529,764	39,480	35
Due dimensioni	0,013	4528,133	35,850	34
<b>Batteria item: come si pensa di essere</b>				
Una dimensione	0,077	5363,635	98,246	35
Due dimensioni (Tafarodi)	0,079	5365,097	97,709	34
Due dimensioni (Tafarodi mod.)	0,056	5333,494	66,105	34

Nota: g.l.=gradi di libertà; CHI<sup>2</sup>=valore del test chi-quadrato.

Dalle analisi svolte emergono quindi le dimensioni<sup>26</sup> che avevamo evidenziato nel corso del secondo paragrafo e che possiamo ricapitolare come segue:

Opinione su se stessi (cinque dimensioni)

1. Fiducia nelle proprie capacità (D1).
2. Capacità di adattamento (D2).
3. Orientamento sociale (D3).
4. Propensione alla comunicazione (D4).
5. Capacità di apertura (D5).

Affrontare i problemi (una dimensione):

1. Abilità nel risolvere i problemi (D6).

Pensare d'essere (due dimensioni):

1. Valutazione oggettiva delle proprie competenze (*self-competence*) (D7).
2. Accettazione delle proprie qualità (*self-liking*) (D8).

Un ultimo elemento che resta da verificare riguarda il problema della cosiddetta invarianza del modello di misurazione tra i diversi gruppi. In altre parole, l'obiettivo è quello di testare se i modelli di misurazione appena esposti sono gli stessi tra beneficiari e non beneficiari. In letteratura si

<sup>26</sup> Va precisato che le etichette date alle diverse dimensioni rappresentano una nostra interpretazione.

possono trovare diversi tipi di invarianza; in questa sede considereremo i principali: configurazionale; metrica (*weak*) e scalare (*strong*).

**Tabella A4** Risultati dell'analisi fattoriale confermativa tra i diversi gruppi

Opinione su se stessi	Beneficiari					Non beneficiari				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
Opinione_fiducia	0,857***					0,857***				
Opinione_utilizzato_cap	0,650***					0,558***				
Opinione_sentirsi_pronto	0,723***					0,664***				
Opinione_soluzioni		0,722***					0,722***			
Opinione_attivita		0,613***					0,804***			
Opinione_amicizie			0,645***					0,645***		
Opinione_incontro			0,795***					0,894***		
Opinione_non_conosco			0,764***					0,638***		
Opinione_non_cap				0,684***					0,684***	
Opinione_non_par				0,730***					0,670***	
Opinione_faccio_err				0,807***					0,684***	
Opinione_disponibile				0,521***					0,468***	
Opinione_comprendere					0,818***					0,818***
Opinione_contatti					0,605***					0,663***
Opinione_nuove_culture					0,505***					0,752***
<b>Affrontare i problemi</b>	<b>D6</b>					<b>D6</b>				
Accordo_risolvere	0,647***					0,647***				
Accordo_contrasto	0,418***					0,434***				
Accordo_raggiungere_ob	0,522***					0,422***				
Accordo_fiducia	0,525***					0,577***				
Accordo_risorse	0,564***					0,584***				
Accordo_impegno	0,597***					0,544***				
Accordo_affrontare	0,620***					0,614***				
Accordo_trovo_sol	0,598***					0,668***				
Accordo_pensare_sol	0,433***					0,513***				
Accordo_gestire	0,552***					0,592***				
<b>Pensare d'essere</b>	<b>D7</b>					<b>D7</b>				
Essere_soddisfatto	0,742***					0,742***				
Essere_non_altezza	0,598***					0,608***				
Essere_aver_qualita		0,455***						0,455***		
Essere_fare_cose		0,622***						0,482***		
Essere_non_orgoglio	0,620***					0,516***				
Essere_inutile	0,584***					0,654***				
Essere_persona_val		0,655***						0,675***		
Essere_rispetto_me	0,420***					0,552***				
Essere_fallimento	0,690***					0,697***				
Essere_atteggiam_pos	0,550***					0,744***				

