

Istituto di Ricerche sulla Popolazione
e le Politiche Sociali del CNR



Giuseppe Gesano

***ricchezza e povertà dei dati demografici
negli studi di popolazione***

Roma, ottobre 2011



Giuseppe Gesano

***ricchezza e povertà dei dati demografici
negli studi di popolazione***

indice

I. In principio	7
II. Individui e collettività: classificazioni, profili, macrodati e l'ecological fallacy	13
III. Della disponibilità e della qualità dei dati demografici	19
IV. Cenni sugli effetti dell'aumento delle potenzialità informatiche	25
V. Le indagini campionarie, o l'illusione di poter indagare su "ogni cosa"	29
VI. I modelli: dall'astrazione matematica alla semplificazione della realtà, al rischio di metodologie standard	35
VII. Spazio e tempo: variabili di contesto o variabili causali?	39
VIII. L'ipercubo delle relazioni	43
IX. De profundis?	47
X. Prospettive di sopravvivenza	51
Riferimenti bibliografici	57

I.

In principio

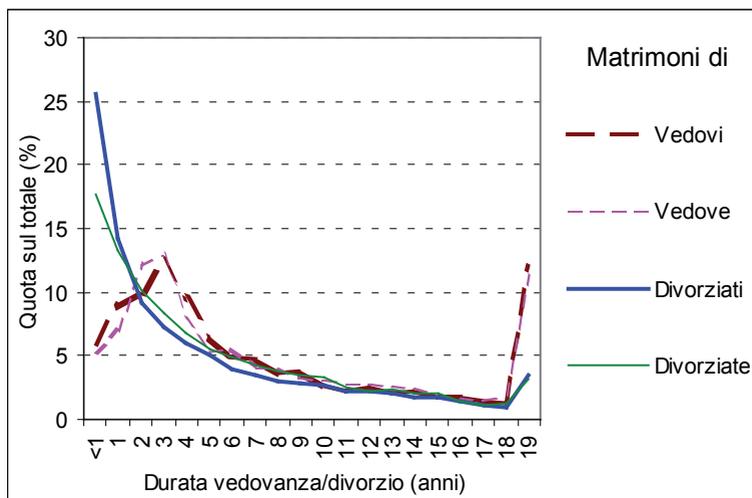
In principio era il dato, quello che contava il *numero* dei sudditi, li classificava per sesso, per fasce di età, per l'essere o non essere abili al lavoro o a portare le armi, per censo, per mestiere, per assolvimento del precetto pasquale; li raccoglieva in famiglie, fuochi, tribù; li collocava sul territorio in accampamenti, casolari, villaggi, contrade, borghi, cittadine e, nelle città, in quartieri, sestieri, rioni.¹ Insomma, singole *persone* con i propri *caratteri individuali*, dai quali ciascuna di esse risulta conformata nella sua inscindibile ma non imm modificabile *identità*, trasformate in *unità statistiche*, le quali di volta in volta fanno scattare il contatore dell'urna di competenza secondo la modalità di uno solo o solamente di alcuni di quei caratteri, così che ne risulti la *struttura della popolazione* rispetto a quell/i carattere/i rilevato/i.

Le statistiche di movimento erano (e ancora sono) altrettanto sommarie: gli *eventi*, una volta raccolti in *fenomeni*, vengono contati e classificati secondo alcuni *caratteri dell'evento* e/o dei soggetti che vi sono coinvolti. In questa operazione, oltre a ridurre la complessità intrinseca dei loro attori, si rischia di perdere la dinamicità dell'evento, il suo inserirsi in una sequenza di fatti (concreti o anche solo mentali; interiori o esterni) e di azioni (individuali o con più attori); si rischia, cioè, di perdere il *processo* attraverso il

1. «... farete il censimento della comunità d'Israele. Compilate un elenco con i nomi di tutti i maschi secondo il proprio gruppo e la propria famiglia. Conterete tutti gli uomini, uno per uno. Registerete i maschi, dall'età di vent'anni in su, in grado di combattere nell'esercito d'Israele. Preparerete l'elenco dei soldati, disposti in ordine, secondo le loro truppe. Vi aiuterà un uomo per ogni tribù: un capofamiglia. [...] Il totale degli Israeliti registrati nel censimento, a partire dall'età di vent'anni in su, in grado di combattere, era dunque di 603.550 uomini.» [Bibbia, Numeri 1: 2-4; 45].

quale si è pervenuti all'evento registrato, oppure di coglierne solo deboli tracce come, ad esempio, nelle tavole sui matrimoni dei vedovi (divorziati) per età, sesso, anno e durata della vedovanza (del divorzio).²

FIGURA 1 – MATRIMONI DI VEDOVÌ E DI DIVORZIATI PER DURATA DELLA VEDOVANZA O DEL DIVORZIO, PER GENERE: ITALIA, 2003 (VALORI PERCENTUALI SUL TOTALE DEI MATRIMONI PER GENERE E STATO CIVILE)



Sta qui, forse, il peccato originale della demografia tradizionale e, più alla larga, della statistica sociale: quello di operare una sorta di rovesciamento della realtà. Perché ciascuno di noi, individualmente preso, è nella sua interezza e *agisce* nella propria storia, mentre nelle statistiche demografiche e in quelle sociali lo si ritrova anonimamente disperso nelle diverse categorie (alcune categorie) che descrivono la struttura della popolazione alla quale appartiene o che riportano le caratteristiche (alcune caratteristiche) dei fenomeni demografici nei quali è coinvolto.³

Non voglio nemmeno sfiorare lo spinoso problema se così facendo non si rischi di trasformare l'uomo, la persona, in un oggetto; il soggetto in una pluralità di oggetti rilevati nei loro caratteri, statici o dinamici che siano. I procedimenti dei primordi dell'informatica inducono a temerlo, con l'unità statistica ricondotta a una serie di fori su una o più schede meccanografiche, le quali venivano poi processate nelle selezionatrici per singola colonna, oppure nelle tabulatrici attraverso una programmazione assai più complicata per incroci di valori in alcune delle colonne: i pacchetti di schede raccolti nei diversi cassetti o i contatori d'incrocio stavano a testimoniare fisicamente il frazionamento dell'unitarietà

2. Cfr. ISTAT [2006a]: Tavv. 2.7 e 2.8, pp. 45-48. V. Figura 1.

3. «Nello stato connesso e continuo della persona vivente malagevole è afferrare un capo filo e ridurre le nozioni ad alcune funzioni capitali.» [Gioia, 1829: 702].

degli individui (o degli eventi) in caratteri di volta in volta rilevanti o del tutto trascurati, mentre altri, i più, non vengono o addirittura non possono essere rilevati.

Ma non voglio nemmeno condividere fino in fondo la sarcastica notazione di Arthur Conan Doyle, secondo il quale «*While the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty. [...] So says the statistician.*» [Doyle, 1892: 174]. Da un lato, così si rischia di sancire l'inconoscibilità dell'essere umano e del suo operare; dall'altro si trascurano tutti gli aspetti della *variabilità* e dell'*aleatorietà* insiti in qualsivoglia operazione statistica, fosse anche una semplice enumerazione.⁴

È però piuttosto curioso che la demografia e la statistica sociale si siano sviluppate come discipline finalizzate allo studio dei gruppi umani in un contesto culturale, quello occidentale, nel quale il singolo individuo e la considerazione della persona stanno o dovrebbero stare al centro di ogni conoscenza.⁵ La pratica dei valori medi e delle distribuzioni normali parrebbe più adatta alle impostazioni collettiviste delle culture orientali, che prediligono l'identità di gruppo a quella individuale.⁶ Ma in questa apparente anomalia si può forse leggere che la statistica è più la scienza della variabilità [Cox, 1968: 791] che quella dell'*uomo medio* di queteletiana memoria [Gould, 2004: 10].

In origine, la demografia e la statistica sociale si fecero carico della missione di trovare delle *regolarità nei comportamenti umani* riducendo la loro variabilità tra i soggetti che fanno parte di un aggregato. Precorrendo un po' troppo arditamente la pratica della valutazione per "campione", nel 1693 Edmond Halley non ebbe remore a intitolare i suoi calcoli, basati sulle rilevazioni della sola città di Breslavia, *An estimate of the degrees of mortality of mankind*; mentre Johann Peter Süssmilch [1741] enunciò come una regolarità statistica dettata addirittura da Dio la prevalenza dei maschi nel rapporto dei sessi alla nascita; un secolo più tardi, Eilert Sundt [1855] generalizzò quanto aveva raccolto sui matrimoni norvegesi nella cosiddetta "legge di Sundt", che stabilisce i legami tra le "onde" nelle variazioni temporali delle strutture demografiche;⁷ il suo epigono Gustav Sundbërg [1906], dall'osservazione comparata della struttura di diverse popolazioni mondiali trasse un'altra presunta regolarità nella costanza della quota di popolazione in età

4. Per quanto riguarda la variabilità, va tenuta distinta la variabilità del carattere osservato tra i soggetti che compongono una popolazione da quella possibile, nel tempo, nella stessa unità statistica (ad es., il livello d'istruzione, lo stato civile, ecc.). Nei conteggi, entrambi gli aspetti generano un'aleatorietà dei risultati aggregati in conseguenza dell'inclusione o meno nell'enumerazione di alcune unità statistiche, oppure del momento storico in cui è stata fatta la rilevazione.

5. «La norma personalistica è il principio supremo degli atti umani, secondo il quale tutta l'azione dell'uomo, in qualsiasi campo, deve essere adeguato alla relazione "con la persona", che è di fondamentale importanza per l'agire umano» [Wojtila, 2002: 183].

6. Ad es., cfr. Poon, 1996: 649.

7. L'idea sarà ripresa da Easterlin [1987], che l'arricchirà con una teorizzazione sulle conseguenze dell'ampiezza numerica delle generazioni sul loro inserimento nel mercato del lavoro, sul loro "reddito relativo" e, quindi, sui loro comportamenti matrimoniali e riproduttivi.

centrale (all'incirca la metà del totale tra i 15 e i 50 anni). La stessa transizione demografica [Notestein, 1945] è stata spesso intesa come uno schema ineluttabile nell'evoluzione delle popolazioni dal regime di demografia "naturale" a quella "controllata", fatte salve le specificazioni dei tempi e dei livelli finali del processo.⁸

Regolarità di struttura o di fenomeno, dunque, oppure regolarità più complesse di funzionamento e di evoluzione dei sistemi demografici. Regolarità che, in ogni caso, sia per essere stabilite sia per essere verificate nella realtà abbisognano di collettivi adeguati in dimensioni e modalità di funzionamento: delle *popolazioni demografiche*, in altri termini. Definizione, questa, apparentemente agevole quando ci si possa riferire ai cittadini di un determinato Stato o agli abitanti di un territorio ben definito o ai fedeli di una certa religione e così via; ma che in ogni caso, proprio in vista di quelle regolarità, deve presentare una sufficiente numerosità statistica e la doppia capacità di identificarsi e di automantenersi, pur variando nel tempo in dimensione e composizione.⁹

Pertanto, diventano caratteristiche fondamentali degli individui che compongono una popolazione la loro *identità*, opportunamente definita, e la loro *partecipazione al funzionamento del sistema demografico*. Così, ad esempio, non fanno parte della popolazione di una nazione né i turisti stranieri né gli immigrati stagionali, mentre ne possono far parte gli immigrati più o meno definitivi, così come i nazionali temporaneamente all'estero o emigrati stagionalmente: in questo caso, il criterio di aggregazione non è solo quello della cittadinanza, ma è supplito dal contributo, effettivo o potenziale che sia, al mantenimento nel tempo della popolazione di assegnazione.

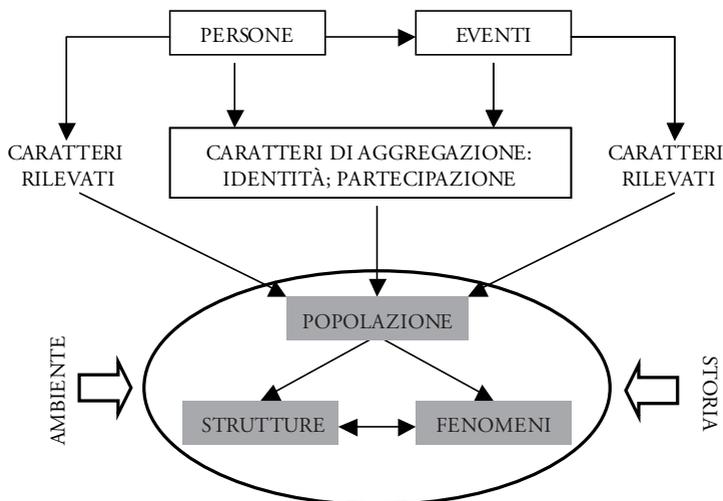
Questi *caratteri di aggregazione* diventano allora prevalenti rispetto a ogni altro carattere dei soggetti coinvolti, tanto da un lato da rappresentare le variabili *pivot* attorno alle quali vengono conformate le strutture e raccolti gli eventi in fenomeni, dall'altro da costituire l'elemento unificante affinché si possa considerare l'aggregato di quei soggetti

8. La cosiddetta "seconda transizione demografica" [Lesthaeghe *et van de Kaa*, 1986; van de Kaa, 1987] ha di per sé un'estensione meno universalistica (Europa o Paesi industrializzati), mentre coinvolge aspetti non strettamente demografici quali il divorzio, la coabitazione *more uxorio*, i modelli famigliari, ecc.

9. «*Coesione e continuità* sono appunto i fatti salienti delle società umane, intorno ai quali si esercita con diverso intento l'indagine dei sociologi e dei demografi». [Benini, 1901: 13; corsivo mio]. Su questi aspetti si è sollevata di recente una diatriba sull'uso "politico" del concetto bio-sociologico di popolazione e su come questo sarebbe stato "scoperto" quando, a partire dalla formazione degli Stati nazionali, i modi del governare passarono dal controllo poliziesco, individuale, a quello liberale, di massa: la popolazione diventa un corpo sociale, meglio biosociale, oggetto delle azioni di governo; le regolarità demografiche, messe in luce dalle statistiche del 18° secolo, riportano la popolazione a un collettivo di soggetti sottoposti a dei fenomeni naturali [Focault, 2004]. Contro la tesi di una "scoperta" così tardiva del concetto di popolazione, v. Curtis, 2002.

come una popolazione demografica.¹⁰ Da questi *criteri di aggregazione* scaturisce un nuovo meta-soggetto, la *popolazione*, da un lato passibile di classificazioni e analisi strutturali, dall'altro capace di manifestare una funzionalità e una vitalità meritevoli di studi sia descrittivi, sia investigativi sulle reciproche connessioni tra i fenomeni demografici e sui condizionamenti che su questi aspetti esercitano l'*ambiente* e la *storia* (Figura 2).

FIGURA 2 – SCHEMA DI PASSAGGIO DALLE UNITÀ STATISTICHE DI RILEVAZIONE AI DATI AGGREGATI DI POPOLAZIONE



In definitiva, all'unitarietà della persona nel complesso dei suoi caratteri viene imposto di "sciogliersi" nella popolazione alla quale risulta assegnata in base ai criteri di aggregazione e nella quale essa porta un patrimonio ristretto di caratteri rilevati o di fatto utilizzati nelle analisi statistiche.

In tutte queste ricerche, la demografia tradizionale, quella basata sui *dati macro*, ha svolto il suo compito in modo egregio, arrivando a notevoli gradi di approfondimento conoscitivo, alla definizione – come ho già accennato – di alcune presunte regolarità e all'intuizione di alcuni meccanismi causali nell'evolvere dei fenomeni demografici. Tuttavia, questi ultimi o, meglio, gli eventi che li compongono sono per lo più frutto di

10. Con questi significati va accolta la definizione di popolazione proposta da Livi Bacci [1981: 1-2], per cui è tale «... un insieme stabilmente costituito di individui [...] legati da vincoli di riproduzione e identificabili da caratteristiche comuni (territoriali, politiche ecc.)». Viene peraltro sottolineato che «... il significato di popolazione è quanto mai elastico», mentre «... la stabilità inerente a una popolazione significa che questa ha una continuità nel tempo, ma non significa che debba essere eterna». A questo proposito, vengono in mente le variazioni territoriali che lo Stato italiano, e quindi la sua popolazione, hanno subito nel corso dei centocinquanta anni di storia unitaria.

comportamenti che si inseriscono in catene di atteggiamenti mentali, preferenze, decisioni, azioni, sia di singoli individui, sia di piccole unità operative demograficamente rilevanti (coppie, famiglie e, in altri contesti antropologici, clan o tribù), che vanno sotto il nome comune di *agenti*. Da qui la necessità di andare oltre l'epifenomeno, con una nuova missione: arrivare alle *determinanti dei comportamenti*, in modo sia da evidenziarne i legami funzionali con i caratteri statistici rilevati sugli agenti sia, se possibile, da risalire la catena decisionale o, quanto meno, connettere l'evento allo studio con la serie di eventi causali che lo hanno preceduto.¹¹

11. Sui tre livelli di *mission* della demografia v. Livi Bacci [2000: 23-4]: «...conoscenza del “sistema demografico” [...] conoscenza [...] del continuo mutamento nei rapporti tra sistema demografico e contesto [...] studio [...] delle determinanti dei comportamenti individuali [...] e del loro combinarsi a secondo delle caratteristiche e delle storie individuali, o gruppi familiari o gruppi associati...».

II.

