

VOLUME LXXII – N. 2

APRILE-GIUGNO 2018

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA



DIRETTORE

CHIARA GIGLIARANO

COMITATO SCIENTIFICO

GIORGIO ALLEVA, GIAN CARLO BLANGIARDO, LUIGI DI COMITE, MAURO GALLEGATI, GIOVANNI MARIA GIORGI, ALBERTO QUADRIO CURZIO, CLAUDIO QUINTANO, SILVANA SCHIFINI D'ANDREA

COMITATO DI DIREZIONE

CHIARA GIGLIARANO, CLAUDIO CECCARELLI, PIERPAOLO D'URSO, SALVATORE STROZZA, ROBERTO ZELLI

REDAZIONE

LIVIA CELARDO, MARIATERESA CIOMMI, ANDREA CUTILLO, GIUSEPPE GABRIELLI, ALESSIO GUANDALINI, SIMONA PACE, GIUSEPPE RICCIARDO LAMONICA, ANDREA SPIZZICHINO

Sede Legale: C/O Studio Associato Cadoni, Via Ravenna n.34 – 00161 ROMA.
sieds.new@gmail.com, rivista.sieds@gmail.com

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

CONSIGLIO DIRETTIVO

Presidenti Onorari: LUIGI DI COMITE, GIOVANNI MARIA GIORGI

Presidente: FRANCESCO MARIA CHELLI

Vice Presidenti: CLAUDIO CECCARELLI, PIERPAOLO D'URSO,
ROBERTO ZELLI

Segretario Generale: MATTEO MAZZIOTTA

Consiglieri: EMMA GALLI, CHIARA GIGLIARANO, STEFANIA GIRONE, LUCIANO NIEDDU,
STEFANIA RIMOLDI, SILVANA MARIA ROBONE, SALVATORE STROZZA, CECILIA VITIELLO

Segretario Amministrativo: ALESSIO GUANDALINI

Revisori dei conti: FABIO FIORINI, SIMONE POLI, DOMENICO SUMMO

Revisori dei conti supplenti: MARGHERITA GEROLIMETTO, GIUSEPPE NOTARSTEFANO

SEDE LEGALE:

C/O Studio Associato Cadoni, Via Ravenna n.34 – 00161 ROMA

sieds.new@gmail.com

rivista.sieds@gmail.com

INDICE

Booth Arlene Karina, Pier Luigi Conti, De Vitiis Claudia, Guandalini Alessio <i>Un metodo alternativo per la selezione delle variabili nella calibrazione</i>	5
Giuseppe Ricciardo Lamonica <i>An analysis of methods for the treatment of autocorrelation in spatial interaction models</i>	17
Diego Chianella, Giuliano Latini, Annalisa Lucarelli, Emilia Matera, Marina Sorrentino <i>Differences in the enterprises' strata classification between sampling and calibration sources: the implications for Istat quarterly indicators on job vacancies and hours worked</i>	29
Miriam De Santis, Alessio Guandalini, Antonella Iorio, Alessandra Lugli <i>L'impatto dei vincoli di calibrazione sull'errore campionario nel Wave Approach</i>	41
Ciro Baldi, Marilena Angela Ciarallo, Sara Gigante, Silvia Pacini <i>The validation process of register based measures of hours paid and worked. How well they fit the international requirements compared to survey based measures?</i>	53
Margherita Gerolimetto, Stefano Magrini <i>Inference for inequality measures: a review</i>	65
Francesco M. Chelli, Mariateresa Ciommi, Francesca Mariani, Maria Cristina Recchioni <i>Opportunity and discrimination in tertiary education: a proposal of aggregation for some European countries</i>	77
Massimiliano Bruno, Alessandro Polli <i>Impatti dell'automazione sul mercato del lavoro. Nuove stime per il caso italiano</i>	89
Nidia Batic <i>Il lavoro ideale</i>	101

Ciro Baldi, Claudio Ceccarelli, Sara Gigante, Silvia Pacini, Francesca Rossetti <i>The labour register in Italy: the new heart of the system of labour statistics</i>	113
Ilaria Franco, Silvia Loriga e Andrea Spizzichino <i>Il caso studio delle ore lavorate in un'ottica di destagionalizzazione e revisioni.....</i>	125
Irene Brunetti, Valentina Ferri <i>Essere NEET in Italia: i principali fattori di rischio</i>	137
Antonio Cappiello <i>Mediation: economic concepts and some examples of rational framework for legal professionals.....</i>	149

UN METODO ALTERNATIVO PER LA SELEZIONE DELLE VARIABILI NELLA CALIBRAZIONE

Booth Arlene Karina, Pier Luigi Conti, De Vitiis Claudia, Guandalini Alessio

Introduzione

I principali istituti nazionali di statistica, tra cui l'ISTAT, utilizzano stimatori di calibrazione (Deville e Särndal, 1992; Särndal, 2007) per produrre stime da indagini campionarie su larga scala.

I motivi che rendono appetibile lo stimatore calibrato sono principalmente la possibilità di produrre stime più efficienti rispetto allo stimatore di Horvitz-Thompson (1956), l'individuazione di un unico sistema di pesi valido per tutte le variabili di interesse e la proprietà di coerenza con dei totali di popolazione noti da fonte amministrativa.

Negli anni il numero di variabili ausiliarie utilizzate nello stimatore calibrato è cresciuto notevolmente. Questo, da un lato, garantisce stime coerenti con un numero sempre più elevato di totali di popolazione ma, dall'altro, porta sempre più frequentemente a problemi di mancata individuazione del sistema di pesi "calibrati". Inoltre, in casi estremi, l'abuso di variabili ausiliarie nella "calibrazione" può portare ad un aumento eccessivo della variabilità dei pesi calibrati. Questo può avere un impatto sull'errore campionario delle stime e può rendere inefficiente lo stimatore.

La scelta delle variabili ausiliarie da considerare nella calibrazione, dunque, ha una notevole importanza. Le variabili ausiliarie da inserire nel sistema di calibrazione generalmente sono scelte in base a valutazioni di esperti. Tuttavia, può essere fatto anche studiando il modello migliore in termini di *fitting* o, meno frequentemente, valutando l'aumento di variabilità dei pesi "calibrati", ad esempio calcolando il loro coefficiente di variazione (Kish, 1992, p. 191).

Nel presente lavoro si propone un approccio in grado di tener conto contemporaneamente di entrambe le componenti che determinano la varianza dello stimatore calibrato. Adattando a questo contesto il metodo di scomposizione di Shapley (Shapley, 1953), mutuato dalla teoria dei giochi cooperativi, è infatti possibile determinare l'apporto marginale, in termini di precisione ma soprattutto di efficienza, di ciascuna variabile ausiliaria sulle stime delle variabili di interesse.

Il lavoro è così strutturato: nel paragrafo 1 viene richiamata brevemente la metodologia alla base dello stimatore di calibrazione e analizzata l'espressione del suo stimatore della varianza campionaria. Nel paragrafo 2 sono presentati alcuni metodi per la scelta delle variabili ausiliarie ed in particolare, nel paragrafo 2.3, il metodo della scomposizione di Shapley viene descritto ed adattato a questo particolare contesto. Nel paragrafo 3 è brevemente presentata l'indagine europea sulla salute (EHIS) che è stata utilizzata come banco di confronto tra i vari metodi. I risultati del confronto sono presentati nel paragrafo 4, mentre nel paragrafo 5 sono discusse le conclusioni.

1. Lo stimatore di calibrazione

Sia U la popolazione di riferimento definita come un insieme finito di N elementi, indicati con $U = (1, \dots, k, \dots, N)$. Sia s un campione selezionato da U secondo un disegno di campionamento $(S, p(\cdot))$, dove S è lo spazio campionario e $p(\cdot)$ è una distribuzione di probabilità su S . La probabilità di inclusione del primo ordine, $\pi_k = P(k \in s)$, e la probabilità di inclusione del secondo ordine, $\pi_{kl} = P(k, l \in s)$, si presume siano note e positive. Sia inoltre, $\pi_{kk} = \pi_k$.

Supponiamo di essere interessati a stimare il totale di una generica variabile Y , $t_y = \sum_{k \in U} y_k$. Per ciascuna unità del campione s vengono rilevate, oltre alla variabile di interesse y_k , una serie di variabili ausiliarie ($j = 1, \dots, J$) rappresentate dal vettore $\mathbf{x}_k = (x_{1k}, \dots, x_{jk}, \dots, x_{Jk})'$. Così per ciascuna unità campionaria si ha un vettore di $J + 1$ osservazioni definito nel seguente modo: $(y_k, \mathbf{x}_k), k \in s$.

Nel caso in cui, per queste variabili ausiliarie, si dispone di totali noti da fonti amministrative, Deville e Särndal (1992) propongono di utilizzare lo stimatore di calibrazione (CAL) definito come:

$$\hat{t}_{y_{CAL}} = \sum_{k \in s} w_k y_k$$

in cui w_k è il peso "calibrato" dell'unità k -ma per $k \in s$. I pesi w_k sono scelti attraverso la risoluzione di un problema di minimo vincolato che, tra i pesi che consentono di ottenere l'uguaglianza tra i totali stimati sul campione e i totali noti da fonte amministrativa, individua quelli più vicini possibili ai pesi da disegno $d_k = 1/\pi_k$ rispetto ad una metrica scelta $G(\cdot)$.

Possono essere considerate diverse funzioni di distanza $G(\cdot)$. Nel caso in cui si utilizza la funzione di distanza del chi-quadro, lo stimatore di calibrazione che si ottiene coincide con lo stimatore di regressione generalizzata ($GREG$) (cfr. Cassel *et al.*, 1979; Särndal, 1980; Isaki e Fuller, 1982; Wright, 1983; Bethlehem e Keller,

1987; Särndal *et al.*, 1989; Fuller, 2002). Deville e Särndal (1992, pp 378-379) hanno dimostrato che in generale utilizzando funzioni di distanza che soddisfano determinate condizioni di regolarità, si ottengono stimatori calibrati asintoticamente equivalenti allo stimatore *GREG*. Quando si utilizza lo stimatore *CAL*, dunque, si assume implicitamente un modello di regressione lineare di super-popolazione $\xi: y_k = \mathbf{x}'_k \boldsymbol{\beta} + \varepsilon_k$ tra la variabile d'interesse Y e le variabili ausiliarie dove ε_k sono delle variabili casuali incorrelate, con valore atteso nullo e varianza σ_k^2 .

Lo stimatore della varianza dello stimatore *CAL* può essere quindi scritto come quello dello stimatore della varianza del *GREG*:

$$\text{var}(\hat{t}_{y_{CAL}}) = \sum_{k \in S} \sum_{l \neq k} \frac{\Delta_{kl}}{\pi_{kl}} (\hat{e}_k w_k) (\hat{e}_l w_l) \quad (1)$$

in cui $\Delta_{kl} = \frac{\pi_{kl} - \pi_k \pi_l}{\pi_k \pi_l}$ e \hat{e}_k e \hat{e}_l sono i residui del modello di regressione lineare di super-popolazione stimati sul campione dove $k, l \in s$ e $k \neq l$. Dalla (1) risulta, quindi, che lo stimatore della varianza campionaria dello stimatore *CAL* dipende dalla bontà di regressione lineare di super-popolazione e dai pesi "calibrati".

2. Importanza della scelta delle variabili ausiliarie e metodi per la loro selezione

I principali motivi che hanno reso appetibile lo stimatore calibrato per gli istituti nazionali di statistica sono la possibilità di produrre stime più efficienti rispetto allo stimatore di Horvitz-Thompson (*HT*, 1956), l'individuazione di un unico sistema di pesi valido per tutte le variabili di interesse e la proprietà di coerenza con dei totali di popolazione noti da fonte amministrativa.

Forse anche a causa di quanto sopra detto, negli anni il numero di variabili ausiliarie utilizzate nello stimatore calibrato è cresciuto notevolmente. Questo, da un lato garantisce stime coerenti con un numero sempre più elevato di totali di popolazione ma, dall'altro, porta sempre più frequentemente a problemi di mancata individuazione del sistema di pesi "calibrati". Inoltre, in casi estremi, l'abuso di variabili ausiliarie può far aumentare in maniera eccessiva la variabilità dei pesi calibrati. Questo può avere un impatto sull'errore campionario delle stime e può rendere inefficiente lo stimatore.

La scelta delle variabili ausiliarie da inserire nel sistema di calibrazione generalmente è affidata alla valutazione di esperti. Volendo attribuire, però, un fondamento metodologico a questa scelta si può seguire un approccio basato sul modello o "*GREG thinking*" (utilizzando la terminologia in Särndal, 2007 p. 102) o un approccio "*design-based*".

2.1. Approccio “GREG thinking”

L’idea alla base di questo approccio è quella di utilizzare nella calibrazione le variabili ausiliare che portano a modelli migliori in termini di *fitting*. Per questo fine si può utilizzare il criterio di informazione di Akaike

$$AIC = 2k - 2 \log(L),$$

in cui k è il numero di parametri (variabili ausiliarie) considerate nel modello e L è la relativa stima di massima verosimiglianza. I modelli con l’*AIC* più basso e quindi con residui più piccoli, per come è definita la (1), dovrebbero portare a stimatori più efficienti.

2.2. Approccio “design-based”

Nell’approccio basato sul disegno l’attenzione si concentra sulla variabilità dei pesi. Si tiene, infatti, conto della variabilità dei pesi attraverso un indice adimensionale proposto da Kish (1992, p. 191) così definito:

$$K = 1 + CV^2$$

in cui il CV è il coefficiente di variazione dei pesi finali. L’idea alla base di questo approccio è che utilizzare pesi campionari troppo variabili può portare a stime meno efficienti. In questo caso la scelta delle variabili ausiliarie da considerare ricade su quelle che determinano i pesi finali w_k tali per cui K è minore. In questo modo si ritiene di tenere sotto controllo la parte di varianza dello stimatore che dipende dalla variabilità dei pesi “calibrati”.

2.3. Metodo di scomposizione di Shapley

Entrambi gli approcci appena visti hanno però dei limiti: quello “GREG thinking” ha lo svantaggio di non tenere sotto controllo la variabilità dei pesi “calibrati”. L’approccio “design-based”, invece, considera come fondamentale l’aumento di variabilità dei pesi “calibrati” ignorando completamente l’impatto dell’assunzione del modello di regressione lineare da super-popolazione sulle stime.

Per risolvere tale inconveniente, in questo lavoro, viene proposto il metodo di scomposizione di Shapley (1953), che consente di tener conto di entrambi gli aspetti e, quindi, di entrambe le componenti che entrano nello stimatore della varianza dello stimatore *CAL*.

Questo metodo, inoltre, consente di determinare l’apporto marginale, in termini di efficienza e di precisione, di ciascuna variabile ausiliaria sulle stime delle variabili di interesse.

Il metodo di scomposizione di Shapley è stato introdotto dallo stesso Shapley nel 1953 e si basa su concetti relativi alla teoria dei giochi cooperativi. L'obiettivo di Shapley era quello di ridistribuire i guadagni generati da una coalizione di giocatori in maniera proporzionale al contributo che ciascun giocatore ha apportato alla coalizione stessa. Il valore Shapley (SV), ovvero la parte del guadagno della coalizione del generico giocatore i , può essere determinato in diversi modi. Uno di questi è:

$$SV(i, v) = \sum_{S \in N \setminus \{i\}} \frac{|S|! (|N| - |S| - 1)!}{|N|} (v(S \cup \{i\}) - v(S)) \quad (2)$$

in cui S è la coalizione contenuta in un insieme di N giocatori e $|S|$ è la sua cardinalità, v è la funzione di utilità che deve soddisfare una serie di proprietà (cfr. Shapley, 1953, pp. 312-313).

Recentemente l'uso del metodo di scomposizione di Shapley, oltre che per ripartire i guadagni utilizzati di ciascun giocatore, è stato utilizzato per determinare il contributo marginale di ciascun regressore nel *fitting* di un modello di regressione lineare (Israeli, 2007), per scomporre alcune misure di povertà (Shorrocks, 2013) ed alcuni indici di disuguaglianza per fattori (Sastre e Trannoy, 2002) e per sotto-gruppi di popolazione (Deutsch e Silber, 2007; Giorgi e Guandalini, 2016).

Con alcuni accorgimenti il metodo Shapley può essere adattato per determinare l'impatto marginale che ciascuna variabile ausiliaria ha sulle stime e sul relativo errore campionario ottenuto con lo stimatore *CAL*.

Ipotizziamo di voler considerare 5 variabili ausiliarie (a, b, c, d, e) nello stimatore *CAL* e di essere interessati a determinare il contributo marginale della prima variabile (a). La (2) diventa:

$$\begin{aligned} SV_a &= \frac{1}{5} (I - I_a) \\ &+ \frac{1}{20} [(I_b - I_{ab}) + (I_c - I_{ac}) + (I_d - I_{ad}) + (I_e - I_{ae})] \\ &+ \frac{1}{30} [(I_{bc} - I_{abc}) + (I_{bd} - I_{abd}) + (I_{be} - I_{abe}) \\ &\quad + (I_{cd} - I_{acd}) + (I_{ce} - I_{ace}) + (I_{de} - I_{ade})] \\ &+ \frac{1}{20} [(I_{bcd} - I_{abcd}) + (I_{bce} - I_{abce}) + (I_{cde} - I_{acde}) + (I_{bde} - I_{abde})] \\ &+ \frac{1}{5} (I_{bcde} - I_{abcde}) \end{aligned}$$

Nell'espressione precedente si può considerare:

- $$I_* = \frac{\hat{t}_{y_{CAL*}}}{\hat{t}_{y_{HT*}}}$$

per determinare l'apporto marginale di (a) sulla stima, ovvero, come varia in proporzione rispetto allo stimatore HT la stima, nel caso in cui la variabile ausiliaria (a) viene inclusa nel sistema di calibrazione;

- $$I_* = \frac{var(\hat{t}_{y_{CAL*}})}{var(\hat{t}_{y_{HT*}})}$$

per determinare l'apporto marginale di (a) sulla varianza, ovvero come varia in proporzione l'errore campionario rispetto a quello dello stimatore HT quando si considera la variabile ausiliaria (a) nel sistema di calibrazione¹.

3. L'indagine europea sulla salute (EHIS)

In questo paragrafo viene illustrata brevemente l'indagine europea sulla salute (EHIS) su cui poi verranno applicati gli approcci presenti nel paragrafo precedente.

L'EHIS è un'indagine armonizzata a livello europeo. Il campione della rilevazione svolta dall'ISTAT nel 2015 ammonta a circa 15900 famiglie selezionate in oltre 550 comuni distribuiti su tutto il territorio italiano. Il disegno di campionamento utilizzato è di tipo complesso a due stadi (comuni, famiglie) con stratificazione dei comuni per ampiezza demografica.

Per ogni famiglia selezionata sono stati intervistati tutti i componenti di almeno 15 anni. La raccolta dei dati è avvenuta tramite delle interviste dirette, condotte con la tecnica PAPI (*Pen And Paper Interview*) e tramite la tecnica dell'auto-compilazione del questionario.

L'obiettivo dell'indagine è produrre stime, individuali e familiari, sullo stato di salute (incidenza di alcune malattie o patologie), sull'accesso e l'utilizzo dei servizi sanitari e sulle determinanti di salute rispetto al contesto socio-demografico. In fase di stima si fa ricorso allo stimatore CAL che utilizza le seguenti variabili ausiliarie:

- Popolazione per 5 aree geografiche²;

¹ L'asterisco (*) sta ad indicare quale combinazione di variabili ausiliarie viene incluso nella calibrazione. Ad esempio I_{abc} sarà il valore (4) o (5), nel caso in cui il sistema di calibrazione consideri le variabili ausiliarie (d) ed (e). Si noti, infatti, che $I_{abcde} = 1$ perché il numeratore corrisponde proprio all'errore o alla stima ottenuta utilizzando come stimatore quello senza nessun vincolo, ovvero lo stimatore HT .

² Nord-Est, Nord-Ovest, Centro, Sud e Isole.

- b) Popolazione per regione³;
- c) Popolazione per 14 classi di età⁴;
- d) Popolazione per regione e 5 classi di età⁵;
- e) popolazione per sesso.