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Al fine di effettuare ipotesi sull'invarianza si usano i modelli multi-gruppo. La strategia si articola attraverso una serie di vincoli che vengono imposti per poter esaminare l'invarianza fattoriale. Si tratta sostanzialmente di una serie di test che partendo da un modello iniziale meno vincolato (invarianza configurazionale) si giunge al modello più vincolato (invarianza scalare). Il primo passo consiste nel presentare l'invarianza configurazionale (tabella A4), che prevede di applicare lo stesso modello ai due gruppi sotto esame. In questo caso l'obiettivo è quello di confermare una similarità qualitativa nella struttura fattoriale tra i due gruppi e di guardare

semplicemente che i *factor loadings* siano significativi e di intensità adeguata tra i due gruppi.<sup>27</sup> Dai dati presentati nella tabella A4 emerge in modo evidente che tutti i parametri sono statisticamente significativi e di intensità adeguata. Il passo successivo consiste nel testare i restanti tipi di invarianza. L'invarianza metrica impone al modello che i *factor loadings* siano uguali tra i diversi gruppi, mentre quella scalare assume che oltre ai *factor loadings* ci sia invarianza anche nelle intercette degli item considerati. In quest'ultimo caso vengono inserite nell'analisi anche le medie delle variabili osservate, che solitamente sono poste pari a 0. L'uguaglianza delle medie permette di testare le differenze tra le medie dei fattori latenti tra i diversi gruppi. In altre parole, se tale modello regge alla prova dei dati allora diviene possibile testare le differenze tra beneficiari e non beneficiari relativamente ai costrutti latenti individuati. In altre parole, si garantisce che i fattori latenti funzionano allo stesso modo tra i diversi gruppi. I confronti tra gli indici di adattamento dei diversi modelli sono riportati nella tabella A5. A prima vista possiamo notare come tutti i modelli mostrino un buon adattamento ai dati e che per i modelli a invarianza scalare l'RMSEA mostra il miglior adattamento ai dati. Diverso è il discorso relativo all'AIC che fa prevalere sempre, anche se di molto poco, il modello a invarianza metrica. In tutti i casi, però, se costruiamo il test basato sul confronto del chi-quadrato troviamo che il modello vincente risulta essere quello a invarianza scalare. Ne consegue che i gruppi dei beneficiari e dei controlli condividono lo stesso modello di misurazione e diviene sensato comparare le medie dei fattori latenti.

**Tabella A5** Indici di adattamento ai dati dei diversi modelli di invarianza

	RMSEA	AIC	CHI2	g.l.
<b>Batteria item: opinione su se stessi</b>				
Invarianza configurazionale	0,088	9780,645	346,876	160
Invarianza metrica	0,086	9775,101	361,331	170
Invarianza scalare	0,083	9821,000	367,231	180
<b>Batteria item: come si affrontano i problemi</b>				
Invarianza configurazionale	0,051	4543,137	97,653	70
Invarianza metrica	0,045	4530,701	103,217	79
Invarianza scalare	0,038	4556,942	107,458	88
<b>Batteria item: come si pensa di essere</b>				
Invarianza configurazionale	0,057	5330,508	101,044	68
Invarianza metrica	0,055	5324,608	111,144	76
Invarianza scalare	0,051	5354,698	117,234	84

Nota: g.l.=gradi di libertà; CHI<sup>2</sup>=valore del test chi-quadrato.

<sup>27</sup> I *factor loadings* rappresentano la correlazione esistente tra il fattore latente e la variabile osservata. Secondo una regola del pollice ampiamente utilizzata in letteratura l'intensità di tali parametri è ritenuta buona se supera il valore di 0,4.

### ***Identificazione e stima***

Il ricorso alla randomizzazione consente di poter identificare l'effetto dell'intervento dal semplice confronto tra il gruppo dei beneficiari e quello di controllo. Più specificatamente, qualora tutti gli studenti rispettino le regole di ammissibilità, la stima dell'effetto su una certa variabile di interesse è data dalle differenza tra le medie della variabile per ammissibili e non ammissibili. Se indichiamo l'outcome con  $Y$  e l'ammissibilità con  $Z$ , che assume valore 1 per gli ammissibili e 0 per i non ammissibili, quanto detto può essere riassunto nella seguente formula, che altro non è che una differenza tra medie:

$$ITT = E(Y | Z = 1) - E(Y | Z = 0). \quad (1)$$

Come esposto nel paragrafo 3, però, nel caso in questione alcuni studenti ammessi a partecipare hanno poi non beneficiato del programma, mentre alcuni non ammessi hanno comunque partecipato a programmi simili. In questa situazione la (1) ci fornisce una stima dell'effetto medio di offrire un trattamento e non di riceverlo. Questa quantità in letteratura è nota come *Intention-To-Treat* (ITT). Si può dimostrare facilmente che, per poter stimare l'effetto medio di aver beneficiato del programma, questa quantità deve essere riscalata per la proporzione di *compliers* all'interno dei due gruppi (Angrist e Pischke 2009: p. 164). Ne consegue che l'effetto del programma può essere calcolato sostituendo le controparti empiriche di questo rapporto:

