

UNIVERSITÀ DEGLI STUDI DI PADOVA



FACOLTÀ DI SCIENZE STATISTICHE

**CORSO DI LAUREA SPECIALISTICA IN
SCIENZE STATISTICHE DEMOGRAFICHE E SOCIALI**

TESI DI LAUREA

ARBITRO, C'È RIGORE?

**Scelte metodologiche per la costruzione
di una graduatoria delle istituzioni universitarie:
approccio del Censis, altri approcci possibili**

RELATORE Ch.mo Prof. Lorenzo Bernardi

LAUREANDA Paola Bolzonello

Anno Accademico 2007 – 2008

*A tutti coloro che
hanno creduto, credono e crederanno in me.*

INDICE

INTRODUZIONE	1
CAPITOLO 1	5
1.1. Il Censis	5
1.2. Le attività del Censis	7
1.3. La Grande Guida all'Università	9
1.3.1 Una graduatoria per Facoltà	10
1.3.2 Chiavi di lettura	12
1.3.3 Le fonti dei dati ed il tempo	14
1.3.4 Le "famiglie"	16
1.3.4.1 Produttività e ricerca	19
1.3.4.2 Produttività e mondo del lavoro	21
1.3.4.3 Didattica e strutture	21
1.3.5 Non solo Facoltà	22
1.3.5.1 Indicatori specifici per Medicina	22
1.3.5.2 Le Facoltà di Atenei non statali	23
1.3.5.3 Gli Atenei	23
1.3.6 Note esplicative	24

1.3.7 Gli indicatori utilizzati	25
1.3.7.1 Produttività	26
1.3.7.2 Didattica	28
1.3.7.3 Ricerca	31
1.3.7.4 Profilo docenti	32
1.3.7.5 Rapporti internazionali	34
1.3.7.6 Attrazione	35
1.3.8 Indicatori e famiglie a confronto: considerazioni finali	36
CAPITOLO 2	41
2.1 Valutare	41
2.2 Misurare	43
2.3 Gli indicatori	46
2.3.1 Le dimensioni di un concetto	47
2.3.2 Proprietà di un indicatore	48
2.3.3 Gli indicatori sociali	49
2.3.3.1 L'informazione statistica per la valutazione dei fenomeni sociali ..	50
2.3.3.2 La nascita degli indicatori sociali	54
2.3.3.3 L'utilizzo degli indicatori sociali	56
2.3.3.4 Alcune considerazioni conclusive	57
2.3.4 La sintesi degli indicatori	58
2.3.4.1 Indicatori semplici e indicatori composti	58
2.3.4.2 Generalità sulla sintesi degli indicatori	59
2.4 Gli indicatori composti	60
2.4.1 Il processo di costruzione di un indicatore composto	60
2.4.1.1 Uno schema sintetico	60

2.4.1.2 Uno schema analitico	61
2.4.2 Considerazioni sugli indicatori composti	69
2.4.2.1 Introduzione	69
2.4.2.2 Pro e contro	71
2.4.3 La qualità degli indicatori composti	72
2.4.3.1 Il profilo di qualità degli indicatori composti	72
2.4.3.2 Le dimensioni della qualità dei dati di base	75
2.4.3.3 Qualità nella costruzione degli indicatori composti	76
2.5 La costruzione degli indicatori composti	78
2.5.1 I "passi chiave"	78
2.5.2 La normalizzazione	79
2.5.2.1 Le trasformazioni lineari	81
2.5.2.2 Le trasformazioni non lineari	84
2.5.2.3 Altre tecniche di trasformazione	87
2.5.2.4 Considerazioni conclusive	87
2.5.3 L'aggregazione	88
2.5.3.1 Considerazioni preliminari	89
2.5.3.2 L'approccio ordinale	89
2.5.3.3 Approcci cardinali additivi	91
2.5.3.4 Approccio multicriterio non compensativo	92
2.5.3.5 Aggregazione geometrica	94
2.5.3.6 Tecniche di aggregazione multivariate	94
2.5.3.7 Altre tecniche di aggregazione	101
2.5.3.8 Metodi di aggregazione a confronto	105
2.5.4 La ponderazione	107

2.5.4.1 Pesi uguali	108
2.5.4.2 Modelli di regressione multipla	109
2.5.4.3 Metodi basati su modelli statistici	110
2.5.4.4 Metodi partecipativi	112
2.5.4.5 Processo gerarchico analitico	114
2.5.4.6 Frontiera di efficienza	116
2.5.4.7 Componenti non osservate	120
2.5.4.8 Altre tecniche di ponderazione	122
2.5.4.9 Considerazioni conclusive	125
2.5.5 Riflessioni finali	128
2.5.5.1 Analisi della correlazione tra indicatori	129
2.5.5.2 Aggregazione e ponderazione: un matrimonio obbligato	130
2.5.5.3 Il significato dei pesi	131
2.5.5.4 Confronto dei risultati	132
2.5.5.5 Con "senso di misura"	134
CAPITOLO 3	139
3.1 I dati del Censis	139
3.1.1 Gli indicatori utilizzati nel 2006	139
3.1.1.1 Produttività	140
3.1.1.2 Didattica	141
3.1.1.3 Ricerca	143
3.1.1.4 Profilo docenti	144
3.1.1.5 Rapporti internazionali	145
3.1.2 Normalizzazione e aggregazione	147
3.2 Il dataset utilizzato per le elaborazioni.....	148

3.3 Analisi preliminare	149
3.3.1 Analisi della correlazione interna alle famiglie	150
3.3.2 Analisi della correlazione tra tutti gli indicatori	154
3.3.3 Analisi della correlazione tra famiglie	158
3.3.4 Considerazioni sui risultati dell'analisi di correlazione	160
3.4 Proposte alternative di analisi dei dati	161
3.4.1 I metodi utilizzati	161
3.4.1.1 Calcolo del punteggio delle famiglie	165
3.4.1.2 Sintesi degli approcci adottati	166
3.4.2 I risultati	169
3.4.2.1 Le graduatorie per famiglia	169
3.4.2.2 Le graduatorie finali	174
3.4.2.3 I confronti tra graduatorie	180
3.4.2.4 Considerazioni conclusive sui risultati	181
CONCLUSIONI	183
BIBLIOGRAFIA	189
APPENDICE	196

INTRODUZIONE

Può risultare imbarazzante vedersi giudicati e collocati in una posizione, non premiante, di una classifica compilata a seguito di una competizione a cui non si sapeva di partecipare. A maggior ragione se chi la costruisce, la propone, la diffonde, non è un organismo ufficiale, deputato e legittimato dai suoi doveri istituzionali, ma un puntuto testimone esterno che, per effetto di una comunque raggiunta credibilità o visibilità, si assume una responsabilità "civile" di informare la collettività sul funzionamento di una istituzione ritenuta decisiva per la crescita di una Nazione.

Deve essere stato questo il coagulo di sentimenti provato dagli Organi dirigenti e dai singoli operatori di ciascuna unità del sistema universitario nazionale all'apparire del primo lavoro del Censis intitolato "La Grande Guida all'Università" sulle graduatorie degli Atenei e delle varie Facoltà d'Italia costruite secondo "personali" criteri di qualità. Perché tale operazione è stata effettuata da un ente di natura privatistica, che in qualche misura anticipava doveri pubblici, su commessa di un organo di stampa che, ovviamente, poteva essere più interessato al clamore che quel prodotto avrebbe suscitato piuttosto che ad avviare, stimolare, imporre, comportamenti virtuosi? Quanto convincenti e affidabili erano disegno generale, criteri definitivi, natura delle fonti a cui si era ricorso, sistema di indicatori e regole di loro congiunzione, adottati? Quale livello di persuasione dell'opinione pubblica sarebbe stato ottenuto? E di conseguenza, quali le possibili ricadute sugli Atenei e sulle Facoltà per quanto riguardava le scelte delle nuove generazioni di candidati all'accesso agli studi superiori? Attaccare o difendersi, abbozzare o rinnegare, confrontarsi o costruire un proprio impianto di osservazione e valutazione? Riteniamo che le reazioni

siano state molte e tra loro talora divaricate, ma non vi è dubbio che il rapporto del Censis e le conseguenti immediate risposte degli Atenei abbiano lasciato molti sedimenti.

In primo luogo è aumentata l'attenzione dell'opinione pubblica che grazie al rapporto ed al potenziato impegno degli Atenei ha potuto godere di più ampia, più incisiva, più trasparente informazione, per scegliere piuttosto che per giudicare.

In secondo luogo ha sollecitato Ministero competente e suoi organismi interni a incentivare ulteriormente un proprio – forse più documentato, certamente più finalizzato – approccio valutativo, potenziando anche la valenza decisionale.

Quindi, forse *obtorto collo* o in modo quasi subliminale, ha spinto soggetti e enti accademici a scovare, talora a riconoscere, i propri punti deboli e a promuovere piccoli o grandi interventi per presentarsi all'esterno con una fisionomia più gradevole, con proposte più accattivanti, con un impegno in precedenza in parte trascurato nel cercare stili più idonei e organici per porsi in relazione non accidentale con la propria utenza potenziale o, semplicemente, con il sistema Paese.

Infine, aspetto per nulla marginale, ha favorito processi di collaborazione tra il Censis, che deve fare obbligato ricorso ai giacimenti informativi costruiti e disponibili presso le istituzioni pubbliche per rendere meno fragile il proprio apparato di analisi, e Referenti del mondo universitario – Ministero, sue strutture interne ed esterne di riferimento, ma anche la Conferenza dei Rettori e i Collegi dei Presidi – che hanno iniziato ad attribuire maggiore significato a quell'impresa, per la quale era opportuno che poggiasse su corretti, affidabili, robusti pilastri di realizzazione.

Non a caso si sono moltiplicati momenti di discussione e verifica di quel prodotto¹ spesso avendo come sfondo l'esame delle esperienze di altri Paesi in cui tali procedimenti godono di storia più lunga e di più ufficiale istituzionalizzazione, non raramente con obiettivi di individuazione di criteri e modalità per il delicato compito della redistribuzione di risorse.

Da questo quadro generale nasce questa tesi: il Rapporto annuale del Censis c'è, ha smosso le acque, si è affermato, potremmo sostenere che ormai è atteso, con speranze o ansie. È parso opportuno tentare di entrare un po' più nel merito delle semplici reazioni di impatto, rivisitandone contenuti e metodi, vagliando alcune possibili alternative,

¹ A titolo di esempio citiamo i seguenti lavori:

- Checchi D., Biolcati Rinaldi F., Guglielmetti C., Turri M. (2007), *Classifiche delle università (University ranking) – Report intermedio* presentato presso la Facoltà di Scienze Politiche dell'Università degli Studi di Milano il 30 ottobre 2007.
- Bittante G. e Masarotto G. (2004), *Interventi in occasione dell'incontro su "Indicatori di Ateneo e qualità dell'offerta formativa"* tenutosi presso l'Università degli Studi di Padova il 22 aprile 2004.

approfondendo la conoscenza delle oggettive difficoltà esistenti nel voler rispondere in modo efficace all'obiettivo che lo promuove e lo pervade. Non vi è alcun intento di rifiuto o di condanna, ma solo un impegno a contribuire, qualora le considerazioni che emergeranno risultassero convincenti, ad affinare il modello attualmente adottato anche passando per la sperimentazione, discussione e verifica, di parziali soluzioni diverse nei vari *step* che lo caratterizzano.

Siamo grati al Censis che, con spirito di corretta, aperta e responsabile collaborazione, ci ha fornito un *set* di dati elementari con cui effettuare le simulazioni di approcci alternativi; contiamo che il nostro lavoro, che mantiene ampi margini di problematicità in quanto, anche contro le nostre intenzioni, sottolinea i larghi spazi che sussistono nell'autonomia delle scelte, implicitamente ricordi non solo la laboriosità sistematicamente insita nei processi valutativi ma soprattutto la discrezionalità che naturalmente li caratterizza e che può essere compensata o contenuta solamente da chiari indirizzi di politica culturale e da giustificati e coerenti metodi per misurarli e sostanziarli.

La tesi si articola quindi secondo tre capitoli. Nel primo di essi abbiamo voluto illustrare chi è il Censis e cosa fa, dedicando gran parte del capitolo ad un'analisi analitica della struttura de "La Grande Guida all'Università" allo scopo di far comprendere anche al lettore più profano i cambiamenti che nel tempo hanno interessato lo sviluppo del rapporto nonché di dare fondamento alle considerazioni di merito espresse a fine capitolo stesso. Abbiamo poi voluto introdurre in maniera approfondita il passaggio e le relazioni esistenti tra "il valutare" ed "il misurare" individuando negli indicatori quello strumento che permette la realizzazione della traduzione di volontà conoscitiva in misurazione oggettiva, confrontabile e passibile – nonché adeguata – di analisi e critiche. Nel secondo capitolo poi abbiamo dedicato ampio spazio a quelli che abbiamo voluto definire "passi chiave" del processo di costruzione degli indicatori composti, illustrando diverse tecniche di normalizzazione, di aggregazione e di ponderazione di indicatori elementari tese alla misurazione di un concetto o fenomeno sociale. Il terzo capitolo raccoglie la sezione più strettamente operativa della tesi: in esso, dopo un'analisi preliminare dei dati, abbiamo proposto alcune elaborazioni alternative a quelle utilizzate dal Censis, giungendo a definire 12 diverse graduatorie. Tali graduatorie sono state sintetizzate mediante opportuni indici di sintesi e sono state poi confrontate attraverso l'analisi della cograduazione, evidenziando alcuni possibili approcci alternativi applicabili dal Censis, con risultati ancor più convincenti, all'intero insieme di dati. L'elaborato è concluso da alcune considerazioni

di merito sulla struttura degli indicatori elementari e delle famiglie su cui il Censis basa le proprie applicazioni. Inoltre, ad un'analisi complessiva dei risultati, segue la presentazione di alcune possibili ed interessanti applicazioni sull'intero insieme di dati di cui lavora il Censis, le quali non hanno potuto trovare concretizzazione per la ridotta numerosità dei dati avuti a nostra disposizione.

1.1 IL CENSIS²

La crescita economica che interessava l'Italia negli anni Cinquanta sollecitava una più marcata attenzione ai risvolti ed alle implicazioni sociali e della convivenza civile che quella generava. In altre parole si imponeva l'urgenza di studi ed analisi sociali condotte con nuove ipotesi, con nuove strategie, con orizzonte d'indagine più ampio capaci di penetrare problemi e temi emergenti in maniera più affinata.

In questo contesto nasce, nel 1964, il Censis, un istituto di ricerca socioeconomica il cui acronimo significa Centro Studi Investimenti Sociali. Dopo circa una decina d'anni, ottiene il riconoscimento a Fondazione (DPR num. 712/1993) grazie anche alla partecipazione di grandi organismi pubblici e privati.

L'istituto svolge attività di studio e consulenza; si occupa inoltre di valutazione nei settori della formazione, del lavoro, del welfare, delle reti territoriali, dell'ambiente, dell'economia, dello sviluppo locale e urbano, del governo pubblico, della comunicazione e della cultura, ambiti essenziali nella realtà sociale.

A livello nazionale opera per amministrazioni pubbliche quali Ministeri ed amministrazioni regionali, provinciali e comunali ma anche per Camere di Commercio, associazioni imprenditoriali e professionali, istituti di credito, aziende private, gestori di reti; lavora inoltre per organismi internazionali nonché nell'ambito dei programmi dell'Unione Europea.

Gli obiettivi del Censis sono riassumibili nei due punti seguenti:

² Il presente paragrafo è liberamente tratto dalla home page del sito del Censis (www.censis.it).

- cogliere e approfondire i temi fondamentali della società civile quali evoluzione della struttura sociale e produttiva, spinte all'innovazione, comportamenti individuali e collettivi, soggetti organizzati e corpi intermedi, protagonismo dei sistemi locali e meccanismi di iniziativa imprenditoriale;
- porre costante attenzione al governo dei processi, alla progettazione delle politiche, al coinvolgimento dei soggetti, all'azione dei poteri istituzionali e alla cultura di governo.

Dalla sua fondazione, il Censis ha esercitato la propria presenza e la propria capacità di analisi su campi e tematiche, poi divenuti emblematici: ciò significa che i vari impegni del Censis hanno accresciuto la loro varietà con il passare del tempo per venire incontro alle esigenze di un sistema sociale di complessità crescente. In particolare negli anni 60 si è posta forte attenzione al welfare ed alle strategie che ne governano l'impostazione e la gestione, con particolare analisi delle dinamiche della domanda e della natura dell'offerta. Nel decennio successivo l'interesse si è spostato sull'economia "sommersa" e sul suo ruolo all'interno del generale sviluppo socioeconomico dell'Italia. La modernizzazione dell'economia e i suoi intrecci con lo sviluppo complessivo della società italiana hanno caratterizzato il lavoro dell'istituto negli anni 80. Durante gli anni 90, è passato ad analizzare, in maniera più approfondita, la struttura sociale ed economica del Paese cercando di legare la storia alle trasformazioni che avvenivano in quel periodo. Dal 2000 l'istituto ha avviato una riflessione approfondita sul tema del "modello" socio-economico italiano con l'obiettivo di mettere in luce gli elementi più innovativi e gli ambiti emergenti in termini di vitalità.

L'attività del Censis ha ottenuto, nel corso degli anni, un continuo riconoscimento di ruolo ed autorevolezza che trova traduzione nel crescente apprezzamento da parte degli osservatori economici e sociali, dei soggetti imprenditoriali ed istituzionali e degli organi d'informazione, sia italiani che stranieri.

È possibile individuare sei ambiti di ricerca:

- **processi formativi**, scolastici e universitari, e loro valutazione;
- **mercato del lavoro**, lavoro sommerso, professioni e rappresentanza;
- **sanità e previdenza**, welfare e politiche sociali;
- **sviluppo locale e processi economici** con attenzione anche ai trasporti e alla logistica;
- **territorio, mercato immobiliare ed istituzioni locali**;

- **comunicazione**, consumo di **media** e **politiche culturali**, con riguardo alle diete mediatiche degli italiani.

Possono essere poi individuati tre settori con competenze di tipo trasversale:

- la **comunicazione**, i **rapporti con i media** e l'**editoria**;
- l'**elaborazione dati**;
- il settore dei **progetti speciali e internazionali**, che ha avuto particolare sviluppo negli ultimi anni, in stretto rapporto con la Commissione Europea.

1.2 LE ATTIVITÀ DEL CENSIS³

Le attività dei diversi settori elencati nel precedente paragrafo, oltre alla produzione di ricerche su temi specifici, sono realizzati in maniera integrata, al fine di rispondere ad esigenze di committenti diversi che richiedono competenze ed approcci metodologici distinti.

Nello specifico è interessante elencare alcune attività tra le più recenti realizzate dal Censis per poter comprendere concretamente quali lavori produce tale istituto.

A livello d'istituto viene fornita assistenza tecnica e vengono studiate azioni di sistema e di supporto operativo ai *decision maker* nella preparazione di atti programmatici, normativi e di pianificazione dello sviluppo (ad esempio le attività svolte per il Ministero del Lavoro e delle Politiche Sociali nell'ambito delle politiche per la trasparenza delle nuove forme di lavoro e per l'emersione del lavoro non regolare, l'Osservatorio permanente sulla condizione degli immigrati e sullo stato dei processi di accoglienza e di integrazione nelle regioni del Mezzogiorno, l'assistenza al Dipartimento della Funzione Pubblica - Presidenza del Consiglio dei Ministri per la realizzazione di interventi formativi di supporto all'innovazione organizzativa e allo sviluppo dell'*e-government* negli Enti Locali). A livello locale, in particolare su commissione degli enti Regione, svolge attività di valutazione delle politiche pubbliche e di supporto alla programmazione (come la realizzazione della ricerca sui fabbisogni formativi del territorio della Regione Sardegna, la realizzazione di un modello di formazione continua per gli operatori della Regione Puglia, la valutazione

³ Il presente paragrafo è liberamente tratto dal sito web del Censis (www.censis.it).

indipendente degli interventi del Fondo Sociale Europeo realizzati dalla Regione Friuli Venezia Giulia, l'assistenza tecnica alla Regione Campania nell'analisi dei fabbisogni di infrastrutture materiali e immateriali nei distretti e nei sistemi locali del lavoro). Per enti nazionali (quali in Ministero degli Affari Esteri) ed internazionali (quali la World Bank), realizza progetti di cooperazione allo sviluppo a favore di Paesi europei ed extraeuropei (come il progetto "Analisi e strategie per lo sviluppo nella transizione della Slovacchia Centrale e per lo scambio con il sistema economico italiano" o il progetto "Sviluppo del sistema di istruzione superiore tecnica in Libano"). Infine realizza e/o partecipa a progetti di ricerca, consulenza e supporto alle politiche e agli interventi della Commissione Europea e di altri organismi comunitari su temi di giustizia, relazioni estere, immigrazione ed integrazione, imprese sociali e valorizzazione del patrimonio locale.

Il Censis svolge oltre 60 ricerche ogni anno per molteplici committenti; il loro originale modello è frutto di un equilibrio tra un tradizionale stile accademico ed un'impostazione operativa degli obiettivi di ricerca. Questo ha implicato due importanti conseguenze: da un lato il fatto che l'istituto ha accumulato un ricco patrimonio conoscitivo dei processi evolutivi della società italiana, dall'altro l'aver a disposizione un apparato metodologico sofisticato che permette di affrontare i singoli obiettivi di ricerca inserendoli in un unico schema interpretativo di carattere generale.

Oltre alle attività di ricerca commissionate da vari enti, il Censis, in quanto fondazione, promuove autonomamente una serie di iniziative per approfondire le logiche interpretative da utilizzare nella ricerca e per esplorare nuovi percorsi conoscitivi della realtà socioeconomica dell'Italia. Alcune attività fondazionali hanno periodicità fissa mentre altre rispondono alle esigenze congiunturali dello sviluppo socioeconomico del Paese. In particolare dal 1967 viene elaborato il *Rapporto annuale sulla situazione sociale del Paese*, nato dall'esigenza di rendere disponibile uno strumento di analisi e d'interpretazione dei fenomeni, dei processi, delle tensioni e dei bisogni emergenti. La struttura del lavoro è costituita dalle seguenti sezioni:

- le considerazioni generali forniscono alcune chiavi di lettura dell'evoluzione della realtà italiana attraverso un modello interpretativo che lega i risultati annuali al mutamento sociale degli anni precedenti;

- la parte dedicata alle fenomenologie dell'anno analizza l'analisi dei fenomeni di maggior rilievo che hanno caratterizzato significativamente l'evoluzione del sistema sociale;
- la presentazione – nelle sezioni denominate “settori e soggetti del sociale” e “mezzi e processi” – di un bilancio settoriale annuo degli ambiti della formazione, del lavoro, del welfare, del territorio e delle reti, dei soggetti economici, del governo locale, della comunicazione e della cultura.

Il Rapporto rappresenta una fotografia sufficientemente completa ed equilibrata dello stato della società italiana, analizza le relazioni tra i diversi aspetti e problemi della realtà sociale, illustra e, spesso, anticipa linee di tendenza e percepibili cambiamenti⁴.

1.3 LA GRANDE GUIDA ALL'UNIVERSITÀ

Dall'anno 2000, per il quotidiano *La Repubblica*, il Censis Servizi realizza “La Grande Guida all'Università” che fornisce un'analisi del sistema universitario suddiviso per facoltà. Tale lavoro si configura come uno studio comparativo che ha l'obiettivo di essere strumento di servizio e di orientamento per gli studenti che si accingono a scegliere quale percorso di studi universitari compiere al termine della scuola secondaria superiore. Vuole fornire, inoltre, un aiuto alle famiglie perché possano supportare i propri figli. La guida, negli ultimi anni, è stata pubblicata in quattro puntate settimanali successive nel mese di luglio, periodo nel quale i neo-diplomati si apprestano alla scelta degli studi futuri.

Quella che il Censis definisce “filosofia di analisi” si riassume propriamente nell'obiettivo sopra indicato di tale studio, obiettivo che indirizza di fatto le scelte sugli indicatori, le metodologie utilizzate e gli strumenti e le piste di lavoro collaterali⁵. La letteratura conferma tale impostazione identificando negli obiettivi di studio il punto di partenza dal quale scegliere le metodologie operative di lavoro in particolare quando si tratta di indicatori composti. Come sarà approfondito nel capitolo 2, per poter diffondere i risultati di elaborazioni che coinvolgono la costruzione di indicatori composti è indispensabile

⁴ Zajczyk F. (1997), *Il mondo degli indicatori sociali*, NIS, Roma, pag. 39.

⁵ Nota metodologica anno 2000.

un'aprioristica trasparenza degli obiettivi, al fine di evitare erronee manipolazioni dei dati o improprie interpretazioni.

Per comprendere a fondo la struttura de "La Grande Guida all'Università" sono quindi state analizzate le note metodologiche che corredano ogni sua edizione. Molte modifiche sono intervenute nella metodologia impiegata, in particolare per quanto riguarda gli indicatori utilizzati i quali differiscono di anno in anno. Più avanti nel capitolo verranno illustrate tali variazioni che, come vedremo, dipendono fortemente dai mutamenti del sistema universitario intercorsi in questi anni nonché dalla disponibilità – o meno – di dati. Vogliamo ora focalizzare l'attenzione proprio sullo spirito e sulle linee guida che hanno guidato la realizzazione di quest'analisi comparativa.

1.3.1 Una graduatoria per Facoltà

Nella prima nota metodologica è illustrata l'articolazione complessiva delle tecniche di analisi che di seguito riportiamo:

- sondaggio sugli studenti del V anno delle scuole superiori (2.000 interviste);
- sondaggio sugli studenti che frequentano l'Università (2.000 interviste);
- sondaggio sulle famiglie italiane con figli iscritti al V anno di scuola inferiore⁶ (1.200 famiglie);
- sondaggio sulle famiglie con figli iscritti all'Università (1.200 famiglie);
- 6 focus group con studenti universitari;
- sondaggio presidenti/segretari di 11 ordini professionali a scala provinciali (circa 500 questionari elaborati al 20.06.2000);
- sondaggio sui segretari delle Unione Industriali Provinciali (55 risposte al 20.06.2000);
- sondaggio sui segretari delle Camere di Commercio (65 risposte al 20.06.2000);
- sondaggio sui Presidi delle facoltà istituzionali (215 risposte al 20.06.2000);
- sondaggio sulle società specializzate in selezione del personale (40 risposte al 20.06.2000);
- ranking per facoltà con indicatori.

⁶ Presumibilmente scuola media inferiore.

Gli esiti di questa molteplice attività di formazione di dati di base e della successiva loro elaborazione sono tradotti e presentati solo con riferimento all'ultimo punto ovvero al ranking (graduatoria) delle facoltà italiane suddivise nei seguenti raggruppamenti:

- Agraria
- Architettura
- Beni Culturali
- Economia
- Farmacia
- Giurisprudenza
- Ingegneria
- Lettere
- Lingue
- Medicina
- Psicologia
- Scienze
- Scienze della Formazione
- Scienze Politiche
- Scienze Statistiche
- Sociologia
- Veterinaria.

La scelta di una graduatoria stilata avendo come unità di riferimento le facoltà ha in sé legittime motivazioni ma anche alcuni limiti.

Sostanzialmente il ranking è elaborato per facoltà perché, secondo il Censis, essa rappresenta l'unità di misura delle scelte di iscrizione per studenti e famiglia. Organizzare l'analisi secondo i Corsi di Laurea avrebbe portato ad un lavoro molto oneroso per il loro elevato numero; inoltre alcuni di essi sono accomunati da periodi di studio iniziale uguali (biennio comune nel vecchio ordinamento, un anno nell'ordinamento attuale) che sfociano nella seconda parte del percorso di studi in una differenziazione contenutistica. L'opzione alternativa di redigere graduatorie relative ai dipartimenti è innanzitutto ingiustificata per il fatto che si tratta di strutture istituzionalmente responsabili dell'organizzazione delle attività di ricerca (mentre spetta alle facoltà la realizzazione dell'offerta formativa) e in parte anche per essere non del tutto note, nella loro natura e nei loro compiti, agli

studenti.⁷ Tale scelta avrebbe inoltre creato difficoltà di definizione degli ambiti di riferimento essendo spesso i dipartimenti aggregazioni di docenti afferenti a più facoltà ed a più Corsi di Laurea.

Aver scelto la facoltà come unità di riferimento per la graduatoria ha però alcuni limiti, primo tra tutti la diversità nel numero e nella natura dei Corsi di Laurea appartenenti a facoltà della stessa specie. Inoltre nello scenario italiano ci sono facoltà che nascondono Corsi di Laurea equiparabili alla facoltà istituzionale (Psicologia, ad esempio, in certi Atenei o in certi periodi storici è o è stata uno dei Corsi di Laurea di altre facoltà); parimenti esistono facoltà prive dei Corsi di Laurea predominante (ad esempio a Napoli nella Università II Lettere ha Psicologia e Beni Culturali).

È inoltre utile aggiungere un'ulteriore considerazione di merito rispetto alla scelta della facoltà come unità di riferimento per la stesura delle graduatorie relativamente all'aspetto della ricerca. Indubbiamente la ricerca viene svolta in maniera preminente dai dipartimenti ma è anche vero che ricerca e facoltà sono concretamente, quasi "empiricamente", dimostrabili, come condiviso da molti presidi:

- la qualità della ricerca si trasferisce quasi automaticamente nella didattica contribuendo a renderla significativamente più articolata ed aggiornata;
- il reclutamento dei docenti, deciso dalle facoltà, avviene quasi sempre con riferimento alle attività di ricerca effettuate nella carriera e che entrano a pieno titolo nei curricula vitae dei partecipanti ai concorsi;
- le relazioni tra ricerca e didattica sono crescenti in relazione all'ampliamento dell'offerta di master, dottorati, etc.⁸

1.3.2 Chiavi di lettura

La graduatoria ha due chiavi di lettura:

- in verticale: è possibile un confronto per facoltà individuando nelle posizioni più alte quelle facoltà con maggiore equilibrio relativo ai diversi aspetti esaminati
- in orizzontale: si possono riconoscere i punti di forza e le debolezze di ciascuna facoltà; questo può favorire per studenti e famiglie – probabilmente in maniera più rilevante – nella scelta relazionata agli obiettivi di studio.

⁷ Va ulteriormente considerato che dipartimento e facoltà sono grandezze non commensurabili per la loro diversa natura e finalità.

⁸ Nota metodologica 2001.

Di fatto si tratta di una vera e propria **valutazione** che ha l'obiettivo di dare ai futuri immatricolati ed alle loro famiglie uno strumento che li aiuti a scegliere in modo consapevole sulla base delle proprie esigenze, delle proprie inclinazioni, delle proprie attese⁹. Vengono valutati diversi elementi del sistema università e la graduatoria attribuisce il rango a ciascuna facoltà secondo il miglior equilibrio che la stessa possiede rispetto ai vari aspetti considerati. In altre parole viene prodotto un giudizio di sintesi del maggior equilibrio raggiunto dalle facoltà. Questo implica quindi una sottintesa – e non esplicita – logica compensativa per cui punteggi elevati per un aspetto possono bilanciare valori negativi relativi ad un diverso elemento valutativo. Tale assunto impedisce, come vedremo in maggior dettaglio nel capitolo 2, l'utilizzo – almeno a livello teorico – di alcune tecniche operative di aggregazione degli indicatori.

Le linee guida che hanno determinato il *modus operandi* del Censis¹⁰ per la produzione di questo rapporto possono essere sintetizzate come segue:

- indossare i panni degli studenti e della domanda di formazione universitaria;
- elaborare indicatori e macrofamiglie di indicatori che diano al lettore la possibilità di orientarsi in relazione alle proprie personali scelte o attese (standard di offerta, spinta alla ricerca, grado di internazionalizzazione, etc.);
- utilizzare in modo combinato dati ed informazioni generati da istituzioni o agenzie con una logica che garantisca uniformità, ufficialità e certificazione dei dati;
- esplicitare con chiarezza e trasparenza gli elementi di criticità e le lacune del modello valutativo, criticità e lacune derivanti molto spesso dalla mancanza di dati, a volte dalla eterogeneità delle fonti informative.

Relativamente all'ultimo punto, pare utile sottolineare come le modifiche intervenute nel lavoro del Censis siano state, negli anni, frutto di produttivi e costruttivi scambi, di critiche e di suggerimenti offerti soprattutto da presidi e docenti universitari. Questo aspetto è stato reso noto agli attenti lettori nelle note metodologiche dei diversi anni nella cui parte introduttiva sono presenti considerazioni e delucidazioni relative ai mutamenti avvenuti. Vedremo più avanti in dettaglio il risultato di tali scambi, sia a livello operativo, sia a livello semantico: contributi nell'interpretazione del cambiamento sociale del sistema universitario ed approfondimenti nello studio delle componenti del sistema stesso hanno portato, nel tempo, a considerevoli modifiche nell'impianto strutturale dell'indagine, sostanzialmente in

⁹ Nota metodologica 2001.

¹⁰ Nota metodologica 2002.

termini di indicatori utilizzati. In altre parole le modifiche subite dalla struttura del sistema valutativo si traducono in una variazione degli indicatori utilizzati per misurare i diversi aspetti dell'università che tiene conto dei cambiamenti sociali ed istituzionali dei contributi interpretativi intervenuti nel tempo.

1.3.3 Le fonti dei dati ed il tempo

Nelle linee guida definite dal Censis – e sinteticamente ricordate nel paragrafo precedente – si fa riferimento alla natura dei dati utilizzati per lo studio comparativo, in particolare per quanto concerne la loro fonte. Dal momento che i dati sono ricavati da soggetti, enti diversi, diversi sono i metodi di raccolta delle informazioni, lo stato di disponibilità dei dati, il grado di aggiornamento (auspicabilmente massimo) e la natura stessa dei dati. È necessario perciò un oneroso lavoro di omogeneizzazione effettuato sia ex ante sia ex post la raccolta dei dati. Ciò significa che, in fase di raccolta delle informazioni necessarie alla costruzione degli indicatori e quindi delle graduatorie, si rendono necessarie operazioni tese a verificare ed incrociare dati provenienti da fonti diverse; in fase ex post, invece, una volta in possesso degli indicatori, è necessario applicare opportune procedure di standardizzazione che rendano tali dati comparabili.

Il Censis ha scelto di utilizzare solamente i dati che rispondono ai tre criteri seguenti¹¹:

- ufficialità: i dati devono essere forniti da un soggetto istituzionale;
- omogeneità: vengono utilizzati dati rilevati con identica metodologia;
- certificazione: i dati devono essere definitivi¹².

Indubbiamente negli anni quantità e qualità delle informazioni sono migliorate in qualsiasi campo del complesso sistema sociale. Ciò ha comportato, nello specifico, come prima detto, una variazione del sistema di indicatori utilizzati per la produzione della graduatoria delle facoltà: alcune lacune nel sistema valutativo sono pertanto attribuibili alla mancanza di dati che rispondano ai tre criteri illustrati nonché alla disomogeneità delle fonti informative.

Se da un lato, quindi, comprendiamo come l'innovazione ed il mutamento sociale abbiano contribuito ad un doveroso adattamento degli strumenti di analisi utilizzati dal

¹¹ I criteri sono presenti in tutte le note metodologiche da quella dell'anno 2000 a quella del 2006.

¹² Un dato è definitivo quando è diffuso a conclusione dell'evento che lo ha generato.

Censis, siamo anche consapevoli che tali modificazioni compromettono i confronti temporali delle graduatorie annuali. È pur vero che dal 2002 gli aspetti esaminati sono rimasti i medesimi ma la modifica degli indicatori semplici¹³ utilizzati per la costruzione dei punteggi dei singoli elementi potrebbe "non rendere giustizia" di quelli che sono i reali mutamenti avvenuti all'interno delle singole unità di analisi, le facoltà, ovvero non rende il lettore in grado di individuare i veri plus ed i veri minus¹⁴ della realtà osservata.

La critica agli improvvisi mutamenti nelle gerarchie – dal Censis prodotte – delle diverse facoltà mina il fondamento dell'assunto di sostanziale stabilità dell'università italiana: si tratta di un principio apparentemente teorico ma che, se si guarda alla sua storia decennale, può difficilmente essere smentito. Tale stabilità è stata considerata dal Censis come icona dell'università italiana del passato¹⁵. Si tratta di un'affermazione che non ha però trovato riscontro nelle analisi, anche elementari, del sistema universitario: non è infatti dubbio che i grandi mutamenti generati, in particolare da interventi normativi o da grandi trasformazioni sociali, pur in periodi molto lunghi, possano aver determinato forti mutamenti anche nel panorama del sistema universitario che pure rimane caratterizzato da elementi di forte viscosità (quali la permanenza per molti anni dei docenti nella medesima facoltà, la ripetizione degli stessi contenuti degli insegnamenti che si ripetono per lunghi periodi di tempo, etc.). In altre parole, quindi, l'università italiana non ha mai goduto di "stabilità statica" ma ha vissuto continui mutamenti ed adeguamenti alla realtà sociale in cui era inserita. Nei tempi più recenti, un forte dinamismo caratterizzato il sistema universitario del nostro Paese: basta pensare alle pesanti modificazioni intervenute con la riforma universitaria che ha portato alla crescita del numero di immatricolati e ad un "gonfiamento" dell'offerta didattica¹⁶.

Per ottenere quindi una graduatoria statisticamente definibile "affidabile"¹⁷, il Censis ha operato, in taluni casi, scelte di carattere tecnico mirate ad attutire eventuali effetti congiunturali e/o di disturbo mediando gli indicatori relativi a diversi anni¹⁸. Questo è avvenuto applicando alla gran parte degli indicatori il principio delle medie mobili, che

¹³ Il significato di "indicatore semplice" verrà illustrato in dettaglio nel capitolo 2 del presente elaborato.

¹⁴ Nota metodologica 2004.

¹⁵ Nota metodologica 2002.

¹⁶ La crescita dell'offerta didattica delle singole facoltà si è tradotta in un incremento del numero di Corsi di Laurea e degli insegnamenti impartiti. Ciò non ha però sempre trovato reale corrispondenza con un maggiore volume di conoscenze e competenze, anche in termini di profili professionali acquisiti dai laureati.

¹⁷ Si definisce affidabilità di uno strumento di misurazione la sua capacità di ottenere, nelle stesse condizioni, le stesse misure.

¹⁸ Cfr. Nota metodologica 2007.

consente di ammortizzare tendenze al cambiamento troppo repentine¹⁹. Si tratta però di una dichiarazione del Censis che suscita ragionevoli sospetti: riteniamo che l'applicazione del metodo delle medie mobili possa essere ritenuta vera solo in parte a causa dei grandi cambiamenti intervenuti nella natura stessa degli indicatori usati. Una giustificazione a tale affermazione può essere rapportata al fatto che, in particolare negli anni iniziali, il Censis può essersi trovato di fronte a situazioni eccessivamente mutevoli dovute anche alla diversa qualità della produzione dei dati per cui ha ritenuto necessario trovare un metodo per "ammorbidire" le grosse variazioni che le graduatorie hanno subito di anno in anno.

1.3.4 Le "famiglie"

La costruzione del punteggio finale su cui si basa la classifica delle facoltà è un processo costituito da diversi passi. Il punto di partenza sono naturalmente indicatori semplici tratti da fonti istituzionali che rispettano, come prima illustrato, i requisiti di ufficialità, omogeneità e certificazione. Tali indicatori sono raggruppati secondo accorpamenti denominati "famiglie": si tratta sostanzialmente dei fattori che il Censis ha individuato per la descrizione del sistema università. Non è chiaramente specificato come tali famiglie siano state individuate: nella nota metodologica del 2007 viene unicamente specificato che le famiglie di indicatori rappresentano altrettante funzioni che il Censis ritiene essenziali per la vita universitaria²⁰. Come siano state definite tali fattori è quindi ignoto al lettore: si è partiti dagli indicatori disponibili ricavandone le aggregazioni o, al contrario, dalle funzioni individuate a priori sono stati scelti gli indicatori semplici? Si tratta di un passaggio non banale nella definizione delle graduatorie finali: il punto di partenza dell'analisi ne determina i risultati. Possiamo ipotizzare diverse strade attraverso le quali il Censis è giunto a stabilire le diverse famiglie:

- empiricamente si può essere partiti da un set di indicatori – magari scelto in base a quanto riportato dalla letteratura di riferimento – individuando, attraverso opportune procedure di statistica multivariata, le categorie di raggruppamento;

¹⁹ Nota metodologica 2003.

²⁰ Cfr. Nota metodologica 2007.

- un'indagine pilota su testimoni privilegiati – quali potrebbero essere, ad esempio, i presidi delle facoltà – potrebbe aver aiutato a stabilire le funzioni valutabili e valutanti dell'università;
- stabilite ex ante dal Censis stesso le famiglie considerate aspetti cruciali del sistema universitario, gli indicatori sono stati selezionati, tra tutti i dati disponibili relativi all'università, secondo principi di pertinenza;
- la scelta, infine, potrebbe derivare da semplici criteri di disponibilità e fruibilità dei dati esistenti.

È ragionevole ipotizzare che la struttura finale sia il risultato di un mix di tali approcci.

Come vedremo in maniera approfondita nel secondo capitolo di questo lavoro, diverse tecniche statistiche sono applicabili per determinare le famiglie secondo cui aggregare gli indicatori elementari. Un'unica considerazione nella nota metodologica del 2007 ci dice che l'articolazione in famiglie di indicatori viene attentamente studiata anno per anno alla luce, come già detto, dei dati disponibili e delle trasformazioni in atto nel modello di offerta del sistema universitario²¹. Ci limitiamo qui a concludere che l'assenza di un'informazione di tale rilievo, pur potendo non assumere rilevanza per i destinatari ultimi – ovvero i futuri immatricolati – de "La Grande Guida all'Università", è valutabile come lacuna dell'impianto metodologico reso al lettore.

Presentiamo di seguito una tabella che sintetizza le famiglie utilizzate dal Censis dal 2000 al 2007.

Tabella 1. Famiglie di indicatori utilizzate per il ranking delle facoltà per anno.

2000	2001	2002	2003	2004	2005	2006	2007
Produttività	Produttività	Produttività	Produttività	Produttività	Produttività	Produttività	Produttività
Didattica	Didattica e standard d'offerta	Didattica e standard d'offerta	Didattica e standard d'offerta	Didattica e standard d'offerta	Didattica e standard d'offerta	Didattica	Didattica
Ricerca	Ricerca	Ricerca	Ricerca	Ricerca	Ricerca	Ricerca	Ricerca
		Profilo docenti	Profilo docenti	Profilo docenti	Profilo docenti	Profilo docenti	Profilo docenti
Internazionalizzazione	Rapporti e cooperazione internazionale	Rapporti e cooperazione internazionale	Rapporti e cooperazione internazionale	Rapporti e cooperazione internazionale	Rapporti e cooperazione internazionale	Rapporti internazionali	Rapporti internazionali
Capacità di attrazione	Attrazione ed influenza sul contesto locale e nazionale						

²¹ Nota metodologica 2007.

Abbiamo riportato nella tabella 1 le famiglie con il nome specifico con cui vengono utilizzate nelle note metodologiche. Possiamo però riassumerle nelle sei seguenti categorie:

- la **produttività** che misura le capacità delle facoltà di garantire un percorso regolare del ciclo di studi (limitare la dispersione, monitorare l'apprendimento degli studenti mediante gli esami e l'attribuzione dei crediti, laureare in corso);
- la valutazione della **didattica** come garanzia di un rapporto equilibrato tra docenti e studenti, offrire strutture adeguate ed un'adeguata offerta formativa, etc.;
- la **ricerca** per valutare la capacità dei docenti di progettare per la ricerca e stabilire quante possibilità abbia uno studente di confrontarsi con professori con tali caratteristiche;
- con il **profilo docenti** si mira a far emergere quelle facoltà che puntano sul ringiovanimento e sulle relazioni internazionali del proprio organico;
- i **rapporti internazionali** per stabilire il grado di apertura delle facoltà alle esperienze internazionali sia dei propri studenti che del corpo docente²²;
- l'**attrazione** delle singole facoltà a confronto con l'Ateneo o le facoltà omogenee.

Un ulteriore fattore valutato dal Censis ma che non è entrato nella costruzione del punteggio finale per la graduatoria delle facoltà è quello relativo al **prestigio**. Una misura di tale aspetto è stata calcolata per gli anni dal 2001 al 2005: il metodo di calcolo prevede l'attribuzione di un punteggio pari a 1, 2, 3 o 4 stelle a tre diversi indicatori di prestigio. Si tratta di un ulteriore criterio di valutazione del sistema universitario segnalato più volte nel mondo accademico da parte dei presidi: esso individua alcuni plus relativi al prestigio del corpo docente o delle facoltà²³. I tre aspetti presi in considerazione sono valutati da altrettanti indicatori che misurano quanto segue:

- la produzione "editoriale universitaria" dei docenti delle diverse facoltà (calcolata su tutte le pubblicazioni edite in Italia con codice ISBN);
- le performance "concorsuali" (misurate attraverso il tasso di idoneità nelle sessioni annue dei docenti delle diverse facoltà);
- il rango delle facoltà derivante dal giudizio espresso dai presidi ai quali è stato chiesto di indicare le prime tre facoltà per qualità complessiva offerta²⁴.

²² Cfr. Nota metodologica 2007.

²³ Nota metodologica 2001.

²⁴ Nota metodologica 2001.

Riportiamo di seguito una tabella che evidenzia come le principali famiglie siano presenti/assenti nel lavoro dei diversi anni.

Tabella 2. Presenza delle famiglie di indicatori presenti nei lavori annuali.

Famiglie	2000	2001	2002	2003	2004	2005	2006	2007
Produttività	x	x	x	x	x	x	x	x
Didattica	x	x	x	x	x	x	x	x
Ricerca	x	x	x	x	x	x	x	x
Profilo docenti			x	x	x	x	x	x
Rapporti internazionali	x	x	x	x	x	x	x	x
Attrazione	x	x						

Come chiaramente evidenzia la tabella sopra riportata, le famiglie della produttività, della didattica, della ricerca e dei rapporti internazionali sono presenti tutti gli anni. L'attrazione è stata valutata solamente nei primi due anni di realizzazione del rapporto mentre il profilo docenti è un fattore entrato stabilmente nella struttura valutativa dall'anno 2002.

Ogni anno il set di indicatori facente parte delle diverse famiglie ha subito molteplici modificazioni per la disponibilità nel patrimonio informativo di nuovi dati utili all'analisi in oggetto ma anche perché fosse preso in considerazione il mutamento sociale, culturale e normativo con cui il sistema universitario è intrecciato.

1.3.4.1 Produttività e ricerca

Un nodo ancora irrisolto resta la valutazione della produttività scientifica, importante fattore che conferisce indiscutibili plus alle facoltà più attive in questo senso ma che non ha trovato, nella ricerca del Censis di questi anni, un adeguato criterio di misurazione. Il calcolo dovrebbe tenere in considerazione il numero di pubblicazioni, il rango delle riviste dove siano apparsi i risultati, gli articoli e le citazioni raccolte. Un indicatore che includa i requisiti esposti è noto in letteratura come Science Citation Index (SCI)²⁵: esso raccoglie i dati necessari al confronto dell'*impact factor* (IF)²⁶. Le scuole di pensiero, sull'utilità e sulle potenzialità dell'analisi proposte attraverso l'IF sono spesso divergenti nel mondo accademico: vi è chi considera l'IF come l'unico valido strumento di comparazione e di

²⁵ L'indicatore è commercializzato dall'istituto americano Institute of Scientific Information.

²⁶ L'impact factor è un elemento molto importante nell'ambito delle pubblicazioni scientifiche tanto che spesso viene usato come indice d'importanza di una pubblicazione in uno specifico campo. L'impact factor di una rivista misura la frequenza con cui un articolo è stato citato mediamente in un anno, fino a due anni dalla prima pubblicazione e per un periodo triennale. Esso è utile nel valutare l'importanza relativa di una rivista, specialmente se confrontata con altre dello stesso settore. L'impact factor viene calcolato dividendo il numero corrente di citazioni di articoli pubblicati nei precedenti due anni per il numero totale di articoli pubblicati nei precedenti due anni.

valutazione della ricerca scientifica e chi lo rubrica come una scorciatoia commerciale di basso significato scientifico²⁷. Inoltre va considerato che l'indice SCI non considera tutta la produzione scientifica: viene infatti attribuito minor peso all'area umanistica, il cui lavoro non può essere misurato strettamente dal numero di articoli scientifici elaborati e pubblicati. Di fatto risulta impossibile raccogliere a livello centrale la totalità delle pubblicazioni scientifiche realizzate dai docenti universitari (nella nota del 2001 si fa riferimento a circa 55.000 pubblicazioni). Volendo però avere una valutazione quantitativa della ricerca in termini di produzione scientifica, il Censis ha avviato nel 2001 un tentativo di raccolta di informazioni attraverso i presidi delle facoltà. L'intento era quindi quello di raccogliere dati relativi alle pubblicazioni dei docenti interni alle facoltà attraverso un processo metodologico standard. Di fatto però, sommando una non completa condivisione dell'impianto metodologico e la ristrettezza dei tempi, non si hanno avute sufficienti risposte per far entrare nel calcolo della graduatoria una misura della produzione scientifica.

Cinque anni più tardi, nel 2006, alla luce di alcuni cambiamenti avvenuti nel sistema universitario, vi era la disponibilità dei primi risultati sulla produttività scientifica elaborati dal CIVR (Comitato di Indirizzo per la Valutazione della Ricerca). Tali informazioni non sono però state utilizzate dal Censis per tre ordini di motivazioni. Anzitutto il CIVR svolge analisi con obiettivi naturalmente diversi da quelli de "La Grande Guida all'Università". Esso prende in esame le aree scientifiche ovvero le attività dipartimentali mentre la graduatoria del Censis è organizzata per facoltà²⁸. In secondo luogo, poi, l'analisi del CIVR si traduce in graduatorie delle strutture secondo la dimensione mentre il Censis non prevede una distinzione sulla base di tale criterio ma semplicemente raggruppa facoltà omogenee. Un ultimo motivo di non utilizzo dei dati CIVR risiede nel periodo di riferimento da esso preso in esame (2001-2003) il quale potrebbe non tenere conto di eventuali trasferimenti di ricercatori, da un ateneo ad un altro, avvenuti successivamente a tale periodo²⁹.

Non è quindi mai stato inserito nella struttura degli indicatori del Censis, nonostante il buon auspicio espresso in termini di future collaborazioni nel 2001³⁰, un indicatore che,

²⁷ Cfr. Nota metodologica 2001.

²⁸ Si rimanda al paragrafo 3.1 dove sono spiegati vantaggi e limiti della scelta di una graduatoria stilata per l'unità "facoltà".

²⁹ Cfr. Nota metodologica 2006.

³⁰ "Il tentativo è stato avviato con coordinatori di alcune facoltà ma i tempi ristretti - e in qualche caso l'impossibilità di condividere l'impianto metodologico per il momento - hanno fatto sì che l'area della valutazione della ricerca scientifica sia stata accantonata. Sarà, speriamo ripresa nel prossimo anno, come sembra dimostrare la volontà di collaborazione di molti coordinatori." (Nota metodologica 2001).

nella famiglia della ricerca, descrivesse il livello di produttività scientifica delle singole facoltà.

1.3.4.2 Produttività e mondo del lavoro

Nella nota del 2004, il Censis fa riferimento alla produttività universitaria in termini di inserimento nel mondo del lavoro dei laureati.

I dati relativi alla carriera dei laureati sono raccolti dal consorzio Almalaurea: nonostante il miglioramento dei meccanismi di analisi e di interpretazione posti in essere, da cui derivano dati completi ed affidabili, non c'è completa copertura per l'intero sistema universitario italiano (alcuni Atenei infatti non vi collaborano) né è stata annullata l'intempestività delle informazioni³¹. Almalaurea è un servizio che dal 1994, su iniziativa dell'Osservatorio Statistico dell'Università di Bologna, rende disponibili on line i curricula dei laureati ponendosi come punto di incontro fra laureati, università ed aziende. Oggi circa 50 atenei aderiscono ad Almalaurea permettendo al 67% dei laureati italiani di entrare nel suo circuito³²: si tratta quindi di una copertura ancora parziale che non permette ad Almalaurea di essere fonte ufficiale di dati necessari alla costruzione delle graduatorie di facoltà e/o atenei.

Oltre ad Almalaurea, è nata nel 2002 in Italia una seconda iniziativa, denominata Stella (Statistica in TEma di Laureati e LAvoro), nato nel 2002 quale raggruppamento interuniversitario con l'obiettivo di realizzare una banca dati per il monitoraggio dei percorsi di studio, delle aspirazioni dei laureati e dei loro comportamenti nella ricerca di occupazione, compresa l'osservazione delle modalità di inserimento nel mercato del lavoro³³. Le finalità sono quindi le medesime di Almalaurea: al consorzio interuniversitario Stella aderiscono 12 Atenei (diversi da quelli aderenti ad Almalaurea). Le metodologie di analisi dei dati utilizzate dai due consorzi sono differenti per cui i dati generati non sono confrontabili, non rendono possibile la creazione di un'unica banca di dati omogenei né tanto meno fanno risultare sufficiente la copertura.

1.3.4.3 Didattica e strutture

Anche nella valutazione della didattica e delle strutture utilizzate dalle facoltà c'è carenza di dati disponibili all'elaborazione: fino a quando non sarà possibile accedere a dati tra loro

³¹ Cfr. Nota metodologica 2004.

³² Cfr. home page del sito web di Almalaurea (www.almalaurea.it).

³³ Tratto da <http://www.unica.it/pub/7/show.jsp?id=3904&iso=21&is=7>.

omogenei e confrontabili non sarà possibile pervenire a misure più accurate di valutazione di tali fattori³⁴.

In molti atenei sono gli studenti frequentanti a valutare l'aspetto della didattica attraverso appositi questionari: la metodologia di indagine – ovvero lo strumento di rilevazione, la modalità ed il periodo di raccolta dati, gli strumenti di analisi e di comunicazione dei risultati – varia però da ateneo ad ateneo e questo rende i dati non comparabili tra loro.

Relativamente alle strutture, oltre a quanto detto per la didattica, è da tenere in considerazione come sia spesso difficile attribuire aule, biblioteche, laboratori ed altre strutture ad una facoltà piuttosto che ad un'altra. È questo un problema sentito in prima istanza dagli Atenei che faticano a mettere a punto modelli di rilevazione consolidati.

1.3.5 Non solo facoltà

Negli ultimi due anni il Censis ha introdotto, unitamente ed a completamento delle graduatorie per facoltà, tre novità:

- un insieme di indicatori specifici per le facoltà di medicina e chirurgia e di medicina veterinaria;
- una graduatoria delle facoltà appartenenti ad atenei non statali;
- una classifica degli atenei statali.

1.3.5.1 Indicatori specifici per Medicina

Per quanto riguarda le facoltà di medicina e chirurgia e di medicina veterinaria, in accordo con le conferenze nazionali delle due facoltà, sono stati introdotti, dal 2005, specifici indicatori ricavati da una scheda di rilevazione ad hoc inviata ai presidi.³⁵ Per medicina e chirurgia è stato inserito nella famiglia della didattica un indicatore relativo al numero di posti letto a disposizione degli studenti. Relativamente al profilo docenti sono stati inseriti tre indicatori che tenessero conto del numero di *visiting professor*, di docenti reclutati da atenei stranieri e del totale dei professori reclutati con finanziamenti esterni. Per quanto concerne la famiglia della ricerca, i tre indicatori specifici entrati in questa

³⁴ Nota metodologica 2001.

³⁵ Cfr. Note metodologiche 2006 e 2007.

categoria sono atti a misurare l'*impact factor*, i progetti finanziati da enti pubblici e privati in rapporto ai docenti di ruolo ed infine contano il numero di brevetti registrati nel corso dell'ultimo triennio per docente di ruolo. Nella costruzione della graduatoria delle facoltà di medicina e veterinaria sono invece stati introdotti nuovi indicatori relativi alla sola famiglia della didattica, indicatori che contano il numero di strutture specifiche presenti e che riguardano il numero di casi clinici e di necroscopie relative agli animali per singolo studente.

1.3.5.2 Le Facoltà di Atenei non statali

La classifica delle facoltà di atenei non statali è stata elaborata confrontando le facoltà non statali con quelle degli atenei statali nelle famiglie della produttività, della ricerca e dei rapporti internazionali (è stata esclusa unicamente la famiglia relativa al "profilo docenti").

1.3.5.3 Gli Atenei

Per produrre invece la graduatoria degli atenei è stato studiato ed implementato un'apposita struttura metodologica che si articola in quattro famiglie. Riportiamo di seguito uno schema che sintetizza gli indicatori utilizzati dal Censis per la graduatoria degli atenei negli anni 2006 e 2007.

Tabella 3. Famiglie e indicatori utilizzati per costruire la graduatoria degli Atenei negli anni 2006 e 2007.

FAMIGLIA	2006	2007
SERVIZI	N. pasti erogati/iscritto	
	N. posti alloggio/1000 iscritti	N. posti e contributi alloggio/1000 iscritti
		N. spazi sportivi per 1000 iscritti
BORSE	N. borse di studio/100 iscritti	
	N. borse di collaborazione/100 iscritti	
STRUTTURE	Posti aula/100 iscritti	
	Posti nelle biblioteche 100 iscritti	
	Posti nei laboratori scientifici/100 iscritti	
	N. spazi sportivi/1000 iscritti	
WEB	Punteggio assegnato dal Censis Servizi ai siti internet degli atenei sulla base della funzionalità e dei contenuti	

Come evidenzia la tabella, le modifiche subite nei due anni riguardano sostanzialmente un unico indicatore: il numero di spazi sportivi per 1000 iscritti ha subito uno spostamento

dalla famiglia delle strutture alla famiglia dei servizi. Non vi è traccia nelle note metodologiche della motivazione che ha spinto a tale cambiamento: possiamo ipotizzare che l'attività sportiva sia stata fatta rientrare tra i servizi perché "accessoria" alla didattica che le strutture, quali aule, biblioteche e laboratori scientifici, supportano e quindi più omogenea, per finalità, ai servizi erogati dalle mense e dagli alloggi universitari. Si potrebbe trattare, in altre parole e comunque sempre secondo ipotesi, di uno di quegli aspetti che aumentano la qualità della vita dello studente ma non ne costituiscono un elemento costitutivo. Nella lettura delle classifiche di ateneo va infine considerato che, accanto ad elementi di giudizio attinenti alle competenze dirette degli Atenei (quali la valutazione del sito web di ateneo), sono valutate dimensioni d'offerta le cui responsabilità sono attribuibili a soggetti diversi (le borse di studio e i posti alloggio erogati dagli Enti per il Diritto allo Studio che a loro volta gestiscono risorse delle Regioni): si tratta quindi di dati che risentono di scelte periodiche – e finanziarie – estranee al sistema universitario, inteso in senso stretto³⁶.