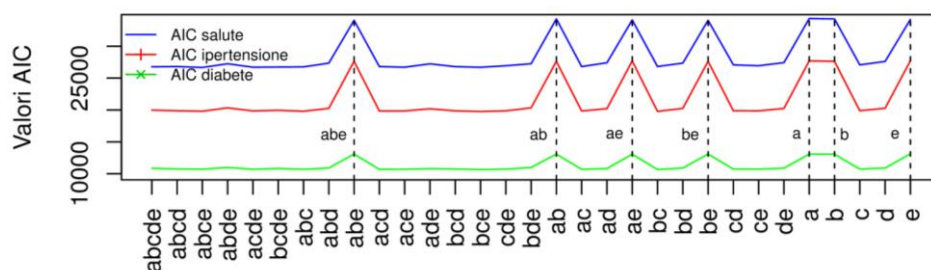
4. Confronto tra i tre approcci

Come già detto, il confronto dei tre approcci illustrati nel paragrafo 3 è svolto sui dati dell'indagine europea sulla salute (EHIS). In particolare, sono state considerate tre variabili d'interesse dicotomiche con una diversa incidenza sulla popolazione:

1. Stato di salute buono o molto buono;
2. Ipertensione;
3. Diabete.

La scelta di queste tre variabili è stata dettata dal fatto che la prima è la variabile principale dell'indagine, la seconda ha una grande incidenza sulla popolazione di riferimento (circa il 20%) e la terza ha un'incidenza bassa (solo il 5%). L'applicazione è stata svolta utilizzando il software statistico R. Per la calibrazione ed il calcolo degli errori campionari si è fatto ricorso al pacchetto ReGenesees (Zardetto, 2015).

Figura 1 – Valore dell'AIC per diversi modelli lineari generalizzati che considerano le possibili combinazioni di variabili ausiliarie.



Nell'approccio basato sul modello, per ognuna delle 3 variabili considerate, sono stati calcolati i modelli lineari generalizzati che sottostanno a tutte le possibili combinazioni di variabili ausiliarie presentate nel paragrafo precedente e che possono costituire un sistema di calibrazione per produrre le stime dell'indagine. Nel caso in cui sono state utilizzate tutte e 5 le variabili ausiliarie, per la generica variabile y , è stato calcolato il seguente modello: $y \sim [a * c + b * d] * e - 1$.

³ Bolzano e Trento pur essendo province sono considerate al pari di regioni per esigenze di stima.

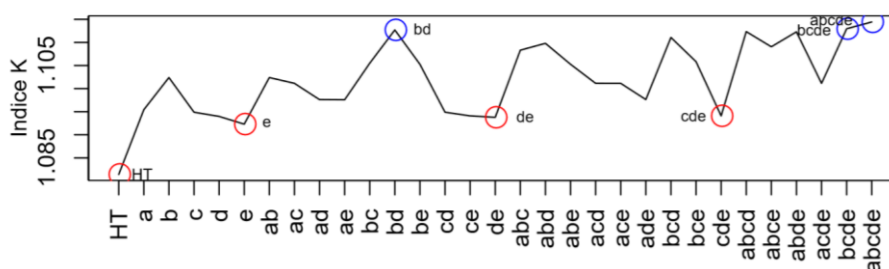
⁴ 0-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-75, 75 e più.

⁵ 0-14, 15-29, 30-49, 50-64, 65 e più;

Per ciascun modello è stato calcolato l'*AIC* per valutare il modello migliore e i valori sono riportati in Figura 1. Come era lecito attendersi, i modelli migliori sono quelli che utilizzano più variabili ausiliarie. I modelli che contengono le variabili popolazione per aree geografiche (*a*) o per regione (*b*) o popolazione per sesso (*e*), sono quelli meno esplicativi. Secondo questo approccio saremmo portati ad inserire nel sistema di calibrazione solamente le variabili ausiliarie relative alla distribuzione della popolazione per 14 classi di età (*c*) e popolazione per regione e 5 classi di età (*d*).

Nell'approccio "design-based" si è analizzata la variabilità dei pesi finali w_k calcolati risolvendo tutti i possibili sistemi di calibrazione ottenuti combinando le diverse variabili ausiliarie.

Figura 2 – Valore dell'indice di Kish (*K*) per i pesi finali ottenuti calibrando rispetto a diverse combinazioni di variabili ausiliarie.



Analizzando i risultati in Figura 2 si nota che all'aumentare del numero di variabili ausiliarie l'indice *K* cresce. Il valore più alto di *K* è ottenuto considerando tutte e 5 le variabili ausiliarie. L'approccio "design based" porterebbe ad utilizzare come variabili ausiliarie nella calibrazione: popolazione per 14 classi di età (*c*), popolazione per regione e 5 classi di età (*d*) e popolazione per sesso (*e*).

Come già detto nel paragrafo precedente, questi due metodi hanno il limite di non considerare contemporaneamente entrambe le componenti della (1) e, come nel caso in questione, possono portare a scelte diverse.

Il metodo di scomposizione di Shapley consente di valutare l'effetto dei diversi sistemi di calibrazione nel loro nel complesso. Il risultato più importante che si ottiene utilizzando questo metodo è, infatti, la possibilità di calcolare per ciascuna variabile ausiliaria considerata l'impatto che questa ha, al netto delle altre, sulle stime e sugli errori campionari.

Nella Tabella 1 si può notare che le 5 variabili ausiliarie hanno un impatto simile tra di loro sulla stima delle variabili di interesse considerate. Le differenze tra una variabile di interesse e l'altra sono dovute alla diversa incidenza di queste nella popolazione. Inoltre, gli impatti sulle stime delle variabili ausiliarie sono sempre positivi in quanto nella EHIS la calibrazione viene utilizzata anche per

Tabella 1 – Valore Shapley delle variabili ausiliarie (a), (b), (c), (d) e (e) in valore assoluto (V.A) e percentuale (%) sulle stime del totale a livello Italia e sui relativi coefficienti di variazione percentuali (CV%) per le variabili “Buona salute”, “Ipertensione”, “Diabete” con relativi intervalli di confidenza (IC) bootstrap al 95%. EHS, 2015.

Variabili di interesse	HT				
	(a)	(b)	(c)	(d)	(e)
	Impatto sulle stime del totale (\hat{t}_y)				
V.A.	34673	1954	1871	1892	1902
IC	[34225; 35026]	[1884; 2058]	[1801; 1940]	[1821; 1997]	[1866; 2007]
%	100	5,6	5,4	5,4	5,4
IC	[5,4; 5,8]	[5,4; 5,9]	[5,2; 5,6]	[5,2; 5,7]	[5,3; 5,7]
V.A.	8862	345	310	345	345
IC	[8696; 9006]	[309; 389]	[274; 345]	[318; 371]	[318; 380]
%	100	3,9	3,5	3,9	3,9
IC	[3,6; 4,1]	[3,5; 4,4]	[3,1; 3,9]	[3,6; 4,2]	[3,6; 4,3]
V.A.	2874	106	106	112	106
IC	[2793; 2949]	[88; 128]	[91; 120]	[100; 126]	[88; 123]
%	100	3,7	3,7	3,9	3,7
IC	[3,2; 4,1]	[3,1; 4,5]	[3,2; 4,2]	[3,5; 4,4]	[3,1; 4,3]
	Impatto sui coefficienti di variazione percentuali (CV%)				
V.A.	1,72	-0,29	-0,30	-0,30	-0,27
IC	[1,64; 1,82]	[-0,29; -0,27]	[-0,30; -0,27]	[-0,32; -0,28]	[-0,28; -0,26]
%	100	-16,2	-17,5	-17,4	-16,0
IC	[-16,9; -15,5]	[-17,6; -15,9]	[-18,6; -16,4]	[-18,3; -16,4]	[-16,6; -15,3]
V.A.	2,03	-0,20	-0,27	-0,27	-0,16
IC	[0,91; 3,12]	[-0,21; -0,13]	[-0,25; -0,15]	[-0,33; -0,19]	[-0,21; -0,12]
%	100	-8,4	-9,7	-13,4	-8,1
IC	[-10,5; -6,3]	[-12,2; -7,2]	[-16,1; -9,3]	[-16,2; -10,7]	[-10,1; -6,1]
V.A.	3,11	0,21	0,25	0,25	0,12
IC	[3,45; 3,82]	[-0,25; -0,08]	[-0,30; -0,10]	[-0,33; -0,16]	[-0,20; -0,05]
%	100	-5,3	-7,9	-8,0	-4,0
IC	[-7,9; -2,7]	[-9,8; -3,3]	[-10,6; -5,3]	[-10,6; -5,4]	[-6,5; -1,6]

(a) Popolazione per 5 aree geografiche; (b) Popolazione per regione; (c) Popolazione per 14 classi di età;

(d) Popolazione per regione e 5 classi di età; (e) Popolazione per sesso.

correggere la mancata risposta e quindi le stime *CAL* sono sempre maggiori di quelle *HT*.

Riguardo l'impatto sull'efficienza delle stime, misurata in termini di $CV\% = \sqrt{\text{var}(\hat{t}_Y)} / \hat{t}_Y \cdot 100$, i risultati in Tabella 1 evidenziano la notevole importanza della variabile popolazione per 14 classi di età (*c*) e popolazione per regione e 5 classi di età (*d*). L'impegno di queste due variabili contribuisce a ridurre notevolmente l'errore campionario delle stime di tutte e tre le variabili studiate. L'impatto delle altre variabili ausiliare è comunque importante.

Analizzando gli estremi degli intervalli di confidenza stimati attraverso il bootstrap non-parametrico ($M=1000$) si ha un'ulteriore conferma che tutte le variabili ausiliarie contribuiscono a correggere verso l'alto le stime e a ridurre l'errore campionario.

Risultati leggermente diversi si ottengono su alcuni domini subnazionali che, per brevità, non sono stati riportati qui. La calibrazione, infatti, può avere effetti diversi soprattutto in termini di efficienza a seconda del dominio di stima. Ad esempio, a livello regionale la variabile ausiliaria popolazione per regione (*b*) assume maggiore importanza. Mentre la popolazione per sesso (*e*), in alcuni casi, comporta addirittura un lieve peggioramento dell'efficienza delle stime. Tuttavia, considerare questa variabile ausiliaria nella calibrazione consente di garantire la coerenza con una variabile molto importate nelle indagini ISTAT.

5. Conclusioni

I principali istituti nazionali di statistica, tra cui l'ISTAT, utilizzano lo stimatore di calibrazione (Deville e Särndal, 1992; Särndal, 2007) per produrre stime da indagini campionarie su larga scala. Negli anni il numero di variabili ausiliarie utilizzate nello stimatore calibrato è cresciuto notevolmente. Questo ha fatto sì che sempre più frequentemente sorgessero problemi di mancata individuazione del sistema di pesi "calibrati". Inoltre, l'elevato numero di variabili ausiliarie nella calibrazione comporta un aumento della variabilità dei pesi che, in alcuni casi, può riflettersi in una perdita di efficienza delle stime. La scelta delle variabili ausiliarie da inserire nella calibrazione, dunque, assume notevole importanza.

In questo lavoro, per selezionare le variabili ausiliarie da inserire nel processo di calibrazione, è stato adattato a questo contesto il metodo di scomposizione di Shapley (1953). Questo metodo, rispetto ad altri, consente di valutare l'effetto dei diversi sistemi di calibrazione nel complesso. Il risultato più importante che si ottiene utilizzandolo è, infatti, la possibilità di calcolare per ciascuna variabile ausiliaria considerata l'impatto che questa ha, al netto delle altre, sulle stime e sugli errori campionari.

L'applicazione su un'indagine su larga scala come l'indagine europea sulla salute (EHIS) ha mostrato l'applicabilità del metodo e le numerose considerazioni che se ne possono trarre.

Ringraziamenti

Questo articolo è frutto della ricerca svolta dalla dott.ssa Arlene Karina Booth durante il tirocinio presso la Direzione Centrale per le Metodologie dell'Istituto Nazionale di Statistica per la stesura della tesi di Laurea Magistrale dal titolo "Scelta delle variabili ausiliarie negli stimatori per calibrazione, con applicazione all'indagine Istat sulla salute".

Riferimenti bibliografici

- BETHLEHEM J.G., KELLER J.W. 1987. Linear weighting of sample survey data. *Journal of Official Statistics*, Vol. 3, No. 2, pp. 141-153.
- CASSEL C.M., SÄRNDAL C.-E., WRETMAN J.H. 1979. Prediction theory on finite population when model-based and design-based principle are combined. *Scandinavian Journal of Statistics*, Vol. 6, No. 3, pp. 97-106.
- DEUTSCH J., SILBER J. 2007. Decomposing income inequality by population subgroups: a generalization. In *Inequality and Poverty*, pp. 237-253. Emerald Group Publishing Limited.
- DEVILLE J.C., SÄRNDAL C.-E. 1992. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, Vol. 87, pp. 376-382.
- FULLER W.A. 2002. Regression estimation for survey samples. *Survey Methodology*, Vol. 28, No. 1, 5-23.
- GIORGI G.M., GUANDALINI A. 2016. Bonferroni index decomposition and the shapley method. *Rivista Italiana di Economia Demografia e Statistica*, Vol. 70, No. 4.
- HORVITZ D.G., THOMPSON D.J. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, Vol. 47, No. 260, pp. 663-685.
- ISAKI C.T. & FULLER W.A. 1982. Survey design under the regression superpopulation model. *Journal of the American Statistical Association*, Vol. 77, No. 377, pp. 89-96.
- ISRAELI O. 2007. A Shapley based decomposition of the R-square of a linear regression. *The Journal of Economic Inequality*, Vol. 5, No.2, pp.199-212.
- SÄRNDAL C.-E. 1980. On π -inverse weighting versus best linear unbiased weighting in probability sampling. *Biometrika*, Vol. 67, No. 3, pp. 639-650.
- SÄRNDAL C.-E. 2007. The calibration approach in survey theory and practice. *Survey Methodology*, Vol. 33, No. 2, pp. 99-119.

- SASTRE M., TRANNOY A. 2002. Shapley inequality decomposition by factor components: Some methodological issues. *Journal of Economics*, Vol. 9, pp. 51-89.
- SHORROCKS A.F. 2013. Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality*, pp. 1-28.
- WRIGHT R.L. 1983. Finite population sampling with multivariate auxiliary information. *Journal of the American Statistical Association*, Vol. 78, No. 384, pp. 879-884.
- ZARDETTO D. 2105. Regenesees: an advanced r system for calibration, estimation and sampling error assessment in complex sample surveys. *Journal of Official Statistics*, Vol. 31, No. 2, pp. 177-203.

SUMMARY

An alternative method for the variable selection in calibration

The main national institutes of statistics use calibration estimator (Deville e Särndal, 1992; Särndal, 2007). During the years the numbers of auxiliary totals considered in calibration is strongly increased. Because of this, two main problems arise. Very often there is failure to identify "calibrated" weights and, furthermore, the variability of "calibrated" weights can make the efficiency of estimates worse. Therefore, the choice of auxiliary variables in calibration is of a greatest importance.

In the present paper the Shapley decomposition method (Shapley, 1953) has been adapted to this context. Compared to other methods, the Shapley decomposition enables the evaluation of the impact of different calibration systems as a whole. Moreover, with the Shapley decomposition method is possible to derive the marginal impact, net of others auxiliary variables, on the estimates and on the sampling error of each auxiliary variable used in calibration.

The method has been used on the European Health Interview Survey (EHIS) data and it has shown its applicability and, furthermore, a lot consideration that scholars can draw from it.

Arlene Karina BOOTH, Università di Roma "Sapienza", arlene_93@hotmail.it
Pier Luigi CONTI, Università di Roma "Sapienza", pierluigi.conti@uniroma1.it
Claudia DE VITIIS, ISTAT, devitiis@istat.it
Alessio GUANDALINI, ISTAT, alessio.guandalini@istat.it

AN ANALYSIS OF METHODS FOR THE TREATMENT OF AUTOCORRELATION IN SPATIAL INTERACTION MODELS

Giuseppe Ricciardo Lamonica

1. Introduction

In regard to spatial regression models where the units subject to analysis are of territorial type, a problem of fundamental importance, often ignored, is that the observations may be dependent which entails that the value of a variable in an elementary territorial unit is determined by those observed in neighbouring localities. When this situation occurs, the phenomenon is spatially autocorrelated (or spatially dependent). The presence of spatial autocorrelation is a problem because the traditional regression models are based on the hypothesis of independence among observations in the different localities of the territory analysed. Hence, when this hypothesis is violated, the estimates of the regression model parameters are bias and/or inefficient. The spatial dependence occurs in two different forms. In the first case, it affects only the error term in the regression model and is mostly considered to be a nuisance which needs to be eliminated. In this case, the spatial error autocorrelation does not cause ordinary least squares (OLS) estimates to be biased, but it alters their efficiency. In the second case, the value of the dependent variable in a generic unit is affected by those of the neighbours units. In this case, OLS estimates are bias and no longer consistent.

The spatial econometrics that deals with treatment of spatial autocorrelation in regression models has developed various techniques to solve this problem. Among them, the most popular approaches used are the Spatial Filtering Model (Fisher and Griffith 2008; Griffith 2009; Chun and Griffith 2011), the Spatial Autoregressive Model and Spatial Error Model (LeSage 1997a; Anselin 1988; LeSage and Pace 2008). While in literature exists a voluminous series of empirical analysis to assess the utility of these methods for cross-section and panel data a scant attention was paid for interactions or flows data. In this last context, the spatial autocorrelation analysis is exacerbated because contrary to classical panel data where the sample involves n spatial unit, with each unit being an observation, the interaction spatial data involve n^2 origin-destination pairs.

Limiting the attention to the interaction data, this paper, by means of a real phenomenon, outline and compare the three methods previously quoted in order to identify their shared characteristics and those specific to each of them. It should be pointed that the literature, excluding some studies, see for example Fisher and Griffith (2008), is not particularly rich in analyses of this type. Consequently, it is interesting furnish a contribution which explores the potentialities of these methods, even if in a limited situation like the one considered by this study.

Moreover, given the particular area of inquiry, the analysis makes no claim to being exhaustive, and the results are to be considered preliminary.

2. The methods for modelling the spatial autocorrelation

In this section, the three techniques to consider the spatial autocorrelation in the interaction models previously quoted, are briefly exposed.

Suppose we have a set of n spatial unit and let \mathbf{F} an $n \times n$ square matrix whose generic entry f_{ij} ($i, j=1, \dots, n$) is the interaction (flow) whose origin is the i -th unit whose destination is the j -th unit.

We restrict the discussion in the context of gravity models which are considered one of the most important spatial interaction model (Everett and Keller, 2002).

The interaction between two generic units depends by three type of explicative variables (covariates). Those that characterize the origin unit of flow (push), those that characterize the destination unit of flow (pull) and finally the covariates that measure the separation between the origin and destination units. Using matrix notation:

$$\mathbf{f} = \alpha \mathbf{1} + \mathbf{X}_O \otimes \mathbf{1} \beta_O + \mathbf{1} \otimes \mathbf{X}_D \beta_D + \mathbf{d} \gamma + \boldsymbol{\varepsilon} \quad (1)$$

where: \mathbf{f} is the vector whose elements are the logarithmic of the flows observed; $\mathbf{1}$ is the $n(n-1) \times 1$ unit vector; \mathbf{X}_O and \mathbf{X}_D are two $n(n-1) \times k$ matrices containing respectively the logarithmic of the covariates relating to the origin spatial unit and destination spatial unit of the flows; \mathbf{d} is the vector whose elements are the logarithmic of the distance between two generic spatial unit; $\boldsymbol{\varepsilon}$ is the vector of the residual variable. Finally, \otimes is the Kronecker product.

Referring to Fisher and Griffith (2008) and LeSage and Pace (2008) for details, the interaction phenomena may exhibit various types of autocorrelation. Autocorrelation at the origin: given a generic flow f_{ij} , also the units contiguous to the i -th unit have similar flows to the j -th unit. Autocorrelation at the destination: given a generic flow f_{ij} , the i -th unit has flows similar with respect to the units bordering on the j -th unit. Autocorrelation at both the origin and the destination: which simultaneously concerns the two previous formulations.

