

NOTA METODOLOGICA

Misurare la povertà educativa. Il lavoro della Commissione scientifica interistituzionale

Il lavoro della Commissione

Nel 2023, l'Istat ha istituito una Commissione scientifica interistituzionale con l'obiettivo di definire e misurare la povertà educativa e individuare le aree territoriali maggiormente interessate dal fenomeno.

La Commissione ha concluso i suoi lavori nel gennaio 2026. I principali risultati vengono presentati in occasione del convegno Misurare la povertà educativa. Risultati e prospettive del lavoro della Commissione scientifica interistituzionale organizzato a Napoli il 16 e il 17 aprile 2026.

Dato il carattere multidimensionale della povertà educativa, la Commissione ha selezionato 78 indicatori statistici per misurarla, rappresentativi delle componenti più rilevanti del fenomeno.

Per il calcolo degli indicatori sono state utilizzate fonti amministrative (Ministero dell'istruzione e del merito, SIAE), indagini Istat (Indagine sull'inserimento degli alunni con disabilità, Indagine sui servizi educativi per la prima infanzia, Indagine sulle biblioteche e Indagine sui musei e le istituzioni similari, Indagine multiscopo sugli aspetti della vita quotidiana, Indagine sul reddito e le condizioni di vita, Rilevazione sulle forze di lavoro, Indagine I cittadini e il tempo libero e l'Indagine Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri), il Registro statistico Istat delle istituzioni non profit e le Rilevazioni nazionali degli apprendimenti delle studentesse e degli studenti (INVALSI).

La Tabella 1 distingue le fonti campionarie da quelle censuarie e riporta per ciascuna il numero di indicatori di povertà educativa selezionati dalla Commissione.

Tabella 1 - Fonti, tipo di indagine e numero di indicatori

| Fonte | Tipo | Numero indicatori |
|--|-------------|-------------------|
| Istat, Indagine multiscopo sugli aspetti della vita quotidiana (AVQ) | Campionaria | 21 |
| Istat, Indagine Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri (B&R) | Campionaria | 17 |
| Istat, Rilevazione sulle Forze di lavoro (FdL) | Campionaria | 4 |
| Istat, Indagine sul reddito e le condizioni di vita (Eu-Sile) | Campionaria | 2 |
| Istat, Indagine sull'inserimento scolastico degli alunni con disabilità nelle scuole statali e non statali (a) | Censuaria | 6 |
| Istat, Indagine su nidi e servizi integrativi per la prima infanzia | Censuaria | 3 |
| Istat, Indagine sulle biblioteche | Censuaria | 1 |
| Istat, Indagine sui musei e le istituzioni similari | Censuaria | 1 |
| Istat, Registro statistico delle istituzioni non profit | Censuaria | 1 |
| Ministero dell'istruzione e del merito (MIM) | Censuaria | 18 |
| INVALSI, Rilevazioni nazionali degli apprendimenti delle studentesse e degli studenti | Censuaria | 3 |
| SIAE | Censuaria | 1 |
| Totale | | 78 |

(a) Per questa fonte, gli indicatori di povertà educativa sono stati calcolati sulla componente censuaria dell'indagine, ovvero quella relativa alle scuole. Non sono stati calcolati indicatori sulla componente campionaria dell'indagine, ovvero quella sugli alunni con disabilità.

Per il livello territoriale la scelta è ricaduta sull'incrocio tra regione e grado di urbanizzazione del comune. È stata utilizzata la classificazione Grado di urbanizzazione (Degree of Urbanisation, DEGURBA), sviluppata da Eurostat in collaborazione con OECD, che distingue tra città, piccole città e sobborghi, zone scarsamente popolate. Sono state così individuate 62 unità territoriali di analisi: 19 regioni e 2 Province autonome disaggregate per 3 gradi di urbanizzazione, con l'eccezione della Valle d'Aosta dove non ci sono comuni classificati come "città".

Per gli indicatori costruiti sulla base dei risultati di indagini campionarie, per le quali l'incrocio regione per DEGURBA non è un dominio pianificato di stima, l'Istat ha proceduto al calcolo delle stime dirette (paragrafo

3), a cui il Centro DAGUM - Istituzione membro della Commissione - ha applicato metodi di stima per piccole aree al fine di migliorarne l'accuratezza (paragrafo 4).