1.3.6 Note esplicative

A partire dal 2001, ogni nota metodologica è stata corredata da alcune note esplicative che forniscono un aiuto – ed anche un chiarimento – nella lettura delle graduatorie prodotte. L'impianto delle note metodologiche è rimasto il medesimo tutti gli anni ad eccezione di alcuni cambiamenti relativi alle facoltà di recente istituzione.

La graduatoria viene stilata con riferimento all'anno accademico precedente; così, ad esempio, la classifica del 2004 prende in esame l'anno accademico 2002-2003. I dati, come detto nel paragrafo 3.3, sono soggetti ai criteri di ufficialità, omogeneità e certificazione.

La valutazione prende come unità statistica di analisi la facoltà: sono state considerate solo le facoltà istituzionali e non i corsi di laurea, seppure denominati come le stesse facoltà. Le facoltà presenti in meno di tre atenei sono state escluse dalla valutazione (è il caso, ad esempio, di Chimica Industriale di Bologna). Inoltre le facoltà di nuova generazione sono state escluse o parzialmente valutate, sostanzialmente per l'assenza di dati fondamentali per la costruzione della maggior parte degli indicatori (ad esempio

³⁶ Cfr. Nota metodologica 2007.

Architettura di Trieste e Parma, Economia di Bolzano). Un'ulteriore nota riferisce alle sedi distaccate che sono state valutate riaggregandole all'ateneo di appartenenza: la scelta è motivata dall'assenza di dati fondamentali (ad esempio per la Facoltà di Lettere di Arezzo, o per Medicina 1 e 2 della Sapienza di Roma) oppure dalla volontà di non spezzare "l'unicum" costituito dalla somma dei corsi di Laurea distribuiti in più sedi di una Facoltà (come per Ingegneria del Politecnico di Milano).

Relativamente alle facoltà considerate come *new-entry* nella graduatoria, le scelte di inclusione ovvero di esclusione sono mutate negli ultimi due anni. Dal 2001 al 2005 sono state considerate *new-entry* le Facoltà istituite da un numero di anni inferiore alla durata istituzionale incrementata di 2 anni. Nel 2006, dato il progressivo ingresso "a regime" di facoltà precedentemente escluse perché troppo "giovani", è stato rivisto il criterio di inclusione per cui le *new-entry* hanno ora la sola caratteristica di essere state istituite da un numero di anni inferiore alla propria durata istituzionale (senza i due anni aggiuntivi).

1.3.7 Gli indicatori utilizzati

Vogliamo ora andare al cuore della metodologia utilizzata dal Censis illustrando la struttura d'analisi impiegata, di anno in anno, per la costruzione delle graduatorie, in termini di indicatori semplici utilizzati.

La strada scelta per illustrare gli indicatori utilizzati per classificare le facoltà è articolata come segue:

- per ogni famiglia sono state individuate delle "classi" in cui sono stati inclusi gli indicatori che misurano il medesimo aspetto;
- è stata costruita una tabella nelle cui righe sono riportate le "classi" di indicatori e nelle cui colonne sono indicati gli anni di produzione della graduatoria (quindi dal 2000 al 2007); le celle vuote indicano l'assenza dell'indicatore per l'anno corrispondente mentre, se la cella è colorata, l'indicatore segnalato nella riga è presente nella struttura degli indicatori dell'anno in colonna.

Gli indicatori possono variare di anno in anno; nel caso in cui più celle – usualmente consecutive – siano presenti nella stessa riga, indicano che per quegli anni l'indicatore utilizzato è il medesimo (cambia naturalmente il solo riferimento temporale). Così, ad esempio, per la famiglia della produttività, l'indicatore che misura gli anni di laurea fuori

corso è presente in tutti gli anni dal 2001 al 2005 ed ha medesima struttura per gli anni 2003, 2004 e 2005. L'indicatore relativo invece al tasso di iscritti in corso è entrato nella struttura della famiglia della produttività dal 2001 con medesima forma per gli anni dal 2001 al 2005 mentre cambia, rispetto al passato, la formulazione per gli anni 2006 e 2007 per i quali l'indicatore è lo stesso. I valori numerici presenti in alcune celle delle tabelle corrispondono al numero di indicatori del medesimo tipo presenti nella famiglia considerata nell'anno indicato in intestazione di colonna.

1.3.7.1 Produttività

	2000	2001	2002	2003	2004	2005	2006	2007
laureati in corso (tasso)								
tasso dei laureati triennali								
età media dei laureati fuori corso								
anni di laurea fuori corso								
tasso di iscritti in corso								
sopravvivenza tra il 1° e il 2° anno (tasso)								
persistenza tra il 1° e il 2° anno (tasso)								
tasso studenti equivalenti/attivi								
iscritti con almeno un'annualità/crediti								
indice di regolarità degli studenti								

Analizzando la famiglia della produttività, sono individuabili due macro categorie relative ai laureati ed alla regolarità del percorso di studi.

Ogni anno sono presenti due indicatori che riferiscono ai laureati quali misura dell'efficienza dell'università nel raggiungere uno dei suoi obiettivi istituzionali: conferire il titolo di studio ai suoi studenti³⁷. Alcuni autori hanno cercato di elaborare complessi indicatori per valutare l'efficienza/inefficienza del processo didattico universitario. D'Ovidio propone un indicatore di efficienza che tenga conto non solo degli esiti finali (laurea o abbandono) ma anche delle situazioni intermedie (trasferimenti, iscrizioni fuori corso, mancate iscrizioni): elabora quindi una sola misura che, utilizzando appositi punteggi attribuiti alle varie situazioni di esito degli studi e ponderati con la frequenza di studenti che vi si trovano, misura la difformità media degli esiti dei vari percorsi di studi

³⁷ Porcu M, Puggioni G. (2002), Due classi di indicatori di inefficienza dei percorsi universitari, in Carli Sardi L., Delvecchio F., Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari, 2002, Cleup, Padova.

dall'obiettivo ottimale (laurea in corso di tutti gli studenti iscritti)³⁸. Porcu e Puggioni propongono alcuni indicatori che possano essere impiegati per valutare congiuntamente diverse componenti relative agli studenti che abbandonano gli studi o che conseguono il titolo in un tempo maggiore alla durata legale degli stessi; tale approccio si basa su considerazioni di costo (investimento sostenuto dall'istituzione università) e risultato/prodotto (laurea degli iscritti)³⁹. Si tratta indubbiamente di alcune elaborazioni molto interessanti ma la cui complessità non può rientrare in una metodologia di analisi i cui risultati sono destinati, quantomeno in prima istanza, a lettori non esperti di matematica né di statistica. Il tasso dei laureati in corso è presente nella famiglia della produttività ogni anno. L'età media dei laureati fuori corso utilizzata come indicatore nel 2000 è stata sostituita negli anni successivi, e fino al 2005, dal numero di anni di laurea dei fuori corso, indicatore indubbiamente più appropriato perché non è condizionato dall'età d'immatricolazione del singolo studente. Il tasso dei laureati triennali compare solo negli ultimi due anni: questo indicatore premia le facoltà che nell'ultimo triennio hanno portato il maggior numero di studenti alla laurea entro la durata regolare del corso e penalizza pertanto le facoltà dove la presenza di studenti fuori corso è maggiore.

Il gruppo di indicatori che misurano la regolarità dei corsi non è entrato nella struttura di analisi del primo anno (2000) mentre ha mantenuto una struttura simile negli anni successivi:

- il tasso di iscritti in corso calcolato ogni anno per premiare le facoltà con il maggior numero di iscrizioni depurato, negli ultimi due anni, degli immatricolati;
- un indicatore che valuta le iscrizioni alla facoltà nel momento più critico, ovvero dal primo al secondo anno, momento noto per l'elevata percentuale di abbandoni (rispetto agli altri passaggi da un anno di corso al successivo);
- un indice di regolarità basato sugli esami sostenuti (negli anni 2001, 2002 e 2003) o in base ai crediti acquisiti (dal 2004) durante l'anno accademico rispetto a quelli previsti.

³⁸ D'Ovidio F. (2002), Un indicatore di efficienza del processo didattico universitario, in Carli Sardi L., Delvecchio F., Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari, 2002, Cleup, Padova.

³⁹ Porcu M, Puggioni G. (2002), Due classi di indicatori di inefficienza dei percorsi universitari, in Carli Sardi L., Delvecchio F., Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari, 2002, Cleup, Padova.

1.3.7.2 Didattica

	2000	2001	2002	2003	2004	2005	2006	2007
rapporto docenti/studenti								
iscritti in complesso/docenti in complesso								
docenti di ruolo/iscritti								
docenti/corsi offerti								
docenti/insegnamenti attivi								
docenti di ruolo/insegnamenti								
somma CdL e CDU offerti dalla facoltà								
somma dei corsi con immatricolati								
num. classi di laurea nell'offerta di corsi								
somma degli insegnamenti offerti o attivi								
posti aula/iscritti						2	2	2
ore posto aula/studente in corso								
campus ONE (almeno un corso finanziato)								
valutazione didattica								
stage per iscritto								
rapporto ricercatori/docenti								
età media del corpo docente								

Gli indicatori appartenenti alla famiglia della didattica riguardano sostanzialmente cinque aspetti:

- il rapporto tra docenti e studenti iscritti;
- il rapporto tra docenti ed insegnamenti;
- l'offerta didattica in termini di Corsi di Laurea o di insegnamenti;
- la disponibilità di posti aula in rapporto agli studenti della facoltà;
- altri elementi complementari all'offerta delle facoltà/degli atenei quali il servizio stage, la valutazione della didattica, i corsi finanziati nell'ambito del progetto Campus One⁴⁰.

Ogni anno ha visto nella struttura metodologica un indicatore che valuta l'adeguatezza del numero di docenti per gli studenti iscritti alla specifica facoltà. Nel 2000 e negli ultimi due anni il rapporto ha avuto a numeratore il numero di docenti ed a denominatore il numero di iscritti premiando quindi le facoltà con il più alto numero di docenti (di ruolo nel

⁴⁰ Campus One è il progetto sperimentale della Conferenza dei Rettori delle Università Italiane che sperimenta la riforma universitaria; ha durata triennale (triennio accademico 2001/2004) e si rivolge in modo specifico ai nuovi corsi di laurea per sostenere e diffondere l'innovazione tecnologica e formativa conseguente alla riforma didattica universitaria. Finanziato dal Governo, intende accompagnare e sostenere i primi passi della riforma in tutte le università, con particolare attenzione a quelle del Mezzogiorno, e per tutte le aree disciplinari. Il progetto Campus One è gestito dalla CRUI. Offre agli atenei italiani la possibilità di applicare concretamente tutti gli aspetti della riforma attraverso azioni distinte rivolte al sistema universitario nazionale, al sistema di ateneo ed al singolo corso di laurea. (www.campusone.it)

2006 e 2007) per studente. Negli anni dal 2001 al 2005, invece, il rapporto risulta invertito premiando quindi le facoltà con il migliore – ovvero il più basso – rapporto iscritti/docenti: l'indicatore misurava quindi l'adeguatezza, in termini quantitativi, del corpo docente rispetto al numero di iscritti. L'inversione applicata dal 2006 rispetto agli anni precedenti (2000 escluso) potrebbe trovare giustificazione nella maggiore semplicità d'interpretazione derivante da un rapporto con "orientamento positivo": in altre parole, pur avendo i due indicatori nel merito il medesimo significato, risulta più immediato un giudizio positivo per indicatori il cui valore crescente dice una situazione migliore anziché attribuire valutazioni migliori per indicatori con valori minimi.

Dal 2001 è presente nella struttura di analisi un indicatore relativo al rapporto tra docenti ed insegnamenti offerti: tale indicatore premia le facoltà con il maggior numero di docenti per corso attivato ovvero la cui adeguatezza, in termini quantitativi, del corpo docente rispetto agli insegnamenti offerti agli studenti è migliore.

Ulteriori indicatori relativi alla didattica sono quelli che "quantificano" l'offerta formativa in termini di numero di Corsi di Laurea (CdL) e Corsi di Diploma Universitario (CDU) (dal 2000 al 2002), come somma dei Corsi con immatricolati (per il 2003, 2004 e 2005), come numero di classi di laurea presenti nell'offerta di Corsi (solo nel 2006). L'obiettivo di questi indicatori è quello di premiare le facoltà con il maggior numero di corsi offerti. Dal 2005 inoltre è stato introdotto un ulteriore valore pari alla somma degli insegnamenti offerti dalla facoltà per attribuire una valutazione migliore alle facoltà che non solo hanno un ampio numero di Corsi di Laurea da offrire agli studenti ma che offrono anche un'ampia gamma di insegnamenti. A tal proposito riteniamo utile portare una nostra considerazione. Non sempre un elevato numero di CdL accompagnato da un ampio numero di insegnamenti offerti corrisponde a situazioni realmente migliori rispetto ad altre i cui valori di questi due indicatori sono più bassi. Indagando sui numeri – e senza necessitare di strumenti molto complessi ma anche semplicemente utilizzando gli indicatori stessi che il Censis propone – troveremmo probabilmente che le stesse facoltà con molti CdL e molti insegnamenti offerti sono quelle con il maggior numero di docenti e con il maggior rapporto docenti/studenti. È realmente, questo, sintomo di una situazione positiva? Abbiamo ragione di credere che in una condizione caratterizzata da questi valori vi sia la tendenza ad impegnare l'intero corpo docente nella didattica magari sottraendo risorse alla ricerca; inoltre è facile supporre che la proliferazione di CdL e di insegnamenti sia risposta

all'elevato numero di docenti presenti in facoltà piuttosto che ad una reale domanda o necessità culturale-sociale dell'indirizzo di studi.

Indicatori che valutano l'adeguatezza delle strutture sono stati introdotti dal 2002 e premiano le facoltà con il maggiore rapporto posti aula/iscritti. Si tratta di indicatori "semplicistici" che mirano a fornire una valutazione delle strutture unicamente attraverso un rapporto tra il numero di posti nelle aule della facoltà ed il numero di iscritti alla facoltà stessa. Si tratta indubbiamente di una dimensione difficile da valutare per la grande assenza di informazioni: da un lato non vengono conteggiati altri spazi quali i laboratori e le biblioteche e dall'altro non è considerato che non tutti gli iscritti utilizzano le strutture perché, magari, sono studenti non frequentanti. Inoltre è importante valutare le strutture a disposizione degli studenti non solo in termini quantitativi ma anche – e, forse potremmo dire, soprattutto – a livello qualitativo. È quindi questo un aspetto indubbiamente da approfondire ed arricchire allorquando vi siano a disposizione dati sufficienti per poter valutare le strutture di tutte le facoltà.

Ulteriori indicatori sono quelli non relativi alla didattica in senso stretto ma che determinano la qualità della stessa premiando le facoltà con almeno un corso finanziato nell'ambito del progetto Campus One, quelle che monitorano l'offerta didattica attraverso diversificati sistemi di valutazione, le facoltà con il maggior numero di stages per studente.

Infine un breve commento per due indicatori utilizzati solo nei primi anni della realizzazione de "La Grande Guida all'Università". Il rapporto tra ricercatori e docenti del 2000 fornisce un giudizio positivo alle facoltà in cui tale valore è maggiore. L'età media del corpo docente, entrata nel 2001 nella famiglia della didattica, premia le facoltà con un corpo docente più giovane: tale indicatore negli anni successivi non è stato eliminato ma è stato, a ragione veduta, spostato all'interno della famiglia "profilo docenti".

1.3.7.3 Ricerca

	2000	2001	2002	2003	2004	2005	2006	2007
progetti realizzati con/per il CNR								
progetti cofinanziati Ministero								
progetti ministeriali/docenti facoltà		2						
num. unità di ricerca finanziate COFIN e FIRB/docente							3	3
tasso di successo partecipazione programma COFIN								
progetti presentati al cofinanziamento COFIN/docente								
progetti finanziati dai programmi comunitari/docente di ruolo								
progetti finanziati dai programmi comunitari								
centri di eccellenza cofinanziati MIUR								
finanziamento medio ministeriale per i progetti								
contributi alla ricerca CNR/docenti facoltà								
finanziamento medio ottenuto dalle unità di ricerca finanziate COFIN							3	3

La famiglia della ricerca in due grandi gruppi:

- un primo gruppo tiene conto dei progetti o unità di ricerca finanziati;
- il secondo gruppo riguarda i finanziamenti ottenuti.

Negli ultimi due anni è aumentato il numero di indicatori relativi alla misurazione dei progetti presentati e finanziati nonché al numero di unità di ricerca finanziate per docente (tre indicatori sia per il 2006 che per il 2007). Un indicatore che premia le facoltà con docenti partecipanti ai centri di eccellenza cofinanziati dal Ministero è comune agli anni dal 2002 al 2005 ma non ha trovato continuità negli anni successivi.

Per quanto riguarda la valutazione dei finanziamenti ottenuti dalle facoltà, dal primo anno sono stati considerati i finanziamenti ministeriali, a cui sono stati affiancati nel solo 2001 i contributi alla ricerca Centro Nazionale Ricerche (CNR) per premiare quelle facoltà con docenti in grado di accedere ai contributi alla ricerca del CNR. Nel 2006 e 2007 sono invece stati considerati tre indicatori relativi al finanziamento medio ottenuto dalle unità di ricerca.

Confrontando la struttura che la famiglia della ricerca ha assunto nei diversi anni in termini di indicatori è anzitutto evidente come negli ultimi due anni sia notevolmente incrementato il numero di indicatori utilizzati in questa famiglia: ciò significa una maggiore attenzione ed una più forte considerazione dell'aspetto della ricerca come fattore chiave di

qualità di una facoltà. Inoltre va registrato che negli anni dal 2002 al 2005 l'analisi è stata compiuta sui medesimi tre indicatori: i dati di questi anni hanno quindi la caratteristica di essere completamente confrontabili.

1.3.7.4 Profilo docenti

	2002	2003	2004	2005	2006	2007
età media del corpo docente						
indice di invecchiamento del corpo docente						
età media degli idonei nei concorsi						
tasso di idonei/docente						
tasso insegnamenti tenuti da docenti extra accademici						
num. università ospitanti studenti Erasmus/docente						
studenti Erasmus in uscita/docente						
finanziamenti progetto "rientro dei cervelli"						
docenti coinvolti nel programma per il c.d. "rientro dei cervelli"						
docenti che hanno ottenuto finanziamenti per il c.d. "rientro dei cervelli"						

Il profilo docenti è entrata nell'analisi solo nell'anno 2002 mantenendo fissi alcuni indicatori: in particolare l'età media del corpo docente presente nel 2001 all'interno della famiglia della didattica ha trovato corretta collocazione all'interno del fattore relativo al profilo docenti. Inoltre l'indice di invecchiamento del corpo docente, sempre presente dal 2002, premia le facoltà che hanno registrato il minor invecchiamento del corpo docente nel periodo considerato. Negli anni dal 2003 al 2005 è poi stata considerata anche l'età media degli idonei valutando così positivamente le facoltà con età media degli idonei più bassa. Un ulteriore indicatore riferito agli idonei ai concorsi è il tasso di docenti idonei ai concorsi sui docenti di ruolo della facoltà al netto degli ordinari: naturalmente più questo valore è elevato, miglior giudizio viene attribuito alla facoltà. Non affrontiamo in questa sede, poiché esula dagli obiettivi dell'elaborato, considerazioni e critiche relative ai concorsi: rimandiamo alla vasta letteratura in merito⁴¹ e ci limitiamo ad osservare che probabilmente sarebbero necessari ulteriori indicatori per valutare accuratamente questo aspetto. L'ultimo indicatore, presente negli anni 2006 e 2007, è il tasso di insegnamenti tenuti da docenti extra accademici che conferisce un punteggio favorevole alle facoltà la

⁴¹ Riportiamo due utili testi in merito:

- Perotti R. (2002), The italian university system: rules vs incentives, in "Annual Report on Monitoring Italy 2002", Istituto di Studi e Analisi Economica (ISAE), Roma.
- Pellini P. (2006), La riforma Moratti non esiste, Il Saggiatore, Milano.

cui offerta in termini di corsi prevede una composizione equilibrata tra insegnamenti tenuti da docenti di ruolo e da docenti non di ruolo extra accademici: si tratta infatti di un elemento di ricchezza per le facoltà che traduce un'ottica di positiva apertura al mondo esterno all'università.⁴²

Anche l'aspetto relativo al progetto Erasmus⁴³ rientra nella famiglia che stiamo ora esaminando rapportato al totale docenti di ruolo della facoltà: vengono quindi premiate le facoltà con apertura maggiore verso esperienze internazionali quale l'Erasmus. Tale aspetto viene considerato anche nella famiglia dell'internazionalizzazione ma in rapporto al totale degli studenti.

Un ultimo gruppo di indicatori comprende quelli relativi al progetto "rientro dei cervelli" che dal 2002 caratterizza la famiglia del profilo docenti: essi calcolano i finanziamenti (dal 2002 al 2005) per progetti di rientro dei docenti di prestigio – secondo relativa legge ministeriale⁴⁴ – oppure forniscono il numero di docenti (anni 2006 e 2007) che hanno partecipato al programma di incentivazione della mobilità degli studiosi italiani e stranieri.

⁴² Personalmente riteniamo che questo indicatore sia adatto a dare valutazioni positive quando riferito a Corsi di Laurea triennale mentre può assumere letture più delicate se riferito a Corsi di Laurea magistrale sostanzialmente per due motivi: un'elevata qualità dei docenti interni – in termini di competenze, conoscenze ed esperienza – avvalorata in modo forte i Corsi di Laurea magistrale; in secondo luogo la diversa mission delle due tipologie di Corsi di Laurea (più professionalizzante nelle triennali e specialistica nelle magistrali) richiede docenti scientificamente più attrezzati nelle lauree di secondo livello e docenti professionalmente più esperti nelle lauree triennali.

⁴³ Il progetto Erasmus (acronimo di European Community Action Scheme for the Mobility of University Students) nasce nel 1987 per opera della Comunità Europea e sancisce la possibilità di uno studente universitario europeo di effettuare in un'università straniera un periodo di studio legalmente riconosciuto dalla propria università.

⁴⁴ MIUR, Decreto Ministeriale 26 gennaio 2001 n. 13, "Incentivi a favore della mobilità di studiosi italiani e stranieri impegnati all'estero".

1.3.7.5 Rapporti internazionali

	2000	2001	2002	2003	2004	2005	2006	2007
progetti finanziati e coordinati 4° Programma Quadro								
somma attività/totale docenti								
percentuale docenti coinvolti progetti 5° e 6° Programma Quadro								
num. partners internazionali coordinati								
finanziamenti per la ricerca internazionale								
mobilità docenti		2						
opportunità internazionali								
borse Erasmus attivate								
mobilità studenti (borse Erasmus/iscritti meno immatricolati)								
studenti che hanno trascorso un periodo all'estero/aventi diritto								
Erasmus in uscita/studente								
Erasmus in entrata/studente								
borse Erasmus/docenti facoltà								
num. università ospitanti/docente								

La famiglia dell'internazionalizzazione è stata definita negli anni dal 2001 al 2005 "Rapporti e cooperazione internazionale" mentre ha assunto la denominazione "Rapporti internazionali" negli ultimi due anni. Le similarità maggiori si riscontrano nelle strutture degli anni dal 2001 al 2005 mentre altri aspetti accomunano le ultime due versioni del lavoro (anni 2006 e 2007).

Ad eccezione degli ultimi due anni, c'è un riferimento esplicito ai progetti coordinati dai Programmi Quadro con indicatori che premiano le facoltà con maggiore partecipazione – ovvero un numero più elevato di progetti – ai programmi quadro sul totale docenti della facoltà. Nel 2004 e 2005 l'indicatore coincideva con la percentuale di docenti coinvolti nei progetti dei programmi quadro V e VI.

Nel 2000 è entrato nell'analisi il numero di partners internazionali coordinati nel IV Programma Quadro assieme all'ammontare dei finanziamenti per la ricerca internazionali. La mobilità docenti, entrata nell'analisi degli anni dal 2001 al 2005, attribuisce valutazione positiva alle facoltà che hanno ricevuto finanziamenti per opportunità internazionali. Analogamente negli ultimi due anni, l'indicatore relativo alle opportunità internazionali è stato elaborato per premiare le facoltà che hanno ricevuto finanziamenti per attività di cooperazione internazionale.

Un blocco consistente di indicatori è quello relativo all'esperienza Erasmus: dal calcolo delle borse Erasmus attivate del 2000, si passa a calcolare un indicatore della mobilità studenti (come rapporto tra il numero di borse attivate ed il numero di iscritti al netto degli immatricolati) negli anni dal 2001 al 2005 che indica la possibilità degli studenti di partecipare ad esperienze internazionali con i progetti Erasmus. Tale indicatore trova un'estensione nel 2007 per cui viene calcolato il rapporto tra gli studenti che hanno trascorso un periodo di studio all'estero finanziato mediante Erasmus ma anche con altri programmi ed il numero di iscritti al netto degli immatricolati. Per premiare le facoltà con un alto numero di studenti Erasmus in uscita ed in entrata sono stati elaborati due specifici indicatori: il valore degli Erasmus in entrata per studente è entrato nell'analisi del 2006 e del 2007 mentre per gli Erasmus in uscita l'indicatore è stato elaborato solo per il 2006. Per il 2001 e fino al 2005 è stato calcolato il medesimo indicatore che premia le facoltà con più elevato rapporto tra il numero di borse Erasmus attive ed il numero di docenti della facoltà. Infine, nel 2006 e nel 2007, il numero di università straniere ospitanti studenti Erasmus sul totale dei docenti di ruolo della facoltà è stato l'indicatore che ha mirato a premiare le facoltà con il maggior rapporto tra numero di università ospitanti e corpo docente di ruolo.

1.3.7.6 Attrazione

	2000	2001
docenti facoltà/docenti ateneo		
docenti facoltà/docenti facoltà omogenee		
studenti facoltà/ studenti ateneo		
studenti facoltà/ studenti facoltà omogenee		
quote di mercato		
variazione quote immatricolati su totale facoltà omogenee		
immatricolati con voto 54-60/totale immatricolati		
immatricolati fuori regione/totale immatricolati		

La famiglia dell'attrazione comprende indicatori che nei primi due anni il Censis ha utilizzato per valutare la capacità delle facoltà di avere influenza sul contesto locale e nazionale confrontando quote di studenti o docenti delle singole facoltà con il complesso dell'ateneo oppure con facoltà omogenee. Così un rapporto elevato tra il numero di docenti della facoltà sul totale docenti dell'ateneo indica maggior influenza della facoltà rispetto all'ateneo in cui è inserita; analogamente va interpretato l'indicatore dato dal rapporto tra gli studenti della facoltà ed il totale degli studenti iscritti all'ateneo. L'influenza

sul contesto nazionale è invece dato dai due indicatori che confrontano il numero di docenti e di studenti della singola facoltà con il totale studenti e docenti delle facoltà omogenee.

Nel 2000 il Censis ha inserito un indicatore tratto dall'Istat relativo alle quote di mercato: nella relativa nota metodologica non viene però illustrato il significato di tale valore.

Ulteriori tre indicatori sono stati inseriti nel 2001 per premiare

- la migliore variazione di immatricolati rispetto al contesto nazionale di una data facoltà (variazione quote immatricolati),
- le facoltà con un tasso maggiore di immatricolati "eccellenti" ovvero con un valutazione all'esame di maturità compresa tra 54 e 60,
- la maggior capacità di attrarre immatricolati provenienti da regioni diverse da quelle di appartenenza della facoltà rispetto al totale degli immatricolati.

La famiglia relativa all'attrattività delle facoltà è stata valutata solo nei primi due anni di elaborazione de "La Grande Guida all'Università": a nostro parere si tratta di una valutazione indubbiamente necessaria per definire la qualità totale di una facoltà. Considerazioni ulteriori in merito sono esposte nel paragrafo successivo.

1.3.8 Indicatori e famiglie a confronto: considerazioni finali

Dalla nota metodologica del 2002 si evince quale vuol essere – e quale si dice sia – il ruolo di questo lavoro:

- accrescere il livello di conoscenza del mondo universitario presso l'opinione pubblica;
- stimolare, dentro il sistema universitario, una maggiore sensibilità ai temi della valutazione;
- spingere attori importanti, quali i presidi delle facoltà, ad avviare un percorso collaborativo per il miglioramento del loro lavoro e del lavoro del Censis.⁴⁵

Indubbiamente "La Grande Guida all'Università" ha generato ricche discussioni, talvolta costruttive, a volte critiche in senso stretto, attorno ai risultati che il lavoro di anno in anno evidenzia. L'opinione pubblica è stata quindi realmente stimolata in questo senso e lo è stato ancor più il sistema universitario, nonostante la connaturata reticenza all'essere oggetto di valutazione. Il quotidiano *The Times*, accompagnando una recente guida

⁴⁵ Cfr. Nota metodologica 2002.

all'università britannica, scrive quando segue: "Ogni anno le classifiche sono state oggetto di un vivace dibattito tra gli accademici che ha aperto la strada ad un numero di articoli e servizi giornalistici che presentavano simili classifiche e che hanno inevitabilmente generato una certa confusione [...] è [...] sorprendente che molti atenei si oppongano all'idea di essere confrontati ad altri atenei quando ciò è quello che gli studenti fanno da sempre". Si tratta di considerazioni sicuramente estendibili anche al panorama italiano e che sono, in certo senso, "consolanti": anche, infatti, dove la cultura della valutazione è più consolidata dell'Italia – come appunto succede in Gran Bretagna –, le graduatorie sono lette come semplici classificazioni di merito e prestigio, con un primo ed un ultimo posto, anziché fungere in primo luogo da suggerimento per gli studenti rispetto ad alcune funzionalità delle facoltà stesse⁴⁶. Crediamo si tratti, però, di un processo che necessita di un tempo adeguato per entrare a pieno titolo nel panorama concernente l'università italiana e per diventare elemento consolidato dell'opinione pubblica al fine di ottenere il riconoscimento prefissato negli obiettivi del lavoro stesso: essere uno studio comparativo sulle facoltà italiane mirato ad essere "di servizio" e di orientamento agli studenti ed alle loro famiglie⁴⁷.

Il coinvolgimento crescente di anno in anno di attori importanti del sistema universitario, quali i presidi delle singole facoltà, va letto come un fatto senza dubbio positivo. In particolare la costruzione di indicatori ad hoc per alcune facoltà – medicina veterinaria e medicina e chirurgia, dal 2005 ad oggi – aspira ad essere di auspicio ad una collaborazione più stretta tra gli organi accademici, come la Conferenza Nazionale dei Presidi di Facoltà, ed il Censis al fine di ottenere graduatorie specifiche che tengano in considerazione aspetti assolutamente peculiari delle diverse facoltà per fornire una valutazione "a tutto tondo" dei raggruppamenti di facoltà individuati.

Il lavoro del Censis per *La Repubblica*, come abbiamo visto nell'analisi delle famiglie e degli indicatori utilizzati nelle edizioni annue della guida all'università, vuole avere l'ambizioso – e quanto mai legittimo – intento di percepire le trasformazioni interne al sistema universitario che, come detto, è un sistema di reale dinamismo non "affetto" da ipotetica stabilità. Molti sono i cambiamenti, intervenuti negli ultimi anni, che hanno trasformato l'università, modificazioni che hanno avuto riflesso anche nella struttura di indicatori e famiglie utilizzati ne "La Grande Guida all'Università":

⁴⁶ Cfr. Nota metodologica 2007.

⁴⁷ Cfr. Nota metodologica 2000.

- una notevole crescita di investimenti pubblicitari che dal 2000 hanno portato i singoli atenei a finanziare campagne di comunicazione e di promozione della propria immagine e della propria offerta didattica;
- la riforma universitaria⁴⁸, cambiamento di grossa portata con effetti ancora difficili da valutare;
- una forte spinta all'internazionalizzazione, sia nella didattica che nella ricerca;
- la crescente autonomia degli atenei che ha portato a modelli gestionali estremamente differenziati relativamente alle risorse per il reclutamento dei docenti, alla semplificazione della discussione delle tesi di laurea, alla costituzione di consorzi, società e fondazioni, etc.;
- l'utilizzo sempre maggiore del sito di ateneo o di facoltà per informare e offrire servizi on-line agli studenti;
- la proliferazione dell'offerta formativa – effetto della riforma universitaria – tesa ad accrescere l'accesso di immatricolati.

In questo senso vanno quindi letti gli "adeguamenti" che di anno in anno ha subito la struttura degli indicatori utilizzati dal Censis.

Nonostante i continui e premianti sforzi di miglioramento, permangono però alcuni elementi critici. La famiglia dell'attrazione, come anticipato, è presente solo nel 2000 e nel 2001. Riteniamo che gli indicatori rientranti in questa famiglia abbiano davvero una notevole importanza: il fatto che non siano stati utilizzati – né sostituiti da altri indicatori simili – negli anni successivi comporta, a nostro parere, una lacuna nell'impianto valutativo realizzato dal Censis. Indubbiamente l'introduzione negli anni di nuovi indicatori atti a valutare in maniera più approfondita la dimensione della ricerca e dei rapporti internazionali traduce da un lato la necessità – del sistema Paese – di investire nella ricerca e dall'altro la rinnovata spinta del sistema universitario verso un'apertura sempre maggiore all'Europa. Siamo però convinti che la capacità delle singole facoltà di avere un'influenza notevole sul territorio nazionale ne conferisce migliore qualità e prestigio: sarebbe pertanto quantomeno conveniente reintrodurre nella struttura degli indicatori e delle famiglie l'aspetto relativo alla capacità delle singole facoltà di esercitare maggiore o minore attrazione rispetto all'ateneo, conferendo, di riflesso, maggior prestigio anche ad esso, e rispetto alle facoltà omogenee risultando la più prestigiosa tra cui scegliere. Un

⁴⁸ La riforma universitaria è stata varata con Decreto Legislativo num. 509 del 1999; ha iniziato ad operare, in generale, tra il 2000 ed il 2001.

ulteriore fattore non adeguatamente approfondito è quello relativo alle strutture messe a disposizione degli studenti. Molte sono le difficoltà nel ricavare informazioni sui posti disponibili nelle aule adibite alla didattica, nei laboratori (linguistici, informatici, etc.), nelle biblioteche e negli spazi riservati allo studio ma certamente valutazioni positive di questi aspetti si traducono in forte capacità di attrazione. A completamento del giudizio sulla potenziale influenza nel contesto locale e nazionale, andrebbero considerati anche tutti i servizi allo studente quali la disponibilità di luoghi di ristorazione, di alloggi per studenti, buoni collegamenti viari e/o ferroviari, accessibilità alle strutture per i disabili, etc.: si tratta di elementi molto difficile da valutare con parametri oggettivi ma la cui validità nell'aver un peso notevole anche nella scelta della facoltà è indiscutibile. Anche la disponibilità di un elevato patrimonio culturale, misurato, ad esempio, attraverso il numero di volumi e riviste scientifiche presenti nelle biblioteche di facoltà influisce positivamente nel giudizio attribuito alla facoltà dallo studente chiamato ad una scelta.

Tra gli elementi positivi va assolutamente menzionata l'introduzione, negli ultimi due anni, dell'idea di stilare una graduatoria degli atenei ed una per le facoltà appartenenti ad atenei non statali è risultata convincente. Si tratta infatti di ulteriori indicatori di qualità che possono condizionare la scelta di un futuro studente universitario e della sua famiglia quanto diventano utili per gli studenti che scelgono, ad esempio, di proseguire gli studi specialistici in un ateneo diverso da quello di ottenimento della laurea triennale. Altro aspetto positivo è la migliore analisi elaborata per la famiglia della ricerca, elemento qualificante di una facoltà rispetto ad un'altra omogenea magari più votata alla didattica. Come il Censis stesso rileva, molti sono i problemi connessi alla misurazione della ricerca: la produttività scientifica, ad esempio, soffre di ostacoli connaturati alla natura stessa dell'oggetto per cui diventa molto complesso semplicemente "contare" quante e quali siano le pubblicazioni di carattere scientifico elaborate dai docenti della facoltà.

Infine vogliamo porre un'osservazione di metodo: è assente dall'impianto metodologico una qualsiasi misura delle variazioni che nel tempo assumono le facoltà nella graduatoria ma anche i punteggi delle famiglie considerate. Un semplice simbolo che indichi se la posizione in graduatoria della facoltà sia aumentata (▲), peggiorata (▼) o rimasta stabile (=) non può essere sufficiente ad illustrare e motivare i cambiamenti intervenuti durante l'anno che hanno determinato più o meno forti variazioni. Sarebbe infatti interessante capire – probabilmente in modo particolare per i lettori più direttamente coinvolti nel sistema universitario – che cosa ha provocato uno slittamento, in basso o in alto, della

singola facoltà nella graduatoria: quale famiglia ha migliorato o peggiorato il proprio punteggio? Cosa è peggiorato? E quale aspetto può essere migliorato? Le tabelle di sintesi che il Censis fornisce ne "La Grande Guida all'Università" non permettono da sole un confronto "serio" col passato: è quindi, secondo noi, auspicabile che nel futuro venga considerata l'ipotesi di elaborare uno strumento statistico-matematico di analisi che fornisca appunto una misura dei cambiamenti avvenuti di anno in anno nella graduatoria rispetto alle singole facoltà valutate.

Nel complesso la struttura utilizzata risulta ricca di molti interessanti aspetti ma evidenzia, come noto al Censis stesso, alcune lacune, come sopra indicato. Siamo anche però coscienti che, dato l'obiettivo del lavoro, una strutturazione troppo complessa non potrebbe avere come destinatari quelli che il Censis ha pensato per "La Grande Guida all'Università": si tratta quindi di cercare sempre il miglior compromesso per valutare accuratamente il sistema universitario senza però ostacolarne una chiara lettura.

CAPITOLO 2

*"When you can measure what you are speaking about,
and express it in numbers, you know something about it;
but when you cannot measure it, when you cannot express it in numbers,
your knowledge is of a meagre and unsatisfactory kind:
it may be the beginning of knowledge,
but you have scarcely in your thoughts advanced to the stage of science."
Lord Kelvin*

2.1 VALUTARE

Nell'analisi delle relazioni tra stratificazione sociale e riuscita scolastica, Trivellato sottolinea come nuove dinamiche sociali abbiano contribuito a dare nuovo impulso all'attenzione verso la costruzione di indicatori originali di partecipazione e risultato dei processi scolastici. In tal senso hanno infatti agito l'evoluzione della situazione scolastica e l'intenso sviluppo della scolarità – e quindi la frazione crescente di risorse, pubbliche e nazionali, impiegate nella scuola – che hanno determinato un più esteso interesse ai problemi di valutazione e di misura sia relativamente ai risultati scolastici, sia per quanto riguarda la produttività del sistema scolastico nel perseguirli⁴⁹. Giustificata altresì dal salto temporale e dai mutamenti sociali degli ultimi decenni, l'idea di un parallelismo con il mondo universitario appare oggi immediata: il notevole incremento nell'accesso all'istruzione universitaria ha portato in sé la necessità di sviluppare adeguati sistemi valutativi che fornissero una fotografia della realtà dell'università.

Ci chiediamo quindi cosa significhi valutare, fare valutazione. La valutazione è sicuramente una delle attività umane che accompagnano la vita di ogni individuo.

⁴⁹ Trivellato U. (1975), Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine sociale sulla riuscita scolastica, Fondazione "Emanuela Zancan", Padova, pag. 23.

Ciascuno di noi, prima di compiere una scelta, valuta proposte alternative in base a giudizi basati sul confronto degli effetti prodotti – o producibili – da una decisione con il sistema di preferenze che orienta il comportamento del decisore⁵⁰. Si attua quindi un processo che determina l'espressione di un giudizio, conseguenza di valutazioni oggettive, fondato su una serie di informazioni acquisite e di criteri soggettivi basati su percezioni personali.

Diversi autori assegnano al termine valutazione definizioni differenti. Secondo Schulberg e altri autori, la valutazione è il processo di determinazione del livello di successo ottenuto nel perseguire un obiettivo prefissato; per Weiss valutare significa studiare gli effetti di un programma rispetto alle conseguenze attese e non attese per la popolazione e le organizzazioni coinvolte. Sintetizzando queste due definizioni potremmo dire che la valutazione si incentra sulla messa a disposizione di informazioni atte a permettere di definire in che misura un determinato intervento ha prodotto l'effetto sperato, cercando di capire quali effetti non previsti sono stati prodotti⁵¹. In quest'ottica, la valutazione del sistema universitario può venire letta come un'analisi tesa a definire – ex post – lo scostamento degli esiti ottenuti, in termini di titoli di studio conferiti, dall'obiettivo ottimale dell'università ovvero l'attribuzione della laurea a tutti gli studenti iscritti nei tempi previsti dalla durata legale del Corso di Laurea.

Si tratta però di un approccio limitato – e limitante – che esclude la possibilità di valutare lo "stato d'arte" dell'università, tenendo in considerazione fattori ex-ante e non solo ex-post, premesse e non solo risultati.

L'analisi della qualità di un servizio deve avvenire attraverso un approccio multidimensionale che definisce un insieme di criteri valutativi costruiti su indicatori relativi alle diverse fasi di produzione di un servizio:

- input: insieme delle risorse materiali ed umane immesse nell'organizzazione per ottenere un determinato risultato;
- processo: insieme delle modalità operative ed organizzative adottate per produrre il servizio;
- output: prestazioni concretamente erogate;
- outcome: effetti concretamente prodotti.

⁵⁰ Cfr. Bertin G. (a cura di) (1995), Valutazione e sapere sociologico: metodi e tecniche di gestione dei processi decisionali, FrancoAngeli, Milano.

⁵¹ Cfr. Bertin G., (1989), Decidere nel pubblico – tecniche di decisione e valutazione nella gestione dei servizi pubblici, Gruppo Editoriale Fabbri, Bompiani, Sonzogno, Etas libri, pag. 130.

Per questo Shaw⁵² ha definito sei criteri di valutazione della qualità:

- **efficacia** per misurare il livello di realizzazione degli obiettivi: si tratta di una misura dell'outcome;
- **efficienza** per definire la qualità dell'output come rapporto tra servizi erogati e risorse utilizzate;
- **equità** come criterio per evitare discriminazioni nell'accesso ai servizi basate su costo e modalità di erogazione;
- **sensibilità o soddisfazione** per valutare la rispondenza alle aspettative;
- **appropriatezza** per verificare la congruenza, qualitativa e quantitativa, delle risorse disponibili alla domanda;
- **accessibilità** come distribuzione adeguata delle risorse disponibili rispetto agli obiettivi.

Nessuno di questi criteri da solo è sufficiente a rappresentare compiutamente la qualità di un servizio ma vanno considerati nella loro totalità per una valutazione completa dell'oggetto in esame.

Rispetto quindi al lavoro che in questo elaborato si intende esaminare, è necessario supporre che il processo di costruzione dell'impianto valutativo, realizzato a priori dal Censis, abbia preso in considerazione tutti i criteri sopra esposti. Naturalmente, come già sottolineato nel capitolo 1, si sarà ricorso al miglior compromesso tra la completezza e la complessità richiesta dalla teoria della valutazione e la reale disponibilità di informazioni ufficiali. "La Grande Guida all'Università" fornisce infatti una vera e propria valutazione delle facoltà italiane (ed anche degli atenei, negli ultimi due anni) producendo l'espressione di un giudizio da parte dei lettori destinatari primi che attuano il personale processo decisionale anche – ma, auspichiamo, non solo – in base alle informazioni che le graduatorie del Censis forniscono.

2.2 MISURARE

Nell'effettuare una ricerca scientifica tramite indagine, due sono le grandi dimensioni di cui l'indagine si compone: l'astrazione e il metodo. La fase dell'astrazione è molto

⁵² Shaw C. D. (1986), Quality Assurance – What the Colleges Are Doing, King's Fund Centre, London.

articolata e da essa dipendono i risultati ottenuti dall'applicazione del metodo: viene fissato l'obiettivo della ricerca, vengono raccolte dalla letteratura le conoscenze già utili alla realizzazione dell'indagine (anche per evitare di duplicare le informazioni; viene cioè elaborato il razionale della ricerca), viene stabilito l'aggregato su cui indagare e soprattutto viene definito l'oggetto della ricerca. In quella che è definita dimensione del metodo, invece, vengono raccolte le diverse fasi operative, metodologiche appunto, attraverso le quali elaborare le informazioni e produrre risultati rispondenti agli obiettivi di ricerca. Le scelte non neutrali del ricercatore vengono sviluppate durante la fase di astrazione mentre nella parte metodologica della ricerca si cerca di ottenere la massima riproducibilità attraverso elaborazioni scientifiche oggettive.

Il primo passo da compiere è quindi quello di stabilire l'oggetto della valutazione, della misurazione. Si tratta cioè di definire ciò che è definito **concetto** ovvero un costrutto logico assunto come simbolo di un fenomeno, un'astrazione mentale con cui rappresentare un insieme di entità, di oggetti, di proprietà. Con il termine concetto ci si riferisce al contenuto semantico, cioè al significato, dei segni linguistici e delle immagini mentali; etimologicamente inoltre (dal latino *cum capio* = prendere assieme) indica l'azione di ordinare molteplici aspetti sotto un unico pensiero. Il concetto è quindi il mezzo attraverso cui l'uomo può conoscere e pensare, è il fondamento di ogni disciplina scientifica.⁵³ Esistono concetti che riferiscono ad entità direttamente osservabili e misurabili (come ad esempio il colore degli occhi) ed altri che invece fanno riferimento a costruzioni mentali astratte non direttamente osservabili (come la giustizia, la felicità). La definizione di un concetto non è fissa nel tempo ma evolve secondo la maturazione di un gruppo, della società: dipende quindi da un maggior affinamento degli strumenti di analisi nonché da un aumento della sensibilità che la società mostra verso quel determinato argomento. Così è naturale che il processo valutativo del sistema universitario sia mutato nel tempo: gli strumenti – ovvero gli indicatori – sono diventati più accurati oltre ad essere stati resi disponibili in quantità maggiore da più fonti ed ha subito un'evoluzione anche il significato che l'università ha nella società.

È necessario quindi tradurre i concetti in modo empirico: secondo tale procedura, definita **operazionalizzazione**, il significato di un concetto viene ricondotto ad un numero finito di operazioni e di predicati empirici osservabili. L'operazionalizzazione ha fondamentale importanza perché è ciò che permette di passare dalla teoria alla realtà

⁵³ Corbetta P. G., (1999), Metodologia e tecniche della ricerca sociale, Il Mulino, Bologna, pag. 91.

empirica osservabile: infatti, come scrisse Blumer, "la teoria ha un senso per la scienza empirica solo nella misura in cui riesce a connettersi proficuamente col mondo empirico; i concetti sono gli strumenti, e gli unici strumenti, per stabilire una tale connessione"⁵⁴.

Quanto detto finora è particolarmente vero in ambito sociale: spesso infatti l'oggetto, il concetto da valutare non ha immediata possibilità di essere misurato né è direttamente osservabile. Al fine di poter quindi applicare i metodi quantitativi che la statistica offre per analizzare i concetti, è necessario tradurre le caratteristiche qualitative in numeri: operationalizzare il concetto permette di acquisire conoscenza su esso attraverso la sua traduzione empirica in aspetti "quantificabili"⁵⁵. Il ricercatore deve quindi individuare una grandezza X , detta **variabile latente**, in stretta relazione con il concetto originario che sia misurabile attraverso uno o più valori osservabili: si produce una misura indiretta del concetto originario per mezzo di un valore X ottenuto come sintesi o combinazione di più misurazioni. Si passa pertanto ad individuare una **variabile** che è l'operationalizzazione del concetto iniziale, la quale assume un determinato valore X . La variabile può assumere diverse **modalità** che corrispondono agli "stati operativizzati" del concetto; per quantificare le diverse modalità con cui un concetto si manifesta, vengono loro assegnati dei **valori** simbolici, solitamente costituiti da numeri. Di fatto quindi operationalizzare un concetto significa compiere procedure di misurazione, ordinamento e classificazione.

Possiamo quindi riassumere nel modo seguente quanto sopra esposto: in campo sociale la valutazione, quantitativa e qualitativa, di un dato fenomeno designato da un concetto, è intesa come la procedura mediante la quale si determina il livello di un carattere posseduto da una determinata unità statistica, livello che può essere sia quantitativo – diventando così un valore – sia qualitativo⁵⁶. L'informazione quantitativa si trova ad un livello superiore rispetto a quella qualitativa che fa da base: l'interrogativo *quanto* presuppone, infatti, spesso implicitamente, un *quale*, mentre il viceversa non è così scontato. Il compito dello statistico è quindi quello di tradurre le caratteristiche quantitative in numeri: il processo di misurazione fa sì che la proprietà misurata sia rappresentata e quantificata e riflette le relazioni empiriche di interesse in relazioni algebriche tra i valori numerici assegnati⁵⁷. Si utilizzano quindi due tipi di strumenti:

⁵⁴ Blumer H., (1939) An appraisal of Thomas and Znaniecki's "The Polish Peasant in Europe and America", New York, Social Science Research Council, Bulletin 44.

⁵⁵ Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano, pag. 197.

⁵⁶ Cfr. Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari, pag. 9.

⁵⁷ Cfr. Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies

- le **scale di misura**, per misurare variabili endogene al fenomeno ovvero variabili che riguardano la struttura e la natura del fatto in analisi e non di fatti simili o correlati ad esso;
- gli **indicatori sociali**, per valutare fenomeni non direttamente osservabili: si osservano infatti fenomeni simili o correlati con il fenomeno sociale di interesse, le cui misure siano direttamente osservabili ed indichino il concetto che si intende valutare⁵⁸.

2.3 GLI INDICATORI

Come solitamente avviene nelle scienze sociali, i concetti per i quali si intende produrre una misurazione, ovvero una valutazione, non sono direttamente misurabili: è necessario ricorrere all'uso di **indicatori**. Essi sono concetti più semplici, specifici, traducibili in termini osservativi, legati a concetti generali attraverso un legame di rappresentanza semantica⁵⁹. È però chiaro che il legame tra concetto ed indicatore è parziale. Un concetto generale non può essere esaurito da un solo indicatore specifico per cui risulta necessario ricorrere a più indicatori per rilevare operativamente un unico concetto; dall'altra parte un indicatore può essere connesso con più concetti⁶⁰. La relazione che lega un indicatore ad un concetto complesso non è quindi mai univoca: si rende necessario operare una scelta tra una molteplicità di indicatori, scelta che deve essere basata su considerazioni scientifiche ma spesso vincolata dalla disponibilità dei dati nonché frutto di selezioni arbitrarie operate secondo la soggettiva sensibilità del ricercatore. In poche parole, quando siamo di fronte ad un concetto che non è misurabile direttamente, aggiriamo l'ostacolo cercando alcune variabili osservabili che hanno una forte relazione semantica con il concetto che vogliamo misurare o con una parte di esso. Queste variabili, opportunamente trasformate, svolgono il ruolo di indicatori⁶¹.

J., (2006), *Introduzione alla statistica per le scienze sociali*, Mc Graw-Hill, Milano, pagg. 198-199.

⁵⁸ Cfr. Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari, pag. 9.

⁵⁹ Cfr. Corbetta P. G., (1999), *Metodologia e tecniche della ricerca sociale*, Il Mulino, Bologna, pag. 115.

⁶⁰ Cfr. Corbetta P. G., (1999), *Metodologia e tecniche della ricerca sociale*, Il Mulino, Bologna, pagg. 115-166.

⁶¹ Cfr. Aiello F., Librizzi L. (2006), *Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo*, in Diamond I., Jefferies J., (2006), *Introduzione alla statistica per le scienze sociali*, Mc Graw-Hill, Milano, pag. 202.

2.3.1 Le dimensioni di un concetto

Un primo aspetto da prendere in esame è la dimensionalità del fenomeno sociale in valutazione: essa non rappresenta una caratteristica intrinseca del concetto ma dipende dall'obiettivo che si persegue, dalla definizione e dagli aspetti che si prendono in considerazione⁶². Se il concetto ricade lungo una dimensione (o *continuum*) individuabile e spiegabile dove può assumere diverse modalità, si parla di concetto **unidimensionale** mentre si parla di concetto **multidimensionale** quando è necessario ricorrere a più sottoconcetti per averne una rappresentazione completa (o il più completa possibile).

Nel caso in cui il concetto sia multidimensionale, la costruzione degli indicatori come traduzione empirica del concetto non direttamente osservabile si realizza in diverse fasi. Anzitutto è necessario definire il concetto in modo chiaro ed univoco, cercando di catturare il suo contenuto semantico. Naturalmente tale operazione risente di quelli che sono gli obiettivi della ricerca, come è stato precedentemente affermato. Usualmente si tratta di una fase affrontata da studiosi ed esperti della disciplina a cui il concetto in analisi afferisce; deve essere posta molta cura nella definizione del concetto in esame perché caratterizzazioni errate possono condurre a misurare un concetto diverso da quello in studio. Talvolta si definisce il concetto a partire da un insieme di indicatori in relazione con il concetto analizzato, individuati a partire dalla disponibilità di dati: in tal modo l'indicatore assume il duplice ruolo di definizione e di misura. Si tratta di "scappatoie" percorse per limitare le difficoltà che il ricercatore incontra nell'individuare una corretta definizione del concetto. È quindi necessario scomporre il concetto in componenti, meglio definite come **dimensioni**, più agevolmente misurabili ed osservabili.

Per determinare tali dimensioni si può far riferimento a due approcci. Nel primo approccio, di tipo teorico, la ricerca delle componenti avviene attraverso un'analisi dello spazio semantico del concetto. È auspicabile che la scelta delle componenti sia fatta in modo tale da coprire l'intero spazio semantico, evitando la sovrapposizione tra una componente e l'altra, ovvero la ridondanza di informazioni su un medesimo aspetto. Il secondo approccio, di tipo empirico, consiste nell'utilizzare le osservazioni su un insieme di variabili – che sono legate sotto l'aspetto semantico al concetto in esame – ed indagare sulla struttura multidimensionale del concetto attraverso l'applicazione di tecniche di

⁶² Idem.

statistica multivariata, come, per esempio, l'analisi delle componenti principali o l'analisi fattoriale.⁶³

2.3.2 Proprietà di un indicatore

In questo paragrafo vorremmo indicare alcune proprietà desiderabili per gli indicatori che vengono utilizzati nelle ricerche su concetti complessi, in particolar modo in campo sociale. La verifica della sussistenza di queste proprietà e requisiti mira a vagliare la qualità degli indicatori proposti, accertando che siano idonei a misurare ed interpretare il fenomeno che si intende valutare⁶⁴.

La **validità** di uno strumento fa riferimento alla capacità di un indicatore di rilevare effettivamente la dimensione del concetto a cui si riferisce. Generalmente per valutare la validità di uno strumento di misurazione si considerano diversi suoi aspetti che vengono testati attraverso apposite procedure di convalida⁶⁵. La validità di contenuto, che si colloca su un piano teorico, indica quanto l'indicatore o gli indicatori scelti coprono l'intero spazio semantico del concetto. Un secondo tipo di validità è definita come validità per criterio e mira a convalidare la corrispondenza tra l'indicatore e criteri esterni ad esso ma correlati con il concetto. Si distinguono in essa la validità predittiva, come capacità dello strumento di prevedere l'evoluzione futura della condizione del fenomeno, e la validità concomitante che prevede la correlazione dell'indicatore considerato con altri alternativi rilevati nello stesso momento temporale. La validità di costrutto intende invece verificare la coerenza tra la valutazione effettuata e le ipotesi teoriche formulate: in altre parole viene esaminata la rispondenza di un indicatore alle assunzioni teoriche in termini di relazioni con le altre variabili. Sono state inoltre individuate ulteriori tipologie di validità:

- la validità interna come grado di certezza dei risultati della ricerca e, in particolare, della relazione causa-effetto che lega la variabile dipendente alla variabile indipendente;

⁶³ Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano, pagg. 202-203.

⁶⁴ Cfr. Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari, pag. 23.

⁶⁵ Per approfondire le procedure di convalida: Corbetta P. G., (1999), Metodologia e tecniche della ricerca sociale, Il Mulino, Bologna.

- la validità statistica che valuta il grado di probabilità che i risultati ottenuti nel campione non siano causali ma riflettano il vero – ed ignoto – andamento del fenomeno nella popolazione di riferimento;
- la validità esterna misura il grado di applicabilità dei risultati a soggetti e a contesti diversi da quelli considerati nella ricerca;
- la validità ecologica, parte della validità esterna, come grado di generalizzabilità dei risultati alla vita reale.

La seconda importante proprietà auspicabile per un indicatore è l'**affidabilità** o **attendibilità** e ha a che fare con la riproducibilità dei risultati segnalando quanto gli strumenti di misura sono in grado di produrre i medesimi risultati in valutazioni ripetute sotto condizioni equivalenti.

2.3.3 Gli indicatori sociali

Il termine "indicatore sociale" è largamente e comunemente diffuso ma anche abusato: viene utilizzato convenzionalmente per definire qualunque tipo di statistica sociale senza alcuno sforzo che miri a coniugare insieme singole statistiche sociali ad un sistema, ad un modello complessivo di indicatori orientato all'interpretazione dei fenomeni o alla loro conoscenza e descrizione.⁶⁶

La locuzione di indicatore sociale nasce attorno agli anni Sessanta per analogia a quella di indicatore economico utilizzata già da alcuni decenni da economisti e statistici. Non trova però una definizione univoca: diversi autori sono giunti a differenti definizioni, scaturite quindi da punti di vista diversi, connessi al ruolo ed agli scopi che vengono assegnati all'indicatore nonché legati alla natura del problema da studiare. Si tratta di uno strumento che valuta il livello di un fenomeno sociale complesso non direttamente misurabile attraverso la misura diretta di fenomeni che hanno un alto contenuto semantico in comune con il concetto che si vuole misurare. È dunque un'informazione quantitativa su un fatto sociale; il livello assunto dall'indicatore e le sue variazioni nel tempo devono essere in grado di indicare non solo la situazione in cui la collettività si trova ma anche la

⁶⁶ Cfr. Zajczyk F. (1997), Il mondo degli indicatori sociali, NIS, Roma, pag. 25.

direzione verso cui si muove. L'indicatore quindi è messo in relazione con un fenomeno complesso che può anche non essere valutato in un'unica dimensione.⁶⁷

2.3.3.1 L'informazione statistica per la valutazione di fenomeni sociali⁶⁸

La valutazione del livello dei fenomeni sociali, tramite misure di fenomeni correlati, occupa, da tempo, sia sociologi che statistici, con un naturale divario di impostazione e di vedute.