Individui e collettività: classificazioni, profili, macrodati e l'ecological fallacy

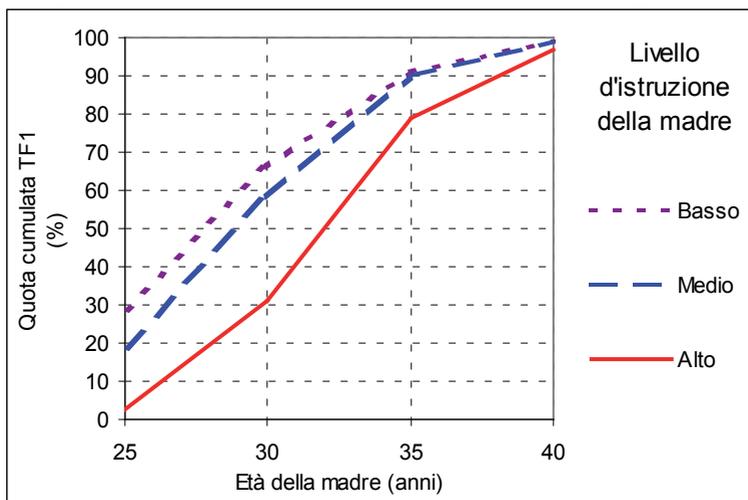
Il primo passo in questa direzione fu compiuto attraverso le sotto-classificazioni, la *demografia differenziale* [Henry, 1959: 24-25; Corsini, 1997] e la *demografia sociale* [Ford et DeJong, 1970; Sonnino, 1979; Billari, 1998]. La comparazione delle strutture e delle dinamiche condotta su opportuni e significativi (in senso gnoseologico) *sottogruppi* di una popolazione permette di apprezzare eventuali differenze e di ipotizzare che queste “dipendano” dai caratteri differenziali tra gli aggregati o, più in generale e con maggiore contenuto euristico, dal “significato” che viene attribuito a ciascun sottogruppo. La misura dell'età al primo parto per livello d'istruzione della madre è un esempio della ricerca sia delle influenze dirette della durata degli studi sul calendario delle scelte riproduttive, sia dei possibili differenziali nei comportamenti riproduttivi tra “classi” socioeconomiche qui delineate in base all'istruzione della donna.

Tuttavia, mentre la dipendenza diretta può essere valutata con sufficiente rigore da opportuni indici di connessione della variabile dipendente “età al primo parto” con la variabile esplicativa “livello d'istruzione della donna”,¹ quest'ultima variabile può risultare insufficiente ai fini della stratificazione socioeconomica e, talora, perfino fuorviante, in

1. Per la verità, per valutare l'entità del rinvio nel processo riproduttivo avrebbe più senso trasformare i titoli nelle corrispondenti durate degli studi, mentre la connessione con l'età al primo parto della singola donna dovrebbe poter tenere conto degli eventuali anticipi o ritardi nel suo percorso formativo. Ma tutto ciò assume un senso non tanto nell'ottica dei raggruppamenti per classi, quanto in quella delle correlazioni dirette tra variabili individuali.

quanto essa è solo uno dei molti *parametri* che servirebbero per definire lo strato sociale al quale appartiene la donna o la coppia di cui essa fa parte.² In altri termini, il livello d'istruzione è solo uno degli *indicatori* attraverso i quali si possono sviluppare confronti di demografia differenziale impostati sulla *variabile latente e complessa* "condizione socioeconomica"; utilizzarne uno solo o solo pochi allo scopo di classificare le unità di rilevazione apre al rischio di conglobare situazioni molto diverse tra loro e di attribuire a un aggregato di fatto eterogeneo i valori risultanti dalla media della variabile dipendente e di quelle esplicative rilevate nei soggetti che vi ricadono.

FIGURA 3 – NASCITE PRIMOGENITE CUMULATE FINO ALLE ETÀ DI 25, 30, 35 E 40 ANNI PER LIVELLO D'ISTRUZIONE DELLA MADRE: ITALIA, LUGLIO 2000 – GIUGNO 2001 (VALORI PERCENTUALI SUL TOTALE PER LIVELLO D'ISTRUZIONE)



FONTE: ELABORAZIONI SU DATI ISTAT [2006B], TAV. 1.4.

Sul fronte opposto, muoversi nella direzione di una classificazione molto minuta per rispettare il più possibile le sfaccettature della realtà presenta il rischio crescente di una ridotta significatività statistica dei gruppi così costituiti, specie quando l'origine dei dati sia un'indagine campionaria, che per ciò stesso ammette solo sottocampioni di un'ampiezza sufficiente a garantire che le differenze osservate tra loro non siano dovute al solo effetto

2. Problemi dello stesso tipo, ma ancora più gravi, presentano i casi in cui la variabile di stratificazione per qualche motivo (inaffidabilità delle risposte, delicatezza dell'argomento, *privacy*, ecc.) non è direttamente rilevabile ed è quindi sostituita da una o più *variabili proxy* che la rappresentano in modo approssimato. Ad es., il livello dei consumi o di certi consumi viene utilizzato per stimare il reddito delle famiglie trascurando le possibili differenze nelle loro propensioni al consumo e al risparmio, oppure tenendone conto in modo standard attraverso funzioni che collegano quelle propensioni al numero e all'età dei loro componenti.

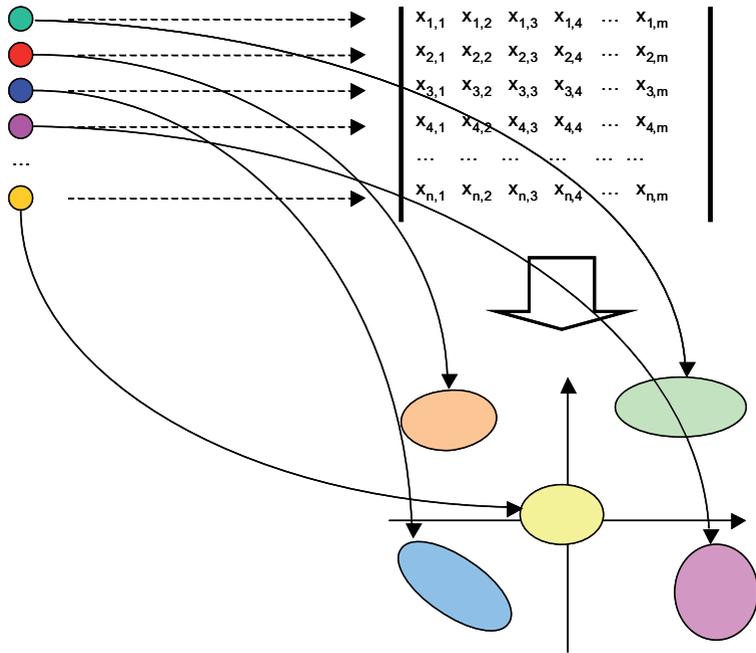
del caso. Anche mantenendosi all'interno di tali limiti (e non sempre ciò accade nelle ricerche che circolano e nei risultati pubblicati), i confronti tra valori medi di gruppo andrebbero proposti con i competenti intervalli di stima, i quali crescono al diminuire della dimensione dei sottocampioni, così che aumenta la probabilità che le differenze tra quei valori medi risultino statisticamente non significative.

Lo sviluppo della *statistica multivariata* ha permesso di arricchire il quadro dei parametri sui quali procedere alla classificazione dei gruppi. Le varie tecniche di raggruppamento (*clustering*), l'analisi delle corrispondenze multiple e, da ultimo, l'analisi congiunta (*conjoint analysis*) consentono di definire aggregati formati su più attributi qualificanti scelti in base alla disponibilità dei dati rilevati e applicando un modello interpretativo e classificatorio della realtà che fa fulcro su di essi.

Lo schema procedurale è quello rappresentato in Figura 4, dove l'insieme dei valori x_{ij} delle m variabili j raccolte ed elaborate per ogni unità i delle n rilevate concorre a definire un sistema di classificazione complesso nel quale trovano collocazione quelle stesse unità statistiche. Ai valori medi di gruppo, utilizzati nel caso di classificazioni su una sola o solo poche variabili, si sostituiscono i punteggi sugli assi principali e delle procedure identificative di *profili* di gruppi che li caratterizzano e li contrappongono tra loro in modo efficace. Le singole unità statistiche vengono quindi ricondotte ai quei profili che hanno contribuito a definire, ma nei quali difficilmente si possono riconoscere in modo puntuale. Per di più, nelle procedure di identificazione dei profili mancano criteri oggettivi e assoluti, per cui in larga misura questi dipendono dalla "sensibilità" e dalla "esperienza" del ricercatore. In definitiva, alla fine del processo di classificazione i singoli soggetti si trovano incasellati in profili che colgono solo in parte le loro caratteristiche e che, piuttosto, rappresentano da un lato la struttura complessiva della popolazione di appartenenza, dall'altro le regole di classificazione e le definizioni fissate dal ricercatore.³

3. Senza voler proporre analogie che risulterebbero insostenibili, va però notato che anche in fisica quantistica il principio d'indeterminazione, almeno nella sua prima interpretazione autentica [Heisenberg, 1927], attribuisce un vincolo decisivo all'intervento dell'osservatore in qualsiasi misura della posizione e della quantità di moto di una particella elementare. Sulle influenze anche del principio di indeterminazione sull'evoluzione della ricerca statistica v. Rizzi, 2002.

FIGURA 4 – SCHEMA DI PASSAGGIO DALLE UNITÀ STATISTICHE AL LORO PROFILO

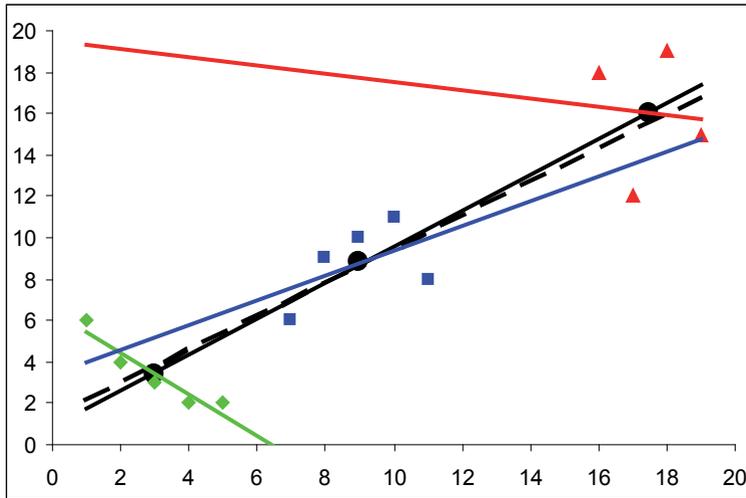


A questo punto sono i raggruppamenti a sostituirsi alle unità statistiche originarie e su questi verranno poi condotte tutte le indagini investigative, comprese le ricerche sulle connessioni tra le variabili esplicative e quelle che descrivono il fenomeno allo studio. Ne consegue che il verso e la misura della connessione sono valutati in base ai valori qualificativi dei gruppi, che vengono a sostituirsi ai valori originari che presenta ciascuna unità statistica che ne fa parte. In altri termini le caratteristiche del gruppo diventano le caratteristiche comuni a tutte le unità statistiche che vi appartengono. Questa sostituzione può indurre a errori di valutazione che vanno sotto il nome di *ecological fallacy*:⁴ In particolare, le relazioni trovate a livello dei valori di gruppo potrebbero risultare non valide e anzi opposte all'interno dei gruppi e in qualche misura diverse anche se si passa attraverso i dati individuali, così come è esemplificato in via teorica in Figura 5⁵.

4. Il problema fu posto da W.S. Robinson nel 1950, che utilizzò il termine “*ecological correlation*”; fu poi ripreso da L. Goodman nel 1953 come “*ecological regression*” [Jargowsky, 2005]. Il termine “*ecological fallacy*” sembra entrato negli studi demografici alla fine degli anni Settanta con alcuni articoli sui comportamenti riproduttivi differenziali pubblicati su *Biodemography and Social Biology*.

5. Ancor più fuorviante, del resto, sarebbe l'errore opposto, la *fallacy of composition*, quando si volesse attribuire la relazione trovata nei singoli gruppi all'insieme della popolazione. Su questo dovremmo riflettere quando ci trovassimo nella necessità di stimare delle relazioni in base a dati relativi a solo una parte della popolazione: ad es., gli immigrati regolari rispetto alla totalità della presenza straniera in Italia.

FIGURA 5 – ESEMPIO DI DIFFORMITÀ DELLE RELAZIONI TRA GRUPPI E NEI GRUPPI



Considerate tutte queste limitazioni e possibili distorsioni, la domanda è: si possono fare buoni studi di popolazione attraverso dati demografici aggregati? Vi sono almeno due aspetti che sembrerebbero negarlo: l'omissione dell'*eterogeneità* delle unità statistiche di base, eterogeneità che, peraltro, per alcuni versi è stata utilizzata proprio per definire i sottogruppi d'appartenenza;⁶ e l'*ottica trasversale* che, in prevalenza o in esclusiva, definisce i gruppi e analizza le eventuali relazioni tra variabili.⁷

Al primo problema sembrerebbe possibile venire incontro moltiplicando le partizioni, fatti salvi, ovviamente, i limiti di significatività delle misure sugli aggregati. Per il secondo, invece, sarebbe necessario modificare l'ottica per impostarla in una prospettiva longitudinale, così che si possa tenere conto anche della dinamica dei caratteri variabili nel tempo e dell'eterogeneità che ne consegue. Quest'ultima operazione porta direttamente alla considerazione delle *coorti*, identificate dalla comunanza di uno o più *eventi origine*, e ai loro effetti sulle misure in atto (*cohort bias*) [Lombardo, 1994; Salvini et San-

6. «[...] la *démographie différentielle* se préoccupe surtout de différences entre groupes, plutôt que des différences entre individus. Or le problème que l'hétérogénéité des populations pose à l'analyse démographique ne serait pas résolu, si l'on était parfaitement informé des différences entre groupes et de leurs corrélations ; il faudrait encore être informé des différences et des corrélations à l'intérieur des groupes, c'est-à-dire de tout ce qui dans les caractéristiques individuelles, physiques et psychiques, est susceptible d'influer sur les aptitudes, sur les comportements, sur les réactions, sur les résistances, sur les susceptibilités» [Henry, 1959: 25].

7. «Applying the classic demographic analysis to each sub-population makes it possible to compare specific behaviour patterns. This analysis is essentially cross-sectional. [...] In cross-sectional analysis, the individual's category at the time of observation is taken into account» [Courgeau et Lelièvre, 1992: 37]

tini, 1999]. Entrambi i problemi indicano quanto sia preferibile sviluppare delle analisi demografiche basate sui dati individuali rispetto a quelli di gruppo: solo i dati individuali, infatti, possono fornire tutto il patrimonio informativo e relazionale necessario a evitare l'*ecological fallacy* e il *cohort bias*. Così facendo, però, si esce fatalmente dalla *demografia macro* per entrare nella *micro-demografia*.

Tuttavia, a mio parere quella domanda è mal posta; almeno nei tempi verbali. È stata fatta della buona, anzi, dell'ottima demografia con i dati macro. Le limitazioni (delle quali peraltro si era perfettamente consapevoli) derivavano meno dall'ampiezza del patrimonio informativo e molto più dalle ridotte capacità di elaborazione disponibili al tempo. Ma, entro quei limiti, è stato sviluppato molto lavoro euristico, che da un lato ha trovato conferme nei successivi studi su base micro-demografica, dall'altro ha continuato a indirizzare gli sviluppi attuali e futuri degli studi di popolazione.

III.

Della disponibilità e della qualità dei dati demografici

Prendo a prestito la frase che l'anno scorso “*vielle*” ha inviato all'ISTAT per la Giornata mondiale della statistica: «I dati che hai non sono quelli che vuoi; i dati che vuoi non sono quelli di cui hai bisogno; i dati di cui hai bisogno non potrai mai averli» [ISTAT, 2010: *tweet* n. 46], ma lo faccio per contestargli/le sia la realtà dei fatti, sia le sue prospettive pessimistiche.

Non abbiamo mai avuto a disposizione così tanti dati come ora, tanto da doverci preoccupare da un lato dell'eccesso d'informazione (*information overload* [Himma, 2007]), dall'altro del rispetto della *privacy* e delle libertà personali messe in discussione da una raccolta, palese e occulta, dei dati più disparati che riguardano noi e la nostra vita privata [Langheinrich, 2001]. Sono però pronto ad ammettere che non è vero che molti dati comportino necessariamente una buona informazione statistica. Sono convinto piuttosto del contrario, per le crescenti difficoltà nel gestirli (a ogni nuova variabile m le possibili relazioni R_i tra le variabili crescono secondo la serie dei numeri triangolari: $R_m=R_{m-1}+(m-1)=[m \cdot (m+1)]/2$) e per le incongruenze che si possono generare tra loro.

Pertanto sarebbe erroneo, o almeno limitativo, far corrispondere “abbondanza e scarsità” alle parole “ricchezza e povertà” nel titolo di questa mia comunicazione; mentre un'altra antitesi è possibile: “pregevolezza e mediocrità”, la quale attiene più agli aspetti qualitativi che a quelli quantitativi. Entrambe queste valutazioni vanno però storicizzate e contestualizzate, in modo da tener conto delle difficoltà nella raccolta dei dati di base e di quelle nella loro gestione. Ad esempio, considerato il contesto amministrativo e culturale di quei tempi ancora burrascosi, sembrano quasi un miracolo i risultati del primo censimento del Regno d'Italia (31 gennaio 1861), pubblicati con dovizia di analisi e sensatezza dei commenti già a partire da sedici mesi

dopo la data della rilevazione.¹ Altrettanto ammirevoli appaiono i contenuti delle schede di nascita (ISTAT D.1 e ISTAT D.2), in uso fin dal 1926, ma cancellate nel 1998 da un ottuso regolamento di attuazione della “Bassanini-bis”.²

In contesti statisticamente meno sviluppati o anche solo con una tradizione amministrativa diversa (ad esempio, dove mancano le anagrafi della popolazione), una parte importante della ricerca statistico-demografica consiste nella stima dei dati di popolazione [Bryan, 2004], mentre per le aree sottosviluppate dal punto di vista statistico, dove sono ancora carenti le registrazioni degli eventi vitali e gli eventuali censimenti soffrono di gravi distorsioni, sono stati elaborati specifici metodi di stima dei dati demografici [Popof *et* Judson, 2004]. Ma possiamo noi chiamarci del tutto fuori da queste pratiche vantando la lunga tradizione delle nostre statistiche demografiche e la loro rispondenza ai criteri EUROSTAT?³ Una serie di ragioni induce a pensare che così non sia, né possa essere:

- a) Fa parte della popolazione che vive in Italia un numero imprecisato di immigrati “irregolari” che sfuggono all’anagrafe e alle altre forme di registrazione statistica. La loro origine non deriva solo (come qualcuno vorrebbe far credere) dagli ingressi clandestini (magari solo via mare), ma piuttosto dallo scadere di precedenti permessi di soggiorno non seguiti dall’abbandono del territorio nazionale. Le loro caratteristiche strutturali e i loro comportamenti demografici con ogni probabilità sono influenzati proprio dalla loro condizione d’irregolarità e, pertanto, non è solo la stima del loro numero,⁴ ma anche la valutazione della loro demografia differenziale a costituire un problema per la qualità dei dati sull’immigrazione e sulla presenza straniera in Italia.⁵

1. V. i relativi volumi a partire dal primo [MINISTRO D’AGRICOLTURA, INDUSTRIA E COMMERCIO, 1864].

2. Per un preciso resoconto di questa irrecuperabile perdita, v. ISTAT, 2006b: 9-10. Sul tentativo di recupero alla statistica del patrimonio informativo contenuto nei CEDAP, v. CISIS, 2004. Per dei metodi di ricostruzione delle serie storiche della fecondità troncate al 1996, cfr.: Giorgi *et* Viola, 2003; Gallo, aa. 2006/07.