A partire dai 78 indicatori, sono stati costruiti alcuni indici composti per sintetizzare le informazioni e realizzare una mappatura del territorio nazionale. Per la costruzione degli indici si è utilizzata la metodologia dell'Adjusted Mazziotta-Pareto Index (AMPI¹) (paragrafo 5).

Gli indicatori di povertà educativa

La Commissione ha definito la povertà educativa un fenomeno multidimensionale complesso - frutto del contesto familiare e sociale in cui i bambini e i ragazzi (0-19 anni) vivono - distinguendola in povertà di risorse e povertà di esiti.

La Commissione ha poi sviluppato un quadro concettuale funzionale alla misurazione quantitativa del fenomeno. Il quadro è articolato in due domini: risorse ed esiti, ciascuno dei due domini è, poi, articolato in una serie di dimensioni e sottodimensioni che danno conto della molteplicità di aspetti legati al fenomeno.

In totale sono stati selezionati 78 indicatori: 60 indicatori di risorse e 18 indicatori di esiti.

Gli indicatori sono stati concepiti e calcolati come misure di deprivazione con l'obiettivo di dare un quadro delle situazioni di disagio in termini di risorse e di esiti già a partire dalla lettura della matrice di dati. Per questo tutti gli indicatori (tranne poche eccezioni) sono hanno polarità positiva: al crescere del valore dell'indicatore aumenta la povertà educativa negli esiti e il rischio di povertà educativa nelle risorse.

Tabella 2 - Framework di povertà educativa e numero di indicatori

| | Dominio | Dimensione | Sottodimensione | Numero indicatori | | |
|---------------------------------------|---|---|--|---------------------|----------------|-------------|
| | | | | Per sottodimensione | Per dimensione | Per dominio |
| Povertà educativa | Risorse | Contesto familiare | Status socio-economico e culturale | 6 | 16 | 60 |
| | | | Abitazione e beni materiali | 3 | | |
| | | | Relazioni, partecipazione sociale e culturale dei genitori | 7 | | |
| | | Contesto scolastico | Offerta dei servizi educativi | 9 | 24 | |
| | | | Adeguatezza dei servizi educativi | 10 | | |
| | | | Fruizione dei servizi educativi | 5 | | |
| | Contesto territoriale, sociale e culturale | Luogo di vita | 10 | 20 | | |
| | | Relazioni, partecipazione sociale e culturale dei bambini e ragazzi | 10 | | | |
| | Esiti | Competenze cognitive | Alfabetiche, numeriche, linguistiche | 3 | 4 | 18 |
| | | | Digitali, scientifiche, finanziarie, civiche e di cittadinanza | 1 | | |
| Competenze personali e sociali | | Relazionali | 3 | 14 | | |
| | | Emotive | 3 | | | |
| | | Interazione fiduciaria | 3 | | | |
| | | Regolarità del percorso scolastico | 5 | | | |

Le stime dirette

Per gli indicatori di fonte amministrativa il livello territoriale regione per grado di urbanizzazione è facilmente ricavabile dall'aggregazione dei dati comunali. Ciò non vale per le indagini di fonte campionaria. Per gli indicatori derivati da tali fonti, Istat ha proceduto al calcolo delle stime dirette, e successivamente il centro DAGUM - Istituzione membro della Commissione - ha poi applicato metodi di stima per piccole aree al fine di migliorarne l'accuratezza.

Gli indicatori di fonte campionaria erano in origine 53 provenienti da cinque indagini campionarie realizzate dall'Istat: indagine Aspetti della vita quotidiana (AVQ), Indagine sul reddito e le condizioni di vita (Eu-SILC), Rilevazione sulle forze di lavoro (FdL), Indagine I cittadini e il tempo libero (CTL) e l'indagine Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri (B&R).

¹ Mazziotta, M., Pareto, A. Statistica per gli indici composti. Giappichelli Editore. 2024.

Tabella 3 – Fonti campionarie per numero di indicatori sui quali sono state calcolate le stime dirette e successive stime per piccole aree

| Fonte | Numero indicatori |
|--|-------------------|
| Indagine multiscopo sugli aspetti della vita quotidiana (AVQ) | 24 |
| Indagine Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri (B&R) | 18 |
| Rilevazione sulle Forze di lavoro (FdL) | 4 |
| Istat, Indagine sul reddito e le condizioni di vita (Eu-Silc) | 2 |
| Indagine I cittadini e il tempo libero (CTL) | 5 |
| Totale | 53 |

Le stime dirette dei 53 indicatori di fonte campionaria sono state calcolate per un livello territoriale così composto: Regione (comprese le Province autonome di Bolzano/Bozen e Trento) incrociata per il grado di urbanizzazione, un dominio di stima non pianificato nel disegno di campionamento di queste 5 indagini.