Per i sociologi una teoria del sistema sociale dovrebbe rappresentare la "guida" alla scelta degli indicatori sociali idonei a mostrare la realtà e le relazioni significative nella società, nonché il "sistema" entro il quale collocare razionalmente le batterie di indicatori empirici raccolti. Mancando però un unico e condiviso quadro teorico, viene meno l'esistenza di un riferimento per la creazione di indicatori validi in assoluto.

Per gli statistici, invece, la costruzione di indicatori sociali non ha per solo scopo la conoscenza della realtà ma intende anche permettere un controllo efficace su di essa. Questo pragmatismo ha prodotto da un lato una poderosa quantità di dati, dall'altro ha contribuito ad accrescere la crisi degli indicatori a causa della nascita di rilevanti problemi sul piano statistico: assenza di una definizione comunemente accettata di indicatore sociale, mancanza di criteri collaudati di scelta degli indicatori, carenza di sperimentazione di procedimenti di sintesi e ponderazione delle batterie di indicatori sociali.

Vi è quindi, indubbiamente, una confusione terminologica e concettuale di fondo che fatica a trovare chiarimento in un quadro teorico costituito da punti di vista eterogenei e plurali tra loro non integrati.

Ma come e quando nasce l'esigenza di approfondire la conoscenza della realtà esaminando lo stato della società? Lo scenario è quello del periodo in cui i Paesi più sviluppati giungono alla fase di post-industrializzazione: essi non sono riusciti a riassorbire i fermenti dovuti al rapido mutamento sociale che l'industrializzazione, accanto all'avvio di un processo di modernizzazione, ha comportato. Si rende necessaria perciò una conoscenza sufficientemente dettagliata della consistenza e dell'evoluzione di fabbisogni ed aspirazioni individuali che richiede soddisfacenti interventi di carattere sociale da parte

⁶⁷ Cfr. Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari, pagg. 53-54.

⁶⁸ Per la stesura di questo ampio paragrafo sono state utilizzate le seguenti fonti:

- Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari.
- Zajczyk F. (1997), *Il mondo degli indicatori sociali*, NIS, Roma.
- Curatolo R. (1972), *Indicatori sociali*. Atti della XXVII Riunione Scientifica della SIS (Palermo, 29-31 maggio), vol. 1.
- Trivellato U. (1975), *Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine sociale sulla riuscita scolastica*, Fondazione "Emanuela Zancan", Padova.

delle istituzioni pubbliche. Ma per lo scarto, dato da difetto d'informazione, tra interventi ed effettivi bisogni, accompagnato allo scollamento tra enunciazioni di propositi ed effettive capacità realizzatrici da parte della pubblica amministrazione, scetticismo e scontento caratterizzano tale periodo rispetto alle capacità dello Stato di venire incontro a bisogni ed attese della popolazione. La crisi che accompagna tale momento storico induce, inoltre, nuovi problemi sia di tipo quantitativo (esplosione demografica, limiti delle risorse naturali, etc.) sia di carattere qualitativo (nuove aspirazioni tese a migliorare la qualità della vita). Dopo decenni di assenza di interventi nel sociale ci si è accorti che non bastava operare esclusivamente lungo la direzione economica: inizia infatti a vacillare l'equazione che eguagliava sviluppo economico a progresso sociale e nasce la necessità di valutare i costi che la collettività affronta per fruire dei benefici dello sviluppo economico. Si sviluppa la consapevolezza che il benessere non può essere soltanto funzione di variabili economiche ma dovrà essere concepito piuttosto quale funzione anche di quelle non economiche, recuperando, così dimensioni puramente sociali. Emerge quindi la necessità di disporre di nuovi indicatori del benessere quali misure statistiche in campo sociale.

Gli studi sugli indicatori sociali nascono attorno ai concetti di "benessere" e di "qualità della vita", termini che hanno costituito il punto di partenza di molte ricerche tese ad individuare una qualche metodologia che ne consentisse la misurazione ovvero la valutazione. Alla loro diffusione non ha però corrisposto una piena legittimazione scientifica, negando una precisa definizione sia sul piano concettuale sia su quello della sua operativizzazione dovuta all'assenza del necessario studio di un approccio teorico-metodologico che combinasse stili di vita, sistemi di valori e condizioni di vita degli individui: in altre parole la lacuna dell'impianto d'analisi ha avuto sede nella presenza di barriere disciplinari tra campi d'investigazione vicini. È infatti dall'integrazione di più approcci teorici ed empirici che può derivare anche un miglioramento nella comprensione dei fattori che determinano cambiamenti oggettivi. L'obiettivo, pertanto, diventa quello di individuare nuovi segmenti concettuali, di sperimentare e mettere a punto nuove strategie di misurazione per poter operationalizzare indicatori sempre più consoni ed efficaci per la lettura di processi sempre più complessi, in una prospettiva che deve tener conto in misura crescente della comparabilità internazionale tra paesi.

Ma vogliamo ora brevemente illustrare il processo di analisi e di sviluppo degli indicatori sociali che nasce e si diffonde, attraverso quello che è stato denominato "movimento degli indicatori sociali", secondo tre grandi momenti. Dagli anni Sessanta alla fine degli anni

Settanta nasce ed ha i primi sviluppi la ricerca sugli indicatori sociali, i quali vedono ampia diffusione sia in ambienti pubblici che nel mondo accademico. Succede quindi, da metà degli anni Settanta a metà degli anni Ottanta, un periodo definito da alcuni – e soprattutto negli Stati Uniti e in Europa – come fasi di crisi, disillusione e delusione del movimento mentre secondo altri studiosi si tratta di un momento di consolidamento e maturazione del “movimento degli indicatori sociali”. Dagli anni Ottanta, infine, è partita una rivitalizzazione della ricerca sugli indicatori sociali, particolarmente tesa a trovare una comparabilità internazionale delle statistiche sociali esistenti. Per tutti gli anni Ottanta, quindi, il movimento degli indicatori sociali non ha avuto una particolare evoluzione – dal punto di vista dell’elaborazione concettuale –; tuttavia ha visto in molti paesi industrializzati un’esplosione di rapporti e compendi di statistiche sociali volte ad analizzare lo stato di benessere socio-economico della popolazione al fine di contribuire a meglio definire i programmi di politica sociale. A metà degli anni Novanta, dopo una fase di entusiasmo e di forse eccessive aspettative, si è giunti a vivere un periodo di ricorso acritico agli indicatori sociali: l’ambizioso progetto del “movimento degli indicatori sociali” di utilizzo degli stessi ai fini diretti delle decisioni politiche e della programmazione sociale ha subito, nel tempo, un ridimensionamento, portando ad attribuire a questi strumenti statistici una funzione sostanzialmente solo conoscitiva ma empiricamente più realizzabile. È in questo periodo storico che cresce la consapevolezza dell’importanza di “conoscere”, nonostante risulti sempre più evidente che la relazione tra conoscenza e decisione politica non è di tipo deterministico. Nonostante venga attribuita al dato empirico, quantitativo, una capacità quasi maieutica⁶⁹ – che in realtà non gli appartiene –, diventa chiaro che un sistema di indicatori sociali, anche se ben costruito, di per sé non può contribuire alla razionalizzazione dei processi politici, né a stabilirne obiettivi e proprietà, non è in grado di valutare i programmi politici, né di costruire un sistema di allertamento sociale preventivo. La conoscenza dei fenomeni diventa l’obiettivo centrale di ogni azione sociale⁷⁰ ed assume il ruolo di componente essenziale della politica e dell’amministrazione sociale: a questo scopo conoscono un forte sviluppo i rapporti e le inchieste sociali, strumenti che aiutano a leggere le emergenze sociali ed economiche. In questo senso aumentano anche gli sforzi delle istituzioni statistiche e di ricerca verso il miglioramento delle basi di dati necessarie

⁶⁹ Per maieutica si intende la capacità di mettere in grado un soggetto di acquisire chiara coscienza delle conoscenze che si formano nella sua mente.

⁷⁰ A tal proposito, Curatolo esprime ragionevoli perplessità sottolineando la mancanza del consueto meccanismo consapevolezza-azione dovuto ad una consapevolezza tutt’altro che acquisita sullo sviluppo dei processi sociali per cui vediamo i fatti ma non sappiamo interpretarli.

alla costruzione degli indicatori sociali nonché degli aspetti concettuali e definatori ad essi connessi. Parte quindi un processo di democratizzazione dell'informazione statistica che si traduce in trasparenza ed ampia diffusione della stessa: l'informazione, sempre più analitica, diventa strumento di controllo e valutazione delle decisioni politiche tra Stati e all'interno degli Stati stessi.

In Italia l'attenzione per gli indicatori sociali si è sviluppata considerevolmente più tardi che in altri Paesi, sia da parte del mondo accademico che da parte dell'apparato pubblico di produzione dei dati (Istat). In particolare, i primi elaborati sugli indicatori sociali pubblicati in Italia sono i seguenti:

- nel 1972 Curatolo cura il lavoro intitolato *Indicatori sociali*, presentato alla XXVII Riunione Scientifica della SIS (Società Italiana di Statistica);
- nel 1975 l'Istat pubblica il primo volume *Statistiche sociali*.⁷¹

Relativamente allo sviluppo passato, presente ed auspicabile per gli indicatori sociali, riteniamo utile riportare alcune sottolineature di importanti aspetti su cui porre attenzione. Negli anni Ottanta, gli indicatori sociali hanno causato una generale delusione determinata dal fatto di non essere riusciti né a disegnare né a prevedere i mutamenti sociali di fine anni Settanta. Muovendo da questo punto, risulta evidente l'importanza di poter osservare l'andamento dinamico, diacronico, dei diversi fenomeni sociali, non solo sotto il profilo del cambiamento di opinioni e percezioni soggettive, ma sempre più rispetto al percorso oggettivo che determina, nel tempo, il costituirsi di certe condizioni sociali. Ma la continuità e la regolarità nel tempo delle rilevazioni sono una carenza del panorama della ricerca sociale legata agli aspetti soggettivi, soprattutto in Italia, nonostante gli sforzi dell'Istat. Si rende pertanto necessario migliorare ed ampliare le basi di dati e di indicatori per permettere ricostruzioni di trend più approfondite ed efficaci, unitamente alla possibilità di utilizzare modelli di simulazione e previsione. Si tratta, in qualche modo, di una "richiesta" implicita derivante dalla complessità dei fenomeni sociali, economici, politici, territoriali e dalla necessità di tenere sotto controllo le diverse tipologie di emergenze sociali: la conoscenza sui più probabili andamenti futuri dei fenomeni, al modificarsi di determinati fattori, sarà infatti sempre più urgente, e gli indicatori sociali dovranno assumere sempre più il ruolo di strumenti di modelli sociali e analisi revisionali.

⁷¹ Il volume è poi apparso nell'anno 1981.

2.3.3.2 La nascita degli indicatori sociali

Per cogliere la poliedricità degli aspetti attinenti agli indicatori sociali, è importante inquadrare questi ultimi nella loro dimensione spaziale e temporale.

Il primo impulso a sostegno di un nuovo approccio allo studio statistico dello stato di una società che approfondisca la conoscenza del sociale, accanto ai più tradizionali settori demografico ed economico, fu dato, nel 1962, dalla NASA allorché incaricò un gruppo di ricercatori di studiare le conseguenze che i programmi spaziali, nonché le innovazioni tecnologiche indotte da questi, avrebbero prodotto sulla comunità statunitense. In quella occasione, uno dei responsabili della ricerca, Raymond Bauer, avvertì per primo l'inadeguatezza a presentare l'immagine della nazione sia attraverso gli indicatori e sia, più in generale, con le informazioni e i dati quantitativi tradizionali, di natura prevalentemente economica (reddito, investimenti, consumi, etc.). Si afferma così un nuovo obiettivo per l'intero sistema statistico nazionale e, in particolare, per le statistiche sociali: fornire funzioni conoscitive, previsive e valutative (tanto per scopo scientifici, quanto per scopi politico-amministrativi) delle condizioni di vita della popolazione e della ineguaglianza – e quindi del "malessere" – sociale. Gli indicatori rappresentano, perciò, lo strumento di lavoro più adatto a perseguire questi scopi in quanto, stando a Bauer, essi sono "statistiche, serie statistiche ed ogni altra forma di dati che ci mettono in condizione di stimare a che punto siamo e dove stiamo andando relativamente ai nostri valori ed obiettivi, e di valutare specifici programmi e determinarne l'impatto". In seguito ai risultati dell'esperienza americana, anche altri Paesi hanno manifestato l'esigenza di una più approfondita conoscenza del sociale rilevando, perciò, statistiche sempre più dettagliate in detto campo, nel tentativo di ampliare il quadro e migliorare la qualità delle informazioni per mezzo, ad esempio, di indagini più frequenti che non fossero solo semplici sottoprodotti di altra attività informativa quali le statistiche economiche e demografiche. Nonostante, però, le esperienze acquisite in queste ultime discipline, non si sono potute evitare – nelle rilevazioni delle statistiche sociali – frammentazione di dati, disparità di sviluppo e confusione terminologica. Dalla medesima istanza però scaturirono anche tentativi di concepire *Sistemi integrati di statistiche demografiche, sociali ed economiche*, di costruire *Sistemi di contabilità sociale (Social accounting)*, di compilare *Rapporti e Relazioni sociali (Social Reports)*, che possono tutti essere riconosciuti quali schemi con cui assegnare una semantica agli indicatori sociali. Queste esperienze tendono, perciò, tutte ad integrare in un solo quadro le statistiche e gli indicatori sociali, apportando indiscutibili

benefici. Si avverte subito, allora, la necessità di moltiplicare gli sforzi per definire un modello, ancor più vasto di quello impegnato nei sistemi dei conti economici nazionale, capace di rappresentare l'intera realtà sociale.⁷²

Un contributo in tale direzione è stato offerto da Stone⁷³ che ipotizza un grande sistema di contabilità socio-demografica per seguire gli individui durante il corso della loro vita: tale sistema consente di raggruppare, attraverso il ricorso a matrici, i dati riguardanti i principali aspetti della vita e delle condizioni sociali della popolazione individuando anche i legami tra condizioni di vita demografica e sociale. Si tratta però di un modello di difficile realizzazione pratica perché, se da un lato diventa agevole costruire il corso di vita a livello individuale, scomposizione e ricomposizione di determinati caratteri socio-demografici a livello aggregato risultano molto limitate. Permane inoltre un certo grado di confusione terminologica e concettuale che non aiuta a spiegare natura e ruolo degli indicatori sociali nell'informazione sullo stato e sull'evoluzione della società.

Molti diversi approcci sono diventati interpretazione di altrettanti studi teorici ed applicazione pratiche tese a risolvere il problema relativo all'uso degli indicatori sociali. Tra questi vogliamo qui dar particolare rilievo all'approccio elaborato da un gruppo di quarantuno scienziati del Department of Health, Education and Welfare in *Toward a Social Report*⁷⁴. La definizione riportata in questo lavoro intende l'indicatore come "una statistica di diretto interesse normativo, che favorisce un giudizio conciso, globale ed equilibrato sullo stato degli aspetti più importanti di una società". L'indicatore sociale assume quindi il valore di dato quantitativo di sintesi di determinati aspetti della vita sociale. Si tratta però di un tentativo debole di concettualizzazione degli indicatori poiché presume di prescindere da un'illustrazione delle interrelazioni tra fenomeni osservati, limitando la definizione di indicatore sociale a componente di modelli del sistema sociale o di qualche particolare segmento o processo del sistema sociale. Risulta pertanto preferibile, seguendo Land⁷⁵, intendere il modello sociale come rappresentazione schematica delle relazioni fra le variabili del processo sociale: è poi dal nesso concettualmente stabilito ed empiricamente verificato con un modello del sistema sociale che l'indicatore sociale trae il proprio valore informativo.⁷⁶

⁷² Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari, pagg. 61-62.

⁷³ Stone R., (1970), *An Integrated System of Demographic, Manpower and Social Statistics and its Links with the System of National Economic Accounts*, in "Memorandum E/CN", United Nations Economic and Social Council UNESCO.

⁷⁴ U.S. Department of Health, Education and Welfare, *Toward a Social Report*, 1969.

⁷⁵ Land K. C. (1971), *On the Definition of Social Indicators*, *The American Sociologist*, n. 6, pag. 323.

⁷⁶ Cfr. Trivellato U. (1975), *Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine*

Molti autori si basano e si sono basati sulla definizione di indicatore sociale elaborata da Land. Tra questi rientra anche Trivellato, che ha ripreso tale definizione nel suo saggio sul sistema scolastico; si tratta di un lavoro che trova immediata equivalenza nel sistema universitario: diventa cioè naturale concepire ed interpretare gli indicatori relativi all'università nel contesto dei modelli di funzionamento del sistema universitario. Inoltre trova conferma anche quanto detto rispetto agli adattamenti che gli indicatori debbono subire: vincoli imposti dalla documentazione statistica disponibile e carenza di una teoria del processo di formazione universitaria operativamente utilizzabile producono necessariamente delle semplificazioni nella struttura degli indicatori utilizzati per valutare l'università.⁷⁷

2.3.3.3 L'utilizzo degli indicatori sociali⁷⁸

Dato il carattere flessibile della definizione del concetto di indicatore sociale, esistono diversi approcci relativi agli impieghi che essi hanno per penetrare la realtà sociale. In questo elaborato esponiamo tre grandi scuole che definiscono tre diversi scopi di utilizzo degli indicatori sociali: la scuola normativa, quella oggettivista e la scuola soggettivista.

La scuola normativa si basa sulla necessità di avere una conoscenza della società di tipo operativo ovvero che trovi traduzione in azioni concrete relative agli interventi nel sistema sociale. Gli indicatori sociali normativi mirano quindi a fissare scopi e priorità della politica sociale, a costruire una contabilità sociale, a valutare l'efficacia dei programmi politici e, in generale, a razionalizzare la politica sociale.

Contrari a questa impostazione sono i ricercatori appartenenti alla scuola oggettivista che non concordano sull'utilizzo di strumenti imperfetti e non scientifici, quali gli indicatori sociali, per dedurre gli obiettivi di una politica e per valutare gli effetti di determinati interventi. Tali utilizzi sono infatti inappropriati per l'incapacità degli indicatori sociali di costruire un'immagine di insieme della realtà sociale e per il non dimostrabile nesso di causalità tra programmi e risultati⁷⁹ (non è infatti statisticamente dimostrabile che siano i programmi, piuttosto che altre variabili incontrollate, a determinare i risultati misurati dagli indicatori). La scuola oggettivista sollecita quindi la pura costruzione di statistiche ed

sociale sulla riuscita scolastica, Fondazione "Emanuela Zancan", Padova, pagg. 26-27.

⁷⁷ Cfr. Trivellato U. (1975), Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine sociale sulla riuscita scolastica, Fondazione "Emanuela Zancan", Padova, pagg. 28-29.

⁷⁸ Cfr. Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari, pagg. 73-80.

⁷⁹ Non è infatti statisticamente dimostrabile che siano i programmi, piuttosto che variabili non controllate, a determinare i risultati misurati dagli indicatori.

indicatori sociali unitamente alla definizione di modelli di istituzioni e processi sociali che fungano da schemi di riferimento per un'adeguata analisi della sfera sociale, legando logicamente gli indicatori in un modello di sistema sociale.

L'emergere del bisogno di valutazione della condizione dell'uomo, non solo da un punto di vista oggettivo ma tenendo conto delle sue esperienze e delle sue esigenze, conduce all'elaborazione di un approccio comportamentale che caratterizza la terza scuola, quella soggettivista.

Considerati i diversi approcci, è possibile individuare le finalità essenziali che gli indicatori sociali si prefiggono:

- essere di supporto alla programmazione degli interventi in campo sociale;
- consentire il controllo del sistema sociale tramite "segnali rapidi" di allarme;

La realizzazione di queste finalità si esplica attraverso:

- la misura dei mutamenti sociali;
- la conoscenza delle relazioni esistenti nel sistema sociale;
- la messa in evidenza dei problemi sociali più importanti perché diventino le priorità e gli obiettivi delle politiche sociali.

Crediamo però sia identificabile un ulteriore compito degli indicatori sociali, in funzione con la loro nascita e con la loro stessa natura: fornire informazioni – facilmente accessibili al corpo sociale – per una partecipazione democratica alle decisioni politiche e per un controllo sull'operato degli organi di governo.

2.3.3.4 Alcune considerazioni conclusive

Gli indicatori sociali nascono quindi per rispondere alla crescente esigenza di ottenere informazioni statistiche in campo sociale. Il loro obiettivo è quello di essere degli strumenti di misura dei fenomeni sociali che siano di supporto alla conoscenza ed alla previsione dei mutamenti del complesso sistema sociale. Gli approcci al loro utilizzo sono diversificati come sono diverse le loro definizioni, la loro natura e gli obiettivi prefissati dai ricercatori che ne fanno uso. La teoria alla base della costruzione degli indicatori sociali non è univocamente definita, né generalmente condivisa: da questo deriva quindi una pluralità di impianti metodologici che fatica a definire un'unica prospettiva di utilizzazione.

In questa direzione gioca – e gioca a favore – l'ampliamento delle fonti da cui provengono gli indicatori. Essendosi ampliato il ruolo dei produttori ufficiali ed istituzionali di dati (primo fra tutti l'Istat), la situazione di monopolio esistente ha ceduto il passo ad un

contesto di pluralità di fonti che rispondono a obiettivi diversi, con criteri e procedure non coordinate. Questo ha ingenerato, e tuttora ingenera, confusione di ruoli, sovrapposizione di compiti e duplicazione di risultati rendendo necessaria, da parte del ricercare, l'acquisizione di informazioni sulle fonti disponibili per supplire alla difficoltà di orientamento in quello che è stato definito "arcipelago delle statistiche".⁸⁰

2.3.4 La sintesi degli indicatori

2.3.4.1 Indicatori semplici e indicatori composti

Per valutare un fenomeno sociale viene raccolta una collezione di dati grezzi (da questionari, da registri, etc.) non direttamente utilizzabili che vengono elaborati prima di diventare indicatori. Si tratta di un processo che sottende a specifici obiettivi conoscitivi: in funzione di questi ultimi si orientano le scelte arbitrarie che caratterizzano il percorso di costruzione degli indicatori. Gli indicatori sociali sono quindi qualcosa in più dei semplici dati: mentre i dati, da soli, non forniscono nessuna indicazione, gli indicatori costituiscono un ponte ipotetico tra ciò che è osservabile e misurabile ed i concetti di cui si vuole fornire una misura.⁸¹

Gli indicatori sociali vengono classificati secondo diversi criteri. In base all'origine dell'informazione si identificano indicatori oggettivi (relativi ad eventi direttamente rilevabili) o soggettivi (costruiti tramite strumenti di rilevazione quali un questionario); secondo gli scopi di utilizzo gli indicatori vengono classificati come descrittivi (per esplicitare lo stato dei fenomeni), esplicativi (interpretativi della realtà sociale), normativi (orientati su problemi da risolvere e di guida per programmare interventi di politica sociale) e predittivi (forniscono indicazione della tendenza evolutiva del fenomeno in esame).⁸²

Un'importante distinzione è quella che scinde il mondo degli indicatori in indicatori semplici e composti. L'**indicatore semplice** o elementare riferisce a fenomeni "singoli" ovvero ad una delle dimensioni, immediatamente quantificabile, nelle quali è stato scomposto un fenomeno sociale complesso. In altre parole è una variabile in grado di

⁸⁰ Cfr. Zajczyk F. (1997), Il mondo degli indicatori sociali, NIS, Roma, pag. 51.

⁸¹ Cfr. Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano, pag. 206.

⁸² Cfr. Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari, pagg. 80-85.

misurare in termini appropriati le variazioni di un determinato oggetto o parte di un concetto. L'aggregazione e l'eventuale ponderazione di più indicatori semplici forma quello che è definito **indicatore composto**.

2.3.4.2 Generalità sulla sintesi degli indicatori

Se un indicatore sociale è costruito tramite un'opportuna aggregazione di indicatori semplici per misurare un determinato concetto, diventa immediato pensare che un fenomeno complesso viene rappresentato tramite una sintesi di indicatori elementari. Secondo alcuni studiosi ed anche secondo alcuni organismi internazionali, è però preferibile utilizzare una lista di indicatori elementari per interpretare un fenomeno sociale perché la sintesi degli indicatori fa perdere di vista la struttura organica del fenomeno in esame.

Il problema della sintesi degli indicatori consiste sia nell'opportunità o meno di effettuarla sia nell'aspetto tecnico di individuazione della metodologia di aggregazione da applicare. La scelta di sintetizzare gli indicatori elementari in un indicatore composto dipende dall'obiettivo dell'analisi: per il monitoraggio e la programmazione di interventi di politica sociale – ovvero per una conoscenza di tipo strumentale – è opportuno avere indicatori disaggregati; se invece lo scopo è di tipo descrittivo o comparativo, sia nello tempo che nello spazio, è preferibile utilizzare indicatori composti.

Il valore di sintesi di un indicatore composto non possiede il contenuto informativo dei singoli indicatori perché può essere ottenuto da molteplici combinazioni di indicatori elementari che, a loro volta, possono subire diversificati processi di trasformazione, ponderazione ed aggregazione.

Gli indicatori semplici sono legati ad aspetti parziali del concetto ovvero alle sue dimensioni: se l'obiettivo è quindi quello di ottenere una misura indiretta complessiva del concetto, bisogna aggregare le informazioni quantitative espresse dai singoli indicatori semplici. In ogni caso, non è possibile pretendere una perfetta corrispondenza tra indicatore composto e concetto complesso: possono esservi aspetti del concetto non facilmente esprimibili in termini quantitativi o variabili difficilmente rilevabili. Tuttavia, se è vero che indicare la complessità di un fenomeno con un numero è una forzatura, è vero anche che è questo il modo più ragionevole per semplificare una realtà complessa e tradurla "pragmaticamente" perché possa essere valutata, confrontata e perché se ne possano rilevare le variazioni nel tempo e nello spazio. È inoltre importante non

dimenticare che il significato di un numero dipende in modo cruciale da come è stato ottenuto ed una interpretazione trasparente dello stesso non può prescindere dalla conoscenza delle scelte effettuate lungo tutto il percorso che va dalla definizione del concetto alla costruzione dell'indicatore composto.⁸³

Va comunque sempre tenuto in considerazione il reale obiettivo della sintesi di indicatori che resta quello di valutare unitariamente un fenomeno complesso, composto di una molteplicità di dimensioni diverse, ma concettualmente unitario⁸⁴.

2.4 GLI INDICATORI COMPOSTI

La scomposizione di un concetto avviene a livello astratto e porta all'identificazione di alcune dimensioni di cui il concetto in esame si compone; l'aggregazione degli indicatori semplici per formare un indicatore composto si realizza invece a livello statistico-matematico.

2.4.1 Il processo di costruzione di un indicatore composto

2.4.1.1 Uno schema sintetico⁸⁵

La costruzione di un qualsiasi indicatore composto deve essere basata su metodologie solide che siano facilmente comprensibili anche a lettori non esperti. Diversi sono i passi del processo di costruzione degli indicatori composti. Un sintetico schema generale individua i passaggi di seguito illustrati.

Scelta del fenomeno da misurare. È necessario capire se questo trae beneficio dall'uso di indicatori composti.

Selezione degli indicatori elementari. È necessario selezionare quegli indicatori che siano realmente rilevanti rispetto al fenomeno che si intende misurare; non è comunque ipotizzabile completa obiettività nella selezione.

⁸³ Cfr. Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano, pagg. 212-213.

⁸⁴ Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari, pag. 120.

⁸⁵ Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).

Valutazione della qualità dei dati. È indispensabile un'elevata qualità dei dati per gli indicatori elementari: in caso contrario l'analista deve decidere se eventualmente eliminare i dati poco affidabili o se deve definire un metodo di ri-costruzione dei dati mancanti. In quest'ultimo caso possono essere applicate molte tecniche alternative (sostituzione con la media dei valori, uso di risultati basati sulla correlazione, utilizzo delle serie temporali, etc.): è comunque necessario poi valutare come la selezione del metodo abbia effetto sui risultati finali.

Valutare la relazione tra gli indicatori elementari. Si tratta di un prerequisito dell'analisi da verificare in fase preliminare; particolari tecniche di analisi⁸⁶ forniscono direttamente una misura della relazione tra indicatori.

Normalizzazione e pesatura degli indicatori. Esistono molti metodi in letteratura per normalizzare– ovvero applicare trasformazioni che consentano di ottenere dati omogenei – e pesare gli indicatori: la loro selezione dipende dai dati stessi e dal ricercatore.

Test di robustezza e sensibilità. Naturalmente la scelta del sistema dei pesi e la selezione degli indicatori influenza i risultati. È comunque importante verificare il grado di sensibilità dei metodi usati rispetto ai risultati per evitare di fornire graduatorie non robuste sulle quali basare forti messaggi o importanti interventi di politica sociale che siano sensibili a minimi cambiamenti nella costruzione degli indicatori composti. Per questo viene suggerito di visualizzare il valore dell'indicatore composto nella forma di intervallo di confidenza.

2.4.1.2 Uno schema analitico⁸⁷

Nello scorso 2005 alcuni autori hanno prodotto una sorta di manuale per la costruzione di indicatori composti allo scopo di fornire una serie di raccomandazioni su come disegnare, sviluppare e divulgare un indicatore composto. Questo perché le questioni metodologiche necessitano di essere avviate in modo trasparente prima della costruzione e dell'uso degli indicatori composti per evitare la manipolazione dei dati e la loro errata presentazione. Per guidare quindi il costruttore e l'utilizzatore degli indicatori composti, è necessario in questa sede presentare i passi da seguire per la loro costruzione. Si tratta di una "sequenza ideale" di dieci passi, ciascuno dei quali è estremamente importante; è

⁸⁶ Metodi come l'analisi delle componenti principali, come sarà più avanti approfondito, consentono di valutare l'entità della relazione sussistente tra gli indicatori elementari candidati alla costruzione di indicatori composti.

⁸⁷ Per la stesura di questo paragrafo sono stati utilizzati i due seguenti documenti:

- Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).

però altrettanto importante la coerenza dell'intero processo: le scelte compiute in un passaggio, infatti, possono avere importanti implicazioni nei passi successivi.

DEFINIZIONE DELLA STRUTTURA TEORICA

Il punto di partenza è lo sviluppo di una solida struttura teorica che fa da base alla selezione ed alla combinazione degli indicatori semplici secondo un principio di pertinenza.

In questo primo passo devono essere eseguite le seguenti operazioni:

- definizione puntuale del concetto da misurare
- determinazione delle dimensioni del concetto, qualora esso sia multidimensionale; le dimensioni che non sono statisticamente indipendenti ed i legami tra esse esistenti dovrebbero essere descritti teoricamente ed empiricamente;
- identificazione dei criteri di selezioni degli indicatori elementari.

SELEZIONE DEI DATI

Gli indicatori devono essere selezionati sulla base della loro analitica stabilità, misurabilità, copertura territoriale, rilevanza con il fenomeno in misurazione e secondo la relazione che tra essi sussiste. Nel caso i dati siano scarsi è possibile utilizzare variabili proxy⁸⁸.

ANALISI MULTIVARIATA

Un ulteriore passo nel processo di costruzione degli indicatori composti consiste nell'effettuare un'analisi esplorativa sui dati tesa ad investigarne la loro struttura generale, a valutare la loro adeguatezza per giustificare le scelte metodologiche che il ricercatore adotterà nel seguito (ad esempio nelle fasi di pesatura ed aggregazione). Gli indicatori sono talvolta selezionati in modo arbitrario con scarsa attenzione alla relazione tra essi esistente: gli strumenti della statistica multivariata permettono di eseguire un'analisi sui dati di natura esplorativa utile a valutare l'adeguatezza dei dati selezionati e la comprensione delle implicazioni delle scelte metodologiche compiute durante la costruzione degli indicatori composti. Si può procedere in questa fase raggruppando gli indicatori elementari oppure raccogliendo in gruppi le unità di analisi (le facoltà nel presente lavoro).

⁸⁸ Si definisce variabile proxy una variabile la cui misura serve per studiare fenomeni ad essa correlati e non direttamente misurabili.

Raggruppamento di indicatori semplici

L'analista deve anzitutto decidere se la struttura a dimensioni del concetto in esame è ben definita e se la collezione di indicatori disponibile è sufficiente o adatta a descrivere il fenomeno. Questa decisione può essere basata sull'opinione di esperti o sulla struttura statistica dei dati. Differenti approcci analitici, come l'analisi delle componenti principali, possono essere utilizzati per verificare che le dimensioni del fenomeno siano ben equilibrate nell'indicatore composto; nel caso non lo siano si rende necessaria una revisione degli indicatori elementari.

Lo scopo dell'analisi delle componenti principali (PCA) è quello di rivelare come variabili diverse cambiano in relazione alle altre e come sono tra loro associate. Questo si ottiene trasformando le variabili correlate in una nuova collezione di variabili incorrelate attraverso una matrice di covarianza o la sua forma standardizzata (la matrice di correlazione). L'analisi fattoriale (FA) è simile alla PCA ma è basata su un particolare modello statistico. Una via alternativa per esaminare il grado di correlazione tra una collezione di variabili consiste nell'utilizzo del coefficiente alfa di Cronbach, comune strumento di stima della consistenza interna degli item in una scala.

Raggruppamento di unità di analisi

L'analisi cluster è un altro strumento per classificare grandi quantità di informazioni in gruppi facilmente manipolabili. Viene utilizzata anche nello sviluppo di indicatori composti per raggruppare informazioni sulle unità di analisi basate sulla loro somiglianza rispetto a diversi indicatori elementari. Questo tipo di analisi può essere adoperata per una molteplicità di scopi: tra questi l'analisi cluster può trovare utilizzo come puro metodo statistico di aggregazione di indicatori o come strumento diagnostico per esplorare l'impatto delle scelte metodologiche compiute durante la costruzione dell'indicatore composto.

Quando il numero di variabili è elevato o quando è verosimile supporre che alcune di queste non contribuiscano ad identificare la struttura a gruppi esistente nell'insieme dei dati, modelli continui e discreti possono essere applicati sequenzialmente. I ricercatori frequentemente eseguono una PCA e successivamente applicano un algoritmo per il raggruppamento (clustering) sui punteggi delle prime poche componenti estratte, chiamando questo tipo di analisi "tandem".

Sono stati proposti vari metodi alternativi che combinano analisi cluster e ricerca di una rappresentazione con dimensionalità più bassa di quella inizialmente ipotizzata. Tra queste segnaliamo l'analisi fattoriale delle k medie che combina l'analisi cluster delle k medie⁸⁹ con aspetti della FA e della PCA.

Riportiamo nella tabella seguente i punti di forza e di debolezza dei metodi sopra presentati per fornire un ulteriore contributo esplicativo.

Tabella 1. Punti di forza e di debolezza dei principali metodi di analisi multivariata.

METODO	PUNTI DI FORZA	PUNTI DI DEBOLEZZA
Analisi delle componenti principali e analisi fattoriale	<ul style="list-style-type: none"> - può riassumere un insieme di indicatori preservando la massima proporzione possibile di variabilità dei dati originali - i punteggi fattoriali maggiori sono assegnati agli indicatori elementari che posseggono la maggior variabilità tra unità di analisi 	<ul style="list-style-type: none"> - la correlazione non necessariamente rappresenta la reale influenza degli indicatori sul fenomeno che viene misurato - sensibilità alle modifiche sui dati originali - sensibilità alla presenza di outliers - sensibilità ai problemi relativi ai piccoli campioni
Coefficiente alfa di Cronbach	<ul style="list-style-type: none"> - misura la consistenza interna all'insieme degli indicatori ovvero quanto bene descrivono un concetto unidimensionale 	<ul style="list-style-type: none"> - la correlazione non necessariamente rappresenta la reale influenza degli indicatori sul fenomeno che viene misurato - ha senso solo quando l'indicatore composto è una scala
Analisi cluster	<ul style="list-style-type: none"> - offre diversi modi per raggruppare le unità di analisi - aiuta ad esaminare la struttura dei dati 	<ul style="list-style-type: none"> - è uno strumento puramente descrittivo

A conclusione di questa terza fase del processo di costruzione di un indicatore composto, devono quindi essere stati identificati sottogruppi di indicatori o gruppi di unità di analisi statisticamente simili⁹⁰ tra loro.

IMPUTAZIONE DEI DATI MANCANTI

La mancanza di dati spesso è di ostacolo allo sviluppo di indicatori composti che siano robusti. L'assenza di dati può essere

⁸⁹ Il metodo di analisi cluster non gerarchico delle k medie si basa su tecniche di ricollocamento secondo le quali, una volta fissato a priori il numero di gruppi (cluster) ed assegnata una partizione iniziale casuale dei dati, questi vengono spostati da un cluster all'altro e riassegnati fino a soddisfare un criterio prefissato. Il metodo consente perciò di muovere i dati nei k cluster fissati secondo due criteri: minimizzare la variabilità tra gli elementi all'interno di un singolo cluster, raggruppare elementi simili, e massimizzare la variabilità tra gli altri cluster, separare gli elementi diversi in gruppi diversi.

(Boschetti M., Brivio P. A., Brogiolo G. P., Crosato A., Soffia E. (2003), Applicazioni GIS per l'archeologia: lo studio della distribuzione dei siti archeologici nel Garda meridionale, Edizioni all'Insegna del Giglio)

⁹⁰ Due o più unità si definiscono statisticamente simili quando, mediante opportuni test, può essere verificata l'ipotesi di uguaglianza tra le variabili che li caratterizzano (ad esempio, attraverso la media).

- *completamente casuale* quando i valori mancanti non dipendono dalla variabile di interesse o da altre variabili osservate facenti parte del dataset;
- *casuale* se non dipendono dalla variabile di interesse ma sono condizionati ad altre variabili;
- *non casuale* nel caso in cui la mancanza di dati dipenda dal valore degli stessi.

Il problema dell'analista risiede però nel fatto che, molto spesso, non ci sono informazioni di base per giudicare se l'assenza di alcuni valori sia casuale o sistematica per cui diventa difficile applicare meccanismi di imputazione di valori mancanti che possiedono come requisito la casualità.

I tre metodi generali per l'imputazione dei dati mancanti sono i seguenti:

- la *cancellazione del record* prevede l'eliminazione dall'analisi dell'unità di analisi per cui manca il dato; questa tecnica ignora la possibilità di sistematiche differenze tra le unità con dati completi ed incompleti e può quindi produrre stime distorte;
- *l'imputazione del valore mancante* attraverso la singola imputazione (il valore mancante può venire sostituito con la media, la mediana o la moda dei valori oppure attraverso metodi di regressione, hot-deck oppure cold-deck) oppure tramite imputazione multipla (ad esempio con l'utilizzo delle catene di Markov dell'algoritmo Monte Carlo⁹¹); questo approccio consente di minimizzare la distorsione.

Devono comunque essere fornite considerazioni appropriate rispetto all'approccio adottato per l'imputazione dei valori mancanti. I valori estremi dovrebbero essere esaminati nel caso in cui diventino valori discriminanti.

NORMALIZZAZIONE

Poiché spesso gli indicatori sono espressi in unità di misura diverse, è necessario normalizzare i dati per renderli omogenei e quindi comparabili.

Esistono numerosi metodi di normalizzazione: esamineremo le diverse tecniche proposte in letteratura successivamente. Vogliamo però qui sottolineare come la scelta del metodo di normalizzazione adatto necessiti di particolare attenzione: esso deve tenere in considerazione le proprietà dei dati come pure l'obiettivo dell'indicatore composto. Diversi

⁹¹ Un processo stocastico markoviano o processo di Markov è un processo stocastico (nella teoria della probabilità, una generalizzazione dell'idea di variabile casuale che prende valori in spazi più generali dei numeri reali) nel quale la probabilità di transizione che determina il passaggio da uno stato di sistema ad un altro stato dipende unicamente dallo stato di sistema immediatamente precedente. Nel caso in cui le probabilità di passaggio da uno stato all'altro siano stazionarie nel tempo – ovvero non dipendono da esso – si dà luogo ad una catena di Markov. (Kemeny J. G., Snell J. L. (1960), Finite Markov Chain, Princeton.)

metodi di normalizzazione producono diversi risultati: opportuni test di robustezza dovrebbero quindi valutare l'entità dell'impatto della tecnica scelta.

PONDERAZIONE ED AGGREGAZIONE

Centrale alla costruzione degli indicatori composti è la necessità di combinare insieme le diverse dimensioni del fenomeno in studio, con decisioni che riguardano sia l'assegnazione dei pesi, sia la procedura di aggregazione degli indicatori elementari. Si tratta, in realtà, di due fasi operativamente distinte nonostante la letteratura faticosi a scindere in modo chiaro ponderazione ed aggregazione: esse, infatti, hanno forti influenze reciproche sia in termini decisionali rispetto alla tecnica da adottare, sia in termini di risultati finali. Naturalmente le scelte, per quanto possibile, non dovrebbero assumere il carattere di arbitrarietà ma gli indicatori semplici dovrebbero essere pesati ed aggregati secondo la struttura teorica ad essi sottostante.

Trattandosi della parte centrale e più strettamente statistico-matematica del processo di costruzione degli indicatori composti, ampio sarà lo spazio dedicato alle tecniche di pesatura e di aggregazione. Ci limitiamo, in questo frangente, ad esporre qualche breve considerazione.

Esistono diverse tecniche per la determinazione dei pesi alcune basate su modelli statistici, altre su metodi partecipativi, altre ancora sul giudizio di esperti, i quali conoscono le priorità politiche e possiedono le conoscenze teoriche rappresentando il punto di vista di una molteplicità di *stakeholders*⁹². Spesso i pesi riflettono la qualità dei dati: pesi maggiori sono così assegnati a dati statisticamente più affidabili ovvero con basse percentuali di valori mancanti, con ampia copertura, etc. L'attribuzione dei pesi ha un forte impatto sui risultati: è perciò necessario che il modello di ponderazione sia esplicito e trasparente. Usualmente ai pesi è attribuito il valore di giudizio: in quest'ottica hanno quindi la proprietà di esplicitare gli obiettivi sottostanti la costruzione degli indicatori composti. La questione del significato dei pesi è però molto dibattuta: vedremo in uno specifico paragrafo quale significato invece i pesi possono assumere.

La fase dell'aggregazione delle informazioni, che ciascuna dimensione del fenomeno in esame porta con sé nella determinazione dell'indicatore composto, si accompagna alla pesatura. Molti diversi metodi di aggregazione sono stati elaborati: gli indicatori semplici

⁹² Con il termine *stakeholders* si individuano i "portatori di interessi", di un'azienda o di un'organizzazione, nei confronti di un'iniziativa o un progetto nel quale sono attivamente coinvolti e la cui soddisfazione influenza il progetto stesso.

possono essere sommati, moltiplicati o combinati utilizzando tecniche non lineari. Ogni approccio implica differenti assunzioni e specifiche conseguenze per cui, anche in questo caso, è importante illustrare in modo chiaro le ragioni che hanno determinato la scelta di una tecnica di aggregazione piuttosto che un'altra ad essa alternativa.

Concludendo, vogliamo unicamente sottolineare come l'assenza di modi "oggettivi" di stabilire pesi e aggregazioni non debba necessariamente indurre a rifiutare la validità del processo operato fintantoché l'intero processo è trasparente.

ROBUSTEZZA E SENSIBILITÀ

Spesso sorgono dubbi sulla robustezza dei risultati espressi dagli indicatori composti e sui messaggi socio-politici che ad essi vengono associati. La combinazione di analisi di incertezza e sensibilità può aiutare a valutare la robustezza e la qualità degli indicatori composti costruiti.

L'analisi dell'incertezza (uncertainty analysis) si focalizza su come l'incertezza nei fattori di input si propaga attraverso la struttura dell'indicatore composto ed ha effetto sul valore dell'indicatore stesso. L'analisi della sensibilità (sensitivity analysis) valuta il contributo di ogni singola fonte di incertezza sulla varianza finale. L'uso iterato di questi due tipi di analisi durante lo sviluppo degli indicatori composti può migliorarne la struttura: idealmente, infatti, tutte le potenziali fonti di incertezza dovrebbero essere esaminate: la selezione degli indicatori semplici, la qualità dei dati, lo schema di normalizzazione, l'imputazione dei dati mancanti, la scelta dei pesi e dei metodi di aggregazione, passaggi nei quali l'arbitrarietà delle scelte e la presenza di giudizi introducono un certo grado di incertezza. Le strade per valutare l'incertezza presente nei diversi passi di costruzione degli indicatori composti sono di seguito indicate:

- includere o escludere i singoli indicatori elementari;
- stimare l'entità dell'errore presente nei dati sulla base delle informazioni disponibili sulla loro variabilità;
- utilizzazione di metodi diversi da quello scelto per l'imputazione dei dati mancanti;
- utilizzazione di schemi di normalizzazioni alternativi a quello adottato;
- utilizzazione di sistemi di aggregazione diversi dal metodo selezionato;
- utilizzazione di valori diversi – ma ugualmente plausibili – per i pesi.

Lo scopo è quindi quello di rendere la costruzione degli indicatori composti più robusta: nessun metodo è, infatti, a priori migliore di uno ad esso alternativo che ugualmente assicura la coerenza interna dei diversi passaggi attuati ma, magari, serve obiettivi diversi.

RELAZIONI CON ALTRE VARIABILI

Gli indicatori composti spesso misurano concetti in relazione con altri fenomeni noti e misurabili: tali legami possono essere utilizzati per testare il potere esplicativo degli indicatori stessi. Un semplice metodo per illustrare la relazione esistente tra un indicatore composto ed un secondo indicatore, ad esso correlato, è quello di costruire un grafico che mappa le unità d'analisi in un piano cartesiano i cui assi rappresentano le due variabili in esame. Un'elevata correlazione indica unicamente che le variazioni dei valori dei due indicatori è simile. Il cambiamento del valore di un indicatore semplice però non necessariamente implica un cambiamento nell'indicatore composto e viceversa: la causalità, infatti, non è confermata dall'analisi della correlazione.

La correlazione tra l'indicatore composto ed il fenomeno misurabile può essere massimizzata o minimizzata, ad esempio, permettendo ai pesi di variare.

È importante sottolineare come gli indicatori composti spesso includano indicatori elementari tra loro correlati inducendo, quindi, un doppio conteggio: un metodo per evitare questo problema consiste nell'impegno del ricercatore ad analizzare l'entità della correlazione esistente tra l'indicatore composto risultante dall'analisi ed altri indicatori già pubblicati in letteratura.

RITORNO AI DATI ELEMENTARI

Gli indicatori composti diventano anche un punto di partenza per l'analisi operando un passaggio a ritroso di scomposizione in indicatori semplici che permette di individuarne il singolo contributo.

PRESENTAZIONE

Il modo con cui gli indicatori composti vengono visualizzati o presentati può influenzare la loro interpretazione: essi infatti dovrebbero essere in grado di comunicare il risultato dell'analisi ai destinatari/committenti in modo rapido ed accurato.

Solitamente l'uso del formato tabellare accompagna le graduatorie, presentate in ordine decrescente; istogrammi a barre orizzontali sono invece utili per confrontare visivamente il

valore delle diverse unità d'esame indicate in ordinata; per mostrare le variazioni subite nel tempo si utilizzano linee di tendenza. Si tratta quindi di diverse tipologie di tecniche di visualizzazione accomunate però da un obiettivo unico: mostrare una "foto" del fenomeno in esame, rilevando eventuali campanelli d'allarme che richiedano interventi politici specifici.

2.4.2 Considerazioni sugli indicatori composti

2.4.2.1 Introduzione⁹³

Gli indicatori sono pezzi d'informazione che sintetizzano le caratteristiche di un sistema o sottolineano cosa in esso si sta verificando. Spesso si configurano come un compromesso tra accuratezza scientifica e disponibilità di informazioni ad un costo ragionevole.

In termini generali, un indicatore è una misura quantitativa o qualitativa che deriva da una serie di "stati" osservati che possono rivelare la situazioni di una determinata area. Quando valutato ad intervalli temporali regolari, un indicatore può segnalare la direzione del cambiamento per le diverse unità d'analisi e nel tempo. Nel contesto dell'analisi politica, gli indicatori sono utilizzati per identificare tendenze di sviluppo ed allo scopo di focalizzare l'attenzione su particolari problemi; possono inoltre essere utili nello stabilire le priorità politiche e nel confronto e monitoraggio delle performance delle unità studiate.

Un indicatore composto si ottiene attraverso un'opportuna combinazione matematica, o aggregazione, di indicatori elementari che rappresentano diverse dimensioni di un concetto la cui descrizione è l'obiettivo dell'analisi.

L'utilizzo degli indicatori composti è ampiamente diffuso per fornire ad esperti e decisori politici indicazioni in merito ai seguenti elementi:

- la direzione dello sviluppo;
- il confronto tra unità di analisi, quali aree geografiche o singoli contesti;
- la valutazione dello status e del trend di variazione in relazione ad obiettivi prefissati;

⁹³ Per la stesura di questo paragrafo sono stati utilizzati i due seguenti documenti:

- Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).
- Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).

- rapidi segnali di allarme;
- l'identificazione di aree di azione;
- l'anticipazione attraverso previsioni dello sviluppo futuro del fenomeno in esame;
- un canale di comunicazione per ogni individuo e per i decisori.

Gli indicatori composti sono basati su indicatori elementari che non hanno unità di misura semantica comune: questo implica la necessità di scegliere le tecniche di analisi che più sono adatte ai dati a disposizione. Così il ricercatore, nelle diverse fasi del processo di costruzione di un indicatore composto, effettua alcune valutazioni che inducono a scelte non sempre caratterizzate dall'oggettività (ad esempio nella selezione degli indicatori, nella scelta del sistema di pesi e della tecnica di aggregazione). Qualora le scelte siano soggettive, esse diventano potenziali giustificazioni alla manipolazione dei risultati: è perciò importante individuare la fonte delle valutazioni soggettive o imprecise attraverso l'utilizzo di analisi di sensibilità e d'incertezza ma anche attraverso una trasparente dichiarazione delle motivazioni che hanno portato all'utilizzo delle metodologie scelte.

Gli indicatori composti tesi a comparare la performance di diversi Paesi trovano crescente riconoscimento come strumenti di analisi politica e di comunicazione pubblica. Essi sono più facilmente interpretabili rispetto al tentativo di definire un quadro generale a partire da una pluralità di indicatori semplici. Nonostante però sia necessario un elevato livello di aggregazione delle informazioni di base, allo scopo di accrescere la consapevolezza dell'esistenza di situazioni problematiche, l'esistenza di valori disaggregati è essenziale per tratteggiare conclusioni pertinenti sul possibile sviluppo del fenomeno analizzato.⁹⁴

Infine ci pare utile sottolineare come l'indicatore composto non sia di per sé solo l'esito di un accurato processo di valutazione: esso può infatti rappresentare il punto di partenza per l'avvio di discussioni d'interesse politico attorno al fenomeno analizzato.

⁹⁴ *"The aggregators believe there are two major reasons that there is value in combining indicators in some manner to produce a bottom line. They believe that such a summary statistic can indeed capture reality and is meaningful, and that stressing the bottom line is extremely useful in garnering media interest and hence the attention of policy makers. The second school, the non-aggregators, believe one should stop once an appropriate set of indicators has been created and not go the further step of producing a composite index. Their key objection to aggregation is what they see as the arbitrary nature of the weighting process by which the variables are combined."*

Sharpe A. (2004), Literature Review of Frameworks for Macro-indicators, Centre for the Study of Living Standards, Ottawa, CAN.

2.4.2.2 Pro e contro⁹⁵

Come nella parte introduttiva di alcuni lavori sugli indicatori composti, anche in questo elaborato vogliamo riportare un elenco di aspetti favorevoli (pro) e sfavorevoli (contro) al loro utilizzo.

Pro

- Possono essere usati per sintetizzare fenomeni complessi o multi-dimensionali nell'ottica di supportare i decisori politici.
- Forniscono una "fotografia" ad ampio raggio del fenomeno; la loro interpretazione è più facile rispetto a quella necessaria nel caso in cui si voglia tentare di creare un quadro generale del fenomeno partendo da un insieme di indicatori semplici.
- Facilitano la creazione di graduatorie relative a fenomeni complessi in esercizi di confronto.
- Sono di supporto nell'attrarre l'interesse pubblico fornendo un'immagine generale attraverso la quale è possibile comparare la performance delle diverse unità di analisi ed il loro progresso nel tempo per fenomeni complessi; i problemi esaminati entrano quindi al centro dell'arena politica.
- Permettono di ridurre una lista di indicatori o aiutano ad includere maggiori informazioni entro i limiti esistenti⁹⁶.
- Facilitano la comunicazione dei risultati al pubblico (ai cittadini, ai mass media, etc.) e promuovono la responsabilità⁹⁷.

Contro

- Possono portare a messaggi politici fuorvianti e non robusti se sono poveramente costruiti o malinterpretati. L'analisi di sensibilità può essere utilizzata per testare la robustezza degli indicatori composti.

⁹⁵ Per la stesura di questo paragrafo sono stati utilizzati i due seguenti documenti:

- Nardo M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).
- Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).

⁹⁶ In altre parole si intende affermare come l'utilizzo degli indicatori composti aiuti a "contenere" una pluralità di informazioni riconducibili a più convincenti misure di concetti/fenomeni complessi.

⁹⁷ L'utilizzo di indicatori composti evita la dispersione e la frammentazione, favorendo una più immediata comprensione da parte del profano e quindi la potenzialità di un diretto coinvolgimento.

- Gli indicatori composti possono invitare coloro che si occupano delle scelte politiche a trarre conclusioni politiche semplicistiche: per questo motivo debbono essere utilizzati in combinazione con gli indicatori elementari che li compongono al fine di tracciare conclusioni politiche sofisticate.
- Nei diversi passi previsti per la costruzione di un indicatore composto è prevista l'espressione di una scelta che può essere definibile come un "giudizio": la selezione degli indicatori elementari, la scelta del modello, i pesi, il trattamento dei valori mancanti, etc. Tali scelte dovrebbero essere trasparenti e basate su principi statistici solidi per evitare che ne venga fatto un uso strumentale (ad esempio vengono costruiti per supportare politiche desiderate da determinati soggetti).
- Gli indicatori composti incrementano la quantità di dati necessari alle elaborazioni perché occorre avere i dati sia per gli indicatori elementari sia per le analisi di significatività statistica.
- Possono nascondere situazioni negative relative ad alcune dimensioni del fenomeno in esame ed incrementano, in tal modo, la difficoltà di identificare azioni risolutive appropriate.
- Possono indurre a politiche sbagliate se alcune dimensioni difficili da misurare vengono ignorate.

2.4.3 La qualità degli indicatori composti⁹⁸

2.4.3.1 Il profilo di qualità degli indicatori composti

Il concetto di qualità degli indicatori composti non è solo una funzione della qualità dei dati (in termini di rilevanza, accuratezza, credibilità, etc.) ma anche della qualità del processo metodologico seguito per la costruzione degli indicatori composti stessi⁹⁹. Data la pluralità di aspetti da considerare, l'analisi della qualità degli indicatori composti non è un compito facile da assolvere. Si tratta quindi di un concetto con più sfaccettature le cui principali caratteristiche dipendono da prospettive, necessità e priorità dell'utente.

⁹⁸ Cfr. Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.

⁹⁹ Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).

Molte organizzazioni hanno operato per l'identificazione delle varie dimensioni della qualità dei prodotti statistici.

Particolarmente importante è la struttura elaborata da Fondo Monetario Internazionale (IMF, International Monetary Found): si tratta del lavoro che è stato definito "Data Quality Assurance Framework" (DQAF), il quale raccoglie un'ampia serie di risposte attraverso la definizione dei prerequisiti della qualità e di cinque sue dimensioni. Per il DQFA, i prerequisiti della qualità riguardano il modo in cui la qualità delle statistiche è influenzata dall'ambiente legale ed istituzionale e dalle risorse disponibili; valuta inoltre l'esistenza della consapevolezza della qualità dei dati nelle attività di gestione delle statistiche. Le cinque dimensioni della qualità stabilite dal Fondo Monetario Internazionale sono le seguenti:

- *assicurazione di integrità* (quali sono le caratteristiche che supportano l'aderenza all'obiettività nella produzione di statistiche, così da mantenere la fiducia degli utilizzatori?)
- *efficacia metodologica* (come le correnti pratiche si mettono in relazione con le pratiche metodologiche accordate a livello internazionale per specifiche attività statistiche?)
- *accuratezza ed affidabilità* (le fonti dei dati, le tecniche statistiche, etc., sono adeguate a ritrarre la realtà da descrivere?)
- *manutenzione* (come si va incontro alle necessità degli utenti in termini di tempestività dei prodotti statistici, della loro frequenza, consistenza e del loro ciclo di revisione?)
- *accessibilità* (i dati e i metadati¹⁰⁰ sono effettivamente disponibili agli utenti? Ed è fornita loro assistenza?)

L'Eurostat ha invece prodotto un lavoro focalizzato sulle produzioni statistiche così come sono visibili agli utenti finali, analizzando il processo che ne ha determinato la costruzione nei casi in cui non venga prodotta una misura diretta del fenomeno. È basato su sette dimensioni, le quali cercano di rispondere alle seguenti domande:

- *attinenza* : i dati rispondo alle aspettative dell'utente?

¹⁰⁰ Un metadato (dal greco meta- "oltre, dopo" e dal latino datum "informazione" - plurale: data), letteralmente "dato su un (altro) dato", è l'informazione che descrive un insieme di dati. I Metadati costituiscono in qualche modo il curriculum vitae dei dati, ovvero raccolgono le informazioni relative al *dove*, al *quando*, al *come* e da *chi* i dati sono stati ottenuti. Si tratta quindi di un corredo indispensabile per rendere tali dati fruibili correttamente da chiunque, anche a distanza di tempo e di spazio.