If the autocorrelation occurs in the dependent variable in order to capture the first type of the autocorrelation, that at the origin, it is possible to consider the following model named Origin Spatial autoregressive model (also O-SAR):

$$\mathbf{f} = \alpha \mathbf{1} + \rho \mathbf{W}_O \mathbf{f} + \mathbf{X}_O \boldsymbol{\beta}_O + \mathbf{X}_D \boldsymbol{\beta}_D + \gamma \mathbf{d} + \boldsymbol{\varepsilon} \quad (2)$$

Where $\mathbf{W}_O = \mathbf{I} \otimes \mathbf{C}^*$ and \mathbf{C}^* is the \mathbf{C} matrix standardized by row (\mathbf{C} is the $n \times n$ matrix of the first-order contiguities). In this case, the vector $\mathbf{W}_O \mathbf{f}$ is such that the generic element is the mean flow from the units contiguous to i -th unit to the j -th unit, and it represents the interaction that would be observed between the two units on the hypothesis of an autocorrelation at the origin.

As regards the second type of autocorrelation, it is sufficient considers in the (2) the following \mathbf{W}_D matrix instead of \mathbf{W}_O :

$$\mathbf{W}_D = \mathbf{C}^* \otimes \mathbf{I} \quad (3)$$

In this case, the generic element of the vector $\mathbf{W}_D \mathbf{f}$ is a mean of the flows from the i -th unit to the units contiguous with the j -th element. The corresponding spatial model (4) that reflects the destination-based dependence is named Destination-SAR (also: D-SAR). Finally, by multiplying \mathbf{W}_O and \mathbf{W}_D (i.e. $\mathbf{W}_{OD} = \mathbf{W}_O \mathbf{W}_D$), it is possible to consider, the third type of autocorrelation. In this last case, the spatial model is the Origin-Destination-SAR (also: OD-SAR).

On the other hand, if the dependence occurs in the error term of the (2), depending on the type of dependence, one of the following model can be utilized:

$$\mathbf{f} = \alpha \mathbf{1} + \mathbf{X}_O \boldsymbol{\beta}_O + \mathbf{X}_D \boldsymbol{\beta}_D + \gamma \mathbf{d} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} = \lambda \mathbf{W}_j \boldsymbol{\varepsilon} + \mathbf{u} \quad \text{and} \quad \mathbf{u} \approx N(0; \sigma^2 \mathbf{I}) \quad (4)$$

If $j=O$ then the model is the Origin Spatial Error Model (O-SEM). For $j=D$ we have the Destination Spatial Error Model (D-SEM). Finally, if $j=OD$ we have the Origin-Destination Spatial Error Model (OD-SEM).

The previous three types of autocorrelation can be jointly considered in a regression model. Restricting, but not limiting the attention to SAR models we can take account the following specifications:

- $\mathbf{f} = \rho_O \mathbf{W}_O \mathbf{f} + \rho_D \mathbf{W}_D \mathbf{f} + \rho_{OD} \mathbf{W}_{OD} \mathbf{f} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon};$
- $\mathbf{f} = \rho_O \mathbf{W}_O \mathbf{f} + \rho_D \mathbf{W}_D \mathbf{f} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon};$
- $\mathbf{f} = \rho (\mathbf{W}_O + \mathbf{W}_D) \mathbf{f} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon};$
- $\mathbf{f} = \rho (\mathbf{W}_O + \mathbf{W}_D + \mathbf{W}_{OD}) \mathbf{f} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon};$

An alternative methodology is the spatial filtering method (FS). See for example Griffith (2003) for details. It presumes that the spatial dependence in the dependent variable is due to one or more spatially autocorrelated not directly

observable variables. As surrogates for the latter the method considers the Moran's I autocorrelation index:

$$I = \frac{\mathbf{f}' \left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \mathbf{W}_j \left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \mathbf{f}}{\mathbf{1}' \mathbf{W}_j \mathbf{1} \mathbf{f}' \left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \mathbf{f}} \quad (5)$$

Referring to De Jong et al. (1984) for details, if e^{\min} and e^{\max} respectively denote the highest and smallest eigenvalues of \mathbf{W}_j , the following inequality holds:

$$e^{\min} \leq \frac{1}{\frac{\mathbf{1}' \mathbf{W}_j \mathbf{1}}{n(n-1)}} \leq e^{\max} \quad (6)$$

Moreover, these extreme eigenvalues also coincide with those of the following matrix:

$$\left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \mathbf{W}_j \left(\mathbf{I} - \frac{1}{n(n-1)} \mathbf{1}\mathbf{1}' \right) \quad (7)$$

As Griffith (2009), reported the eigenvector associated with the largest eigenvalue of (7) is the one possessing the highest autocorrelation, and it is orthogonal to the other eigenvectors. Likewise, the eigenvector associated with the second largest eigenvalue is the one possessing the second highest autocorrelation. The remaining eigenvectors can be considered in similar fashion.

The set of all the eigenvectors of (7) can be regarded as distinct and uncorrelated spatial maps, each exhibiting a certain degree of autocorrelation coincident with the corresponding eigenvalue. These spatial configurations are therefore likened to proxy variables depicting all the possible forms of autocorrelation which, starting from matrix \mathbf{W}_j , are latent in the variable subject to analysis. The spatial filtering method uses these artificial indicators as surrogates for the unobservable variables that cause the spatial autocorrelation. For empirical purposes, it is not reasonable to add the full set of eigenvectors as spatial proxy variables to the model (1), but only the dominant eigenvectors. The latter may be chosen in various way. For example in a stepwise procedure regressing the dependent variable \mathbf{f} on the set of the eigenvectors and using the conventional R^2 maximization criterion. As a consequence, the regression model with spatial filtering can be formalized in the following way, where \mathbf{A} is the matrix of the dominant eigenvectors of (7):

$$\mathbf{f} = \alpha \mathbf{1} + \mathbf{X}_O \otimes \mathbf{1} \boldsymbol{\beta}_O + \mathbf{1} \otimes \mathbf{X}_D \boldsymbol{\beta}_D + \mathbf{d} \boldsymbol{\gamma} + \mathbf{A} \boldsymbol{\rho} + \boldsymbol{\varepsilon} \quad (8)$$

Whereas the regression model with spatial filtering does not require particular techniques except for the usual least squares method, in the cases of the SAR and SEM it is necessary to use iterative procedures such as that developed by LeSage (1997).

3. Data and experimentation results

To compare the methods previously exposed, the migration flows for change of residence between the twenty Italian regions collected by the ISTAT for the year 2006 were used and analysed by a gravity model.

In according with the literature, the version applied this paper considers not only the classic determinants of migration (i.e., the size of populations and the distance between places), but also the effects attributable to a set of push and pull variables explaining the economic and demographic differences between the Italian regions. To this regard, we consider 17 indicators (see Appendix). As a preliminary examination of the indexes showed the presence of correlations that rendered them unsuitable for use in a regression model, they have been synthesized by means of factor analysis

The results of this analysis are set out in extreme synthesis in Appendix. The high and positive coefficients of correlation between the first factor and all the variables of economic nature suggest identification of this factor as a complex index of the economic structure, while the close correlations of the second factor with the remaining indexes suggest its identification as a complex index of the demographic structure. Then, the following 18 models were considered:

Table 1 – Models considered in the analysis.

SAR1: $f = X\beta + \rho W_O f + \epsilon$	SAR2: $f = X\beta + \rho W_D f + \epsilon$	SAR3: $f = X\beta + \rho W_{OD} f + \epsilon$
SAR4: $f = X\beta + \rho (W_O + W_D) f + \epsilon$	SAR5: $f = X\beta + \rho (W_O + W_D + W_{OD}) f + \epsilon$	SAR6: $f = X\beta + (\rho_O W_O + \rho_D W_D) f + \epsilon$
SAR7: $f = X\beta + (\rho_O W_O + \rho_D W_D + \rho_{OD} W_{OD}) f + \epsilon$	SAR8: $f = X\beta + (\rho_O W_O + \rho_D W_D + \rho_{OD} W_{OD}) f + \epsilon$	
SEM1: $f = X\beta + \epsilon$ and $\epsilon = \lambda W_O \epsilon + u$	SEM2: $f = X\beta + \epsilon$ and $\epsilon = \lambda W_D \epsilon + u$	SEM3: $f = X\beta + \epsilon$ and $\epsilon = \lambda W_{OD} \epsilon + u$
SEM4: $f = X\beta + \epsilon$ and $\epsilon = \lambda (W_O + W_D) \epsilon + u$	SEM5: $f = X\beta + \epsilon$ and $\epsilon = \lambda (W_O + W_D + W_{OD}) \epsilon + u$	
SF1: $f = X\beta + A_1 \rho + \epsilon$	SF2: $f = X\beta + A_2 \rho + \epsilon$	SF3: $f = X\beta + A_3 \rho + \epsilon$
SF4: $f = X\beta + A_4 \rho + \epsilon$	SF5: $f = X\beta + A_5 \rho + \epsilon$	

where, $X\beta = \beta_0 + \beta_1 p_i + \beta_2 p_j + \beta_3 d + \beta_4 f_1 + \beta_5 f_1_j + \beta_6 f_2 + \beta_7 f_2_j$, and: f is the vector that contain the annual flows between two generic regions; p_i and p_j the vectors that contains the logarithmic amounts of the populations resident respectively in the origin and destination of the flows; d the vector that contain the logarithmic geographical distances between two regions; f_1 , f_1_j and f_2 , f_2_j the vectors that contains the values of first two factors extracted by the factor analysis in the regions of the origin and destination of the flows; A_j (for $j=1, \dots, 5$) are the matrices of the predominant eigenvectors respectively of the W_O , W_D , W_{OD} , $(W_O + W_D)$ and

($\mathbf{W}_O + \mathbf{W}_D + \mathbf{W}_{OD}$) matrices. In this enquiry, the predominant eigenvectors were chosen in a stepwise procedure, regressing, the observed logarithmic of flows on the set of the eigenvectors of the previously \mathbf{W} matrices, using the conventional R^2 as the maximisation criterion. Finally, $\boldsymbol{\varepsilon}$ and \mathbf{u} two residual vectors.

It should be noted that, given the nature of the dependent variable (count data), the linear regression model was chosen instead of the Poisson regression model, because: the data showed the presence of the well-known problem of overdispersion. the investigation was intended to be explanatory, not predictive, the mean flow take a value such that as Baltagi (2011) and Lejenne (2010) reported, the Poisson variable can be well approximated by the normal variable. Finally, the use of this model is in full accordance with the literature. See: Black (1992), Egger (2005), Lewer and Van den Berg (2008), Griffith (2009), Kim and Koen (2010), Mayda (2010), Ludo (2012) and LeSage and Agnan (2015).

A preliminary analysis was conducted with the goal to verifying the presence of spatial correlation in the data. To this regard in the following Table 2 are set out the results of the non-spatial interaction model (i.e. $\mathbf{f} = \mathbf{X}\boldsymbol{\beta}$).

Table 2 – Results of classical log gravity model.

Parameters	Estimates	p-value
Intercept	-21.06	<.001
\mathbf{p}_i	0.96	<.001
\mathbf{p}_j	0.98	<.001
\mathbf{d}	-0.41	<.001
$\mathbf{f1}_i$	-0.11	<.001
$\mathbf{f1}_j$	0.15	<.001
$\mathbf{f2}_i$	-0.09	0.013
$\mathbf{f2}_j$	0.003	0.934
Statistics	Value	
F-Fisher of the model	292.34	<.001
R^2	0.84	
Breush-Pagan test of heteroscedasticity	17.69	0.01
Kolmogorov test of normality	0.046	0.05
Statistics	Value	p-value
Moran (using \mathbf{W}_O)	9.54	<.001
Moran (using \mathbf{W}_D)	9.63	<.001
Moran (using \mathbf{W}_{OD})	10.02	<.001

As is possible to shows the estimates of the constant and the parameters associated with the population size of the regions, as well as the parameter relative to the distance, were highly significant. The estimates of the parameters associated with the economic factor in the regions of origin and in the destination regions of flows ($\mathbf{f1}_i$ and $\mathbf{f1}_j$ respectively) were also significant. When we looked at the

demographic factor of the places of origin ($f2_i$) and of the places of destination ($f2_j$), the estimates of the associated parameters were found to have been non-significant especially for the $f2_j$ factor. Finally, the tests of Breush-Pagan, Kolmogorov-Smirnov, and the Moran Index showed that the regression residuals were respectively: homoscedastic, normally distributed and spatially correlated at the origin and the destination of flows. Consequently, the previously spatial interaction models were considered and estimated. In Appendix are set out the results of this analysis.

First to be noted a particular coincidence in the estimation of the parameters of the various models. However, the estimates are more stable in the various spatial filtering and spatial error models respect to the spatial autoregressive models. Relating the economic factor, to be noted that this always has a negative effect in the origin regions of the flows ($f1_i$). It should be noted that this factor is always significant in the spatial filtering models and in some spatial autoregressive models, while it is never significant in the spatial error models. In the destination regions ($f1_j$), this factor is significant only in the spatial filtering models.

Similar considerations apply to the demographic factor, which in the spatial filtering models, always has a negative estimate in the origin regions of flows ($f2_i$) and a positive estimate in the destination ones ($f2_j$). On the contrary, in the remaining models, $f2_j$ has a negative estimate. However, the effect of the demographic factor is not uniform in the various models. This complex variable is never significant in the destination regions, whilst in the origin regions of the flows, it is generally significant.

To recapitulate, it emerges quite clearly that the three approaches considered and the various types of autocorrelation, do not exhibit either among themselves, substantial differences in the estimates and signs of the parameters. All the models verifies the hypothesis of normal distribution, the SAR models and SEM, generates heteroscedastic residuals. All the models considered are able to capture the effects of the spatial autocorrelation present but the Moran test of autocorrelation shows that the residuals of some SAR models (SAR1, SAR2, SAR4 and SAR5) are still correlated, while those of the SEM and the SF models are still uncorrelated.

The Root Mean Square Error (RMSE) has values between 0.38 (SEM 3) and 0.63 (SAR 7) and the Spatial Filtering models are those with the lowest value in the index considered. They are followed by the Spatial Error Models (SEM) and then by the SAR models. Moreover, the SAR models that involves separately more than one weight matrix (SAR6, SAR7 and SAR8) have a more goodness of fit respect the models that involves a single weight matrix (SAR1, SAR2 and SAR3) and respect those involving cumulative weight matrices (SAR4 and SAR5). On the contrary, the SEM do not shows substantial differences between the various alternatives considered. Finally, the spatial filtering models that involves more than

one weight matrix (SF4 and SF5) have a more goodness of fit. Particularly, the SF4 is the model with the highest goodness of fit.

4. Conclusions

Spatial interaction models of the gravity type are widely used to model origin destination flows. This, model is misspecified if the residuals are spatially correlated. This problem arises when the observations are dependent over the space. To solve the problem same methods were proposed in literature. The most popular approaches are the Spatial Filtering Model, the Spatial Autoregressive Model and Spatial Error Model. Using real data based on entries and cancellation for change of residence between the twenty Italian regions, corresponding to the second level of the Nomenclature of Territorial Units for Statistics (NUTS 2) this enquiry has illustrated and compared the previously three methods. As is obvious the results of the inquiry lay no claim to exhaustiveness indeed, they are to be regarded as only preliminary but they seem very interesting.

The most important results obtained highlights that the three approach do not exhibit substantial differences in the estimation of the parameters and in the goodness of fit with the reality.

Consequently, from this point of view the choice of one or other method is indifferent. On the contrary, analysing the regression residuals, substantial differences are clear. If on the one hand the three approaches generates normally distributed and homoscedastic residuals on the other hand, the residuals of the SAR, in several situations are heteroscedastic and still correlated. The analysis was repeated also for the years 1995 to 2005. The results perfectly coincides with those of the present study.

Appendix

Table 3 – Factorial analysis results.

	Factor 1	Factor 2
Variance explained	0.60	0.19
Correlations between the variables and the first two factors.		
Variables	Factor 1	Factor 2
Employment rate	0.946	0.002
Per capita added value	0.420	0.409
Per person employed added value	0.965	-0.174
GDP per capita	0.895	-0.366
GDP per person employed	0.965	-0.198
% of employed in industry	0.862	-0.416
% of employed in agriculture	-0.836	0.055
% of employed in other activities	0.816	-0.240
Consumption per capita	-0.853	0.215
Income per capita	0.674	-0.607
Units of labour per inhabitant	0.935	0.014
Size of unit of labour	0.275	0.793
Age dependency ratio	-0.930	-0.186
Index of turnover in the active population	0.589	0.707
% of persons aged 65 and over	0.514	0.796
Old-age dependency ratio	-0.010	-0.465
Index of active population structure	0.854	0.340

Table 4 – Results of the SAR Models.

Var.	SAR 1: ρ_{W_O}	SAR2: ρ_{W_D}	SAR3: $\rho_{W_{OD}}$	SAR4: $\rho_{(W_O+W_D)}$	SAR5: $\rho_{(W_O+W_D+W_{OD})}$	SAR6: $\rho_{O}W_O+\rho_D W_D$
Estimates						
Int.	-20.52*	-20.39*	-23.88*	-19.38*	-24.16*	-20.80*
p_i	0.67*	1.08*	1.02*	0.80*	0.98*	0.97*
p_j	1.08*	0.65*	1.03*	0.80*	0.99*	0.93*
d	-0.36*	-0.36*	-0.39*	-0.36*	-0.36*	-0.46*
$f1_i$	-0.06	-0.13*	-0.10*	-0.04	-0.08*	-0.11*
$f1_j$	0.06	0.11*	0.11*	0.09*	0.08	0.14*
$f2_i$	-0.07	-0.08*	-0.08	-0.12*	-0.10*	-0.09*
$f2_j$	-0.01	-0.02	-0.01	-0.06	-0.03	-0.01
ρ	0.31*	0.35*	0.18*	0.47*	0.38*	
ρ_O						0.06
ρ_D						0.01
Statistics						
	Value					
B-P	19.39	22.09	19.21	21.63*	20.83*	5.88
K-N	0.041	0.04	0.04	0.04	0.03	0.04
log lik	-341.18	-337.30	-360.44	-333.22	-348.65	-942.97
RMSE	0.58	0.58	0.62	0.57	0.60	0.62
Moran	3.43*	3.14*	6.28*	8.26*	0.17*	0.83

Legend: B-P= Breush-Pagan test of heteroscedasticity; K-N= Kolmogorov test of normality; log-lik=Log-likelihood; Moran=Moran test of autocorrelation on residuals; RMSE=Root Mean Square Errors; AIC=Akaike Information Criterium; * p-value<0.01.

Table 5 – Results of the SEM. Year 2006.