1. "Città" o "Zone densamente popolate";
2. "Piccole città e sobborghi" o "Zone a densità intermedia di popolazione";
3. "Zone rurali" o "Zone scarsamente popolate".

Per valutarne l'accuratezza, ogni stima è accompagnata dall'intervallo di confidenza al 95% che rappresenta un intervallo di valori entro il quale, con il livello di fiducia prefissato, si ritiene sia compreso il vero valore della popolazione. Per il calcolo degli errori di campionamento delle indagini condotte dall'Istat sulle famiglie viene correntemente utilizzata una procedura informatica sviluppata nell'ambito dell'Istituto ².

Di seguito le informazioni specifiche e i riferimenti metodologici per ogni singola indagine campionaria.

Indagine multiscopo sugli aspetti della vita quotidiana

L'indagine Aspetti della Vita Quotidiana (AVQ) fa parte di un sistema integrato di indagini sociali – le Indagini Multiscopo sulle famiglie – e rileva informazioni fondamentali relative alla vita quotidiana degli individui e delle famiglie. Le informazioni raccolte da un campione di circa 20.000 famiglie consentono di conoscere le abitudini dei cittadini e i problemi che essi affrontano ogni giorno.

A livello territoriale i domini di stima pianificati sono le regioni e le province autonome di Trento e Bolzano/Bozen.

Per approfondimenti sugli aspetti metodologici dell'indagine si veda il seguente link:

<https://www.istat.it/wp-content/themes/EGPbs5-child/microdata/download.php?id=%2F60%2F2023%2F01%2FNota.pdf>

La popolazione di riferimento per il calcolo degli indicatori di povertà educativa è il sotto-insieme di individui con età fino a 19 anni. Sono state calcolate le stime dirette regione per grado di urbanizzazione di 22 indicatori al 2024, 1 indicatore al 2023 e 1 indicatore come media triennale 2022-2024. Per quanto riguarda quest'ultimo indicatore, la stima della varianza ha richiesto un calcolo approssimato in serie di Taylor sul rapporto di due stime di varianza (numeratore e denominatore).

² Zardetto D. (2015). ReGenesees: an Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys, (extended version). *Journal of Official Statistics*, 31(2):177-203.

Gli indicatori selezionati per misurare la povertà educativa sono i seguenti.

Bambini e ragazzi di 0-19 anni con genitori che vedono raramente o mai amici
Bambini e ragazzi di 0-19 anni con genitori che non svolgono attività di partecipazione sociale
Bambini e ragazzi di 0-19 anni con genitori che non svolgono attività di partecipazione civica e politica
Bambini e ragazzi di 0-19 anni con genitori che non leggono libri nel tempo libero
Bambini e ragazzi di 0-19 anni con genitori che non hanno visto spettacoli fuori casa
Bambini e ragazzi di 0-19 anni con genitori che non vanno in biblioteca
Bambini e ragazzi di 0-19 anni con genitori che non praticano sport
Bambini e ragazzi di 0-19 anni che vivono in abitazioni dove ci sono al massimo 25 libri
Bambini e ragazzi di 0-19 anni che vivono in abitazioni che presentano 3 o più problemi su 6 considerati
Bambini e ragazzi di 0-19 anni che vivono in famiglie che dichiarano condizioni economiche scarse o insufficienti
Bambini e ragazzi di 0-19 anni che vivono in una zona dove non ci sono piste ciclabili
Bambini e ragazzi di 0-19 anni che vivono in una zona con 3 o più problemi su 10 considerati
Studenti fino a 19 anni iscritti ad una scuola secondaria di II grado che impiegano più di 30 minuti per andare a scuola
Ragazzi di 14-19 anni che non svolgono attività di partecipazione civica e politica
Bambini e ragazzi di 6-19 anni che non leggono libri nel tempo libero
Bambini e ragazzi di 3-19 anni che non hanno visto spettacoli fuori casa
Bambini e ragazzi di 3-19 anni che non vanno in biblioteca
Bambini e ragazzi di 3-19 anni che non praticano sport
Ragazzi di 14-19 anni che non svolgono attività di volontariato
Ragazzi di 14-19 anni che non svolgono attività di partecipazione sociale
Bambini e ragazzi di 0-19 anni che vivono in una zona dove non ci sono parchi, giardini o altro verde pubblico
Ragazzi di 16-19 anni con competenze digitali almeno di base
Ragazzi di 14-19 anni che esprimono un voto di soddisfazione per la vita compreso tra 0-7
Ragazzi di 14-19 anni con scarsa fiducia interpersonale: pensano sia necessario stare molto attenti agli altri