$$LATE = \frac{E(Y | Z = 1) - E(Y | Z = 0)}{P(D = 1 | Z = 1) - P(D = 1 | Z = 0)}; \quad (2)$$

noto anche come *Local Average Treatment Effect* (LATE). Tale effetto è definito locale, poiché identifica l'effetto dell'aver partecipato al programma limitatamente ai *compliers*<sup>28</sup>. In altre parole, la differenza tra medie presentata in (1) viene ponderata per la proporzione di *compliers*. Questa operazione può essere svolta per via parametrica, ricorrendo al metodo delle variabili strumentali. Seguendo questo approccio lo strumento  $Z$  è l'ammissibilità al trattamento, mentre la variabile causale di interesse  $D$  è il beneficiare o meno del programma.

In questa logica la tabella 1 di pagina 5 rappresenta la distribuzione congiunta dello strumento e della variabile causale di interesse. Questa consente di dire che, dato che gli studenti nel 90% dei casi hanno seguito la regola di assegnazione data dall'estrazione, la variabile strumentale utilizzata è adeguata e che quindi le stime che mostreremo possono considerarsi robuste.

---

<sup>28</sup> Per maggiori approfondimenti si veda anche Bloom (2008).

## Riferimenti bibliografici

- Angrist, J. e Pischke, J. S. (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, Princeton University Press.
- Berk, R. A. (2005) 'Randomized Experiments as the Bronze Standard', *Journal of Experimental Criminology*, 1(4): 417–33.
- Bertrand, M. e Mullainathan, S. (2001) 'Do people mean what they say? Implication for subjective survey data', *The American Economic Review*, 91(2): 67-72.
- Bloom, H.S. (2008) The core analytics of randomized experiments for social research, in P. Alasuutari, L. Bickman e J. Brannen (a cura di), *The SAGE Handbook of Social Research Methods*, Thousand Oaks(CA), Sage Publications.
- Corbetta, P. (2002) *Metodi di analisi multivariata per le scienze sociali. I modelli di equazioni strutturali*, Bologna, Il Mulino.
- European Commission (2000) *The Lisbon European Council - an agenda of economic and social renewal for Europe*, Brussels, Commission of the European Communities [disponibile al sito: [http://www.europarl.europa.eu/summits/lis1\\_en.htm](http://www.europarl.europa.eu/summits/lis1_en.htm)]
- Heckman, J.J. e Smith, J. (1995) 'Assessing the case for social experiments', *Journal of Economic Perspectives*, 9 (2): 85-110.
- Houser, C., Brannstrom, C., Quiring, S.M. e Lemmons, K.K. (2011) 'Study Abroad Field Trip Improves Test Performance through Engagement and New Social Networks', *Journal of Geography in Higher Education*, 35(4): 513–528.
- Ismail, B., Morgan, M., e Hayes, K. (2006) 'Effect of short study abroad course on student openness to diversity', *Journal of Food Science Education*, 5(1): 15-18.
- Martini, A. e Trivellato, U. (2011) *Sono soldi ben spesi? Perché e come valutare l'efficacia delle politiche pubbliche*, Venezia, Marsilio Editori.
- McLeod, M. e Wainwright, P. (2008) 'Researching the Study Abroad Experience', *International Journal of Intercultural Relations*, 13(1): 66–71.
- Sutton, R.C. e Rubin, D.L. (2010) *Documenting the Academic Impact of Study Abroad: Final Report of the GLOSSARI Project*, paper presented at the NAFSA Annual Conference.
- Tafarodi, R. e Milne, A. (2002) 'Decomposing global self-esteem', *Journal of Personality*, 70(4): 443–483.
- Tafarodi, R. e Swann, W. (2001) 'Two-dimensional self-esteem: theory and measurement', *Personality and Individual Differences* 31(5): 653–673.

Waxin, M.F. (2004) 'Expatriates' interaction adjustment: the direct and moderator effects of culture of origin', *International Journal of Intercultural Relations*, 28(1): 61–79.

Williams, T.R. (2005) 'Sensitivity Exploring the Impact of Study Abroad on Students' Intercultural Communication Skills: Adaptability and Sensitivity', *Journal of Studies in International Education*, 9(4): 356–371.