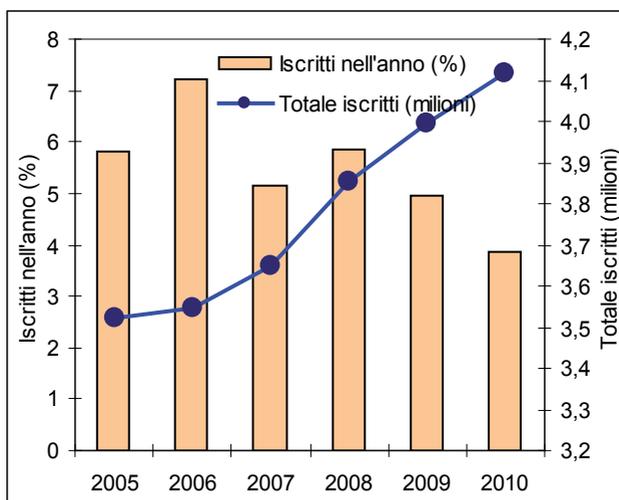
3. Per una sintesi, ma estesa a tutta la statistica pubblica, v. Rondini, 2003; v. anche Golini, 1991. Tuttavia, la tradizione di statistiche sulla popolazione ampie e accurate risale addietro nei secoli e costituisce ancora la fonte per i numerosi studi di demografia storica. Su questa, si vedano i due volumi dell’apposito Seminario del 1971-72 [Comitato italiano per lo studio della demografia storica, 1974]. Si vedano i buoni risultati della *peer review* dell’ISTAT alla pagina dell’EUROSTAT <http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=PEER_REVIEW_IT_2006>

4. La Fondazione ISMU (di fatto, Gian Carlo Blangiardo) fornisce ormai ogni anno una stima della presenza straniera in Italia, distinguendo le presenze regolari da quelle irregolari [ad es., Fondazione ISMU, 2011: 29 ss.gg.]. Per la metodologia della stima v. Blangiardo *et* Tanturri, 2006.

5. Tra i primi a porsi il problema della qualità dei dati relativi all’immigrazione in Italia, nell’ambito della ricerca interuniversitaria su “La presenza straniera in Italia”, finanziata dal CNR e diretta da Nora Federici, o al di fuori di essa: Gesano, 1983; Cariani, 1985; Cortese, 1987; Maciotti *et* Pugliese, 1991.

- b) Sul versante opposto, l'interesse quasi nullo dei migranti che lasciano il nostro Paese a cancellarsi dai registri che li includono fa conservare nella popolazione una componente non più presente, che andrebbe sottratta per quantità ignote al dato demografico statico, mentre non può più contribuire a quelli dinamici. Anche in questo caso, la particolarità di questi migranti richiederebbe la stima non solo del loro numero, ma anche delle loro strutture, per poterli sottrarre alle sub-popolazioni di competenza che, altrimenti, nei confronti con le altre rischiano di vedere sottostimate le proprie statistiche di movimento.
- c) Problemi simili, ma di minore dimensione, derivano dall'assenza di nostri concittadini che si trovano all'estero per un tempo più o meno lungo. Le registrazioni anagrafiche di coloro che lo sono in modo "permanente", infatti, dovrebbero essere trasferite nella anagrafe degli italiani residenti all'estero (AIRE). Dopo la riforma del 1988, ciò avviene qualora la residenza all'estero si prolunghi oltre i dodici mesi, ma solo su dichiarazione spontanea dell'interessato o per accertamento d'ufficio, mentre il mantenimento o il recupero per questa via del diritto di voto alle elezioni politiche potrebbe influire sul calendario delle iscrizioni e, quindi, sull'ammontare totale.⁶ I problemi sulle statistiche demografiche nascono dunque da coloro che, pur risiedendo stabilmente o quasi all'estero, continuano a permanere nell'anagrafe corrente e dalla possibilità che si verifichino ritardi o mancate comunicazioni degli eventi demografici che riguardano costoro.

FIGURA 6 – TOTALE DEI CITTADINI ITALIANI ISCRITTI ALL'AIRE E QUOTA DEGLI ISCRITTI NELL'ANNO: ITALIA, 2005-2010



FONTE: ELABORAZIONI SU DATI <[HTTP://INFOAIRE.INTERNO.IT/STAT_NOTE.HTM](http://infoaire.interno.it/stat_note.htm)>

6. Sia la crescita degli iscritti all'AIRE, sia il calendario delle iscrizioni, curiosamente inflazionate negli anni di consultazione elettorale come riportato in Figura 6, pongono «ragionevoli dubbi sull'attendibilità del dato» [Pugliese, 2002: 15].

- d) Il mantenere o il fissare la propria residenza più secondo i propri interessi che nel rispetto delle norme che la vorrebbero presso il Comune di dimora abituale possono creare analoghi problemi con effetti sulle statistiche relative alla distribuzione territoriale della popolazione e dei relativi fenomeni demografici. Talora, agli interessi individuali si possono sommare quelli della amministrazione locale o della collettività quando determinati vantaggi o trattamenti economici siano collegati a soglie basate sul numero degli iscritti nell'anagrafe comunale [Gesano *et al.*, 1993].
- e) Sotto questo aspetto, però, è forse proprio lo stesso concetto di *dimora abituale* a non essere più in grado di cogliere l'ampiezza degli spazi di vita entro i quali una consistente parte della popolazione ormai si muove e si identifica: abitazione principale, luogo di lavoro, abitazione di origine, seconda/terza casa, eccetera; luoghi spesso situati ciascuno in un Comune diverso e magari tra loro lontani, ma che includono tutti una parte delle nostre abitudini residenziali.⁷
- f) Allo stesso modo, il modificarsi delle modalità e delle forme di vita sociale può rendere gli strumenti di rilevazione e le categorie statistiche utilizzate non più adeguate a cogliere alcuni fenomeni e, pertanto, i relativi dati possono risultare distorti o addirittura carenti. Si ha la netta impressione, ad esempio, che le "convivenze" affettive non vengano pienamente rilevate nelle varie forme che esse ora assumono [Costa, 2007], sopravvalutando di conseguenza il numero di "famiglie" unipersonali anche quando sussistono almeno parte dei presupposti di fatto per non dover considerare una persona (anzi, spesso due persone) come *single*.

Tutto ciò impone un'attenta sorveglianza sulla *qualità statistica* (in particolare sulla *rilevanza*) dei dati che si utilizzano nella ricerca e il costante impegno da parte dei produttori dei dati a «tenere il passo dei mutamenti che avvengono nella società e dei miglioramenti delle metodologie di produzione» [ISTAT, 2009: 10]. Da qui l'importanza per il ricercatore di disporre dei *metadati* che chiariscano il reale contenuto del materiale statistico che viene messo a disposizione; ma anche l'utilità dell'analisi critica delle fonti, dalla quale cogliere lo stimolo al miglioramento evitando di rifugiarsi in "difese d'ufficio", spesso giustificate in modo più o meno pretestuoso da vincoli di legge o da pur innegabili difficoltà burocratiche.

Vi è tuttavia una qualità ormai fondamentale delle statistiche che in qualche modo potrebbe rischiare di limitarne la capacità rappresentativa di contesti o di aspetti particolari: si tratta dell'esigenza di poter comparare i dati tra i diversi Paesi; *comparabilità internazionale*, dunque, sia di definizioni e di contenuti, sia dei metodi di rilevazione e di

7. Non ha dato i frutti sperati il tentativo fatto con il 14° Censimento (2001) di rilevare nella Lista B ("Persone che NON hanno dimora abituale nell'alloggio, cioè che vivono temporaneamente, o che sono occasionalmente presenti, nell'alloggio") le "Persone non residenti in alloggio da più di 90 giorni" tramite i quesiti B/3: una volta eliminati coloro che avevano dimora abituale all'estero o nello stesso Comune in cui si trovava l'alloggio censito, essi si riducono a 683.347, di cui 25.064 presenti per vacanza e 134.310 per motivi di studio.

elaborazione. Il problema fu affrontato fin dall'immediato dopoguerra cogliendo quanto già era in atto nella rilevazione delle cause di morte e delle malattie e analizzando lo stato dell'arte e le concrete possibilità nelle diverse branche e statistiche specifiche [MILBANK MEMORIAL FUND, 1949]. La crescente importanza degli organismi internazionali (ONU e sue agenzie, OCSE, eccetera) e le deleghe soprannazionali (COMMISSIONE EUROPEA attraverso la sua agenzia EUROSTAT)⁸ hanno ampliato di molto la sfera della comparabilità internazionale delle statistiche [Franchet, 1991]. Nel campo di quelle sulla popolazione: i censimenti sono ampiamente normati, pur lasciando ai singoli sistemi nazionali più o meno larghi margini di autonomia nel soddisfare alcune esigenze conoscitive comuni a tutti; altrettanto comuni sono i principi che regolano la tenuta delle statistiche sul movimento della popolazione, anche se nella fondamentale distinzione tra i sistemi statistico-demografici basati sulle registrazioni anagrafiche e quelli che si limitano alle registrazioni degli eventi di nascita e di morte (*vital statistics*); le indagini sulle forze di lavoro hanno da tempo un ampio substrato comune in tutti i Paesi sviluppati; notevoli sforzi si vanno facendo per rendere comparabili le statistiche sulle migrazioni e sull'integrazione dei migranti [Kelly, 1987; Šteinbuka, 2009].

Il quadro delle statistiche comparabili in campo internazionale tende ad ampliarsi ad altri aspetti. Ma la novità più interessante in campo europeo è venuta dagli accordi su alcune indagini campionarie che sono svolte dagli Istituti di statistica dei diversi Paesi su una base comune (oltre all'indagine ormai tradizionale sulle forze di lavoro, ora utilizzata anche come fonte di dati sulla presenza straniera). *L'indagine europea sui redditi e le condizioni di vita delle famiglie* (EU-SILC) [ISTAT, 2008] costituisce al momento l'esempio più promettente nel campo degli studi di popolazione,⁹ ma altre indagini internazionali si vanno prospettando, anche se c'è da temere che l'attuale crisi economico-finanziaria rimanderà le effettive possibilità di una loro attuazione a breve.

Del resto, la comunità scientifica si era da tempo organizzata per sviluppare indagini e basi dati sulla popolazione che fossero internazionalmente comparabili. La *World Fertility Survey* (WFS) ha fornito nell'ultimo quarto del secolo scorso un materiale statistico di grande valore sul quale sono stati sviluppati anche nel nostro Paese studi basilari per la comprensione del processo riproduttivo dal punto di vista demografico e sociale.¹⁰ L'Italia è

8. Si consultino, in particolare, i materiali d'interesse indicizzati nelle pagine all'indirizzo: <http://epp.eurostat.ec.europa.eu/portal/page/portal/publications/collections/methodologies_working_papers>

9. Google Scholar, ad es., elenca più di 3.000 pubblicazioni nelle quali si fa riferimento alla sigla dell'indagine, di cui più di 200 in italiano.

10. Google Scholar in questo caso elenca più di 500.000 pubblicazioni nelle quali si fa riferimento alla WFS, di cui solo 44 in italiano. Nell'ambito della WFS vanno tuttavia ricordati i fondamentali contributi forniti dall'Indagine nazionale sulla fecondità (INF-1), eseguita da ricercatori delle Università di Padova, Firenze e Roma sotto la guida di Paolo De Sandre [1982], e dalla successiva seconda indagine INF-2, che raccolse anche il contributo scientifico di ricercatori delle Università di Bologna e di Trieste, dell'ISTAT e del CNR-IRP [De Sandre *et al.*, 1999].

invece rimasta fuori del progetto dell' *European Social Survey* (Ess), che si è svolta in quattro *round* tra il 2002 e il 2008; mentre sono stati raccolti ma non sono ancora disponibili i dati dell'indagine italiana che rientra nel *Generations and Gender Survey* (Ggs) dell'omonimo programma UNECE. Questo ultimo programma era stato preceduto dal progetto europeo *Network for Integrated European Population Studies* (NIEPS), che aveva proposto una serie di tematiche d'analisi, da affrontare prevalentemente via indagini integrate, allo scopo di incrementare la conoscenza e favorire l'intervento politico su alcuni temi di popolazione.¹¹

A conclusione di questo paragrafo un accenno va fatto all'accessibilità effettiva, per il ricercatore, a dati demografici così ampi e ricchi. Prassi correnti, politiche editoriali, concorrenza e imitazione tra produttori di statistiche si sono rincorse con i progressi tecnici nella trasferibilità di grosse masse di dati e nella possibilità di elaborarli in remoto. Rimandando al prossimo paragrafo la trattazione di questi ultimi, essenziali aspetti, va qui ribadito che «*Publicly funded research data are a public good, produced in the public interest. As such they should remain in the public realm*» e, ancora, che «*[their] availability should be restricted only by legitimate considerations of national security restrictions; protection of confidentiality and privacy; intellectual property rights; and time-limited exclusive use by principal investigators*» [Arzeberger *et al.*, 2004: 136]. Non sempre questi principi di pubblicità dei dati vengono rispettati, ma va riconosciuto che si sono compiuti importanti passi in avanti rispetto al passato, soprattutto riducendone l'uso esclusivo, e ciò anche nel rispetto dei difficili ostacoli posti dalle norme sul trattamento dei dati personali e da loro interpretazioni eccessivamente restrittive.¹² Ora, mentre è fuori discussione che il diritto alla riservatezza rientra tra i diritti fondamentali della persona, come è sancito dalla Carta Europea approvata nel 2000,¹³ è altrettanto indiscutibile che la ricerca statistica bandisce ogni interesse agli aspetti personali (quelli, cioè, che portano all'identificazione nominativa) dei casi trattati e per ciò stesso andrebbe tutelata con la stessa determinazione.

11. «*The general and main objective was to create an all-European platform of national population institutes in Europe that would promote a dialogue and foster research cooperation on policy relevant interactive domains of population and family dynamics on the one hand and socio-economic processes on the other. Three interconnected key structural demographic changes were addressed: Gender relations, family-building and patterns of work; Ageing, intergenerational solidarity and age-specific vulnerabilities; Demographic and cultural specificity and integration of migrants*» [EUROPEAN COMMISSION, 2005: 7]. Per una riflessione di sintesi, v. Avramov et Cliquet, 2005.

12. V. il parere del Garante per la protezione dei dati personali del 15 novembre 2007 <1464806> e il capitolo su “La ricerca statistica e storica” nella sua Relazione 2008 <<http://www.garanteprivacy.it/garante/doc.jsp?ID=1637700>>, oltre al Codice di deontologia e di buona condotta per i trattamenti di dati personali a scopi statistici e di ricerca scientifica <<http://www.garanteprivacy.it/garante/doc.jsp?ID=1556573>>, che però limita la sua applicabilità al Sistema statistico nazionale e non include né la ricerca accademica né quella di privati studiosi, regolate invece dal codice deontologico del 2004, disponibile alla pagina <<http://www.garanteprivacy.it/garante/doc.jsp?ID=1556635>>.

13. Carta Europea dei diritti fondamentali (GUCE 2000/C 364/01), art. 8 c. 1: «Ogni individuo ha diritto alla protezione dei dati di carattere personale che lo riguardano».

IV.

Cenni sugli effetti dell'aumento delle potenzialità informatiche

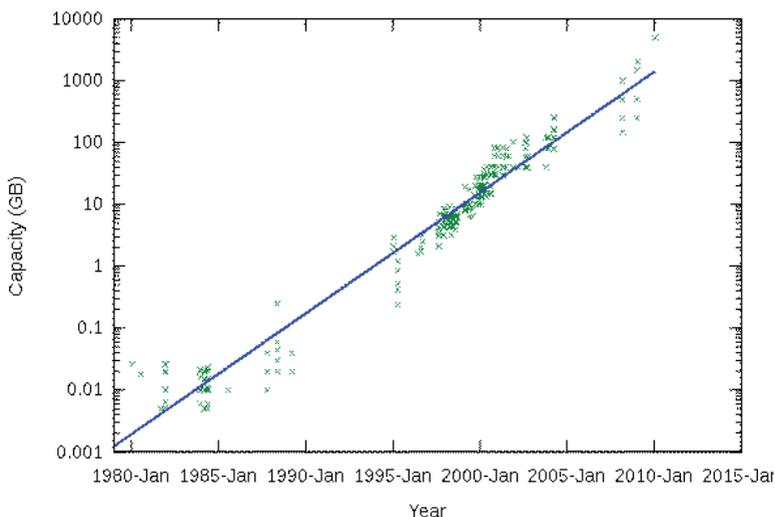
Quando, a metà degli anni Settanta del secolo scorso, portavamo al Centro di calcolo le schede relative agli 8.100 Comuni sui quali studiavamo le vicende dello spopolamento in Italia potevamo sperare di avere i risultati solo la mattina successiva; risultati che poi avremmo ulteriormente elaborato con le OLIVETTI DIVISUMMA e le HEWLETT-PACKARD. Qualche anno prima, nel preparare la tesi di laurea presso l'Istituto di statistica della Facoltà di Economia e commercio dell'Università di Firenze, mi ero impratichito nell'uso dell'OLIVETTI PROGRAMMA 101, il primo "calcolatore" personale da tavolo, che andava programmato passo, passo nelle operazioni aritmetiche e negli spostamenti dei dati tra i registri di cui disponeva.