(<http://www.istitutoveneto.it/venezia/dati/metadati/metadati1.htm>)

- *accuratezza* : i dati sono affidabili?
- *comparabilità* : i dati sono comparabili tra i vari Paesi in tutti gli aspetti necessari?
- *completezza* : gli aspetti per i quali sono disponibili dati statistici comprendono tutte e solo le esigenze espresse dagli utenti?
- *coerenza* : i dati sono coerenti con altri dati?
- *tempestività e puntualità* : l'utente riceve i dati in tempo e in base alle date pre-stabilite?
- *accessibilità e chiarezza* : il quadro complessivo configurato è accessibile e comprensibile?

I due approcci descritti sono diversamente costruiti: mentre il lavoro dell'IMF è focalizzato sul processo di costruzione degli indicatori ed è quindi principalmente basato su valutazioni qualitative, l'Eurostat è orientato ad analizzare la qualità degli indicatori finali forniti all'utente elaborando misure di tipo quantitativo.

Anche l'OECD (Organisation for Economic Co-operation and Development) ha pubblicato nel 2003 una prima versione di un lavoro sulla qualità dei dati da sé stesso prodotti definito "Quality Framework and Guidelines for OECD Statistics". In esso vengono sottolineati due aspetti di qualità che devono caratterizzare i dati divulgati da un'organizzazione internazionale: la qualità delle statistiche nazionali ricevute e la qualità del processo interno di raccolta, elaborazione, analisi e diffusione di dati e metadati. Da questo punto di vista, ci sono molte similarità tra quanto l'OECD ha elaborato rispetto alla qualità delle sue stesse statistiche e quelle che sono le caratteristiche degli indicatori composti, la cui qualità complessiva dipende da due aspetti: la qualità dei dati di base¹⁰¹ e la qualità delle procedure utilizzate per la costruzione e la diffusione dell'indicatore composto. Entrambi gli elementi sono ugualmente importanti: l'applicazione di tecniche avanzate allo sviluppo di indicatori composti basati su dati non accurati o incoerenti non produrrà risultati di elevata qualità. Specularmente, un indicatore composto che combina dati di base molto buoni utilizzando procedure povere produrrà risultati non affidabili ed instabili. Infine è bene ricordare che la diffusione di indicatori composti senza appropriati metadati può portare facilmente ad interpretazioni imprecise.

¹⁰¹ Per dati di base si intendono tutte le informazioni a disposizione all'inizio del processo di costruzione di un indicatore composto le quali vengono poi selezionate, esaminate ed elaborate per ottenere indicatori elementari ed indicatori composti.

2.4.3.2 Le dimensioni della qualità dei dati di base

Riportiamo ora un elenco di dimensioni relative alla qualità degli indicatori di base, la cui selezione deve mirare a massimizzare la qualità totale dei risultati finali.

Rilevanza

Si tratta di una valutazione qualitativa del valore apportato dagli indicatori di base. Nel contesto degli indicatori composti, questo aspetto deve essere valutato considerando lo scopo generale dell'indicatore.

Accuratezza

L'accuratezza dei dati di base è il grado con cui essi stimano o descrivono correttamente le quantità o le caratteristiche che debbono misurare. Si tratta quindi di una valutazione relativa alla vicinanza tra i valori a disposizione ed i veri (sconosciuti) valori. Questo aspetto è estremamente importante nella costruzione degli indicatori composti in cui la credibilità della fonte dei dati diventa cruciale unitamente al carattere di oggettività dei dati forniti.

Tempestività

La valutazione di questa dimensione riflette la lunghezza del tempo trascorso tra l'accadimento dell'evento descritto dal fenomeno e la disponibilità dei dati su di esso. La puntualità dei dati prodotti implica la predisposizione di uno "scadenario" e riflette il grado con cui i dati sono rilasciati in conformità ad esso. Relativamente agli indicatori composti, la tempestività è particolarmente importante per ridurre al minimo la necessità di stimare dati mancanti e per revisionare dati pubblicati in precedenza. Come singole fonti di dati di base stabiliscono un individuale compromesso tra precisione e tempestività che tenga conto di vincoli istituzionali, organizzativi e di risorse, anche i dati relativi a dimensioni diverse del fenomeno in esame sono spesso rilasciati in tempi diversi, assicurando però particolare attenzione alla coerenza globale.

Accessibilità

L'accessibilità dei dati prodotti riflette quanto prontamente i dati possono essere posizionati e resi accessibili dalla fonte originale. Nel contesto degli indicatori composti, l'accessibilità dei dati di base è in grado di influenzare il costo globale di produzione e di aggiornamento degli indicatori nel tempo.

Interpretabilità

Questa dimensione riflette la facilità con la quale l'utente può capire, utilizzare propriamente ed analizzare i dati. L'aver una gamma di utenti diversi conduce a

predispone metadati in livelli di dettaglio crescente. La disponibilità di definizioni e classificazioni relative ai dati di base è essenziale per valutare la comparabilità dei dati nel tempo e tra unità (territoriali). Ribadiamo quindi l'importanza di avere adeguati metadati per valutare la qualità globale degli indicatori.

Coerenza

Si tratta dell'aspetto della qualità relativo al grado con cui i dati prodotti sono logicamente connessi e reciprocamente coerenti, permettendo confronti temporali e spaziali.

2.4.3.3 Qualità nella costruzione degli indicatori composti

Vogliamo ora collegare ciascuna fase del processo di costruzione di un indicatore composto con le dimensioni della qualità sopra descritte.

La corretta definizione del quadro teorico riguarda la rilevanza degli indicatori composti ma anche la loro credibilità – o attendibilità – ed interpretabilità. La rilevanza di un indicatore composto è usualmente valutata tenendo conto delle esigenze di analisi e politiche, ma anche del suo fondamento teorico. Da questo punto di vista, alcuni indicatori composti sono piuttosto deboli e tale debolezza spesso è adottata come critica all'uso di indicatori composti.

La qualità dei dati di base scelti per costruire gli indicatori composti influisce fortemente la loro accuratezza e la loro credibilità. Anche la tempestività può essere influenzata in larga misura dalla scelta di dati adeguati.

L'utilizzo dell'analisi multivariata per identificare la struttura dei dati può accrescere l'accuratezza e l'interpretabilità dei risultati finali. Questo passaggio è molto importante anche per identificare ridondanze tra i fenomeni selezionati e valutare possibili lacune nei dati di base.

La scelta della tecnica d'imputazione dei dati mancanti ha effetto sull'accuratezza dell'indicatore composto e sulla sua attendibilità. Inoltre, un uso massiccio di tecniche di imputazione può compromettere la qualità globale dell'indicatore e la sua importanza, anche se ne può migliorare la tempestività.

La fase di normalizzazione è cruciale sia per l'accuratezza sia per la coerenza dei risultati finali: una procedura di normalizzazione inappropriata può condurre a risultati non affidabili o distorti. D'altro canto, l'interpretabilità di un indicatore composto si basa pesantemente sulla correttezza dell'approccio seguito nella fase di normalizzazione.

Una delle questioni chiave nella costruzione di indicatori composti è la scelta del modello di aggregazione e di ponderazione. Quasi tutte le dimensioni della qualità sono influenzate dalle scelte effettuate in questa fase, ma in modo particolare coinvolgono l'accuratezza, la coerenza e l'interpretabilità. Questa è anche una delle caratteristiche degli indicatori composti più criticata: per questo colui che costruisce un indicatore deve prestare particolare attenzione al fine di evitare contraddizioni interne ed errori nella ponderazione e nell'aggregazione di indicatori elementari.

Per ridurre al minimo i rischi di produrre indicatori composti privi di senso, è necessario effettuare analisi di sensibilità e robustezza, che possono migliorare l'accuratezza, l'attendibilità e l'interpretabilità dei risultati finali. Dato il forte interesse pubblico e dei media rispetto alla costruzione di graduatorie, controlli sulla robustezza dei risultati possono aiutare a distinguere le differenze significative da quelle poco rilevanti riducendo al minimo il rischio di interpretazioni ed utilizzi scorretti.

Il confronto tra l'indicatore composto costruito ed altri indicatori già noti in letteratura su fenomeni correlati può essere molto utile per valutare la capacità di produrre risultati significativi. La rilevanza e l'interpretabilità dei risultati possono essere fortemente rafforzati da tale confronto. Inoltre, la credibilità di un indicatore può trarre beneficio dalla sua capacità di produrre risultati altamente correlati con altri indicatori di riferimento.

Data la complessità degli indicatori composti e considerando che il pubblico (mass media, cittadini, ecc), come pure i responsabili politici, in genere non prestano attenzione alle note metodologiche, è necessario curare in modo particolare la diffusione e la presentazione dei risultati per garantirne una corretta interpretazione. La comprensione dei risultati sarà infatti in gran parte basata sui messaggi forniti attraverso tabelle o grafici di sintesi.

Come già evidenziato in precedenza, gli indicatori composti possono assumere il ruolo di "punto di partenza" per l'analisi, che deve essere approfondita risalendo al dettaglio. Questa fase di analisi influenza la rilevanza degli indicatori e anche la loro capacità di essere correttamente interpretati. Inoltre, se il modo in cui l'indicatore è costruito o diffuso non permette agli utenti ed agli analisti di scendere nei dettagli, la credibilità generale dell'analisi può essere compromessa.

Infine, la fase di comunicazione dei risultati è essenziale per garantire la rilevanza degli indicatori, la loro attendibilità, l'accessibilità e l'interpretabilità. Troppo spesso gli statistici non prestano sufficiente attenzione a questa fase fondamentale, limitando così la

diffusione dei risultati prodotti ed il loro impatto globale. Recentemente l'OCSE ha messo a punto un manuale per la presentazione dei dati e dei metadati contenente linee guida utili per migliorare la diffusione dei prodotti statistici.

Tabella 2. Dimensioni della qualità dei dati coinvolte nelle fasi del processo di costruzione di un indicatore composto.

FASI DI COSTRUZIONE DEGLI I. C.	Rilevanza	Accuratezza	Credibilità	Tempestività	Accessibilità	Interpretabilità	Coerenza
Definizione del quadro teorico	X		X			X	
Selezione dei dati		X	X	X			
Analisi multivariata		X				X	X
Imputazione dei dati mancanti	X	X	X	X			
Normalizzazione		X				X	X
Ponderazione ed aggregazione	X	X	X			X	X
Analisi di robustezza e sensibilità		X	X			X	
Legame con altre variabili	X		X			X	X
Presentazione	X					X	
Ritorno al dettaglio	X		X			X	
Diffusione	X		X		X	X	

2.5 LA COSTRUZIONE DEGLI INDICATORI COMPOSTI

2.5.1 I "passi chiave"¹⁰²

I passi strettamente matematico-statistici che portano alla determinazione di un indicatore composto sono quelli relativi alla normalizzazione, alla ponderazione ed all'aggregazione.

Con il primo termine si intendono raccogliere tutte le operazioni di trasformazione degli indicatori elementari affinché essi siano resi confrontabili tra loro in termini di direzione, di unità di misura e di ordine di grandezza. La normalizzazione prevede anzitutto

¹⁰² Cfr. Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano, pag. 214-215.

l'eliminazione dell'effetto della direzione attraverso il ribaltamento del verso di quegli indicatori semplici che hanno direzione opposta al contenuto semantico del concetto. Si procede poi nel rendere omogenei i valori rispetto all'unità di misura e all'ordine di grandezza.

La ponderazione determina un insieme di pesi da applicare alle diverse dimensioni del fenomeno in esame. Ci sono diverse posizioni rispetto al significato attribuito ai pesi: spesso essi determinano il grado di importanza delle diverse dimensioni del concetto per cui viene assegnato peso maggiore alle dimensioni considerate più rilevanti. Vedremo nel seguito in dettaglio posizioni non conformi a questa teoria, secondo le quali esiste inconsistenza teorica nel correlare il concetto di importanza alla quantificazione di una variabile¹⁰³. Spesso i pesi sono determinati in base a giudizi caratterizzati da soggettività; metodi statistici alternativi (quali l'analisi delle componenti principali e l'analisi fattoriale) consentono di determinare i pesi sulla base dei dati disponibili.

La fase dell'aggregazione è strettamente legata alla ponderazione: è a questo punto che il ricercatore è chiamato a scegliere un'opportuna funzione matematica che combini tra loro gli indicatori semplici al fine di determinare un solo valore che costituisce l'indicatore composto.

2.5.2 La normalizzazione

Come precedentemente detto più volte, un indicatore composto misura un concetto complesso suddivisibile in dimensioni misurabili. Dette dimensioni vengono valutate quantitativamente da un insieme di indicatori elementari i quali hanno spesso unità di misura, ordine di grandezza e direzione diversi.

L'obiettivo della fase di normalizzazione è pertanto quello di rendere omogenei gli indicatori semplici al fine di poterli combinare in un unico valore nel tentativo di esprimere un fenomeno – o una parte di esso – in una forma più facilmente comprensibile. La normalizzazione, inoltre, è eseguita per evitare di avere valori estremi che dominino sull'insieme dei dati e mira parzialmente anche a correggere i problemi di qualità dei dati stessi. Non vi è motivo di ritenere che valori molto lontani dalla media o esterni ai valori

¹⁰³ Munda G., Nardo M. (2005), Constructing consistent composite indicators: the issue of weights, Institute for the Protection and Security of the Citizen.

centrali della distribuzione riflettano dati sottostanti poco accurati: se alcune variabili hanno distribuzioni altamente distorte, esse possono essere livellate attraverso trasformazioni logaritmiche e i dati possono essere troncati se ci sono valori estremi isolati (outliers).

La trasformazione deve quindi eliminare gli effetti dipendenti dall'unità di misura dei dati, dal loro ordine di grandezza e dalla loro direzione, mirando ad ottenere insieme di dati dove non sia forte il peso dei valori estremi e dove siano eliminati eventuali effetti ciclici.

In termini matematici il nostro costrutto sarà identificato con X : abbiamo bisogno di una funzione T che, attraverso una funzione link f , aggrega un insieme di dati diversi in un unico valore che è misura di X . La relazione tra T ed f può essere scritta così:

$$X = f\left[T_1(x_1), T_2(x_2), \dots, T_k(x_k)\right]$$

dove x_i è il valore assunto dall' i -esimo *indicatore elementare*, T_i è la i -esima *trasformazione* ed f è la funzione *link*.

La funzione T dovrebbe possedere le seguenti caratteristiche:

- *lisciata*: dovrebbe essere elementare e ben conosciuta, di ampio utilizzo e deve preservare l'ordine dei dati; deve prevedere l'eliminazione di eventuali effetti ciclici relativi ai fenomeni misurati in modo longitudinale¹⁰⁴;
- *facile computazionalmente* ovvero deve prevedere calcoli elementari;
- *comparabilità con i dati originali*: deve ri-esprimere un insieme di dati in un modo quasi paragonabile al set di dati originale;
- *resistenza*: T è resistente se contiene parametri resistenti; in altre parole ciò significa che le operazioni di trasformazione non devono lasciare che forti asimmetrie o outliers abbiano un pesante effetto sul nuovo insieme di dati.

Ma definiamo in maniera specifica cos'è una trasformazione T sui dati. T è una trasformazione sui dati x_1, x_2, \dots, x_k che mappa ogni x_i in nuovi valori $T(x_i)$; si ottengono quindi i nuovi valori $T_1(x_1), T_2(x_2), \dots, T_k(x_k)$. Solitamente è una funzione elementare, strettamente crescente (o strettamente decrescente), continua e differenziabile.

Ogni funzione T deve essere descritta in modo completo relativamente alle sue proprietà statistiche e matematiche nonché al suo campo di applicazione. Deve implicare una

¹⁰⁴ Quando la valutazione di un fenomeno avviene eseguendo ripetute misurazioni nel tempo, i valori possono risentire di effetti congiunturali dipendenti dalla stagionalità e/o dalla ciclicità degli eventi. La funzione T che noi cerchiamo deve "lisciare" le serie temporali di dati al fine di evitare l'influenza di tali fenomeni sui valori degli indicatori che il ricercatore utilizza nell'analisi.

riduzione della variabilità nel nuovo insieme di dati rispetto ai dati originali unitamente al garantire la proprietà di resistenza rispetto alla presenza di valori estremi.

2.5.2.1 Le trasformazioni lineari¹⁰⁵

Le trasformazioni lineari (LT) permettono di cambiare l'origine, la scala e l'unità di misura dei dati originali ma non cambiano la loro forma. I valori $x \{x: x \in \mathbb{R}^+\}$ sono riespressi nella forma seguente:

$$T(x) = y = a + bx \quad \text{con } a, b \in \mathbb{R}^+$$

La principale caratteristica delle trasformazioni lineari è la **proporzionalità**: essa permette di mantenere lo stesso rapporto tra osservazioni con diversa origine ($a \neq 0$) e diversa scala ($b \neq 0$).

(1) CONFRONTO CON L'UNITÀ LEADER

La prima funzione è molto comune in ogni campo di applicazione per la sua facilità di calcolo; ha inoltre significato chiaro e univoco. Dividendo l'indicatore per il valore massimo da esso assumibile, permette di annullare l'unità di misura e di modificare i valori originali forzandoli in un intervallo più ristretto.

$$T(x) = \frac{x}{\max(x)}$$

con $a=0$ e $b = \frac{1}{\max(x)}$.

In una sua seconda formulazione, tale trasformazione definisce la distanza di un indicatore dall'unità "leader" assegnando il valore 100 all'unità che possiede il valore più elevato dell'indicatore ed ordinando le altre unità in termini di punti percentuali di distanza dal leader.

$$T(x) = 100 \frac{x}{\max(x)}$$

Per la sua formulazione, si dice che tale trasformazione è dominata dal valore massimo: dal momento che talvolta i valori massimi possono essere outliers della distribuzione, l'utilizzo di questo approccio per la normalizzazione può portare a risultati non convincenti.

¹⁰⁵ Aiello F., Attanasio M. (2004), How to transform a batch of simple indicators to make up a unique one?, Atti del Convegno SIS giugno 2004, Bari – Sessioni Specializzate.

(2) VALORI RELATIVIZZATI AL CAMPO DI VARIAZIONE

Partendo dalla precedente, si arriva ad una seconda trasformazione che mappa i dati nel semplice intervallo $[0,1]$. È spesso utilizzata in economia applicata perché offre un efficace modo di confrontare i dati a livello spaziale e/o temporale con una riduzione del range.

$$T(x) = \frac{x - \min(x)}{\max(x) - \min(x)}$$

con $a = -\frac{\min(x)}{\max(x) - \min(x)}$ e $b = \frac{1}{\max(x) - \min(x)}$.

Riportando la scala da $[0,1]$ ad una nuova, ma del tutto analoga, scala da 0 a 100, la trasformazione in oggetto può essere vista come la misura della distanza dall'unità peggiore, a cui viene attribuito valore 0, e dall'unità migliore a cui si assegna valore 100.

$$T(x) = 100 \frac{x - \min(x)}{\max(x) - \min(x)}$$

Per la sua formulazione tale trasformazione è definita dominata sia dal massimo sia dal minimo. Anche in questo caso la presenza di forti asimmetrie nelle distribuzioni caratterizzate da valori estremi (sia molto elevati, sia molto piccoli) può portare ad indicatori trasformati distorti. Può però anche accadere che tale ridimensionamento ampli il campo di variazione degli indicatori che, in origine, giacciono entro un piccolo intervallo, incrementando quindi l'effetto sull'indicatore composto finale.

(3) NUMERI INDICE

La terza trasformazione valuta quanto ogni unità dista dall'unità di riferimento che viene fatta corrispondere all'unità di analisi con valore medio dell'indicatore. Al valore medio viene attribuito valore 100 ed ogni unità riceve un punteggio a seconda della sua distanza dalla media. I nuovi indicatori superiori a 100 indicano quindi prestazioni superiori alla media.

$$T(x) = 100 \frac{x}{M(x)}$$

Nella formula $M(x)$ indica il valore medio della distribuzione dell'indicatore.

(4) STANDARDIZZAZIONE (O PUNTEGGIO Z)

Standardizzare un insieme di indicatori significa convertire la loro distribuzione nella distribuzione normale standard avente media uguale a zero e varianza (o deviazione standard) unitaria. Si tratta di un approccio di ampio utilizzo per la sua facilità

d'interpretazione e perché inserisce nel calcolo una misura di variabilità. Misurando, infatti, la distanza dalla media attraverso la deviazione standard, tale approccio è molto utile nel confronto di diversi insiemi di dati.

$$T(x) = \frac{x - M(x)}{\sqrt{\text{Var}(x)}}$$

con $a = -\frac{M(x)}{\sqrt{\text{Var}(x)}}$ e $b = \frac{1}{\sqrt{\text{Var}(x)}}$.

L'utilizzo di tale trasformazione porta gli indicatori con valori estremi ad avere un maggiore effetto sull'indicatore composto, proprietà desiderabile nel caso in cui l'intento sia quello di premiare i comportamenti eccezionali. In tal modo, un risultato estremamente positivo di pochi indicatori può risultare migliore rispetto all'averne molte valutazioni con valore medio: questo effetto può essere corretto con la scelta di un'opportuna tecnica di aggregazione, che prevede, ad esempio, l'esclusione dell'indicatore migliore e di quello peggiore dal calcolo dell'indicatore composto oppure l'assegnazione di pesi diversi basati sulla "desiderabilità" del valore dell'indicatore.

(5) COMPARAZIONE CON UN GRUPPO DI CONTROLLO

Un'altra trasformazione, basta sul punteggio z , può essere sviluppata quando l'obiettivo è quello di valutare un insieme di dati rispetto ad un gruppo "normale", preso come riferimento, per il quale sono noti i valori della media e della deviazione standard. I dati risultanti saranno riespressi e misurati secondo la scala del gruppo di riferimento, con media e deviazione standard dello stesso. Si tratta di una trasformazione ampiamente usata nei test psicometrici.

Data la distribuzione – nota – del gruppo "normativo" $y \sim [M(y), \text{Var}(y)]$, la funzione di normalizzazione ha la seguente formula:

$$T(x) = M(y) + \frac{\text{Var}(y) \times [x - M(x)]}{\text{Var}(x)}$$

Anche questa trasformazione può risentire della presenza di outliers nei dati originali.

(6) DISTANZA DALLA MEDIANA

Un'ulteriore trasformazione utilizza la mediana come parametro di posizione ed il MAD (*median absolute deviation*)¹⁰⁶ come parametro di scala.

$$T(x) = \frac{x - \text{Med}(x)}{\text{MAD}(x)}$$

con $a = -\frac{\text{Med}(x)}{\text{MAD}(x)}$ e $b = \frac{1}{\text{MAD}(x)}$.

Si tratta di una funzione molto resistente perché la presenza di outliers è controllata dalla presenza della mediana e del MAD.

2.5.2.2 Le trasformazioni non lineari¹⁰⁷

Una delle ragioni più naturali che portano all'utilizzo di normalizzazioni di tipo non lineare è quella di "linearizzare" i dati attraverso trasformazioni logaritmiche, le quali cambiano la forma dei dati originali. Il principale vantaggio che ne determina l'utilizzo risiede nella necessità di non stabilire alcuna assunzione specifica per l'applicazione di un modello, elemento previsto invece nei modelli lineari.

Prima di analizzare alcune tecniche di trasformazione non lineare, è necessario un appunto terminologico. La graduatoria di un insieme di N osservazioni è l'ordine con cui vengono disposte le unità secondo le caratteristiche in esame; il rango individuale denota la posizione che ciascuna unità assume nella graduatoria.

Le trasformazioni non lineari sono basate sull'assegnazione di valori ordinali ai diversi livelli dell'indicatore: in altre parole i valori numerici assunti dagli indicatori vengono ordinati in modo crescente/decescente e viene ad essi attribuito un valore ordinale crescente/decescente. Diventa quindi possibile sommare e derivare il punteggio complessivo di ogni unità per tutti gli indicatori e/o il punteggio di ogni indicatore per l'insieme delle unità di analisi. Talvolta questi punteggi vengono pesati per costruire una media ponderata; le misure di variabilità usualmente non vengono invece utilizzate perché la loro interpretazione risulta piuttosto complessa.

¹⁰⁶ Il *median absolute deviation* è definito come $\text{MAD} = \text{median}(|Y_i - \tilde{Y}|)$ dove \tilde{Y} è la mediana dei dati e $|Y|$ indica il valore assoluto di Y . Si tratta di una variazione della deviazione media assoluta affetta in misura minore dalla presenza di valori estremi nelle code della distribuzione perché i dati e le code influenzano meno il calcolo della mediana rispetto a quanto succede con la media aritmetica.

(<http://www.itl.nist.gov/div898/handbook/eda/section3/eda356.htm>)

¹⁰⁷ Aiello F., Attanasio M. (2006), Some Issues in Constructing Composite Indicators, VIII International Meeting on Quantitative Methods for Applied Sciences, Certosa di Pontignano, 11-13 settembre 2006.

(1) TRASFORMAZIONI POTENZA

Le trasformazioni potenza (PT = Power Transformation) sono spesso utilizzate nei modelli statistici per l'analisi di dati sperimentali poiché permettono la stabilizzazione della varianza, il ripristino della normalità e la rimozione della non additività.

$$T_p(x_i) = \begin{cases} ax^p + b & \text{se } p \neq 0 \\ a \log(x) + b & \text{se } p = 0 \end{cases}$$

In essa il parametro p assume usualmente valori compresi tra -1 e 2.

Viene usata prevalentemente per ottenere cambiamenti nella forma della distribuzione originale dei dati, nonostante non mantenga la proporzionalità con i dati originali. Considerando che l'applicazione di una PT modifica sempre la forma dei dati, per cambiare l'origine o la scala dei dati è necessario assumere rispettivamente valori di $b \neq 0$ oppure di $a \neq 0$. Il codominio della funzione non è limitato: i risultati variano infatti nell'insieme \mathbb{R}^+ o in \mathbb{R}^∞ .

La scelta di un opportuno valore di p è condizionata allo scopo dello studio ed alla natura dei dati originali. Generalmente p viene desunto dai metodi grafici ed assume, come già detto, valori nell'intervallo $[-1,2]$. Quando $p \in [-1,1[$ si verifica sempre una riduzione del range e quindi una riduzione della variabilità; nel caso in cui invece $p \geq 1$ il range incrementa facendo aumentare anche la variabilità dei dati. Per costruire indicatori composti, i dati grezzi sono spesso trasformati con la trasformazione logaritmica in cui viene assunto $p = 0$: i dati essenzialmente vengono linearizzati e ne viene ridotta l'asimmetria. È però svantaggiosa quando i dati grezzi sono concentrati nell'intervallo $[0,1]$ perché i dati trasformati diventano molto negativi. Altri valori comunemente utilizzati sono $p = -1$ e $p = 2$: si tratta però di valori che non mantengono la proporzionalità tra i dati trasformati e i dati grezzi.

(2) TRASFORMAZIONI "RANGO"¹⁰⁸

Riportiamo di seguito tre tipi di trasformazioni RT (Rank Transformation) al fine di descrivere i più comuni usi nelle applicazioni pratiche. Si tratta di tecniche di normalizzazione molto semplici; non è possibile però una valutazione in termini assoluti dei risultati delle diverse unità perché tale informazione viene persa, unitamente alle informazioni relative alla variabilità dei dati.

¹⁰⁸ La traduzione dei termini inglesi utilizzati per i tipi di trasformazioni rientranti in questa categoria è nostra.

Rango

Definita RT (Rank Transformation), indica un classe di funzioni monotone che mappano i dati originali in dati ordinali, usualmente etichettati con numeri o lettere.

$$T(x) = \text{rank} \{x\}$$

Si tratta di un modo facile e veloce per comprendere e descrivere i dati, i quali vengono semplicemente ordinati; non è affetto dalla presenza di outliers e permette confronti nel tempo.

Rango a livelli (graduatoria a livelli)

La trasformazione RS (Ranking Score) è una funzione monotona che mappa i dati in intervalli.

$$T(x) = \text{score} \{x\}$$

Le applicazioni più frequenti di tale tecnica prevedono l'assegnazione di punteggi allo scopo di ordinare categorie di risposte nei questionari.

Scala categoriale

L'ultima tipologia di trasformazioni, definite CSTs (Categorical Scale Transformations), rientranti nella categoria delle RT, corrisponde all'usuale operazione di raggruppamento di un insieme di dati in categorie al fine di sintetizzarli e di visualizzarli in una forma più semplice e comprensibile.

Per molte variabili categoriali è ragionevole immaginare l'esistenza di una variabile continua sottostante: per approssimarla, è spesso utile assegnare un adeguato insieme di punteggi alle categorie. Attualmente sono utilizzati due tipi di CST: il primo assegna punteggi o valori ordinali basati sui percentili per assicurare un'adeguata rappresentanza di ogni categoria mentre il secondo basa l'assegnazione su punti di cambiamento fissati a priori di significato familiare.

Vengono quindi premiate le unità con punteggio più elevato e penalizzate quelle con performance via via peggiorare. Il limite maggiore di questo approccio è l'esclusione delle informazioni relative alla variabilità.

2.5.2.3 Altre tecniche di trasformazione

(1) INDICATORI SOPRA O SOTTO LA MEDIA

Si tratta di una tecnica che trasforma gli indicatori in modo tale che i valori attorno alla media ricevano valore 0, mentre quelli al di sopra o al di sotto di una soglia prefissata ricevano rispettivamente i valori 1 e -1. Questo metodo di normalizzazione è semplice e non è influenzato dalla presenza di outliers; tuttavia viene criticato per l'arbitrarietà che caratterizza lo stabilire il livello del valore soglia e l'omissione del livello assoluto delle informazioni. Se, ad esempio, il valore di un dato indicatore per l'unità A è 3 volte (300%) superiore alla media ed il valore dell'unità B è pari al 25% sopra la media, sia A che B sarebbero conteggiate semplicemente "sopra la media" nonostante la distanza in termini assoluti dal valore medio sia molto diversa. Oltre alla semplicità di applicazione, un ulteriore elemento positivo consiste nella non incidenza sui risultati degli outliers.

(2) METODI PER INDICATORI CONGIUNTURALI

I risultati delle indagini sulle tendenze del mercato sono in genere combinati in indicatori composti al fine di ridurre il rischio di falsi segnali e per fornire una migliore previsione dei cicli delle attività economiche. Questo metodo assegna implicitamente minor peso alle serie relative ai movimenti ciclici più irregolari dell'indicatore composto, ad eccezione dei casi in cui viene precedentemente applicato un lisciamiento ai valori della serie temporale.

(3) PERCENTUALE DI DIFFERENZE ANNUE

L'ultima tecnica di normalizzazione che qui presentiamo riguarda i casi in cui determinati indicatori siano disponibili per un certo numero di anni consecutivi. La percentuale di differenza annua su più anni consecutivi considera come valore dell'indicatore, anziché il valore assoluto, la percentuale di crescita dell'indicatore rispetto all'anno precedente.

2.5.2.4 Considerazioni conclusive

Ognuno dei metodi sopra presentati ha propri vantaggi e svantaggi. La trasformazione lineare definita standardizzazione è la più comunemente usata perché possiede caratteristiche desiderabili quando si devono poi aggregare gli indicatori elementari. Le variabili infatti sono convertite in una scala comune; i dati assumono una distribuzione normale, a media zero, evitando di introdurre distorsioni nell'aggregazione. In altri

approcci, il fattore di scala corrisponde al campo di variazione dei dati anziché alla deviazione standard: in questi casi i valori estremi possono avere un forte effetto sull'indicatore composto. La distanza dalla mediana, piuttosto che dalla media, può essere meno soggetta a distorsioni derivanti da outliers o da variabili con elevata variabilità. La trasformazione in scala categoriale ha un elevato grado di soggettività poiché i livelli e le soglie¹⁰⁹ sono determinate in gran parte arbitrariamente. Gli approcci basati invece sulla costruzione di graduatorie in base ai ranghi assegnati alle unità omettono una grande quantità di informazioni sull'entità della variabilità tra unità.

La selezione del metodo adatto necessita di particolare attenzione: devono essere prese in considerazione sia le proprietà dei dati sia l'obiettivo dell'indicatore composto. Diversi metodi di normalizzazione producono risultati diversi: per valutare il loro impatto sui risultati finali possono essere applicati opportuni test di robustezza.

2.5.3 L'aggregazione¹¹⁰

Il processo di scelta di un'appropriata funzione di aggregazione che combini in modo significativo diverse dimensioni è legato alla funzione di normalizzazione dei dati adottata.

La letteratura sugli indicatori composti offre numerosi esempi di tecniche di aggregazione.

È possibile applicare ai dati diversi metodi di aggregazione: gli indicatori elementari possono essere sommati (ad esempio attraverso l'aggregazione lineare), moltiplicati (con l'aggregazione geometrica) o aggregati utilizzando tecniche non lineari (come l'analisi multicriterio). Naturalmente ogni tecnica implica diverse ipotesi e ha specifiche conseguenze. Le più utilizzate sono quelle di tipo additivo che determinano la costruzione di una graduatoria delle unità sommando il rango che ciascuna assume per i diversi

¹⁰⁹ Le soglie corrispondono ai valori che determinano il passaggio da una categoria/livello alla/al successiva/o.

¹¹⁰ L'intero paragrafo è stato tratto dai seguenti contributi:

- Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).
- Allegra F. S., La Rocca A. (2004), Sintetizzare misure elementari: una sperimentazione di alcuni criteri per la definizione di un indice composto, Istat, Roma.
- Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).
- Delvecchio F. (1995), Scale di misura e indicatori sociali, Cacucci Editore, Bari.

indicatori ed aggregando, secondo un opportuno sistema di ponderazione, gli indicatori elementari normalizzati. Tuttavia, le aggregazioni additive implicano il soddisfacimento di requisiti e proprietà, sia relativamente agli indicatori elementari sia per i pesi ad essi associati, spesso non sono desiderabili, a volte difficili da rispettare o onerosi da verificare. Per superare queste difficoltà, la letteratura propone altri, e meno diffusi, metodi di aggregazione quali quelli moltiplicativi (o geometrici) o non lineari come l'analisi multicriterio o l'analisi cluster.

2.5.3.1 Condizioni preliminari

La prima operazione da compiere consiste nell'accertare la direzione verso cui si muovono gli indicatori rispetto al fenomeno che essi misurano. È necessario infatti che le graduatorie relative a ciascun indicatore elementare siano tutte crescenti o tutte decrescenti: occorre perciò stabilire se l'indicatore deve essere considerato di tipo crescente o positivo ossia se a variazioni positive o negative dell'indicatore corrispondono variazioni nella stessa direzione del fenomeno. Allo scopo di attribuire a tutti gli indicatori elementari il medesimo verso, occorre ribaltare la direzione degli indicatori di segno diverso; volendo, ad esempio, costruire indicatori crescenti rispetto al fenomeno in esame, sarà necessario ribaltare il verso degli indicatori negativi (ovvero di verso contrario al fenomeno) facendone il complemento rispetto all'unità di riferimento. Sintetizzando possiamo dire che si tratta di effettuare un'operazione di ri-orientamento delle misure elementari di verso negativo.

Prima di procedere con l'aggregazione, è opportuno verificare anche che gli indicatori elementari siano comparabili: ciò avviene applicando una delle tecniche di trasformazione presentate nel precedente paragrafo, attraverso le quali vengono eliminate le unità di misura riducendo gli indicatori a numeri puri ovvero svincolati dall'unità di misura.

2.5.3.2 L'approccio ordinale

I metodi di sintesi degli indicatori elementari effettuati secondo un approccio di tipo ordinale consistono nell'ordinare dapprima le unità statistiche in graduatorie rispetto ai singoli indicatori e, quindi, nel sostituire il valore assunto dall'indicatore con il rango (o numero d'ordine) con cui l'unità si colloca nella corrispondente graduatoria. Nel caso in cui due o più unità assumano lo stesso valore rispetto ad un dato indicatore, sarà ad esse attribuito il rango medio.

Operando secondo l'approccio ordinale si ottiene una graduatoria, espressa mediante una scala di tipo ordinale, che evidentemente non consente di valutare in alcun modo le distanze che separano un'unità dall'altra. È quindi evidente che queste tecniche hanno significato solamente a livello di analisi comparative nello spazio: ci si limita ad ordinare le unità – spesso territoriali – in base al livello raggiunto dal fenomeno senza poter precisare in quale misura ciascuna di esse lo possieda e a che distanza si trovino le une dalle altre.

(1) SOMMA DEI RANGHI

Il metodo della somma dei ranghi prevede – per effettuare una sintesi delle graduatorie relative ai diversi indicatori semplici – che si sommino, per ogni unità, i valori corrispondenti al rango ottenuto da ciascun indicatore.

Formalmente, considerando la generica unità statistica i ed indicando il rango assegnato come $rank\{x_{ij}\}$ nella graduatoria rispetto al j -esimo indicatore, otterremo l'indicatore sintetico

$$I_{SR} = \sum_{j=1}^m rank\{x_{ij}\}$$

Tale indicatore varia tra m (numero degli indicatori elementari impiegati) ed $m \cdot n$. Per ottenere un indicatore che vari tra 0 e 1 è sufficiente applicare la seguente formula:

$$I'_{SR} = \frac{I_{SR} - m}{mn - m}$$

che non richiede alcuni tipo di standardizzazione dei valori assunti dagli indicatori elementari dal momento che I_{SR} non tiene conto di tali valori ma solo della loro posizione relativa.

Si tratta del metodo di aggregazione più semplice; i principali vantaggi risiedono nella sua estrema facilità di calcolo e nell'indipendenza dalla presenza di outliers ma ha il grosso limite di perdere l'informazione relativa al valore assoluto assunto dalle unità.

(2) MEDIA DEI RANGHI

Alternativo al metodo della somma dei ranghi è il calcolo della media aritmetica degli stessi associati a ciascuna unità statistica. Ovviamente, in questo caso, il valore dell'indicatore sintetico si ottiene dividendo per m la somma dei ranghi attribuiti a ciascuna unità.

$$I_{MR} = \frac{I_{SR}}{m}$$

2.5.3.3 Approcci cardinali additivi

Allorché una graduatoria ottenuta secondo l'approccio ordinale non sia sufficientemente descrittiva della realtà in esame, per cui occorrono delle misure metriche che evidenzino, nei confronti temporali e spaziali, non soltanto se il fenomeno sia aumentato o diminuito, ma anche di quanto esso sia aumentato o diminuito, dovremo passare ad adottare metodi di sintesi effettuati con l'approccio cardinale.

(1) MEDIA DEI NUMERI INDICE

Dopo che ogni indicatore è stato trasformato in distanza dalla performance media (dividendo cioè il valore dell'indicatore per la media della sua distribuzione), l'indicatore sintetico può essere calcolato semplicemente come media aritmetica dei valori relativi al j -esimo indicatore registrati presso le N unità statistiche.

Si tratta di un metodo di aggregazione semplice a cui la letteratura fa ampio ricorso soprattutto in virtù dell'ipotesi, sempre conveniente in assenza di altre informazioni, di relazioni di tipo additivo tra le variabili rappresentate dagli indicatori elementari.

Vale la pena sottolineare che dopo la trasformazione gli indicatori semplici sono liberi dall'unità di misura ma non svincolati dalla loro variabilità: operando quindi con la media aritmetica pesano di più sul risultato finale gli indicatori con variabilità maggiore.

(2) MEDIA DEI VALORI RELATIVIZZATI COL CAMPO DI VARIAZIONE

Tale tecnica di aggregazione costruisce l'indicatore composto come media aritmetica dei valori relativizzati al proprio campo di variazione: gli indicatori elementari sono svincolati dall'unità di misura ma anche dalla loro variabilità dal momento che vengono appunto relativizzati alla differenza tra il valore massimo e il minimo della distribuzione.

(3) MEDIA DEI VALORI STANDARDIZZATI

Un'ulteriore metodo di aggregazione, costruito in maniera analoga ai precedenti, prevede la sintesi di un insieme di indicatori elementari come media aritmetica dei valori standardizzati (o punteggi z). L'utilizzo di indicatori elementari così trasformati rende questa tecnica di aggregazione robusta in presenza di eventuali outliers dal momento che la deviazione standard è una buona misura di variabilità.

2.5.3.4 Approccio multicriterio non compensativo

Utilizzando tecniche di aggregazione additive o moltiplicative, i pesi vanno ad assumere il significato di "tassi di scambio" (trade-off¹¹¹) e non indicano l'importanza degli indicatori ad essi associati. Ciò implica una logica di compensazione per cui è possibile compensare situazioni negative rispetto ad alcuni aspetti con valutazioni positive in altri. Approfondiremo tale questione in seguito: ci siamo qui limitati a riportarla brevemente per introdurre un ulteriore approccio utilizzato nella fase di aggregazione di indicatori elementari, caratterizzato dall'assunzione del significato di "coefficienti d'importanza" dei pesi attribuiti agli indicatori. Si tratta del criterio denominato approccio multicriterio non compensativo, il quale risolve il problema dato dalla logica della compensazione alla luce dell'assenza dell'indipendenza preferenziale¹¹².

Dato un insieme di indicatori $G = \{x_i\}$ ($i = 1, \dots, m$) ed un insieme finito di N unità, si assume che la valutazione di ciascuna unità rispetto ad un singolo indicatore x_i si basi su una scala di misura intervallo o rapporto. Per semplicità di esposizione, si ipotizza che un valore più elevato dell'indicatore sia preferibile a un valore inferiore: in altre parole significa che l'unità con valore dell'indicatore più alto è la migliore. Si suppone inoltre l'esistenza di un insieme di pesi $w = \{w_i\}$ che sommano a 1 ($\sum_{i=1}^m w_i = 1$) interpretati come coefficienti d'importanza. Questa informazione costituisce la *matrice di impatto*.

Dal punto di vista matematico, il problema si traduce nella ricerca di un modo per utilizzare queste informazioni al fine di creare una graduatoria delle unità dalla migliore alla peggiore senza alcuna relazione di incomparabilità. Quattro sono i punti particolarmente importanti:

- intensità di preferenza (quanto un'unità a è meglio dell'unità b rispetto all'indicatore i);
- numero di indicatori a favore di una determinata unità;
- peso attribuito a ciascun indicatore;
- rapporto tra ogni unità e le altre appartenenti all'insieme valutato.

L'operazione matematica di aggregazione può essere divisa in due fasi:

¹¹¹ Il trade-off indica la situazione in cui un guadagno rispetto a un dato obiettivo implica necessariamente una perdita riguardo a un altro.

¹¹² L'indipendenza preferenziale è una condizione molto forte sui dati: implica che rapporto trade-off tra due variabili sia indipendente dei valori delle restanti $m-2$ variabili. Da un punto di vista operativo, ciò significa che una funzione di aggregazione additiva permette la valutazione del contributo marginale di ciascuna variabile separatamente. Questi contributi marginali possono quindi essere sommati per ottenere un valore sintetico complessivo.

- confronto a coppie delle unità in base all'insieme completo di indicatori;
- classifica delle unità in ordine preliminare completo.

Il risultato del primo passo è una matrice E di dimensioni $N \times N$ chiamata matrice outranking (matrice di posizionamento) il cui generico elemento e_{jk} con $j \neq k$ è il risultato del confronto a coppie tra l'unità j e l'unità k rispetto all'insieme completo degli m indicatori elementari. Ogni confronto tra coppie di unità è ottenuto dalla seguente equazione:

$$e_{jk} = \sum_{i=1}^m \left(w_i (\text{Pr}_{jk}) + \frac{1}{2} w_i (\text{In}_{jk}) \right)$$

dove $w_i(\text{Pr}_{jk})$ e $w_i(\text{In}_{jk})$ sono i pesi assegnati agli indicatori che presentano una relazione rispettivamente di preferenza (Pr) e di indifferenza (In). In altre parole, il punteggio dell'unità j è dato dalla somma dei pesi degli indicatori elementari per cui j è migliore di k e dalla metà dei pesi associati agli eventuali indicatori per cui le due unità risultano avere la stessa performance. Chiaramente $e_{jk} + e_{kj} = 1$.

Il problema è quindi quello di valutare se l' i -esimo indicatore è più elevato per l'unità A o per l'unità B; se realmente, ad esempio, l'unità A ha una valutazione superiore, si ottiene il peso dell'indicatore i che entra nel calcolo dell'importanza globale dell'unità A, coerentemente con la definizione di pesi quali misure d'importanza.

Il modo in cui le informazioni vengono combinate genera una pluralità di possibili graduatorie, ciascuna delle quali ha propri limiti e vantaggi. Un possibile algoritmo da applicare per la costruzione di tali graduatorie è il Condorcet-Kemeny-Young-Levenglick (CKYL)¹¹³. Secondo tale procedura, la classificazione delle unità con la maggiore verosimiglianza è quella sostenuta dal massimo numero di indicatori per ogni confronto a coppia, cumulativamente su tutte le coppie di unità considerate. Formalmente, tutti gli $N(N-1)$ confronti a coppie compongono la matrice outranking E . Definiamo R l'insieme di tutte le $N!$ possibili graduatorie alternative, tale che $R = \{r_s\}$ con $s=1, 2, \dots, N!$. Per ogni r_s viene calcolato il corrispondente punteggio φ_s come sommatoria degli elementi e_{jk} di tutte le $\binom{N}{2}$ coppie di confronto j,k . Formalmente avremo

$$\varphi_s = \sum e_{jk} \cdot$$

¹¹³ Munda, G. and Nardo, M. (2003), On the methodological foundations of composite indicators used for ranking countries, OECD/JRC Workshop on Composite Indicators of Country Performance, Ispra, Italy, May 12.

La graduatoria finale (r^*) è la soluzione di

$$r^* \Leftrightarrow \varphi_s = \max \sum e_{jk}$$

ovvero la permutazione con il punteggio maggiore.

Questo metodo di aggregazione ha il vantaggio di superare alcuni dei problemi sollevati dalle aggregazioni additive o moltiplicative, quali la dipendenza preferenziale, l'uso di diverse scale rapporto o intervallo per esprimere lo stesso indicatore ed il significato di trade-off dato ai pesi. Con questo metodo, inoltre, le informazioni qualitative e quantitative possono essere trattate congiuntamente. Inoltre, non necessita l'applicazione di manipolazioni o di normalizzazioni sugli indicatori per garantirne la comparabilità.

D'altra parte, tale tecnica presenta anche alcuni inconvenienti quali, prima fra tutti, la dipendenza delle alternative irrilevanti, vale a dire la possibile presenza di inversioni nei cicli/graduatorie per cui nella graduatoria finale l'unità A risulta migliore di B, B è preferibile a C, ma C è preferibile ad A. Inoltre, l'informazione sull'intensità di preferenza delle variabili non viene mai utilizzata: se un indicatore per l'unità A è di molto inferiore al valore dello stesso indicatore per l'unità B, la graduatoria sarà uguale a quella che viene prodotta nel caso in cui tale differenza sia molto piccola.

Infine vogliamo far notare come il punto focale di questo metodo sia spostato, piuttosto che alla mera aggregazione degli indicatori, sulla determinazione dei pesi, cruciale per la determinazione della graduatoria finale delle unità.

2.5.3.5 Aggregazione geometrica

Una caratteristica non desiderabile nelle aggregazioni di tipo additivo è la piena compensabilità: bassi valori di alcuni indicatori possono essere compensati da valori sufficientemente elevati di altri indicatori. Se l'analisi multicriterio comporta la piena non compensazione, l'uso di un'aggregazione geometrica (chiamata anche "indice di deprivazione") è una soluzione intermedia.

$$I = \prod_{i=1}^m x_i^{w_i}$$

Si tratta quindi della produttoria degli m indicatori elementari elevati al loro peso.

2.5.3.6 Tecniche di aggregazione multivariate

I metodi di aggregazione di indicatori elementari di tipo multivariato si basano sull'analisi delle componenti principali e sull'analisi fattoriale. Intendiamo quindi, anzitutto, spiegare

brevemente tali modelli statistici per poter poi comprendere in maniera più chiara le tecniche di aggregazione da esse dipendenti.

Analisi delle Componenti Principali

Uno degli obiettivi più rilevanti che ci si propone di cogliere mediante il metodo delle componenti principali (PCA = Principal Components Analysis) è quello di eliminare la ridondanza che può esservi nel complesso delle informazioni disponibili su una popolazione, sostituendo alle variabili originali – fra loro correlate – un minor numero di nuove variabili latenti, derivate da queste ultime e fra loro non correlate, in grado di fornire una quota sufficiente dell'informazione complessiva contenuta nelle variabili originarie. Nello specifico si tratta quindi di sintetizzare le informazioni contenute nell'insieme degli m indicatori elementari raccolti mediante un numero p di variabili latenti dette componenti principali (con $p < m$).

Formalmente, la prima operazione da eseguire consiste nell'esprimere le variabili iniziali, ossia gli indicatori elementari, sotto forma di scarti standardizzati z_i allo scopo di svincolarle dall'unità di misura e di omogeneizzarne le varianze. Le componenti principali W_i sono combinazioni lineari degli indicatori standardizzati

$$W_i = a_{i1} z_1 + a_{i2} z_2 + \dots + a_{im} z_m$$

dove $i = 1, 2, \dots, m$ con m numero degli indicatori elementari selezionati; esse sono linearmente incorrelate tra loro e spiegano in proporzioni progressivamente decrescenti la varianza totale. La prima componente principale W_1 va perciò determinata in modo che spieghi il massimo della variabilità degli indicatori con la condizione che il vettore $\underline{a}_1^T = (a_{11}, a_{12}, \dots, a_{1m})$ sia normalizzato ovvero $\underline{a}_1^T \underline{a}_1 = a_{11}^2 + a_{12}^2 + \dots + a_{1m}^2 = 1$. La seconda componente, incorrelata con la prima già estratta, si ottiene imponendo che essa spieghi il massimo della variabilità residua; le successive componenti principali vengono ottenute iterando via via il procedimento.

Il problema si riduce, quindi, alla ricerca dei valori del vettore $\underline{a}_1^T = (a_{11}, a_{12}, \dots, a_{1m})$.

Matematicamente la soluzione risiede nella ricerca degli **autovalori** λ_j della matrice di correlazione¹¹⁴ R : una volta ricavati e posti in ordine decrescente, il vettore \underline{a}_j è proprio

¹¹⁴ La matrice di correlazione R è costituita dai coefficienti di correlazione tra le variabili considerate nell'analisi. Le radici λ_j dell'equazione caratteristica $|R - \lambda I| = 0$ si denominano autovalori di R . Il vettore \underline{a}_j soluzione del sistema vettoriale $R \underline{a}_j = \lambda_j \underline{a}_j$ si denomina autovettore di R .

l' i -esimo **autovettore** normalizzato della anzidetta matrice. Inoltre l'autovalore λ_i è uguale alla varianza della componente latente:

$$\text{Var}(W_i) = \lambda_i.$$

Poiché $\text{Var}(z_j) = 1$ (essendo z_j standardizzata), per cui la varianza complessiva spiegata dalle variabili originarie è m , risulta che $\frac{\lambda_i}{m}$ esprime quanta parte della varianza totale è spiegata dalla i -esima componente latente. Un'ulteriore proprietà soddisfatta dall'analisi consiste nell'uguaglianza tra somma degli autovalori e numero degli indicatori elementari:

$$\sum_{i=1}^m \lambda_i = m.$$

La mancanza di correlazione tra componenti principali è un'importante proprietà: significa infatti che esse misurano diverse "dimensioni statistiche" dei dati.

Uno degli obiettivi di tale analisi, come abbiamo detto, mira alla parsimonia ed alla chiarezza della struttura dei dati: se però gli indicatori elementari sono originariamente tra loro non correlati, l'applicazione dell'analisi delle componenti principali non porterà a alcun risultato in questo senso.

I coefficienti di correlazione $r(W_i, z_j)$ tra le componenti principali W_i e le variabili originali z_j sono definiti **pesi fattoriali**; nel caso in cui gli indicatori elementari siano incorrelati, essi sono uguali ai coefficienti a_{ji} . Il quadrato del valore dei pesi esprime la percentuale di varianza della i -esima variabile spiegata da ciascuna componente principale.

Il **punteggio fattoriale** per una data unità statistica rispetto ad una determinata componente principale è calcolato moltiplicando il valore standardizzato assunto dall' i -esima variabile per il corrispondente punteggio della stessa variabile nella data componente principale e sommando tali prodotti.

Analisi Fattoriale

Questo strumento statistico ha scopi molto simili all'analisi delle componenti principali. L'analisi fattoriale (FA = Factor Analysis) è in grado di offrire risultati significativi in ordine alla sintesi di indicatori: consente di estrarre pochi fattori latenti - dall'insieme dei dati a disposizione del ricercatore - che esprimano la maggior parte delle informazioni contenute nelle variabili originarie, consentendo l'individuazione delle dimensioni effettivamente diverse di cui si compone il concetto in esame.

Il metodo consiste nel supporre che ciascuna delle variabili standardizzate originarie z_j sia funzione lineare di alcune variabili latenti F_i , dette **fattori comuni**, e di un fattore presente per ciascuna variabile detto **fattore unico**:

$$z_j = b_{j1} F_1 + b_{j2} F_2 + \dots + b_{jp} F_p + c_j U_j$$

in cui il coefficiente b_{ji} , denominato peso fattoriale, è il coefficiente di correlazione tra la variabile z_j ed il fattore F_i .

La somma delle parti di varianza della variabile osservata z_j spiegata dai fattori comuni F_i è detta **comunanza** mentre il fattore unico spiega la varianza rimanente, detta **unicità**. In questa analisi si ipotizza che i fattori unici siano incorrelati tra loro e che i fattori comuni siano indipendenti dai fattori unici ed abbiano varianza unitaria.

Una volta individuati i fattori comuni, nasce il problema della loro identificazione o denominazione per potersi riferire ad essi come dimensioni di cui il concetto in esame si compone.

I passi operativi previsti dall'analisi fattoriale sono riassumibili come segue:

- identificazione della matrice su cui eseguire l'analisi, usualmente R ;
- estrazione dei fattori, comunemente operata ricorrendo al metodo delle componenti principali;
- rotazione dei fattori estratti per ottenere una soluzione più facilmente interpretabile;
- calcolo dei punteggi fattoriali f_{hi} che sono i valori assunti, per l'unità h -esima, dall'indicatore che esprime la dimensione i -esima del concetto.

Se si estraggono i fattori con il metodo delle componenti principali, l' i -esimo fattore è dato dalla seguente formula

$$F_i = \frac{W_i}{\sqrt{\lambda_i}}$$

dove W_i è l' i -esima componente principale estratta e λ_i è l' i -esimo autovalore.

Per identificare e denominare i fattori ci si serve della matrice di correlazione $B = [b_{ij}]$ (per $i=1, 2, \dots, p$ e $j=1, 2, \dots, m$) fra le m variabili originarie ed i p fattori: in tal modo, infatti, osservando con quali indicatori elementari originali il fattore presenta maggiore correlazione, è possibile spiegarne con più facilità il significato. Nel caso in cui tali correlazioni non siano molto marcate, ci viene in aiuto la rotazione dei fattori che fa aumentare alcune correlazioni e ne fa diminuire altre migliorando l'interpretabilità dei risultati.

Dopo tali operazioni, è possibile determinare il valore dei punteggi fattoriali f_{hi} , ovvero i valori che l'indicatore sintetico assume presso la generica unità statistica h ($h = 1, 2, \dots, n$), attraverso la formula di seguito riportata:

$$f_{hi} = \frac{z_{h1} a_{1i}}{\sqrt{\lambda_i}} + \frac{z_{h2} a_{2i}}{\sqrt{\lambda_i}} + \dots + \frac{z_{hm} a_{mi}}{\sqrt{\lambda_i}} .$$

Scritta la matrice M dei coefficienti fattoriali, ognuna delle dimensioni individuate può spiegarsi in una scala unidimensionale: è sufficiente calcolare i pesi fattoriali f_{hi} considerando solo le variabili z_j correlate con il fattore considerato.

Questo metodo attribuisce un punteggio fattoriale alle unità statistiche pari al valore dell'indicatore sintetico che consente il confronto fra unità ma che fornisce anche il contributo che ciascuna dimensione apporta nel confronto.

Considerazioni su PCA e FA¹¹⁵

Esiste un'importante differenza tra questi due tipi di analisi: l'analisi fattoriale è basata su un modello statistico specifico che definisce un'ipotesi sulla struttura della variabilità comune alle osservazioni; l'analisi delle componenti principali, invece, non è basata su alcun modello particolare ovvero su un modello di "non scelta" per cui tutte le componenti sono comunque considerate, anche se soltanto alcune saranno poi utilizzate ai fini interpretativi.

Nello sviluppo di indicatori composti è preferibile l'utilizzo dell'analisi delle componenti principali per la sua virtù di semplicità e perché permette la costruzione di pesi rappresentativi dell'informazione contenuta negli indicatori elementari. Va comunque notato che diversi metodi di estrazione forniscono valori diversi dei fattori così come dei pesi, influenzando il valore finale dell'indicatore composto e la corrispondente graduatoria delle unità.

Relativamente al problema di quanti fattori tenere nell'analisi senza perdere troppa informazione, esistono molte diverse tecniche sulle quali le opinioni dei metodologi divergono. L'interrompere l'estrazione dei fattori dipende fondamentalmente dal momento in cui rimane poca variabilità residua casuale, decisione comunque piuttosto arbitraria. Tuttavia sono stati sviluppati diversi orientamenti ("stopping rules"): di seguito sono riportati approssimativamente nell'ordine di frequenza d'uso.

¹¹⁵ Le considerazioni sull'analisi delle componenti principali e sull'analisi fattoriale sono state tratte, oltre ai lavori già considerati nella nota ad inizio del paragrafo sull'aggregazione, dal seguente testo:
Fabbris L. (1997), *Statistica multivariata – Analisi esplorativa dei dati*, McGraw-Hill, Milano.

Criterio di Kaiser. Vengono eliminati tutti i fattori con autovalori inferiori ad 1. La più semplice giustificazione a questa regola risiede nel fatto che non ha senso aggiungere un fattore che spieghi meno della varianza introdotta nel modello da ogni singolo indicatore elementare (che è appunto pari a 1).

Rappresentazione grafica degli autovalori. Questo metodo prevede la costruzione di un grafico – detto *scree plot* – che mappa il valore degli autovalori (in ascissa) rispetto alle relative componenti principali o ai fattori (in ordinata). Unendo i punti individuati nel piano cartesiano, si ottiene una spezzata: se questa mostra due tendenze, una forte inclinazione all'altezza dei primi fattori ed un successivo appiattimento che la porta ad essere quasi parallela all'asse delle ascisse, i fattori che appartengono a quest'ultima parte della spezzata possono essere ignorati. Si considerano quindi rilevanti solo i fattori il cui autovalori, stando più in alto del "gomito" (o "flesso") descritto dalle due tendenze, si stacca visibilmente dagli altri.