Indagine I cittadini e il tempo libero

L'indagine "I cittadini e il tempo libero" è una rilevazione campionaria realizzata dall'Istat nel 2024 con l'obiettivo principale di studiare i comportamenti degli individui nel tempo libero. In particolare, vengono osservate attività culturali, pratiche sportive, uso dei media e delle tecnologie, oltre ad altre forme di svago. La popolazione di riferimento è costituita da tutte le famiglie residenti in Italia e dagli individui che le compongono a partire da 0 anni (sono escluse le persone che risiedono in istituti di convivenza).

A livello territoriale i domini di stima pianificati sono le regioni e le province autonome di Trento e Bolzano/Bozen.

Per approfondimenti sugli aspetti metodologici dell'indagine si veda il seguente link:

<https://www.istat.it/wp-content/uploads/2025/12/Nota-metodologica-1.pdf>

La popolazione di riferimento per il calcolo degli indicatori di povertà educativa è quella degli individui fino a 19 anni. Sono state calcolate le stime dirette regione per grado di urbanizzazione di 5 indicatori riferiti al 2024.

Gli indicatori selezionati per misurare la povertà educativa sono i seguenti.

Bambini di 2-5 anni che, al di fuori dell'orario scolastico, non leggono, non sfogliano, non colorano libri o albi illustrati
Bambini di 2-5 anni che non fanno mai giochi di movimento
Bambini e ragazzi di 6-19 anni che dichiarano di non avere (o giudicano difficile raggiungere) alcuni luoghi, spazi culturali
Ragazzi di 14-19 anni che non hanno mai sentito parlare di intelligenza artificiale
Bambini di 3-10 anni ai quali entrambi i genitori non leggono fiabe, storie e non raccontano/inventano fiabe storie.

Indagine Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri

L'indagine "Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri" è condotta dall'Istat con l'obiettivo di raccogliere informazioni su alcuni aspetti fondamentali della vita quotidiana di ragazzi e ragazze tra gli 11 e i 19 anni residenti in Italia. Specifica attenzione viene dedicata ai ragazzi di cittadinanza straniera.

A livello territoriale i domini di stima pianificati sono le regioni e le province autonome di Trento e Bolzano/Bozen.

Per approfondimenti sugli aspetti metodologici dell'indagine si veda il seguente link:

<https://www.istat.it/wp-content/uploads/2025/06/Nota-metodologica-1.pdf>

La popolazione di riferimento per il calcolo degli indicatori di povertà educativa è quella dei bambini e ragazzi con età compresa fra 11 e 19 anni. Sono state calcolate le stime dirette regione per grado di urbanizzazione di 18 indicatori riferiti al 2023.

Gli indicatori selezionati per misurare la povertà educativa sono i seguenti.