Var.	SEM1:	SEM2:	SEM3:	SEM4:	SEM5:
	λW_O	λW_D	λW_{OD}	$\lambda (W_O+W_D)$	$\lambda (W_O+W_D+W_{OD})$
	Estimates				
Int.	-20.34*	-20.14*	-20.08*	-19.50*	-19.79*
p_i	0.96*	0.94*	0.94*	0.96*	0.95*
p_j	0.95*	0.96*	0.94*	0.94*	0.94*
d	-0.65*	-0.70*	-0.60*	-1.01*	-0.73*
$f1_i$	-0.04	-0.06	0.004	-0.01	0.01
$f1_j$	0.12*	0.22*	0.17*	0.17	0.15
$f2_i$	-0.11	-0.11*	-0.11*	-0.11	-0.11*
$f2_j$	-0.05	-0.03	-0.06	-0.05	-0.07
λ	0.53*	0.59*	0.64*	0.91*	0.87*
	Statistics				
	Value				
B-P	22.27*	24.43*	23.25*	34.39*	23.09*
K-N	0.06	0.04	0.04	0.04	0.05
log lik	-318.84	-311.23	-335.04	-255.81	-294.15
RMSE	0.53	0.52	0.57	0.43	0.51
Moran	-0.69	-0.56	-3.42*	0.01	-0.003

*Legend: B-P= Breush-Pagan test of heteroscedasticity; K-N= Kolmogorov test of normality; log-lik=Log-likelihood; Moran=Moran test of autocorrelation on residuals; RMSE=Root Mean Square Errors; AIC=Akaike Information Criterium; * p-value<0.01.*

Table 6 – Results of the Spatial Filtering Models. Year 2006

Var.	SF1: W_O	SF2: W_D	SF3: W_{OD}	SF4: W_O+W_D	SF5: $W_O+W_D+W_{OD}$	
	Estimates					
Int.	-21.06*	-21.06*	-21.06*	-21.06*	-21.06*	
p_i	0.96*	0.96*	0.96*	0.96*	0.96*	
p_j	0.98*	0.98*	0.98*	0.98*	0.98*	
d	-0.41*	-0.41*	-0.41*	-0.41*	-0.41*	
$f1_i$	-0.11*	-0.11*	-0.11*	-0.11*	-0.11*	
$f1_j$	0.15*	0.15*	0.15*	0.15*	0.15*	
$f2_i$	-0.09*	-0.09*	-0.09*	-0.09*	-0.09*	
$f2_j$	0.00	0.00	0.00	0.003	0.003	
	Dominant eigenvectors					
	a_3	$1.40^* a_{11}$	$-2.03^* a_4$	$3.34^* a_1$	$-1.14^* a_1$	-1.15^*
	a_8	$1.96^* a_{18}$	$-1.28^* a_5$	$2.96^* a_2$	$0.97^* a_4$	3.56^*
	a_{11}	$1.85^* a_{22}$	$-1.32^* a_{14}$	$-3.04^* a_4$	$3.76^* a_5$	-2.62^*
	a_{15}	$1.67^* a_{27}$	$-3.38^* a_{18}$	$-4.58^* a_8$	$3.38^* a_7$	2.08^*
	a_{27}	$2.50^* a_{38}$	-1.52^*	a_9	$-1.23^* a_9$	1.30^*
	a_{35}	$1.28^* a_{44}$	-1.69^*	a_{12}	$-1.54^* a_{10}$	-1.46^*
	a_{37}	$-1.34^* a_{45}$	1.71^*	a_{14}	$-2.94^* a_{14}$	-3.72^*
	a_{38}	$2.45^* a_{49}$	1.76^*	a_{18}	$2.85^* a_{17}$	-2.59^*
	a_{44}	$1.46^* a_{51}$	-1.52^*	a_{20}	$0.88^* a_{18}$	3.71^*
	a_{45}	$-2.24^* a_{55}$	-1.31^*	a_{23}	$3.76^* a_{28}$	2.31^*
	a_{49}	$1.30^* a_{57}$	-1.42^*	a_{26}	$2.45^* a_{39}$	-2.22^*
	a_{51}	$1.56^* a_{60}$	-1.21^*	a_{28}	1.46^*	
	a_{55}	$1.26^* a_{62}$	-1.34^*	a_{33}	-1.11^*	
	a_{57}	$-1.75^* a_{70}$	-2.36^*	a_{37}	1.45^*	
	a_{73}	$2.07^* a_{71}$	1.38^*	a_{42}	-1.981^*	
	a_{88}	$1.53^* a_{72}$	1.79^*	a_{52}	1.05^*	
		a_{73}	2.52^*	a_{59}	-1.28^*	

	a ₇₅	-1.92*	a ₈₇	-1.48*	
	a ₇₇	1.39*	a ₁₀₇	1.60*	
	a ₈₅	-1.22*	a ₁₂₀	1.66*	
	a ₉₃	1.33*			
Statistics	Value				
F-Fisher	90.85*	105.4*	291.20*	183.7*	233.8*
R ²	0.91	0.92	0.89	0.94	0.92
B-P	67.64	42.84	48.84*	64.37*	23.41
K-N	0.03	0.03	0.05	0.05	0.04
log-lik	-243.42	-222.17	-288.78	-172.04	-238.31
Moran	1.61	1.63	0.54	-0.018	-0.02
RMSE	0.46	0.43	0.52	0.38	0.45

*Legend: B-P= Breush-Pagan test of heteroscedasticity; K-N= Kolmogorov test of normality; log-lik=Log-likelihood; Moran=Moran test of autocorrelation on residuals; RMSE=Root Mean Square Errors; AIC=Akaike Information Criterion; * p-value<0.01.*

References

- ANSELIN L. 1988. *Spatial Econometrics: methods and model*. Dondrecht, Kluwer Academic Publishers.
- BALTAGI H.B. 2011. *Econometrics 5th edition*. Berlin/Heidelberg: Springer.
- BLACK W.R. 1992. Network autocorrelation in transport network and flow system, *Geographical Analysis*, Vol. 24, No. 3, pp. 207-222.
- CHUN Y., GRIFFITH D.A. 2011. Modeling network autocorrelation in space-time migration flow data: an eigenvector spatial filtering approach, *Annals of the association of American Geographer*, Vol. 101, No. 3, pp. 523-536.
- DE JONG P., SPRENGER C., VAN VEEN F. 1984. On the extreme value of Moran's I and Gary's c, *Geographical Analysis*, Vol. 16, No. 1, pp. 17-24.
- EVENETT S.J., KELLER W. 2002. On the Theories Explaining the Success of the Gravity Equation, *Journal of Political Economy*, Vol. 110, pp. 281-316.
- FISCHER M.M., GRIFFITH D.A. 2008a. Modelling spatial autocorrelation in spatial interaction data: An application to patent citation in the European Union, *Journal of Regional Science*, Vol. 48, No. 5, pp.969-89.
- FORREST R., MURIE A. 1990. Moving strategies among home owners. In JOHNSON J. H. and SALT J. (Eds.) *Labour Migration: the internal geographical mobility of labour in the developed world*, London David Fulton Publisher, pp. 191-209.
- GRIFFITH D.A. 2003. *Spatial autocorrelation and spatial filtering*. Springer-Verlang, Berlin.

- GRIFFITH D.A. 2009. Modelling spatial autocorrelation in spatial interaction data: Empirical evidence from 2002 Germany journey-to-work flows, *Journal of Geographical System*, Vol. 11, pp. 117-40.
- KIM K., KOEN E.J. 2010. Determinants of international migration flows to and from industrialized countries: A panel data approach beyond gravity, *International Migration Review*, Vol. 44, No. 4, pp. 899-932.
- LEJENNE M. 2010. *Statistique la théorie e ses application*. Paris, Springer.
- LESAGE J.P. 1997a. Regression analysis of spatial data, *Regional Analysis and Policy*, Vol. 27, No. 2, pp. 83-94.
- LESAGE J.P., PACE R.K. 2008. Spatial econometric modeling of origin-destination flows, *Journal of Regional Science*, Vol. 48, No. 5, pp. 941-967.
- LESAGE J.P., AGNAN C.T. 2015. Interpreting spatial econometric origin-destination flow models, *Journal of Regional Science*, Vol. 55, No. 2, pp. 188-208.
- LEWER J.J., VAN DEN BERG H. 2008. A gravity model of immigration, *Economic Letters*, Vol. 99, pp. 164-167.
- LUDO P. 2012. Gravity and spatial structure: The case of interstate migration in Mexico, *Journal of Regional Science*, Vol. 52, No. 5, pp. 819-856.
- MAYDA A.M. 2010. International migration: A panel data analysis of the determinant of bilateral flows, *Journal of Population Economics*, Vol. 23, pp. 1249-1274.

SUMMARY

An analysis of methods for the treatment of autocorrelation in spatial interaction models

Using real data, this paper sets out the results of an analysis of the methods developed in the literature for the treatment of spatial autocorrelation in the spatial interactions models. In particular, the inquiry compares the autoregressive method (SAR), the method with autoregressive errors (SEM), and the spatial filtering method (SF). The results shows a substantial uniform behaviour among the considered approaches but the residuals of the SAR in same situations remains spatial correlated.

Giuseppe RICCIARDO LAMONICA, Polytechnic University of Marche – Faculty of Economics “G. Fuà”, g.ricciardo@staff.univpm.it

**DIFFERENCES IN THE ENTERPRISES' STRATA
CLASSIFICATION BETWEEN SAMPLING AND CALIBRATION
SOURCES:
THE IMPLICATIONS FOR ISTAT QUARTERLY INDICATORS ON
JOB VACANCIES AND HOURS WORKED**

Diego Chianella, Giuliano Latini, Annalisa Lucarelli,
Emilia Matera, Marina Sorrentino

1. Introduction

This work is the result of the reweighting and calibration process used to obtain the quarterly hours worked and job vacancies indicators.

Istat disseminates hours worked and job vacancies indicators on a quarterly basis.

To produce these indicators, data on the target variables collected by two surveys are used together with those on jobs of an auxiliary administrative source (OROS) which are employed for editing and imputation and calibration. The first survey is the ISTAT quarterly survey on job vacancies and hours worked (VELA), based on a stratified random sample of about 26,000 enterprises with less than 500 employees in Nace Rev. 2 sections B to S. The sample is drawn from the Istat register of active enterprises (ASIA), updated to two years before the survey reference period. The stratification variables are economic activity, size and geographical area (Cochran 1977; Conti and Marella 2011). The second survey is the ISTAT monthly survey on employment, hours worked, wages and labour costs in large enterprises (GI), which is a census survey of around 1,400 enterprises with at least 500 employees in sections B to S.

A calibration estimator is used to obtain the quarterly estimates, and the calibration constraints within each calibration domain are derived from OROS for the reference quarter. The calibration domains are obtained aggregating one or more sampling strata. The aggregation usually takes place for different geographical areas and size classes within the same economic activity, but it can involve also this third variable.

Due to the different reference periods of ASIA and OROS, enterprises may be classified in different calibration domains depending on whether we consider the information in ASIA or in OROS for the reference quarter, as a consequence of

changes in size and/or economic activity over time. These movements across domains, if not properly handled, can result in high calibration weights variability within the calibration domains.

The purpose of this paper is to present and explain the methodology developed in order to handle these movements across domains, between the sampling design and calibration population, due to the different reference periods, reducing the weights variability within the calibration domains and standard errors of the estimates.

2. The estimation procedure

A calibration estimator (Deville and Särndal 1992; Särndal 2007; Bacchini *et al.* 2014) is used to obtain Italian hours worked and job vacancies estimates. The general procedure is the following:

1. First, design weights δ_j , are obtained as the inverse of the inclusion probabilities of any enterprise in the sample.
2. Then, correction factors k_j , are worked out as the reciprocal of the response rates, in order to take into account enterprises non-response. Intermediate weights, corrected for non-response, are then computed multiplying design weights by these correction factors $d_j = \delta_j * k_j$.
3. Then, starting from intermediate weights d_j , final weights w_j are obtained solving a constrained minimization problem. The function to be minimized is a (logit, in this case) distance between final and intermediate weights. The constraints are based on the number of jobs in the calibration domains, as measured by OROS for the reference quarter.

The sampling design is simple random without replacement (srswo), so the design weights δ_j are given by the ratio between the number of units in the population and the number of units in the sample within each stratum h ($\forall h = 1, \dots, l$):

$$\delta_j = \frac{N_h}{n_h^s}$$

where n_h^s is strictly smaller than and equal to N_h respectively in VELA and GI.

Design weights are corrected with the reciprocal of the response rates in each calibration domain H ($\forall H = 1, \dots, L$ where $L \leq l$):

$$k_j = \frac{1}{n_H^r/n_H^s}$$

where n_H^r is the number of respondents in the calibration domain H. Hence, the intermediate weights are obtained as:

$$d_j = \delta_j * k_j = \frac{N_h}{n_h^s} * \frac{1}{n_H^r/n_H^s}$$

Therefore, intermediate weights are the product between design weights depending on the sampling information (based on ASIA) and the reciprocal of the response rates in the calibration domains, depending on the more updated information from OROS.

Calibration domains are obtained aggregating one or more sampling strata, and a single stratum h belongs to a single calibration domain H, as shown in Table 1.

Table 1 - Correspondence between strata and calibration domains, based on Asia only

Strata	Calibration Domains	Design Weights
h_1		N_1/n_1
h_2	H_1	N_2/n_2
h_3		N_3/n_3
h_4	H_2	N_4/n_4
h_5		N_5/n_5
...
h_{l-1}	H_L	N_{l-1}/n_{l-1}
h_l		N_l/n_l

The different sources used for sampling and calibration and the distance between the two sources reference periods imply that an enterprise can be classified differently according to the two sources, if its stratification variable values have changed over time. As a consequence, the original correspondence between a design stratum h and a calibration domain H may be broken, because an enterprise can be assigned during the estimation phase to a calibration domain that is not obtained as an aggregation of strata including the one in which the enterprise

was classified on the basis of Asia. In the example in Table 2, considering the OROS updated information for the enterprises' stratification variables, some enterprises belonging to the h_1 stratum according to ASIA can be classified within calibration domain H_1 , while others in H_2 . This increases the weights variability within calibration domains. In the example below the weights variability increases within the calibration domain H_2 .

Table 2 - Correspondence between strata and calibration domains, considering the updated information for the calibration domains

Strata	Calibration Domains	Design Weights
h_1		N_1/n_1
h_2	H_1	N_2/n_2
h_3		N_3/n_3
h_1		N_1/n_1
h_4	H_2	N_4/n_4
h_5		N_5/n_5
...
h_{l-1}		N_{l-1}/n_{l-1}
h_l	H_L	N_l/n_l

The problem mainly affects the enterprises that in the estimation phase are classified in census calibration domains, while in the sampling phase were classified in strata with inclusion probabilities strictly lower than one. The different classification implies that in some census calibration domains there can be most enterprises with design weight equal to one and a small number of enterprises coming from sample strata with larger design weights. In this situation, the final weights can be often near to zero for the enterprises that had design weight equal to one, while much larger weights are attributed to the enterprises with larger design weights.

3. Design weights corrections to adjust for enterprise changes

3.1 Old reweighting procedure

Until the estimates for the fourth quarter 2017, to solve the problem of the different stratum classification over time, the design weights were calculated at the moment of the quarterly estimation, at the level of calibration domains rather than design strata and considering only the information contained in OROS (Bellisai *et al.* 2013). Thus, the design weights were:

$$\delta_j = N_H/n_H^s.$$

Design weights were corrected with the reciprocal of the response rates in each calibration domain H, with the following correction factors:

$$k_j = \frac{1}{n_H^r/n_H^s},$$

where n_H^r is the number of respondents in the calibration domain H. Therefore, the intermediate weights were equal to:

$$d_j = \delta_j k_j = N_H/n_H^r$$

and no problem described at the end of Section 2 and related to the change in the variables stratification values over time did not occur. However, this procedure implies significant differences between the thus calculated design weights and the original inclusion probabilities within design strata. A revised procedure was therefore studied and implemented, which is closer to the one described in Section 2.

3.2 A new reweighting procedure

Starting from the estimates for the first quarter of 2018 and the revisions of the three previous years, the formula for the design weights has been revised. In particular, the formula described in Section 2 has been used, in which the weights are calculated within the sampling stratum h , but with the difference that the information on the stratification variables, concerning only the enterprises in the sample, are updated with OROS:

$$\delta_j = \frac{N_h^{updated}}{n_h^{s,updated}}$$

Hence, the values of numerator and denominator are very similar to the N_h and n_h^s that would be calculated if only the sampling design and ASIA information were used. While the j -th enterprise in the sample is classified as belonging to a stratum, on the basis of the most recent information about the stratification variable values available in OROS.

As a consequence, the correspondence between strata and calibration domains is maintained as in Table 1.

4. Empirical results: weights variability and standard error estimates

The impact of the reclassification on the basis of OROS of only the enterprises in the sample, can be highlighted by a comparison of the final weights distributions obtained applying the procedures described in Section 2 and in Section 3.2.

Table 3 - Percentile distribution of the final weights – 4th quarter 2015 estimates

Level	No correction of the design weights	New reweighting procedure
100% - Max	80.718	83.122
99%	60.351	58.588
95%	41.740	41.020
90%	31.979	32.348
75% - Q3	21.152	20.378
50% - Median	10.853	10.540
25% - Q1	5.877	5.605
10%	1.046	1.122
5%	0.904	1.050
1%	0.466	1.000
0% - Min	0.001	0.578

Source: ISTAT, VELA

In Table 3, the comparison is made on the estimates for the fourth quarter 2015. It can be observed that if the formula in Section 2 is applied, the final weights at the first percentiles are much lower than if the respondent enterprises are reclassified according to OROS. In fact, without correction of the design weights

for the new information on respondents, the lowest final weights values are close to zero, and between 5% and 10% are lower than one. With the new reweighting procedure, the minimum value of the final weights is 0.57 and already at the first percentile we have weights equals to one.

Table 4 - A real example of the weights distribution within a census calibration domain

Calibrat. Domain	X	No correction of the design weights			Old reweighting procedure			New reweighting procedure		
		Strata	Design weight	Final weight	Strata	Design weight	Final weight	Strata	Design weight	Final weight
DOM2	304	A-1	8.325	4.981	DOM2	1	1	A-2	1	1
DOM2	350	A-1	8.325	4.561	DOM2	1	1	A-2	1	1
DOM2	420	A-1	8.325	3.992	DOM2	1	1	A-2	1	1
DOM2	483	A-1	8.325	3.524	DOM2	1	1	A-2	1	1
DOM2	520	A-2	1	0.385	DOM2	1	1	A-2	1	1
DOM2	527	A-2	1	0.38	DOM2	1	1	A-2	1	1
DOM2	586	A-2	1	0.337	DOM2	1	1	A-2	1	1
DOM2	633	A-2	1	0.307	DOM2	1	1	A-2	1	1
DOM2	637	A-2	1	0.304	DOM2	1	1	A-2	1	1
DOM2	707	A-2	1	0.264	DOM2	1	1	A-2	1	1
DOM2	728	A-2	1	0.252	DOM2	1	1	A-2	1	1
DOM2	760	A-1	8.325	0.236	DOM2	1	1	A-2	1	1
DOM2	1445	A-2	1	0.053	DOM2	1	1	A-2	1	1

Source: ISTAT, VELA

Although the constrained minimization problem is solved in both cases, the second case is preferable, because it allows the estimates to represent all the respondent enterprises. This is even more relevant in the light of the fact that the final weights with a value similar to zero are recorded in the largest dimensional class, corresponding to the census strata, with initial weights equal to 1. In Table 4, we have a real example of the weight distribution for all enterprises within a generic census calibration domain (DOM2).

X is the calibration variable (jobs). Stratum is the aggregation of economic activity (A) and dimensional class (1 for sample class and 2 for census class). As we can see, in the basic method without correction of the design weights, we have some units that, with the updated information, are classified in the census domain, although they were classified in the sample stratum, at the time of the sampling.

Hence, for these units we have an initial weight larger than one. This means that to meet the calibration constraints on the total of X while minimizing the distance between initial and final weights, the enterprises in the census stratum (A-2) will have a final weight much smaller than one. This can be a problem, for example when estimating job vacancies. In fact, on a given reference date, most respondent enterprises have no vacancies, and the probability of an enterprise having at least a vacancy and the number of vacancies increase with the enterprise size. In the example, larger enterprises, which are more likely to have job vacancies, are assigned a final weight near to zero if no correction of the design weights is implemented. Moreover, with the old reweighting procedure, for calibration purposes, the stratum is defined as equal to the calibration domain. Therefore, final weights are equal to one for all enterprises. The same weight distribution is generated by the new reweighting procedure. In the analysed case, in the calibration domain DOM2, with the new procedure all enterprises are classified in stratum A-2.

Table 5 - A real example of the weights distribution within a sample calibration domain

Calibrat. Domain	X	No correction of the design weights			Old reweighting procedure			New reweighting procedure		
		Strata	Design weight	Final weight	Strata	Design weight	Final weight	Strata	Design weight	Final weight
DOM1	479	A1-1	9.5	13.7	DOM1	14.2	8.7	A1-1	10.3	13.7
DOM1	346	A1-1	9.5	12.8	DOM1	14.2	10.1	A1-1	10.3	12.8
DOM1	236	A1-1	9.5	12.1	DOM1	14.2	11.3	A1-1	10.3	12
DOM1	161	A1-1	9.5	11.6	DOM1	14.2	12.2	A1-1	10.3	11.4
DOM1	125	A1-1	9.5	11.3	DOM1	14.2	12.6	A1-1	10.3	11.2
DOM1	194	A2-1	7.3	8.3	DOM1	14.2	11.8	A2-1	7.5	8.4
DOM1	158	A2-1	7.3	8.2	DOM1	14.2	12.2	A2-1	7.5	8.3
DOM1	139	A2-1	7.3	8.1	DOM1	14.2	12.5	A2-1	7.5	8.2
DOM1	146	A3-1	3.1	3.3	DOM1	14.2	12.4	A3-1	3.1	3.4

Source: ISTAT, VELA

However, usually each calibration domain is an aggregation of more than one stratum. This is the case represented in Table 5.