Criterio della varianza spiegata. Alcuni ricercatori utilizzano semplicemente la regola del mantenere il numero di fattori che spiega complessivamente una predeterminata percentuale di varianza (ad esempio, non inferiore all'80%).

Criterio di Joliffe. Secondo tale regola, non vengono considerati tutti i fattori con autovalori inferiori a 0.70; si tratta però di un criterio poco utilizzato soprattutto perché può portare a considerare un numero troppo elevato di fattori.

Comprensibilità. Tale criterio, non strettamente matematico, limita il numero di fattori, da tenere in considerazione nell'analisi, in funzione della facilità d'interpretazione che ne deriva; spesso si tratta solamente dei primi – al massimo i primi tre – fattori estratti.

Può risultare delicato scegliere la modalità di utilizzazione dell'analisi delle componenti principali e dell'analisi fattoriale per aggregare indicatori elementari al fine di costruire un indicatore composto. Allegra e La Rocca¹¹⁶ hanno proposto alcune tecniche di aggregazione multivariata per sintetizzare misure elementari al fine di definire un indicatore composto, dalle quali si possono ottenere i seguenti risultati:

- un solo indice sintetico per gli indicatori considerati, in corrispondenza dell' i -esima unità statistica del collettivo;
- due o più indici sintetici, in corrispondenza dell' i -esima unità statistica.

¹¹⁶ Allegra F. S., La Rocca A. (2004), Sintetizzare misure elementari: una sperimentazione di alcuni criteri per la definizione di un indice composto, Istat, Roma.

Tra le tecniche multivariate sono state scelte quelle di fattorizzazione, proposte nelle seguenti varianti, in relazione al metodo di estrazione e di rotazione degli assi fattoriali. L'ipotesi di base, comune a tutte le varianti, è che la correlazione tra le misure elementari è spiegata da fattori non osservabili (latenti). Punto di partenza è la matrice di correlazione R degli indicatori elementari standardizzati. Dalla matrice dei coefficienti fattoriali, che si ricava dalla matrice dei pesi fattoriali, è possibile ottenere uno o più indici sintetici, a seconda del numero di fattori selezionati.

1) ANALISI DELLE COMPONENTI PRINCIPALI CON UN SOLO FATTORE ESTRATTO

L'utilizzo di questa tecnica prevede la determinazione dell'indicatore sintetico attraverso il calcolo della somma dei prodotti tra il generico valore z_{ij} rappresentato dal fattore ed il relativo punteggio fattoriale secondo la formula che segue:

$${}^a I = \sum_{j=1}^m z_{ij} {}^a F_j$$

dove z_{ij} è la determinazione standardizzata assunta dalla i -esima unità osservata per il j -esimo indicatore e f_j definisce il punteggio fattoriale per il j -esimo indicatore. Fissando un unico fattore, si ottiene quindi un unico indicatore composto che, tuttavia, può non essere in grado di rappresentare tutti gli indicatori elementari: la sintesi in un solo fattore comporta inevitabilmente una perdita di informazioni per cui le performance delle unità sono sottostimate o sovrastimate.

Sottolineiamo infine come ordinare le unità lungo la dimensione estratta significhi graduare le unità in relazione alle misure che più hanno contribuito a formarla: operativamente questo avviene attraverso l'analisi delle correlazioni degli indicatori originali con l'unica componente estratta.

2) ANALISI DELLE COMPONENTI PRINCIPALI CON DUE FATTORI ESTRATTI E CONSIDERATI ORTOGONALI, MA CON L'ADOZIONE DEL PRIMO FATTORE

Per prendere in considerazione altre misure non rappresentate dall'unico asse fattoriale è necessario utilizzare un secondo asse. Si ottiene in questo modo un secondo indicatore di sintesi analogo al primo:

$${}^b I = \sum_{j=1}^m z_{ij} {}^b F_j$$

Per semplificare la struttura dei fattori, consentendone la separazione, si estraggono quindi due fattori ottenendo, per ogni unità, due indici composti aI e bI : attraverso la rotazione degli assi, muteranno i contributi degli indicatori elementari ed il punteggio fattoriale del primo fattore si modificherà in virtù della presenza del secondo asse.

3) ANALISI DELLE COMPONENTI PRINCIPALI CON SOMMA O MEDIA DEI PUNTEGGI FATTORIALI DEI DUE FATTORI ORTOGONALI

L'esigenza di avere un unico indicatore sintetico può comportare che gli indicatori aI e bI possano essere accorpati attraverso un'operazione di somma dei punteggi fattoriali dei due fattori ortogonali oppure mediante la media dei punteggi. In altri termini, i criteri di sintesi adottati considerano i primi due assi fattoriali estratti dall'analisi delle componenti principali. Si dà conto così di una quantità maggiore di varianza spiegata, rappresentando gli indicatori "sfuggiti" alla prima componente.

4) ANALISI FATTORIALE CON UN FATTORE ESTRATTO

È infine possibile ottenere un ulteriore criterio di sintesi mutando il metodo di estrazione dei fattori: attraverso l'analisi fattoriale, analizzando solo la varianza attribuibile ai fattori comuni (la comunanza).

Le differenze – in termini di posizione delle unità in una graduatoria rispetto ai precedenti criteri – sono interamente attribuibili al diverso contributo che le misure elementari danno alla formazione dei fattori.

2.5.3.7 Altre tecniche di aggregazione

(1) INDICATORI SOPRA E SOTTO LA MEDIA

Tale metodo è basato sulla differenza tra il numero di indicatori che sono sopra o sotto il una prefissata soglia usualmente vicina – o coincidente – alla media della distribuzione.

$$I = \sum_{i=1}^N \operatorname{sgn} \left[\frac{x_i}{M(x)} - (1 + p) \right]$$

La soglia p può essere scelta arbitrariamente. Tale metodo di normalizzazione degli indicatori, e, di conseguenza, la relativa tecnica di aggregazione, non risente dell'eventuale presenza di outliers.

(2) METODO TASSONOMICO DI WROCLAW

Nato in campo economico, il metodo di sintesi degli indicatori di Wroclaw ha lo scopo di costruire indicatori composti espressivi del livello di sviluppo. Può tuttavia essere applicato in tutti i casi in cui sia utile la determinazione di una misura complessiva e sintetica di indicatori e risulta di grande utilità come approccio di analisi comparativa delle unità – in genere territoriali – osservate.

Due sono le tecniche operative di applicazione di questo metodo: la prima prevede la sintesi degli indicatori elementari per mezzo di un indice di distanza di ogni unità territoriale da una "unità leader"; la seconda tecnica raggruppa le unità territoriali omogenee analizzando le mutue distanze di ogni unità dalle altre tramite grafici tassonomici.

Distanza da una unità ideale

La base di partenza è costituita dalla matrice degli indicatori elementari normalizzati in punteggi z (ovvero standardizzati): essa viene utilizzata per costruire un indice sintetico ottenuto mediante aggregazione delle distanze delle singole unità da un valore ideale che, per ciascun indicatore, è il valore massimo raggiunto. Le misure elementari migliori, convertite in scarti standardizzati z , serviranno a costruire un vettore di valori ideali che identifica appunto l'unità ideale. Il confronto di ciascuna unità con l'unità ideale si effettua mediante il calcolo della loro distanza euclidea, definita *misura dello sviluppo* del modello di Hellewig, $D_{i,0}$ che diverrà il numeratore dell'indicatore composto.

$$D_{i,0} = \sqrt{\sum_{j=1}^m (z_{ij} - z_{0j})^2}$$

Considerando poi la media aritmetica \bar{D}_0 delle distanze eucleree delle singole unità statistiche dal valore ideale e il loro scarto quadratico medio σ_0 , è possibile determinare il denominatore dell'indicatore composto come segue:

$$D_0 = \bar{D}_0 + 2\sigma_0$$

dove $\bar{D}_0 = \frac{1}{N} \sum_{i=1}^N D_{i,0}$ e $\sigma_0 = \sqrt{\frac{1}{N} \sum_{i=1}^N (D_{i,0} - \bar{D}_0)^2}$.

L'indicatore sintetico finale, che si ottiene dalla seguente formula,

$$I = \frac{D_{i,0}}{D_0}$$

indica il livello di sviluppo dell' i -esima unità, che sarà tanto più elevato quanto più l'indicatore è vicino a 0 ovvero quanto più piccola è la distanza dall'unità ideale.

Grafici tassonomici

L'utilizzo dei grafici tassonomici per il raggruppamento di unità – territoriali – omogenee offre un notevole contributo all'analisi dei fenomeni complessi mediante indicatori elementari. Essi vengono costruiti in base agli indici di mutua distanza tra le unità, e non, perciò, da un'unità ideale.

Anzitutto definiamo la distanza di una generica unità r da un'altra unità s :

$$D_{rs} = \sqrt{\sum_j (z_{rj} - z_{sj})^2}$$

essendo valide le seguenti relazioni che caratterizzano la distanza:

$$D_{rr} = 0 \qquad D_{rs} = D_{sr} \qquad D_{rs} \leq D_{rp} + D_{ps} \quad .$$

Calcolate le distanze tra unità, si costruisce la matrice delle distanze $D_{n,n}$, la quale è simmetrica con valori nulli sulla diagonale principale: essa servirà di base per l'individuazione di gruppi omogenei (o tipologici) di unità.

Il procedimento ha inizio determinando nella matrice D gli elementi più piccoli di ciascuna riga che denoteremo, al variare di j , come segue:

$$d_i^* = \min(D_{ij}) \quad .$$

Indicheremo poi con j_i l'indice della colonna a cui appartiene l'elemento più piccolo della riga i : al termine del procedimento si otterranno N coppie ordinate di indici $(1, j_1), (1, j_2), \dots, (1, j_N)$ dall'analisi delle quali si individueranno i raggruppamenti di unità. Per ogni raggruppamento viene poi identificato un centro che corrisponde all'unità la cui distanza media dalle altre unità del gruppo è minore.

Si tratta di un procedimento di analisi multidimensionale applicato a batterie di indicatori per la semplicità di calcolo.

(3) DIFFERENZE ANNUALI

Anzitutto si normalizza l'indicatore elementare sostituendolo con la differenza tra l'ultimo valore rilevato e il valore dell'anno precedente divisa per il valore dell'anno precedente. Successivamente, l'indicatore composto viene costruito come somma degli indicatori normalizzati.

$$I = \frac{\sum_{i=1}^N w_i y_i^t}{\sum_{i=1}^N w_i}$$

con $y_i^t = \frac{x_i^t - x_i^{t-1}}{x_i^t}$ dove x_i^t è il valore assunto dall'indicatore nell'anno corrente (t) mentre x_i^{t-1} identifica il valore che l'indicatore x ha assunto nell'anno precedente (t-1).

(4) METODO RIZZI

Com'è noto, fissato un sistema di riferimento (costituito da un punto O detto origine, da m assi orientati che partono da esso e da m unità di misura, una per ciascuna dimensione), c'è una corrispondenza biunivoca fra i punti dello spazio (che, essendo costituito da più dimensioni viene definito iperspazio) e le m -ple ordinate di numeri reali. Dal punto di vista geometrico, quindi, un'unità statistica, su cui sono state osservate m variabili, può essere considerata come un punto di un iperspazio ad m dimensioni. In questa ottica, anche i valori assunti presso le N unità statistiche dalle m variabili latenti dell'analisi delle componenti principali possono essere considerati le coordinate di N punti di un iperspazio a m dimensioni, i cui assi ortogonali sono costituiti proprio dalle m variabili latenti dell'analisi delle componenti principali.

Ciò premesso, il Rizzi¹¹⁷, per tener conto di tutte le informazioni contenute nei dati (e non solo di quelle contenute nella prima componente principale), propone come indicatore sintetico degli indicatori elementari, rilevati per ciascuna unità statistica, la distanza euclidea dall'origine di detta unità rappresentata nello spazio individuato dalle variabili latenti anzidette.

Assunto il segno della prima componente principale ($\text{sgn } c_{i1}$) dal momento che essa, per definizione, è quella che spiega più variabilità rispetto alle altre componenti, l'indicatore sintetico viene calcolato secondo la seguente formula:

$$I = \text{sgn } c_{i1} \sqrt{\sum_j c_{ij}^2} .$$

¹¹⁷ Rizzi A. (1988), Un metodo di graduazione di più unità statistiche, Rivista di Statistica Applicata, vol. 21, n. 1, marzo.

2.5.3.8 Metodi di aggregazione a confronto

Come abbiamo visto, esistono diversi metodi di aggregazione; qualora le graduatorie ottenute con i vari metodi presentassero un'elevata concordanza, essi offrono comunque un diverso grado di informazione e prevedono diversi contesti di utilizzo preferenziale.

Il metodo ordinale non consente la valutazione delle distanze del livello del fenomeno nelle unità indagate. A svantaggio del metodo di trasformazione degli indicatori elementari in numeri indici gioca il fatto che, pur svincolati dall'unità di misura, gli indicatori sintetici sono sensibili alla variabilità degli indicatori di base, poiché pesano di più gli indicatori che, in termini di numeri indici, hanno variabilità maggiore. A questo inconveniente si sottrae la tecnica che prevede il calcolo della media dei valori standardizzati, la quale offre un indice di sintesi espresso da un numero puro, svincolato dalla variabilità degli indicatori originari, di facile determinazione. Queste le ragioni del suo largo impiego, giustificato pure dalla possibilità di fornire la base per l'applicazione di metodologie statistiche multivariate.

Il metodo tassonomico offre una valutazione del livello del fenomeno in esame assimilabile ad una scala ad intervalli con origine definita dalla regione ideale. Grosso limite del metodo però è il criterio, spesso arbitrario e soggettivo, di determinazione dell'unità di riferimento. Per quanto invece riguarda il metodo Rizzi, fondato sulla prima componente principale, occorre sottolineare che la validità dei risultati ai quali conduce è funzione del grado di correlazione che essa ha con le variabili originarie: la sua applicazione è, infatti, giustificata soltanto allorché tale grado sia sufficientemente elevato.

Va inoltre ricordato un difetto che caratterizza quei metodi che utilizzano la media: essa, non essendo un indice robusto, risente notevolmente dei valori che nella distribuzione occupano posizioni estreme e può quindi condurre a risultati devianti.

L'aggregazione lineare può essere applicata quando tutti gli indicatori hanno la stessa unità di misura e quando sono state neutralizzati ulteriori ambiguità legate agli effetti della scala di misura. Inoltre implica la piena (e costante) compensabilità: in altre parole prestazioni scadenti rispetto ad alcuni aspetti – misurate dai corrispondenti indicatori – possono essere compensate da valori sufficientemente elevati di altri indicatori.

L'aggregazione geometrica è invece appropriata quando gli indicatori, non comparabili e strettamente positivi, sono espressi in scale diverse; essa comporta compensabilità parziale (non costante), ovvero una compensazione più bassa nei casi in cui l'indicatore composto è costituito da indicatori con valori bassi.

L'assenza di effetti di sinergia o conflitto è una condizione necessaria nell'applicazione di aggregazioni sia lineari che geometriche anche se è comunque difficile da raggiungere.

Il tipo di aggregazione utilizzato è comunque correlato con la tecnica scelta per la normalizzazione dei dati originali. In particolare Ebert and Welsch¹¹⁸ hanno provato che l'uso di aggregazioni lineari permette di ottenere indicatori composti più rilevanti solo se tutti i dati sono espressi in scala di misura di tipo intervallo (come la temperatura) parzialmente comparabile ($f: x \rightarrow \alpha x + \beta_j$ con, ad esempio, α fissato e β_j che varia tra indicatori elementari) o totalmente comparabile (β costante). I dati misurati invece in scala rapporto (come il peso in chilogrammi) $f: x \rightarrow \alpha_j x$ dove $\alpha_j > 0$ (ad esempio con α_j che varia tra indicatori elementari) possono essere aggregate in maniera sensata solo utilizzando funzioni geometriche, previa verifica che x sia strettamente positivo. In altri termini, ad eccezione del caso in cui tutti gli indicatori sono misurati secondo scale rapporto diverse, la scala di misura deve essere la stessa per tutti gli indicatori per poter effettuare l'aggregazione. Per questo motivo deve essere posta molta attenzione quando nello stesso indicatore composto coesistono indicatori valutati secondo scale diverse: per rimuovere l'effetto della scala di misura deve essere propriamente affrontata la fase di normalizzazione.

È utile, a questo punto, trarre alcune osservazioni conclusive. La fase di sintesi, qualunque sia il metodo impiegato, se è fondamentale nelle indagini condotte sugli indicatori sociali, poiché offre un contributo di chiarezza insostituibile, può portare, essendo sempre affidata a metodi che appiattiscono e selezionano l'informazione di base, ad una lettura miope della realtà, in modo particolare se non è sostenuta, a monte, da un'adeguata selezione ed interpretazione degli indicatori elementari. È questa la ragione per cui è indispensabile, per ottenere risultati validi ed affidabili, sostenere la fase di scelta della batteria di indicatori di base con un'impalcatura teorica che definisca la realtà sociale in ciascuna delle sue dimensioni, che possono essere messe bene in luce da un'analisi fattoriale.

Un'ultima considerazione è relativa a quanto ancora c'è da fare in questo campo applicativo. Riteniamo che sia opportuno un nuovo e più adeguato sviluppo delle tecniche di sintesi degli indicatori sociali sostenendo, ad esempio, la produzione di un maggior numero di applicazioni pratiche delle tecniche fondate sulle più articolate e verificate

¹¹⁸ Ebert U. and Welsch H., (2004), Meaningful environmental indices: a social choice approach, Journal of Environmental Economics and Management, vol. 47, pp. 270-283.

metodologie di analisi multivariata dei dati, così da scoprire che la sintesi degli indicatori, oltre ad essere una tecnica descrittiva di impatto immediato, potrebbe rappresentare anche un mezzo efficace per penetrare conoscitivamente la realtà sociale nella sua struttura interna.

2.5.4 La ponderazione¹¹⁹

Centrale alla costruzione di un indicatore composto è la necessità di combinare in modo significativo le diverse dimensioni: questo implica la scelta di una specifica procedura di aggregazione unitamente alla scelta di un adeguato modello di ponderazione. Gli indicatori elementari, necessari alla determinazione di un unico indicatore composto, devono quindi prima essere ponderati: può essere dato a tutti lo stesso peso oppure possono essere assegnati pesi differenti che riflettono il significato, l'affidabilità o altre caratteristiche dei dati sottostanti.

Sono disponibili diverse tecniche di pesatura degli indicatori elementari: quelle basate su modelli statistici (come, ad esempio, l'analisi fattoriale o i modelli delle componenti non osservate) o su metodi partecipativi (ad esempio, l'allocatione di un budget o i processi gerarchici analitici). Per esempio, i pesi possono essere determinati sulla base dei coefficienti di correlazione o sull'analisi delle componenti principali per evitare il problema del doppio conteggio quando due o più indicatori misurano, anche solo parzialmente, uno stesso aspetto. I pesi dovrebbero inoltre riflettere la qualità statistica dei dati: pesi più elevati andrebbero perciò assegnati a dati statisticamente attendibili (con bassa percentuale di valori mancanti, con ampia copertura, con valori robusti). Secondo questa prospettiva, l'intento è quello di premiare solo gli indicatori facili da misurare e prontamente disponibili, penalizzando le informazioni più problematiche da individuare e misurare. Gli indicatori possono anche essere ponderati con pesi basati sull'opinione di

¹¹⁹ L'intero paragrafo è stato tratto dai seguenti lavori:

- Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).
- Freudenberg, M. (2003), Composite indicators of country performance: a critical assessment, OECD, Paris.
- Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).
- http://composite-indicators.jrc.ec.europa.eu/S6_weighting.htm

esperti, i quali, conoscendo le priorità politiche e le interpretazioni teoriche più affermate, possono far riflettere ai pesi la molteplicità di punti di vista dei diversi *stakeholders*.

Il peso attribuito alle diverse variabili influenza pesantemente i risultati finali degli indicatori composti: il rango di un'unità in un data graduatoria può facilmente essere modificato con sistemi di ponderazione alternativi. Per questo motivo i pesi dovrebbero idealmente essere selezionati in base ad un quadro teorico di base o al razionale concettuale relativo all'indicatore composto. La metodologia utilizzata per la determinazione dei pesi deve essere dichiarata in maniera trasparente: questo è utile anche per poter illustrare come risultati e classifiche cambiano con l'uso di sistemi alternativi di ponderazione. Infine è bene tener presente che, indipendentemente dal metodo utilizzato, i pesi sono essenzialmente giudizi di valore e hanno la proprietà di rendere espliciti gli obiettivi che sottendono la costruzione di un indicatore composto.

2.5.4.1 Pesi uguali

In molti indicatori composti viene assegnato lo stesso peso a tutte le variabili per ragioni di semplicità dal momento che è difficile costruire un quadro teorico che porti alla determinazione di sistemi di ponderazione coerenti. Spesso questo accade quando non ci sono motivi statistici o empirici per la scelta di un diverso sistema di ponderazione, comportando il riconoscimento di un uguale status per tutti gli indicatori; secondo una logica interpretativa alternativa, potrebbero invece essere il risultato di un'insufficiente conoscenza dei rapporti di causalità, dell'ignoranza circa il corretto modello da applicare, oppure addirittura la mancanza di consenso su soluzioni alternative.

Questa scelta implica però che tutti gli indicatori elementari abbiano la medesima importanza: non si può trattare, quindi, di una scelta casuale. Applicare pesi uguali non significa infatti non ponderare perché questo approccio implica l'espressione di un giudizio implicito sugli indicatori elementari.

Con l'applicazione di un approccio di ponderazione che assegna pesi uguali ad ogni variabile, nel caso in cui due o più indicatori misurino lo stesso aspetto relativo al fenomeno in esame, vi è il rischio che questo aspetto venga conteggiato più volte. Per evitare tale situazione, un possibile rimedio consiste nel testare la correlazione statistica esistente tra gli indicatori elementari assegnando pesi più bassi ad indicatori fortemente correlati tra loro. Infatti, elevate correlazioni (o esistenza di collinearità) tra indicatori possono mostrare che la *performance* delle unità rispetto all'aspetto considerato sono

simili. Un'ulteriore soluzione consiste nella possibilità di considerare solo gli indicatori elementari con basso grado di correlazione: la minimizzazione del numero di variabili sintetizzate in un indicatore composto è desiderabile anche in termini di trasparenza e parsimonia.

Infine va sottolineato che l'effetto di un'uguale ponderazione dipende anche da come gli indicatori composti sono divisi in categorie o gruppi: pesando in maniera uguale categorie che raggruppano un numero diverso di indicatori elementari, si può mascherare l'assegnazione di pesi diversi ai singoli indicatori semplici.

2.5.4.2 Modelli di regressione multipla

I modelli di regressione lineare possono dirci qualcosa in merito al legame tra un gran numero di indicatori X_1, X_2, \dots, X_m ed un'unica misura di output \hat{Y} che rappresenta l'obiettivo da raggiungere. Viene quindi stimato un modello – usualmente lineare di regressione multipla – per calcolare il peso degli indicatori elementari:

$$\hat{Y} = a + b_1 X_1 + \dots + b_m X_m$$

dove \hat{Y} è la misura del fenomeno espressa dall'indicatore elementare, a è una costante e b_1, \dots, b_m sono i coefficienti di regressione (pesi) associati agli indicatori X_1, X_2, \dots, X_m .

Questo approccio, anche se adatto ad un gran numero di indicatori di diversa tipologia, implica l'assunzione di comportamenti lineari e richiede l'indipendenza degli indicatori esplicativi. Se questi indicatori sono correlati, infatti, gli stimatori avranno alta varianza: ciò significa che i parametri stimati non saranno precisi e le ipotesi testate non saranno convincenti. Nel caso estremo di perfetta collinearità tra gli indicatori, il modello non può nemmeno essere identificato. Inoltre, se il concetto che s'intende misurare può essere rappresentato da un'unica misura \hat{Y}_c , non vi è alcun bisogno di costruire un indicatore composto. Questo approccio potrebbe essere utile per verificare ed aggiustare i pesi oppure qualora l'interpretazione degli indicatori si debba tradurre in possibili azioni politiche.

VANTAGGI

- Può essere usato anche se gli indicatori non sono correlati.
- Non implica alcuna manipolazione dei pesi attraverso restrizioni ad hoc.
- È utile per aggiornare o convalidare l'applicazione di un insieme di pesi.

SVANTAGGI

- Fornisce risultati modesti in caso di forte correlazione tra indicatori per l'esistenza di problemi di multicollinearità; rimedi in tal senso possono essere trovati associando l'analisi di regressione con l'analisi delle componenti principali.
- Richiede una grande quantità di dati per produrre stime con note proprietà statistiche di stabilità e robustezza.

2.5.4.3 Metodi basati su modelli statistici

Nell'assegnazione di pesi uguali per ciascun indicatore, può accadere che, combinando indicatori fortemente correlati tra loro, si introduca un elemento di doppio conteggio nell'indicatore composto finale. Una soluzione in tal senso è stata spesso fornita dalla verifica dell'esistenza di correlazione statistica (ad esempio attraverso il calcolo del coefficiente di correlazione di Pearson) scegliendo i soli indicatori con basso grado di correlazione oppure aggiustando i pesi in modo che ad indicatori correlati fosse attribuito un peso inferiore. Ridurre al minimo il numero di indicatori che vanno a formare l'unico indicatore composto può inoltre essere auspicabile anche per motivi di parsimonia.

Rimarrà comunque quasi sempre una certa correlazione positiva tra le diverse misure dello stesso aggregato. Deve quindi essere utilizzata una regola di tipo empirico, in modo da stabilire la soglia oltre la quale la correlazione comporta un doppio conteggio. Se i pesi dovessero riflettere idealmente il contributo che ciascun indicatore elementare dà all'indicatore composto, il doppio conteggio non può essere determinato solo da un'analisi statistica ma anche da un'analisi di ogni singolo indicatore rispetto al resto degli indicatori e rispetto al fenomeno che essi intendono misurare.

(1) ANALISI DELLE COMPONENTI PRINCIPALI E ANALISI FATTORIALE

L'applicazione di questi due tipi di analisi legate alla costruzione di indicatori composti sono le seguenti:

- identificazione della dimensionalità del fenomeno;
- raggruppamento degli indicatori;
- definizione dei pesi.

L'analisi delle componenti principali e più specificamente l'analisi fattoriale, infatti, raggruppano insieme indicatori collineari per formare un indicatore composto in grado di

catturare la maggior quantità possibile di informazione comune a tali indicatori. Ogni fattore assorbe l'insieme di indicatori che hanno la più alta associazione con esso. L'idea sottostante questo approccio è quella di tener conto della maggior variabilità possibile tra gli indicatori utilizzando il minor numero possibile di fattori. Pertanto l'indicatore non dipende dalla dimensionalità¹²⁰ dei dati ma è invece basata sulle dimensioni "statistiche" interne nei dati. Secondo questo metodo, la ponderazione interviene solo per correggere la sovrapposizione d'informazioni data da due o più indicatori tra loro correlati: i pesi non sono quindi una misura d'importanza degli indicatori a cui sono associati.

Il primo passo consiste nella verifica della struttura di correlazione dei dati: se la correlazione tra gli indicatori è bassa, allora è poco probabile che essi condividano fattori comuni.

Il secondo passo della procedura prevede l'identificazione di un certo numero di fattori latenti, inferiore al numero di indicatori elementari considerati. Ogni fattore dipende da un insieme di coefficienti (punteggi) ciascuno dei quali misura la correlazione tra il singolo indicatore ed il fattore latente. Per estrarre i fattori si utilizza – usualmente - l'analisi delle componenti principali. Per l'analisi fattoriale viene mantenuto solo un sottoinsieme delle componenti principali estratte (diciamo p), ovvero quelle che rappresentano la maggior quota di varianza. Secondo gli standard più diffusi, si scelgono i fattori a cui sono associati autovalori maggiori di 1, oppure i fattori che individualmente contribuiscono a spiegare più del 10% della varianza oppure vengono considerati i fattori che, cumulativamente, spiegano oltre il 60% della varianza totale.

L'operazione successiva – la terza del processo di analisi – riguarda la rotazione dei fattori. La rotazione (solitamente è utilizzata la rotazione secondo il criterio varimax) è usata per minimizzare il numero di indicatori con elevato punteggio sul medesimo stesso fattore. L'idea della trasformazione degli assi fattoriali è quella di ottenere una "struttura più semplice" dei fattori (idealmente una struttura in cui ciascun indicatore fa riferimento ad uno solo dei fattori conservati). La rotazione è un passo standard nell'analisi fattoriale: essa cambia i punteggi fattoriali e, quindi, l'interpretazione dei fattori, lasciando invariate le soluzioni analitiche ottenute ex-ante ed ex-post la rotazione.

L'ultimo passaggio riguarda la costruzione dei pesi dalla matrice dei punteggi fattoriali dopo la rotazione; questo è possibile perché il quadrato dei punteggi fattoriali rappresenta

¹²⁰ Il numero delle componenti principali estraibili è detto dimensionalità.
(Fabbris L. (1997), *Statistica multivariata – Analisi esplorativa dei dati*, McGraw-Hill Libri Italia srl, Milano, pag. 165.)

la proporzione di varianza totale dell'indicatore spiegata dal fattore. Si noti che diversi metodi di estrazione delle componenti principali implicano diversi sistemi di pesi e quindi diversi valori dell'indicatore composto finale (e possibili posizionamenti diversi delle unità nella relativa graduatoria).

VANTAGGI

- Non implicano alcuna manipolazione dei pesi non ponendo su di essi alcuna restrizione ad hoc.
- Risolvono il problema del doppio conteggio.

SVANTAGGI

- Possono essere utilizzati solo con indicatori elementari correlati.
- Sono sensibili alle modifiche dei dati di base: revisioni ed aggiornamenti (ad esempio nuove osservazioni e nuove unità) possono modificare il sistema di pesi individuato (ovvero cambiano la stima dei punteggi fattoriali) utilizzati nell'indicatore composto.
- Sono sensibili alla presenza di outliers, i quali possono introdurre variabilità spuria¹²¹ nei dati.
- Sono sensibili ai problemi derivanti dalle piccole dimensioni del campione e dalla carenza di dati: questi possono rendere difficile l'identificazione o l'interpretazione dei valori risultati.
- Minimizzano il contributo degli indicatori non correlati con altri.
- Sono sensibili alle scelte effettuate nella tecnica di estrazione dei fattori e nei metodi utilizzati per la rotazione.

2.5.4.4 Metodi partecipativi

Gli approcci partecipativi, i quali coinvolgono il giudizio di esperti o giudizi pubblici, vengono spesso utilizzati per la determinazione dei pesi, al fine di esprimere l'importanza relativa degli indicatori dal punto di vista della società. Questi metodi, tuttavia, raggiungono un loro limite quando alcuni indicatori assumono significato minimo o nullo

¹²¹ Il metodo matematico delle componenti principali è estensibile ai fattori latenti in virtù del fatto che i fattori sono ripuliti dalla variabilità ridondante, ossia già compresa in altre variabili osservate, ma anche dalla variabilità spuria, ossia introdotta da variabili che stanno ai margini dei fenomeni fondamentali su cui si fa ricerca ma non hanno diretta attinenza con essi.

(Fabbris L. (1997), *Statistica multivariata – Analisi esplorativa dei dati*, McGraw-Hill, Milano)

per gli intervistati. Ad esempio, mentre un qualsiasi cittadino potrebbe avere un'opinione circa l'importanza dell'aria più pulita o di un ambiente più silenzioso, i pesi assegnati agli indicatori non saranno gli stessi se la persona è invitata a giudicare l'importanza degli ossidi di azoto rispetto alle emissioni di biossido di zolfo. Ovviamente, in tali casi, è necessario il parere di esperti che abbiano un ampio spettro di conoscenze, di esperienze e di approcci, in modo da garantire che venga individuato un adeguato sistema di ponderazione per una determinata applicazione. Essi infatti possono cooperare alla determinazione dei pesi perché hanno sensibilità ed autorevolezza per includere e dare valore e significato ai punti di forza, alle debolezze e alle sfumature dei dati entro un dato contesto teorico.

(1) ALLOCAZIONE DI UN BUDGET

Secondo tale metodo un gruppo di esperti è dotato di un budget di Q punti da distribuire su un certo numero di indicatori, assegnando una quota maggiore di punti agli indicatori la cui importanza l'esperto vuole sottolineare.

L'allocazione del budget si realizza secondo quattro diverse fasi:

- selezione degli esperti per la valutazione;
- attribuzione dei punteggi agli indicatori;
- calcolo dei pesi;
- iterazione dell'assegnazione dei punti fino a convergenza (opzionale).

L'assegnazione di un budget su un numero troppo elevato di indicatori può comportare un forte stress cognitivo per gli esperti: per questo motivo il numero adeguato di indicatori non deve essere superiore a 10.

VANTAGGI

- La ponderazione è basata sull'opinione di esperti e non su manipolazioni tecniche.
- Il parere degli esperti può aumentare la legittimità dell'indicatore composto e creare un forum di discussione attorno al quale formare consenso per le azioni politiche.

SVANTAGGI

- I pesi possono riflettere specifiche condizioni locali (ad esempio in materia di problemi ambientali) per cui il sistema di pesi individuato dagli esperti non è trasferibile in contesti diversi.

- Distribuire un determinato insieme di punti su un numero troppo grande di indicatori diventa pesante per gli esperti; il metodo può produrre incongruenze per un numero di indicatori superiori a 10.
- La ponderazione può non misurare l'importanza di ogni indicatore elementare quanto piuttosto l'urgenza o la necessità di interventi politici relativi alla dimensione misurata dagli indicatori (ad esempio può essere assegnato un elevato peso alle emissioni di ozono qualora l'esperto ritenga che non sia stato fatto abbastanza per abatterle).

(2) SONDAGGI D'OPINIONE

Nei sondaggi di opinione pubblica, le questioni da sottoporre agli intervistati sono selezionate dai problemi all'ordine del giorno, per cui godono già di una certa popolarità dovuta all'attenzione ad essi dedicata dai media.

Da un punto di vista metodologico, i sondaggi d'opinione si fondano sul principio di interesse per cui le persone sono invitate ad esprimere se sono "molto" o "poco" interessati rispetto ai problemi misurati dagli indicatori.

Come per le valutazioni di esperti, il metodo dell'allocatione di un budget può essere applicato anche ai sondaggi d'opinione, anche se risulta più faticoso chiedere al pubblico di distribuire un ammontare fissato di punti ad un insieme di indicatori anziché chiedere loro di esprimere il grado di preoccupazione rispetto ai problemi che gli indicatori misurano.

VANTAGGI

- Riguardano problematiche note alla collettività.
- Permette a tutti gli *stakeholders* di esprimere le loro preferenze, creando così consenso sull'azione politica.

SVANTAGGI

- Implica la misura soggettiva dell'interesse.
- Può produrre incongruenze quando il numero di indicatori è troppo elevato.

2.5.4.5 Processo gerarchico analitico

Il processo gerarchico analitico è ampiamente utilizzato nei processi decisionali multipli ovvero con più alternative. Esso consente la decomposizione di un problema secondo

gerarchia ed assicura che sia gli aspetti qualitativi che quelli quantitativi siano incorporati nel processo di valutazione.

La migliore descrizione di questo metodo probabilmente è quella data da Forman et al.¹²²: "AHP (Analytic Hierarchy Process) è una metodologia di decisione compensativa perché le alternative, efficienti rispetto ad uno o più obiettivi, possono compensare con le loro performance altri obiettivi. Essa consente l'applicazione di dati, esperienza ed intuizione in modo logico ed approfondito entro un sistema gerarchico. In particolare, AHP come metodo di ponderazione permette di derivare i pesi anziché assegnarli in modo arbitrario."

Tale metodo consiste nel raccogliere sistematicamente l'opinione sugli indicatori attraverso confronti a coppie, chiedendo, in primo luogo, quale tra due indicatori è il più importante e domandando, successivamente, in quale misura. L'intensità di preferenza per coppie di indicatori è espressa su una scala semantica da 1 (uguaglianza) a 9 (quando un indicatore è ritenuto 9 volte più importante di quello con cui è messo a confronto). Viene quindi costruita una matrice di confronti tra indicatori; il valore relativo al confronto tra un indicatore e se stesso è 1 (diagonale principale della matrice) mentre il confronto tra l'indicatore X_1 e X_2 sarà l'inverso del confronto fra X_2 e X_1 (se il primo confronto vale, ad esempio, 3, il secondo varrà $1/3$). Ogni giudizio riflette quindi la percezione del rapporto – in termini di contributo relativo – (peso) tra due indicatori. Il peso degli indicatori viene calcolato utilizzando una tecnica che adopera un autovettore, il quale permette di risolvere eventuali incongruenze (come, ad esempio, il caso in cui X_1 è migliore di X_2 , X_2 è meglio di X_3 che è migliore di X_1).

Uno dei vantaggi di questo metodo consiste nel suo essere in grado di verificare la coerenza della matrice di confronto attraverso il calcolo degli autovalori.

Spesso ci si trova però nella situazione in cui non c'è coerenza nelle opinioni espresse dalle persone. Per esempio, se si sostiene che X_1 è molto più importante di X_2 , X_2 è leggermente più importante di X_3 , e X_3 è di poco più importante di X_1 , si è di fronte a giudizi incoerenti e le decisioni prese in questi contesti sono meno affidabili. L'incoerenza, tuttavia, fa parte della natura umana e, quindi, in realtà, è sufficiente per valutare il grado

¹²² "AHP is a compensatory decision methodology because alternatives that are efficient with respect to one or more objectives can compensate by their performance with respect to other objectives. AHP allows for the application of data, experience, insight, and intuition in a logical and thorough way within a hierarchy as a whole. In particular, AHP as weighting method enables decision-maker to derive weights as opposed to arbitrarily assign them."

Forman, E.H. (1983) The Analytic Hierarchy Process as a decision support system, Proceedings of the IEEE Computer Society.

di incoerenza: questa sembra essere l'unica spiegazione del perché i risultati sono difendibili e giustificati di fronte alla comunità.

La metodologia AHP tollera una modesta incoerenza attraverso la misura dell'ammontare della ridondanza. Per una matrice di dimensioni $m \times m$ è necessario effettuare solo $m-1$ confronti per stabilire i pesi da assegnare agli m indicatori. Tuttavia il numero effettivo di confronti eseguiti è pari a $m(m-1)/2$; si tratta di un procedimento analogo alla stima di un valore attraverso il calcolo della media di osservazioni ripetute. Questo si traduce in un insieme di pesi che sono meno sensibili agli errori di giudizio. Inoltre, questa ridondanza consente di misurare gli errori derivanti dai giudizi, fornendo uno strumento di calcolo del rapporto di incongruenza¹²³. Rapporti di incoerenza piccoli (usualmente minori di 0.1 oppure di 0.2) non incidono comunque drasticamente sui pesi.

2.5.4.6 Frontiera di efficienza

La principale caratteristica di questo approccio consiste, secondo Storrie e Bjurek¹²⁴, che ne hanno dato un'approfondita descrizione, nella sua estrema parsimonia relativamente alla ponderazione, in quanto lascia che siano i dati a risolvere il problema dell'individuazione di un adeguato sistema di ponderazione.

Il metodo, definito DEA (Data Envelopment Analysis), utilizza la programmazione lineare per individuare una frontiera di efficienza che viene utilizzata come termine di confronto per valutare il livello del fenomeno di un gruppo di unità. L'insieme dei pesi deriva da questo confronto. Due sono le questioni principali coinvolte in questa metodologia: la costruzione di un parametro di riferimento (la frontiera) e la misurazione della distanza tra unità in un quadro multidimensionale.

La costruzione del parametro di confronto è effettuata assumendo

- che i pesi siano positivi (più elevato è il valore dell'indicatore, migliore è la performance dell'unità corrispondente),
- che non vi sia discriminazione delle unità migliori per ogni singolo indicatore, per cui vengono graduate in modo uniforme, e
- che sia possibile disporre di una combinazione lineare delle migliori performance (questo determina la convessità della frontiera).

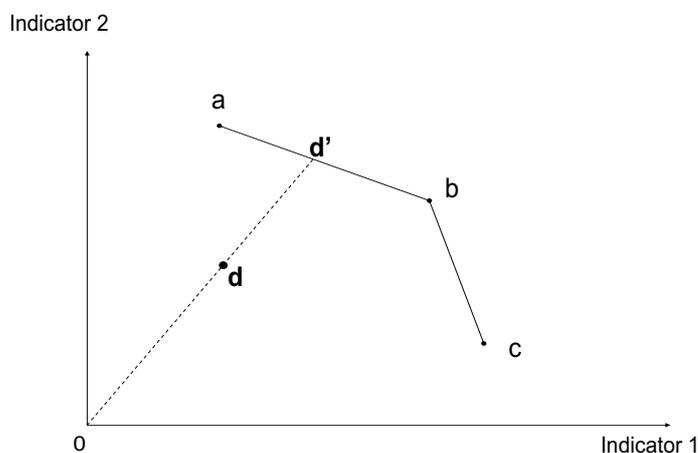
¹²³ Saaty, T. L. (1980) *The Analytic Hierarchy Process*, McGraw-Hill Book Co., N.Y.

Karlsson J. (1998), A systematic approach for prioritizing software requirements, Ph.D. dissertation No. 526, Linköping, Sverige.

¹²⁴ Storrie D., Bjurek H. (2000), Benchmarking European labour market performance with efficiency frontier techniques. Discussion paper FS I 00-211.

La distanza di ciascuna unità dal termine di riferimento è determinata dal posizionamento della stessa e dalla sua immagine sulla linea della frontiera.

Il grafico sotto proposto aiuta la comprensione del metodo esemplificandone l'utilizzo per un insieme di dati relativo a due indicatori misurati per quattro unità (*a*, *b*, *c* e *d*).



I due indicatori sono rappresentati nei due assi e le quattro unità sono classificate in base al valore degli indicatori considerati. La linea che collega le unità *a*, *b* e *c* costituisce la frontiera ed è il termine di riferimento per l'unità *d*, che si trova al di là della frontiera. Le unità di frontiera corrispondono ai risultati migliori, mentre l'unità *d* presenta un livello inferiore del fenomeno in esame. Il grado di rendimento è definito dal rapporto della distanza tra l'origine ed il punto effettivamente osservato e la distanza tra l'origine e la proiezione del punto sulla frontiera: $\overline{0d} / \overline{0d'}$. Il valore dell'indicatore composto per le unità appartenenti alla frontiera sarà pari ad 1, mentre i rendimenti peggiori saranno inferiori a questo valore. L'insieme dei pesi di ciascuna unità dipenderà pertanto dalla sua posizione rispetto alla frontiera.

L'idea, illustrata graficamente in due dimensioni, può essere estesa, in linea di principio, ad un qualsiasi numero di indicatori, rimanendo fissa l'idea di base di una frontiera e della distanza da essa.

(1) APPROCCIO DEL "BENEFICIO DEL DUBBIO"

L'approccio del "beneficio del dubbio" (BOD = Benefit Of the Doubt) è un caso particolare del metodo DEA che consente alle unità di enfatizzare e dare priorità a quegli aspetti per i quali esse mostrano una buona performance (identificazione del target nella frontiera di efficienza). I pesi, in questo approccio, sono "unità-dipendente" e sono sensibili ai valori di riferimento. Il processo di ottimizzazione, se non viene imposta alcuna

restrizione, può portare molti pesi ad assumere valore nullo: in questi casi, molte unità sarebbero considerate come termini di riferimento. L'imposizione di alcune restrizioni sui pesi è perciò necessaria, anche per rendere questo metodo di uso pratico.

Tale tecnica è definita anche come l'applicazione dell'approccio DEA al campo degli indicatori composti ed è stata in origine proposta per valutare performance a livello macroeconomico.

L'indicatore composto è definito come il rapporto tra la performance di un'unità e la performance dell'unità di riferimento:

$$CI_u = \frac{\sum_{j=1}^m I_{ju} w_{ju}}{\sum_{j=1}^m I'_{ju} w_{ju}}$$

dove I_{ju} è il valore normalizzato (con il metodo dei valori relativizzati al campo di variazione) del j -esimo indicatore elementare ($j = 1, 2, \dots, m$) per l'unità u ($u = 1, 2, \dots, N$) e w_{ju} è il peso corrispondente.

L'indicatore di riferimento viene ottenuto come soluzione del seguente problema di massimizzazione (anche se sono possibili determinazioni diverse):

$$I' = I'(w) = \arg \max_{I_k, k \in \{1, \dots, N\}} \sum_{j=1}^m I_{jk} w_j$$

dove I' è il valore dell'unità ipotetica che massimizza la performance complessiva, dato un insieme non noto di pesi w .

Si noti che

- i pesi sono specifici per unità: diversi pesi possono portare a scegliere unità di riferimento diverse dal momento che non vi è un'unità con il punteggio più elevato per tutti gli indicatori elementari;
- il riferimento potrebbe essere, in generale, dipendente dall'unità: potrebbe non esistere un unico parametro di riferimento, a meno che un'unità non sia in assoluto migliore delle altre per tutti gli indicatori;
- gli indicatori devono essere comparabili ovvero devono avere la stessa unità di misura.

Il secondo passo della procedura prevede la specificazione di un sistema di pesi per ciascuna unità; l'insieme di pesi ottimale (se esiste) garantisce il migliore posizionamento in graduatoria dell'unità considerata rispetto a tutte le altre unità del campione. I pesi ottimali si ottengono dalla risoluzione del seguente problema

$$CI'_u = \arg \max_{w_{ju}, j=1, \dots, m} \frac{\sum_{j=1}^m I_{jk} w_{ju}}{\max_{I_k, k \in \{1, \dots, N\}} \sum_{j=1}^m I_{jk} w_{ju}}$$

soggetto al vincolo di non negatività dei pesi.

L'indice composto risultante varierà tra zero (performance più bassa possibile) e 1 (per l'unità ideale di riferimento). Operativamente, l'espressione può essere ridotta ad un problema di programmazione lineare moltiplicando tutti i pesi per un fattore comune (che non altera il valore dell'indicatore) e risolvendo attraverso l'utilizzo di algoritmi di ottimizzazione:

$$CI'_u = \arg \max_{w_{ju}} \left(\sum_{j=1}^m I_{ju} w_{ju} \right)$$

con i vincoli $\sum_{j=1}^m I_{jk} w_{jk} \leq 1$ e $w_{jk} \geq 0$.

VANTAGGI

- L'indicatore sarà sensibile alle priorità politiche relative al fenomeno in esame, in quanto i pesi sono determinati in maniera endogena da quanto effettivamente osservato presso le unità.
- Il parametro di riferimento non è basato su vincoli teorici ma è una combinazione lineare delle prestazioni migliori osservate.
- È un approccio apprezzato nella scena politica dal momento che i responsabili politici non possono lamentare ponderazioni ingiustificate: qualsiasi altro insieme di pesi genererebbe valori inferiori dell'indicatore composto.
- Tale metodo tende ad "incentivare" le unità con performance inferiori anziché "punirle".
- I pesi, rivelando informazioni circa le priorità politiche, possono aiutare a definire i trade-off, superando le difficoltà legate alle aggregazioni lineari.

SVANTAGGI

- Poiché i pesi sono specifici per unità, il confronto incrociato tra unità non è possibile.
- Senza imporre vincoli sui pesi (eccetto la non negatività), la soluzione più probabile è quella di avere per tutte le unità un indicatore composto pari a 1. Quando

esistono invece vincoli sui pesi, può non esserci, per alcune unità, una soluzione al problema di massimizzazione come può esistere una molteplicità di soluzioni che rendono indeterminato l'insieme ottimale di pesi.

- Può accadere che l'indicatore premi lo status-quo delle unità in quanto, per ciascuna unità, il problema di massimizzazione assegna pesi superiori alle misure del fenomeno più alte.
- Con una tecnica di ponderazione endogena si rischia di sostituire l'opinione di esperti con la manipolazione dei pesi da parte dell'analista (attraverso l'imposizione di vincoli): in tal caso viene a cadere la trasparenza della procedura operativa applicata.
- Il valore della valutazione dipende dal valore dell'unità ideale di riferimento: se questo cambia, cambierà anche il valore dell'indicatore composto così come l'insieme dei pesi (e quindi la graduatoria).
- L'unità con la misura migliore del fenomeno in esame non vedrà riflesso del suo eventuale sviluppo nell'indicatore composto (il suo apporto resta infatti unitario): questo inconveniente può essere risolto con l'imposizione di un riferimento con cui confrontare le unità esterne all'insieme esaminato.

2.5.4.7 Componenti non osservate

Nel modello delle componenti non osservate (UCM = Unobserved Components Model) i pesi sono ottenuti stimando, attraverso il metodo della massima verosimiglianza, una funzione che combina gli indicatori elementari. Questi ultimi sono assunti dipendenti da una variabile non osservata più un termine di errore. La stima della componente sconosciuta pone quindi luce sulla relazione tra l'indicatore composto e le sue componenti. Il peso ottenuto sarà determinato in modo da ridurre al minimo l'errore nell'indicatore composto. Questo metodo è simile all'analisi di regressione: la principale differenza risiede nella variabile dipendente, che in questo modello non è nota.

Sia ph_u il fenomeno ignoto da misurare. I dati osservati sono costituiti da gruppi di indicatori; ciascun gruppo misura un aspetto del fenomeno ph_u . Il valore osservato dell'unità u per l'indicatore j , I_{uj} , può essere scritto come funzione lineare del fenomeno non osservato e di un termine di errore ε_{uj} :

$$I_{uj} = \alpha_j + \beta_j [ph_u + \varepsilon_{uj}]$$

dove α_j e β_j sono parametri non noti che mappano ph_u in I_{uj} .

Il termine di errore cattura due fonti di incertezza nel rapporto tra il fenomeno e gli indicatori. Anzitutto il fenomeno potrebbe essere misurato od osservato in maniera imperfetta per ciascuna unità (ad esempio a causa di errori di misurazione). In secondo luogo il rapporto tra ph_u e I_{uj} è imperfetto: se vi sono differenze tra unità rispetto a quanto l'indicatore rileva, esso potrebbe essere solo una misura connotata da "rumore" del fenomeno). Il termine di errore ε_{uj} si assume avere media nulla, $E[\varepsilon_{uj}] = 0$, e la stessa varianza tra unità rispetto ad un determinato indicatore (ma diversa varianza tra indicatori diversi), $E[\varepsilon_{uj}^2] = \sigma_u^2$; inoltre $E[\varepsilon_{uj} \varepsilon_{vk}] = 0$ con $u \neq v$ o $j \neq k$.

Il termine di errore è assunto indipendente tra gli indicatori elementari dal momento che ciascun indicatore misura idealmente una dimensione del fenomeno indipendente dalle altre. Inoltre, per facilitare l'aspetto computazionale, si assume spesso ph_u come variabile casuale con media zero e varianza unitaria; gli indicatori sono ri-scalati nell'intervallo [0, 1]. L'assunzione che sia ph_u che ε_{uj} siano congiuntamente distribuite in modo normale semplifica la stima del livello del fenomeno per la singola unità u , dato dalla media della distribuzione condizionata della componente non osservata (una volta che il valore osservato è stato opportunamente riscalato):

$$E\left[\frac{ph_u}{I_{u1}, \dots, I_{uN}}\right] = \sum_{j=1}^m w_{uj} \frac{I_{uj} - \alpha_j}{\beta_j} . \quad [1]$$

I pesi si ottengono dalla seguente formula:

$$w_{uj} = \frac{\sigma_j^{-2}}{1 + \sum_{j=1}^m \sigma_j^{-2}} .$$

Il peso w_{uj} è una funzione decrescente della varianza dell'indicatore j (esprimendo l'idea che minore è la precisione del j -esimo indicatore, minore sarà il peso ad esso assegnato) ed è funzione crescente della varianza degli altri indicatori. Il peso w_{uj} dipende dall'unità presa in considerazione: dipende dalla varianza dell'indicatore elementare j (numeratore) e dalla somma delle varianze di tutti gli indicatori, incluso j (denominatore). Tuttavia, dal momento che non tutte le unità dispongono di dati per tutti gli indicatori, il denominatore di w_{uj} può essere specifico per unità: questo può produrre una non comparabilità, come avviene nell'utilizzo dell'approccio BOD. Ovviamente ogni qual volta l'insieme di indicatori è uguale per tutte le unità, i pesi non saranno più specifici per unità e la comparabilità tra

unità sarà assicurata. La varianza della distribuzione condizionale è dato dalla seguente formula:

$$\text{Var} \left[\frac{ph_u}{I_{u1}, \dots, I_{uN}} \right] = \left[1 + \sum_{j=1}^m \sigma_j^{-2} \right]^{-1}$$

e può essere vista come una misura di precisione dell'indicatore composto, utile per la costruzione di intervalli di confidenza.

La stima del modello è facilitata dall'assunzione di normalità per ph_u e per ε_{uj} . La funzione di verosimiglianza dei dati osservati viene massimizzata rispetto agli ignoti parametri α_j , β_j e σ_j^2 ; le stime dei loro valori sostituite nell'equazione [1] permettono di ottenere il valore dell'indicatore composto e dei pesi.

VANTAGGI

- I pesi non dipendono da restrizioni ad hoc.

SVANTAGGI

- L'affidabilità e la robustezza dei risultati dipende da una disponibilità di dati sufficiente.
- Con indicatori fortemente correlati potrebbero sorgere problemi di identificazione: il metodo, infatti, funziona in modo migliore con indicatori indipendenti.
- Il metodo premia l'assenza di outliers nei dati, poiché i pesi sono funzione decrescente della varianza degli indicatori elementari.
- Nel caso in cui per ogni unità si abbia a disposizione un numero di indicatori elementari diverso, i pesi sono specifici per unità.

2.5.4.8 Altre tecniche di ponderazione

Riportiamo in questo paragrafo ulteriori tecniche di determinazione del sistema di ponderazione per un insieme di indicatori elementari. Si tratta di metodi meno diffusi dei precedenti ai quali, anche la letteratura, riserva un interesse minore.

(1) ANALISI CONGIUNTA

Chiedendo semplicemente ad un gruppo di rispondenti quanta importanza essi attribuiscono ad un indicatore, è improbabile ottenere effettive valutazioni sulla loro "disponibilità a pagare" ovvero non è chiaramente determinabile quanto ciascuno è in

grado di "scommettere" su un indicatore anziché su un altro. Queste valutazioni possono essere dedotte mediante l'analisi congiunta (CA = Conjoint Analysis) dalle graduatorie che gli intervistati offrono per scenari alternativi.

Si tratta di una tecnica di analisi dei dati multivariata decompositiva frequentemente utilizzata nelle ricerche di marketing e sui consumatori: essa richiede una valutazione (preferenza) su una serie di scenari alternativi (uno scenario può essere pensato come un insieme di valori per gli indicatori). Questa preferenza viene poi decomposta ricavando il valore degli indicatori di un determinato scenario. Se la tecnica AHP determina il "valore" di un'alternativa addizionando il valore dei singoli indicatori, la CA, al contrario, disaggrega le preferenze.

Sebbene questo metodo utilizzi l'analisi statistica per il trattamento di dati, esso si realizza attraverso il contributo di persone (esperti, politici, cittadini) alle quali viene chiesto di scegliere un insieme di indicatori secondo criteri di preferenza. Il valore assoluto (o livello) degli indicatori potrebbe variare sia all'interno dell'insieme scelto da un individuo sia tra individui. La stima della funzione di preferenza è ottenuta utilizzando le informazioni provenienti dai diversi scenari: la probabilità di una preferenza può quindi essere stimata in funzione del livello degli indicatori che definiscono i diversi scenari alternativi

$$pref_u = P(I_{1u}, I_{2u}, \dots, I_{mu})$$

dove I_{ju} è il valore dell'indicatore elementare j per l'unità u .

Dopo la stima di questa probabilità (spesso utilizzando modelli di scelta discreti), gli indicatori della funzione di preferenza possono essere utilizzati come pesi per aggregare gli indicatori elementari in un indicatore composto:

$$CI_u = \sum_{j=1}^m \frac{\partial P}{\partial I_{ju}} I_{ju} .$$

L'idea è quella di calcolare il differenziale totale della funzione P presso il punto di indifferenza tra stati alternativi. Risolvendo l'equazione rispetto all'indicatore j , si ottiene il tasso marginale di scambio di I_{ju} : $\frac{\partial P}{\partial I_{ju}}$ (quindi il peso) indica quindi un trade-off (come la preferenza cambia al variare dell'indicatore). Ciò implica la compensabilità tra indicatori, ovvero la possibilità di compensare carenze in una certa dimensione con eccellenti performance in altre dimensioni. Come per altri approcci sopra descritti, questa è una caratteristica importante e deve essere attentamente valutata rispetto agli obiettivi dell'intera analisi.

VANTAGGI

- Si ottengono pesi con il significato di trade-off.
- Tiene conto del contesto socio-politico e del valore dato dai rispondenti.

SVANTAGGI

- Necessita della specificazione, a priori, di una funzione di utilità ed implica la compensabilità.
- Dipende dal campione di rispondenti scelto e dalla formulazione delle domande.
- Richiede un grande campione di intervistati e ad ogni partecipante può essere richiesto di esprimere un elevato numero di preferenze.
- Il processo di stima è complesso.

(2) NEUTRALIZZAZIONE DELLA CORRELAZIONE

Questo metodo viene utilizzato per aggregare in maniera graduale indicatori elementari nel tentativo di rimuovere l'effetto della correlazione tra essi esistenti.

Riportiamo la procedura da applicare nel caso di aggregazione di tre indicatori: essa si basa sull'idea di base di considerare la correlazione esistente tra coppie di indicatori.

Supponiamo quindi di avere tre indicatori X_1 , X_2 e X_3 ed ipotizziamo che i primi due siano fortemente correlati. Il primo passo consiste nello standardizzare gli indicatori (sottraendo la loro media e dividendo per la deviazione standard); indicheremo quindi i due indicatori standardizzati correlati con Z_1 e Z_2 mentre Z_3 corrisponderà alla standardizzazione di X_3 . Viene quindi calcolato un sotto-indicatore Y come media di Z_1 e Z_2 secondo la formula seguente:

$$Y = \left[2(1 + r_1) \right]^{-\frac{1}{2}} (Z_1 + Z_2) = \frac{Z_1 + Z_2}{\sqrt{2(1 + r_1)}}$$

dove r_1 indica il coefficiente di correlazione tra Z_1 e Z_2 . Il passaggio successivo prevede quindi l'aggregazione di Y e di X_3 .

$$Z = \left[2(1 + r_2) \right]^{-\frac{1}{2}} (Y + Z_3) = \frac{Y + Z_3}{\sqrt{2(1 + r_2)}}$$

dove r_2 indica il coefficiente di correlazione tra Y e Z_3 .