| |
|---|
| Bambini e ragazzi di 11-19 anni che vivono in una zona senza spazi verdi, parchi e giardini |
| Bambini e ragazzi di 11-19 anni che vivono in una zona con difficoltà di collegamento con i mezzi pubblici |
| Bambini e ragazzi di 11-19 anni che vivono in una zona con mancanza di cinema, teatri, luoghi per spettacoli dal vivo |
| Bambini e ragazzi di 11-19 anni che frequentano gli amici meno di una volta a settimana |
| Bambini e ragazzi di 11-19 anni che non fanno parte o non frequentano associazioni/organizzazioni |
| Bambini e ragazzi di 11-19 anni che al di fuori dell'orario scolastico non svolgono lezioni di inglese o di altra lingua |
| Bambini e ragazzi di 11-19 anni che non partecipano a feste, cene o altre occasioni di incontro organizzate da loro o dai loro amici |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco/per niente soddisfatti delle relazioni amicali |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco/per niente soddisfatti delle relazioni familiari |
| Bambini e ragazzi di 11-19 anni che non hanno amici con cui confidarsi |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco/per niente soddisfatti della vita |
| Bambini e ragazzi di 11-19 anni che si dichiarano molto d'accordo con l'affermazione: Quando fallisco ho paura di non avere abbastanza talento |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco/per niente d'accordo con l'affermazione: Non ho paura di esprimere i miei pensieri e le mie idee anche di fronte a chi la pensa diversamente |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco o per niente d'accordo con l'affermazione: La fiducia in me stesso/a mi ha permesso di superare momenti difficili |
| Bambini e ragazzi di 11-19 anni che si dichiarano poco/per niente d'accordo con l'affermazione: Quando sono in una situazione difficile di solito riesco a trovare la via d'uscita |
| Bambini e ragazzi di 11-19 anni che non si ritengono molto bravi a scuola |
| Bambini e ragazzi di 11-19 anni che non praticano sport al di fuori dell'orario scolastico |
| Bambini e ragazzi di 11-19 anni che non hanno fatto vacanze negli ultimi 12 mesi |

Rilevazione sulle Forze di lavoro

La “Rilevazione sulle Forze di Lavoro” è un’indagine campionaria realizzata dall’Istat con l’obiettivo principale di fornire informazioni sul mercato del lavoro. Nel tempo l’indagine ha adottato le linee guida internazionali e dal 2004 segue i regolamenti europei per garantire la comparabilità e l’armonizzazione delle informazioni prodotte a livello internazionale e in particolare a livello comunitario.

Le informazioni rilevate presso la popolazione costituiscono la base sulla quale vengono derivate le stime ufficiali degli occupati e dei disoccupati, nonché le informazioni sui principali aggregati dell’offerta di lavoro – professione, settore di attività economica, ore lavorate, tipologia e durata dei contratti, istruzione e formazione.

La popolazione di riferimento è costituita da tutte le famiglie residenti in Italia e dagli individui che le compongono a partire da 0 anni (sono escluse le persone che risiedono in istituti di convivenza). A livello territoriale i domini di stima pianificati sono le regioni, le province autonome di Trento e Bolzano/Bozen, su base trimestrale e le province e i 14 grandi comuni per la media annua.

Per approfondimenti sugli aspetti metodologici dell’indagine si veda il seguente link:

<https://www.istat.it/informazioni-sulla-rilevazione/forze-lavoro/>

Nell’ambito di questa indagine sono stati individuati come rilevanti per lo studio della povertà educativa i seguenti 4 indicatori di cui sono state calcolate le stime dirette regione per grado di urbanizzazione.

| |
|---|
| Bambini e ragazzi di 0-19 anni con genitori con basso titolo di studio |
| Bambini e ragazzi di 0-19 anni con genitori non occupati |
| Bambini e ragazzi di 0-19 anni con genitori con occupazioni vulnerabili |
| Uscita precoce dal sistema di istruzione e formazione (dispersione esplicita) (ragazzi di 18-24 anni) |

Indagine sul reddito e le condizioni di vita

Il progetto Eu-Silc (European Union Statistics on Income and Living Conditions) costituisce una delle principali fonti di dati per i rapporti periodici dell’Unione europea sulla situazione sociale e sulla diffusione del disagio economico nei Paesi membri. Gli indicatori previsti dal Regolamento sono incentrati sul reddito e sull’esclusione sociale, in un approccio multidimensionale al problema e con una particolare attenzione agli

aspetti di deprivazione materiale. L'Italia partecipa al progetto con l'indagine su "Reddito e condizioni di vita delle famiglie", svolta a cadenza annuale a partire dal 2004.

A livello territoriale i domini di stima pianificati sono le regioni e le province autonome di Trento e Bolzano/Bozen.

Per approfondimenti sugli aspetti metodologici dell'indagine si veda il seguente link:

<https://www.istat.it/informazioni-sulla-rilevazione/eu-silc/>

La popolazione di riferimento per il calcolo degli indicatori di povertà educativa è quella degli individui fino a 19 anni. Sono state calcolate le stime dirette regione per grado di urbanizzazione di 2 indicatori riferiti al 2024.