Qualche giorno fa ho chiesto un'elaborazione piuttosto complessa a partire dai microdati delle indagini sulle Forze di lavoro. Compresa la programmazione in R, in un paio d'ore ho avuto i risultati direttamente sul mio iPad che avevo portato con me a qualche centinaio di chilometri di distanza dall'Istituto.

Il progresso che ha registrato l'informatica nell'*hardware* e nel *software* viene spesso descritto da una curva approssimativamente esponenziale, che va sotto il nome di "legge di Moore" (Figura 7). È certo, però, che la crescita delle capacità di calcolo ha da subito distanziato di molto le necessità correnti negli studi demografici ed è servita piuttosto all'informatizzazione della produzione dei dati di base, migliorandone la qualità attraverso appositi metodi di controllo e di integrazione dei dati mancanti, ma non sempre riuscendo ad accelerarne i tempi di pubblicazione. Tuttavia, le accresciute disponibilità informatiche e l'appropriazione da parte dei ricercatori di capacità e di autonomia operativa tramite i *personal computer* hanno permesso da un lato l'utilizzazione sempre più dif-

fusa delle banche dati e dei dati di indagini sempre più ricche di informazioni individuali, dall'altro lo sviluppo di modelli sia descrittivi, sia previsivi o di simulazione.

FIGURA 7 – LA CORSA DEI MILIARDI... DI BYTE: CRESCITA DELLA CAPACITÀ DEL DISCO RIGIDO NEL TEMPO



FONTE: WIKIMEDIA COMMONS <[HTTP://EN.WIKIPEDIA.ORG/WIKI/FILE:HARD_DRIVE_CAPACITY_OVER_TIME.SVG](http://en.wikipedia.org/wiki/File:Hard_drive_capacity_over_time.svg)>

Sotto il primo aspetto, riprendendo quanto accennato nel precedente paragrafo, per chi come me ha cominciato la sua carriera di ricercatore imputando su una calcolatrice caterve di dati tratti da tabelle pubblicate su carta¹ rimane uno strabenedetto miracolo il poter accedere liberamente a una banca dati per interrogarla a seconda delle variabili e dei caratteri che di volta in volta ci interessano. In un recente lavoro su una realtà territoriale inserita in una lontana regione transfrontaliera [Gesano, 2010] ho potuto eseguire tutte le elaborazioni necessarie senza muovermi dal mio tavolo di lavoro, semplicemente accedendo alle banche dati degli Istituti di statistica della Svezia, della Danimarca e a quella specifica della regione *Öresund*.

Anche l'ISTAT si sta finalmente attrezzando in questa direzione,² dopo aver messo ampiamente a disposizione dati già elaborati, ma per ciò stesso rigidi rispetto alle possibilità di sviluppare nuovi incroci [Golini *et al.*, 1991]. Per quanto riguarda poi l'accesso

1. S'immagini quanti errori potevano cumularsi in un percorso che dal dato grezzo andava ai conteggi, spesso manuali, alla composizione tipografica, alla lettura, alla digitazione del dato e, infine, alla trascrizione del risultato.

2. V. I.Stat – Guida utente all'indirizzo <[http://dati.istat.it/modalhelp/Istat/WBOS%20User%20Guide%20\(IT\).PDF](http://dati.istat.it/modalhelp/Istat/WBOS%20User%20Guide%20(IT).PDF)>

ai micro-dati di indagini, le diverse modalità offerte dall'ISTAT garantiscono oggi buone possibilità di ricerca a costi nulli o assai ridotti. C'è piuttosto da lamentare lo scarso uso dei ricercatori di alcune forme d'accesso ai dati (soprattutto i "file per la ricerca" del Censimento 2001), forse per difficoltà burocratico-amministrative, forse per l'eccessivo ritardo con cui sono stati messi a disposizione.³

L'esplosiva espansione del ricorso alle indagini campionarie e lo sviluppo travolgente dell'uso della modellistica negli studi di popolazione meritano ciascuno un paragrafo a parte. Ma va ribadito che né l'una né l'altro avrebbe mai potuto avere luogo in quelle dimensioni e con quei ritmi in mancanza dell'aumento esponenziale delle potenzialità informatiche.

3. Secondo i dati che mi sono stati messi gentilmente a disposizione dai funzionari ISTAT, nel corso del 2010 vi sono stati 790 contatti al *call center* per richiedere dati relativi al 14° censimento della popolazione, poco più del 5% del totale dei contatti; meno di un quinto proveniva da ricercatori dell'università o degli enti di ricerca. I contatti per altri dati sulla popolazione hanno coperto quasi l'8%.

V.

Le indagini campionarie, o l'illusione di poter indagare su "ogni cosa"

Tutto cominciò nel 1955 con le interviste a 2.713 donne americane sposate, che furono interrogate su una serie molto vasta di dati personali, di fattori potenzialmente influenti, di aspetti di contesto e retrospettivi, nonché su atteggiamenti e desideri, il tutto in relazione alla fecondità attesa.¹ La metodologia dei sondaggi d'opinione, sviluppatasi tra le due guerre negli Stati Uniti con finalità di previsioni elettorali e di ricerche di mercato, nel dopoguerra si era sposata con la radicata tradizione delle *social surveys* [Eaton, 1930] generando *indagini su base campionaria rappresentativa*: la statistica inferenziale entrava così negli studi sociali mentre, tra i vari aspetti indagati, venivano affrontati alcuni temi più propriamente demografici, soprattutto connessi con i processi riproduttivi e con la mobilità della popolazione sul territorio.

Anche in Italia, intorno al 1970, una serie di indagini campionarie venne ad arricchire il patrimonio informativo degli studi di popolazione. Tra le prime va ricordato lo studio "Fecondità e lavoro della donna: un'indagine in quattro zone tipiche italiane", diretto da Nora Federici [1973] e curato da Bielli, Pinnelli e Russo. Il CNR ed altri enti finanziarono a quei tempi ampie indagini sul campo permettendo di studiare diversi aspetti socio-demografici, come le famiglie degli emigrati [Corsini

1. V. Freedman *et al.*, 1973. Accurate informazioni e l'accesso ai dati della 1955 *Growth of American Families* si trovano alla pagina web <<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/20000>>. Per una rassegna delle indagini del tempo e, in generale, degli studi sulla fecondità differenziale, v. Pinnelli, 1967-68.

et Sonnino, 1972] e l'integrazione degli immigrati in ambiente urbano [Ascolani *et al.*, 1974]. Negli anni successivi, furono diversi gli studi di popolazione nei quali l'indagine sul campo costituì uno dei fulcri principali, se non l'unico. La diversità delle tematiche d'indagine, le specificità di somministrazione dei questionari e dei loro contenuti e, soprattutto, i particolari universi indagati costrinsero spesso a adattare i *disegni di campionamento* e perfino a inventarsi schemi nuovi, come fu per lo studio dell'ignoto universo degli stranieri presenti in Italia [Blangiardo, 1996]. Le indagini si dilatarono a considerare gli antefatti, le intenzioni, le preferenze e le opinioni;² si estesero nel passato degli intervistati, come anche nello spazio delle loro relazioni, talvolta arrivando a raccogliere informazioni (quando non pareri!) per interposta persona.

Il punto di svolta fu segnato dall'ISTAT nel 1993 quando, sulla base di una sperimentazione avviata già nel 1987 e raccogliendo l'esperienza e i contenuti di alcune altre indagini e metodologie di campionamento ormai consolidate, prese forma il sistema delle *indagini multiscopo sulle famiglie* (IMF), in una logica d'integrazione e di ripetizione a cicli dei diversi aspetti indagati [Sabbadini *et Sgritta*, 1996]: all'indagine annuale sugli "Aspetti della vita quotidiana" furono affiancate e ancora si affiancano le indagini quinquennali sulle "Condizioni di salute e ricorso ai servizi sanitari", su "I cittadini e il tempo libero", sulla "Sicurezza dei cittadini", sulla "Famiglia e soggetti sociali" e sull'"Uso del tempo". Ma il sistema delle indagini sociali multiscopo è diventato nel tempo una sorta di *omnibus* sul quale sono stati fatti salire i temi e gli approfondimenti più diversi tramite l'inserimento di moduli *ad hoc*: in particolare, gli approfondimenti sui problemi della disabilità, sulla violenza contro le donne e sulle criticità dei percorsi lavorativi in un'ottica di genere [ISTAT, 2006c].

Va riconosciuto che il sistema delle indagini multiscopo, per quanto non specificamente puntato su aspetti demografici, ha permesso la raccolta di una vasta serie di informazioni sui fattori che possono intervenire nei comportamenti rilevanti dal punto di vista demografico, così come ha fornito un complesso quadro delle variabili sociali. C'è di nuovo da chiedersi quanto sia stato sfruttato veramente questo patrimonio informativo e se indagini mirate non avrebbero potuto soddisfare meglio e a costi inferiori quelle esigenze specifiche di ricerca. Non va però dimenticato che il sistema, proprio per la sua duttilità, ha consentito di attuare agevolmente alcune direttive d'informazione armonizzata prese a livello della Comunità Europea o di altri organismi internazionali.

Del resto, è stato proprio l'aggancio con le direttive della statistica europea a introdurre anche in Italia delle *indagini ufficiali di tipo longitudinale*, che dovrebbero costituire la spina dorsale della ricerca demografica. Agli studi di popolazione italiani

2. Ad es.: per antefatti, intenzioni e preferenze, v. un confronto tra INF-1 e INF-2 in De Rose *et* Racioppi, 1999; sulle intenzioni vs. realizzazioni, v. Menniti, 2005, per la quale si è ricorsi a delle indagini su campione ripetuto; per delle indagini sulle opinioni degli italiani su alcuni argomenti demografici, v. Palomba, 1987; 1991.

è sempre mancata, infatti, una base dati come quella fornita dal *British Cohort Study*, che si rinnova ogni anno fin dal 1970. Vero è che le nostre registrazioni anagrafiche e quelle di stato civile sarebbero in grado di fornire i dati classificati anche per generazione di appartenenza;³ ma questo non è sufficiente per delle ricostruzioni individuali seguite. Vi ha sopperito prima l'indagine *European Community Household Panel* (ECHP), disponibile per gli anni dal 1994 al 1997, poi, dal 2004, l'indagine *Statistics on Income and Living Conditions* (EU-SILC), dalle quali, se pure non propriamente destinate agli studi demografici, sono state estratte informazioni assai utili per analisi longitudinali.⁴

Se si escludono alcune specifiche sezioni dei questionari INF per quel che riguarda la storia riproduttiva e di quelli in alcune indagini più corpose sugli immigrati, in Italia mancano anche delle *indagini retrospettive* sociodemografiche di largo respiro, alla Courgeau – per intendersi – che con la sua indagine del 1981 *Trois Biographies* (familiare, professionale e migratoria) provò a ricostruire la sequenza degli eventi demograficamente rilevanti nella vita delle persone.⁵ Va detto che, per quanto il progetto fosse seducente nella sua pretesa di collegare la mobilità sui tre piani, risultò impacciato proprio dalla complessità del ventaglio degli eventi e delle loro sequenze, così da non poter essere gestito in modo soddisfacente attraverso una indagine campionaria di “soli” 4.602 intervistati, ai quali fu peraltro somministrato un questionario di ben quattordici pagine fitte di domande.

Già, i *questionari*, croce e delizia dei ricercatori sociali, i quali, quando ne costruiscono uno, combattono un conflitto interiore tra eccesso e rigore: il primo dominato dalla loro curiosità e dalla frenesia di “non perdere l'occasione per...”, il secondo imposto dai limiti di spesa e dalla sostenibilità delle interviste a fronte del *disturbo statistico* inflitto all'intervistato. Credo che ognuno di noi abbia per ciò una *Mea culpa* da recitare; per parte mia, ricordo il questionario di una delle prime indagini sugli stranieri in Italia [Gesano, 1986], che costringeva intervistato e intervistatore a un colloquio della durata di quasi un'ora. In effetti, la somministrazione faccia a faccia spinge ad ampliare i questionari, perché il costo della rilevazione è dominato dai costi del contatto, mentre è solo debolmente proporzionale alla durata dell'intervista che, tra l'altro, si può ora ridurre con le tecniche CAPI. La prospettiva è però cambiata da quando la somministrazione avviene via telefono, specie nella versione CATI, il

3. A questo proposito, c'è da chiedersi per quale ragione la data di riferimento del Censimento attualmente in corso sia stata anticipata di una decina di giorni rispetto a quelle adottate negli ultimi tre censimenti, tutte intorno al 20 ottobre, così che si potevano fare agevolmente delle ricostruzioni per generazione che ora risulteranno per ciò meno precise.

4. Va ricordato lo sviluppo in atto di studi basati sul *record linkage*, nelle sue diverse accezioni, risultati e potenzialità, comprese quelle di ricavare informazioni *panel* anche da indagini su soggetti diversi o dall'incrocio di indagini campionarie e dati amministrativi [Castagnaro *et al.*, 2011].

5. V. la scheda dell'indagine, i questionari e i documenti relativi alla pagina web <http://www.ined.fr/fr/ressources_documentation/enquetes/liste_enquetes/bdd/enquete/3B_1981/>.

che impone interviste brevi, ma abbatta notevolmente i costi e, quindi, permette la moltiplicazione delle indagini su qualsivoglia argomento di momentaneo interesse.⁶

In definitiva, le indagini campionarie hanno sicuramente colmato molte lacune informative nelle ricerche sulla popolazione e, soprattutto, ne hanno modificato l'impostazione da macro a micro. Tuttavia, una qualche riflessione va fatta al riguardo. In primo luogo c'è da chiedersi se non abbiamo sofferto e ancora soffriamo di una sorta di bulimia nei confronti delle potenzialità offerte dalle interviste dirette, con qualche eccesso nel ricorso allo strumento delle indagini campionarie e con l'abuso della pazienza degli intervistati nell'interrogarli sugli argomenti più disparati, intimi e delicati. Vi sarebbe poi da verificare quanto del patrimonio informativo raccolto a così caro prezzo (in denaro e in tempo) sia stato effettivamente utilizzato, anche perché non reso pubblico, ovvero pubblicato troppo tardi o in forme non utilizzabili da altri ricercatori, così che è più volte capitato che venissero condotte indagini in pratica sovrapponibili negli obiettivi e nei contenuti. Sotto il profilo del metodo e della pratica sussiste poi qualche dubbio circa la generale, rigorosa applicazione delle regole inferenziali, sia nel disegno campionario, sia nella sua realizzazione, sia nell'interpretazione dei risultati, troppo spesso analizzati per aggregati non altrettanto rappresentativi quanto il campione nel suo complesso.

Vorrei infine fare qualche considerazione sugli sviluppi che sembrano avere ora le *interviste in profondità*, tese a interrogare solo pochi soggetti sulle matrici complesse di comportamenti o di atteggiamenti demograficamente rilevanti, oppure sulle loro storie di vita. Si tratta di sviluppi certamente promettenti al fine di cogliere la natura dei meccanismi alla base di scelte e di atti che possono avere riflessi sulle dinamiche della popolazione e che, quindi, sono utili soprattutto nell'indirizzare la ricerca da sviluppare poi attraverso la tradizionale strumentazione statistica propria degli studi demografici: in altri termini, degli strumenti di sondaggio preventivo, così come lo sono i *focus group*, sicuramente capaci di far emergere l'esistenza di relazioni profonde e di condizionamenti non percepiti a priori dal ricercatore, ma in ogni caso da verificare e misurare attraverso una rigorosa indagine quantitativa condotta secondo adeguati canoni di numerosità e di rappresentatività statistica.⁷ I pareri motivati di un piccolo gruppo di persone, per quanto assortito con sagacia, rimangono i pareri di *quelle* persone, e non potranno mai rappresentare la varietà

6. Non pretendo di contestare la validità di metodologie d'indagine sulle quali si è a lungo dibattuto e per le quali sono state elaborate tecniche di correzione ed integrazione. Pur tuttavia, le interviste telefoniche sollevano non poche perplessità sulla rappresentatività dell'insieme dei soggetti di fatto contattabili rispetto all'universo di volta in volta indagato, così come sulla controllabilità dei parametri di stratificazione e sulle selezioni che fatalmente si verificano sia nel reperimento dei soggetti estratti, sia nell'accettazione di rispondere, sia nell'eventuale troncamento delle interviste. Sui problemi di *privacy* nella pratica del CATI e, più in generale, nella selezione di piccoli campioni, v. Cerbara, 2005.

7. «*Une tradition de critique des sources et de rigueur dans la collecte des données aide peut-être les démographes à accepter l'apport de méthodes qualitatives dans les phases préliminaires de la construction des recherches ; l'exigence de réflexivité devrait les conduire plus systématiquement à renforcer les collectes quantitatives par des dispositifs qualitatifs*» [Bozon, 2006 : 459].

dei pareri che potremmo trovare nella popolazione. Allo stesso modo, le matrici mentali e fattuali delle scelte demografiche riportate dal signor X o la biografia narrata dalla signora Y hanno il valore limitato delle loro personali esperienze, né la considerazione congiunta delle esperienze di poche decine di intervistati in profondità (il costo di tali interviste è di diverse volte superiore a quello di un'intervista CATI) ci potrà mai dare il quadro complessivo dei processi decisionali o delle sequenze nella vita del complesso dei componenti di una popolazione.