Procedendo in modo analogo a quanto presentato, la procedura è estendibile ad un numero più elevato di indicatori.

(3) DISTANZA DAGLI OBIETTIVI

Un modo per evitare l'immediata selezione dei pesi è quello di misurare la necessità di un intervento politico e l'urgenza di un problema attraverso la distanza da un prefissato obiettivo. L'urgenza è tanto più elevata tanto più si è lontani dagli obiettivi.

La ponderazione si realizza dividendo l'indicatore elementare per il corrispondente valore obiettivo, entrambi espressi nella stessa unità di misura. I parametri che si ottengono attraverso l'applicazione di questo metodo possono essere sintetizzati, al fine di produrre l'indicatore composto, da una semplice media.

L'utilizzo dei target politici come valori obiettivo convince i responsabili delle decisioni politiche sulla "solidità" del metodo di ponderazione, almeno fintantoché tali obiettivi sono stati definiti. Questo approccio è tecnicamente realizzabile solo quando vi è una base politica ben definita relativa ad un certo problema o argomento. Un'argomentazione contraria all'uso di tale metodo risiede nel fatto che i benefici di una determinata politica debbono essere valutati indipendentemente dagli obiettivi politici esistenti: un'alternativa può essere quella di sostituire i target politici con livelli di sostenibilità, con effetti quantificabili sull'ambiente o con la performance migliore tra tutte le unità.

(4) PESI BASATI SULLA QUALITÀ

Un ulteriore approccio prevede l'assegnazione di pesi minori alle variabili che soffrono di più della presenza di valori mancanti nel tentativo di correggere parzialmente i problemi relativi ai dati. L'affidabilità di un indicatore composto può essere migliorata attribuendo peso maggiore alle componenti di qualità maggiore e con disponibilità di dati più ampia. Tuttavia, questo può sviare la lettura delle graduatorie finali degli indicatori composti verso i fattori che risultano più semplici, più facilmente identificabili e di più immediata interpretazione.

2.5.4.9 Considerazioni conclusive

La scelta del metodo di determinazione del sistema di pesi da applicare agli indicatori elementari al fine di costruire un indicatore composto deve tenere conto della natura dei dati e degli obiettivi dell'analisi.

L'applicazione di pesi uguali a ciascun indicatore può essere applicata dopo un'appropriata normalizzazione degli indicatori stessi. Tale scelta in termini di ponderazione – e ricordiamo che si tratta di una scelta ben precisa, non di una “non scelta” - funziona in maniera ottimale nei casi in cui tutti gli indicatori elementari sono assolutamente incorrelati oppure quando, al contrario, sono fortemente correlati. Tuttavia, quando sono coinvolti nel processo di analisi solo pochi indicatori fortemente correlati, questo metodo, pur semplice, non diventa il migliore strumento di supporto all'aggregazione.

I modelli di regressione multipla sono in grado di gestire un elevato numero di indicatori elementari. Questo approccio può essere applicato nei casi in cui gli indicatori, considerati come input per il modello (variabili indipendenti), sono legati a diverse azioni politiche; il risultato del modello assume il ruolo di obiettivo (variabile dipendente). Un uso alternativo di questo metodo può essere legato a scopi di previsione. Permane, tuttavia, l'incertezza che le relazioni desunte dal modello di regressione, per un determinato insieme di indicatori di input e di output, possano essere valide per insiemi di indicatori diversi.

L'analisi delle componenti principali è una tecnica esplorativa molto interessante per esaminare la struttura della correlazione esistente in un insieme di variabili. Nello sviluppo di indicatori composti, l'applicazione dell'analisi delle componenti principali può identificare le dimensioni sottostanti i dati e/o definire i pesi per gli indicatori elementari. L'analisi fattoriale, invece, è solitamente utilizzata come un metodo supplementare, al fine di esaminare approfonditamente le relazioni tra gli indicatori di base. Tuttavia, vi sono importanti problemi legati a queste tecniche. In primo luogo i pesi assegnati agli indicatori elementari sono basati sulla correlazione tra essi esistente, la quale non corrisponde necessariamente alle reali relazioni esistenti tra gli indicatori semplici ed il fenomeno che si intende misurare: in altre parole c'è confusione tra correlazione e causalità. Non è possibile conoscere (o stimare) il peso reale fintanto che si ha bisogno di una variabile dipendente: se infatti esistesse una variabile dipendente soddisfacente, non ci sarebbe alcun bisogno di costruire un indicatore composto. Inoltre è non consigliabile utilizzare l'analisi delle componenti principali quando gli indicatori di base si riferiscono a cicli temporali diversi: questo può ridurre l'affidabilità dell'indicatore composto perché alcuni indicatori possono indicare performance migliori in un ciclo ed altri in un ciclo diverso¹²⁵.

¹²⁵ Nilsson R. (2000) .Calculation of Composite leading Indicators: A comparison of two different methods., Paper presented at the CIRET conference, Paris, October 2000.

I metodi partecipativi che richiedono il giudizio di esperti necessitano il riunire un gruppo di persone esperte con ampio spettro di conoscenze e di esperienze, per garantire la determinazione di un adeguato sistema di ponderazione. L'allocazione del budget è una tecnica ottimale per un massimo di 10-12 indicatori elementari; se il numero di misure di base è elevato, questo metodo può provocare forte stress cognitivo agli esperti ai quali è chiesto di distribuire un punteggio totale tra gli indicatori.

I sondaggi di opinione sono stati impiegati diffusamente per molti anni per la determinazione dei pesi. Le tematiche trattate sono selezionate tra problematiche già note e discusse dall'opinione pubblica. In molti casi, i sondaggi d'opinione, in diversi Paesi ed in diversi anni hanno portato a schemi di ponderazione simili relativamente ad alcuni problemi ambientali: questo indica che l'opinione pubblica, circa le principali minacce per l'ambiente, è notevolmente stabile, sia nello spazio che nel tempo. Alla luce di questo, i timori che l'opinione pubblica valuti, ad esempio, le questioni ambientali su una base irrazionale, e che quindi i pesi basati su tale metodologia siano instabili, sembrano perciò essere infondati.

Il processo gerarchico analitico è una tecnica ampiamente utilizzata nei processi decisionali multipli; come metodo di ponderazione permette ai decisori – politici – di derivare i pesi anziché assegnarli arbitrariamente. Un vantaggio di tale metodologia consiste nel fatto che, a differenza di molti altri metodi basati sulla teoria dell'utilità, il suo uso a fini di confronto non richiede una scala di riferimento. Inoltre questo metodo tollera una certa entità di incoerenza nelle risposte delle persone valutata attraverso la ridondanza (per cui sono disponibili più equazioni rispetto al numero di pesi da definire): i pesi risultanti sono perciò meno sensibili agli errori di valutazione. Questi vantaggi rendono i pesi derivanti dall'applicazione della tecnica AHP difendibili e giustificabili di fronte all'opinione pubblica.

L'approccio della frontiera di efficienza è estremamente parsimonioso per quanto riguarda gli assunti della ponderazione, in quanto lascia che siano i dati a decidere il sistema dei pesi da assegnare agli indicatori elementari. Si tratterebbe, però, di un approccio empirico che potrebbe non indicare la giusta direzione verso cui intervenire con un'azione politica, ad esempio, in una data area territoriale al fine di migliorarne la situazione.

Nel modello delle componenti non osservate i pesi sono ottenuti stimando, attraverso il metodo della massima verosimiglianza, una funzione che combina gli indicatori elementari.

Questi ultimi sono assunti dipendenti da una variabile non osservata più un termine di errore: la stima della componente sconosciuta mette in luce la relazione tra l'indicatore composto e le sue componenti. Il peso ottenuto sarà determinato in modo da ridurre al minimo l'errore nell'indicatore composto. Questo metodo è simile all'analisi di regressione: la principale differenza risiede nella variabile dipendente, che in questo modello non è nota.

La neutralizzazione degli effetti della correlazione è un metodo adatto ai casi in cui gli indicatori elementari da combinare sono pochi, alcuni dei quali fortemente correlati o forme diverse del medesimo problema. Si tratta di un metodo empirico che ha avuto applicazioni limitate.

Un modo per evitare l'immediata selezione dei pesi è quello di misurare la necessità di un intervento politico e l'urgenza di un problema attraverso la distanza da un determinato obiettivo. Questo approccio è tecnicamente realizzabile solo quando c'è una base ben definita relativamente ad una certa politica. L'utilizzo di tale tecnica vede come principale contro-argomentazione l'indipendenza che deve esistere tra i risultati ottenuti da una determinata politica e gli obiettivi politici esistenti.

2.5.5 Riflessioni finali

Dopo aver descritto il processo di costruzione di un indicatore composto, ci siamo soffermati in maniera dettagliata ed approfondita ad esaminare quelli che abbiamo definito "passi chiave": abbiamo infatti focalizzato l'attenzione sulle tecniche di normalizzazione degli indicatori elementari e sulle metodologie di aggregazione e di ponderazione degli stessi.

In questo ultimo paragrafo vogliamo raccogliere alcuni aspetti che riteniamo possano fornire un contributo prezioso e complementare a quanto finora visto relativamente all'analisi degli indicatori.

2.5.5.1 Analisi della correlazione tra indicatori¹²⁶

Oltre all'analisi delle componenti principali e all'analisi fattoriale – precedentemente descritte –, un altro modo per studiare il grado di correlazione di un insieme di indicatori elementari è quello di utilizzare un coefficiente di affidabilità (o coerenza) chiamato alpha (α) di Cronbach. Questo coefficiente valuta quanto bene un insieme di variabili (o indicatori) misura lo stesso costrutto sottostante. L'alfa di Cronbach può essere scritto come una funzione del numero m di indicatori e la correlazione media \bar{r} esistente tra gli indicatori:

$$\alpha = \frac{p\bar{r}}{1 + (p-1)\bar{r}}$$

Una formulazione diversa e più complessa del coefficiente, lo definisce come segue:

$$\alpha_u = \frac{m}{m-1} \frac{\sum_{i \neq j} \text{Cov}(x_i, x_j)}{\text{Var}(x_0)} = \frac{m}{m-1} \left(1 - \frac{\sum_j \text{Var}(x_j)}{\text{Var}(x_0)} \right)$$

con $u = 1, 2, \dots, N$ e $j = 1, 2, \dots, m$; il numero di unità considerate è N mentre m indica il numero di indicatori elementari. Inoltre $x_0 = \sum_{j=1}^m x_j$ corrisponde alla somma di tutti gli indicatori. Il coefficiente α di Cronbach misura la proporzione di variabilità totale del campione di indicatori dovuta alla correlazione tra gli indicatori stessi. Esso cresce all'aumentare del numero di indicatori ed al crescere della covarianza di ogni coppia di essi. Se non esiste correlazione e gli indicatori sono indipendenti, il coefficiente vale 0, mentre assume valore 1 quando gli indicatori sono perfettamente correlati. Un valore del coefficiente maggiore o uguale a 0.80 è interpretato, nella maggior parte delle applicazioni, come evidenza che gli indicatori misurano lo stesso costrutto. Se invece il valore del coefficiente è basso per un determinato insieme di indicatori, ciò significa che i dati sono in realtà multidimensionali.

Il coefficiente di Cronbach non è una statistica di test ma è un coefficiente di affidabilità basato sulle correlazioni tra indicatori: un valore elevato potrebbe significare che gli indicatori misurano il medesimo costrutto. Nonostante sia spesso interpretato come tale, il coefficiente α di Cronbach non è strettamente una misura di unidimensionalità; un

¹²⁶ Per la stesura di questo paragrafo sono stati utilizzati i due seguenti documenti:

- Saisana M., Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).
- http://composite-indicators.jrc.ec.europa.eu/S3_multivariate_analysis.htm

insieme di indicatori può determinare un elevato valore di α ma essere comunque multidimensionale. Questo accade quando distinti gruppi di indicatori elementari (dimensioni separate) sono caratterizzati da un'elevata correlazione interna ma i gruppi non sono tra loro altamente correlati. Relativamente al valore auspicabile per il coefficiente, alcuni autori suggeriscono il valore 0.70 come soglia accettabile di affidabilità mentre altri, più indulgenti, propongono 0.60; in generale, comunque, il valore accettabile varia a seconda del campo di applicazione.

2.5.5.2 Aggregazione e ponderazione: un matrimonio obbligato¹²⁷

L'aggregazione delle informazioni, portate dalle diverse dimensioni dentro ad un unico indicatore composto, va in coppia con la questione relativa alla loro ponderazione.

Sia nelle aggregazioni lineari che in quelle geometriche, i pesi esprimono il trade-off tra indicatori: la carenza in una dimensione può quindi essere compensata da un sovrappiù in un'altra. Ciò implica una contraddizione tra la concezione usuale dei pesi (ovvero come misura d'importanza della variabile associata) e l'effettivo significato che essi assumono quando si utilizzano aggregazioni geometriche o lineari. In un'aggregazione lineare la compensabilità è costante, mentre nelle aggregazioni geometriche la compensabilità è inferiore per gli indicatori composti con bassi valori.

In termini di scelte pertinenti all'azione politica, quando viene utilizzata l'aggregazione geometrica, se è ammessa la compensabilità (come nel caso di puri indicatori economici), un'unità con una bassa valutazione su un indicatore avrà bisogno di un valore molto più elevato sugli altri per migliorare la propria situazione. Per questo motivo nelle analisi comparative le unità con valori bassi prediligono tecniche di aggregazione di tipo lineare piuttosto che geometrico. D'altro canto, nell'aggregazione geometrica l'utilità marginale derivante da un aumento in termini assoluti del valore di indicatori bassi sarebbe molto superiore a quella derivante da un incremento assoluto di valori già elevati. Di conseguenza, quando viene utilizzata un'aggregazione geometrica anziché una lineare, un'unità, quale, ad esempio un Paese, sarebbe più interessata ad aumentare i settori e/o le attività con punteggio inferiore, al fine di avere la massima possibilità di migliorare la propria posizione nella classifica finale.

¹²⁷ Per la stesura di questo paragrafo sono stati utilizzati i due seguenti documenti:

- Nardo, M. *et al.* (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.
- Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).

Se si vuole garantire che i pesi rimangano una misura d'importanza, devono quindi essere utilizzati altri metodi di aggregazione, in particolare metodi che non consentano la compensabilità. Inoltre, se differenti obiettivi sono ugualmente legittimi ed importanti, è necessario garantire una logica di non compensazione tra indicatori, come nel caso di indici ambientali che includono dimensioni diverse quali quella fisica, quella sociale e quella economica. L'approccio multicriterio non compensativo potrebbe garantire la compensabilità trovando un compromesso tra due o più obiettivi. Nella sua forma base, questo approccio non premia gli outliers dal momento che conserva solo informazioni di tipo ordinale. Questo metodo, tuttavia, potrebbe essere computazionalmente costoso quando il numero di unità è elevato dal momento che il numero di permutazioni da calcolare aumenta in modo esponenziale.

L'assenza di una via "obiettiva" di determinazione di pesi e di metodi di aggregazione non conduce necessariamente al rifiuto della validità degli indicatori composti, almeno fino a quando è garantita la trasparenza dell'intero processo operativo. Gli obiettivi di ricerca devono pertanto essere chiaramente indicati all'inizio dell'analisi ed il modello scelto deve essere controllato per verificare fino a che punto esso soddisfa gli intenti del ricercatore.

2.5.5.3 Il significato dei pesi¹²⁸

Si sta assistendo ad una proliferazione della produzione di indicatori composti da parte delle maggiori organizzazioni internazionali: questo è un chiaro sintomo dell'importanza politica e della rilevanza operativa che essi hanno nella programmazione politica. Come conseguenza, la ricerca tesa a migliorare il modo in cui questi indicatori vengono costruiti sembra essere molto importante sia dal punto di vista teorico che operativo.

Focalizzando l'attenzione sulla fase di ponderazione, è necessario chiarire il significato che viene attribuito ai pesi. Esiste infatti incoerenza tra il significato teorico dei pesi ed il significato che è loro generalmente attribuito dall'utilizzo pratico che se ne fa nella costruzione di indicatori composti, generando, in modo ricorsivo, in molte applicazioni empiriche, un'importante alterazione di senso che deriva dalla debolezza tecnica degli indicatori composti dovuta all'incoerenza tra formule di aggregazione matematica e significato dei pesi.

¹²⁸ Munda G., Nardo M. (2005), Constructing consistent composite indicators: the issue of weights, Institute for the Protection and Security of the Citizen.

A illustrazione di questa affermazione, è opportuno ricordare che, in generale, un indicatore composto I , utilizzato spesso per definire una graduatoria di un insieme di unità, può essere considerato un'aggregazione lineare pesata applicata a un set di variabili:

$$I = \sum_{i=1}^N w_i x_i$$

dove x_i indica l'indicatore normalizzato mentre i pesi sono identificato come w_i ; questi ultimi, usualmente, sommano a 1 ed il loro valore appartiene all'intervallo $[0,1]$.

La regola comunemente seguita per l'assegnazione dei pesi stabilisce che il peso maggiore sia attribuito alle componenti considerate più rilevanti nel contesto del particolare tema affrontato¹²⁹. Il concetto di importanza è però incompatibile con i metodi di aggregazione lineari: in essi, infatti, i pesi assumono il significato di trade-off a causa della logica compensativa ad essi sottostante.

Chiaramente i trade-off possono essere valutati solo nel caso in cui siano noti i punteggi quantitativi delle variabili coinvolte; al contrario, invece, il concetto di importanza è connesso alla variabile stessa e non alla sua quantificazione. Per questo motivo, quindi, l'interpretazione dei pesi come misura della percezione dell'importanza di ogni singola dimensione contenuta nel concetto è sempre inappropriato quando i pesi dipendono dal campo di variazione dei valori assunti dalle variabili, come accade nel contesto delle aggregazioni lineari.

L'utilizzo dei pesi come coefficienti di importanza è più difficile rispetto al valutarli come trade-off: questa è probabilmente la maggior ragione per cui in pratica si tende a usare i pesi attribuendo loro il significato di importanza, nonostante questa pratica non sia difendibile dal punto di vista teorico. Affinché i pesi siano interpretati come "coefficienti d'importanza" devono essere utilizzate procedure di aggregazione non compensative.

2.5.5.4 Confronto dei risultati¹³⁰

Una volta elaborati i dati raccolti su un insieme di unità statistiche relativamente ad un certo numero di indicatori per misurare un dato fenomeno, spesso il ricercatore ordina le unità secondo una graduatoria. L'utilizzo di metodi di normalizzazione, aggregazione e

¹²⁹ Freudenberg, M. (2003), *Composite indicators of country performance: a critical assessment*, OECD, Paris.

¹³⁰ L'intero paragrafo è stato tratto dai seguenti contributi:

- Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari.
- Fabbris L., Capiluppi C., Giancotti G., Meneghello A. (1993), *Statree 1.0 – Sistema esperto per la scelta del metodo di analisi statistica*, Manuale per l'uso, Edizioni Summa, Padova.
- Attanasio M., Capursi V. (1997). *Graduatorie sulla qualità della vita: prime analisi di sensibilità delle tecniche adottate*. Atti XXXV Riunione Scientifica SIEDS, Alghero.

ponderazione diversi produce, naturalmente, graduatorie diverse. Si rende pertanto necessario valutare secondo un qualche metodo matematico – e quindi oggettivo – la distanza tra graduatorie diverse ovvero misurare la loro cograduazione. Si tratta, in altre parole, di verificare se le graduatorie ottenute con diversi metodi concordano o meno.

Per fare questo ci si può avvalere, ad esempio, di due indici statistici: l'uno di Spearman che misura la concordanza tra graduatorie ed il secondo di Kendall che misura la concordanza tra più graduatorie.

Quando codomini e punteggi finali degli indicatori sono diversi – perché diverse sono le tecniche applicate per la determinazione del valore dell'indicatore composto finale –, il confronto relativo alla posizione della medesima unità in due graduatorie diverse può essere fatto solo in termini di ranghi. L'indice ρ di cograduazione di Spearman, che varia tra -1 (massima discordanza) e 1 (massima concordanza) si ottiene come segue:

$$\rho = 1 - \frac{6 \sum_{i=1}^n \delta_i^2}{n(n^2 - 1)}$$

dove δ_i indica la differenza tra i ranghi attribuiti all'unità i in due graduatorie ed n è il numero di unità considerate nell'analisi. In particolare la distanza risolve la formula seguente:

$$\delta_i = \text{rank}\{x_i\} - \text{rank}\{y_i\}$$

dove i due termini della differenza sono rispettivamente il rango dell'unità i -esima nella graduatoria X e nella graduatoria Y . L'indice misura quanto gli ordinamenti delle unità per le due graduatorie concordano o differiscono.

L'indice W di cograduazione multipla di Kendall misura invece la concordanza esistente tra k graduatorie (paragonabili a k osservatori diversi) che ordinano n unità. Il valore di W si ottiene mediante la seguente formula:

$$W = \frac{12 \sum T_i^2}{k^2 n(n^2 - 1)} - \frac{3(n+1)}{n-1}$$

dove T_i coincide con la somma dei ranghi attribuiti all'unità i -esima nella k graduatorie.

Qualora ci siano ranghi ripetuti – ovvero a più unità è attribuita la stessa posizione d'ordine – la formula deve essere corretta e diventa la seguente:

$$W = \frac{12 \sum T_i^2 - 3(n+1)^2}{k^2 n(n^2 - 1) - n \sum C_j}$$

dove $C_j = \sum t_h (t_h^2 - 1)$ è il fattore di correzione da apportare per tener conto dei ranghi ripetuti nella graduatoria j -esima e t è il numero di volte in cui è ripetuto un rango. Esso il coefficiente W di Kendall varia tra 0 e 1: in particolare assume un valore tra 0 e $1/n$ nel caso di discordanza mentre ha valori compresi tra $1/n$ e 1 se c'è concordanza.

Spesso – e ciò è auspicabile – alcune posizioni nella graduatoria delle unità rimangono stabili a prescindere dal metodo di aggregazione applicato agli indicatori elementari: questo fa supporre che le posizioni di unità con caratteristiche assai distinte emergano comunque, mentre le differenze più grandi possono essere attribuite a situazioni non ben definite e/o non correttamente rilevate a livello di indicatori elementari.

2.5.5.5 Con "senso di misura"¹³¹

Secondo Land¹³², un indicatore assume significato solo quando possiede un valore informativo all'interno di un modello teorico comunque definito – matematicamente, operazionalmente, logicamente, verbalmente, etc. – di analisi ed interpretazione degli eventi sociali (ovvero "di funzionamento dei settori o dei segmenti sociali"). Alcuni principi che ne devono caratterizzare la costruzione, sovente dichiarati ma raramente acquisiti dalle soluzioni apportate, sono i seguenti:

- proprietà di costituire parametri utili al confronto spazio-temporale attraverso l'adozione di metodologie riproducibili ed idonee a rappresentare contestualmente realtà multiformi;
- abilità di "raccontare" il divenire dei sistemi sociali;
- impegno ad assicurare l'analisi integrata dei sistemi, puntando sull'interazione e sulle relazioni esistenti tra le varie aree di interesse;
- collocazione in modelli concettualmente definiti che anche teoricamente ed operativamente servano da guida ed orientamento nell'attività di astrazione scientifica, di selezione degli aspetti pertinenti e rilevanti, di modalità di costruzione delle informazioni elementari e di loro trattamento a fini di determinazione di indicatori.

¹³¹ L'intero paragrafo è stato tratto dai seguenti lavori:

- Bernardi L., Capursi V., Librizzi L. (2004). Measurement awareness: the use of indicators between expectations and opportunities, SIS: Sezione Specializzata, Atti della XLIII Riunione Scientifica, Bari.
- Bernardi L., Torelli N. (1990), Le potenzialità offerte alle statistiche sociali ufficiali dalla definizione di modelli e indicatori, ISTAT, Avanzamenti metodologici e statistiche ufficiali, Atti delle prime giornate di studio, Roma, 13-14 dicembre.

¹³² Land K. (1971), On the definition of social indicators, *The American Sociologist*, 6, 322-325.

Focalizzando l'attenzione su quest'ultimo aspetto e ponendoci in relazione a quanto Land suggerisce, è possibile proporre una preliminare classificazione funzionale di modelli che traduce il valore che acquistano la costruzione, il ruolo giocato, il significato stesso degli indicatori in modelli che prevedono, per diversa natura, diverse procedure operative.

I **modelli descrittivi** hanno una duplice funzione: in prima istanza essi sono rivolti a mostrare ammontare, composizione, dinamiche delle grandezze relative ai singoli sistemi; in secondo luogo il loro obiettivo è quello di spiegare le relazioni esistenti tra le stesse grandezze considerate e fra queste ed altre appartenenti ad altri sistemi, con i quali sia ipotizzabile un intreccio organico, la cui mancata comprensione potrebbe rendere inintelligibili comportamenti ed esiti dei processi sotto esame.

L'adozione di **modelli previsionali** è spesso strettamente condizionata dalla qualità della risoluzione conseguita con i modelli relazionali precedenti, anche se non mancano casi in cui una corretta estrapolazione semplice – e a condizioni controllate – delle dinamiche osservate può risultare efficace a fornire indicazioni di prospettiva.

La funzione dei **modelli decisionali** riguarda la capacità di stabilire la natura, la misura, la cadenza con cui devono intervenire le intensità e i valori delle variabili che agiscono nei processi, in modo che questi possano essere governati ed indirizzati verso obiettivi giudicati socialmente desiderabili.

Rivolti a "prevedere" gli esiti di scelte possibili nei settori di interesse sono i **modelli simulativi**, che si realizzano, secondo logiche distinte, ponendo in alternativa due o più soluzioni adottabili o sulla scorta degli atteggiamenti desunti dall'analisi di un passato con condizioni di alta similarità.

Infine i **modelli valutativi** costituiscono strumenti, fondati sull'uso di indicatori di processo o di risultato, il cui ruolo è quello di interpretare il contributo netto determinato dall'azione politica promossa.

Alla preziosa affermazione di Land, tuttavia, la storia degli indicatori sviluppata in questi ultimi decenni sembra aver aggiunto un ulteriore principio: lo stesso singolo indicatore è solitamente il risultato di un processo di disarticolazione di un evento complesso ovvero di un fenomeno sociale che, visto nelle sue componenti elementari, deve essere successivamente riassembleto attraverso giustificate procedure di normalizzazione, aggregazione e ponderazione degli indicatori che sono stati individuati. Chiaramente questo deve avvenire utilizzando modalità di osservazione/descrizione/analisi/interpretazione qualitative e quantitative, soggettive ed

oggettive, da fonti o da indagini originali, con la ricerca di connessioni (casuali o non) e secondo punti di vista ragionevoli.

L'indicatore esiste quindi entro un modello ed è anche prodotto del modello stesso: secondo questo principio, esso spesso accresce contenuto e significato di un concetto complesso in esame interno al modello.

Data la copiosa produzione di informazioni statistiche in campo sociale, è possibile affermare che è diffuso un certo entusiasmo per la costruzione e la produzione di indicatori sociali, spesso utili e facili al consumo, frequentemente stimulate da approcci volontaristici e definiti negli esiti attesi anziché basati su solidi modelli teorici e su una buona progettazione ex ante. Molti sono quindi gli interrogativi che nascono rispetto alla legittimità delle scelte di merito e di metodo effettuate in fase di costruzione di un indicatore. Il punto cruciale, come segnalato da Curatolo¹³³, sembra essere dato da una mancanza di "convenienza" nella costruzione di indicatori complessi o "aggregati". L'autore ha ampiamente delineato le ragioni per cui privilegiare la costruzione di indicatori semplici o "disaggregati": questo è il motivo per cui nel suo lavoro è quasi del tutto assente l'analisi del passaggio da qualità a quantità. Tuttavia, pur condividendo quanto espresso da Curatolo, riteniamo che il problema della qualità-quantità non possa essere ignorato. Gli obiettivi che permettono l'identificazione di un indicatore condizionano le scelte di merito e di metodo: si possono costruire indicatori per creare semplici graduatorie, per valutazioni, per monitoraggi, per effettuare confronti.

Ma quale utilizzo si può fare del confronto fra singole dimensioni se l'obiettivo è quello di confrontare sistemi e non singoli elementi? È legittimo quindi chiedersi quali indicatori utilizzare e per quale scopo. Una risposta, legata alla difficoltà del concetto da misurare attraverso l'indicatore potrebbe essere la seguente: cerchiamo di trovare un modo per misurare il concetto che, prendendo in considerazione gli obiettivi, risulti essere il più funzionale e miri a descrivere quantitativamente e qualitativamente il processo d'interesse, avendo comunque una mappa di lettura del fenomeno che ne rappresenta un modello d'interpretazione.

Nell'ipotesi che siano gli obiettivi di ricerca a determinare gli indicatori, non è possibile stabilire una definizione chiara di indicatore: vogliamo per questa via affermare che essi sono il legame tra le osservazioni ed il fenomeno che si vuole misurare; sono quindi metadati (dati che descrivono dati), statistiche con uno scopo. Inoltre gli indicatori aiutano

¹³³ Curatolo R. (1972), Indicatori sociali. Atti della XXVII Riunione Scientifica della SIS (Palermo, 29-31 maggio), vol. 1.

a far luce sul concetto da misurare e finiscono per assumere il duplice ruolo di specificazione e di misurazione del concetto.

È necessario specificare che l'obiettività nella misura di un indicatore complesso è più formale che sostanziale: nel processo valutativo si oppongono quindi misure basate su evidenze fattuali a misure basate su opinioni soggettive.

Un problema da affrontare¹³⁴ riguarda la costruzione di una scala per misurare un concetto unidimensionale o multidimensionale, soluzione giustapposta alla costruzione di un indicatore semplice o complesso. La misurazione può essere concepita come la costruzione di scale che traducono strutture relazionali empiriche in strutture relazionali numeriche sulla base di specifiche assunzioni teoriche. Se le scale devono quindi riflettere la relazione tra il concetto (costrutto) e la sua misurazione, anche l'indicatore deve riflettere questa relazione.

In generale, il costrutto deve essere rappresentato dalle misure di variabili correlate che danno origine a singoli indicatori.

Quando il costrutto è unidimensionale e direttamente misurabile, il problema consiste nel rispettare le proprietà matematico-statistiche per il confronto (ad esempio nella normalizzazione); qualora il costrutto invece non sia direttamente misurabile, l'attenzione va posta sull'identificazione della relazione specifica tra costrutto ed indicatore, valutando se la trasformazione applicata ai dati originali lascia invariata la scala.

Se il costrutto ha carattere multidimensionale e se sono note la forma generale della relazione che lega ogni variabile ed il relativo indicatore e la funzione tra il costrutto e le sue componenti, dobbiamo porre attenzione alle trasformazioni applicate alle singole variabili al fine di ottenere indicatori semplici ed al processo di aggregazione di questi. Il problema è quindi complesso ma può essere risolto purché le proprietà matematiche siano rispettate dalle operazioni di aggregazione.

Nei casi in cui il costrutto è complesso e di difficile definizione, c'è un ampio margine di arbitrarietà nell'identificazione sia delle dimensioni e degli indicatori ad essere relativi, sia degli indicatori complessi, che spesso misurano variabili latenti.

Tutto ciò premesso, anche qualora non ci siano errori procedurali, è difficile identificare il miglior approccio per misurare un costrutto multidimensionale ed astratto, sia dal punto di

¹³⁴ L'analisi di questo problema deriva da quanto suggerito da Bartholomew nei seguenti lavori:
- Bartholomew D. J. (1996) *The statistical approach to social measurement*, Academic Press, San Diego.
- Bartholomew D. J. (2002) Discussion on the paper by Fayers P. M., Hand D. J., p.253.

vista del senso della misurazione sia relativamente al campo di applicazione. Si tratta, in altre parole, di capire quali misure utilizzare relativamente agli obiettivi definiti.

Se l'obiettivo è quello di ricevere segnali d'allarme, l'attenzione va centrata sulle misure delle singole componenti mantenendo le informazioni disaggregate. Nei casi in cui lo scopo dell'analisi consista nel confronto tra più situazioni, la sintetizzazione, con le dovute cautele, è invece necessaria.

Concludendo possiamo definire l'indicatore come uno strumento per tradurre la misurazione di fenomeni complessi in una fusione sistematica di congetture interpretative, di identificazione di relazioni, di collocazione entro un modello funzionale di riferimento.

Tuttora vi è una certa distanza che separa volontà conoscitiva e fattibilità operativa: si tratta, in altre parole, della separazione parziale ancora esistente tra l'intenzione convinta, indispensabile, talvolta normativa/legislativa di arrivare a un giudizio su una dimensione complessa e la sua realistica misurazione, ampiamente determinata dal sistema di condizioni operative effettivamente adottate, anche se con la protezione di una singolare ed esplicitata riflessione sulla rigosità metodologica.

Concludiamo il capitolo riportando un pensiero di Vianelli¹³⁵:

"Il pensiero semplificatore privilegia la disgiunzione o la riduzione. Il pensiero complesso deve praticare nello stesso tempo distinzioni e congiunzioni e deve tentare di concepire il molteplice".

¹³⁵ Vianelli S. (1989), Statistica e Scienze della complessità, in Atti della XXXI Riunione scientifica della Società Italiana di Economia Demografia e Statistica, Palermo, Vol. XLIII, nn. 1-2, pp.15-67.

3.1 I DATI DEL CENSIS

Prima di metter mano ai dati, è necessario descrivere la struttura che essi hanno assunto relativamente ai metodi utilizzati dal Censis per la normalizzazione, l'aggregazione e la ponderazione nonché illustrare quali dati ci sono stati dal Censis forniti.

Al fine di permettere qualche elaborazione, il Censis ha scelto di inviarci i dati raccolti per la costruzione della graduatoria del 2006 delle facoltà di Scienze Politiche.

Riteniamo quindi opportuno, innanzitutto, illustrare gli indicatori elementari utilizzati per la determinazione della classifica delle facoltà pubblicata nel 2006, a cavallo tra i mesi di giugno e di luglio, a cadenza settimanale sul quotidiano *La Repubblica*.

3.1.1 Gli indicatori utilizzati nel 2006

Le famiglie considerate per la valutazione delle facoltà dell'anno 2006 sono state le seguenti:

- produttività per misurare la capacità di garantire un percorso regolare del ciclo di studi;
- didattica per valutare appunto la didattica offerta nel suo complesso;
- ricerca per esaminare la capacità della facoltà di progettare la ricerca;
- profilo docenti per determinare quanto la facoltà punta sul ringiovanimento e sulle relazioni internazionali del proprio organico;

- rapporti internazionali per stabilire il grado di apertura della facoltà alle esperienze internazionali di studenti e docenti.

Per ciascuna famiglia sono stati poi utilizzati un numero variabile di indicatori elementari.

Di seguito riportiamo ciascuno indicatore semplice in uno schema analitico indicandone

- il nominativo assegnato dal Censis,
- i dati che sono stati utilizzati per la sua determinazione,
- la fonte da cui i dati sono stati ricavati,
- l'aspetto che ciascun indicatore premia,

al fine di agevolare la comprensione delle successive elaborazioni.

Per ciascuna famiglia illustreremo, inoltre, la metodologia adottata per la creazione di un indicatore composto unico per ogni dimensione.¹³⁶

3.1.1.1 Produttività

La famiglia della produttività è stata valutata attraverso i cinque indicatori elementari seguenti.

(P1) TASSO DI PERSISTENZA TRA IL I ED IL II ANNO

iscritti all'anno accademico (A.A.) 2004-2005 che si sono immatricolati nell'anno accademico precedente / immatricolati nell'A.A. 2003-2004

Fonte: Miur - Ufficio di Statistica

Premia le facoltà con il maggiore tasso di passaggio dal primo al secondo anno.

(P2) INDICE DI REGOLARITÀ DEGLI STUDENTI

crediti acquisiti nell'anno solare (A.S.) 2004 dagli studenti iscritti alle lauree triennali e a ciclo unico / studenti iscritti alle lauree triennali e a ciclo unico nell'A.A. 2003-2004 * 60

Fonte: Miur - Ufficio di Statistica

Premia le facoltà dove è maggiore la quota dei crediti effettivamente acquisiti dagli studenti del nuovo ordinamento sul totale dei crediti acquisibili nell'anno accademico considerato.

(P3) TASSO DI ISCRITTI IN CORSO A.A. 2004-2005

(iscritti in totale - immatricolati - iscritti fuori corso) / (iscritti in totale - immatricolati)

¹³⁶ I paragrafi relativi alla descrizione degli indicatori elementari di ogni famiglia ed alla loro aggregazione (interna alla famiglia) sono tratti dalla nota metodologica del 2006.

Fonte: Miur - Ufficio di Statistica

Premia le facoltà con il maggior tasso di iscritti in corso al netto degli immatricolati.

(P4) TASSO DI LAUREATI TRIENNALI

laureati nell'A.S.2004 nei corsi di laurea triennale immatricolati nell'A.A. 01-02/
immatricolati ai corsi di laurea triennale nell'A.A. 01-02

Fonte: Miur - Ufficio di Statistica

Premia le facoltà che nell'ultimo triennio hanno portato il maggior numero di studenti alla laurea entro la durata regolare del corso.

(P5) TASSO DI LAUREATI IN CORSO

laureati in corso nell'A.S. 2004 nei corsi di laurea a ciclo unico e del vecchio ordinamento / totale dei laureati nei corsi citati

Fonte: Miur- Ufficio di Statistica

Premia le facoltà con il maggior tasso di laureati in corso del vecchio ordinamento e dei corsi a ciclo unico nei due anni solari considerati.

Questi indicatori semplici sono poi stati aggregati secondo la seguente formula:

$$P = \frac{\text{std}(P1) + \text{std}(P2) + \text{std}(P3) + \text{std}\left(\frac{\text{std}(P4) * n1 + \text{std}(P5) * n2}{n1 + n2}\right)}{4}$$

dove $\text{std}(X)$ indica il valore normalizzato dell'indicatore X.

Inoltre, nella formula

- $n1$ è il numero dei laureati nell'A.S. 2004 nei corsi di laurea triennale immatricolati nell'A.A. 2001-2002
- $n2$ è il totale dei laureati nei corsi di laurea a ciclo unico e del vecchio ordinamento nell'A.S. 2004

3.1.1.2 Didattica

Gli indicatori elementari utilizzati per valutare la famiglia della didattica sono gli otto indicati di seguito.

(D1) NUMERO DI CLASSI DI LAUREA PRESENTI NELL'OFFERTA DI CORSI A.A. 2004-2005

Fonte: Miur - Ufficio di Statistica

Premia le facoltà che contano il maggior numero di classi di laurea

(D2) NUMERO DI INSEGNAMENTI PRESENTI NELL'OFFERTA DI CORSI A.A. 2003-2004

Fonte: CNVSU- NUCLEI 2005

Premia le facoltà che hanno attivato il maggior numero di insegnamenti.

(D3) DOCENTI DI RUOLO / INSEGNAMENTI

(A.S. 2004 - A.A. 2004-2005)

Fonte: Miur - Ufficio di Statistica ; CNVSU- NUCLEI 2005

Premia le facoltà con il più alto numero di docenti di ruolo per insegnamento attivato.

(D4) DOCENTI DI RUOLO / ISCRITTI

docenti di ruolo al 31.12.2004 / iscritti nell' A.A. 2004-2005

Fonte: Miur - Ufficio di Statistica

Premia le facoltà con il più alto numero di docenti di ruolo per studente.

(D5) POSTI AULA ISCRITTO

n. di posti aula NUCLEI 2004 / iscritti nell'A.A. 2002-2003

Fonte: Miur - Ufficio di Statistica ; CNVSU- NUCLEI 2004

Premia la facoltà con il maggior numero di posti aula per studente (l'indice assume un valore massimo pari a 1).

(D6) POSTI AULA ISCRITTO

n. di posti aula NUCLEI 2005 / iscritti nell'A.A. 2003-2004

Fonte: Miur - Ufficio di Statistica ; CNVSU- NUCLEI 2005

Premia la facoltà con il maggior numero di posti aula per studente (l'indice assume un valore massimo pari a 1).

(D7) STAGES PER STUDENTE A.A. 2003-2004

Fonte: Miur - Ufficio di Statistica ; CNVSU- NUCLEI 2005

Premia la facoltà con il maggior numero di stages per studente

(D8) ATTIVITÀ DI MONITORAGGIO E DI VALUTAZIONE DEI CORSI A.A. 2003-2004

Fonte: Miur - Ufficio di Statistica ; CNVSU- NUCLEI 2005

Premia le facoltà che hanno realizzato attività di programmazione e monitoraggio dell'offerta didattica.

Questi indicatori semplici sono poi stati aggregati secondo la seguente formula:

$$D = \frac{\text{std} \left(\frac{\text{std} (D1) + \text{std} (D2) + \text{std} (D3)}{3} \right) + \text{std} (D4) + \text{std} \left(\frac{\text{std} (D5) + \text{std} (D6)}{2} \right) + 0.5 * \text{std} (D7)}{4} * k$$

dove $\text{std}(X)$ indica il valore normalizzato dell'indicatore X ed il valore di k è determinato come segue:

$$k = \begin{cases} 1.00 & \text{se } D8 < 75 \\ 1.05 & \text{se } D8 \geq 75 \end{cases}$$

3.1.1.3 Ricerca

La famiglia della ricerca conta sette indicatori semplici.

(R1) NUMERO DI UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2003 PER DOCENTE
n. di unità di ricerca finanziate / docenti di ruolo al 31.12.2002

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà con il maggior numero di unità di ricerca finanziate per docente di ruolo.

(R2) NUMERO DI UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN NEL 2004 PER DOCENTE
n. di unità di ricerca finanziate / docenti di ruolo al 31.12.2003

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà con il maggior numero di unità di ricerca finanziate per docente di ruolo.

(R3) NUMERO DI UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2005 PER DOCENTE
n. di unità di ricerca finanziate / docenti di ruolo al 31.12.2004

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà con il maggior numero di unità di ricerca finanziate per docente di ruolo.

(R4) FINANZIAMENTO MEDIO OTTENUTO DALLE UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E
FIRB NEL 2003

totale dei finanziamenti ottenuti dalle unità di ricerca / n. delle unità finanziate

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà che hanno ottenuto il maggior finanziamento medio per unità di ricerca.

(R5) FINANZIAMENTO MEDIO OTTENUTO DALLE UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN
NEL 2004

totale dei finanziamenti ottenuti dalle unità di ricerca / n. delle unità finanziate

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà che hanno ottenuto il maggior finanziamento medio per unità di ricerca.

(R6) FINANZIAMENTO MEDIO OTTENUTO DALLE UNITÀ DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E
FIRB NEL 2005

totale dei finanziamenti ottenuti dalle unità di ricerca / n. delle unità finanziate

Fonte: Miur - Ufficio di Statistica ; CINECA

Premia le facoltà che hanno ottenuto il maggior finanziamento medio per unità di ricerca.

(R7) PROGETTI DI RICERCA FINANZIATI DAI PROGRAMMI COMUNITARI

V e VI Programma Quadro; Programma Tempus

Fonte: CORDIS - UNIONE EUROPEA - EFT

Premia le facoltà con progetti di ricerca finanziati dai programmi comunitari citati.

Questi indicatori semplici sono poi stati aggregati secondo la seguente formula:

$$R = \frac{\text{std} \left(\frac{\text{std} (R1) + \text{std} (R2) + \text{std} (R3)}{3} \right) + \text{std} \left(\frac{\text{std} (R4) + \text{std} (R5) + \text{std} (R6)}{3} \right)}{2} * k$$

dove $\text{std}(X)$ indica il valore normalizzato dell'indicatore X ed il valore di k è determinato come segue:

$$k = \begin{cases} 1.00 & \text{se } R7 = 0 \\ 1.05 & \text{se } R7 > 0 \end{cases}$$

3.1.1.4 Profilo docenti

La valutazione della famiglia "profilo docenti" viene effettuata attraverso cinque indicatori elementari.

(PD1) ETÀ MEDIA DEI DOCENTI AL 2005

età media dei docenti di ruolo al 2005

Fonte: Miur - Ufficio di Statistica

Premia le facoltà con i docenti mediamente più giovani.

(PD2) INVECCHIAMENTO

età media dei docenti di ruolo al 2005 -età media dei docenti di ruolo al 2001

Fonte: Miur - Ufficio di Statistica

Premia le facoltà che hanno registrato il minor invecchiamento del corpo docente nel periodo considerato.

(PD3) ERASMUS IN USCITA PER DOCENTE

n. di studenti con borse Erasmus nell'a.a.04-05 / docenti di ruolo al 31.12.2004

Fonte: Agenzia Nazionale Socrates Italia – Ufficio Erasmus ; Uffici Relazioni Internazionali degli Atenei ; MIUR- Ufficio di Statistica

Premia le facoltà dove è più alto il rapporto tra borse erasmus / docenti della facoltà.

(PD4) TASSO DI INSEGNAMENTI TENUTI DA DOCENTI NON DI RUOLO EXTRA ACCADEMICI

insegnamenti tenuti da docenti non di ruolo "extra accademici"/ totale degli insegnamenti attivati nell'a.a.03-04

Fonte: CNVSU- NUCLEI 2005

Premia le facoltà la cui offerta di insegnamenti presenti una composizione equilibrata tra insegnamenti tenuti da docenti di ruolo e da docenti non di ruolo "extra accademici" (il benchmark di riferimento è stato individuato nel valore mediano).

(PD5) DOCENTI COINVOLTI NEL PROGRAMMA PER IL C.D. "RIENTRO DEI CERVELLI"

n. di docenti che hanno partecipato al programma di incentivazione della mobilità degli studiosi italiani e stranieri nel triennio 2004-2006

Fonte: CINECA

Premia le facoltà che hanno avuto finanziamenti per il c.d. "rientro dei cervelli"

Questi indicatori semplici sono poi stati aggregati secondo la seguente formula:

$$PD = \frac{\text{std} \left(\frac{\text{std} (PD1) + \text{std} (PD2)}{2} \right) + 0.5 * \text{std} (PD3) + 0.5 * \text{std} (PD4)}{3} * k$$

dove $\text{std}(X)$ indica il valore normalizzato dell'indicatore X ed il valore di k è determinato come segue:

$$k = \begin{cases} 1.00 & \text{se } PD5 = 0 \\ 1.05 & \text{se } PD5 > 0 \end{cases}$$

3.1.1.5 Rapporti internazionali

I rapporti internazionali delle facoltà sono misurati tramite quattro diversi indicatori elementari.

(RI1) ERASMUS IN USCITA PER STUDENTE

studenti con borse erasmus in uscita nell'A.A. 2004-2005 / iscritti al netto degli immatricolati A.A. 2004-2005

Fonte: Agenzia Nazionale Socrates Italia – Ufficio Erasmus ; Uffici Relazioni Internazionali degli Atenei ; MIUR- Ufficio di Statistica

Premia le facoltà dove è più alto il rapporto tra borse erasmus in uscita ed iscritti (al netto degli immatricolati).

(RI2) ERASMUS IN ENTRATA PER STUDENTE

media degli studenti stranieri che hanno ottenuto una borsa Erasmus nella facoltà negli anni accademici 2003-04 e 2004-05 / iscritti A.A. 2004-2005

Fonte: Agenzia Nazionale Socrates Italia – Ufficio Erasmus ; Uffici Relazioni Internazionali degli Atenei ; MIUR- Ufficio di Statistica

Premia le facoltà dove è più alto il rapporto tra borse erasmus in entrata ed iscritti.

(RI3) UNIVERSITA' OSPITANTI PER DOCENTE

numero di università straniere che hanno ospitato gli studenti Erasmus nell' A.A. 2004-2005 / docenti di ruolo al 31.12.2004

Fonte: Agenzia Nazionale Socrates Italia – Ufficio Erasmus ; Uffici Relazioni Internazionali degli Atenei ; MIUR- Ufficio di Statistica

Premia le facoltà nelle quali è più alto il rapporto tra il numero di università ospitanti ed il corpo docente di ruolo.

(RI4) OPPORTUNITA' INTERNAZIONALI

numero di contributi ottenuti dalla facoltà per attività di cooperazione internazionale 2003-2006

Visite per scambio docenti finanziate dal Miur nel 2004 ; Programma Vigoni 2003-2004 ; Programma Italia-Germania 2003-2004 ; Azioni Italia-Spagna 2004-2005 ; Programma Italia-Germania 2004-2005 ; Programma Galileo Italia-Francia 2004-2005; Cooperazione Internazionale finanziata dal Ministero degli Esteri - Accordi Bilaterali 2002-2006)

Fonte: MIUR; CRUI ;MINISTERO AFFARI ESTERI

Premia le facoltà che hanno ricevuto finanziamenti per opportunità internazionali.

Questi indicatori semplici sono poi stati aggregati secondo la seguente formula:

$$RI = \frac{\text{std}(RI1) + \text{std}(RI2) + \text{std}(RI3)}{3} * k$$

dove $\text{std}(X)$ indica il valore normalizzato dell'indicatore X ed il valore di k è determinato come segue:

$$k = \begin{cases} 1.00 & \text{se } RI4 = 0 \\ 1.05 & \text{se } RI4 > 0 \end{cases}$$

3.1.2 Normalizzazione e aggregazione

Una volta raccolte le variabili necessarie al calcolo degli indicatori elementari, si ottengono valori non comparabili perché misurati con diverse unità di misura ed aventi diverso ordine di grandezza. È quindi necessario normalizzare tali indicatori grezzi riconducendoli entro uno stesso campo di variazione perché siano comparabili nonché aggregabili in un unico indicatore composto.

La tecnica di normalizzazione che il Censis utilizza è quella che trasforma i valori in indicatori relativizzandoli al campo di variazione secondo la formula seguente:

$$I = \frac{X - \min(X)}{\max(X) - \min(X)} 1000$$

dove X è il valore dell'indicatore grezzo, mentre $\min(X)$ e $\max(X)$ sono rispettivamente il valore minimo ed il massimo assunti che l'indicatore assume nell'insieme delle facoltà omogenee considerate.¹³⁷ I valori trasformati varieranno quindi da un minimo di 0 ad un massimo pari a 1000 e saranno confrontabili all'interno di ogni raggruppamento di facoltà: non sono infatti possibili confronti tra facoltà diverse per lo stesso indicatore.

I punteggi di sintesi delle famiglie sono infine riportati alla scala 66-110, come consuetudine del voto di laurea. In nessuna delle note metodologiche de "La Grande Guida all'Università" è stato chiarito come questa operazione sia stata effettuata: ipotizziamo quindi che la formula utilizzata sia la seguente:

$$66 + \left(\frac{I}{1000} * 44 \right) .$$

dove, attribuito 110 al valore massimo e 66 al valore minimo della singola famiglia (e attribuendo quindi tali valori alle facoltà corrispondenti), i singoli indicatori sono stati relativizzati ai 44 punti di scarto tra massimo e minimo.

Il voto medio finale M attribuito a ciascuna facoltà è stato calcolato come media aritmetica dei punteggi normalizzati delle cinque famiglie considerate:

$$M = \frac{\text{std}(P) + \text{std}(D) + \text{std}(R) + \text{std}(PD) + \text{std}(RI)}{5}$$

¹³⁷ Un'interessante applicazione potrebbe prevedere l'utilizzo del massimo e del minimo teorico per gli indicatori elementari per cui tali valori sono determinabili: questa tecnica permetterebbe di ridurre le distanze tra unità osservate in termini di scarti tra valori normalizzati degli indicatori elementari.

dove P indica il punteggio della famiglia della produttività, D è il punteggio della didattica, R della ricerca, PD della famiglia relativa al profilo docenti ed infine RI dei rapporti internazionali.

3.2 IL DATASET UTILIZZATO PER LE ELABORAZIONI

Naturalmente non abbiamo potuto essere in possesso di tutti i dati grezzi per il calcolo di tutti gli indicatori elementari sopra illustrati per tutte le facoltà. Il Censis ha perciò acconsentito ad inviarci una selezione della grande banca dati in suo possesso: le informazioni che ci sono state fornire sono quelle relative ai dati raccolti per tutte le facoltà di Scienze Politiche ed utilizzati per la stesura de "La Grande Guida all'Università" dell'anno 2006. Il numero totale di facoltà di Scienze Politiche in Italia era pari a 28 unità nel 2006: poiché però la facoltà appartenente all'Ateneo Tuscia di Viterbo è stata istituita nel medesimo anno, non avendo a disposizione i valori di ogni indicatore elementare – e quindi delle famiglie – è stata esclusa dalla valutazione.

La struttura dei dati che ci è pervenuta ha richiesto un attento esame ed un'altrettanto accurata interpretazione al fine di determinare da quali variabili, tra tutte le misure del dataset¹³⁸, era composto ogni singolo indicatore elementare. Rimandiamo all'appendice del presente elaborato una breve trattazione di questa fase del nostro lavoro. È doveroso però chiarire che, utilizzando le informazioni fornite dal Censis, operando secondo le medesime metodologie di normalizzazione, di aggregazione e di ponderazione, tenendo comunque in considerazione quella che è stata la nostra interpretazione dei dati inviatici, la graduatoria ottenuta e per le singole famiglie e a livello complessivo non ha coinciso con quella pubblicata dal Censis.

Resta da chiarire su quale insieme di dati abbiamo svolto le nostre elaborazioni. Per poter garantire la confrontabilità tra graduatorie, abbiamo scelto di utilizzare i valori ottenuti dalle nostre elaborazioni sui dati inviati dal Censis. Sono stati mantenuti i punteggi complessivi della famiglia ottenuti come aggregazione dei singoli indicatori elementari così come risultati dall'applicazione della formula prevista in nota metodologica, senza riportarli

¹³⁸ Un dataset è una collezione di dati presentati usualmente in forma tabellare: ogni colonna rappresenta una variabile ed ogni riga corrisponde ad un'unità del collettivo.

alla scala 110-66 per due ordini di motivi: anzitutto per l'invarianza che questa operazione determina in termini di graduatoria finale (la classifica delle facoltà è infatti la medesima) e, in secondo luogo, per l'incertezza sulla coerenza tra la formula di riproporzionamento da noi ipotizzata e quella utilizzata dal Censis.

3.3 ANALISI PRELIMINARE¹³⁹

Nel secondo capitolo abbiamo sottolineato come l'analisi esplorativa dei dati rappresenti un passaggio importante precedente l'elaborazione degli stessi per il suo scopo di indagine della struttura dei dati in termini di relazioni esistenti tra le variabili considerate.

In quest'ottica abbiamo voluto partire da questo *step* operativo con l'obiettivo di stabilire l'esistenza o meno di relazioni tra gli indicatori elementari considerati per la normalizzazione, l'aggregazione e la ponderazione ed, eventualmente, il verso e l'entità di queste relazioni.

Due sono i tipi di relazione possibile esistenti tra due variabili: asimmetrica o simmetrica. Nel primo caso rientrano le variabili vengono divise in due gruppi, uno di variabili dipendenti ed uno di variabili esplicative (o indipendenti o predittive). I modelli basati su questo tipo di relazione, definita $Y \leftarrow X$: essi ipotizzano una relazione di causa-effetto tra i due gruppi di variabili. Quando invece le variabili analizzate si ipotizzano essere sullo stesso piano, il modello di analisi è simmetrico e si indica con $Y \leftrightarrow X$. Se le variabili X e Y sono quantitative, la relazione asimmetrica si può valutare mediante la regressione; se la relazione è invece di tipo simmetrico si utilizza la correlazione.

La relazione di indipendenza tra due variabili X e Y è misurabile attraverso il coefficiente di correlazione di Bravais-Pearson:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

dove σ_{xy} è la covarianza¹⁴⁰ di X e Y ; a denominatore troviamo il prodotto degli scarti quadratici medi delle due variabili. Il coefficiente varia nell'intervallo $[-1, 1]$ e in particolare

¹³⁹ Fabbris L., Capiluppi C., Giancotti G., Meneghello A. (1993), Statree 1.0 – Sistema esperto per la scelta del metodo di analisi statistica, Manuale per l'uso, Edizioni Summa, Padova.

raggiunge il massimo quando i valori di Y sono una combinazione lineare dei valori di X , cioè $y = \alpha + \beta x$ con β positivo, ed il minimo quando i punti sono ugualmente allineati ma con β negativo; il coefficiente è nullo quando non esiste relazione lineare tra le due variabili.

La determinazione dell'entità della correlazione avviene quindi attraverso il calcolo di questo coefficiente; per stabilire però se tale relazione è statisticamente significativa, il valore viene sottoposto ad una verifica di ipotesi che lo uguaglia a zero:

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases}$$

Se il test rifiuta l'ipotesi nulla (H_0) significa che, stabilito un livello di significatività del 5% (ovvero pari a 0.05), esiste relazione lineare tra le due variabili; nel caso contrario X e Y sono incorrelate.

3.3.1 Analisi della correlazione interna alle famiglie¹⁴¹

Abbiamo esaminato, anzitutto, la correlazione tra gli indicatori elementari appartenenti ad una stessa famiglia: l'evidenza di forti correlazioni tra i diversi indicatori indicherebbe che alcuni aspetti sono misurati – e quindi conteggiati – più volte all'interno della stessa famiglia, non rispondendo al criterio di parsimonia che deve guidare la costruzione di indicatori composti.

Riportiamo di seguito, per ogni famiglia, la matrice di correlazione tra gli indicatori ad essa appartenenti. Gli asterischi riportati accanto ad alcuni valori ne indicano la significatività: due asterischi (**) indicano che la correlazione è statisticamente diversa da 0 a livello 5% mentre un solo asterisco (*) indica la significatività fissato il livello dell'1%. Le tabelle – ovvero le matrici – sono simmetriche rispetto alla diagonale principale nella quale troviamo tutti valori unitari dal momento che la correlazione tra una variabile e sé stessa è perfetta (e quindi determina un coefficiente di correlazione pari ad 1¹⁴²).

¹⁴⁰ La covarianza di due variabili si calcola secondo la seguente formula: $\sigma_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$ con \bar{x} e \bar{y} media della distribuzione rispettivamente della variabile X e della variabile Y .