Bambini e ragazzi di 0-19 anni che vivono in famiglie che non possono sostenere spese impreviste
Bambini e ragazzi di 0-19 anni che vivono in abitazioni sovraffollate

La stima per piccole aree

Una volta calcolate le stime dirette degli indicatori di fonte campionaria il centro DAGUM - Istituzione membro della Commissione - ha applicato metodi di stima per piccole aree al fine di migliorarne l'accuratezza.

È stato applicato il modello di Fay-Herriot (FH), introdotto nel 1979 (Fay e Herriot, 1979) per stimare il reddito medio pro capite a livello locale negli Stati Uniti, ed ancora oggi uno dei modelli di riferimento nella stima per piccole aree (SAE). Questo modello viene utilizzato quando la dimensione campionaria ridotta non consente stime affidabili per alcune aree, spesso non pianificate nel disegno campionario, come in questo caso. Poiché richiede informazione ausiliaria a livello aggregato, il modello FH è particolarmente utile in contesti complessi come quello della povertà educativa, dove spesso mancano dati individuali o coerenti tra le fonti. Il modello prevede numerose estensioni (Rao e Molina, 2015). In questo caso è stata applicata la versione base del modello, spiegata nel seguito, poiché l'estensione spaziale del modello di Fay-Herriot non risulta direttamente applicabile alle aree subregionali individuate secondo la classificazione DEGURBA, e l'estensione temporale non è altresì applicabile a dati riferiti ad un'unica occasione di misurazione.

Si supponga che vengano considerate k piccole aree (o domini non pianificati), e che θ_i rappresenti uno degli indicatori di interesse per la popolazione nell'area i , $i=1, \dots, k$. Il modello FH combina le stime dirette θ_i , basate solo sui dati di indagine, e le stime da modello utilizzando un vettore di covariate, x_i per ogni area. Si ottiene il modello lineare ad effetti misti:

$$\hat{\theta}_i = x_i^t \beta + v_i + e_i$$

dove β rappresenta il vettore di parametri associati alle covariate, v_i sono effetti random definiti a livello di area per tenere conto di specifiche caratteristiche di ogni area non osservate tramite i dati a disposizione, per cui si assume distribuzione normale con media 0 e varianza σ_u^2 . Il valore e_i rappresenta l'errore nella stima dell'indicatore relativo all'area i e anche per questo si assume essere distribuito normalmente con media 0 e varianza $\sigma_{e_i}^2$. Nel modello, dunque, le covariate utilizzate per "spiegare" l'indicatore di interesse devono essere disponibili a livello di area. Nel modello classico FH tali covariate non devono essere affette da errore (per esempio, errore campionario), ipotesi che è solitamente soddisfatta quando le covariate provengono da archivi amministrativi o censimenti.

Sotto questo modello, la stima della media della piccola area è data dall'Empirical Best Linear Unbiased Predictor (EBLUP), che possiamo scrivere come

$$\hat{\theta}_i^{FH-EBLUP} = \hat{\gamma}_i \hat{\theta}_i + (1 - \hat{\gamma}_i)(x_i^t \hat{\beta} + \hat{v}_i)$$

dove $\hat{\beta}$ è lo stimatore dei minimi quadrati ponderati di β , $\hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \sigma_{e_i}^2}$ è il fattore shrinkage che misura la varianza del modello rispetto alla varianza totale, e $\hat{\sigma}_u^2$ è uno stimatore consistente di σ_u^2 . Lo stimatore FH-EBLUP attribuisce più peso alla stima sintetica quando la varianza campionaria è elevata o quando $\hat{\sigma}_u^2$ è piccola, e più peso alla stima diretta quando $\sigma_{e_i}^2$ diminuisce o $\hat{\sigma}_u^2$ aumenta.

La stima dell'errore quadratico medio (MSE) dello stimatore FH-EBLUP è stata proposta per la prima volta da Prasad e Rao (1990). Per una discussione dettagliata sulla stima del modello di FH, dell'EBLUP e del suo MSE, si rimanda a Rao e Molina (2015).

Nei modelli di stima per i 53 indicatori le variabili ausiliarie sono state selezionate dal sistema informativo “A misura di Comune”, sviluppato da Istat per supportare analisi territoriali e processi decisionali a livello locale. Il sistema integra fonti ufficiali e sperimentali, tra cui i dati del progetto ARCH.I.M.E.DE. e open data provenienti da altri enti del Sistan, fornendo indicatori comunali organizzati secondo i domini del Benessere equo e sostenibile (Bes) e gli Obiettivi di sviluppo sostenibile (SDG). I dati, aggiornati periodicamente a partire dal 2014, coprono numerosi ambiti, tra cui istruzione, lavoro, benessere economico, infrastrutture, mobilità e ambiente. Al momento dell’applicazione delle stime per piccole aree agli indicatori selezionati per misurare la povertà educativa gli indicatori erano aggiornati all’anno 2021.