Ma, a ben guardare, questo *caveat* vale anche per le indagini su base campionaria, disegnate secondo le regole più rigorose. Ad esempio, un campione di donne in età tra 18 e 49 anni, opportunamente stratificato in modo da rappresentare le varie classi di età e le diverse latitudini e tipologie di residenza, ci fornirà una stima corretta (con il suo bravo intervallo di confidenza) della distribuzione per figli avuti nella popolazione di riferimento, e potrà anche dirci come il complesso delle donne di quella fascia d'età si distribuisce probabilisticamente rispetto alle intenzioni di avere un altro figlio. Potremo anche analizzare come su quelle stime intervengano, differenziandole, le diverse variabili introdotte, con l'avvertenza, però, che il disegno campionario non assicura la significatività statistica ai sottoinsiemi costruiti sulle modalità di variabili che non siano parametri di stratificazione. Più in là non dovrebbe essere concesso di spingerci, perché già un'analisi dei risultati per livello di istruzione e per classe di età non gode dei vantaggi della rappresentatività statistica, dal momento che non può essere garantito, ad esempio, che le donne di 30-34 anni che abbiamo intervistato siano distribuite per istruzione come lo sono nell'universo di riferimento. Non è invece infrequente imbattersi in ricerche che dimenticano per strada il disegno campionario di partenza, e procedono in analisi e sotto-analisi affidando la significatività dei risultati ottenuti alla sola dimensione dei sub-aggregati e agli schemi usuali di calcolo degli intervalli di confidenza. Ma non vanno così le cose e più si scende nell'intrico degli incroci più si fa solo della statistica descrittiva del sottogruppo intervistato e meno si è sicuri di rappresentare l'universo di riferimento per via inferenziale. I soggetti, nella loro individualità e specificità, finiscono così con il dominare nei risultati, con il rischio di attribuire a questi ultimi un valore collettivo che essi non hanno né possono avere.

VI.

I modelli: dall'astrazione matematica alla semplificazione della realtà, al rischio di metodologie standard

Per molto tempo non avemmo che un solo modello, quello della *popolazione stabile*, tanto utopico da volerci delle centinaia d'anni d'inedefessa monotonia nei comportamenti demografici per poterci avvicinare. Il risultato finale, almeno nella sua sottospecie *stazionaria*, sarebbe una popolazione sempre uguale a sé stessa, in dimensioni e struttura. Non però nelle persone che la popolano, le quali si porterebbero dietro tutte le loro diversità fisiche, mentali e di atteggiamenti e comportamenti. E similmente, così come continuano a variare gli agenti, continuerà a modificarsi anche il contesto, che sui quei comportamenti e quegli atteggiamenti di certo influisce. Pertanto, è ben difficile immaginare una popolazione stabile se non nella dimensione astratta di un teorema matematico, e ciò nonostante tutti gli sforzi che sono stati fatti per arricchire il modello con le componenti migratorie della dinamica demografica, per aprirlo al suo sviluppo contestuale in più di una popolazione e per ipotizzare percorsi di avvicinamento meno perentori.¹

Sono apparsi successivamente altri modelli o, meglio, delle *curve interpolatrici* che cercano di descrivere spicchi della dinamica demografica da interpretare attraverso apposite formulazioni matematiche, quasi sempre in funzione dell'età dei soggetti e di adeguati parametri da stimare attraverso opportune operazioni statistiche.² Quei parametri

1. Per una sintesi sugli avanzamenti che ha avuto il modello di Lotka a partire dalla sua formulazione del 1939, v. Caselli *et al.*, 2006. Per i modelli multistato, v. Rogers, 1986.

2. Per una magistrale sintesi di tali metodi, se pur datata, v. Keyfitz, 1982, successivamente ampliata ed aggiornata, ma con un taglio più applicativo, fino all'ultima edizione disponibile [Keyfitz *et al.* Caswell, 2005]. V. anche Newell, 1988.

vengono spesso utilizzati come indicatori sintetici di confronto tra popolazioni e, specie quando è possibile dare loro un significato concreto rispetto al fenomeno descritto, si possono anche produrre delle simulazioni tramite la loro regolazione, operando modifiche su specifici caratteri o sezioni del fenomeno stesso. Il limite, ma anche il pregio di quei modelli, sta proprio nella loro astrattezza, nella possibilità di applicarli alle situazioni più diverse, quasi che nella loro formulazione matematica essi costituissero delle leggi immanenti rispetto alla variabilità dei fenomeni demografici.

In una logica del tutto opposta, di tipo deduttivo, si sono invece mossi i *modelli causali* che originano da delle ipotesi circa il funzionamento di un sistema che comprenda variabili di popolazione poste in relazione ad altri insiemi di variabili socioeconomiche: il tentativo è quello di simulare la realtà semplificandola. La loro configurazione è pertanto collegata con il “qui e ora” del fenomeno e del suo contesto,³ mentre la loro generalizzazione ad altre situazioni può avvenire solo per analogia, sotto la vincolante ipotesi di somiglianza delle condizioni e del tipo di relazione. Modelli di questo genere si svilupparono soprattutto negli anni Settanta e Ottanta, ancora una volta in concomitanza con il diffondersi dell’uso del calcolatore elettronico negli studi di popolazione.

Se si guarda ora con spirito critico a quei tentativi è facile imputare loro l’illusoria pretesa di voler ridurre dei sistemi complessi a poche equazioni, il più delle volte di tipo lineare. L’aspetto più discutibile sta però nella scelta delle variabili indipendenti, scelta molto spesso limitata dalla disponibilità di dati accoppiabili con quelli relativi alle variabili demografiche che vengono studiate.⁴ In teoria, la scelta dovrebbe essere dettata da uno schema a priori di relazioni che colleghino razionalmente le variabili in un *rapporto di causa ed effetto*. Ma quelle che poi si riescono a misurare di fatto sono solo delle relazioni statistiche, nelle quali il verso della causazione rimane indefinito. Il che può forse rappresentare anche più correttamente (ma senza riuscire a pesarne le opposte componenti) l’ambivalenza delle relazioni nei sistemi complessi, relazioni che non sono quasi mai unidirezionali, mentre implicano *fenomeni di retroazione* spesso trascurati dalla

3. Per la verità, la determinazione dei loro parametri può avvenire solo tramite la considerazione congiunta di un numero di situazioni non inferiore al numero dei parametri da stimare, così che sono necessarie estensioni in senso spaziale (più aree geografiche o più aggregati demografici), oppure temporale (più momenti di rilevazione). Questo comporta che fenomeno, contesto e relazione diventino medi rispetto alle situazioni puntuali utilizzate per la stima.

Una chiara verifica di quanto le relazioni trovate dipendano dalla situazione spazio-temporale analizzata la si può avere nel caso delle dipendenze reciproche tra livello di fecondità e tasso di attività extradomestica delle donne: negativo un tempo nei Paesi sviluppati, è diventato progressivamente positivo, a riprova delle possibili modifiche temporali delle relazioni e ancor più dell’insufficienza esplicativa di un’ipotesi di schema causale troppo semplice rispetto alla realtà.

4. A quei limiti di disponibilità e a una certa disinvoltura nel correlare variabili “come che sia” si devono alcune “scoperte” assai discutibili del passato, come l’esistenza di forti relazioni statistiche, calcolate su dati nazionali o regionali, tra il livello di fecondità e il numero di apparecchi telefonici o di televisori ogni mille abitanti. In questo caso, al problema dell’*ecological fallacy* si aggiungono tutti i problemi delle correlazioni spurie.

modellistica. Un altro limite delle variabili indipendenti si manifesta quando si vuole utilizzare il modello per ipotizzare realtà proiettate nel futuro: la natura della maggior parte di quelle variabili soffre di una limitata prevedibilità, di certo assai inferiore a quella delle variabili demografiche che attraverso quelle si vorrebbero poter controllare.

In definitiva, l'eleganza formale nelle costruzioni teoriche dei grandi modelli macro di quel periodo si esaurì davanti alle loro difficoltà interpretative e ai loro limiti applicativi in contesti diversi; in altre parole, davanti alla loro insufficienza nel descrivere dei sistemi complessi nei quali intervengono troppe variabili contemporaneamente e dove le relazioni sono spesso di natura circolare, non lineari né unidirezionali.

Eppure quell'esperienza servì anche nella prospettiva più ridotta delle sole variabili demografiche, per le quali fu posta una particolare attenzione al gioco congiunto dei fattori che agiscono nelle tre dimensioni dell'età, del periodo e delle coorti. I *modelli APC* [Hobcraft *et al.*, 1982] affrontano, in una chiave ancora macro, l'intricata questione della simultanea dipendenza dei comportamenti demografici e delle loro variazioni dall'età raggiunta dai *coetanei*, dall'esperienza vissuta dai *coevi* e dalla congiuntura sperimentata dai *contemporanei*. Nell'identificazione del modello si deve fare i conti con il problema della sua iperspecificazione ($P = C + A$!); devono dunque essere adottate delle strategie di stima dei parametri che passano attraverso la fissazione di un ordine arbitrario nelle tre dimensioni, ordine che in genere si cerca di far corrispondere a un prevalente schema interpretativo-causale del fenomeno allo studio.

Con i modelli APC arriviamo al limite della specificazione di dati aggregati: in essi, gli individui sono raccolti contemporaneamente nelle tre dimensioni collegate al tempo, se pure per classi, annuali o pluriennali che siano. Ed è qui che si palesa la forzata approssimazione dei dati collettivi, perché per tutti gli individui della popolazione si ipotizza: dietro l'età una somiglianza di effetti del suo procedere; dietro il periodo una sostanziale uniformità di effetti trasversali; e – quel che forse più contrasta con l'esperienza e la logica – dietro le coorti delle storie individuali sostanzialmente simili, almeno sotto il profilo dei principali fattori che influiranno sui successivi comportamenti demografici. Ma è proprio la diversità delle biografie individuali rispetto a quei fattori a differenziare gli individui di una stessa coorte, sia nel comportamento demografico alle diverse età, sia nel modo d'affrontare e reagire alle congiunture del momento. Anzi, più ancora, l'accumulo delle diversità individuali sui tre piani temporali finisce con il modificare il collettivo d'appartenenza dei soggetti, soprattutto se il fenomeno indagato è di tipo ad eliminazione. *Eterogeneità e selezione* entrano così nell'analisi demografica come ultimo baluardo nei modelli macro, ma introducono già agli effetti della diversità delle esperienze individuali sui comportamenti demografici.⁵

5. L'eterogeneità di una popolazione (di una coorte) non deriva solo dalla diversità delle esperienze cumulate dai soggetti che ne fanno parte nel corso delle loro vite, ma possono essere originate da fattori biologici o genetici che rendono gli individui diversi tra loro rispetto ai "rischi" di sperimentare un determinato evento demografico. Sull'argomento, v. la trattazione di Vaupel *et Yashin*, 2006.

Da qui in poi bisogna affidarsi ai modelli micro, nei quali gli individui mantengono compatto nelle analisi tutto il loro patrimonio informativo indagato, mentre i risultati ne tengono simultaneamente conto, così da poter attribuire a ciascuna variabile, esplicativa o supplementare che sia, il segno e il contributo alla variabilità del fenomeno in esame.

Questo schema di analisi ha ormai invaso il panorama attuale degli studi di popolazione, con gli eventuali adattamenti del modello alla tipologia dei dati disponibili e al tipo di relazione di base ipotizzata. In generale, i risultati si dimostrano d'immediata lettura, corredati come sono dagli asterischi che segnalano la corrispettiva significatività.⁶ Tuttavia, la loro interpretazione non è sempre altrettanto chiara, specie quando il loro segno contrasta con quello atteso nelle ipotesi teoriche. Insomma, si ha l'impressione che un uso talvolta acritico e standardizzato di tali modelli conduca a conclusioni semplicemente confermatrice quando i coefficienti concordano con le attese, mentre si rischia la mancanza di spiegazioni logiche (o s'induce una loro affannosa e discutibile ricerca) non appena vi sia un contrasto tra quanto atteso e quanto "misurato". Si obietterà che, specie nel primo caso, il valore aggiunto di tali modelli sta nel fornire una graduatoria dei fattori o, meglio, delle loro modalità che intervengono a determinare un dato fenomeno: è però almeno evidente che l'oggettività di quella graduatoria dipende dall'insieme delle variabili utilizzate e dalla loro articolazione in valori o modalità.

Se mi è permessa una notazione critica, ho l'impressione che talvolta la tecnica di tali modelli tenda a prendere il sopravvento sull'analisi dei fenomeni e che il ricercatore spesso fatichi a governare la "strambata" che ne consegue. Del resto, l'attuale prevalenza delle tecniche rispetto ai contenuti sta cambiando il modo di affrontare i problemi della ricerca e della conoscenza. Può darsi che ciò possa aprire nuovi orizzonti di metodo e interpretativi, ma rimane il dubbio che la realtà, multiforme nelle sue fondamenta e nel suo divenire, possa essere indagata davvero attraverso un limitato repertorio di ricette. Inoltre, proprio un'eccessiva standardizzazione modellistica potrebbe rischiare di ridurre la specificità di contenuto degli studi di popolazione secondo la prospettiva e le competenze demografiche, con ulteriori perdite di visibilità e di rilevanza disciplinare della materia [Golini, 2000].

6. Alla luce di quanto notato sopra circa i limiti della significatività calcolata su variabili campionarie che non siano parametri di stratificazione, siamo sicuri che basti la numerosità dei sottogruppi a definire un livello di confidenza?

VII.

Spazio e tempo: variabili di contesto o variabili causali?

Il *tempo*, infatti, e lo *spazio* sono parametri essenziali negli studi di popolazione: non solo perché vengono a delimitare l'aggregato demografico allo studio, ma perché agiscono direttamente nella sua dinamica come variabili intervenienti [Gutmann *et al.*, 2011].

Del territorio ho già detto nell'approccio macro per i connessi problemi dell'*ecological fallacy*. Ma anche in un'impostazione micro il territorio di riferimento mantiene tutta la sua importanza, sia come contenitore dei soggetti allo studio (e di quelli eventualmente rappresentati per via campionaria), sia come portatore di alcuni caratteri che potrebbero influire sugli eventi demografici e/o sui comportamenti retrostanti. Sotto il primo profilo, la questione è se il territorio di riferimento sia configurato in modo appropriato rispetto alle caratteristiche dell'aggregato, ai fenomeni demografici allo studio o, più in generale, alla significatività geografica (soprattutto nel senso della geografia umana ed economica) del territorio: troppo spesso i confini amministrativi, definiti per lo più secondo criteri geo-morfologici e/o storici, condizionano i territori d'analisi e, quindi, gli aggregati demografici sui quali si opera, con il rischio da un lato di dovere fare i conti con un'elevata disomogeneità della popolazione allo studio, dall'altro di avere dei risultati riferiti a un territorio di assai scarsa significatività funzionale [Vitali, 1992].¹

1. A partire dagli studi di Ornello Vitali e di Marco Martini, il problema della definizione di aree funzionali è stato affrontato dall'ISTAT, che da tempo fornisce diversi dati anche per i *Sistemi locali del lavoro*, aggregati di comuni contigui che "rappresentano i luoghi della vita quotidiana della popolazione che vi risiede e lavora". Il loro problema è che sarebbe necessario aggiornarli continuamente seguendo il modificarsi delle reti di circolazione, mentre i dati di base per il loro disegno poggiano sulla rilevazione censuaria che avviene ogni dieci anni.

Un territorio che sia significativo dal punto di vista geo-statistico, però, è probabile che mantenga delle relazioni con i fenomeni demografici allo studio, almeno implicitamente. Si tratta allora di scegliere a quale livello far intervenire quelle relazioni nell'analisi esplicativa, se come *variabili di contesto* o come delle effettive *variabili causali*. Ad esempio, la salubrità dell'ambiente potrà essere assunta come variabile di contesto rispetto ai comportamenti riproduttivi, ma rischia di diventare una variabile causale estremamente esplicativa negli studi sulla salute e la mortalità, specie per alcune patologie. I *modelli multilevel*, che combinano informazioni ai due livelli macro e micro, rispondono bene soprattutto alla prima esigenza [Racioppi *et al.*, 1997]. Anche in questo caso, però, si presenta il problema di dover attribuire all'intero territorio in analisi i valori di parametri ambientali che spesso sono rilevati in punti specifici del territorio. In particolare, quando si studiano gli effetti di una specifica fonte d'inquinamento (atmosferaico, elettromagnetico, o altro) un parametro di solito trascurato è la distanza che separa i soggetti indagati da quella fonte. In prospettiva, si può pensare di associare i micro-dati a dei parametri di localizzazione anche molto precisa (tramite Gps, ad esempio) o, ancor meglio, alle rispettive aree di prevalente circolazione: ciò potrebbe permettere migliori deduzioni causali dalle relazioni tra i parametri ambientali e i fenomeni demografici.

Anche del tempo ho già detto e delle sue diverse prospettive negli studi di popolazione.² Qui voglio sollevare il problema se il tempo degli individui sia un tempo congruente con i tempi della demografia, dovendo dare ormai per scontato che una società che corre all'attuale velocità non ha più il tempo di preoccuparsi di ciò che avviene nell'arco delle generazioni, tanto è vero che l'unico fenomeno demografico di cui la politica si occupa attivamente sono le migrazioni e, anche di quelle, lo fa in una prevalente ottica attualistica (e molto spesso elettoralistica), ben poco attenta alle sue determinanti e alle implicazioni demografiche di medio e di lungo periodo. Gli individui, proprio perché vivono le tre dimensioni del tempo della demografia, si trovano schiacciati tra il galoppare dell'attualità e la dimensione distesa del proprio passato e dei programmi del loro futuro: le loro scelte e i loro comportamenti demograficamente rilevanti risentono di entrambe le dimensioni, alle quali si aggiunge l'età, che per tutti procede con la stessa cadenza cronologica, ma per ognuno ha un valore (reale o anche solo percepito) diverso.