In this case, we have a sample calibration domain (DOM1), obtained aggregating strata across economic activity. With the old reweighting procedure, design weights are very different from the inverse of the inclusion probabilities, because they were recalculated at the calibration domain level rather than at the design stratum one. On the other hand, with the new reweighting procedure, initial

weights are similar to those obtained without updating the classification information for respondent enterprises.

The coefficients of variation (CVs) of the estimates of the job vacancy rate and hours worked per capita obtained with the old and the new reweighting procedure for the four quarters 2016 are presented in Table 6. The CVs with the new method tend to be lower and more stable over time, in particular for the estimates of the job vacancy rate (Zardetto 2015).

Table 6 - CVs for the total of the economic activities – Year 2016

Year 2016 Quarter	Job vancancy rate		Hours worked per capita	
	Old procedure	New procedure	Old procedure	New procedure
1	0.042	0.053	0.003	0.003
2	0.193	0.059	0.003	0.003
3	0.047	0.048	0.003	0.003
4	0.072	0.044	0.003	0.003

Source: ISTAT, VELA

5. Conclusions

The new reweighting procedure used for the job vacancy and hours worked estimation takes into account changes in the size class and economic activity of the sample enterprises, between the sampling and estimation phases. The results show less variability of the final weights compared to the method that does not reclassify the sample enterprises with the information available at the time of the calibration and greater or similar accuracy of the estimates of the target variables, as measured by their standard errors.

6. References

- BACCHINI F., CECCARELLI C., CHIANELLA D. & IANNACCONE R. 2014. Improving quarterly index of turnover by means of a calibration estimator, *Rieds*, Vol. LXVIII, pp. 15-22.
- BELLISAI D, FIVIZZANI S. & SORRENTINO M. 2013. A business survey on job vacancies: integration with other sources and calibration. In DAVINO C., FABBRIS L. (Eds) *Survey Data Collection and Integration*, Heidelberg: Springer-Verlag, pp. 139-152.

- COCHRAN W. 1977. *Sampling Techniques. Third edition*. New York: Wiley & Sons.
- CONTI P., MARELLA D. 2011. *Campionamento da popolazioni finite: teoria e tecnica*. Milano: Springer.
- DEVILLE J.C., SÄRNDAL C.E. 1992. Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, Vol. 87, pp. 376–382.
- SÄRNDAL C.E. 2007. The calibration approach in survey theory and practice, *Survey Methodology*, Vol. 33, No. 2, pp. 99–119.
- ZARDETTO D. 2015. ReGenesees: an Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys, *Journal of Official Statistics*, Vol. 31, No. 2, pp. 177–203.

SUMMARY

Differences in the enterprises' strata classification between sampling and calibration sources: the implications for Istat quarterly indicators on job vacancies and hours worked

The Italian National Institute of Statistics produces quarterly estimates of the job vacancy rate and hours worked for EU Regulations and national dissemination through calibrated weights. Job vacancy and hours worked data are collected by two direct business surveys. The survey samples are drawn from the Italian statistical business register (ASIA), updated to two years before the reference quarter of the survey data, while the calibration constraints are derived from an administrative based source for the reference quarter.

The distance in time between the reference periods of sampling frame and source for calibration constraints implies that sample units may be classified as belonging to different strata depending on whether the information in the sample frame or in the source used for calibration constraints is used, due to changes in the enterprises' economic activity and/or size between the two reference periods.

As a consequence, if the initial weights are based on the sample frame units classification, calibrated weights are found to show a high variability within some calibration strata. In particular, in a given calibration stratum, weights can be concentrated only on few sample units while the remaining ones can be near to zero.

To ensure a more homogeneous distribution of the calibration weights within strata, the initial weights have been recalculated, updating the sample units classification with the information used in the calibration strata construction and based on the source used for the calibration constraints.

Furthermore, the initial weights are calculated so as to correct also for non-response by the inverse of the response rate in the calibration stratum.

The first empirical results obtained applying this method have shown a reduction in the estimated sampling variance with respect to those obtained using initial weights based only on the not updated information of the sampling frame.

Diego CHIANELLA, ISTAT, chianella@istat.it
Giuliano LATINI, ISTAT, latini@istat.it
Annalisa LUCARELLI, ISTAT, anlucare@istat.it
Emila MATERA, ISTAT, ematera@istat.it
Marina SORRENTINO, ISTAT, mrsorren@istat.it

L'IMPATTO DEI VINCOLI DI CALIBRAZIONE SULL'ERRORE CAMPIONARIO NEL WAVE APPROACH

Miriam De Santis, Alessio Guandalini, Antonella Iorio, Alessandra Lugli

Introduzione

L'indagine continua sulle Forze di Lavoro condotta in Italia (FoL) è la principale fonte statistica sul mercato del lavoro italiano (Istat, 2006). L'approccio metodologico utilizzato per la produzione di stime dell'indagine FoL tiene conto della coerenza rispetto ad alcuni aggregati di popolazione noti principalmente da fonte amministrativa (ad esempio, tra gli altri, popolazione per età, sesso e stato civile). Per la produzione delle stime si ricorre, dunque, allo stimatore di calibrazione che soddisfa la proprietà di coerenza con un set di totali opportunamente scelti (Deville & Särndal, 1992; Särndal, 2007).

Eurostat, nel programma di modernizzazione, al fine di ridurre il carico sui rispondenti, ha proposto un nuovo approccio basato sulla modularizzazione dell'indagine. Questo approccio, che va sotto il nome di Wave Approach (WA), prevede che, per alcuni quesiti, i dati siano rilevati solo su un sotto-campione. Poi per garantire maggiore robustezza e coerenza, il sotto-campione deve essere vincolato alle stime sulla condizione occupazionale stimate sull'intero campione annuale della stessa indagine FoL.

L'applicazione del WA apre diversi interrogativi dal punto di vista metodologico: che impatto ha sulle stime imporre la coerenza su altre stime ottenute utilizzando un campione di cui lo stesso sotto-campione fa parte? L'impatto si limita alle sole stime o interessa anche il relativo errore campionario? L'impatto sull'errore campionario è positivo o negativo? Come si può valutare in maniera opportuna l'errore campionario nel caso in cui si applichi il WA all'indagine FoL?

Il presente lavoro si pone l'obiettivo di rispondere a questi quesiti. Il lavoro è strutturato come segue: nel paragrafo 1 viene brevemente illustrato il disegno campionario ed il processo di stima dell'indagine FoL, mentre nel paragrafo 2 viene presentato in maniera più dettagliata il WA e la sperimentazione effettuata sul modulo *ad hoc* "Garanzia Giovani". Nel paragrafo 3.1 viene richiamata brevemente la metodologia alla base dello stimatore di calibrazione e nel paragrafo

3.2 l'espressione di uno stimatore della varianza campionaria adatto al caso del WA in FoL viene illustrato. Nel paragrafo 4 sono presentati alcuni risultati della sperimentazione effettuata, mentre nel paragrafo 5 sono discusse le conclusioni.

1. L'indagine continua sulle Forze di Lavoro condotta in Italia (FoL)

L'indagine sulle Forze di Lavoro condotta in Italia (FoL) è una rilevazione continua effettuata in tutte le settimane dell'anno e rappresenta la fonte ufficiale del mercato del lavoro in Italia.

Il principale obiettivo dell'indagine FoL è la produzione delle stime ufficiali degli occupati, delle persone in cerca di occupazione e degli inattivi. Questi tre aggregati di interesse rappresentano una partizione della popolazione in età lavorativa (15 anni e oltre) in tre gruppi esaustivi e mutuamente esclusivi. Le stime di questi tre aggregati vengono pubblicate con cadenza trimestrale e con dettaglio regionale, ma, per i principali parametri, vengono prodotte anche delle stime annuali con dettaglio provinciale e delle stime mensili a livello nazionale.

Figura 1 – Schema di rotazione dell'indagine continua sulle Forze di Lavoro (FoL) per l'anno 2015.

Anno		2015			
Trimestre		1	2	3	4
Gruppo di Rotazione	A	A4			
	B	B3	B4		
	C		C3	C4	
	D			D3	D4
	E	E2			E3
	F	F1	F2		
	G		G1	G2	
	H			H1	H2
	I				I1

Il disegno campionario per ciascuna rilevazione trimestrale è di tipo complesso a due stadi (comuni-famiglie) con stratificazione delle unità di primo stadio a livello provinciale in base all'ampiezza demografica e prevede uno schema di rotazione del campione del tipo $2_T - 2_T - 2_T$.

Lo schema di rotazione $2_T - 2_T - 2_T$ vuol dire che ciascuna famiglia rimane nella rilevazione per 18 mesi consecutivi: viene intervistata per due trimestri consecutivi¹, poi esce temporaneamente dal campione per due trimestri e, infine, rientra per essere intervistata per gli ultimi due trimestri consecutivi. La rotazione dei campioni (Figura 1) fa sì che i campioni trimestrali siano parzialmente

¹ Ciascuna famiglia viene intervistata una sola volta in una specifica settimana del trimestre (sempre la stessa per tutti i trimestri in cui la famiglia entra nel campione).

sovrapposti (50% a un trimestre di distanza, 25% a tre trimestri di distanza, 50% a quattro trimestri e 25% a cinque trimestri). Inoltre, il campione di ogni trimestre è composto da quattro diversi gruppi di famiglie (gruppi di rotazione) che si trovano in occasioni diverse di rilevazione (*wave*).

I quesiti vengono sottoposti all'intero campione trimestrale. Lo stimatore utilizzato per le procedure di stima è lo stimatore calibrato (Deville & Särndal, 1992; Särndal, 2007) che verrà illustrato in dettaglio nel paragrafo 3.

Le normative Europee (E.C. n. 577/1998 del 9 marzo 1998) richiedono che le stime dell'indagine FoL siano coerenti con degli ammontare di popolazione noti da fonte amministrativa (detti anche *vincoli*). L'Istituto Nazionale di Statistica (Istat), per soddisfare specifici bisogni nazionali, ha preso in considerazione anche altri vincoli. Nel complesso il numero di vincoli applicati nella calibrazione trimestrale è di 283 totali. In particolare:

- distribuzione della popolazione per sesso e quattordici fasce di età (0-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75 e più), a livello NUTS II;
- distribuzione della popolazione straniera per sesso e cittadinanza (UE, non UE), a livello NUTS II;
- numero di famiglie per ciascun gruppo di rotazione (1/4 del totale), a livello NUTS II;
- distribuzione della popolazione per sesso e cinque fasce di età (0-14, 15-29, 30-49, 50-64, 65 e più), a livello NUTS III;
- distribuzione della popolazione per sesso e cinque fasce di età (0-14, 15-29, 30-49, 50-64, 65 e più) per i tredici grandi comuni (comuni con più di 250.000 abitanti);
- distribuzione della popolazione per sesso, per ciascun mese del trimestre, a livello NUTS II.

Nel seguito si farà riferimento a questo sistema di vincoli come S_{283} .

2. Il Wave Approach (WA) ed una sua prima applicazione nell'indagine FoL

Eurostat, nel programma di modernizzazione, al fine di ridurre il carico sui rispondenti, ha proposto un nuovo approccio basato sulla modularizzazione dell'indagine FoL. Questo approccio, che va sotto il nome di wave approach (WA), prevede l'utilizzo solo di una parte del campione per produrre stime annuali.

In particolare, il WA applicato all'indagine FoL prevede che alcune domande siano somministrate solamente ai gruppi entranti nei trimestri dell'anno, ovvero solo i gruppi di rotazione in prima *wave*. Quindi, con riferimento alla Figura 1; il WA riguarda i gruppi di rotazione F, G, H, I evidenziati in rosso. Il sotto-campione che ne deriva è, dunque, costituito da osservazioni indipendenti – ciascun gruppo di

rotazione è estratto in maniera indipendente rispetto agli altri e non ha unità campionarie in comune con gli altri gruppi di rotazione – ed ha la stessa dimensione del campione teorico trimestrale – è composto da quattro gruppi di rotazione come il campione trimestrale.

Per garantire maggiore robustezza e coerenza delle stime le normative Europee (E.C. n. 377/2008 del 25 aprile 2008) richiedono che il sotto-campione riproduca, come minimo, alcune distribuzioni per condizione occupazionale (occupati, persone in cerca di occupazione e inattivi), sesso e classi d'età (15-24, 25-34, 35-44, 45-54, 55 e più). Questo, tuttavia, potrebbe non essere sufficiente a ottenere risultati pienamente coerenti, quindi l'Istat ha deciso di includere, in fase di stima, un totale di 62 nuovi vincoli questa volta non provenienti da fonte amministrativa, ma stimati sull'intero campione annuale dell'indagine. Ovvero:

- distribuzione della popolazione per condizione lavorativa, sesso e cinque fasce di età (15-24, 25-34, 35-44, 45-54, 55 e più anni);
- distribuzione degli occupati per carattere dell'occupazione (tempo determinato, tempo indeterminato), sesso e cinque fasce di età (15-24, 25-34, 35-44, 45-54, 55 e più);
- distribuzione della popolazione 15 e più e 15-29 anni per sesso e condizione lavorativa;

Nel seguito questo sistema di vincoli, che tiene conto di questi 62 vincoli stimati sull'intero campione FoL, insieme ai vincoli già illustrati nel sistema S_{283} , sarà chiamato S_{345} .

Facendo sempre riferimento alla Figura 1, si può vedere come il campione annuale relativo all'anno 2015 è composto dal gruppo di rotazione A in quarta wave (A4), il gruppo di rotazione B, C e D in terza e quarta wave (B3, B4, C3, C4, D3, D4), il gruppo di rotazione E in seconda e terza wave (E2, E3), il gruppo di rotazione F, G e H in prima e seconda wave (F1, F2, G1, G2, H1, H2) ed infine il gruppo di rotazione I in prima wave (I1). Le stime ottenute sul campione annuale quindi sono basate anche sui gruppi di rotazione in prima wave (F1, G1, H1 e I1) che costituiscono il sotto-campione su cui il WA viene applicato. In questo senso il sotto-campione del WA dovrà riprodurre le stime che sono state ottenute sul campione annuale di cui anche lui fa parte². In pratica, il campione annuale ed il sotto-campione del WA non sono indipendenti.

Attualmente non sono ancora prodotte stime dall'indagine FoL utilizzando il WA. La prima applicazione ufficiale riguarderà il modulo *ad hoc* 2018 sulla "Riconciliazione tra lavoro e vita familiare".

² In pratica ¼ del campione annuale è costituito dal sotto-campione considerato nel WA.

L'Istat, nel frattempo, ha deciso di svolgere un test per valutare l'impatto che questo approccio può avere sulle stime in termini di precisione ed efficienza in modo da riuscire a prevedere l'impatto sulle stime e predisporre una metodologia opportuna per il calcolo dell'errore campionario.

Questo test ha riguardato le stime derivanti da una batteria di nove domande che sono presenti nel questionario FoL a partire dal primo trimestre 2015 relative al programma "Garanzia Giovani"³. Il programma "Garanzia Giovani" è stato introdotto dal Governo italiano all'interno dell'intervento legislativo noto come "Jobs Act" per combattere la disoccupazione giovanile. Le domande contenute in questo modulo hanno l'obiettivo di fornire indicazioni circa la conoscenza ed efficacia del programma "Garanzia Giovani" così come richiesto dalla Comunità Europea.

La popolazione di interesse, relativamente alle domande del modulo "Garanzia Giovani", è costituita da individui di età compresa tra 15 e 29 anni definiti NEET (*Not engaged in Education, Employment or Training*), cioè giovani che non lavorano e non studiano, residenti in Italia, cittadini UE o non comunitari, regolarmente soggiornanti.

3. Lo stimatore calibrato nel Wave Approach

Come è stato illustrato nel paragrafo precedente nel WA, per ridurre il carico sui rispondenti, alcuni quesiti vengono rilevati solo su un sotto-campione di famiglie. Per dare maggiore robustezza e coerenza alle stime, attraverso lo stimatore di calibrazione (1992), si impone che alcuni ammontare di popolazione stimati sul campione annuale siano riprodotte dal sotto-campione del WA. Il campione annuale ed il sotto-campione del WA, però, non sono indipendenti in quanto è composto in parte dalle stesse famiglie che compongono il sotto-campione del WA. Di questo aspetto si deve tener conto in fase di calcolo della varianza campionaria soprattutto per valutare opportunamente la bontà delle stime ottenute.

Il WA non richiede accorgimenti metodologici per quanto riguarda la produzione delle stime con lo stimatore di calibrazione. La complessità è, invece, rappresentata dal calcolo della varianza campionaria. Infatti, con il WA si rende necessario tener conto del fatto che alcuni dei vincoli inseriti nel sistema di calibrazione sono a loro volta stime. L'aspetto più importante, però, è rappresentato dal fatto che queste stime utilizzate come vincoli sono ottenute su un campione che in parte è formato dalle stesse unità del sotto-campione del WA.

³ La "Garanzia Giovani" è il Piano Europeo per la lotta alla disoccupazione giovanile. In sostanza l'Unione Europea fornisce fondi specifici per gli Stati membri con tassi di disoccupazione superiori al 25%. Questi fondi devono essere investiti da ogni Paese con attività di formazione, sostegno e aiuti per l'inserimento nel mondo del lavoro (European Council Recommendation 2013/C, 120/01).

Per questo motivo nel paragrafo 3.1 sarà illustrato lo stimatore di calibrazione nella sua definizione originale, mentre nel paragrafo 3.2 l'attenzione sarà concentrata su degli stimatori della varianza in grado di tener conto della particolarità del WA applicato all'indagine FoL.

3.1. La definizione originale dello stimatore di calibrazione

Lo stimatore di calibrazione è stata proposto da Deville e Särndal (1992) in base all'idea che se sono disponibili variabili ausiliarie fortemente correlate con le variabili di interesse rilevate sul campione, per cui si conoscono dei totali di popolazione da fonti amministrative e si è in grado di individuare un sistema di pesi che si "comportano bene" per le variabili ausiliarie, allora questo sistema di pesi dovrebbe "comportarsi bene" anche per le variabili di studio.

Si consideri una popolazione finita $U = \{1, \dots, k, \dots, N\}$. Sia il campione s di dimensione fissa n selezionato dalla popolazione U secondo un disegno di campionamento $(S, p(\cdot))$, dove S è lo spazio campionario e $p(\cdot)$ è una distribuzione di probabilità su S . La probabilità di inclusione del primo ordine, $\pi_k = Pr(k \in s)$, e la probabilità di inclusione del secondo ordine, $\pi_{kl} = Pr(k, l \in s)$, si presume siano note e positive. Poniamo $\pi_{kk} = \pi_k$.

Supponiamo di essere interessati a stimare il totale della variabile Y , $t_Y = \sum_{k \in U} y_k$ e di aver rilevato su ogni unità $k \in s$, oltre al valore delle variabili di interesse y_k , i valori di un vettore ausiliario, $x_k = (x_{k1} \dots x_{kp} \dots x_{kP})$.

Inoltre, i totali di popolazione di X_s , $t_{X_p} = \sum_{k \in U} x_{kp}$ con $p = 1, \dots, P$, sono noti senza errore di campionamento.

Deville e Särndal hanno dimostrato che risolvendo il problema

$$\left\{ \begin{array}{l} \min_{w_k} \left\{ \sum_{k \in s} G(w_k, d_k) / q_k \right\} \\ \sum_{k \in s} w_k x_k = t_x \end{array} \right\},$$

se questo ha soluzione, si individua un sistema di pesi, w_k , calibrati (coerenti) con i totali noti che possono essere utilizzati per ottenere lo stimatore calibrato (CAL):

$$\hat{t}_{Y_{CAL}} = \sum_{k \in s} y_k w_k.$$

I pesi w_k sono, inoltre, i più vicini possibile ai pesi del disegno campionario, $d_k = 1/\pi_k$, rispetto ad una funzione di distanza $G(\cdot)$.