Questa ampia disponibilità di dati ha permesso di costruire una base informativa robusta per la selezione delle covariate, scegliendo fra 51 variabili ausiliarie in grado di catturare le caratteristiche socio-economiche e demografiche interne alle regioni italiane. La selezione del miglior modello per ciascun indicatore è stata effettuata attraverso l’indice Adjusted R^2 . Si sottolinea che, mentre le covariate, come detto, fanno riferimento all’anno 2021, i dati da indagine utilizzati per il calcolo delle stime dirette sono relativi prevalentemente all’anno 2024. Si ritiene che la differenza di tre anni dei due riferimenti temporali non incida sulla bontà di predizione dei modelli, trattandosi di indicatori – per esempio riferiti alla struttura della popolazione – che non mostrano variazioni nel breve periodo. Dopo aver identificato il miglior set di covariate, il modello FH è stato applicato utilizzando la funzione *fh* del package *emdi* del software R.

L’applicazione del modello FH, combinando informazioni ausiliarie da stime dirette e variabili ausiliarie, ha consentito un miglioramento significativo della precisione delle stime. In particolare, per tutti i 53 indicatori le stime da modello mostrano un errore di stima inferiore rispetto a quello delle stime dirette, con una conseguente riduzione dell’intervallo di stima.

Per assicurare la coerenza territoriale, un ultimo passo ha riguardato il *bechmarking* delle stime DEGURBA interne ad ogni regione rispetto alla corrispondente stima regionale. In particolare, con tale operazione la somma pesata delle tre stime DEGURBA interne ad ogni regione è stata vincolata alla stima diretta regionale, più precisa e affidabile. I pesi utilizzati sono stati selezionati, per ciascun indicatore, in modo da rappresentare la proporzione della popolazione regionale di riferimento nelle aree DEGURBA.

Per le stime finali FH ottenute dopo l’operazione di *bechmarking* è risultata in generale una correlazione sempre molto elevata con le corrispondenti stime dirette, indice questo di un buon adattamento del modello di stima per piccole aree.

Dei 53 indicatori di fonte campionaria, ne sono stati selezionati 44. I motivi di esclusione degli indicatori sono i seguenti:

1) tra indicatori “simili”, calcolati da due fonti campionarie diverse, e che sono risultati fortemente correlati è stato selezionato l’indicatore affetto da un errore campionario minore (ad esempio all’indicatore “Bambini e ragazzi di 0-19 anni che vivono in una zona dove non ci sono parchi, giardini o altro verde pubblico” di fonte AVQ è stato preferito l’indicatore “Bambini e ragazzi di 11-19 anni che vivono in una zona senza spazi verdi, parchi e giardini” di fonte B&R) o è stato selezionato l’indicatore con la classe d’età più ampia (ad esempio all’indicatore “Ragazzi di 14-19 anni che non svolgono attività di partecipazione sociale” di fonte AVQ è stato preferito l’indicatore “Bambini e ragazzi di 11-19 anni che non fanno parte o non frequentano associazioni/organizzazioni” di fonte B&R)

2) quando l’indicatore presenta più del 25% di casi (celle) con un errore campionario, misurato tramite il coefficiente di variazione percentuale, superiore al 33%. Questo caso si è sempre verificato per i 5 indicatori dell’indagine CTL che quindi sono stati esclusi dal set finale degli indicatori di povertà educativa.

Gli indici compositi

Un modo per diffondere efficacemente le informazioni contenute in un insieme più o meno ampio di indicatori è sintetizzare gli stessi in un'unica misura, detta indice composito. Gli indici compositi possono raggiungere questo obiettivo dapprima trasformando gli indicatori nella stessa unità di misura (normalizzazione) e, successivamente, aggregando gli indicatori normalizzati che rappresentano le diverse dimensioni concettuali di un fenomeno multidimensionale. Un indice composito, quindi, si costruisce quando i singoli indicatori (detti indicatori elementari) vengono aggregati in un'unica misura basata su un modello concettuale (framework teorico). In tale contesto, per sintetizzare gli indicatori elementari in un'unica misura, si applica l'Indice Mazziotta-Pareto Aggiustato (AMPI). Si tratta di un indice composito parzialmente non compensativo, per il cui calcolo sono previste le seguenti procedure: gli indicatori elementari vengono standardizzati, rendendoli così indipendenti dall'unità di misura e trasformati in modo che varino approssimativamente all'interno dell'intervallo aperto (70; 130). Il valore base 100 è attribuito ad un'unità statistica di riferimento (in questo caso la media Italia). Le fasi di calcolo dell'AMPI sono descritte di seguito.