Per queste ragioni i singoli individui (oppure gli agenti pluripersonali) reagiscono diversamente a una medesima congiuntura e mantengono o cambiano in modo diverso le loro strategie demografiche al medesimo variare del contesto o delle prospettive. C'è da porsi il problema, allora, di come si possano ricondurre tutte queste diversità oggettive e soggettive a degli studi che devono forzatamente semplificare rispetto agli effetti delle

2. Ho sorvolato, ma almeno in nota li devo ricordare, sui problemi delle *serie storiche*, che sono soprattutto problemi relativi ai dati macro: la loro produzione, infatti, varia nel tempo per definizioni e per modalità di rilevazione e di pubblicazione, cosicché diventa necessario eseguire complesse e talvolta assai artificiose ricostruzioni al fine di rendere confrontabili i dati delle serie. In occasione delle modifiche, sarebbe sempre utile poter disporre per qualche rilevazione dei dati secondo sia i criteri nuovi sia di quelli precedenti, in modo da poter operare dei confronti e basare su di essi la ricostruzione delle serie.

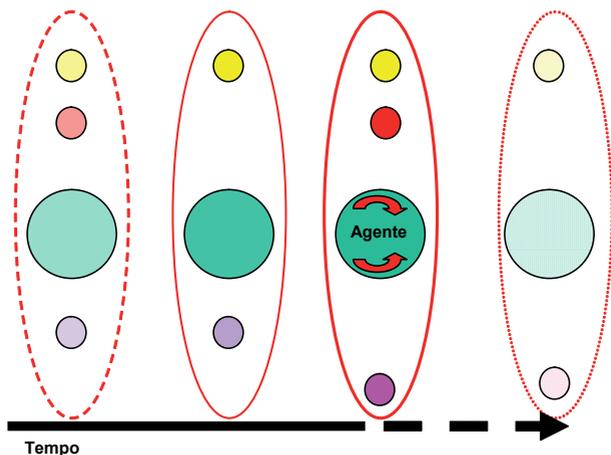
coniunture, ai percorsi delle età e alle storie delle coorti. Ad esempio, bisognerà chiedersi quanto correttamente interpretino i *modelli ad eliminazione* gli effetti dei comportamenti e degli atteggiamenti dei diversi soggetti appartenenti a un aggregato demografico, e se non si rischi così di tornare allo studio di un epifenomeno trascurando di nuovo le sue matrici e le sue componenti e motivazioni individuali.

VIII.

L'ipercubo delle relazioni

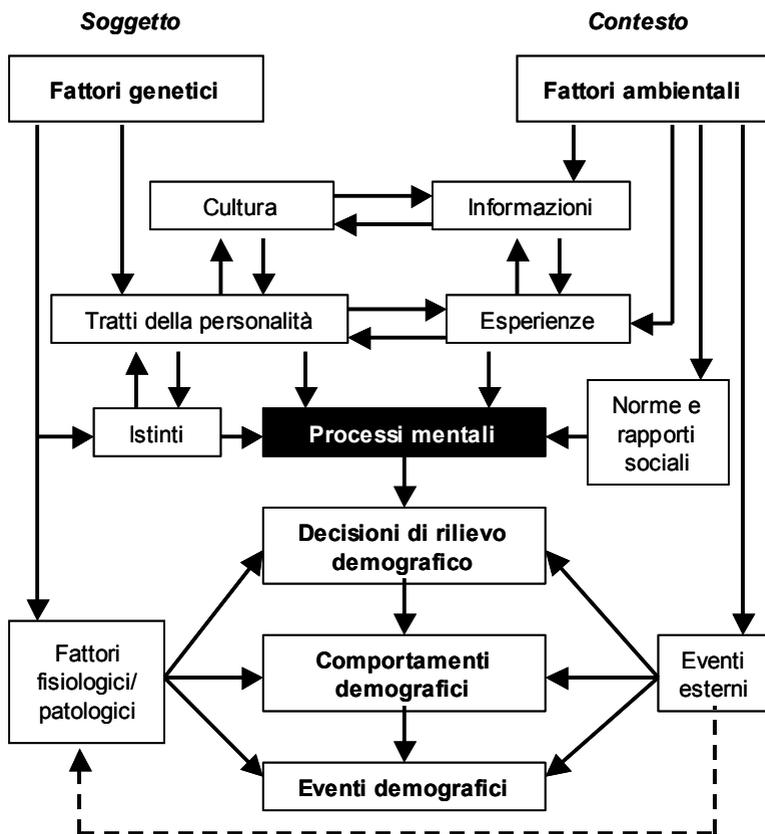
La complessità delle relazioni potenzialmente rilevanti ai fini dei comportamenti demografici avrebbe bisogno di una rappresentazione in uno spazio a molte dimensioni. Lo schema semplificato che ne do in Figura 8 fornisce solo uno sprazzo di quella complessità, che si estende sia nello spazio (da intendere come territorio e ambiente, ma anche come spazio delle relazioni interpersonali o con enti e istituzioni umane), sia nel tempo (soprattutto nel passato, nella storia dell'agente e delle relazioni che questo ha tenuto con gli altri e con l'ambiente; ma anche nel futuro, almeno per le influenze che attese e intenzioni possono avere sulle scelte e i comportamenti dell'oggi).

FIGURA 8 - L'INTRECCIO DELLE RELAZIONI



Evito di esemplificare perché ognuno di noi ha sperimentato come le proprie decisioni circa l'uscire dalla famiglia d'origine, spostare la propria residenza, iniziare un rapporto di coppia o mettersi fine, generare un figlio, sottoporsi a una determinata cura od operazione chirurgica, eccetera, siano dipese dalle relazioni che egli/ella teneva in quel momento con la rete dei referenti e dalle condizioni ambientali nelle quali si trovava a vivere, ma anche con la storia di quelle relazioni e con gli effetti che le esperienze e gli ambienti vissuti nel passato mantenevano su di lui/lei. Su certi argomenti, soprattutto, le attese circa i rapporti futuri e sulle condizioni a venire hanno poi costituito la ragione prima delle decisioni e delle conseguenti azioni di rilevanza demografica [Golini *et* Iacoucci, 2006].

FIGURA 9 – LO SCRIGNO DEI FENOMENI DEMOGRAFICI



Alla complessità di quelle relazioni esterne va aggiunta quella delle relazioni interne all'agente, esplicite quando questo sia pluripersonale, come nel caso di un nucleo familiare, o implicite e di ancor più difficile tracciamento quando si tratti di decisioni o azioni a livello individuale. Anche in questo caso il tempo esercita un ruolo di primo piano, forse anche più evidente nei condizionamenti che provengono dal passato sul

diverso modo dei soggetti di percepire e reagire agli stimoli esterni e, per il futuro, dal loro diverso atteggiamento nei confronti delle prospettive a venire e dalla capacità di programmare il proprio percorso di vita.

In Figura 9 ho provato a disporre l'insieme di tutte queste relazioni in un sistema centrato sull'individuo o, meglio, sui suoi meccanismi mentali che presiedono alle decisioni e alle azioni (o mancate azioni) che abbiano un rilievo demografico. L'apparente contrapposizione tra soggetto e contesto (ossia tra fattori biologici e fattori ambientali, entrambi colti nella più ampia accezione possibile) si riduce da un lato attraverso il circuito della conoscenza (fatta di informazione ed esperienze, come anche dalla capacità di elaborare quegli input attraverso i tratti della personalità e la cultura), dall'altro lato attraverso la mediazione del conflitto tra istinti e norme sociali. L'elaborazione mentale di tutti questi elementi conduce l'agente alle decisioni di rilevanza demografica. Su queste e sui successivi processi che eventualmente portano l'agente all'evento demografico intervengono o possono intervenire da un lato alcuni fattori fisiologici o patologici dell'agente stesso, dall'altro degli eventi esterni che intervengono a modificare il contesto in cui l'agente opera (oltre che, se è il caso, anche le sue condizioni fisiologiche o patologiche).

La complessa realtà di queste relazioni è però nulla rispetto all'incognita che si cela nei processi mentali, nella vera centrale di elaborazione e di decisione rispetto all'insieme del nostro essere, come siamo e come ci siamo andati formando, e degli input esterni, come realmente sono e come noi li percepiamo: quella che Thomas Burch ha definito "scatola nera" del processo decisionale nei comportamenti sociali quando ci richiama all'insufficienza della modellistica usata di solito negli studi demografici per arrivare ai meccanismi e ai processi mentali che fanno muovere una persona da uno stato a un altro,¹ e reclama la necessità di una o più teorie in grado di spiegare quei processi, per poi, *solo poi*, verificarne l'ammissibilità alla luce delle prove empiriche disponibili [Burch, 2003a].

Con Burch [2003b], con l'*agent-based demography* [Billari *et al.*, 2003; Hobcraft, 2006] e con la *casual analysis* [Engelhardt *et al.*, 2009] tornano in auge i modelli su base razionale e ipotetica, piuttosto che empirica e fondata principalmente sui dati statistici e sulle loro relazioni matematico-probabilistiche. Dico tornano, perché dello stesso tipo erano o volevano essere i grandi modelli demo-economici degli anni Ottanta, solo che allora le ipotesi e le teorie retrostanti riguardavano i comportamenti di grossi aggregati e i dati erano tipicamente dati macro; ora, invece, il comportamento analizzato e previsto è quello dei singoli agenti e i dati per la successiva conferma sono raccolti ed elaborati a

1. «*It urges us to dig deeper, below our multistate analyses to get at the mechanisms and mental processes that move a person from a state to another. Clearly, at the individual level, these will be decision mechanisms, including the decision to imitate others or to conform to fad, fashion, and cultural norms [...]. To the extent that longitudinal data sets do not directly measure such mental or psychological variables [...], explanations can only be done by theory. Multivariate statistical models, even dynamic models, will not suffice. Too many relevant factors will remain in the black box, with no sense of their role in the transformation of measured inputs into measured outputs*» [Burch, 2001: 278].

livello micro. In entrambi i modelli, però, la prevalenza è data alla teoria, a uno schema mentale che proponga delle *relazioni logiche* nei “comportamenti” (là macro, qui micro; là solo fattuali, qui articolati nella catena che va dalle intenzioni all’azione) che abbiano un interesse ai fini degli studi di popolazione.

La strada tracciata è certamente più interessante di quella delineata dalla sola analisi di dati statisticamente significativi. Come allora, però, i quesiti sono: se abbiamo la capacità di modellare processi così complessi; se disponiamo di tutto l’input necessario; e se siamo in grado di tradurre le relazioni logiche ipotizzate in equazioni matematiche. Talvolta, la scorciatoia dei modelli empirici sulla sola base statistico-probabilistica può sembrare allora una soluzione preferibile, anche se assai meno soddisfacente sotto il profilo intellettuale. In fondo, tra le capacità intellettuali vanno annoverate, oltre alla razionalità, anche l’intuizione, che certamente presiede alla scelta delle variabili da mettere in relazione nei modelli empirici, e lo spirito di adattamento, che cerca di sopperire là dove i dati necessari siano carenti o approssimativi rispetto alla realtà indagata.

IX.

De profundis?

Dopo il libro dei *Numeri*, lasciatemi aprire, avviandomi alla conclusione, quello delle *Lamentazioni*, dal quale traggio in nota qualche passo che mi è parso più appropriato per piangere la presente condizione degli studi demografici in Italia:¹

- 1) Un censimento che in parecchie aree, probabilmente quelle più interessanti perché immerse nelle realtà più popolose e dinamiche, per le domande più ficcanti si è ridotto a un'indagine al trenta per cento, con tutti i problemi di sostituzione dei non rispondenti, di stima campionaria, d'errore della stima, eccetera, eccetera. Non che i precedenti censimenti universali fossero esenti da errori nell'impostazione, nell'esecuzione o nella diffusione dei dati. È ciò che avviene normalmente nei censimenti di tutto il mondo, perché nessuno è senza peccato e gli statistici sono portati a commetterne più di altri proprio per la natura aleatoria e incerta del loro procedere. Ma da noi ogni volta quelle iatture (ineliminabili, ma se si vuole largamente gestibili, emendabili o quantomeno quantificabili) sono state vissute da un lato come intollerabili scandali, dall'altro come colpe da ammettere in pubblico solo se del caso e solamente in parte.
- 2) Un'anagrafe che ancora stenta a integrarsi nella sua dispersione su 8.098 registri di-

1. «Come sta solitaria la città un tempo ricca di popolo! [1/1]; Tutti i suoi amici l'hanno tradita, le sono divenuti nemici [1/2]; L'avversario ha steso la mano su tutte le sue cose più preziose [1/10]; Osserva, Signore, e considera come sono disprezzata! [1/11]; Siedono a terra in silenzio gli anziani della figlia di Sion, hanno cosperso di cenere il capo [2/10]; La nostra eredità è passata a stranieri, le nostre case a estranei [5/2]; È caduta la corona dalla nostra testa [5/16]».

versi, più o meno affollati, tenuti in modo più o meno accurato e aggiornato in tempi ragionevoli, registri che si presentano più o meno veritieri in relazione all'iscrizione in essi di residenti stranieri (dubbia la cancellazione in caso di ritorno in patria o d'emigrazione verso un altro paese estero) e alla accondiscendenza degli ufficiali d'anagrafe nei confronti degli interessi di singoli cittadini o degli stessi amministratori a far risultare realtà virtuali di comodo collegate con la residenza anagrafica. A questo proposito, che cosa ci si può aspettare dalle conseguenze del federalismo fiscale sui dati delle anagrafi? Ne verrà, forse, un ulteriore stimolo alla diaspora delle residenze e al moltiplicarsi di quelle fittizie in ragione della convenienza nelle imposte applicate e nei servizi resi?

- 3) Dei dati di movimento che si sono impoveriti drasticamente grazie a un'interpretazione restrittiva della legge sul trattamento dei dati personali e a una semplificazione amministrativa che non è basata sull'integrazione delle notizie già presenti nei diversi archivi amministrativi, ma che mira unicamente ad abbattere il "disturbo statistico" del cittadino e a diminuire il carico di lavoro (e quindi i costi) dell'amministrazione. Impoveriti così tanto da dover essere affiancati (ad esempio per le nascite) da indagini *ad hoc*, con i relativi limiti di rappresentatività e gli indubbi costi aggiuntivi.
- 4) Una posizione della demografia nell'Accademia che è sempre più su una linea difensiva, in una ritirata che ha ben poco di strategico e molto più di una sconfitta che passa attraverso forzati apparentamenti e discutibili conglobazioni. E negli enti di ricerca nei quali ci si occupa(va) di studi di popolazione le cose non vanno meglio: scarsa considerazione, emarginazione, ristrutturazioni riduttive dei ruoli, delle funzioni e del personale addetto; in ciò complici le difficoltà di bilancio degli stessi enti (quando si deve tagliare, gli studi demografici non sono certo in testa alle priorità!) e quelle dei potenziali committenti della ricerca applicata (ministeri, enti locali, eccetera).
- 5) Una figura nella società, quella del demografo, che ricorda da vicino quella del Grillo parlante di collodiana memoria.² Per di più, i demografi sono delle Cassandre a lunga scadenza, che vengono giudicati del tutto inutili in logiche politiche ed economiche (per non dire elettoristiche e speculative) di periodo sempre più breve.

Sul versante opposto si potrebbe però osservare una presenza delle problematiche demografiche sugli organi d'informazione più frequente di un tempo; ma il più delle volte episodica, in collegamento con qualche notizia sconcertante di cronaca, ovvero in occasione della pubblicazione di dati di per sé poco o per nulla innovativi. Presenza, peraltro, spesso macchiata da imprecisioni (ad esempio: "età media" per vita media; "tasso di fecondità dell'1,4%" ; e così via) o da plateali errori, fatto salvo quando sono alcuni dei

2. *Le avventure di Pinocchio*, Capitolo IV «... dove si vede come i ragazzi cattivi hanno a noia di sentirsi correggere da chi ne sa più di loro. [...] Io non me ne andrò di qui, – rispose il Grillo – se prima non ti avrò detto una gran verità. [...] – Canta pure, Grillo mio, come ti pare e piace: [...] io, a dirtela in confidenza, di studiare non ne ho punto voglia, e mi diverto più a correre dietro alle farfalle e a salire su per gli alberi a prendere gli uccellini di nido».

nostri maggiori demografi a scrivere gli articoli o a parlare nei programmi radiofonici o televisivi.³

Del resto, l'ampia disponibilità di indicatori e le tecniche demografiche correnti, con i relativi programmi informatici, consentono o, meglio, sembrano consentire l'accesso agli studi di popolazione anche ai non specialisti; così che vi si cimentano volentersamente urbanisti, epidemiologi, sociologi, storici, economisti, geografi, nonché funzionari di enti territoriali e di organizzazioni le più varie per competenze e finalità. Non che si possa pretendere l'esclusiva degli studi demografici e, tanto meno, quelli di popolazione, ma sorge il sospetto che, trattando di fenomeni che rientrano nell'esperienza comune di ognuno, la demografia possa apparire come una materia abordabile e soprattutto interpretabile da chiunque.

La battuta polemica che una certa Marina Turi ha inviato alla Giornata della Statistica – «La statistica è il come e il quanto. Per il perché, se possono, ci pensano i sociologi» [ISTAT, 2010: *tweet* n. 132] – non può valere per gli studi di popolazione, che certamente descrivono il *come* e misurano il *quanto*, ma si sforzano altresì di capire il *perché* [Golini, 2011], magari operando, gli studiosi di popolazione, in collaborazione con altri specialisti, ma mai demandando loro l'intero compito, dal momento che la prospettiva del demografo rimane peculiare e altamente istruttiva, come dimostra, ad esempio, il diffondersi in altre scienze umane dell'approccio longitudinale, quello che segue nel tempo i fenomeni e i loro attori.⁴

3. Certamente importante, per l'attualità dei temi trattati, la serietà della trattazione e la comunicativa degli interventi, è l'esperienza della *Newsletter Neodemos* (<<http://www.neodemos.it/>>). I suoi accessi sono però ancora limitati (in media 150 al giorno nell'ultimo mese, con massimi non superiori a 300), ben lontani da quelli di un sito al quale volutamente si ispira, *Lavoce.info*, che nell'ultimo anno ha avuto più di trecentomila visite; ma (è solo una conferma) l'attenzione alla politica economica è ben più diffusa che quella ai suoi fattori demografici. Del resto, grazie all'impegno di alcuni demografi, anche *Lavoce.info* ha ospitato alcuni buoni articoli su temi di popolazione. Su questi temi, v. l'intervento di Alessandro Rosina al Convegno Intermedio Aisp 2010 "L'evoluzione degli studi di popolazione. Una lettura per generazione", Firenze, 14 ottobre 2010.