Nel caso in cui $G(\cdot)$ è la distanza dal chi quadro, lo stimatore di calibrazione che si ottiene è uguale allo stimatore di regressione generalizzata (GREG) (si veda

Cassel *et al.*, 1979; Särndal, 1980; Isaki & Fuller, 1982; Wright, 1983; Bethlehem & Keller, 1987; Särndal *et al.*, 1989; Fuller, 2002):

$$\hat{t}_{Y_{GREG}} = \hat{t}_{Y_{\pi}} + (t_x - \hat{t}_{x_{\pi}})' \hat{\beta},$$

dove $\hat{t}_{Y_{\pi}} = \sum_{k \in S} y_k d_k$ è la stima di t_Y ottenuta con lo stimatore di Horwitz-Thompson e $\hat{t}_{x_{\pi}} = \sum_{k \in S} x_k d_k$ è il vettore dei totali delle variabili ausiliare calcolate con lo stimatore di Horwitz-Thompson. Mentre

$$\hat{\beta} = \left(\sum_{k \in S} \frac{q_k \mathbf{x}_k \mathbf{x}_k'}{\pi_k} \right)^{-1} \left(\sum_{k \in S} \frac{q_k \mathbf{x}_k y_k'}{\pi_k} \right)$$

è il vettore dei coefficienti di regressione stimati, con q_k sistema di ponderazione tale che $q_k > 0$.

Anche usando altre metriche rispetto alla distanza del chi quadro (si veda Deville & Särndal, 1992 pp. 378-379) lo stimatore calibrato è asintoticamente equivalente allo stimatore GREG. Pertanto, anche la varianza asintotica dello stimatore CAL può essere derivata facendo riferimento allo stimatore GREG. Quindi:

$$\hat{V}(\hat{t}_{Y_{GREG}}) = \sum_{k \in S} \sum_{l \neq k} \frac{\Delta_{kl}}{\pi_{kl}} (w_k \hat{e}_k) (w_l \hat{e}_l) = \hat{V}(\hat{t}_{Y_{CAL}}) \quad (1)$$

con $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$. L'espressione (1) è funzione dei residui del modello di regressione della sovra-popolazione implicitamente ipotizzato dallo stimatore GREG e stimati sul campione $\hat{e}_k = y_k - \mathbf{x}_k' \hat{\beta}$ e $\hat{e}_l = y_l - \mathbf{x}_l' \hat{\beta}$ per $k, l \in S$ e con $k \neq l$.

3.2. L'errore campionario dello stimatore di calibrazione nel Wave Approach (WA) dell'indagine sulle Forze di Lavoro (FoL)

Tutti i principali istituti nazionali di statistica utilizzano lo stimatore di calibrazione con totali noti da fonti amministrative per produrre stime, così come definito originariamente da Deville e Särndal (1992). È sempre più frequente, però, il ricorso a totali noti stimati da altre indagini, ma anche a totali stimati da occasioni precedenti delle stesse indagini per migliorare le stime e renderle meno volatili, soprattutto nel caso in cui si utilizza uno disegno ruotato (si veda Guandalini 2014 per un quadro delle ragioni alla base dell'utilizzo di totali stimati nello stimatore di calibrazione). Il caso del WA nell'indagine FoL è, da un punto di vista metodologico, simile a quest'ultimo caso. Anche nel WA, infatti, si utilizzano totali ottenuti su una parte dello stesso campione per cui si intende ottenere delle stime.

Lo stimatore della varianza nella (1) sotto-stimerebbe l'errore campionario. Infatti, non tiene conto del fatto che i vincoli sono stimati e quindi affetti anche

loro da errore campionario (Berger *et al.*, 2009; Dever and Valliant, 2010). Per avere una valutazione complessiva dell'errore campionario, però, è necessario considerare anche la dipendenza tra il sotto-campione e il campione annuale.

Si consideri lo stimatore di regressione esteso Ballin *et al.* (2000), Rancourt (2001), Berger *et al.* (2009):

$$\hat{t}_{Y_E} = \hat{t}_{Y_\pi} + (\tilde{\mathbf{t}}_{AUX} - \hat{\mathbf{t}}_{AUX_\pi})' \hat{\boldsymbol{\beta}} \quad (2)$$

in cui $\tilde{\mathbf{t}}_{AUX} = (t_{X_1}, \dots, t_{X_P}, \hat{t}_{Z_1}, \dots, \hat{t}_{Z_M}) = (\mathbf{t}_X, \hat{\mathbf{t}}_Z)$, in cui i primi P elementi sono totali noti da fonte amministrativa (come ad esempio quelli utilizzati in S_{283}) e gli altri M elementi sono stime (come ad esempio i 62 vincoli aggiuntivi considerati nel WA dell'indagine FoL). Lo stimatore della varianza dello stimatore nella (2) (Berger *et al.*, 2009, Renssen and Nieuwenbroek, 1997) è dato da:

$$\hat{V}(\hat{t}_{Y_E}) = \hat{V}(\hat{t}_{Y_{GREG}}) + \hat{V}_1 + \hat{V}_2 - \hat{V}_3$$

con:

$$\begin{aligned} \hat{V}_1 &= \boldsymbol{\beta}' \hat{V}(\tilde{\mathbf{t}}_X) \boldsymbol{\beta} = (\beta_{P+1} \dots \beta_{P+M})' \hat{V}(\hat{\mathbf{t}}_Z) (\beta_{P+1} \dots \beta_{P+M}) \\ \hat{V}_2 &= 2\boldsymbol{\beta}' \hat{C}(\hat{t}_{Y_\pi}, \tilde{\mathbf{t}}_X) \boldsymbol{\beta} = (\beta_{P+1} \dots \beta_{P+M})' \hat{C}(\hat{t}_{Y_\pi}, \hat{\mathbf{t}}_Z) (\beta_{P+1} \dots \beta_{P+M}) \\ \hat{V}_3 &= 2\boldsymbol{\beta}' \hat{C}(\hat{\mathbf{t}}_{X_\pi}, \tilde{\mathbf{t}}_X) \boldsymbol{\beta} = (\beta_{P+1} \dots \beta_{P+M})' \hat{C}(\hat{\mathbf{t}}_{AUX_\pi}, \hat{\mathbf{t}}_Z) (\beta_{P+1} \dots \beta_{P+M}) \end{aligned} \quad (3)$$

in cui $\hat{\mathbf{t}}_{AUX_\pi}$ è il vettore dei totali delle $P + M$ variabili ausiliare calcolate con lo stimatore di Horwitz-Thompson e $\hat{V}(\tilde{\mathbf{t}}_X)$ è la matrice di varianze e covarianze campionarie dei totali noti, che si riduce a $\hat{V}(\hat{\mathbf{t}}_X)$, ovvero la matrice di varianze e covarianze campionarie dei soli totali stimati. Inoltre $C(\cdot, \cdot)$ è la matrice delle covarianze campionarie, in cui il generico elemento tiene conto della covarianza dovuta alla dipendenza tra il campione per cui si sta producendo la stima e il campione o i campioni su cui sono stimati i totali noti. Nel caso in cui, i campioni sono indipendenti \hat{V}_2 e \hat{V}_3 sono uguali a zero.

Nel caso del WA, però, come già detto, il sotto-campione del WA ed il campione annuale hanno delle famiglie in comune, quindi non sono indipendenti e \hat{V}_2 e \hat{V}_3 sono diversi da 0. In particolare, \hat{V}_2 tiene conto della correlazione tra lo stimatore della variabile di interesse e lo stimatore dei totali noti, ovviamente solo quelli stimati sul campione annuale. \hat{V}_3 , invece, tiene conto della correlazione tra gli stimatori di Horwitz-Thompson delle variabili ausiliarie sul sotto-campione del WA e gli stimatore dei totali stimati sul campione annuale.

Nell'espressione (3), \hat{V}_3 ha segno negativo e tanto più sarà alta la correlazione tra le variabili ausiliarie sul sotto-campione del WA e gli stimatori dei totali stimati sul campione annuale, tanto maggiore sarà il guadagno di efficienza che si otterrà. Considerando che il sotto-campione del WA rappresenta circa $\frac{1}{4}$ del campione annuale, è lecito attendersi che questa correlazione sia alta e quindi l'utilizzo di questi vincoli possa rendere più efficienti le stime.

4. Risultati della sperimentazione

Come già detto un primo banco di prova per l'applicazione del WA nell'indagine FoL è il modulo del programma "Garanzia Giovani". Questo modulo *ad hoc* ha l'obiettivo di valutare la conoscenza e l'efficacia del programma. Per questa sperimentazione tra le stime prodotte si è considerato il numero di individui che conoscono il programma, il numero di partecipanti al programma e il numero di occupati/formati dal programma.

Nella Tabella 1 sono riportate le stime ottenute con il sistema di calibrazione S_{283} e quello che considera anche i 62 vincoli aggiuntivi richiesti dal WA, quindi il sistema S_{345} . Inoltre, l'errore campionario è calcolato attraverso l'espressione (1) propria dello stimatore calibrato che ricorre a totali noti da fonte amministrativa e l'espressione (3) che tiene conto del fatto che alcuni vincoli sono stime, quindi affetti da errore campionario a loro volta, e della dipendenza tra il campione del WA ed il campione annuale da cui sono ottenuti questi totali.

Per brevità si riportano solo le stime a livello nazionale nonostante lo studio si sia concentrato anche su stime a livello sub-nazionale.

Tabella 1 – *Principali stime del modulo di "Garanzia Giovani" con relativo errore campionario, ottenuto con diversi sistemi di vincoli e con diverse formule della varianza campionaria, Indagine sulle Forze di Lavoro, anno 2015*

Sistema di vincoli	Stimatore della varianza	Variabili	Stima	CV%
S_{283}	$\hat{V}(\hat{t}_{GREG})$	Individui che conoscono il programma	1187,9	2,19
		Partecipanti al programma	379,3	3,48
		Occupati/Formati dal programma	79,5	7,95
Wave Approach S_{345}	$\hat{V}(\hat{t}_{GREG})$	Individui che conoscono il programma	1180,3	1,86
		Partecipanti al programma	372,3	2,94
		Occupati/Formati dal programma	779,4	6,57
S_{345}	$\hat{V}(\hat{t}_{VE})$	Individui che conoscono il programma	1180,3	2,06
		Partecipanti al programma	372,3	3,14
		Occupati/Formati dal programma	79,4	6,77

L'inserimento dei 62 vincoli per il WA comporta una leggera variazione nella quantità dei vari aggregati a livello nazionale. Guardando l'efficienza delle stime in termini di coefficiente di variazione percentuale (CV%), i vincoli richiesti dal WA comportano una miglioramento evidente. Inoltre, si ha la dimostrazione che se si utilizzasse l'espressione (1), ovvero $\hat{V}(\hat{t}_{GREG})$, si sotto-stimerebbe in maniera importante l'errore campionario di queste stime.

Tuttavia, utilizzando lo stimatore per il calcolo della varianza campionaria $\hat{V}(\hat{t}_E)$, ovvero quello opportuno in questo caso, emerge che il guadagno di efficienza avviene anche se i totali inseriti nella calibrazione del WA sono delle

stime. Questo perché i 62 vincoli aggiuntivi richiesti dal WA consentono di definire un modello di regressione migliore e la riduzione dell'errore campionario dovuta a questo compensa il fatto che questi siano stime e quindi affetti da errore campionario.

5. Conclusioni

Eurostat, nel programma di modernizzazione, al fine di ridurre il carico sui rispondenti, ha proposto un nuovo approccio basato sulla modularizzazione dell'indagine sulle Forze di Lavoro (FoL), questo approccio va sotto il nome di Wave Approach (WA). Attualmente non sono ancora prodotte stime dall'Indagine FoL utilizzando il WA. La prima applicazione ufficiale riguarderà il modulo *ad hoc* 2018 sulla "Riconciliazione tra lavoro e vita familiare".

L'Istat, nel frattempo, ha deciso di svolgere un test per valutare l'impatto che questo approccio può avere sulle stime in termini di precisione ed efficienza in modo soprattutto da predisporre una metodologia opportuna per la stima ed il realtivo calcolo dell'errore campionario.

Questo test ha riguardato le stime derivanti da una batteria di nove domande che sono presenti nel questionario FoL a partire dal primo trimestre 2015 relative al programma "Garanzia Giovani". In particolare sono state analizzate le tre variabili principali di questo modulo: numero di individui che conoscono il programma, il numero di partecipanti al programma e il numero di occupati/formati dal programma.

Le stime ottenute seguendo il WA sono molte vicine a quelle ottenute con la procedura standard.

L'errore campionario è stato calcolato ricorrendo ad uno stimatore che tiene conto della natura campionaria dei totali utilizzati e dalla relazione che lega il sotto-campione del WA e del campione annuale FoL dovuta allo schema di rotazione dell'indagine.

L'impatto più evidente sembra essere quello in termini di efficienza. Considerando i 62 vincoli richiesti dal WA si ottengono stime più efficienti nonostante la natura campionaria dei vincoli. Questo perché questi vincoli consentono di definire un modello di super-popolazione migliore e la riduzione dell'errore campionario dovuta a questo compensa il fatto che questi siano stime e quindi affetti da errore campionario.

I risultati dimostrano come con il WA, oltre ad ottenere l'obiettivo di coerenza delle stime, si riescano ad ottenere anche stime più efficienti. Ovviamente questo risultato non è sempre verificato per tutte le variabili, e soprattutto se si considerano diversi sotto-domini, poiché l'efficienza delle stime dipende contemporaneamente dalla bontà del modello di super-popolazione nel sotto-dominio e dall'errore campionario dei vincoli considerati.

Riferimenti bibliografici

- BALLIN M., FALORSI P.D., RUSSO A., 2000. Condizioni di Coerenza e Metodi di Stima per le Indagini Campionarie sulle Imprese. *Rivista di Statistica Ufficiale*, Vol. 2, pp. 31-52.
- BERGER Y.G., MUÑOZ J.F., RANCOURT E., 2009. Variance estimation of survey estimates calibrated on estimated control totals - An Application to the Extended Regression Estimator and the Regression Composite Estimator. *Computational Statistics & Data Analysis*, Vol. 53, pp. 2596-2604.
- BETHLEHEM J.G. & KELLER J.W., 1987. Linear weighting of sample survey data. *Journal of Official Statistics*, Vol. 3, No. 2, pp. 141-153.
- CASSEL C.M., SÄRNDAL C.-E., WRETMAN J.H., 1979. Prediction theory on finite population when model-based and design-based principle are combined. *Scandinavian Journal of Statistics*, Vol. 6, No. 3, pp. 97-106.
- CECCARELLI C., GIORGI G.M., GUANDALINI A., 2011. Varianza dello stimatore calibrato in presenza di informazioni ausiliarie campionarie. *Rivista Italiana di Economia Demografia e Statistica*, Vol. 65, No. 1, pp. 53-60.
- DEVER J.A. & VALLIANT R., 2010. A comparison of variance estimators for poststratification to estimated control totals. *Survey Methodology*, Vol. 36, No. 1, pp. 45-56.
- DEVILLE J.C. & SÄRNDAL C.-E., 1992. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, Vol. 87, pp. 376-382.
- DEVILLE J.C., SÄRNDAL C.-E., SAUTORY O., 1993. Generalized Raking Procedures in Survey Sampling. *Journal of the American Statistical Association*, Vol. 88, pp. 1013-1020.
- DEVILLE J.C., 1999. Simultaneous Calibration of Several Surveys. *Proceedings of Statistics Canada Symposium. Combining data from different sources*, pp. 207-212.
- FULLER W.A., 2002. Regression estimation for survey samples. *Survey Methodology*, Vol. 28, No. 1, 5-23.
- FULLER W.A. & RAO J.N.K., 2001. A Regression Composite Estimator with Application to the Labour Force Survey. *Survey Methodology*, Vol. 27, pp. 45-51.
- ISAKI C.T. & FULLER W.A., 1982. Survey design under the regression superpopulation model. *Journal of the American Statistical Association*, Vol. 77, No. 377, pp. 89-96.
- GUANDALINI A., 2014. Coerenza ed Ottimalità delle Stime Calibrate su Informazioni da Indagini Campionarie. Tesi di Dottorato, XXVI ciclo. Dipartimento di Scienze Statistiche, Università "La Sapienza" di Roma.

- ISTAT, 2006. La Rilevazione sulle Forze di Lavoro: Contenuti, Metodologie, Organizzazione. Collana Metodi e Norme, No. 32.
- RANCOURT E., 2001. La regression èntendue: un esemble de pratiques d'estimation qui poussent constamment la thèorie. Reccuil Enquetes. In J.J. Dreesbeke and I. Lebart (eds.), *Modèles et Application*, pp. 334-343.
- RENSEN R.H. & NIEUWENBROEK N.J., 1997 Aligning estimates for common variables in two or more sample surveys. *Journal of the American Statistical Association*, Vol. 92, No. 437, pp. 368-374.
- SÄRNDAL C.-E., 1980. On π -inverse weighting versus best linear unbiased weighting in probability sampling. *Biometrika*, Vol. 67, No. 3, pp. 639-650.
- SÄRNDAL C.-E., 2007. The calibration approach in survey theory and practice. *Survey Methodology*, Vol. 33, No. 2, pp. 99-119.
- SÄRNDAL C.-E., SWENSSON B., WRETMAN J., 2003. *Model assisted survey sampling*. Springer Science & Business Media.
- WRIGHT R.L., 1983. Finite population sampling with multivariate auxiliary information. *Journal of the American Statistical Association*, Vol. 78, No. 384, pp. 879-884.

SUMMARY

The impact of wave approach consistency constraints on sampling errors evaluation

Eurostat, for reducing the respondent burden in the Labour Force Survey (LFS), proposed the use of Wave Approach (WA). Since 2018 the Italian National Institute of Statistics (Istat) will start to produce some estimates following the WA. In the meanwhile, Istat have started an experimental study for evaluating the impact on estimates and properly evaluates their sampling error. In this work the WA in the Italian LFS and a proper estimator for the variance are presented. Furthermore, some results on the WA estimates are discussed.

Miriam DE SANTIS, ISTAT, mdesantis@istat.it
Alessio GUANDALINI, ISTAT, alessio.guandalini@istat.it
Antonella IORIO, ISTAT, iorio@istat.it
Alessandra LUGLI, ISTAT, lugli@istat.it

THE VALIDATION PROCESS OF REGISTER BASED MEASURES OF HOURS PAID AND WORKED. HOW WELL THEY FIT THE INTERNATIONAL REQUIREMENTS COMPARED TO SURVEY BASED MEASURES?

Ciro Baldi, Marilena Angela Ciarallo, Sara Gigante, Silvia Pacini

1. The use of registers for statistical production: the issue of definitions

The production of official statistics, especially in some domains, is more and more based on administrative data. The use of this data seems to be an optimal solution to the challenge of producing more and better information reducing the direct and indirect costs for the society. However, the administrative data, and the statistical registers built upon them, have their limitations and are a potential source of errors. Among these, the content of the variable has been often indicated as critical in that it might not fulfil the definitions required by the statistical standards.

A common opinion is that while in a survey the definition of a variable may be controlled by the statistician through an accurate design of the questionnaire, in the administrative data this is not usually possible since they are collected for purposes different from the statistical one. It is thus a task for the official statistician to assess whether an administrative variable fulfill the statistical definition and then can be used as such in the figure compilation (direct use) or can be used only as auxiliary variable in a, say, regression based estimator (indirect use). In many cases, the study of the administrative definition is not sufficient to solve the doubts about the type of uses that can be done of a certain variable and it is necessary to compare the actual measurement that can be done from the administrative source with the one from a survey source to complete the judgement.

This paper reports the analysis, both theoretical and empirical, performed to evaluate the statistical uses of the variables related to the working time derived from the Register on employment, earnings and hours (hereafter RACLI). The structure of the paper is the following. In §2 the main concepts of hours worked and hours paid are briefly presented. In §3 and §4 it is described how these variables are measured respectively in the Italian business surveys and in the RACLI register. In §5 the results of micro data comparison of the survey vs register measurements are shown in order to assess the magnitude of the

discrepancies and draw some hypothesis on the their reason. A paragraph of provisional conclusions follows (§6).

2. Indicators on working hours: uses and importance

Indicators on working time are very important within social and economic statistics. The main document that sets the theoretical framework for the international standards is the resolution of 18th International Conference of Labour Statisticians (ILO, 2008) which declares seven concepts of working time associated with the productive activities of a person and performed in a job. Among these the two core concepts are the number of hours (actually) worked and hours paid.