Data la matrice $X = \{x_{ij}\}$ con n righe (unità statistiche) e m colonne (i singoli indicatori), si calcola la matrice normalizzata $R = \{r_{ij}\}$ come segue:

$$r_{ij} = \frac{(x_{ij} - \text{Min}_{x_j})}{(\text{Max}_{x_j} - \text{Min}_{x_j})} 60 + 70$$

dove x_{ij} è il valore dell'indicatore j per la i -esima unità, e Min_{x_j} e Max_{x_j} sono valori calcolati (i cosiddetti *goalposts*) in modo da porre uguale a 100 il valore medio nazionale. L'indice composito si ottiene aggregando gli indicatori normalizzati con peso uguale mediante media aritmetica semplice, una funzione di sintesi additiva che, in quanto tale, presuppone un effetto compensativo fra gli indicatori elementari. Nell'ipotesi in cui gli indicatori elementari siano non sostituibili, ossia che il peso teorico sia equamente distribuito tra gli stessi, l'effetto compensativo della media aritmetica (intensità del fenomeno o effetto medio) è corretto aggiungendo alla media stessa un fattore (coefficiente di penalità) che dipende dalla variabilità dei valori normalizzati di ciascuna unità (variabilità orizzontale), ossia dalla variabilità degli indicatori rispetto ai valori di riferimento utilizzati per la normalizzazione. L'indice composito della i -esima unità (nel nostro caso, l'area geografica sub regionale individuata con la classificazione DEGURBA) si ottiene sommando alla media aritmetica degli indicatori una funzione della variabilità orizzontale che penalizzi gli squilibri (AMPI+):

$$\text{AMPI}_i^{+/-} = M_{r_i} \pm S_{r_i} cv_i$$

Dove M_{r_i} e S_{r_i} sono rispettivamente, la media e la deviazione standard dei valori normalizzati della i -esima unità; $cv_i = S_{r_i}/M_{r_i}$ il coefficiente di variazione della i -esima unità e il segno \pm dipende dal tipo di fenomeno misurato. Nel caso in cui sia negativo, come la povertà, allora la penalità sarà aggiunta alla media poiché il primo è il più povero.

Il metodo AMPI è stato applicato a 78 indicatori e 62 unità di analisi. Sono stati calcolati 7 indici compositi: uno per il dominio esiti, uno per il dominio risorse e cinque indici intermedi per le dimensioni contesti familiare, scolastico, territoriale, sociale e culturale, competenze cognitive e personali e sociali.

Il metodo AMPI richiede che nella matrice dei dati non ci siano valori mancanti, ciò ha comportato delle scelte nei casi delle indagini campionarie in cui il dato della cella regione per DEGURBA, calcolato con la stima per piccole aree aveva un errore campionario, misurato tramite il coefficiente di variazione percentuale, superiore al 33%³. In questi casi è stato inserito il valore medio regionale come proxy.

In altri casi, in cui i dati regionali non erano disponibili per vari motivi, è stato inserito il valore medio della ripartizione geografica disaggregata per DEGURBA come proxy.

I 7 indici variano approssimativamente all'interno dell'intervallo 70-130 e il valore base 100 è la media Italia.

Il 100 può essere interpretato come soglia relativa. Le unità territoriali con valori superiori a 100 mostrano una situazione relativamente peggiore della media nazionale, le unità con valori inferiori a 100 al contrario presentano una situazione relativamente migliore.

Nell'analisi dei risultati degli indici compositi, si distingue la povertà educativa che si manifesta in forma di esiti individuali dal rischio di povertà educativa a cui si è esposti per mancanza di risorse e opportunità, siano esse individuali, sociali o territoriali.

³ Statistics Canada, Guide to the Labour Force Survey 2010, Catalogue no. 71-543-G