4. V., ad es.: Hastings *et* Berry, 1976; Mason, 1985; Willekens, 2001; Glenn, 2005.

X.

Prospettive di sopravvivenza

Al di là di queste miserie nazionali, con le quali però bisogna fare i conti per delineare ipotesi e strategie di sopravvivenza, il quadro complessivo delle prospettive negli studi di popolazione che mi sembra possibile trarre è dunque molto complesso, ma non è privo di interessanti sviluppi.

Ciascun tipo d'approccio presenta limiti e difficoltà, ma anche indubbi vantaggi sotto un profilo o un altro. La grande disponibilità di buoni dati comparabili rimane il punto di forza degli approcci macro, sia nella loro funzione descrittiva della realtà, sia in quella finalizzata a reperire delle "regolarità demografiche" [Federici, 1979: 15], sia nei tentativi di comprovare la validità di modelli deterministici o di simulazione. L'approccio micro, invece, assicura un maggior rispetto dell'eterogeneità insita in qualsiasi aggregato demografico, con tutto ciò che questo comporta in termini di selezione e di diversità rispetto al "rischio" di "subire" un determinato evento demografico. Anche in questo approccio si contrappone (ma fino a che punto?) un uso dei dati a fini investigativi, attraverso una modellistica basata in prevalenza su relazioni di tipo statistico, ad un uso probativo di modelli teorici, disegnati a priori.

I limiti, nel caso dell'approccio macro stanno proprio nella approssimazione del dato medio, per di più spesso calcolato su aggregati di popolazione la cui "significatività" rimane discutibile; in un approccio micro potrebbero venire da una scarsa rappresentatività del campione di soggetti sui quali quasi obbligatoriamente si basa una analisi di tipo micro.¹ Torna immediata l'analogia con un sistema ottico che, tanto più ingrandisce i particolari, tanto meno è in grado

1. In realtà, i file individuali dei censimenti potrebbero costituire un interessante base per degli studi micro di tipo "universale", fatta salva una copertura bilanciata di tutte le componenti della popolazione e un'articolazione del questionario capace di fornire le variabili dipendenti, indipendenti e di contesto necessarie allo studio programmato.

di cogliere l'insieme, mentre, all'inverso, la visione d'assieme si fa confusa e si perdono quei particolari che di fatto compongono la realtà.

La contrapposizione macro/micro, peraltro, potrebbe non essere così netta: nella teoria (o nella "pratica alta") con lo sviluppo dei modelli di previsione *MicMac*, il cui obiettivo è sviluppare una metodologia che faccia da ponte tra la proiezione aggregata delle coorti e la proiezione dei percorsi di vita dei singoli membri di una coorte;² e nella pratica di ogni giorno, con lo sfruttamento del vasto patrimonio informativo disperso tra le decine di archivi amministrativi che contengono dati individuali della più diversa natura.

Forse stanno proprio in questa apparente dicotomia le linee di sviluppo futuro per gli studi di popolazione: da un lato le indagini approfondite, esplorative, sui meccanismi interni ed esterni che guidano gli agenti nei loro comportamenti demograficamente rilevanti; dall'altro la traduzione, per il possibile, di quelle intuizioni e ipotesi in riprove e impieghi su larga scala, essendo questa la sola che permetta di passare dall'astrattezza della ricerca teorica alla concretezza di interventi di politica economica, sociale e della popolazione.

Sui contenuti e sulle prospettive degli studi di popolazione i demografi italiani si sono spesso interrogati in specifici convegni e volumi;³ da ultimo nel Convegno intermedio dell'AIISP di Firenze, circa un anno fa, nel quale le diverse generazioni di studiosi hanno ripercorso la loro esperienza e hanno provato a gettare la sfida sui possibili sviluppi futuri. Nella tavola rotonda finale ("La demografia che vorrei") è emersa la necessità di recuperare nella ricerca la dimensione macro per poter puntare su temi rilevanti e di facile comunicazione, tenendo però elevata la sensibilità alla qualità del dato.⁴

A me paiono, questi, dei suggerimenti importanti che, tra l'altro, vanno incontro all'attuale e, con grande probabilità, prospettiva carenza di fondi per gli studi di popolazione e, più in generale, per la ricerca sociale. Il quadro macro, inteso come generalità della informazione, potrebbe essere recuperato con una certa facilità attraverso la integrazione delle banche di dati individuali già disponibili, mantenendo pertanto la possibilità di analisi micro, ma estendendo l'ampiezza degli studi a seconda delle esigenze territoriali [Egidi *et* Giovannini, 1999].

Il quadro delle banche dati a contenuto demografico, già ampio all'inizio degli anni Novanta [Golini *et al*, 1991], copre ora diversi aspetti rilevanti ai fini delle ricerche sulla popolazione, ma soprattutto la tecnica informatica con il *middleware* e quella statistica con il *rekord linkage* permetterebbero ora dei collegamenti tra gli archivi che un tempo – per la verità, molto tempo fa – potevano essere solo ipotizzati da una mente fervida e proiettata verso il futuro come quella di Bruno De Finetti (Figura 10). Una decina d'anni fa, presiedendo una sessione in un

2. V. <www.micmac-projections.org>.

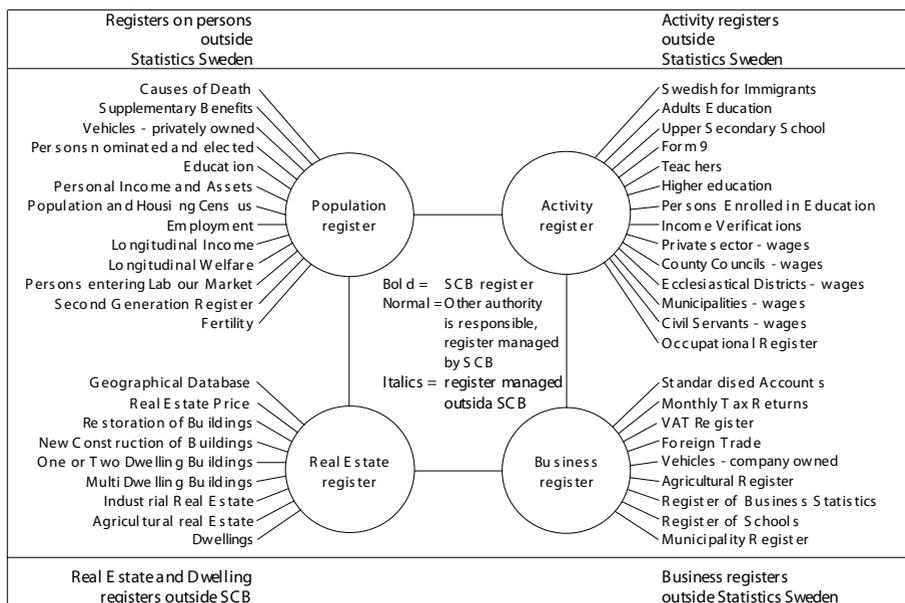
3. Tra i convegni, va ricordato quello di Padova, nel 2000, i cui atti furono pubblicati a cura di Paolo De Sandre e Fausta Ongaro; tra i volumi di riflessione ricordo almeno Billari *et al*, 1999.

4. V. il resoconto della tavola rotonda a cura di Alessandra De Rose in <<http://www.sis-aisp.it/about/15/textpattern-4.0.7/about15iniziative-romosse>>

Sento già le obiezioni a uno schema così ambizioso e omnicomprensivo: da quelle di principio (Ma cinquant'anni fa il concetto di *privacy*, nella particolare accezione di protezione dei dati personali, doveva ancora nascere [Mayer-Schönberger, 1997]); a quelle letterarie (Ma a quei tempi, *1984* di George Orwell veniva letto soprattutto in chiave antitotalitaria); a quelle tecnologiche (Ma i metodi di archiviazione sono spesso diversi e si parlano con difficoltà); alle preoccupazioni circa la qualità degli archivi amministrativi ai fini della informazione statistica [Aimetti *et al*, 2004; Istat, 2009: *passim*]. È quest'ultima la sola obiezione che considero rilevante, ma che potrebbe, dovrebbe, essere ridotta da un insieme di norme, basate più sulle coerenze dei fatti che sui controlli, finalizzate a far combaciare le risultanze documentali con la realtà dei fatti.

FIGURA 11 – IL SISTEMA DEI REGISTRI IN SVEZIA

Chart 7. The register system of Statistics Sweden's and other statistical authorities



FONTE: STATISTIKA CENTRALBYRÅN – STATISTICS SWEDEN, 2001.

Eppure, la conferma che quello schema non fosse solo un sogno irrealizzabile la si ha andando a vedere come funzionano i sistemi dei registri nei Paesi nordici, ad esempio in Svezia (Figura 11), che sono Paesi tra i più garantisti in tema di *privacy* personale, ma dove non si confonde la riservatezza sui dati personali con il diritto al segreto istruttorio e dove l'uso statistico di quei dati – cioè anonimo e collettivo, cioè macro, ma soprattutto finalizzato alla conoscenza della società e all'intervento di apposite politiche su di essa - non viene mai messo in discussione [Johannesson, 2002].

Non è (ancora?) così nel nostro Paese, per una serie di timori, di gelosie, di incapacità operative, e forse anche di mancanza di coraggio nell'adottare sistemi innovativi. Di recente, ad esempio, in una ricerca FAR ampiamente finanziata dal MIUR, abbiamo svilup-

pato un sistema *middleware* finalizzato alla gestione integrata delle basi dati, dei modelli e dei processi per il governo locale.⁵ Come inizialmente previsto, la ricerca ha prodotto il *concept* ed alcuni moduli applicativi in diversi campi, con largo uso di variabili demografiche e sociali. Si tratterebbe ora di passare alla realizzazione del prototipo e, poi, alla commercializzazione del prodotto, ma per avviare queste fasi occorre poter contare su un mercato potenziale costituito soprattutto dagli enti locali, i quali, già restii a investimenti che siano di scarsa visibilità pubblica, trovano ora le loro risorse fortemente decurtate, da riservare necessariamente ad altre priorità di più immediato ritorno.

In conclusione, sono convinto che le ricerche sulla popolazione debbano avvalersi parimenti di analisi macro e di approfondimenti micro: le prime, soprattutto in una prospettiva pragmatica e finalizzata all'intervento, ma non per questo meno valide sotto il profilo scientifico; i secondi, soprattutto in funzione esplorativa e di verifica di modelli concettuali. In entrambi gli approcci è però essenziale che venga mantenuta il più possibile l'unitarietà dell'informazione sulle unità statistiche indagate, il che non contrasta affatto con un approccio macro che si sforzi di tenere conto e gestire l'eterogeneità in modo non solo euristicamente corretto, ma anche finalizzato al rispetto del quadro delle diversità che la realtà presenta.

5. Il progetto "Scenario", da me coordinato, è stato svolto tra il 2006 e il 2009 con il contributo essenziale del gruppo di lavoro di Fernando Ferri e Patrizia Grifoni della sede di Roma del CNR-IRPPS, del gruppo di lavoro guidato da Giuseppe Ponzini nella sede di Penta di Fisciano, e da tre società esterne che operano nel campo dell'informatica.

Riferimenti bibliografici

- AIMETTI Pietro, MARTINI Marco, ZAVANELLA Biancamaria (2004). *Qualità degli archivi amministrativi e qualità dell'informazione statistica*. Milano, Franco Angeli.
- ARZBERGER Peter, SCHROEDER P., BEAULIEU Anne, BOWKER Geoffrey C., CASEY Kipping, LAAKSONEN Leif, MOORMAN David, UHLIR Paul, WOUTERS Paul (2004). Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal*, 3: 135-52.
- ASCOLANI Augusto, BIRINDELLI Anna Maria, GESANO Giuseppe (1974). L'integrazione degli immigrati in ambiente urbano: risultati di un'indagine a Roma e a Torino. Roma, CISP.
- AVRAMOV Dragana, CLIQUET Robert (2005). *Integrated Policies on Gender Relations, Ageing and Migration in Europe: Lessons from the Network for Integrated European Population Studies* (NIEPS). Antwerp, Garant.
- BENINI Rodolfo (1901). *Principii di demografia*. Palermo, Barbera.
- BILLARI Francesco C. (1998). *Appunti di demografia sociale*. Dipartimento di Scienze Statistiche, Università di Padova, *mimeo*.
- BILLARI Francesco C., BONAGUIDI Alberto, ROSINA Alessandro, SALVINI Silvana, SANTINI Antonio (a cura di) (1999). *Quadri concettuali per la ricerca in demografia: quattro saggi*. Serie Ricerche teoriche. Firenze, Dipartimento di Statistica "G. Parenti".

- BILLARI Francesco C., ONGARO Fausta, PRSKAWETZ Alexia (2003). Introduction: Agent-Based Computational Demography. In F.C. Billari, A. Prskawetz (eds), *Agent-Based Computational Demography*. Physica-Verlag: 1-17.
- BLANGIARDO Gian Carlo (1996). Il campionamento per centri o ambienti di aggregazione nelle indagini sulla presenza straniera. In *Atti in onore di Giampiero Landenna*. Milano, Giuffr : 13-30.
- BLANGIARDO Gian Carlo, TANTURRI Maria Letizia (2006). La presenza straniera in Italia. In G.C. Blangiardo, P. Farina (a cura di), *Il Mezzogiorno dopo la grande regolarizzazione. Volume 3. Immagini e problematiche dell'immigrazione*. Milano, Franco Angeli.
- BOZON Michel (2006). L'apport des m thodes qualitatives en d mographie. In G. Caselli, J. Vallin, G. Wunsch (eds), *D mographie : analyse et synth se – VIII Observation, m thodes auxiliaires, enseignement et recherche*. Paris,  dition de l'INED : 439-63.
- BRYAN Thomas (2004). Population Estimates. In J.S. Siegel, D.A. Swanson (eds.), *cit.*: 523-60.
- BURCH, Thomas K. (2001). Longitudinal Research in Social Science: Some Theoretical Challenges. *Canadian Studies in Population*, Vol. 28 (2): 263-83.
- BURCH, Thomas K. (2003a). Data, Models, Theory and Reality: The Structure of Demographic Knowledge. In F.C. Billari, A. Prskawetz (eds), *Agent-Based Computational Demography*. Physica-Verlag: 19-40.
- BURCH, Thomas K. (2003b). Demography in a new key: A theory of population theory. *Demographic Research*, Vol. 9/11: 263-84.
- CARIANI Giovanni (1985). La presenza straniera in Italia: considerazioni sui dati desunti dal censimento demografico del 1981. In *Atti della VI Conferenza italiana di Scienze regionali. Genova, 23-25 ottobre 1985*.
- CASELLI Graziella, VALLIN Jacques, WUNSH Guillaume (eds) (2006). *Demography: Analysis and Synthesis – A Treatise in Population*. San Diego, Elsevier Academic Press
- CASELLI Graziella, VALLIN Jacques, WUNSH Guillaume (2006). Population Models. In G. Caselli, J. Vallin, G. Wunsch (eds), *cit.*: 249-68.
- CASTAGNARO Cinzia, GUARNERI Antonella, PRATI Sabrina, RINESI Francesca (2011). *Building a longitudinal database multi source: new challenges and opportunities*. Convegno intermedio SIS 2011: <<http://www.sis-statistica.it/meetings/index.php/sis2011/2011/paper/view/1328>>.