¹⁴¹ Tutte le elaborazioni presentate in questo paragrafo e nei successivi paragrafi del capitolo sono state ottenute utilizzando i pacchetti statistici SAS e SPSS e l'applicativo Excel di Microsoft Office.

¹⁴² Inoltre la covarianza di una variabile verso sé stessa (σ_{xx}) coincide con la varianza per cui il coefficiente ρ ha numeratore e denominatore uguali.

Tabella 1. Matrice di correlazione degli indicatori elementari appartenenti alla famiglia "produttività".

	P1	P2	P3	P4	P5
P1	1	.303	.130	.391*	-.141
P2	.303	1	.327	.648**	.134
P3	.130	.327	1	.295	-.118
P4	.391*	.648**	.295	1	.280
P5	-.141	.134	-.118	.280	1

Dalla tabella emerge anzitutto la significativa correlazione positiva tra gli indicatori P2 e P4 che misurano rispettivamente la regolarità degli studenti in termini di crediti acquisiti sul totale crediti acquisibili ed il tasso di laureati triennali. Naturalmente si tratta di una relazione facilmente dimostrabile anche da un punto di vista logico nonché causale. Sorprendono però i bassi valori del coefficiente di correlazione tra i due indicatori relativi ai laureati (P4 e P5) e tra P2 e P5. Un ulteriore valore che suscita stupore è quello che indica una relazione negativa tra P3 e P5 ovvero tra il tasso di iscritti e di laureati in corso: tra essi è infatti ovvio pensare ad una forte correlazione positiva, non confermata però, a quanto pare, dai nostri dati. Le altre correlazioni hanno valori bassi per cui è ipotizzabile l'assenza di una reale relazione lineare tra le coppie di variabili interessate.

Naturalmente – e lo diciamo ora ma attribuendo valenza a questa affermazione nell'analisi di ogni matrice di correlazione – i risultati risentono pesantemente del basso numero di unità relativamente alle quali disponiamo di dati: 27 unità infatti non rappresentano un numero sufficiente di osservazioni tale da determinare risultati stabili e convincenti, nonché estendibili all'universo complessivo delle facoltà italiane.

Tabella 2. Matrice di correlazione degli indicatori elementari appartenenti alla famiglia "didattica".

	D1	D2	D3	D4	D5	D6	D7	D8
D1	1	.701**	-.006	-.286	-.016	-.237	.145	-.271
D2	.701**	1	-.435*	-.309	.010	-.131	.052	-.165
D3	-.006	-.435*	1	-.082	-.269	-.344	-.360	.100
D4	-.286	-.309	-.082	1	.184	.348	.155	.167
D5	-.016	.010	-.269	.184	1	.587**	.546**	.223
D6	-.237	-.131	-.344	.348	.587**	1	.617**	.262
D7	.145	.052	-.360	.155	.546**	.617**	1	.289
D8	-.271	-.165	.100	.167	.223	.262	.289	1

Anche dall'analisi delle correlazioni tra indicatori della famiglia della didattica emergono alcune considerazioni interessanti. La forte relazione positiva (0.701) tra D1 e D2 trova spiegazione nell'andamento parallelo che queste due misure hanno: più cresce il numero di classi di laurea, parimenti cresce il numero di Corsi di Laurea offerto dalle facoltà, nonostante questo non sia in assoluto un indice di buona organizzazione¹⁴³. Gli indicatori D5 e D6 mostrano buona correlazione positiva per la loro analoga natura (cambia infatti solo il loro riferimento temporale) nonostante sia presumibile un valore più vicino ad 1 del loro coefficiente di correlazione. La significatività della relazione positiva tra D7 ed i due indicatori relativi al numero di posti aula per iscritto potrebbe essere interpretata come evidenza del fatto che le facoltà più organizzate in termini logistici sono anche quelle che offrono ai loro studenti possibilità di crescita maggiori come quelle fornite da esperienze di stages. Ricordiamo però che le nostre analisi riguardano le facoltà di Scienze Politiche: probabilmente tale relazione assume consistenza e significato maggiore per quelle facoltà che propongono in maniera più forte l'attività di stage (quali Scienze Statistiche, Economia o Ingegneria). L'unica relazione lineare negativa statisticamente significativa è quella esistente tra gli indicatori D2 e D3, relazione facilmente comprensibile dal momento che l'uno è denominatore dell'altro: più alto infatti è il numero di corsi attivati, più basso sarà, a parità di numero dei docenti di ruolo, il valore dell'indicatore D3.

Tabella 3. Matrice di correlazione degli indicatori elementari appartenenti alla famiglia "ricerca".

	R1	R2	R3	R4	R5	R6	R7
R1	1	-.172	.773**	.185	.026	.527**	-.126
R2	-.172	1	-.104	.095	.396*	-.014	.263
R3	.773**	-.104	1	.218	-.022	.237	-.166
R4	.185	.095	.218	1	.163	.491**	.173
R5	.026	.396*	-.022	.163	1	.237	.014
R6	.527**	-.014	.237	.491**	.237	1	.201
R7	-.126	.263	-.166	.173	.014	.201	1

Per gli indicatori elementari appartenenti alla famiglia della didattica ci si attenderebbero forti correlazioni positive tra R1, R2 ed R3 da un lato e tra i successivi tre indicatori dal momento che i due gruppi di indicatori misurano i medesimi aspetti ma con riferimento

¹⁴³ Il decreto ministeriale 270/2004 ha evidenziato la necessità di ridurre il numero di Corsi di Laurea attivati dalla singola facoltà appartenenti alla medesima classe di laurea: l'intento è quello di frenare il rapido proliferare del numero di Corsi di Laurea offerti, i quali sono spesso molto simili tra loro e che, talvolta, mascherano la necessità della facoltà di non lasciare scoperta nessuna area disciplinare.

temporale diverso. I valori dei coefficienti di correlazione non sembrano però confermare queste ipotesi. La più forte relazione lineare lega R1 ed R3, indicatori costruiti in modo analogo che si riferiscono l'uno al 2003 ed il secondo al 2005. Legami positivi ma difficilmente spiegabili esistono tra R1 e R6 e tra R4 e R6. Inoltre una relazione si suppone esistere – secondo senso logico – tra R1 e R4, tra R2 e R5, tra R3 e R6: solo il valore relativo al coefficiente di correlazione tra R2 e R5 è statisticamente significativo ma stabilisce un legame piuttosto debole.

Tabella 4. Matrice di correlazione degli indicatori elementari appartenenti alla famiglia "profilo docenti".

	PD1	PD2	PD3	PD4	PD5
PD1	1	.519**	.074	-.445*	.059
PD2	.519**	1	-.243	-.031	-.086
PD3	.074	-.243	1	.275	.486*
PD4	-.445*	-.031	.275	1	.137
PD5	.059	-.086	.486*	.137	1

Gli indicatori PD1 e PD2 presentano una correlazione positiva statisticamente significativa: entrambi infatti valutano l'età del corpo docente. Una relazione lineare positiva lega anche PD3 e PD5, indicatori che valutano quanti studenti Erasmus in uscita ci sono per docente l'uno ed il numero di docenti che ha beneficiato di finanziamenti per il "rientro dei cervelli": si tratta infatti di due valori che misurano un elemento di apertura delle facoltà rispetto all'internazionalizzazione. Negativo è invece il coefficiente di correlazione che lega PD1 e PD4: l'interpretazione di questo valore può indicare che, a parità di insegnamenti attivati, laddove l'età degli insegnanti è maggiore, il tasso di insegnamenti tenuto da docenti extra accademici è inferiore ovvero le facoltà con un corpo docente di età maggiore affida un numero inferiore di insegnamenti a docenti esterni alla facoltà.

Tabella 5. Matrice di correlazione degli indicatori elementari appartenenti alla famiglia "rapporti internazionali".

	RI1	RI2	RI3	RI4
RI1	1	.894**	.645**	.164
RI2	.894**	1	.528**	.138
RI3	.645**	.528**	1	-.202
RI4	.164	.138	-.202	1

Una correlazione positiva molto forte lega RI1 e RI2, indicatori elementari che valutano il numero di studenti con borse Erasmus rispettivamente in entrata e in uscita, relazione facilmente ipotizzabile e chiaramente interpretabile. Relazioni positive di una certa entità legano anche RI1 con RI3 e RI2 con RI3: tutti questi indicatori misurano aspetti relativi all'esperienza Erasmus per cui è facile immaginare l'esistenza di una relazione positiva tra di essi.

3.3.2 Analisi della correlazione tra tutti gli indicatori

Un passo successivo della fase di analisi delle relazioni sottostanti il dataset prevede l'esame della correlazione tra tutti gli indicatori elementari considerati. Data la numerosità degli indicatori (trattiamo infatti 30 variabili), è difficoltoso riportare la matrice di correlazione completa in un'unica tabella. Presentiamo, pertanto, cinque distinte tabelle che rappresentano la matrice di correlazione degli indicatori appartenenti ad ogni singola famiglia con tutte le restanti variabili; le relazioni tra una famiglia ed un'altra sono presentate in ogni tabella ma vengono commentate sola la prima volta che compaiono¹⁴⁴.

Tabella 6. Matrice di correlazione tra gli indicatori elementari appartenenti alla famiglia "produttività" e gli indicatori restanti.

	P1	P2	P3	P4	P5
D1	-.079	.290	.394*	.144	-.045
D2	.194	.314	.355	.290	.016
D3	-.215	-.168	-.398*	-.426*	.024
D4	.288	-.018	-.124	.061	-.167
D5	.237	.076	.332	.551**	.172
D6	.109	.179	.443*	.422*	.062
D7	-.126	.226	.410*	.424*	.240
D8	-.070	-.319	-.044	-.063	.271
R1	.194	.296	.490**	.361	.005
R2	.415*	.051	.271	.312	.006
R3	.277	.077	.188	.120	.133
R4	.091	-.046	.148	-.061	.339
R5	-.146	-.228	.118	.018	.032
R6	.260	.184	.413*	.291	.153
R7	.509**	.220	.187	.066	-.022
PD1	.047	-.143	-.127	-.246	.051

¹⁴⁴ Questo significa, ad esempio, che le correlazioni tra gli indicatori della famiglia produttività e quelli della famiglia didattica sono riportate in entrambe le tabelle ma i valori dei coefficienti ed il significato delle relazioni viene analizzato solo dopo la tabella della produttività (ovvero la prima tabella in cui compaiono).

	P1	P2	P3	P4	P5
PD2	.308	-.157	-.140	-.107	-.059
PD3	.236	.296	.114	.293	.158
PD4	.354	.490**	.331	.708**	.182
PD5	.163	.202	.064	.015	.131
RI1	.550**	.129	.102	.048	-.045
RI2	.597**	.181	-.011	.071	-.069
RI3	.203	.317	.143	.234	.079
RI4	.122	-.499**	-.145	-.374	-.077

La tabella evidenzia alcune relazioni lineari significative (8 coefficienti di correlazione su 40 sono significativamente diversi da zero) tra gli indicatori della famiglia della produttività e quelli appartenenti alla didattica. In misura minore (solo 4 coefficienti significativi su 35) esistono relazioni lineari statisticamente rilevanti tra gli indicatori P1 e P3 della produttività e gli indicatori della famiglia della ricerca; anche produttività e rapporti internazionali sono legate da correlazioni abbastanza importanti (3 correlazioni significative tra 0.5 e 0.6) che mettono in relazione solo P1 e P2 con gli indicatori di RI. Il legame tra produttività e profilo docenti riguarda solo PD4 per cui esistono 2 correlazioni positive con la famiglia PD.

Le relazioni più forti, ovvero quelle il cui coefficienti di correlazione è più vicino a 1 o a -1, legano P4 e PD4 (0.708), P1 e RI2 (0.597), P4 e D5 (0.551) ed infine P1 e RI1 (0.550): si tratta di relazioni per le quali è difficilmente ravvisabile un significato sostantivo.

Tabella 7. Matrice di correlazione tra gli indicatori elementari appartenenti alla famiglia "didattica" e gli indicatori restanti.

	D1	D2	D3	D4	D5	D6	D7	D8
P1	-.079	.194	-.215	.288	.237	.109	-.126	-.070
P2	.290	.314	-.168	-.018	.076	.179	.226	-.319
P3	.394*	.355	-.398*	-.124	.332	.443*	.410*	-.044
P4	.144	.290	-.426*	.061	.551**	.422*	.424*	-.063
P5	-.045	.016	.024	-.167	.172	.062	.240	.271
R1	.091	.031	-.242	.239	.249	.443*	.574**	-.040
R2	.135	.357	-.430*	-.003	.062	.049	-.143	-.053
R3	-.017	-.066	-.021	.295	.019	.158	.284	.187
R4	.269	-.037	.329	.002	.117	.024	.035	.014
R5	.266	.297	-.313	-.049	.143	.194	.306	-.117
R6	.304	.266	-.149	-.129	.354	.209	.466*	-.236
R7	.284	.477*	.040	-.153	-.005	-.056	-.225	-.113
PD1	.262	.092	.320	.028	-.191	-.366	-.243	.003
PD2	.100	.054	-.072	.353	-.200	-.019	-.034	-.174
PD3	.159	.367	-.043	-.222	.168	-.026	-.157	-.143
PD4	.116	.326	-.513**	.045	.367	.429*	.492**	-.038
PD5	.322	.567**	-.122	-.258	-.126	-.027	-.013	-.045
RI1	-.051	.305	-.130	.217	.048	-.031	-.236	-.006

	D1	D2	D3	D4	D5	D6	D7	D8
RI2	-.099	.234	-.121	.379	.130	.120	-.192	-.008
RI3	-.033	.159	-.014	-.190	.118	-.006	-.206	-.192
RI4	.164	.087	.184	-.160	-.035	-.267	.019	.227

Anche in questa tabella sono rilevabili alcune significative relazioni tra gli indicatori elementari: le famiglie della didattica e della ricerca sono legate da 5 correlazioni statisticamente significative, mentre sono 4 quelle che legano didattica e profilo docenti. Nessun coefficiente di correlazione tra gli indicatori appartenenti alla famiglia della didattica e gli indicatori dei rapporti internazionali è statisticamente diverso da zero per cui si può supporre l'assenza di relazione tra questi due insieme di indicatori. Le relazioni lineari più forti sono quelle che legano positivamente D7 e R1 (0.574) e D2 con PD5 (0.567): anche in questo caso sono però correlazioni non rilevanti dal punto di vista semantico.

Tabella 8. Matrice di correlazione tra gli indicatori elementari appartenenti alla famiglia "ricerca" e gli indicatori restanti.

	R1	R2	R3	R4	R5	R6	R7
P1	.194	.415*	.277	.091	-.146	.260	.509**
P2	.296	.051	.077	-.046	-.228	.184	.220
P3	.490**	.271	.188	.148	.118	.413*	.187
P4	.361	.312	.120	-.061	.018	.291	.066
P5	.005	.006	.133	.339	.032	.153	-.022
D1	.091	.135	-.017	.269	.266	.304	.284
D2	.031	.357	-.066	-.037	.297	.266	.477*
D3	-.242	-.430*	-.021	.329	-.313	-.149	.040
D4	.239	-.003	.295	.002	-.049	-.129	-.153
D5	.249	.062	.019	.117	.143	.354	-.005
D6	.443*	.049	.158	.024	.194	.209	-.056
D7	.574**	-.143	.284	.035	.306	.466*	-.225
D8	-.040	-.053	.187	.014	-.117	-.236	-.113
PD1	-.228	-.071	-.244	.204	-.218	.045	.355
PD2	.136	.090	.174	.094	.147	.133	.183
PD3	-.053	.277	-.162	.042	-.106	.186	.377
PD4	.537**	.257	.364	-.072	.289	.332	.194
PD5	-.044	.055	-.029	.143	.081	.281	.531**
RI1	.006	.418*	.044	-.025	-.071	.041	.480*
RI2	-.108	.312	-.116	-.018	-.119	.033	.544**
RI3	.068	.087	-.126	-.071	-.347	.103	.298
RI4	-.231	.194	-.055	.299	.369	.291	.226

Oltre a quelle precedentemente illustrate, ci sono poche altre relazioni significative tra la famiglia della ricerca e le due famiglie del profilo docenti (2 coefficienti statisticamente

diversi da zero) e dei rapporti internazionali (3 coefficienti significativi). Le relazioni più forti si attestano su valori vicino a 0.5 che dicono un legame di tipo positivo ma di peso non importante: la correlazione tra R7 e RI2 (0.544) e quella tra R1 e PD4 (0.537) non danno particolari indicazioni su evidenti legami tra le famiglie interessate; l'unica relazione significativa dal punto di vista semantico è quella che lega R7 e PD5 (0.531).

Tabella 9. Matrice di correlazione tra gli indicatori elementari appartenenti alla famiglia "profilo docenti" e gli indicatori restanti.

	PD1	PD2	PD3	PD4	PD5
P1	.047	.308	.236	.354	.163
P2	-.143	-.157	.296	.490**	.202
P3	-.127	-.140	.114	.331	.064
P4	-.246	-.107	.293	.708**	.015
P5	.051	-.059	.158	.182	.131
D1	.262	.100	.159	.116	.322
D2	.092	.054	.367	.326	.567**
D3	.320	-.072	-.043	-.513**	-.122
D4	.028	.353	-.222	.045	-.258
D5	-.191	-.200	.168	.367	-.126
D6	-.366	-.019	-.026	.429*	-.027
D7	-.243	-.034	-.157	.492**	-.013
D8	.003	-.174	-.143	-.038	-.045
R1	-.228	.136	-.053	.537**	-.044
R2	-.071	.090	.277	.257	.055
R3	-.244	.174	-.162	.364	-.029
R4	.204	.094	.042	-.072	.143
R5	-.218	.147	-.106	.289	.081
R6	.045	.133	.186	.332	.281
R7	.355	.183	.377	.194	.531**
RI1	.251	.135	.686**	.115	.315
RI2	.280	.228	.598**	.056	.371
RI3	.138	-.246	.893**	.215	.283
RI4	.329	.213	.030	-.250	.307

Analizzando le relazioni tra la famiglia "profilo docenti" e l'ultima famiglia, relativa ai rapporti internazionali, le correlazioni statisticamente diverse da zero sono 3 ed i relativi coefficienti indicano relazioni positive forti. I coefficienti che legano le due famiglie riguardano unicamente l'indicatore elementare PD3, il quale valuta il rapporto tra studenti Erasmus in uscita rispetto al numero di docenti di ruolo della facoltà: è naturale quindi che questo indicatore sia positivamente correlato e con valori elevati del coefficiente con RI3 (0.893), con RI1 (0.686) e con RI2 (0.598).

Tabella 10. Matrice di correlazione tra gli indicatori elementari appartenenti alla famiglia "rapporti internazionali" e gli indicatori restanti.

	RI1	RI2	RI3	RI4
P1	.550**	.597**	.203	.122
P2	.129	.181	.317	-.499**
P3	.102	-.011	.143	-.145
P4	.048	.071	.234	-.374
P5	-.045	-.069	.079	-.077
D1	-.051	-.099	-.033	.164
D2	.305	.234	.159	.087
D3	-.130	-.121	-.014	.184
D4	.217	.379	-.190	-.160
D5	.048	.130	.118	-.035
D6	-.031	.120	-.006	-.267
D7	-.236	-.192	-.206	.019
D8	-.006	-.008	-.192	.227
R1	.006	-.108	.068	-.231
R2	.418*	.312	.087	.194
R3	.044	-.116	-.126	-.055
R4	-.025	-.018	-.071	.299
R5	-.071	-.119	-.347	.369
R6	.041	.033	.103	.291
R7	.480*	.544**	.298	.226
PD1	.251	.280	.138	.329
PD2	.135	.228	-.246	.213
PD3	.686**	.598**	.893**	.030
PD4	.115	.056	.215	-.250
PD5	.315	.371	.283	.307

La tabella 10 è stata riportata a completamento delle precedenti: le relative considerazioni sulla correlazione tra gli indicatori appartenenti alla famiglia "rapporti internazionali" e quelli rientranti nelle altre famiglie sono già state precedentemente illustrate.

3.3.3 Analisi della correlazione tra famiglie

In ultima battuta abbiamo calcolato la matrice di correlazione tra i punteggi complessivi delle famiglie. Presentiamo due tabelle: nella prima i coefficienti di correlazione sono stati calcolati a partire dai dati pubblicati dal Censis relativamente alle cinque famiglie; la seconda parte invece dai punteggi delle famiglie che noi abbiamo costruito, seguendo

medesima metodologia operativa del Censis, a partire dagli indicatori elementari elaborati secondo quanto illustrato in appendice.

Tabella 11a. Matrice di correlazione tra famiglie – dati pubblicati.

	P	D	R	PD	RI
P	1	.432*	.451*	.093	.328
D		1	.381*	-.125	.118
R			1	-.208	.100
PD				1	.307
RI					1

La tabella evidenzia una relazione lineare positiva di verso positivo tra le famiglie P, D e R. I coefficienti di correlazione che legano queste tre variabili sono significativi, fissato un livello pari a 0.05, e portano a rifiutare l'ipotesi di assenza di correlazione (ovvero l'ipotesi che prevede l'uguaglianza a 0 del valore del coefficiente di correlazione). Si tratta comunque di relazioni il cui valore non è elevato per cui la forza che lega le tre famiglie considerate non dovrebbe avere impatto rilevante sui risultati finali. Anche i singoli indicatori elementari appartenenti alle famiglie produttività, didattica e ricerca erano legati da relazioni di intensità comparabile a quelle rilevate dalla matrice di correlazione delle famiglie.

Dal punto di vista sostantivo, tali correlazioni dimostrano che valutazioni positive in una delle tre famiglie si riflette in andamenti positivi anche nelle altre due famiglie considerate: se quindi, ad esempio, una facoltà riesce a laureare un buon numero di studenti in corso, la didattica offerta dalla facoltà sarà di buon livello come lo sarà la ricerca.

Tabella 11b. Matrice di correlazione tra famiglie – dati calcolati secondo strategia operativa del Censis a partire dagli indicatori elementari da noi costruiti.

	P	D	R	PD	RI
P	1	.454*	.423*	.443*	.297
D		1	.419*	.093	.111
R			1	.145	.087
PD				1	.200
RI					1

Tra le relazioni evidenziate, quelle statisticamente significative a livello 5% corrispondono alle correlazioni rilevate nella precedente tabella: i valori dei coefficienti sono leggermente diversi ma l'interpretazione che ne possiamo dare rimane la medesima. Tuttavia la matrice di correlazione calcolata a partire dagli indicatori elementari da noi costruiti mette in

evidenza un'ulteriore relazione positiva tra le famiglie "produttività" e "profilo docenti": tale correlazione è spiegata dagli elevati valori dei coefficienti di correlazione che legano l'indicatore elementare PD4 a due indicatori della famiglia P, P2 e P4.

Non si tratta però di una relazione che trova un nesso semantico: le due famiglie sono legate da indicatori che misurano la regolarità e il tasso di laureati triennali da una parte con il tasso di insegnamenti tenuti da docenti non di ruolo extra accademici dall'altra.

3.3.4 Considerazioni sui risultati dell'analisi di correlazione

Le analisi di correlazione tra indicatori elementari appartenenti alla stessa famiglia ci hanno aiutato nell'esaminare la struttura di relazioni che lega i nostri dati: non ne abbiamo però tratto, tuttavia, un risultato particolarmente significativo in termini di rilevanti correlazioni tra gli indicatori stessi. Tale tipo di indagine sui dati si rende utile per determinare – anche se in modo indiretto – il sistema di pesi da applicare agli indicatori elementari nella loro aggregazione al fine di determinare un punteggio sintetico (indicatore composto) per ciascuna famiglia. Esaminando quindi le formule stabilite dal Censis per la determinazione del punteggio finale di ciascuna famiglia, è immediato notare come non venga attribuito lo stesso peso ad ogni indicatore ma alcune coppie o terne di indicatori vengono raggruppati: si tratta di informazioni che valutano aspetti strettamente connessi da un punto di vista semantico per i quali ci si attenderebbero, quindi, elevati valori del coefficiente di correlazione. Non possiamo però basarci sui nostri risultati per stabilire la struttura di correlazioni sottostante i dati per un'unica ma estremamente importante ragione: le informazioni a nostra disposizione riguardano solamente 27 unità, troppo poche per determinare uno schema relazionale che sia universalmente valido per tutte le facoltà. Tuttavia i dati presentano alcune correlazioni, talvolta anche di un certo peso, tra indicatori appartenenti alla stessa famiglia ma considerati in maniera "separata" nella formula di aggregazione stabilita dal Censis. Questo potrebbe significare una sovrapposizione nelle misurazioni di uno stesso concetto (ovvero di quanto ciascuna famiglia intende valutare) oppure il non soddisfacimento della proprietà di validità da parte di alcuni indicatori.

Ferme restando queste considerazioni sui nostri dati, non riteniamo di aver potuto assicurare – sempre naturalmente per la ridotta numerosità degli stessi – un sufficiente

grado di incisività nell'esaminare le informazioni. Per questi motivi, ritenendo la scelta – comunque vincolata necessaria, utilizzeremo la struttura di aggregazione proposta dal Censis per l'aggregazione degli indicatori elementari della stessa famiglia, senza quindi modificare i pesi che sono stati nella formula attribuiti.

Per poter completare un'analisi preliminare sui dati avremmo voluto dedicare una specifica sezione ad analisi di tipo multivariato tese a valutare la dimensionalità statistica traibile dai singoli indicatori elementari. Un tentativo in tal senso è stato fatto ma data la dimensione effettiva del dataset a nostra disposizione non è stato possibile ricavare utili informazioni che potessero aiutarci nella costruzione di una configurazione eventualmente diversa delle famiglie ovvero delle dimensioni del fenomeno in esame.

3.4 PROPOSTE ALTERNATIVE DI ANALISI DEI DATI

3.4.1 I metodi utilizzati

Tra tutte le tecniche di normalizzazione, aggregazione e ponderazione presentate nel capitolo 2, abbiamo scelto di elaborare i dati a nostra disposizione con quelle che, fra tutte, la letteratura indica come più robuste e più convincenti. Volendo inoltre adottare tecniche tra loro coerenti, basate su approcci diversi, siamo partiti dal selezionare due metodi di normalizzazione, l'uno lineare e l'altro non lineare di tipo ordinale. Abbiamo poi stabilito l'utilizzo di due metodi diversi di aggregazione applicabili a qualsiasi normalizzazione. Infine la scelta delle metodologie di ponderazione da utilizzare si è dovuta ridurre alla selezione di un metodo partecipativo. Vedremo ora in dettaglio secondo quali tecniche sono state compiute nel complesso le nostre elaborazioni.

Premessa ad ogni spiegazione sta la decisione dell'aver voluto applicare ai nostri dati le tecniche di aggregazione e ponderazione stabilite dal Censis stesso al fine di tentare di ottenere gli stessi risultati pubblicati ma soprattutto per permettere un confronto più prudente e congruente.

Il punto di partenza per tutte le nostre elaborazioni è stato costituito dall'insieme degli indicatori elementari ottenuti a partire dalle variabili inviate dal Censis secondo quanto descritto in appendice.

La normalizzazione degli indicatori elementari è avvenuta quindi secondo tre diverse tecniche:

- relativizzazione al campo di variazione (come applicato dal Censis);
- standardizzazione o calcolo del punteggio z;
- trasformazione "rango".

Il calcolo dei valori standardizzati rientra nell'ambito delle trasformazioni lineari: si tratta di un tipo di applicazione – molto frequente in campo statistico – che converte la distribuzione di un insieme di dati nella distribuzione normale standard, avente media nulla e varianza unitaria; esso inoltre prevede l'inserimento nel calcolo di una misura (la deviazione standard che tiene conto della variabilità dei dati).

Tra le normalizzazioni non lineari abbiamo scelto una tecnica caratterizzata da estrema semplicità computazionale, di comprensione ed interpretativa: la trasformazione rango mappa i dati originali in dati ordinali, ciascuno dei quali indica la posizione che l'unità occupa nella graduatoria.

L'aggregazione degli indicatori avviene in due passi distinti: anzitutto vengono aggregati gli indicatori semplici appartenenti alla stessa famiglia per determinare il valore finale della famiglia stessa; in seconda istanza vengono aggregati i punteggi delle cinque famiglie considerate per costruire la graduatoria complessiva finale. Come abbiamo prima anticipato, le nostre scelte coinvolgeranno solo la seconda aggregazione mantenendo inalterata quella formulata dal Censis per il calcolo dei valori finali di ciascuna famiglia. I due metodi scelti per l'aggregazione sono i seguenti:

- la media aritmetica;
- l'aggregazione geometrica.

La prima, molto semplice e di immediata comprensione, consiste nell'effettuare la media aritmetica dei punteggi normalizzati delle cinque famiglie: questa tecnica è stata scelta anche dal Censis per la sintesi delle famiglie.

L'aggregazione geometrica, invece, è stata selezionata perché è un buon compromesso tra approcci di tipo lineare e l'approccio multicriterio non compensativo in termini di compensabilità.

Per la ponderazione, infine, sono stati scelti tre metodi:

- pesi uguali ad ogni famiglia;
- metodo partecipativo con allocazione di un budget da parte di esperti;
- metodo basato sulla regressione.

La prima tecnica è quella che è stata scelta anche dal Censis per ponderare ciascuna famiglia nell'indicatore composto finale complessivo; come già abbiamo avuto modo di dire, non si tratta di una "non scelta" ma di mettere tutte le famiglie sullo stesso piano.

Per la ponderazione mediante giudizio di esperti del fenomeno, ci siamo "arrogati" questo ruolo stabilendo il seguente sistema di pesi:

- alle famiglie Didattica, Ricerca e Profilo Docenti abbiamo assegnato un peso pari a 0.25;
- alle famiglie Produttività e Rapporti Internazionali abbiamo attribuito un peso di 0.125.

È stato assegnato un peso inferiore alla famiglia relativa alla produttività del sistema universitario per due ordini di motivi: anzitutto perché è legato alla composizione della componente studentesca e, in secondo luogo, perché, valutando in maniera positiva le facoltà che sostanzialmente riescono a conferire la laurea ai suoi studenti nei tempi più vicini possibili alla durata legale del corso e senza perdita di iscrizioni, potrebbe premiare quelli che sono, in realtà, comportamenti lassisti. Ai rapporti internazionali abbiamo attribuito peso inferiore perché in questa famiglia è considerato essenzialmente solo il progetto Erasmus.

Il terzo metodo di determinazione dei pesi, basato sulla costruzione di una funzione di regressione lineare, non ha potuto trovare applicazione per la non significatività statistica dei risultati dei modelli. Come variabile indipendente avevamo scelto di utilizzare il valore che ciascuna facoltà ottenne nel 2001 nella famiglia "attrazione e influenza", deputata a valutare la capacità della singola facoltà di attrarre a sé studenti rispetto alle facoltà omogenee, rispetto all'Ateneo ed in termini di iscritti provenienti da regioni diverse da quella di appartenenza dell'Ateneo. In una logica ideale di scelta della facoltà da parte delle famiglie, queste dovrebbero poter avere a disposizione il massimo delle informazioni

su tutte le facoltà per giungere ad una decisione caratterizzata da razionalità: è questo l'obiettivo a cui risponde il lavoro del Censis. Possiamo assumere che una valutazione positiva della capacità di attrazione di una facoltà sia il vero obiettivo di miglioramento ed affinamento degli Organi della singola Facoltà. Questo giustifica la scelta di assumere come variabile dipendente la misura dell'attrattività. Purtroppo non c'è una serie storica di tale dato che giunga fino al 2006: esaminando però la storia della partecipazione universitaria delle facoltà, possiamo assumere che questi indicatori abbiano una certa costanza. Non pare quindi così inadeguato prendere i valori relativi all'ultimo anno in cui questa famiglia è stata valutata, ovvero il 2001. Le variabili indipendenti dovevano coincidere con i punteggi sintetici di ciascuna famiglia secondo quindi il modello seguente:

$$At = \alpha + \beta_1 P + \beta_2 D + \beta_3 R + \beta_4 PD + \beta_5 RI$$

dove At indica il valore pubblicato nel 2001 per la famiglia della capacità di attrazione e α rappresenta l'intercetta del modello; le variabili esplicative P , D , R , PD e RI corrispondono alle cinque famiglie valutate nel rapporto del 2006 e β_1, \dots, β_5 ne costituiscono i coefficienti. Abbiamo calcolato una pluralità di modelli assumendo diverse categorie di valori per le variabili indipendenti: abbiamo utilizzato i punteggi pubblicati dal Censis e quelli da noi ottenuti secondo i tre metodi di normalizzazione sopra descritti ed ogni modello è stato calcolato con o senza intercetta. In nessun caso, però, i coefficienti delle famiglie sono risultati, nel complesso¹⁴⁵, statisticamente diversi da zero; inoltre i valori di β_1, \dots, β_5 erano molto prossimi l'un l'altro e molto bassi per cui non sarebbero stati pesi rilevanti ai nostri fini. Nei modelli nei quali era prevista l'intercetta, essa è risultata essere il solo coefficiente significativo per valori di α compresi tra 75 e 85: questo potrebbe significare l'esistenza di una forte base comune a tutte le famiglie ma potrebbe anche dire, più semplicemente, l'assenza di relazione con quella che noi abbiamo considerato variabile risposta.

In sostanza, quindi, i metodi utilizzati per la determinazione del sistema di pesi sono stati solamente i primi due.

¹⁴⁵ In alcuni modelli era presente un solo coefficiente significativo (ovvero diversi da zero) ma non è stato possibile considerarlo perché, nel complesso, il modello non forniva risultati rilevanti.

3.4.1.1 Calcolo del punteggio delle famiglie

Prima di illustrare i dati utilizzati ed i risultati ottenuti, riteniamo doveroso appuntare alcune osservazioni sulle formule di calcolo utilizzate dal Censis per aggregare gli indicatori elementari appartenenti alla stessa famiglia.

Definiamo con $n(X)$ la normalizzazione dell'indicatore X che coinciderà, di volta in volta, con la relativizzazione al campo di variazione, con il punteggio z e con il rango.

Nella famiglia della produttività (P) gli indicatori previsti sono stati aggregati secondo la formula sopra presentata: i primi tre indicatori sono stati considerati con peso uguale mentre $P4$ e $P5$, entrambi relativi al tasso di laureati, sono stati accorpati e ponderati per il numero dei laureati $n1$ ed $n2$. Poiché $P5$ ci è stato fornito come tasso (ovvero non ci sono state inviate le variabili che ne definiscono numeratore e denominatore), non ci è possibile ricavare il valore di $n2$ utilizzando quindi la medesima formula di calcolo prevista dal Censis per questa famiglia¹⁴⁶. Per questo motivo, il punteggio da noi calcolato per P è stato ottenuto come semplice media aritmetica degli indicatori elementari $P1$, $P2$, $P3$ e $P4$:

$$P = \frac{n(P1) + n(P2) + n(P3) + n(P4)}{4} .$$

La didattica (D) prevede l'utilizzo di 8 indicatori elementari la cui ponderazione nella formula di aggregazione prevista dal Censis ci ha suscitato alcune perplessità, dovute in gran misura al fatto che tutte le scelte compiute dall'ente in termini di pesi da attribuire agli indicatori semplici non sono state chiaramente illustrate in nota metodologica. Ci siamo chiesti per quale motivo a $D7$ è stato attribuito un peso di 0.5; di conseguenza poi, dal momento che il peso totale degli indicatori a numeratore è pari a 3.5 (1 + 1 + 1 + 0.5), non ci spieghiamo perché il denominatore sia pari a 4. Abbiamo però scelto di mantenere la formulazione applicata dal Censis sostanzialmente per mantenere un certo grado di coerenza con i risultati da esso ottenuti. Il punteggio finale della famiglia D è quindi ottenuto come segue:

$$D = \frac{n\left(\frac{n(D1) + n(D2) + n(D3)}{3}\right) + n(D4) + n\left(\frac{n(D5) + n(D6)}{2}\right) + 0.5 * n(D7)}{4} * k$$

$$\text{dove } k = \begin{cases} 1.00 & \text{se } D8 < 75 \\ 1.05 & \text{se } D8 \geq 75 \end{cases} .$$

¹⁴⁶ Considerando che i valori di $P4$ e di $P5$ si differenziano molto tra facoltà, ci rendiamo conto che questa è una decisione opinabile anche dal punto di vista semantico ma è stata, purtroppo, una scelta obbligata.

Anche il calcolo del valore finale di R, che misura la famiglia della ricerca, è rimasto inalterato rispetto a quello utilizzato dal Censis:

$$R = \frac{n\left(\frac{n(R1) + n(R2) + n(R3)}{3}\right) + n\left(\frac{n(R4) + n(R5) + n(R6)}{3}\right)}{2} * k$$

$$\text{con } k = \begin{cases} 1.00 & \text{se } R7 = 0 \\ 1.05 & \text{se } R7 > 0 \end{cases} .$$

Nell'esaminare il profilo docenti (PD), sono emersi dubbi analoghi a quelli rilevati per la famiglia della didattica: a numeratore la somma dei pesi è pari a 2 (1 + 0.5 + 0.5) ma il denominatore della formula è 3. Anche in questo caso, però, conformemente a quanto scelto per la famiglia D, abbiamo mantenuto la formula di calcolo trovata in nota metodologica:

$$PD = \frac{n\left(\frac{n(PD1) + n(PD2)}{2}\right) + 0.5 * n(PD3) + 0.5 * n(PD4)}{3} * k$$

$$\text{dove } k = \begin{cases} 1.00 & \text{se } PD5 = 0 \\ 1.05 & \text{se } PD5 > 0 \end{cases} .$$

Infine il valore della famiglia che valuta i rapporti internazionali delle facoltà (RI) è stato ottenuto, come dal Censis, secondo la seguente formula:

$$RI = \frac{n(RI1) + n(RI2) + n(RI3)}{3} * k$$

$$\text{con } k = \begin{cases} 1.00 & \text{se } RI4 = 0 \\ 1.05 & \text{se } RI4 > 0 \end{cases} .$$

3.4.1.2 Sintesi degli approcci adottati

Volendo sintetizzare quanto illustrato all'inizio del paragrafo, riportiamo un semplice schema per raccogliere le diverse elaborazioni che abbiamo compiuto sui dati.

	Valori relativizzati al campo di variazione	Punteggi z	Ranghi
Media	pesi uguali	pesi uguali	pesi uguali
	pesi da esperti	pesi da esperti	pesi da esperti
Aggregazione geometrica	pesi uguali	pesi uguali	pesi uguali
	pesi da esperti	pesi da esperti	pesi da esperti

Sono quindi 12 le graduatorie che si otterranno dall'applicazione dei tre diversi metodi di normalizzazione, delle due tecniche di aggregazione e dei due sistemi di pesi sopra descritti. Ciascuna di esse ha determinato un unico indicatore di sintesi per facoltà; il punto di partenza sono stati i punteggi di sintesi normalizzati di ciascuna famiglia. Prima di riportare le formule applicate, illustriamo anzitutto le notazioni utilizzate.

Abbiamo denominato cvP il punteggio di sintesi della famiglia della produttività relativizzato al campo di variazione, zP il valore di sintesi della famiglia P trasformato in punteggio z e rP quello trasformato in rango. Analogamente abbiamo definito i punteggi delle altre famiglie (D, R, PD e RI).

Di seguito presentiamo quindi le formule di calcolo che hanno determinato le 12 graduatorie finali.

A1 : media con pesi uguali ai punteggi delle famiglie relativizzate al campo di variazione

$$A1 = \frac{cvP + cvD + cvR + cvPD + cvRI}{5}$$

A2 : media con pesi uguali ai punteggi delle famiglie trasformati in punteggi z

$$A2 = \frac{zP + zD + zR + zPD + zRI}{5}$$

A3 : media con pesi uguali ai punteggi delle famiglie trasformati in ranghi

$$A3 = \frac{rP + rD + rR + rPD + rRI}{5}$$

B1 : media con pesi determinati da esperti ai punteggi delle famiglie relativizzate al campo di variazione

$$B1 = \frac{(cvP * 0.125) + (cvD * 0.25) + (cvR * 0.25) + (cvPD * 0.25) + (cvRI * 0.125)}{5}$$

B2 : media con pesi determinati da esperti ai punteggi delle famiglie trasformati in punteggi z

$$B2 = \frac{(zP * 0.125) + (zD * 0.25) + (zR * 0.25) + (zPD * 0.25) + (zRI * 0.125)}{5}$$

B3 : media con pesi determinati da esperti ai punteggi delle famiglie trasformati in ranghi

variazione

$$B3 = \frac{(rP * 0.125) + (rD * 0.25) + (rR * 0.25) + (rPD * 0.25) + (rRI * 0.125)}{5}$$

C1 : aggregazione geometrica con pesi uguali dei punteggi delle famiglie relativizzate al campo di variazione

$$C1 = \left(cvP^{\frac{1}{5}} \right) * \left(cvD^{\frac{1}{5}} \right) * \left(cvR^{\frac{1}{5}} \right) * \left(cvPD^{\frac{1}{5}} \right) * \left(cvRI^{\frac{1}{5}} \right)$$

C2 : aggregazione geometrica con pesi uguali dei punteggi delle famiglie trasformati in punteggi z

$$C2 = \left(zP^{\frac{1}{5}} \right) * \left(zD^{\frac{1}{5}} \right) * \left(zR^{\frac{1}{5}} \right) * \left(zPD^{\frac{1}{5}} \right) * \left(zRI^{\frac{1}{5}} \right)$$

C3 : aggregazione geometrica con pesi uguali dei punteggi delle famiglie trasformati in ranghi

$$C3 = \left(rP^{\frac{1}{5}} \right) * \left(rD^{\frac{1}{5}} \right) * \left(rR^{\frac{1}{5}} \right) * \left(rPD^{\frac{1}{5}} \right) * \left(rRI^{\frac{1}{5}} \right)$$

D1 : aggregazione geometrica con pesi determinati da esperti dei punteggi delle famiglie relativizzate al campo di variazione

$$D1 = \left(cvP^{\frac{0.125}{5}} \right) * \left(cvD^{\frac{0.25}{5}} \right) * \left(cvR^{\frac{0.25}{5}} \right) * \left(cvPD^{\frac{0.25}{5}} \right) * \left(cvRI^{\frac{0.125}{5}} \right)$$

D2 : aggregazione geometrica con pesi determinati da esperti dei punteggi delle famiglie trasformati in punteggi z

$$D2 = \left(zP^{\frac{0.125}{5}} \right) * \left(zD^{\frac{0.25}{5}} \right) * \left(zR^{\frac{0.25}{5}} \right) * \left(zPD^{\frac{0.25}{5}} \right) * \left(zRI^{\frac{0.125}{5}} \right)$$

D3 : aggregazione geometrica con pesi determinati da esperti dei punteggi delle famiglie trasformati in ranghi

$$D3 = \left(rP^{\frac{0.125}{5}} \right) * \left(rD^{\frac{0.25}{5}} \right) * \left(rR^{\frac{0.25}{5}} \right) * \left(rPD^{\frac{0.25}{5}} \right) * \left(rRI^{\frac{0.125}{5}} \right)$$

Lavoreremo prevalentemente sulle graduatorie degli indicatori composti complessivi che sintetizzano le cinque famiglie: alle 12 graduatorie ottenute dall'applicazione delle tecniche

di elaborazione da noi scelte, aggiungeremo la classifica pubblicata dal Censis e procederemo quindi al confronto come verrà nel seguito descritto.

3.4.2 I risultati

Nei paragrafi seguenti illustreremo i risultati ottenuti dall'applicazione delle tecniche di normalizzazione, aggregazione e ponderazione sopra illustrate. Una prima serie di tabelle analizza le graduatorie per famiglia mentre ci siamo soffermati, con maggior dettaglio, nell'esame delle graduatorie complessive.

Il confronto è stato effettuato attraverso il calcolo dell'indice ρ di Spearman il quale, come spiegato nel secondo capitolo, calcola la cograduazione in base alla posizione che ciascuna unità assume nelle due classifiche considerate. Stabiliamo anzitutto che la graduatoria pubblicata dal Censis sia quella di riferimento: questa scelta trova motivazione nella notorietà che la graduatoria stessa possiede – in particolare nel mondo accademico – perché ampiamente diffusa attraverso la stampa nazionale e non tanto, quindi, per un'assoluta, indiscutibile ed inopinabile veridicità e coincidenza con quanto la realtà ci fa osservare.

3.4.2.1 Le graduatorie per famiglia

Aggregando, attraverso le formule fornite dal Censis, gli indicatori elementari normalizzati secondo le tre diverse tecniche di trasformazione considerate (valori relativizzati al campo di variazione, punteggi z e ranghi), abbiamo ottenuto tre diverse graduatorie per ciascuna famiglia. Di seguito proponiamo una tabella per ciascuna famiglia in cui riportiamo in riga le facoltà ed in ogni cella la posizione che essa occupa nelle tre graduatorie da noi costruite e nella graduatoria pubblicata dal Censis.

Il punteggio **cv** è stato ottenuto aggregando indicatori elementari relativizzati al campo di variazione, il punteggio **z** corrisponde alla standardizzazione dei valori mentre **r** indica i valori trasformati in rango; l'ultima colonna è quella relativa alla graduatoria pubblicata.

L'ultima riga di ogni tabella riporta il valore dell'indice ρ di Spearman, misura del confronto tra la graduatoria stabilita da ogni colonna e quella pubblicata dal Censis.

Tabella 12. Graduatorie per la famiglia "produttività" con relativo indice di Spearman.

FACOLTÀ	cv	z	r	CENSIS
Bari	22	22	20	21
Bologna	7	9	10	7
Cagliari	10	10	8	9
Calabria - Cosenza	6	7	6	6
Catania	16	15	15	16
Firenze	9	6	12	9
Genova	19	19	17	16
Macerata	18	17	17	16
Messina	21	21	21	22
Milano 1	11	11	8	11
Napoli Orientale	13	13	12	12
Napoli 1 - Federico II	24	24	24	22
Padova	3	3	2	2
Palermo	27	27	26	27
Pavia	5	5	6	4
Perugia	8	8	5	7
Piemonte Orientale	1	1	1	1
Pisa	25	25	25	25
Roma 1	26	26	27	26
Roma 3	12	12	12	12
Salerno	20	20	22	20
Sassari	17	16	19	15
Siena	14	14	11	14
Teramo	15	18	15	16
Torino	4	4	4	5
Trieste	2	2	3	3
Urbino - Pesaro	23	23	23	24
ρ di Spearman	0.9902	0.9866	0.9759	

La cograduazione tra le graduatorie calcolate secondo le nostre trasformazioni degli indicatori elementari e la classifica pubblicata dal Censis per la famiglia della produttività è prossima ad 1: il confronto quindi evidenzia alta similarità tra le diverse graduatorie, aspetto reso evidente dalla minima distanza tra il rango di una stessa unità in due graduatorie diverse.

Tabella 13. Graduatorie per la famiglia "didattica" con relativo indice di Spearman.

FACOLTÀ	cv	z	r	CENSIS
Bari	25	24	24	24
Bologna	13	13	17	12
Cagliari	15	15	8	15
Calabria - Cosenza	7	7	15	5
Catania	24	26	26	25
Firenze	12	12	13	12
Genova	6	6	10	6
Macerata	27	27	27	27
Messina	19	19	17	20
Milano 1	4	4	4	4
Napoli Orientale	18	18	14	18
Napoli 1 - Federico II	23	23	25	23
Padova	21	21	19	21
Palermo	20	20	16	18
Pavia	2	2	2	1
Perugia	16	16	7	17
Piemonte Orientale	1	1	1	2
Pisa	17	17	12	16
Roma 1	10	11	21	9
Roma 3	22	22	22	21
Salerno	5	5	11	6
Sassari	26	25	23	25
Siena	8	9	5	8
Teramo	14	14	9	12
Torino	3	3	3	3
Trieste	9	10	6	9
Urbino - Pesaro	11	8	20	9
ρ di Spearman	0.9915	0.9815	0.7958	

Anche per la famiglia della didattica la cograduazione è buona: è molto elevata per il confronto tra classifica pubblicata e le graduatorie costruite con i valori relativizzati al campo di variazione e con i punteggi z mentre l'indice assume un valore più basso (0.7958) nel confronto con la graduatoria costruita trasformando gli indicatori elementari in ranghi.

Tabella 14. Graduatorie per la famiglia "ricerca" con relativo indice di Spearman.

FACOLTÀ	cv	z	r	CENSIS
Bari	23	23	26	21
Bologna	8	8	4	7
Cagliari	24	24	22	21
Calabria - Cosenza	9	9	13	8
Catania	19	19	17	18
Firenze	2	2	1	2
Genova	25	25	24	21
Macerata	22	22	20	20
Messina	17	17	15	15
Milano 1	3	4	2	6
Napoli Orientale	27	27	27	27
Napoli 1 - Federico II	15	16	18	24
Padova	11	11	14	10
Palermo	26	26	23	26
Pavia	13	13	10	15
Perugia	14	14	10	12
Piemonte Orientale	1	1	3	1
Pisa	4	3	9	3
Roma 1	21	21	25	25
Roma 3	7	7	8	12
Salerno	20	20	21	19
Sassari	6	5	7	4
Siena	5	6	5	4
Teramo	18	18	19	15
Torino	16	15	10	14
Trieste	10	12	16	8
Urbino - Pesaro	12	10	5	10
ρ di Spearman	0.9365	0.9420	0.8877	

L'indice ρ di Spearman che confronta le tre graduatorie costruite con i valori normalizzati secondo le tecniche da noi scelte con la graduatoria costruita e pubblicata dal Censis è molto elevato, indicando quindi una concordanza quasi massima.

Tabella 15. Graduatorie per la famiglia "profilo docenti" con relativo indice di Spearman.

FACOLTÀ	CV	Z	r	CENSIS
Bari	24	23	24	20
Bologna	2	2	3	5
Cagliari	5	6	6	3
Calabria - Cosenza	1	1	1	6
Catania	22	22	23	20
Firenze	13	14	15	18
Genova	19	19	19	16
Macerata	14	13	13	11
Messina	26	25	25	25
Milano 1	6	7	7	4
Napoli Orientale	11	11	10	8
Napoli 1 - Federico II	25	26	25	25
Padova	12	12	11	11
Palermo	10	10	9	7
Pavia	17	17	17	16
Perugia	18	18	18	14
Piemonte Orientale	9	9	12	27
Pisa	27	27	27	23
Roma 1	23	24	20	20
Roma 3	3	3	4	1
Salerno	16	16	16	10
Sassari	8	8	5	9
Siena	21	21	21	18
Teramo	4	4	2	2
Torino	20	20	21	23
Trieste	7	5	8	14
Urbino - Pesaro	15	15	14	11
ρ di Spearman	0.8111	0.8025	0.8538	

Analogamente a quanto accade per le prime tre famiglie analizzate, la cograduazione delle graduatorie relative alla famiglia "profilo docenti" rivela, attraverso il valore dell'indice di Spearman, una buona concordanza tra le classifiche da noi calcolate e quella costruita dal Censis.

Tabella 16. Graduatorie per la famiglia "rapporti internazionali" con relativo indice di Spearman.

FACOLTÀ	cv	z	r	CENSIS
Bari	24	25	25	27
Bologna	3	3	3	3
Cagliari	14	14	16	12
Calabria - Cosenza	23	22	22	19
Catania	25	24	24	25
Firenze	1	1	2	3
Genova	2	2	1	1
Macerata	21	21	21	22
Messina	19	18	18	19
Milano 1	18	19	19	16
Napoli Orientale	9	9	8	8
Napoli 1 - Federico II	27	27	26	25
Padova	12	12	11	14
Palermo	20	20	20	18
Pavia	8	8	8	7
Perugia	6	6	4	6
Piemonte Orientale	26	26	26	24
Pisa	22	23	23	19
Roma 1	10	11	10	10
Roma 3	5	5	6	5
Salerno	13	13	14	17
Sassari	11	10	12	10
Siena	7	7	7	8
Teramo	16	16	13	13
Torino	15	15	14	15
Trieste	4	4	4	2
Urbino - Pesaro	17	17	16	23
ρ di Spearman	0.9594	0.9588	0.9542	

Torna invece ad essere positivo e prossimo ad 1 il valore dei coefficienti ρ di Spearman per la famiglia dei rapporti internazionali, evidenziando quindi una buona cograduazione tra le graduatorie da noi costruite e quella pubblicata dal Censis.

3.4.2.2 Le graduatorie finali

Utilizzando le formule sopra riportate, abbiamo ottenuto, per ogni applicazione, un elenco di 27 valori che abbiamo provveduto ad ordinare (secondo valore decrescente per i valori "cv" e "z" ma in ordine crescente per i ranghi "r") al fine di ottenere una graduatoria delle facoltà. Il risultato è presentato nella tabella 12 che presenta in ogni cella la posizione che la singola facoltà – definita in riga – risulta avere nelle diverse graduatorie; nell'ultima colonna è riportata la posizione delle facoltà nella graduatoria pubblicata dal Censis.

Tabella 17. Graduatorie finali per facoltà.

FACOLTÀ	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	Censis
Bari	26	26	27	27	27	27	21	26	27	21	1	27	26
Bologna	2	3	1	2	2	3	1	14	3	1	9	4	1
Cagliari	15	16	14	14	16	13	14	17	15	13	23	13	12
Calabria - Cosenza	7	6	11	3	4	9	8	23	8	6	24	5	8
Catania	25	24	25	25	24	25	20	21	25	20	5	25	25
Firenze	3	2	4	4	3	5	3	12	2	3	9	3	5
Genova	10	13	16	15	15	18	12	25	14	16	25	16	9
Macerata	23	23	23	24	25	23	23	15	23	23	9	23	21
Messina	21	21	21	23	22	22	22	19	22	22	7	22	23
Milano 1	8	8	3	7	7	1	6	22	4	4	26	2	7
Napoli Orientale	17	19	16	19	19	17	23	5	17	23	9	18	17
Napoli 1 - Federico II	27	27	26	26	26	26	23	24	26	23	6	26	27
Padova	12	12	11	13	14	14	11	10	12	12	9	15	15
Palermo	24	25	20	22	23	20	23	2	21	23	3	21	24
Pavia	5	5	4	6	5	5	4	16	6	5	9	6	2
Perugia	9	9	7	11	12	7	7	7	7	10	9	11	10
Piemonte Orientale	1	1	4	1	1	2	17	27	1	9	27	1	6
Pisa	22	20	21	20	20	21	23	1	20	23	2	20	20
Roma 1	20	22	24	21	21	24	19	6	24	19	9	24	22
Roma 3	6	7	9	8	8	9	5	20	11	7	4	12	4
Salerno	18	17	19	17	18	19	16	4	19	17	8	19	18
Sassari	16	15	15	16	13	14	15	9	16	15	9	14	16
Siena	11	10	8	10	10	8	9	11	10	11	9	8	11
Teramo	13	14	13	9	9	12	10	8	13	8	9	10	13
Torino	14	11	9	12	11	9	13	18	9	14	22	9	14
Trieste	4	4	1	5	6	4	2	3	5	2	9	7	3
Urbino - Pesaro	19	18	18	18	17	16	18	13	18	18	9	17	19

Da una prima semplice analisi è immediato verificare come la posizione delle singole facoltà non risulti particolarmente variabile come non si segnalino grosse oscillazioni per la maggior parte delle applicazioni ed anche rispetto alla classificazione stabilita dal Censis. Forti scostamenti si presentano però per la graduatoria C2 e ancor più per D2: questo sicuramente dipende dalle difficoltà di calcolo delle stesse per cui i programmi utilizzati non hanno fornito valori precisi perché troppo piccoli o troppo elevati. Per D2, ad esempio i valori di ben 13 facoltà sono stati valutati pari a 0 assegnando ad esse rango 9 ed alla prima facoltà con valore diverso da 0 è stato attribuito rango 22. Entrambe le applicazioni prevedono l'utilizzo dell'aggregazione geometrica dei punteggi standardizzati z : probabilmente la combinazione di queste due tecniche non porta ad ottenere risultati affidabili per una loro incompatibilità. Anche per le graduatorie C1 e D1 c'è la presenza di ranghi ripetuti: la procedura di calcolo ha assegnato valori pari a 0 per le ultime cinque unità della graduatoria alle quali è stato quindi attribuito rango 23.

Nella pagina successiva riportiamo una tabella analoga alla precedente ma nelle cui celle sono disposte le facoltà ed è fissata per riga la posizione in graduatoria: pensiamo sia utile al lettore perché consente un impatto visivo più immediato nell'individuare la posizione di ogni singola facoltà.

Tabella 18. Graduatorie finali per posizione.