The variable “Hours actually Worked” (HW), defined as “the time spent in a job for the performance of activities that contribute to the production of goods and/or services during a specified short or long reference period”, is probably the most important measure of labour input. Its importance is either per se to contribute to the analysis of the business cycle or as denominator in measure of labour productivity, labour cost, etc... The number of hours worked are requested in a number of European statistical regulations.

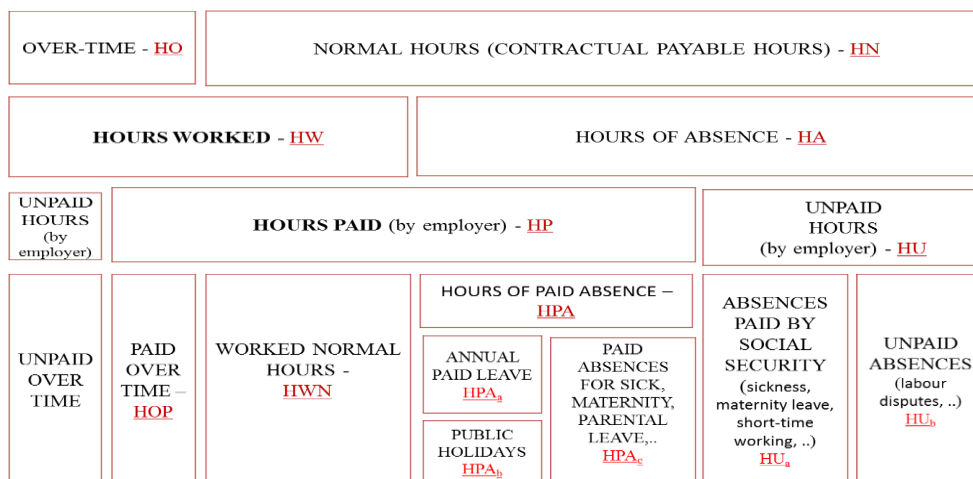
The variable “Hours Paid” (HP) is defined as “the time for which persons have received payment from their employer (at normal or premium rates, in cash or in kind) during a specified short or long reference period, regardless of whether the hours were actually worked or not”. It is mainly used as denominator of earning or labour costs to evaluate the “price of labour”. For instance, the earning per hour paid is the basic indicator to gauge pay differentials between categories of workers, jobs or even economic units. Among the others, it is the basis of a politically sensible indicator such as the gender pay gap.

Figure 1 provides the general framework in which these two concepts can be carved. For what concerns this paper, which refers to the time regularly paid in regular employee jobs¹, the number of hours paid are composed by the hours worked including the overtime and hours paid for absences by the employer (holidays, leave paid – e.g for sickness or maternity etc..).

The definitions and structure of these two variables, sketched up above, have been adopted, virtually unchanged, by Eurostat regulations that constitute the international requirements for Italian statistical system (see for instance Labour Cost Regulation - Commission Regulation n. 1737/2005, and the Structure of Earnings Survey Regulation - Commission Regulation 1738/2005).

¹ That is it excludes irregular jobs. Moreover, wages (and related hours) paid irregularly to regular jobs as well as unpaid hours are also excluded.

Figure 1 – The hours paid and worked: a theoretical framework.



3. Measuring the working time in business surveys: bottom-up and top down approaches

Business surveys generally can use two main approaches to collect and measure working time concepts in order to respond to the described international requirements : top-down and bottom-up.

In a top-down approach, starting from a more comprehensive variable, usually the number of normal hours, the respondent is requested to provide a measure for its components. Eventually the target variable is retrieved by subtracting from the top concept the components not belonging to the target variable and adding those relevant. An example of a top-down approach is presented by the Implementing Regulation of the Labour Cost Survey (Commission Regulation n. 1737/2005) whose basic data is at level of enterprise or local units.

Using the theoretical framework of figure 1, in synthesis, the regulation suggests that the paid time can be calculated starting from the number of normal contractual hours (HN), adding the number of paid overtime hours (HOP) and subtracting the hours unpaid by the employer (HU) for sickness, maternity leave, short-time working (HU_a), and other days not actually worked (HU_b e.g. for labour disputes)

$$HP = (HN + HOP) - (HU_a + HU_b) \tag{1}$$

Consequently the hours worked (HW) may be calculated starting from the HP (1) and subtracting the hours of absences paid by the employer but not worked for holidays/vacation (HPA_a), for public holidays (HPA_b) and for sickness leave, maternity leave etc (HPA_c)

$$HW = HP - (HPA_a + HPA_b + HPA_c) \quad (2)$$

The top-down approach can take advantages from information derived from administrative data to supplement the surveys. This method was used in the Structure of Earnings Survey (SES) 2014 edition (Baldi *et al.*, 2016), where for each sampled employee the employer, starting from the prefilled contractual hours derived from the RACLI register (HN), was asked to break down this top-level variable into its sub-components (Table 1). This approach helps to reduce the burden on respondents by maintaining at the same time a strong link with the register.

Table 1 – Top-down (on the left) and bottom-up (on the right) approaches.

<i>TOP-DOWN</i>		<i>BOTTOM-UP</i>	
Normal hours in the reference month (HN)	-	Worked normal hours (HWN)	+
<i>Hours of absences (HA)</i>	=	<i>Overtime (HOP)</i>	=
Total hours actually worked (HW)	+	Hours actually worked (HW)	+
<i>Hours of paid absences (HPA)</i>	=	<i>Hours of paid absences (HPA)</i>	=
Total hours actually paid (HP)		Total hours actually paid (HP)	

The SES 2014 is the only survey that has used this approach for the measurement of working time in Italy. Traditionally, all other business surveys instead have used a bottom-up approach, where the respondents are requested to provide the components that are later aggregated in the number of hours worked and hours paid. The components might be requested at different degree of details. The two business short-term surveys, the monthly survey on Large Enterprise (GI) and the quarterly survey on job Vacancies and Hours Worked on small and medium enterprises (VELA), require, separately for full-time and part-time employees, only three components: the worked normal hours (HWN), the paid over time hours (HOP) and the hours not worked but paid by the employer (HPA). The sum of the first two components produce the hours worked and adding the third component the hours paid are obtained (Table 1). Instead, the four yearly Labour Cost Survey (LCS) in recent years has requested a greater detail.

4. The estimation of hours paid and worked in the RACLI register

The method to derive the the working time variables adopted in the RACLI register at work relationship level, for the reference year uses a set of input information (Table 2). As shown in the table, the input information, although rich, are not homogeneous respect to time dimensions, since they have diverse reference periods and time metrics. With respect to the last, in fact, some variables refers to calendar days and others to standard (std.) days or weeks.

Table 2 – Main administrative information used for the RACLI Register hours estimate.

Administrative variables	Source*	Reference period	Time metrics
Contractual weekly hours	SSD	Week	Std.
N. of paid days	SSD	Month	Std.
Worked days	SSD	Day	Calendar
Percentage of part-time	SSD	Month	-
Lost wage because of events of absence not payed by employer	SSD	Month	-
N. of short-time-working hours directly payed by Social Security	SSD	Month	-

Note (*): SSD=Social Security Data (UniEmens and other social security declarations).

The method exploits this information jointly to produce a set of coherent variables at the level of microdata, for each work relationship in each month. From that level upward it is possible to obtain, through aggregations, annual estimates for different statistical units (working positions, individuals, employers).

Starting from the available administrative variables, some variables are derived (Table 3) to serve as building blocks in the calculation of final hours variables.

Table 3 – Main derived variables designed for the RACLI Register hours estimate.

Derived variables	Derivation method
Workable days in a week (5/6)	Estimated on the base of the daily worked days in the year.
Hourly theoretical wage	Theoretical wage / contractual working time.
Hours equivalent to the lost wage because of the absences	Lost wage because of events of absence not paid by the employer / hourly theoretical wage.
Wages not paid in each pay period	Actual wage - theoretical wage.
Days of paid absences	Result of a bounded model that harmonizes the declarations related to the paid (standard) and the worked (calendar) days.
Hours Unpaid for absences not paid by Social Security(HU _b)	Duration of the job - the number of the days (transformed in hours) recognized as dependent job.
Hours Unpaid for absences paid by Social Security(HU _a)	Hours equivalent to the lost wage because of the absences + short-time-working hours payed directly by Social Security

Considering the information available, the adopted approach is a top-down one starting from a top-level variable. The top-level variable is the number of normal hours in the reference month (HN). In particular, for each unit in each month the following steps are taken:

- The number of normal hours in a day is calculated dividing the contractual weekly hours by the number of workable days in a week (Table 3). Their summarization over the reference month (reduced in the case of activation and/or cessation of the work relationship during the month) produce the normal hours in the reference month (HN).
- The hours unpaid (HU) in the reference month obtained by summarizing the hours of unpaid absences paid by social security (HU_a) with the ones not paid by social security (HU_b) (Table 3).
- The hours paid overtime (HOP) obtained through a prediction method based on a regression estimator on survey data.
- The hours paid (HP) is then obtained as: $HP = (HN) - (HU) + (HOP)$
- Calculation of hours of paid absences (HPA) obtained transforming into hours the days of paid absences (table 3)
- The hours worked (HW) is then obtained as: $HW = HP - HPA$

5. The differences between register and survey based measurement

5.1. A general overview

In this paragraph we report the empirical differences between the register-based and survey-based variables on “hours paid” and “hours worked”. The three different surveys considered are the two four yearly, the LCS 2016 and the SES 2014, and the Quarterly Surveys on Hours Worked and Job Vacancies (Vela-GI).

Since we are interested in possible measurement errors with respect to the definitions illustrated in §3, we will focus the attention on the distribution of the differences between the survey and the register measures at microdata level. The statistical unit for comparison involving LCS and Vela-GI is the enterprise while for the SES is the employee within the enterprise. The reference periods are the years 2014, 2015, 2016 for Vela-GI, year 2016 for LCS and October 2014 for SES. The data used for the surveys are the raw microdata that is before any editing and imputation intervention for LCS and SES, while it is after E&I for Vela-GI. However, since most of the analysis are based on the medians and quartiles of the differences this choice is unlikely to affect the main results.

The metric we use throughout the comparisons is expressed by the following difference:

$$D = \frac{V_S - V_R}{(V_S + V_R)/2} \quad (3)$$

where V_S , V_R are respectively the variable (hours paid and hours worked, both net of overtime) measured in the Survey and in the Register.

The main advantage of this metric over to the more common percentage difference that at the denominator presents only one item (e.g. V_R), is that it is symmetric. This is in line with the spirit of the present work where there is no a priori benchmark value: in principle, we do not assume that the value of the survey or that of the register is the preferred measure. A second advantage of the chosen metric is that it is bounded between -200 and +200.

Before going in more depth for each comparison, Table 4 shows the main results presenting the median of the difference metric across the surveys and the years.

As it can be seen the difference is rather small with the highest value (in absolute terms) being -2.3%. Most comparisons show a negative sign, signaling that, at the median, the register value is larger than the survey value. A second characteristics is that the difference for hours worked is lower, in absolute terms, respect to the difference for hours paid.

Table 4 – Differences between median values of hours worked and paid between surveys and RACLI register. Percentage values.

Survey	Year	Hours Paid	Hours Worked
LCS	2016	-1.5	-1.1
SES	2014(October)	0.0	0.0
Vela-GI	2014	-1.9	-1.5
Vela-GI	2015	-1.3	-1.3
Vela-GI	2016	-2.3	-0.9

5.2. Structural Earning Survey and RACLI register

The survey whose values are the nearest to those of the register is the SES, which collects the hours and the salaries for every employee in the enterprise with reference to the month of October. Besides, we have to remember that the top-down approach followed in the SES2014 for the estimation of worked and paid hours is a result of the use of the RACLI register for pre-filling the top-level variables (number of workable hours in the month) in the questionnaire.

Table 5 – *Distribution of difference between SES and RACLI register on hours paid and worked by working time and pre-filled information. October 2014*

Working time	Paid hours				Worked hours			
	Mean	Q1	Median	Q3	Mean	Q1	Median	Q3
Total known pre-filled not modified by enterprises								
Full-time	0.0	0.0	0.0	0.0	-0.5	0.0	0.0	3.1
Part-time	-0.2	0.0	0.0	0.0	0.1	0.0	0.0	1.0
Total	-0.1	0.0	0.0	0.0	-0.4	0.0	0.0	3.0
Total known pre-filled modified by enterprises								
Full-time	0.0	-2.3	0.8	8.0	-4.0	-7.8	0.0	8.0
Part-time	-1.9	-2.0	0.0	4.8	-5.2	-10.0	0.0	3.1
Total	-0.6	-2.0	0.2	8.0	-4.4	-8.0	0.0	6.8

As table 5 shows, when the respondents did not modify the pre-filled data the differences in distribution are next to zero; when the respondents found inaccurate the total of workable hours given, the differences, although superior, are quite low. As in the following comparisons, the worked hours are more similar than the paid ones. Besides, there are not relevant differences between part-time and full-time workers.

5.3. The Vela-GI short-term surveys on hours and the RACLI register

Vela-GI are two different surveys: GI is monthly and covers all the enterprises in the private non-agricultural sectors with at least 500 employees; Vela collects information quarterly for a sample of enterprises with 10-499 employees in the same sectors covered by GI. As mentioned in §3, in both surveys the questionnaires, with a bottom-up approach, ask for the number of hours normal worked, overtime hours and hours paid but not worked. Since the comparisons performed are net of overtime hours the hours paid are obtained by summing up the first and the third component.

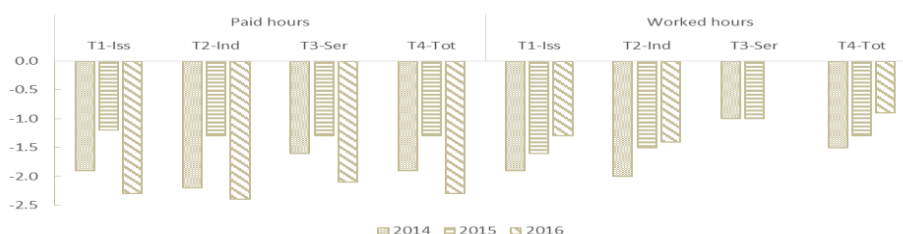
Table 6 – *Distribution of the difference between Vela-GI surveys and RACLI register on paid and worked hours in the main economic activity section. Year 2016*

NACE	Hours Paid				Hours Worked			
	Mean	Q1	Median	Q3	Mean	Q1	Median	Q3
T1-Iss	-3.5	-6.8	-2.3	1.4	-0.4	-5.1	-1.3	2.5
T2-Ind	-3.8	-7.2	-2.4	1.4	-0.7	-5.4	-1.4	2.4
T3-Ser	-4.2	-8.7	-2.1	3.4	2.9	-5.8	0.0	8.4
T4-Tot	-3.9	-7.8	-2.3	2.1	1.0	-5.5	-0.9	4.6

In 2016, the median difference of hours paid over the economic activity sections (Table 6) is negative in all sectors and totally being -2.3%. The normal hours worked, which constitute the large part of hours paid, unsurprisingly follow the same patterns, but show much smaller differences with the median for all the sectors being -0.9%.

Figure 2 compare the median over three years. A glance at the figures highlights that the median of the differences for hours paid is fairly stable over the years. For hours worked, in this short-time span, it seems to emerge an upward trend toward zero difference.

Figure 2 – Differences between Vela-GI and RACLI medians level of paid and worked hours per capita for economic sectors. Percentage values. Year 2014, 2015 and 2016.



5.4. The Labour Cost Survey 2016 and the RACLI Register

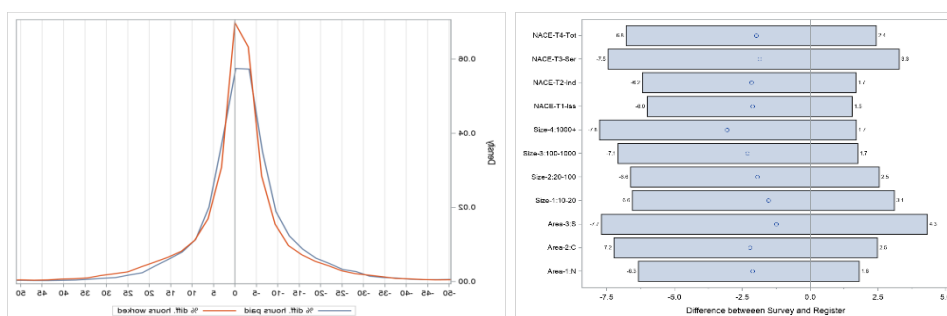
The LCS is a 4-yearly business survey on the structure of the labour cost covering a sample of economic units in NACE rev.2 sections from B to S with at least 10 employees. As mentioned in §-3, LCS uses a bottom-up approach to obtain the hours paid and worked.

Figure 3 (left side) reports the density curves of the difference of the hours paid (blue line) and hours worked (red line) per capita of full-time workers between the survey and the register. The two curves have a similar shape obviously, with a negative median, but with two notable differences. The one for hours worked has a median closer to zero and is more concentrated.

Looking more in depth (Figure 3 right side), the systematic deviation in the differences is found across several dimensions but the size is notably different. For instance, the distribution of the differences for the enterprises located in South Italy, although being wider, has a median which is only -1.2% compared to one of the North and the Center that are closer to -2.5%. Also significant is that the differences grows with the size of the firms.

Moreover the median differences of two variables have about the same magnitude of the comparisons with Vela-GI: for the hours paid it hovers around -2% and for hours worked -1.3%.

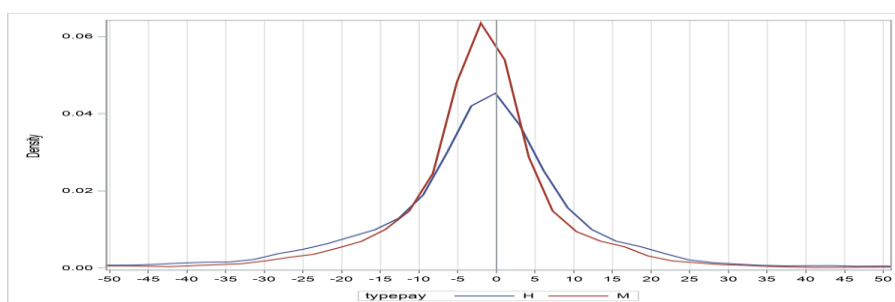
Figure 3 - Density curves (left) and median and quartile across size and geographical areas (right) of the difference of hours between LCS and RACLI. Year 2016.



Note. The graph is limited between -50 and +50.

A possible dimension that can provide some insights on the nature of these differences is the type of payslip. In Italy it can be monthly or hourly. In the first case, which is the most used, the employees are paid on a monthly basis that is the same amount (of regular payment) every month, independently of the number of calendar days in the month. In the second case (used especially for blue collar workers in some sectors such as Construction) the employees are paid on an hourly basis².

Figure 4 - Density curves of the difference of hours between LCS survey and RACLI register according to the type of payslip. Year 2016.



² This implies that, being equal the normal daily hours, the regular payment for short months (e.g. February) will be lower than the one for a longer month (one with 31 days).³

6. Conclusions

The differences on hours paid and worked between business surveys and RACLI register data are not very relevant but systematic, at least when the analysis unit is the enterprise and the reference period is the year. In these cases, in fact, at the median, the register variable is usually higher than the survey one. The reason why this happens is not clear. On one hand in the register there is a possible over-estimation of hours not worked and not paid in days worked. On the other hand, the survey may have systematic measurement errors. For example, the difference in the payslip can affect the survey, according to our hypothesis, because the accountancy of attendance at work is different. The enterprises with monthly payslip have no incentive to monitor the public holidays that occur during the working week, as the worker will be paid the same at the end of the month in presence or not of public holidays in workable days. Instead, the enterprises which pay on an hourly basis have to count these days which have to be paid. So the survey measurement for the first type of enterprises might be easily affected by underreporting.

Another possible source of error that is likely to affect the survey more than the register is the errors derived from aggregation over time and employees. In fact, the comparison of the register with the SES that is for a single month and at employee level show much more coherence between the two measurements.