- CERBARA Loredana (2005). Problemi di privacy nella selezione di piccoli campioni. *Statistica*, LXV (4): 395-405.
- CISIS (2004). *La rilevazione dei dati del Certificato di assistenza al parto: stato di attuazione ed esperienze al confronto*. Atti dell'incontro tecnico Ministero della Salute – Istat – Regioni, Roma, 29 ottobre 2003. Roma, CISIS.
- COMITATO ITALIANO PER LO STUDIO DELLA DEMOGRAFIA STORICA (1974). *Le fonti della demografia storica in Italia: atti del Seminario di demografia storica, 1971-1972*. Roma, CISP.
- CORSINI Carlo Alberto (1997). Stratificazione e demografia differenziale nei processi di mobilità sociale. In SIDES, *Disuguaglianze: stratificazione e mobilità sociale nelle popolazioni italiane (dal sec. XIV agli inizi del secolo XX)*. Collana Società Italiana di demografia Storica. Bologna, CLUEB: 279-83.
- CORSINI Carlo Alberto, SONNINO Eugenio (1972). The Cisp Survey on the Families of Italian Emigrants Abroad. In M. Livi Bacci (ed), *The demographic and social pattern of emigration from the southern European countries*. Firenze, Università – Dipartimento statistico-matematico.
- CORTESE Antonio (1987). La presenza straniera in Italia al 1981: considerazioni sulla qualità dei risultati del censimento demografico. *Sviluppo*, 51-52: ??
- COSTA Marta (a cura di) (2007). *Metamorfosi del matrimonio e altre forme di convivenza affettiva*. Bologna, Libreria Bonomo Editrice.
- COURGEAU Daniel, LELIÈVRE Éva (1992). *Event History Analysis in Demography*. New York, Oxford University Press.
- COX Philip C. (1968). Some Observations on the Teaching of Statistical Consulting. *Biometrics* 24 (4): 789-801.
- CURTIS Bruce (2002). Foucault on Governmentality and Population: The Impossible Discovery. *Canadian Journal of Sociology* 27 (4): 505-533.
- DE FINETTI Bruno (1965). Sull'opportunità di perfezionamenti e di estensione di funzioni dei servizi anagrafici. In AA.VV., *Problemi di rilevazione e classificazione di dati demografici*. Roma, Istituto di demografia, Università di Roma "La Sapienza".
- DE ROSE Alessandra, RACIOPPI Filomena (1999). *Possibilità di analisi comparative INF-1/INF-2: questioni di dati, metodi e ipotesi interpretative*. Relazione provvisoria presentata alle Giornate di Studio sulla Popolazione, III edizione, Firenze, 7-9 gennaio <http://www.ds.unifi.it/gsp99/relazfin/de_rose/de_rose.htm>

- DE SANDRE Paolo (a cura di) (1982). *Indagine sulla fecondità in Italia, 1979. Rapporto generale*. Padova, istituto di statistica dell'Università; Firenze, Dipartimento statistico dell'Università; Roma, Istituto di demografia dell'Università.
- DE SANDRE Paolo, PINNELLI Antonella, SANTINI Antonio (a cura di) (1999). *Nuzialità e fecondità in trasformazione: percorsi e fattori del cambiamento*. Bologna, Il Mulino.
- DE SANDRE Paolo, ONGARO Fausta (a cura di) (2000). *Demografia: presente e futuro*. Padova, Cleup.
- DOYLE Arthur Conan (1892). *The Boys' Sherlock Holmes*. New York, Harper & Brothers.
- EASTERLIN Richard A. (1987). *Birth and Fortune: The Impact of Numbers on Personal Welfare*. 2nd edition. Chicago: University of Chicago Press.
- EATON, Allen Hendershott (1930). *A bibliography of social surveys*. New York, Russel Sage Foundation.
- EGIDI Viviana, GIOVANNINI Enrico (1999). *Il sistema informativo per l'analisi locale*. Iris, Incontri pratesi sullo sviluppo locale, 13-17 settembre.
- ENGELHARDT Henriette, KOHLER Hans-Peter, PRSKAWETZ Alexia (2009). *Causal Analysis in Population Studies: Concepts, Methods, Applications*. Springer.
- EUROPEAN COMMISSION (2005). *Network for integrated European population studies. Final report*. Issued in May 2003.
- FEDERICI Nora (1973). Introduzione. In C. Bielli, A. Pinnelli, A. Russo, *Fecondità e lavoro della donna: un'indagine in quattro zone tipiche italiane*. Roma, Istituto di Demografia.
- FEDERICI Nora (1979). *Istituzioni di Demografia*. Roma, Elia.
- FOCAULT Michel (2004). *Sécurité, territoire, population – Cours du Collège de France 1977-78*. Paris, Seuil/Gallimard.
- FONDAZIONE ISMU (2011). *Sedicesimo Rapporto sulle migrazioni 2010*. Milano, Franco Angeli.
- FORD Thomas R., DE JONG Gordon F. (1970). *Social Demography*. Englewood Cliffs (New Jersey), Prentice Hall.
- FRANCHET Yves (1991). International Comparability of Statistics: Background, Harmonization Principles and Present Issues. *Journal of the Royal Statistical Society. Series A*, 154(1): 19-22.

- FREEDMAN Ronald, WHELPTON Pascal Kidder, CAMPBELL Arthur A. (1973). *1955 growth of American families study documentation*. University of Michigan. Survey Research Center.
- GALLO Margherita (aa. 2006/07). *Fecondità delle coorti e ordine di nascita. Un aggiornamento con dati di indagine*. Tesi di laurea specialistica in Scienze Statistiche, Demografiche e Sociali dell'Università degli Studi di Padova (relatore prof. Maria Castiglioni).
- GESANO Giuseppe (1983). Problemi e caratteri particolari dell'immigrazione straniera in Italia. *Studi Emigrazione – Etudes Migrations*, XX, 71: 416-24.
- GESANO Giuseppe (1986). Il questionario: problemi, criteri, struttura». *Studi Emigrazione – Etudes Migrations*, XXIII, 82-83: 322-42.
- GESANO Giuseppe (2010). *The Sydsverige National Area (Sweden) 'Euro standard' in a cross-border region*. DEMIFER Case Studies, Deliverable 12/10: <http://www.espon.eu/export/sites/default/Documents/Projects/AppliedResearch/DEMIFER/FinalReport/DEMIFER_Deliverable12_10_SE22Sydsverige.pdf>
- GESANO Giuseppe, HEINS Frank, PAGANELLI Fernando (1993). *Differenze anagrafocensimento: verifica di alcune motivazioni politico-amministrative*. Relazione presentata alle Giornate di studio sulla popolazione, Bologna, 6-7 dicembre.
- GIOIA Melchiorre (1829). *Filosofia della statistica*. Milano, Annali universali delle scienze e dell'industria.
- GIORGI Piero, VIOLA Giulia (2003). Il completamento della fecondità delle coorti: un approccio basato sulla modellistica ARIMA. *Quaderni di Statistica*, 5: 189-241.
- GLENN Norval D. (2005). *Cohort Analysis*. Second Edition. London, Sage Publications.
- GOLINI Antonio (1991). Le statistiche demografiche e sociali in Italia nella prospettiva del 1993: il quadro generale. In *La statistica italiana per l'Europa del 1993*, Atti del Convegno di studi organizzato dalla Facoltà di Scienze statistiche dell'Università "La Sapienza", 21-23 maggio 1990, Roma, Edizioni L'Ed.
- GOLINI Antonio (2000). Quale formazione per il demografo del 2000?. In P. De Sandre, F. Ongaro (a cura di), *Demografia: presente e futuro*. Padova, CLEUP Ed.: 95-120.

- GOLINI Antonio (2011). Statistica e demografia. In Istituto dell'Enciclopedia Italiana, *L'unificazione italiana*, volume pubblicato con il contributo di Aspen Institute Italia <<http://www.aspeninstitute.it/interesse-nazionale/altre-iniziativewww.aspeninstitute.it/interesse-nazionale/altre-iniziative>> Pdf: L'unificazione italiana - sezione II.
- GOLINI Antonio, BONIFAZI Corrado, MIGLIORINI Enzo (1991). *Le banche dati demografiche in Italia*. WP03/91. Roma, IRP-Istituto di Ricerche sulla Popolazione del CNR.
- GOLINI Antonio, IACOUCCI Raffaella (2006). Demographic Trends and Relationships between Generations. In G. Caselli, J. Vallin, G. Wunsch (eds), *Demography: Analysis and Synthesis*, Vol. III, Elsevier and Academic Press, 2006: 305-25.
- GOULD Robert (2004). Variability: One Statistician's View. *Statistics Education Research Journal*, 3 (2): 7-16 <[http://www.stat.auckland.ac.nz/~iase/serj/SERJ3\(2\)_Gould.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ3(2)_Gould.pdf)>
- GUTMANN Myron P., DEANE Glenn D., MERCHANT Emily R., SYLVESTER Kennet M. (2011). *Navigating Time and Space in Population Studies*. Springer.
- HASTINGS Donald W., BERRY Linda G. (1979). *Cohort analysis: a collection of interdisciplinary readings*. Scripps Foundation for Research in Population Problems.
- HEISENBERG Werner Karl (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43: 172-98.
- HENRY Luis (1959). D'un problème fondamental de l'analyse démographique. *Population*, 14 (1): 9-32
- HIMMA Kenneth Einar (2007). The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition. *Ethics and Information Technology*, 9: 259-72.
- HOBcraft John (2006). The ABC of demographic behaviour: How the interplays of alleles, brains, and contexts over the life course should shape research aimed at understanding population processes. *Population Studies*, Vol. 60 (2): 153-87.
- HOBcraft John, MENKEN Jane, PRESTON Samuel (1982). Age, Period, and Cohort Effects in Demography: A Review. *Population Index*, 48 (1): 4-43.
- ISTAT (2006a). *Matrimoni, separazioni e divorzi – Anno 2003*. Annuario, 16 – 2006. Roma, ISTAT.
- ISTAT (2006b). *L'indagine campionaria sulle nascite: obiettivi, metodologia e organizzazione*. A cura di M. Lo Conte e S. Prati. Metodi e Norme, 28. Roma, ISTAT.

- ISTAT (2006c). *Il sistema di indagini sociali multiscopo. Contenuti e metodologia delle indagini*. A cura di G. Bagatta. Metodi e norme, 31. Roma, ISTAT.
- ISTAT (2006d). *Avere un figlio in Italia – Approfondimenti tematici dall'indagine campionaria sulle nascite – Anno 2002*. A cura di M. Lo Conte e S. Prati. Collana Informazioni, 32. Roma, ISTAT.
- ISTAT (2007). *Nascere in Italia – Anno 2005*. Statistiche in breve (17 gennaio). Roma, ISTAT.
- ISTAT (2008). *L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc)*. A cura di C. Ceccarelli, M. di Marco e C. Rimaldelli. Metodi e norme, 37. Roma, ISTAT.
- ISTAT (2009). *Navigando tra le fonti demografiche e sociali*. A cura di V. Egidi e A. Ferruzza. Roma, ISTAT.
- ISTAT (2010). *La Statistica è la Statistica – Giornata mondiale della statistica*. <http://www.istat.it/istat/eventi/2010/giornata_mondiale_statistica/volume_Twitter.pdf>
- JARGOWSKY Paul A. (2005). Ecological Fallacy. In K.L. Kimberly (ed. in chief), *Encyclopedia of Social Measurement*. Oxford, Elsevier: I, 715-22.
- JOHANNESSON I, (2002). *Total Population Register of Statistics Sweden: New Possibilities and Better Quality*. Örebro, Statistics Sweden.
- KELLY John J. (1987). Improving the Comparability of International Migration Statistics: Contributions by the Conference of European Statisticians from 1971 to Date. *International Migration Review*, 21, 4: 1017-37.
- KEYFITZ Nathan (1982). Mathematical Demography. In J. Ross (ed.), *International Encyclopaedia of Population*. New York, Free Press: 437-43.
- KEYFITZ Nathan, CASWELL Hal (2005). *Applied Mathematical Demography*. Third edition. New York, Springer Science+Business Media.
- LANGHEINRICH Marc (2001). Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems. *Lecture Notes in Computer Science*, 2201/2001: 273-91.
- LESTHAEGHE Ron J., VAN DE KAA Dirk J. (1986). Twee Demografische Transitie's? In R. Lesthaeghe et D.J. van de Kaa (eds.), *Bevolking: Groei en Krimp*. Deventer, Van Loghum Slaterus: 9-24.

- LIVI BACCI Massimo (1981). *Introduzione alla demografia*. Torino, Loescher.
- LIVI BACCI Massimo (2000). Confini dei fenomeni e confini disciplinari: una divagazione. In P. De Sandre, F. Ongaro (a cura di), *cit.*
- LOMBARDO ENZO (1994). Evoluzione diacronica della demografia. In M. Livi Bacci, G.C. Blangiardo, A. Golini (a cura di). *Demografia*. Guida agli studi di scienze sociali in Italia. Torino, Edizioni della Fondazione Giovanni Agnelli <<http://www.fga.it/uploads/media/Demografia.pdf>>: 28-62.
- MACIOTI Maria Immacolata, PUGLIESE Enrico (1991). *Gli immigrati in Italia*. Bari, Laterza.
- MASON William M. (1985). *Cohort analysis in social research: beyond the identification problem*. Springer-Verlag.
- MAYER-SCHÖNBERGER Viktor (1997). Generational Development of Data Protection in Europe. In P.E. Agre, M. Rotenberg (eds), *Technology and Privacy: The New Landscape*. Cambridge, Massachusetts, The MIT Press.
- MENNITI Adele (2005). *I comportamenti riproduttivi: atteggiamenti, intenzioni e scelte delle donne italiane – I risultati della quinta e sesta indagine dell'Osservatorio sulle aspettative di fecondità*. WP 4/2005. Roma, CNR-IRPPS.
- MILBANK MEMORIAL FUND (1949). Problems in the Collection and Comparability of International Statistics. New York, Milbank Memorial Fund.
- MINISTRO D'AGRICOLTURA, INDUSTRIA E COMMERCIO (a cura di) (1864). *Popolazione – Censimento Generale. Volume primo*. Torino, Tipografia Letteraria.
- NEWELL Colin (1988). *Methods and Models in Demography*. London, Belhaven Press.
- NOTESTEIN Frank W. (1945). Population – The Long View. In Theodore W. Schultz (ed.), *Food for the World*. Chicago, University of Chicago Press: 40-57.
- PALOMBA Rossella (a cura di) (1987). *Vita di coppia e figli. Le opinioni degli italiani degli italiani degli anni Ottanta*. Firenze, La Nuova Italia.
- PALOMBA Rossella (a cura di) (1991). *Crescita zero. Le opinioni degli italiani in un'indagine dell'Istituto di Ricerche sulla Popolazione*. Firenze, La Nuova Italia.
- POON VHK (1996). The Effects of Immigration on Family Health for Hong Kong Chinese Emigrating to North America. Discussion Paper. *Hong Kong Practitioner* 18 (12): 647-54 <<http://sunzi.lib.hku.hk/hkjo/view/23/2301886.pdf>>
- POPOF Carol, JUDSON Dean H. (2004). Some Methods of Estimation for Statistically

- Underserved Areas. In J.S. Siegel, D.A. Swanson (eds.), *cit.*: 603-41.
- PINNELLI Antonella (1967-68). La fecondità differenziale: rassegna delle indagini. *Genus*, XXIII (3-4): 247-305 e XXIV (1-4): 215-48.
- PUGLIESE Enrico (2002). *L'Italia tra migrazioni internazionali e migrazioni interne*. Bologna, Il Mulino.
- RACIOPPI Filomena, RAMPICHINI Carla, ZACCARIN Susanna (1997). L'integrazione delle dimensioni micro e macro negli studi di popolazione: alcuni aspetti dell'approccio di ricerca multilevel. In P. Giorgi, S. Strozza (a cura di), *Secondo Convegno dei Giovani Studiosi dei Problemi di Popolazione*. Studi di Popolazione. Temi di Ricerca Nuova. Roma, Università di Roma: 529-52.
- RIZZI Alfredo (2002). Inferenza statistica: scimmiette e cattivi maestri – Un titolo per sorridere. *Statistica & Società*, 1(1): 6-14.
- ROGERS Andrei (1986). Parameterized Multistate Population Dynamics and Projections. *Journal of the American Statistical Association*, 81 (393): 48-61.
- RONDINI Livia Linda (2003). *La storia della statistica pubblica in Italia*. Milano, Franco Angeli.
- SALVINI Silvana, SANTINI Antonio (1999). Dalle biografie alle coorti, dalle coorti alle biografie. In F. Billari *et al.*, *cit.*: 40-55.
- SIEGEL Jacob s., SWANSON David A. (eds.) (2004). *The Methods and Materials of Demography*. Second Edition. San Diego, Elsevier Academic Press.
- SABBADINI Linda Laura, SGRITTA Giovanni Battista (1996). Significati e tendenze delle statistiche sociali ufficiali: un cambio di paradigma. In *Atti della XXXVIII Riunione Scientifica della Società Italiana di Statistica – Rimini, 9-13 aprile*. Rimini, Maggioli Editore.
- SONNINO Eugenio (1979). *Aspetti e problemi di demografia sociale e di politica della popolazione in Italia*. Facoltà di Scienze Statistiche, Demografiche ed Attuariali, Istituto di Demografia, Università degli Studi di Roma “La Sapienza”.
- STATISTISKA CENTRALBYRÅN – STATISTICS SWEDEN (2001). The future developments of the Swedish register system. Final report of the Register Project. R&D Report 2001/1.
- ŠTEINBUKA Inna (2009). How to improve social surveys to provide better statistics on migrants. 95th DGINS Conference “Migration - Statistical Mainstreaming” 1st October 2009, Malta: <http://epp.eurostat.ec.europa.eu/portal/page/portal/conferences/documents/95th_dgins_conference/Improving_survey_dataI_

- SUNDBÄRG Gustav (1906). *Aperçus statistiques internationaux*. Stockholm.
- SUNDT Eilert (1855/1980). *On Marriage in Norway* (translated and introduced by Michael Drake). Cambridge, Cambridge University Press.
- SÜSSMILCH Johann Peter (1741). *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, Tod, und Fortpflanzung desselben erwiesen*. Berlin, I.C. Spener.
- VAN DE KAA Dirk J. (1987). Europe's Second Demographic Transition. *Population Bulletin* 42 (1): 1-59.
- VAUPEL James V., YASHIN Anatoli I. (2006). Unobserved Population Heterogeneity. In G. Caselli, J. Vallin, G. Wunsch (eds), *cit.*: Vol. 1, 271-78.
- VITALI Ornello (a cura di) (1992). *Cambiamenti demografici e urbanizzazione*. Napoli, Edizioni Scientifiche Italiane.
- WILLEKENS Frans (2001). Theoretical and Technical Orientation Toward Longitudinal Research in the Social Science. *Canadian Studies in Population*, Vol. 28 (2): 189-217.
- WOJTYŁA Karol (2002). *L'uomo nel campo della responsabilità*. Milano. A cura di A. Delogu. Milano, Bompiani.

Nell'ultimo mezzo secolo, gli studi di popolazione hanno conosciuto un'evoluzione profonda per ampiezza del campo d'interesse e per livello d'approfondimento delle matrici socioeconomiche e culturali dei comportamenti demograficamente rilevanti. Alla crescente esigenza di disporre di dati complessi a livello micro hanno risposto da un lato le indagini ad hoc, dall'altro l'uso congiunto del ricco patrimonio informativo disponibile negli archivi amministrativi. In parallelo, le capacità di calcolo, enormemente aumentate, hanno permesso lo sviluppo e l'applicazione diffusa di raffinati metodi d'analisi, così come l'elaborazione di modelli descrittivi ed interpretativi delle diverse componenti delle dinamiche demografiche, spesso coniugando l'analisi dei dati micro con la sintesi dei dati di contesto. L'attualità e le prospettive degli studi di popolazione, tuttavia, non sembrano altrettanto promettenti, almeno nel nostro Paese, dove una serie d'interventi sulle strutture accademiche, di ricerca e di produzione dei dati stanno limitando le possibilità del loro sviluppo. È dunque necessario rivedere le strategie di ricerca a partire dalle fonti dei dati, soprattutto migliorando l'uso di quelli amministrativi, opportunamente controllati nella loro qualità statistica e integrati fra loro a fini di ricerca.