Rango	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	Censis
1	Piemonte	Piemonte	Bologna	Piemonte	Piemonte	Milano 1	Bologna	Pisa	Piemonte	Bologna	Bari	Piemonte	Bologna
2	Bologna	Firenze	Trieste	Bologna	Bologna	Piemonte	Trieste	Palermo	Firenze	Trieste	Pisa	Milano 1	Pavia
3	Firenze	Bologna	Milano 1	Calabria	Firenze	Bologna	Firenze	Trieste	Bologna	Firenze	Palermo	Firenze	Trieste
4	Trieste	Trieste	Firenze	Firenze	Calabria	Trieste	Pavia	Salerno	Milano 1	Milano 1	Roma 3	Bologna	Roma 3
5	Pavia	Pavia	Pavia	Trieste	Pavia	Firenze	Roma 3	Napoli Or.	Trieste	Pavia	Catania	Calabria	Firenze
6	Roma 3	Calabria	Piemonte	Pavia	Trieste	Pavia	Milano 1	Roma 1	Pavia	Calabria	Napoli 1	Pavia	Piemonte
7	Calabria	Roma 3	Perugia	Milano 1	Milano 1	Perugia	Perugia	Perugia	Perugia	Roma 3	Messina	Trieste	Milano 1
8	Milano 1	Milano 1	Siena	Roma 3	Roma 3	Siena	Calabria	Teramo	Calabria	Teramo	Salerno	Siena	Calabria
9	Perugia	Perugia	Roma 3	Teramo	Teramo	Calabria	Siena	Sassari	Torino	Piemonte	Bologna	Torino	Genova
10	Genova	Siena	Torino	Siena	Siena	Roma 3	Teramo	Padova	Siena	Perugia	Firenze	Teramo	Perugia
11	Siena	Torino	Calabria	Perugia	Torino	Torino	Padova	Siena	Roma 3	Siena	Macerata	Perugia	Siena
12	Padova	Padova	Padova	Torino	Perugia	Teramo	Genova	Firenze	Padova	Padova	Napoli Or.	Roma 3	Cagliari
13	Teramo	Genova	Teramo	Padova	Sassari	Cagliari	Torino	Urbino	Teramo	Cagliari	Padova	Cagliari	Teramo
14	Torino	Teramo	Cagliari	Cagliari	Padova	Padova	Cagliari	Bologna	Genova	Torino	Pavia	Sassari	Torino
15	Cagliari	Sassari	Sassari	Genova	Genova	Sassari	Sassari	Macerata	Cagliari	Sassari	Perugia	Padova	Padova
16	Sassari	Cagliari	Genova	Sassari	Cagliari	Urbino	Salerno	Pavia	Sassari	Genova	Roma 1	Genova	Sassari
17	Napoli Or.	Salerno	Napoli Or.	Salerno	Urbino	Napoli Or.	Piemonte	Cagliari	Napoli Or.	Salerno	Sassari	Urbino	Napoli Or.
18	Salerno	Urbino	Urbino	Urbino	Salerno	Genova	Urbino	Torino	Urbino	Urbino	Siena	Napoli Or.	Salerno
19	Urbino	Napoli Or.	Salerno	Napoli Or.	Napoli Or.	Salerno	Roma 1	Messina	Salerno	Roma 1	Teramo	Salerno	Urbino
20	Roma 1	Pisa	Palermo	Pisa	Pisa	Palermo	Catania	Roma 3	Pisa	Catania	Trieste	Pisa	Pisa
21	Messina	Messina	Messina	Roma 1	Roma 1	Pisa	Bari	Catania	Palermo	Bari	Urbino	Palermo	Macerata
22	Pisa	Roma 1	Pisa	Palermo	Messina	Messina	Messina	Milano 1	Messina	Messina	Torino	Messina	Roma 1
23	Macerata	Macerata	Macerata	Messina	Palermo	Macerata	Macerata	Calabria	Macerata	Macerata	Cagliari	Macerata	Messina
24	Palermo	Catania	Roma 1	Macerata	Catania	Roma 1	Napoli Or.	Napoli 1	Roma 1	Napoli Or.	Calabria	Roma 1	Palermo
25	Catania	Palermo	Catania	Catania	Macerata	Catania	Napoli 1	Genova	Catania	Napoli 1	Genova	Catania	Catania
26	Bari	Bari	Napoli 1	Napoli 1	Napoli 1	Napoli 1	Palermo	Bari	Napoli 1	Palermo	Milano 1	Napoli 1	Bari
27	Napoli 1	Napoli 1	Bari	Bari	Bari	Bari	Pisa	Piemonte	Bari	Pisa	Piemonte	Bari	Napoli 1

Per sintetizzare le informazioni relative alle 13 graduatorie, abbiamo scelto l'utilizzo di due indici di posizione non affetti da outliers per cui anche la presenza delle graduatorie C2 e D2 non distorce i risultati. Il primo di questi due indici è la moda: essa fornisce per ciascuna facoltà la posizione che essa assume nella maggioranza delle graduatorie ovvero la posizione assunta con maggior frequenza. Il secondo indice utilizzato è la mediana: ovvero il valore che occupa la posizione centrale nell'insieme delle posizioni assunte da ciascuna facoltà; in altre parole metà dei valori di cella sono superiori alla mediana, mentre l'altra metà è valore inferiore.

Tabella 19. Sintesi delle 12 graduatorie finali e graduatoria pubblicata.

	Moda	Mediana	Graduatoria pubblicata
Bari	27	26.5	26
Bologna	1 / 2 / 3	2.5	1
Cagliari	13 / 14	14.5	12
Calabria - Cosenza	6 / 8	7.5	8
Catania	25	24.5	25
Firenze	3	3	5
Genova	16	15.5	9
Macerata	23	23	21
Messina	22	22	23
Milano 1	4 / 8 / 7	6.5	7
Napoli Orientale	17 / 19	17.5	17
Napoli 1 - Federico II	26	26	27
Padova	12	12	15
Palermo	23	21.5	24
Pavia	5	5	2
Perugia	7	9	10
Piemonte Orientale	1	1.5	6
Pisa	20	20	20
Roma 1	24	21	22
Roma 3	7 / 8 / 9	8	4
Salerno	19	17.5	18
Sassari	15	15	16
Siena	10	10	11
Teramo	9 / 13	10	13
Torino	9	11.5	14
Trieste	4	4	3
Urbino - Pesaro	18	18	19

Nelle colonna relativa alla moda, alcune celle prevedono la presenza di più valori: si tratta di casi in cui non esiste un'unica moda perché le posizioni delle unità riportate nella cella hanno la medesima frequenza nel complesso delle 12 graduatorie considerate. La mediana evita questo problema ma presenta valori decimali perché le graduatorie prese in

considerazione sono in numero pari. Il confronto tra le tre colonne rileva una distanza contenuta tra le posizioni della medesima unità rispetto a moda e mediana mentre lo scarto diventa, in generale, più ampio quando questi due indici vengono comparati alla graduatoria pubblicata dal Censis. Con riguardo a tale confronto si noti la generale congruenza dei risultati fatte salve alcune realtà per cui la distanza tra Censis e valore mediano sembrano piuttosto rilevanti: Genova, Piemonte Orientale, Roma 3 e in misura minore Padova e Torino. Il coefficiente di cograduazione di Spearman tra la graduatoria calcolata come mediana delle 12 graduatorie da noi costruite e la classifica pubblicata è molto buono: un valore di 0.956 indica una relazione positiva tra le due graduatorie.

Dopo aver confrontato la graduatoria pubblicata dal Censis con la sintesi delle 12 graduatorie da noi elaborate, vogliamo presentare una tabella i cui valori sintetizzano tutte le 13 classifiche considerate.

Tabella 20. Sintesi di tutte le graduatorie finali.

	Moda	Mediana
Bari	27	26
Bologna	1	2
Cagliari	13 / 14	14
Calabria - Cosenza	8	8
Catania	25	25
Firenze	3	3
Genova	16	15
Macerata	23	23
Messina	22	22
Milano 1	7	7
Napoli Orientale	17	17
Napoli 1 - Federico II	26	26
Padova	12	12
Palermo	23	22
Pavia	5	5
Perugia	7	9
Piemonte Orientale	1	2
Pisa	20	20
Roma 1	24	21
Roma 3	4 / 7 / 8 / 9	8
Salerno	19	18
Sassari	15 / 16	15
Siena	10 / 11	10
Teramo	13	10
Torino	9	12
Trieste	4	4
Urbino - Pesaro	18	18

Anche in questa occasione, nei casi in cui non esiste un unico valore per la moda, abbiamo riportato tutte le posizioni ugualmente frequenti per la medesima facoltà nelle 13 graduatorie considerate. I valori della mediana non presentano valori decimali perché il numero delle classifiche che ne hanno determinato il valore è dispari: è quindi possibile individuare un valore centrale alla distribuzione di frequenza che sia superiore al 50% dei valori ed inferiore all'altra metà dei valori.

3.4.2.3 I confronti tra graduatorie

Anche il confronto tra le 12 graduatorie da noi costruite con la classifica pubblicata dal Censis è avvenuto utilizzando l'indice ρ di Spearman che appunto valuta la cograduazione di una coppia di graduatorie.

Tabella 21. Indici di cograduazione di Spearman tra le graduatorie finali da noi prodotte e la graduatoria complessiva pubblicata dal Censis.

Graduatoria	ρ di Spearman
A1	0.975
A2	0.960
B1	0.946
B2	0.938
A3	0.936
C3	0.934
D1	0.921
B3	0.920
D3	0.911
C1	0.904
C2	-0.074
D2	-0.476

La cograduazione più elevata (0.975) è quella relativa al confronto tra A1 e la graduatoria pubblicata dal Censis. Le tecniche di normalizzazione, aggregazione e ponderazione utilizzate nella costruzione delle due graduatorie sono le medesime; cambiano unicamente gli indicatori elementari di partenza, per cui però non c'è riscontro sulla corrispondenza tra quelli da noi costruiti e quelli effettivamente utilizzati dal Censis. In ogni caso, considerando la premessa appena illustrata, è assolutamente verosimile che il valore dell'indice ρ sia prossimo ad 1, indicando una cograduazione quasi massima.

Dalla tabella sopra riportata possiamo evidenziare l'elevato valore dell'indice di Spearman (mai inferiore a 0.9) che indica un notevole grado di cograduazione tra la graduatoria del Censis e tutte le graduatorie che utilizzano dati normalizzati in ranghi (nell'ordine A3, C3, B3 e D3).

Valori comunque prossimi ad 1 segnalano una buona concordanza della classifica del Censis con le graduatorie costruite utilizzando la media aritmetica come tecnica di aggregazione (tutte le graduatorie A e B) mentre più bassi risultano i coefficienti di cograduazione delle graduatorie ottenute con l'utilizzo dell'aggregazione geometrica (di tipo cioè C e D).

Gli indici di cograduazione più bassi si registrano per le graduatorie C2 e D2, sulla cui valenza abbiamo già espresso notevoli perplessità. Il valore del ρ di Spearman per C2 è prossimo allo 0: questo significa che tra le due graduatorie non è ipotizzabile alcun grado di relazione lineare¹⁴⁷. In particolare, poi l'indice associato a D2 è negativo (-0.476) indicando che le due graduatorie sono molto diverse tra loro ed, anzi, esprimono in generale un orientamento opposto l'una all'altra.

Infine l'attribuzione di pesi diversi alle famiglie diventa rilevante nella determinazione dei risultati finali: le graduatorie del tipo D, presentano, in generale, coefficienti di Spearman inferiori a quelli relativi alle tecniche che prevedono l'utilizzo di pesi uguali per ciascuna famiglia. Questo dimostra che anche in presenza di pochi dati, come nel nostro caso, la ponderazione gioca un ruolo importante nella definizione delle graduatorie finali.

3.4.2.4 Considerazioni conclusive sui risultati

Dati i risultati fin qui illustrati in termini di graduatorie, per famiglia e complessive, vogliamo presentare alcune interessanti considerazioni emersi dalle nostre elaborazioni.

Per quanto riguarda le graduatorie per famiglia, le elaborazioni non sono scese in profondità perché abbiamo voluto focalizzare l'attenzione sulla costruzione di alternative graduatorie finali soprattutto in ragione del fatto che abbiamo scelto di mantenere le formule di aggregazione degli indicatori elementari – interni alle famiglie – proposte dal Censis senza proporre di alternativi. Questa scelta è stata motivata dal nostro voler rendere le graduatorie finali maggiormente comparabili nonché da un'analisi di correlazione interna alle famiglie non del tutto soddisfacente, anche per la ridotta numerosità dei dati. Abbiamo comunque rilevato una buona concordanza tra le

¹⁴⁷ L'indice ρ di Spearman, in assenza di ranghi ripetuti, coincide infatti con il coefficiente di correlazione.

graduatorie da noi costruite – secondo i tre metodi di normalizzazione adottati per le nostre elaborazioni – e la classifica per famiglia pubblicata dal Censis.

Soffermando invece la nostra attenzione alla costruzione delle graduatorie complessive, due sono sostanzialmente gli ordini di osservazioni che illustreremo.

Tentando una combinazione delle graduatorie, utilizzando diversi metodi di normalizzazione, di aggregazione e di ponderazione dei dati, abbiamo potuto rilevare come la mediana diventi un buon indice di sintesi¹⁴⁸. Una possibilità di elaborazione alternativa che si profila può consistere quindi nell'effettuazione di una molteplicità di elaborazioni – che utilizzano differenti tecniche di trasformazione, aggregazione e ponderazione dei dati – da sintetizzare poi in un'unica graduatoria finale costruita come mediana delle graduatorie determinate dai diversi metodi utilizzati.

Tale approccio prevede quindi la necessità di elaborare più graduatorie ottenute attraverso l'applicazione di diversi metodi da sintetizzare poi in un unico valore complessivo: questo impianto operativo incrementerebbe in maniera esponenziale la complessità computazionale che il Censis dovrebbe sostenere nelle sue analisi – la cui mole è ben più ampia di quella esaminata nelle nostre elaborazioni – ma garantirebbe maggior robustezza e minor sensibilità dei risultati finali.

In secondo luogo, a partire dall'analisi di cograduazione svolta sulle graduatorie da noi costruite rispetto alla classifica pubblicata dal Censis, è possibile rilevare come l'utilizzo dei ranghi porti a risultati vicini a quelli pubblicati: trattandosi di una tecnica di semplice utilizzo, questa potrebbe risultare una buona alternativa di normalizzazione dei dati alla relativizzazione degli indicatori elementari al loro campo di variazione. Visti inoltre i soddisfacenti risultati ottenuti in termini di concordanza nella graduazione delle unità con la classifica delle facoltà proposta dal Censis, riteniamo che l'aggregazione geometrica, quando non utilizza i punteggi z , sia una buona tecnica di aggregazione: essa si basa su una logica di non completa compensabilità tra gli indicatori e, in ragione di questo, permette di attribuire ai pesi significato di importanza (anziché di trade-off come nel caso delle aggregazioni lineari), garantendo maggiore comprensione ed interpretabilità dei risultati.

¹⁴⁸ La mediana è un buon indice di sintesi per le proprietà matematiche ad essa intrinseche e perché la graduatoria prodotta dal suo utilizzo non dista molto da quella pubblicata dal Censis.

CONCLUSIONI

*"Nella vita non sono i segni quelli che mancano.
Quello che manca è il codice."
(D. Pennac)*

La valutazione delle istituzioni universitarie suscita in diverse categorie di soggetti, dalle famiglie agli studenti universitari futuri ma in particolare nel mondo accademico, fermento, attese e critiche. Con questa tesi abbiamo cercato di andare al cuore del problema, partendo dal lavoro che dal 2000 il Censis annualmente propone: "La Grande Guida all'Università" è un rapporto sullo *status* delle università italiane, le quali vengono posizionate in una classifica secondo il punteggio ottenuto in alcuni aspetti ritenuti fondamentali per un'istituzione universitaria. Questo è stato il nostro "via": dall'analisi della struttura di questa graduatoria, siamo partiti per capire quali strumenti utilizzare nella misurazione, come metterli assieme e come poter giungere a risultati complessivi solidi. Abbiamo individuato negli indicatori lo strumento di base su cui operare per valutare un concetto o un fenomeno: naturalmente essi necessitano di opportune elaborazioni perché assumano i caratteri dell'interpretabilità, della facile comprensibilità e della robustezza. Per questo la sintesi degli indicatori assume un ruolo chiave nei processi valutativi: del complesso processo di costruzione di un indicatore composto abbiamo messo in risalto le fasi della normalizzazione, dell'aggregazione e della ponderazione, illustrando, per ciascuna di esse, una molteplicità di tecniche basate su teorie ed adatte a scopi di tipo diverso. Nell'ottica di voler fornire una possibilità di confronto tra più situazioni, oltre a dare segnali d'allarme per i singoli aspetti, l'analisi deve prevedere la sintetizzazione delle informazioni, naturalmente con le dovute cautele.

Questo è quindi quello che abbiamo cercato di fare: applicare tecniche operative diverse ai dati per giungere a risultati che permettano di comparare le istituzioni universitarie sulla

base della sintesi di un ampio campo di applicazioni alternative. È così che l'aggregazione geometrica diventa preferibile alla semplice aggregazione lineare che calcola le media degli elementi; la ponderazione assume, nel contesto della sintesi delle informazioni, un peso davvero significativo; tecniche semplici di normalizzazione – quali la trasformazione ordinale degli indicatori elementari in ranghi – sembrano funzionare alla stregua di standardizzazioni più complesse.

I dati che abbiamo utilizzato nelle nostre elaborazioni sono stati forniti e raccolti dal Censis: è quindi risultato fondamentale definire questi dati "di qualità", attribuendo loro – a posteriori, per ovvie ragioni – caratteristiche di accuratezza, validità e coerenza.

La ridotta numerosità dei dati a disposizione a cui spesso abbiamo dovuto far riferimento nelle nostre elaborazioni, ha limitato la sfera di applicazioni possibili escludendo, in particolare, l'opportunità di analizzare la struttura realmente sottostante ai dati attraverso metodi di analisi multivariata. In particolare un'analisi fattoriale sull'intero insieme di indicatori elementari potrebbe portare a definire famiglie diverse – in numero ed in significato – rispetto a quelle pre-determinate dal Censis; un'analisi cluster ex post può invece essere utile per verificare l'esistenza di aree geografiche o di tipologie di facoltà problematiche ovvero virtuose secondo gli aspetti analizzati. Anche le analisi preliminari basate sulla correlazione, se effettuate su un dataset più ampio, potrebbero evidenziare ridondanze tra gli indicatori appartenenti alla stessa famiglia o sovrapposizioni tra famiglie.

Entrando poi nel merito delle dimensioni del fenomeno analizzate dal Censis, riteniamo doverose alcune considerazioni. Anzitutto, dopo i primi due anni di stesura del rapporto, è stata eliminata la famiglia che valuta la capacità di attrazione delle facoltà: riteniamo questa sia una lacuna dell'impianto valutativo del Censis per la notevole importanza che tale dimensione assume nella realtà, conferendo maggior prestigio ad atenei e facoltà più attrattivi. Sono assenti misure sul destino dei laureati nella famiglia della produttività: questo dipende indubbiamente dall'assenza di un'unica base dati nazionale certificata per cui non sono disponibili informazioni omogenee sul destino professionale di coloro in possesso di un titolo di studio maggiore o uguale alla laurea¹⁴⁹. Non è prevista alcuna misura del grado di apprendimento: a parità di tasso di laureati in corso, ad esempio, una

¹⁴⁹ Bittante G. (2004), Interventi in occasione dell'incontro su "Indicatori di Ateneo e qualità dell'offerta formativa" tenutosi presso l'Università degli Studi di Padova il 22 aprile 2004.

facoltà può conferire titoli a studenti meno preparati rispetto ad un'altra; inoltre possono essere incentivati i comportamenti indulgenti delle facoltà che mirano a laureare un elevato numero di studenti in corso, senza adeguatamente appurare il livello di preparazione. Non è considerata nella struttura della valutazione la dotazione di strutture adeguate delle facoltà: si tratta indubbiamente di un aspetto della didattica per cui non c'è né ampia né diffusa disponibilità di dati per cui quindi va messa in conto una notevole difficoltà di reperimento. Per altri indicatori della didattica, non sempre valori elevati corrispondono ad organizzazioni migliori della facoltà: così un numero di classi di laurea alto in una facoltà può mascherare la necessità di coprire tutte le aree disciplinari anziché evidenziare un'ampia offerta didattica reale. Mancano misure relative alla produzione scientifica, che entrerebbero nella famiglia della ricerca conferendo indiscutibili plus alle facoltà più attive in questo senso. L'indicatore che valuta in maniera positiva le facoltà con il minor invecchiamento del corpo docente non tiene conto di eventuali elevati tassi di turn-over. I rapporti internazionali sono misurati sostanzialmente da indicatori relativi al progetto Erasmus omettendo valutazioni, ad esempio, sullo scambio docenti, sui consorzi tra atenei, sui dottorati internazionali.

Gli indicatori elementari considerati dal Censis sono poi, nella quasi totalità dei casi, indicatori di input e di processo; è difficile assumere che gli indicatori di risultato presenti siano adeguati perché sono numericamente pochi e discrezionali. Sarebbe perciò necessario accompagnare le misure presenti con altre che valutano in maniera più approfondita l'output del sistema universitario: si tratta di indicatori di più difficile costruzione ma che permettono di tener conto degli effetti del risultato conseguito, anche in termini di benefici derivanti all'individuo dal possesso di un livello d'istruzione maggiore¹⁵⁰.

Considerando la struttura complessiva del sistema di valutazione del Censis, un'importante osservazione è quella che denota la mancanza di una misura della variazione nella posizione in graduatoria e dei mutamenti nei punteggi delle famiglie che, di anno in anno, subisce ogni singola facoltà. Naturalmente una qualsiasi valutazione in tal senso deve comunque considerare i cambiamenti che intervengono nella struttura di

¹⁵⁰ Cfr. Trivellato U. (1975), *Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine sociale sulla riuscita scolastica*, Fondazione "Emanuela Zancan", Padova, pag. 32-33.

indicatori e famiglie ogni anno, cambiamenti che non rendono le graduatorie annuali comparabili in assoluto.

Secondo alcuni studiosi¹⁵¹, inoltre, la classifica degli atenei è contraddetta dagli esiti della classifica per facoltà per cui vi è scarsa coerenza interna: è però vero, anche se non giustifica l'incoerenza, che la graduatoria degli atenei è costruita in maniera del tutto diversa da quella delle facoltà per cui, anche in questo caso, confronto e sintesi sono problematici e non attuabili senza opportune accortezze.

Tornando alla valutazione, ci pare utile riportare la definizione che ne dà Patton¹⁵²: "La pratica della valutazione riguarda la raccolta sistematica di informazioni circa le attività, le caratteristiche e i risultati di programmi, personale e strutture destinata a persone interessate a giudicare come specifici aspetti di quei programmi, di quel personale, di quelle strutture abbiano operato e abbiano influito sul sistema. Tale definizione enfatizza: (1) la raccolta sistematica di informazioni (2) con riguardo ad un ampio spettro di oggetti (3) da usare da parte di specifici soggetti. Si usa pertanto il termine valutazione come un termine generale che include valutazione di programmi, valutazione del personale, valutazione di politiche, valutazione di strutture e altri processi valutativi." In questa chiave, le qualità che devono essere possedute dalla valutazione, e quindi, nel nostro caso, dal rapporto che il Censis elabora sulle istituzioni universitarie, possono essere così sintetizzate:

- centrata sul principio di utilità;
- consapevole delle condizioni di fattibilità;
- orientata alla pertinenza e alla appropriatezza rispetto alle finalità;
- basata sull'accuratezza metodologica.

La valutazione assume perciò chiari connotati di valore politico, di mediazione tra istanze scientifiche ed istanze operative, capace di coniugare mondo della ricerca e mondo delle risposte alle esigenze della società e delle sue parti.¹⁵³ A questo deve quindi rispondere anche l'impianto di valutazione del Censis sull'università italiana: fornire informazioni utili, pertinenti ed accurate, basate sulla raccolta sistematica di dati da fonti certificate, assicurando procedure metodologiche solide.

¹⁵¹ Checchi D., Biolcati Rinaldi F., Guglielmetti C, Turri M. (2007), *Classifiche delle università (University ranking)* Report intermedio presentato presso la Facoltà di Scienze Politiche dell'Università degli Studi di Milano il 30 ottobre 2007.

¹⁵² Patton M. Q. (1981), *Creative Evaluation*, Sage Publications, Beverly Hills.

¹⁵³ Bernardi L. (1997), *La valutazione: dai modelli alle applicazioni possibili*, in Corsi M., Fabbris L., Franci A. (a cura di), *La valutazione della qualità dei servizi socio-assistenziali*, Cleup, Padova.

Ma nel campo delle ricerca sociale, una grossa questione è quella che mette in relazione osservatore e osservato in quell'antagonismo che divide il quantitativo dal qualitativo, punto cruciale nella valutazione di fenomeni sociali che, per natura, sono non direttamente misurabili. Da un lato le scienze sociali sono state costrette, fin dalle origini, a fare i conti con la questione dell'osservatore dal momento che lo stesso "oggetto" a cui si applicano rimanda – come uno specchio un po' dispettoso – l'immagine del soggetto o, meglio, della comunità dei soggetti che sta dietro l'azione conoscitiva. Lo scienziato sociale, colui che investiga fenomeni relativi alla società, non può che rivedere sé stesso – con limiti, contraddizioni ed autoinganni che lo caratterizzano – nel referente delle sue analisi. Dall'altro lato i sociologi hanno spesso preferito fingere di non rivedersi nell'oggetto dei loro studi, preferendo considerarli alla stregua delle cose di cui si occupano le scienze naturali.¹⁵⁴ Anche il Censis, quindi, in quanto ricercatore sociale, diventa quell'osservatore che ha il compito di stimolare l'attenzione sull'oggetto della sua indagine, cercando di non riflettere sé stesso – né i committenti – nelle scelte operative compiute nell'esercizio valutativo.

Ci siamo chiesti, come da titolo di questa tesi, se il Censis è un arbitro che opera con rigore. Ci sono una serie di scelte autonome compiute dal Censis che conducono a risultati abbastanza simili a molte delle strategie alternative da noi utilizzate; da questo punto di vista ci pare di poter affermare che gli aspetti tecnico-metodologici di costruzione degli indicatori composti non sembrano determinare grandi cambiamenti negli esiti, per cui due sembrano i principali fattori discriminanti: la natura, l'articolazione e la qualità dei dati di base assunti per rappresentare le sottodimensioni del fenomeno considerato e la strategia di loro ponderazione all'interno della sottodimensione (o famiglia come chiamata dal Censis) e tra famiglie. Il Censis fornisce – nella pubblicazione e nel sito¹⁵⁵ – la possibilità di esaminare e valutare il procedimento complessivo che adotta¹⁵⁶; possiamo denominare questo atteggiamento "ricerca di trasparenza ex-post": riteniamo che uno sforzo congiunto dei soggetti coinvolti e interessati, per stabilire oggetti, rispettiva rilevanza, regole per loro unione, potrebbe rappresentare un viatico, laborioso ma necessario, per un affinamento del processo che lasci meno strascichi, meno insofferenza o meno indifferenza dopo la sua

¹⁵⁴ Bernardi L. (2005), *Percorsi di ricerca sociale – Conoscere, decidere, valutare*, Carocci editore, Roma.

¹⁵⁵ Il sito a cui ci riferiamo è il sito web del quotidiano *La Repubblica* dal quale abbiamo reperito le note metodologiche de "La Grande Guida all'Università" dei diversi anni.

¹⁵⁶ Dobbiamo peraltro ammettere qualche difficoltà nell'applicarlo ai dati che ci sono stati forniti che hanno sollevato qualche sospetto sulla congruenza tra dichiarato e adottato.

presentazione. Ovviamente nella piena consapevolezza che un percorso partecipativo è meno efficiente di un meccanico affidamento di incarico ad esterni ma anche che la delicatezza del tema e delle ricadute dei suoi esiti sul sistema sembra meritare un momento di condivisione più alto per poter divenire più credibile. D'altronde tutta la letteratura sulla valutazione dedica ampio spazio alla considerazione della qualità dell'interazione tra attori, con varie responsabilità, nel percorso di sua assunzione, in particolare in contesti complessi, dinamici e "turbolenti", sottolineando la necessità di una sua corretta gestione affinché, su processi ed esiti, vi sia corresponsabilità e quindi maggior rispetto e, in prospettiva, maggiore utilizzazione delle indicazioni che emergono.

BIBLIOGRAFIA

Aiello F., Attanasio M. (2004), How to transform a batch of simple indicators to make up a unique one?, Atti del Convegno SIS giugno 2004, Bari – Sessioni Specializzate.

Aiello F., Attanasio M. (2006), Some Issues in Constructing Composite Indicators, VIII International Meeting on Quantitative Methods for Applied Sciences, Certosa di Pontignano, 11-13 settembre 2006.

Aiello F., Librizzi L. (2006), Gli indicatori nelle scienze sociali: dal qualitativo al quantitativo, in Diamond I., Jefferies J., (2006), Introduzione alla statistica per le scienze sociali, Mc Graw-Hill, Milano.

Allegra F. S., La Rocca A. (2004), Sintetizzare misure elementari: una sperimentazione di alcuni criteri per la definizione di un indice composto, Istat, Roma.

Attanasio M., Capursi V. (1997). Graduatorie sulla qualità della vita: prime analisi di sensibilità delle tecniche adottate. Atti XXXV Riunione Scientifica SIEDS, Alghero.

Bartholomew D. J. (1996) The statistical approach to social measurement, Academic Press, San Diego.

Bartholomew D. J. (2002) Discussion on the paper by Fayers P. M., Hand D. J..

Bernardi L. (1997), La valutazione: dai modelli alle applicazioni possibili, in Corsi M., Fabbris L., Franci A. (a cura di), La valutazione della qualità dei servizi socio-assistenziali, Cleup, Padova.

Bernardi L. (2005), Percorsi di ricerca sociale – Conoscere, decidere, valutare, Carocci editore, Roma.

Bernardi L., Capursi V., Librizzi L. (2004). Measurement awareness: the use of indicators between expectations and opportunities, SIS: Sezione Specializzata, Atti della XLIII Riunione Scientifica, Bari.

Bernardi L., Torelli N. (1990), Le potenzialità offerte alle statistiche sociali ufficiali dalla definizione di modelli e indicatori, ISTAT, Avanzamenti metodologici e statistiche ufficiali, Atti delle prime giornate di studio, Roma, 13-14 dicembre.

Bertin G., (1989), Decidere nel pubblico – tecniche di decisione e valutazione nella gestione dei servizi pubblici, Gruppo Editoriale Fabbri, Bompiani, Sonzogno, Etas libri.

Bertin G. (a cura di) (1995), Valutazione e sapere sociologico: metodi e tecniche di gestione dei processi decisionali, FrancoAngeli, Milano.

Bittante G. e Masarotto G. (2004), Interventi in occasione dell'incontro su "Indicatori di Ateneo e qualità dell'offerta formativa" tenutosi presso l'Università degli Studi di Padova il 22 aprile 2004.

Boschetti M., Brivio P. A., Brogiolo G. P., Crosato A., Soffia E. (2003), Applicazioni GIS per l'archeologia: lo studio della distribuzione dei siti archeologici nel Garda meridionale, Edizioni all'Insegna del Giglio.

Blumer H., (1939) An appraisal of Thomas and Znaniecki's "The Polish Peasant in Europe and America", New York, Social Science Research Council, Bulletin 44.

Cecchi D., Biolcati Rinaldi F., Guglielmetti C, Turri M. (2007), Classifiche delle università (University ranking) – Report intermedio presentato presso la Facoltà di Scienze Politiche dell'Università degli Studi di Milano il 30 ottobre 2007.

Corbetta P. G., (1999), Metodologia e tecniche della ricerca sociale, Il Mulino, Bologna.

Curatolo R. (1972), *Indicatori sociali*. Atti della XXVII Riunione Scientifica della SIS (Palermo, 29-31 maggio), vol. 1.

Delvecchio F. (1995), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari.

D'Ovidio F. (2002), *Un indicatore di efficienza del processo didattico universitario*, in Carli Sardi L., Delvecchio F., *Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari*, 2002, Cleup, Padova.

Ebert U. and Welsch H., (2004), *Meaningful environmental indices: a social choice approach*, *Journal of Environmental Economics and Management*, vol. 47.

Fabbris L., Capiluppi C., Giancotti G., Meneghello A. (1993), *Statree 1.0 – Sistema esperto per la scelta del metodo di analisi statistica*, Manuale per l'uso, Edizioni Summa, Padova.

Fabbris L. (1997), *Statistica multivariata – Analisi esplorativa dei dati*, McGraw-Hill, Milano.

Forman E.H. (1983) *The Analytic Hierarchy Process as a decision support system*, *Proceedings of the IEEE Computer Society*.

Freudenberg M. (2003), *Composite indicators of country performance: a critical assessment*, OECD, Paris.

Kemeny J. G., Snell J. L. (1960), *Finite Markov Chain*, Princeton.

Karlsson J. (1998), *A systematic approach for prioritizing software requirements*, Ph.D. dissertation No. 526, Linköping, Sverige.

Land K. C. (1971), *On the Definition of Social Indicators*, *The American Sociologist*, n. 6.

Munda G., Nardo M. (2003), On the methodological foundations of composite indicators used for ranking countries, OECD/JRC Workshop on Composite Indicators of Country Performance, Ispra, Italy, May 12.

Munda G., Nardo M. (2005), Constructing consistent composite indicators: the issue of weights, Institute for the Protection and Security of the Citizen.

Nardo M. et al. (2005), Handbook on constructing composite indicators: methodology and user guide, OECD Statistics Working Papers, 2005/3, OECD Publishing. doi:10.1787/533411815016.

Nardo M., Saisana M., Saltelli A., Tarantola S. (2005), Tools for composite indicators building, Report EUR 21682 EN. European Commission-Joint Research Centre, Ispra (Italy).

Nilsson R. (2000) .Calculation of Composite leading Indicators: A comparison of two different methods., Paper presented at the CIRET conference, Paris, October 2000.

Nota metodologica 2000 de "La Grande Guida all'Università".

http://www.repubblica.it/misc/guida/metodologia_01.html

Nota metodologica 2001 de "La Grande Guida all'Università".

<http://www.repubblica.it/speciale/universita/metodologia.html>

Nota metodologica 2002 de "La Grande Guida all'Università".

http://www.polimi.it/ateneo/nucleo_di_valutazione/file/relazione2002/allegato11.pdf

Nota metodologica 2003 de "La Grande Guida all'Università".

<http://www.repubblica.it/speciale/2003/universita/misura.html>

Nota metodologica 2004 de "La Grande Guida all'Università".

http://download.repubblica.it/pdf/metodologia_facolta_2004.pdf

Nota metodologica 2005 de "La Grande Guida all'Università".

http://www.repubblica.it/speciale/2005/guida_universita/metodologie.html

Nota metodologica 2006 de "La Grande Guida all'Università".

<http://download.repubblica.it/pdf/scuola/facolta.pdf>

Nota metodologica 2007 de "La Grande Guida all'Università".

http://download.repubblica.it/pdf/2007/CensisMetodologia_2007.pdf

Patton M. Q. (1981), *Creative Evaluation*, Sage Publications, Beverly Hills.

Pellini P. (2006), *La riforma Moratti non esiste*, Il Saggiatore, Milano.

Perotti R. (2002), *The italian university system: rules vs incentives*, in "Annual Report on Monitoring Italy 2002", Istituto di Studi e Analisi Economica (ISAE), Roma.

Porcu M., Puggioni G. (2002), *Due classi di indicatori di inefficienza dei percorsi universitari*, in Carli Sardi L., Del vecchio F., *Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari*, 2002, Cleup, Padova.

Rizzi A. (1988), *Un metodo di graduazione di più unità statistiche*, *Rivista di Statistica Applicata*, vol. 21, n. 1, marzo.

Saaty T. L. (1980) *The Analytic Hierarchy Process*, McGraw-Hill Book Co., N.Y.

Saisana M., Tarantola S. (2002), *State-of-the-art report on current methodologies and practices for composite indicator development*, Report EUR 20408 EN, European Commission-Joint Research Centre, Ispra (Italy).

Sharpe A. (2004), *Literature Review of Frameworks for Macro-indicators*, Centre for the Study of Living Standards, Ottawa, CAN.

Shaw C. D. (1986), *Quality Assurance – What the Colleges Are Doing*, King's Fund Centre, London.

Stone R., (1970), *An Integrated System of Demographic, Manpower and Social Statistics and its Links with the System of National Economic Accounts*, in "Memorandum E/CN", United Nations Economic and Social Council UNESCO.

Storrie D., Bjurek H. (2000), *Benchmarking European labour market performance with efficiency frontier techniques*. Discussion paper FS I 00-211.

Trivellato U. (1975), *Scuola e stratificazione sociale – Problemi di valutazione e di misura degli effetti dell'origine sociale sulla riuscita scolastica*, Fondazione "Emanuela Zancan", Padova.

U.S. Department of Health, Education and Welfare, *Toward a Social Report*, 1969.

Vianelli S. (1989), *Statistica e Scienze della complessità*, in *Atti della XXXI Riunione scientifica della Società Italiana di Economia Demografia e Statistica*, Palermo, Vol. XLIII, nn. 1-2.

Zajczyk F. (1997), *Il mondo degli indicatori sociali – Una guida alla ricerca sulla qualità della vita*, La Nuova Italia Scientifica, Roma.

SITOLOGIA

www.censis.it

www.almalaurea.it

www.unica.it (Ricerca scientifica)

www.campusone.it

www.istitutoveneto.it/venezia/dati/metadati/metadati1.htm

www.itl.nist.gov/div898/handbook/eda/eda.htm

<http://composite-indicators.jrc.ec.europa.eu>

APPENDICE

Il contenuto di questa appendice ha lo scopo di illustrare il processo che ha ci ha guidati dall'esame del materiale inviatoci dal Censis alla determinazione degli indicatori elementari previsti per la guida del 2006.

Dopo un semplice elenco delle variabili presenti nel dataset inviato dal Censis, integrato dall'interpretazione che noi abbiamo attribuito alle stesse, riportiamo per ciascuna famiglia una tabella che indica, per ciascun indicatore elementare, quale variabili sono state utilizzate per determinarne il valore.¹⁵⁷

¹⁵⁷ I valori degli indicatori elementari, così come calcolati a partire dalle variabili inviateci dal Censi, sono a disposizione presso l'autore.

ELENCO VARIABILI

Tabella 1 Elenco delle variabili del dataset inviato dal Censis con relativa interpretazione.

VARIABILE	INTERPRETAZIONE
FAC_ATE	facoltà
ISCR AA04-05 (ANNO IMM AA03-04)	iscritti A.A. 04-05 immatricolati nell'A.A. 03-04
IMM_AA03-04	immatricolati nell'A.A. 03-04
CREDITI_L_LSCU_AS04	crediti acquisiti nell'A.S. 2004 dagli studenti iscritti alle lauree triennali e a ciclo unico
ISCR_AA03-04_L_LSCU	studenti iscritti alle lauree triennali e a ciclo unico nell'A.A. 03-04
ISCR AA04-05	iscritti all'A.A. 04-05
ISCR FUORI CORSO AA04-05	iscritti fuori corso all'A.A. 04-05
IMM 04-05	immatricolati nell'A.A. 04-05
LAU L AS04 (ANNO IMM AA_01-02)	laureati nell'A.S. 2004 nei corsi di laurea triennale immatricolati nell'A.A. 01-02
IMM L AA01-02	immatricolati ai corsi di laurea triennale nell'A.A. 01-02
TASSO LAU IN CORSO AS03-AS04 (VECCHIO ORDINAMENTO)	tasso dei laureati in corso
CLASSI DI LAUREA AA04-05	numero di classi di laurea presenti nell'offerta di corsi dell'A.A. 04-05
INSEGNAMENTI ATTIVATI AA03-04	numero di insegnamenti presenti nell'offerta di corsi nell'A.A. 03-04
INSEGNAMENTI TENUTI DA DOCENTI NON DI RUOLO EXTRA ACCADEMICI	insegnamenti tenuti da docenti non di ruolo extra accademici
POSTI AULA NUCLEI USO ESCLUSIVO 2004	numero di posti aula NUCLEI 2004
POSTI AULA NUCLEI 2005	numero di posti aula NUCLEI 2005
% DEI CORSI CON ATTIVITA' DI MONITORAGGIO E VALUTAZIONE	percentuale dei corsi con attività di monitoraggio e valutazione
TOT max_stages_N2005	totale attività di stages AS 2005
TOT_ISCR_AA03-04	iscritti AA 03-04
B FIN 2003	numero di unità di ricerca finanziate nel 2003
B FIN 2004	numero di unità di ricerca finanziate nel 2004
B FIN 2005	numero di unità di ricerca finanziate nel 2005
ASSEGNATO MURST 2003 (Euro)	finanziamenti ottenuti dal Murst nel 2003
ASSEGNATO MURST 2004 (Euro)	finanziamenti ottenuti dal Murst nel 2004
ASSEGNATO MURST 2005 (Euro)	finanziamenti ottenuti dal Murst nel 2005
TOT DOCENTI 2002	totale docenti di ruolo al 31.12.2002
TOT DOCENTI 2003	totale docenti di ruolo al 31.12.2003
TOT DOCENTI 2004	totale docenti di ruolo al 31.12.2004
FP5+FP6+TEMPUS_2005	progetti finanziati dai Programmi Quadro V e VI + Programma Tempus
FIRB_u_finanziate_2003	numero di unità di ricerca finanziate FIRB nel 2003
FIRB_finanziamento_2003	finanziamenti FIRB 2003

VARIABILE	INTERPRETAZIONE
FIRB_u_finanziate_2005	numero di unità di ricerca finanziate FIRB nel 2003
FIRB_finanziamento_2005	finanziamenti FIRB 2005
ETA MEDIA 2005	età media dei docenti nel 2005
ETA MEDIA AL 2001	età media dei docenti nel 2001
ERASMUS_UFFICI_AA04-05UFFICI	numero di studenti con borse Erasmus (in uscita) A.A. 04-05
CERVELLI ril 2004-05-06	numero di docenti che hanno beneficiato di finanziamenti per il c.d. "rientro dei cervelli" nel triennio 2004-2006
OPPORTUNITA' INTERNAZIONALI	numero di contributi ottenuti dalla facoltà per attività di cooperazione internazionale 2003-2006
ERASMUS_IN_2003_2004	studenti che hanno ottenuto una borsa Erasmus nell'A.A. 03-04
ERASMUS_IN_2004_2005	studenti con borse Erasmus in uscita nell'A.A. 04-05
UNIVOSP_AA04-05INDIRE	numero di università straniere che hanno ospitato gli studenti Erasmus nell'A.A. 04-05

NOTE

- Il totale docenti dell'anno XXXX abbiamo stabilito coincidere con il numero complessivo di docenti di ruolo al 31.12.XXXX;
- Per la variabile relativa agli insegnamenti tenuti da docenti non di ruolo extra accademici non è presente alcun riferimento temporale: supponiamo esso coincida con l'A.A. a cui l'intera guida del 2006 si riferisce (ovvero all'A.A. 2004-2005).
- Per la variabile relativa alla percentuale dei corsi con attività di monitoraggio e valutazione non è presente alcun riferimento temporale: supponiamo esso coincida con l'A.A. a cui l'intera guida del 2006 si riferisce (ovvero all'A.A. 2004-2005).
- Finanziamenti e numero di unità di ricerca finanziate riguardano COFIN e FIRB per gli anni 2003 e 2004 mentre per il 2004 riguardano solamente il programma FIRB.
- Non c'è una chiara distinzione tra i valori relativi al numero di studenti con borse Erasmus in entrata e in uscita: la nostra interpretazione si è basata sull'indicazione temporale che accompagna il nome delle variabili.
- Non sono chiari i riferimenti agli anni accademici ed agli anni solari.
- È faticoso interpretare a quali i corsi di laurea (triennali, specialistici, a ciclo unico, vecchio ordinamento) si riferiscano i dati del numero di iscritti o laureati: l'interpretazione ha pertanto seguito logiche di esclusione (ovvero le variabili necessarie al calcolo di un determinato indicatore sono state selezionate escludendo quelle che abbiamo ritenuto non pertinenti).
- Le variabili che si riferiscono a Corsi di Laurea del vecchio ordinamento o a ciclo unico misurano unicamente il numero di studenti o laureati iscritti al vecchio ordinamento dal momento che la Facoltà di Scienze Politiche non prevede Corsi di Laurea a ciclo unico.

FAMIGLIA "PRODUTTIVITÀ"

Tabella 2 Variabili utilizzate per la determinazione degli indicatori elementari della famiglia "produttività".

		NUMERATORE	DENOMINATORE
P1	TASSO DI PERSISTENZA TRA IL I ED IL II ANNO iscritti all'A.A. 04-05 che si sono immatricolati nell'anno accademico precedente / immatricolati nell'A.A. 03-04	ISCR AA04-05 (ANNO IMM AA03-04)	IMM_AA03-04
P2	INDICE DI REGOLARITÀ DEGLI STUDENTI ISCRITTI ALLE TRIENNALI crediti acquisiti nell'A.S. 2004 dagli studenti iscritti alle lauree triennali e a ciclo unico / studenti iscritti alle lauree triennali e a ciclo unico nell'A.A. 03-04 * 60	CREDITI_L_LSCU_AS04	ISCR_AA03-04_L_LSCU
P3	TASSO DI ISCRITTI IN CORSO (iscritti in totale - immatricolati - iscritti fuori corso)/(iscritti in totale - immatricolati)	(ISCR AA04-05) - (IMM 04-05) - (ISCR FUORI CORSO AA04-05)	(ISCR AA04-05) - (IMM 04-05)
P4	TASSO DI LAUREATI TRIENNALI laureati nell'A.S. 2004 nei corsi di laurea triennale / immatricolati ai corsi di laurea triennale nell'A.A. 01-02	LAU L AS04 (ANNO IMM AA_01-02)	IMM L AA01-02
P5	TASSO DI LAUREATI IN CORSO laureati in corso nell'A.S. 2004 nei corsi di laurea a ciclo unico e del vecchio ordinamento / totale dei laureati nei corsi citati	TASSO LAU IN CORSO AS03-AS04 (VECCHIO ORDINAMENTO)	

NOTE

- Il valore dell'indicatore P5 è stato fornito direttamente come tasso: non abbiamo quindi a disposizione i valori di numeratore e denominatore.

FAMIGLIA "DIDATTICA"

Tabella 3 Variabili utilizzate per la determinazione degli indicatori elementari della famiglia "didattica".

		NUMERATORE	DENOMINATORE
D1	N. DI CLASSI DI LAUREA PRESENTI NELL'OFFERTA DI CORSI A.A. 04-05	CLASSI DI LAUREA AA04-05	
D2	N. DI INSEGNAMENTI PRESENTI NELL'OFFERTA DI CORSI A.A. 03-04	INSEGNAMENTI ATTIVATI AA03-04	
D3	DOCENTI DI RUOLO / INSEGNAMENTI (A.S. 2004 - A.A. 04-05)	TOT DOCENTI 2004	INSEGNAMENTI ATTIVATI AA03-04
D4	DOCENTI DI RUOLO / ISCRITTI docenti di ruolo al 31.12.2004 / iscritti nell'A.A. 04-05	TOT DOCENTI 2004	ISCR AA 04-05
D5	POSTI AULA ISCRITTO n. di posti aula NUCLEI 2004 / iscritti nell'A.A. 03-04	POSTI AULA NUCLEI USO ESCLUSIVO 2004	TOT_ISCR_AA03-04
D6	POSTI AULA ISCRITTO n. di posti aula NUCLEI 2005 / iscritti nell'A.A. 04-05	POSTI AULA NUCLEI 2005	ISCR AA 04-05
D7	STAGES PER STUDENTE A.A. 03-04	TOT max_stages_N2005	TOT_ISCR_AA03-04
D8	ATTIVITA' DI MONITORAGGIO E DI VALUTAZIONE DEI CORSI A.A. 03-04	% DEI CORSI CON ATTIVITA' DI MONITORAGGIO E VALUTAZIONE	

NOTE:

- Per l'indicatore D3 abbiamo dovuto utilizzare il valore degli insegnamenti attivati nell'A.A. 2003-2004 anziché il numero di insegnamenti attivati nell'A.A. successivo in quanto quest'ultima informazione non ci è stata fornita.
- Nonostante non fosse chiaramente esplicitata la sua natura di rapporto – come invece accade nei casi in cui l'indicatore prevede in maniera esplicita un numeratore ed un denominatore –, per determinare il valore di D7 abbiamo rapportato il numero di stages al totale degli iscritti dell'A.A. 2003-2004.
- Il numero totale di stages pare fare riferimento all'anno 2005: dal momento che l'indicatore prevede il riferimento all'A.A. 2003-2004 – e che la variabile TOT max_stages_N2005 è l'unica che fa riferimento agli stages – abbiamo considerato questa informazione per la costruzione dell'indicatore D7.

FAMIGLIA "RICERCA"

Tabella 4 Variabili utilizzate per la determinazione degli indicatori elementari della famiglia "ricerca".

		NUMERATORE	DENOMINATORE
R1	N. DI UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2003 PER DOCENTE n. di unità di ricerca nazionali o locali finanziate / docenti di ruolo al 31.12.2002	B FIN 2003 + FIRB_u_finanziate _2003	TOT DOCENTI 2002
R2	N. DI UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN NEL 2004 PER DOCENTE n. di unità di ricerca nazionali o locali finanziate / docenti di ruolo al 31.12.2003	B FIN 2004	TOT DOCENTI 2003
R3	N. DI UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2005 PER DOCENTE n. di unità di ricerca nazionali o locali finanziate / docenti di ruolo al 31.12.2004	B FIN 2005 + FIRB_u_finanziate _2005	TOT DOCENTI 2004
R4	FINANZIAMENTO MEDIO OTTENUTO DALLE UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2003 totale dei finanziamenti ottenuti dalle unità di ricerca nazionali o locali della facoltà / n. delle unità finanziate	ASSEGNATO MURST 2003 (Euro) + FIRB_finanziamento _2003	B FIN 2003 + FIRB_u_finanziate _2003
R5	FINANZIAMENTO MEDIO OTTENUTO DALLE UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN NEL 2004 totale dei finanziamenti ottenuti dalle unità di ricerca nazionali o locali della facoltà / n. delle unità finanziate	ASSEGNATO MURST 2004 (Euro)	B_FIN 2004
R6	FINANZIAMENTO MEDIO OTTENUTO DALLE UNITA' DI RICERCA FINANZIATE DAL PROGRAMMA COFIN E FIRB NEL 2005 totale dei finanziamenti ottenuti dalle unità di ricerca nazionali o locali della facoltà / n. delle unità finanziate	ASSEGNATO MURST 2005 (Euro) + FIRB_finanziamento _2005	B FIN 2005 + FIRB_u_finanziate _2005
R7	PROGETTI DI RICERCA FINANZIATI DAI PROGRAMMI COMUNITARI (V e VI Programma Quadro; Programma Tempus)	FP5+FP6+TEMPUS _2005	

FAMIGLIA "PROFILO DOCENTI"

Tabella 5 Variabili utilizzate per la determinazione degli indicatori elementari della famiglia "profilo docenti".

		NUMERATORE	DENOMINATORE
PD1	ETÀ MEDIA DEI DOCENTI AL 2005 età media dei docenti di ruolo al 2005	ETA MEDIA 2005	
PD2	INVECCHIAMENTO età media dei docenti di ruolo al 2005 – età media dei docenti di ruolo al 2001	(ETA MEDIA 2005) - (ETA MEDIA AL 2001)	
PD3	ERASMUS IN USCITA PER DOCENTE n. di studenti con borse Erasmus nell'A.A. 04-05 / docenti di ruolo al 31.12.2004	ERASMUS UFFICI_AA04- 05UFFICI	TOT DOCENTI 2004
PD4	TASSO DI INSEGNAMENTI TENUTI DA DOCENTI NON DI RUOLO EXTRA ACCADEMICI insegnamenti tenuti da docenti non di ruolo "extra accademici" / totale degli insegnamenti attivati nell'A.A. 03-04	INSEGNAMENTI TENUTI DA DOCENTI NON DI RUOLO EXTRA ACCADEMICI	INSEGNAMENTI ATTIVATI AA03-04
PD5	DOCENTI CHE HANNO OTTENUTO FINANZIAMENTI PER IL C.D. "RIENTRO DEI CERVELLI" n. di docenti che hanno beneficiato di finanziamenti per il c.d. "rientro dei cervelli" nel triennio 2004-2006	CERVELLI ril 2004- 05-06	

FAMIGLIA "RAPPORTI INTERNAZIONALI"

Tabella 6 Variabili utilizzate per la determinazione degli indicatori elementari della famiglia "rapporti internazionali".

		NUMERATORE	DENOMINATORE
RI1	ERASMUS IN USCITA PER STUDENTE studenti con borse Erasmus in uscita nell'A.A. 04-05 / iscritti al netto degli immatricolati A.A. 04-05	ERASMUS_IN_2004 _2005	(ISCR AA04-05) - (IMM 04-05)
RI2	ERASMUS IN ENTRATA PER STUDENTE media degli studenti stranieri che hanno ottenuto una borsa Erasmus nella facoltà negli anni accademici 2003-04 e 2004-05 / iscritti A.A. 04-05	ERASMUS_IN_2003 _2004	ISCR AA04-05
RI3	UNIVERSITÀ OSPITANTI PER DOCENTE numero di università straniere che hanno ospitato gli studenti Erasmus nell'A.A. 04-05 / docenti di ruolo al 31.12.2004	UNIVOSP_AA04- 05INDIRE	TOT DOCENTI 2004
RI4	OPPORTUNITÀ INTERNAZIONALI numero di contributi ottenuti dalla facoltà per attività di cooperazione internazionale 2003-2006	OPPORTUNITÀ INTERNAZIONALI	

NOTE:

- Il calcolo dell'indicatore RI2 prevede a numeratore la media tra il numero di studenti Erasmus in entrata dell'A.A. 2003-2004 e quelli dell'A.A. successivo: avendo a disposizione solo il primo di questi due valori, è diventato questo il numeratore del rapporto.

Bene. Eccoci ai ringraziamenti (e quando si scrivono questi vuol dire che davvero sei giunto alla fine).

Anzitutto voglio ringraziare il professor Lorenzo Bernardi per l'avermi dato la possibilità di lavorare su un argomento come questo, che da subito ha suscitato la mia curiosità ed ha stimolato, procedendo via via nella stesura della tesi, l'interesse conoscitivo per la valutazione, per tutte le problematiche ad essa connesse e per la fatica di trovare la "soluzione ottima" ai problemi di misurazione. Lo ringrazio inoltre di cuore per la fiducia che ha dimostrato nei miei confronti, sostenendo questo mio lavoro ed offrendomi opportunità di crescita anche esterne all'università.

Ringrazio la mia famiglia per avermi sostenuto fino a qui con continui sproni e stimoli ad arrivare fino in fondo: grazie a papà perché non ha ostacolato le mie scelte ed ha invece cercato di capirle; grazie a mamma per la pazienza che ha sempre avuto con me nei momenti – troppo frequenti – in cui vien fuori la mia intemperanza; grazie ad Elisa perché, silenziosamente, mi ha sempre dato una mano, all'università ma anche fuori, facendomi sentire – cosa importante per me – un riferimento. Vi voglio bene.

Un grazie immenso a Daniele che da quasi due anni accompagna ogni mio passo con la sua dolcezza, la sua sensibilità, la sua razionalità, la sua pazienza, il suo saper guardare la vita con gli occhi degli altri: sei davvero speciale e sono davvero orgogliosa di poter "camminare" con te, al tuo fianco!

E grazie ad un'altra persona importante, don Claudio, che da lungo tempo mi aiuta a discernere ciò che buono e migliore per me, con gli occhi di chi mi conosce in profondità e sa leggere la mappa della mia vita suggerendomi le strade da percorrere.

Tornando all'ambito accademico, un grazie a tutti i docenti della Facoltà che mi hanno accolta ed ascoltata durante i Consigli di Facoltà e le sedute della Commissione NIM ma anche fuori da queste sedi. In particolare rivolgo un ringraziamento al professor Guido Masarotto per avermi offerto un'importante opportunità formativa in campo medico "agganciandomi" allo IOV (Istituto Oncologico Veneto). Inoltre ringrazio assieme a lui anche il professor Tommaso Di Fonzo perché, in qualità di Presidi della Facoltà, hanno sempre accolto con la massima disponibilità la mia "rappresentanza studentesca". E non posso certo dimenticare Delfina, Fiorenza e Maria Paola che mi hanno sempre fatta sentire a casa ogni volta che ho messo piede in Presidenza: siete state davvero un forte sostegno in questi anni! Indubbiamente se non avessi avuto la possibilità di essere rappresentante degli studenti in Facoltà e in Ateneo il mio interesse per quanto trattato in questa tesi sarebbe stato decisamente inferiore: un grazie quindi a tutti coloro, studenti, docenti,

personale della segreteria, mi hanno fatto "mettere le mani in pasta" nel mondo accademico.

Grazie a tutto il Servizio Sperimentazioni Cliniche e Biostatistica dello IOV dove ho trascorso quasi due anni: Paola, Denise, Ilaria, Laura, Elisa (anche se adesso hai cambiato lavoro, tu sei stata lì la mia collega!), per l'ampia sezione femminile, e l'unico uomo, Gian Luca che, da ottimo responsabile, mi ha insegnato l'importanza della delicatezza delle informazioni e della cura costante da dedicare agli studi clinici. È stata una grossa opportunità quella di lavorare con voi, per la mia crescita professionale e perché ho conosciuto persone davvero belle: mi mancate.

Da Padova a Vicenza: per Demos devo ringraziare Fabio per avermi dato e perché continua a darmi la possibilità di crescere professionalmente soprattutto nella responsabilità del mio lavoro di statistico; grazie a Natascia per i piccoli consigli, per quelle volte in cui mi hai tranquillizzata e perché nei momenti faticosi – come quello passato nella preparazione della tesi – l'ho sentita vicina. E grazie a Cristina che ogni giorno mi accoglie "in casa Demos-Step"! Un ringraziamento anche al professor Ilvo Diamanti perché mi ha offerto l'opportunità di lavorare sulla realtà sociale in cui quotidianamente siamo immersi.

E grazie agli amici universitari che hanno accompagnato, secondo passi ed andature diverse, la mia avventura universitaria: Marianna, Riccardo, Ilaria, Margherita, Nicola, Erica, Enrico, Lella, Alice del mitico SPS (ricordando che SPS = Siamo Proprio Simpatici!), Arianna e Claudio preziosi compagni di treno, e Cristiano a cui è bastato poco per "convertirsi" a statistica!

Grazie a tutti gli amici di Biadene e periferia (sì sì, Caonada e Pederiva sono periferia di Biadene!!!! eheh!), conosciuti nel corso degli anni con i quali ho condiviso – e spero dividerò ancora – davvero molto: Checca, Angela S., Anna e Roberto, Federica, Silvia, Valentina e tutti gli altri che qui non sono scritti nero su bianco ma che sono scritti nei miei ricordi e nel mio cuore. Grazie per aver donato note di colore ai giorni più grigi ma anche per aver aggiunto la vostra firma ai giorni più felici!

Un grazie a tutti gli amici di Azione Cattolica, della mia mitica Diocesi di Treviso ma anche gli amici sparsi in giro per l'Italia (Michela e Fabrizio in primis!): è davvero bello sentirsi a casa in una famiglia in cui ciascuno può trovare il suo posto, in cui camminare assieme in comunione e corresponsabilità, in cui ciascuno sostiene e dà occasione all'altro per poter dire un carico di speranza "duc in altum!".

Vi stimo tutti e vi porterò sempre tra i ricordi più belli!

Infine, il grazie più grande a Colui che ha reso tutto questo possibile, Colui che mi ha sostenuto attraverso le persone che mi ha messo vicino ogni giorno, Colui che mi manda

le prove dandomi assieme la forza necessaria per sopportarle, Colui che prima di tutti mi ama e sa come rendermi davvero felice.

Semplicemente, a tutti e a ciascuno... grazie.

Paola