Overall, considering how difficult is to measure this variable, the register data does not seem to have serious disadvantages compared to survey to satisfy the international regulations, at least for the normal hours (paid and worked). In fact the register measurement still needs an integration with information derived from a survey needed to estimate the overtime.

References

- COMMISSION REGULATION n. 1737/2005, regards the definition and transmission of information on labour costs, implementing the Council Regulation (EC) N. 530/1999 of 9 March 1999 concerning structural statistics on earnings and on labour costs in Official Journal of the European Union L 279/11 of 22.10.2005
- ILO. 2008. 18th International Conference of Labour Statisticians, Geneva, Switzerland, 24 November - 5 December 2008.
- BALDI C., CASCIANO C., CIARALLO M. A., CONGIA M. C., DE SANTIS S., PACINI. S. 2016. Designing the integration of register and survey data in earning statistics. <http://www.ine.es/q2016/docs/q2016Final00268.pdf>

ZHANG L.C. (2012) Topics of statistical theory for register based statistics and data integration *Statistica Neerlandica* (2012) Vol. 66, nr. 1, pp. 41-63

SUMMARY

The validation process of register based measures of hours paid and worked. How well they fit the international requirements compared to survey based measures?

The variables on hours worked and paid are essential measures on a wide set of statistics ranging from labour to business statistics and national accounts. Traditionally this kind of variables are collected through surveys. In Italy, the Employment, earnings and hours (RACLI) register calculates measures of these target variables. As a first step the main use of the register-based variable on hours paid has been as auxiliary information in the estimation process of the number of hours worked in the context of Structural Business Statistics. More recently, consequently to an improvement in the administrative information, the number of hours paid have been used directly in the Earning statistics and a first register based estimate of the number of hours worked produced. This paper describes the estimation method of the two variables, the validation analysis performed at micro level by comparing register, and survey based measures. The surveys used for the comparison are either at enterprise level, such as the short-term surveys on labour input and the four yearly Labour Cost Survey, and at the employee level (the Structure of Earning Survey). The study of, both theoretical and empirical, discrepancies between register and survey based measures of hours paid and worked provides a quite accurate indication of the capacity of the first measures to fit the international standards and at the same time provides insights on the peculiarities and, sometime, fallacies of survey based measurements.

Ciro BALDI, Istat, baldi@istat.it

Marilena Angela CIARALLO, Istat ciarallo@istat.it

Sara GIGANTE, Istat, gigante@istat.it

Silvia PACINI, Istat, pacini@istat.it

INFERENCE FOR INEQUALITY MEASURES: A REVIEW

Margherita Gerolimetto, Stefano Magrini

1. Introduction

After the seminal paper by Schluter and Trede (2002), who introduced concepts and tools for extreme value theory to the econometrics, part of the literature on inequality measures has concentrated on the consequences on inference due to heavy upper tail of the income distribution. More in details, the focus is on inferential problems created by population income distributions whose right tail decays slowly like a power function, *i.e.*

$$(1) \quad P(Y > y) \sim \beta y^{-\alpha} \quad \text{as} \quad y \rightarrow \infty$$

α is the stability index and provides information on moments finiteness, if $\alpha < 1$, the mean is infinite mean and also the variance, if $\alpha < 2$, the variance is infinite.

Examples of this type of distributions, often called “heavy tailed”, are the Singh-Maddala, Pareto, Generalized Beta distributions. For these heavy tailed distributions, standard methods of inference, both asymptotic and bootstrap, are unreliable. The underlying intuition is that as tail heaviness increases, the population moments increase and eventually cease to exist, whilst the (finite) sample moments tend to underestimate them.

In the attempt to overcome severe inference problems, a number of contributors have presented in the literature. Broadly speaking, works on this topic can be divided into two groups. The first group includes methods focusing on point measures of inequality and aiming at handling the limits of conventional inference (in particular, bootstrap inference) in the presence of heavy tails. The second group includes methods based on partially ordering income distributions using stochastic dominance and related criteria, one such is Donald *et al.* (2012).

In the present review, we will concentrate in particular on the first group of methods about which a heat debated has developed in the last 10 years. On the one hand, there are approaches oriented to improve the finite sample performance in

case of very heavy tails, based on bootstrap and permutation (Davidson and Flachaire, 2007; Dufour *et al.* 2017). On the other hand, there are approaches based on deriving asymptotic expansions for the distributions, to transform the statistics opportunely (Schluter, 2012; Schluter and van Garderen, 2009).

This paper is structured as follow. In the second section the contribute by Davidson and Flachaire (2007) is discussed. The third section is devoted to Schluter (2012). The fourth section presents the very recent paper by Dufour *et al.* (2017). The fifth section concludes.

2. Davidson and Flachaire (2007)

Davidson and Flachaire (2007), DF07 hereafter, begin by studying the finite-sample performance of asymptotic and bootstrap inference for inequality measures, in particular they concentrate on the Generalized Entropy class. Their preliminary simulations show that, in spite of large samples of usually *iid* observations, neither asymptotic nor standard bootstrap inference perform well, in particular severe overrejection of the null hypothesis is documented. As anticipated in the first section, the reason of this is that asymptotic and bootstrap inference is very sensitive to the exact nature of the upper tail of the income distribution, especially in cases of infinite variance.

Focussing on the Theil index¹, DF07 propose two alternative bootstrap to deal with these cases of infinite variance: i) a revised version of the m out of n bootstrap (moon bootstrap) and ii) a semiparametric bootstrap.

The m out of n bootstrap (Politis and Romano, 1994; Bickel *et al.* 1997) is based on bootstrap samples of dimension $m < n$, where n is the original sample size, and it is known in literature as a good option when the standard bootstrap fail or it is difficult to prove its consistency.

The performance of the moon bootstrap is shown via simulations, for the Theil index, $n=50$, $m=2, \dots, 50$, from a Singh-Maddala distribution which can successfully mimic observed income distributions in various countries (Brachman *et al.*, 1996) and whose cumulative distribution function is

$$F(y) = 1 - \frac{1}{(1+ay^b)^c} \quad (2)$$

¹ For a random variable y with cumulative distribution function F , the Theil index can be written as $T(F) = \left(\frac{v_F}{\mu_F}\right) - \log(\mu_F)$, where $\mu_F = E_F(y)$ and $v_F = E_F(y \log y)$.

where $a=100$, $b=2.8$, $c=1.7$. The tail parameter $\alpha=bc=4.76$ is a choice that closely mimics the net income distribution of German households. DF07 results show that for $m=22$ the percentage of rejection is very close to the nominal level. However, the moon bootstrap performance is very sensitive to the choice of m . In particular, for $m=n$ the moon bootstrap coincides with the standard one and approaches serious overrejection; by reducing m the number of rejections decreases, but when m is too small, the moon bootstrap approaches underrejection. Given these results, DF07 introduce a different version of the moon bootstrap. To do this, they firstly quantify the bias that leads to overrejection, then they approximate it via the moon bootstrap (with some choice of m), finally, they obtain an adjusted p-value (denoted revised moon bootstrap p-value).

As an alternative to the revised moon bootstrap, DF07 propose a semiparametric bootstrap that combines a parametric bootstrap for the right tail with a standard bootstrap for the main body of the distribution. In a first step, it is estimated the index of stability of the right tail of the distribution (resorting to the k greatest order statistics of a sample of size n , for some integer $k \leq n$)

$$\hat{\alpha} = H_{k,n}^{-1} \text{ and } H_{k,n} = k^{-1} \sum_{i=0}^{k-1} \log y_{n-1} - \log y_{n-k+1} \quad (3)$$

where y_j is the j -th order statistic of the sample, the choice of k is a matter of trade-off between bias and variance. Common choice, usually based on graphical methods, and adopted here, is the square root of the sample size (Coles, 2001).

In a second step, the bootstrap samples are drawn from a distribution defined as a function of a probability p_{tail} that constitutes the tail of the distribution. Each observation of the bootstrap sample is, with probability p_{tail} , a drawing from the distribution with cumulative density function

$$F(y) = 1 - (y/y_0)^{\hat{\alpha}}$$

where $y > y_0$ and y_0 is the order statistic of rank $\bar{n} = n(1 - p_{tail})$ of the sample and, with probability $1 - p_{tail}$, a drawing from the empirical distribution of the sample of the smallest $n(1 - p_{tail})$.

Table 1 – ERP ($\alpha=4.76$, $h=0.4$, m is the closest integer to \sqrt{n})

n	50	100	500	1000	2000	3000
Asymptotic	0.140	0.115	0.071	0.062	0.052	0.051
Standard bootstrap	0.058	0.049	0.032	0.022	0.021	0.020
Moon bootstrap (revised)	-0.043	-0.0018	0.021	0.020	0.020	0.020
Semiparametric bootstrap	0.026	0.022	0.005	0.001	-0.002	0.003

Simulations results by DF07 from Singh-Maddala distribution are summed up in tables 1 and 2, where $p_{tail} = hk/n$, $h=0.3, 0.4, 0.6, 0.8$. The performance is in terms of Error in Rejection Probability, ERP, (left hand tail) for the Theil index, hence the closer the figures are to zero, the better. The revised moon bootstrap yields a slight improvement over the standard bootstrap. For the the semiparametric bootstrap, instead, the performance improves dramatically over the standard bootstrap, with insignificant ERPs for sample sizes greater than around 1000 (table 1). Still ERP are unacceptably high in case of very heavy tails (table 2).

Table 2 – ERP (in case of heavier tail, $h=0.4$, $n=100$, m is the closest integer to \sqrt{n})

	Asy	Stand boot	Moon boot (rev)	Semiparam boot
$\alpha=2.1$	0.41	0.24	0.15	0.16
$\alpha=1.9$	0.48	0.28	0.20	0.18

All in all, the simulation results by DF07 show that by adopting their bootstrap proposals the inferential problem is mitigated, especially with the semiparametric bootstrap, but still performance deteriorates as the tail of the income distribution becomes heavier.

3. Schluter (2012)

Moving from the results by DF07 and by Schluter and van Gardener (2009), Schluter (2012), hereafter SCH12, proposes a normalizing transformation of inequality measures, in particular the Generalized Entropy class. The work is based on Edgeworth expansions to adjust asymptotic Gaussian approximations in order to deal with the inference problem due to the heavy upper tail of the income distribution discussed above.

SCH12 begins by observing a systematic relationship between \hat{I} and $\widehat{var}(\hat{I})$ as potentially responsible of the severity of the inference problem. In particular, by plotting pairs of \hat{I} and $\widehat{var}(\hat{I})$ and observing the corresponding coverage error, the author recognize that the wrong confidence limits are associated to particularly low realizations of both \hat{I} and $\widehat{var}(\hat{I})$. Exploiting this relationship suggests the application of a variance stabilizing transform:

$$H(I) = \int_0^I \frac{du}{[\sigma(u)^2]^{1/2}} \quad (4)$$

where $\sigma^2(I)$ denotes the variance as a function of I . In conjunction with a bootstrap, the transform in (4) reduces the inference problem significantly. Finally, SCH12 develops asymptotic expansions for studentized (based on estimated variance)

$$S_n = \sqrt{n} \left(\frac{I-I}{\hat{\sigma}} \right) \quad (5)$$

where $\hat{\sigma}$ is the asymptotic standard deviation, derived with the delta method, and standardized inequality measures (based on the theoretical variance)

$$S_n = \sqrt{n} \left(\frac{I-I}{\sigma} \right) \quad (6)$$

A finite sample experiment documents (whose detailed results are not reported here) the positive effects of the stabilizing variance transform for various levels of the heaviness of the tail for a Singh-Maddala distribution. Compared to the poor quality of the Gaussian approximation (discussed in the previous section), the performance of the studentized bootstrap, coupled with the stabilizing variance transform, improves. In spite of this improvement in the performance, however for $\alpha = 2$ there is still a substantial difference between the nominal and the actual coverage behaviour.

4. Dufour, Flachaire and Khalaf (2018)

Dufour, Flachaire and Khalaf (2018), hereafter DFK18, propose Monte Carlo permutation and bootstrap methods for the problem of testing the equality of inequality measures between two samples. Their results cover the Generalized Entropy class. In addition to the previously discussed problems of heavy upper tail, DFK18 emphasize that inequality measures inference can also be confounded because those indices, as functionals of the cumulative distribution function, can be equal even in case the underlying distributions differ.

Consider two *iid* samples $X = \{X_1, X_2, \dots, X_n\}$ and $Y = \{Y_1, Y_2, \dots, Y_m\}$, from cumulative distribution functions F_X and F_Y , and the hypothesis testing

$$H_0: \theta(F_X) = \theta(F_Y) \quad (7)$$

where $\theta(\cdot)$ is some functional on some subset F of distributions. Inequality indices are special cases of $\theta(\cdot)$. A natural statistic is:

$$T = \theta(\hat{F}_X) - \theta(\hat{F}_Y) \quad (8)$$

where \hat{F}_X and \hat{F}_Y are the empirical cumulative distributions and the studentized version is

$$S = \frac{\theta(\hat{F}_X) - \theta(\hat{F}_Y)}{\sqrt{\hat{\nu}(\theta(\hat{F}_X)) + \hat{\nu}(\theta(\hat{F}_Y))}} \quad (9)$$

and, for both T and S , DFK18 consider three p-values (in addition to the asymptotic p-value, based on the Gaussian limiting distribution): (i) MC permutation p-value (ii) bootstrap p-value (iii) bootstrap p-value that imposes the null hypothesis.

The permutation p-value (i) is obtained from the distribution derived by permuting in all possible ways the $N = n + m$ observations of the combined sample

$$Z = \{X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m\}$$

under the assumption of *iid* samples. The permutations, in total $(m + n)!$, are all equally probably which in turn determines the permutational distribution of T or S .

B permutations are drawn at random (Dwass, 1957) from the set of all permutations and along with the actual data this yields $B + 1$ random permutations of Z ., \hat{F}_{X^*} and \hat{F}_{Y^*} are the corresponding cumulative distribution function and the value of the test statistic is:

$$T_* = \theta(\hat{F}_X) - \theta(\hat{F}_Y) \quad (10)$$

The following is the permutation p-value function, where $j = 1, \dots, B$ refer to the series of permutation statistics and $\mathbf{1}(\cdot)$ is the indicator function:

$$p_* = 2 \min \left(\frac{\sum_{j=1}^B \mathbf{1}(T_{*j} \leq x) + 1}{B+1}; \frac{\sum_{j=1}^B \mathbf{1}(T_{*j} \geq x) + 1}{B+1} \right) \quad (11)$$

similar arguments hold for S_{*j} , the studentized version of T_{*j} .

The bootstrap p-value (ii) is obtained moving from the bootstrap samples, (X_b, Y_b) and \hat{F}_{X_b} and \hat{F}_{Y_b} , their corresponding empirical cumulative distribution functions; the bootstrap statistic is

$$S_b = \frac{(\theta(\hat{F}_{X_b}) - \theta(\hat{F}_{Y_b})) - (\theta(\hat{F}_X) - \theta(\hat{F}_Y))}{\sqrt{\hat{V}(\theta(\hat{F}_{X_b})) + \hat{V}(\theta(\hat{F}_{Y_b}))}} \quad (12)$$

hence for a two-tailed test, the bootstrap p-value, based on $j = 1, \dots, B$ bootstrap statistics:

$$p_b = 2 \min\left(\frac{1}{B} \sum_{j=1}^B \mathbf{1}(S_{bj} \leq S_0); \frac{1}{B} \sum_{j=1}^B \mathbf{1}(S_{bj} > S_0)\right) \quad (13)$$

It is interesting to observe that the permutation approach does not differ radically from the bootstrap approach. For example, a sample obtained by permuting elements of the combined sample Z is equivalent to resampling without replacement N observations from Z . Thus, resampling with replacement from Z represents an alternative bootstrap sample that respects the null hypothesis from which the bootstrap p-value under the null (iii) can be derived. This bootstrap sample is denoted by (X_0, Y_0) , \hat{F}_{X_0} and \hat{F}_{Y_0} are the corresponding empirical cumulative distribution functions; the bootstrap statistic is:

$$S_0 = \frac{(\theta(\hat{F}_{X_0}) - \theta(\hat{F}_{Y_0}))}{\sqrt{\hat{V}(\theta(\hat{F}_{X_0})) + \hat{V}(\theta(\hat{F}_{Y_0}))}} \quad (14)$$

and for a two-tailed test the bootstrap p-value under the null, based on $j = 1, \dots, B$ bootstrap statistics, is

$$p_0 = 2 \min\left(\frac{1}{B} \sum_{j=1}^B \mathbf{1}(S_0 \leq S_0); \frac{1}{B} \sum_{j=1}^B \mathbf{1}(S_0 > S_0)\right) \quad (15)$$

DFK18 provide asymptotic conditions for the validity² of the proposed methods (Romano, 1990; Chung and Romano, 2013) and they do it for two scenarios. In one case, they consider testing equality of the inequality measures when the population have the same distributions; this is equivalent to test $H_0: F_X = F_Y$ versus the alternative $\theta(F_X) \neq \theta(F_Y)$. Under this circumstance both level and size of the tests can be controlled, irrespective whether the distribution F_X or F_Y is continuous or discrete, without any restriction on the form of the functional θ . Moreover, permutation tests are exact both for the T and S statistic.

² Validity is intended in the sense that under the null hypothesis the rejection frequency tends to the nominal level as the sample size increases.

In the second case, when the populations do not have the same distributions things are more complicated since permutations test are no longer exact. However, such tests can be asymptotically valid if some restrictions are satisfied. In particular, for $\theta(\cdot)$ linear functional, if

$$V_{as}(\theta(\hat{F}_X)) = V_{as}(\theta(\hat{F}_Y))$$

implying that $n^{1/2}((\theta(\hat{F}_X) - \theta(F_X)))$ and $m^{1/2}((\theta(\hat{F}_Y) - \theta(F_Y)))$ have the same asymptotic variance or if $\frac{m}{m+n} \rightarrow \frac{1}{2}$ as $n \rightarrow \infty$ implying sample sizes asymptotically equal, asymptotic validity of the permutation test is guaranteed.

Note that the (arithmetic) mean is a linear functional, but the quantile is not. So comparing means from samples of similar size is then asymptotically valid when the sample sizes are similar even if the underlying distributions are not identical, while comparing quantiles with a permutation test is no longer valid, in general, if the underlying distribution are not identical.

The Generalized Entropy class considered by the authors is not a linear functional, unless the mean in the group is the same. However, given the scale invariance property of the Generalized Entropy class, it is possible to base a permutation test on the rescaled samples

$$\left\{ \frac{X_1}{\mu(F_X)}, \dots, \frac{X_n}{\mu(F_X)} \right\} \text{ and } \left\{ \frac{Y_1}{\mu(F_Y)}, \dots, \frac{Y_m}{\mu(F_Y)} \right\}$$

Comparing these indices from the rescaled samples makes no difference, while it validates (asymptotically) the use of permutation test. In practice, the following combined samples will be used where the sample means \bar{X} and \bar{Y} are adopted:

$$Z_S = \left\{ \frac{X_1}{\bar{X}}, \dots, \frac{X_n}{\bar{X}}, \frac{Y_1}{\bar{Y}}, \dots, \frac{Y_m}{\bar{Y}} \right\}$$

and the same holds for Gini index.

DFK18 carry out a simulations study focusing on extreme cases of heavy-tailed distributions in small samples to stress-test the methods employed in testing for Theil and Gini index. Competitors p-values are: asymptotic, standard bootstrap test S_b , permutation T rescaled (T_* based on Z_S), permutation S rescaled (S_* based on Z_S), permutation S standard (T_* based on Z), bootstrap S rescaled (bootstrap S° based on Z_S), bootstrap S standard (bootstrap S° based on Z).

Table 3 – *Theil index – Empirical size (same distribution, (r) stands for rescaled sample)*

α	asy	Boot S_b	Perm $T_*(r)$	Perm $S_*(r)$	Perm S_*	boot $S_\circ(r)$	boot S_\circ
2.9	0.083	0.082	0.130	0.065	0.051	0.041	0.038
3.2	0.075	0.081	0.115	0.061	0.051	0.042	0.038
3.5	0.071	0.082	0.092	0.060	0.050	0.043	0.041
4	0.068	0.080	0.075	0.055	0.050	0.043	0.041
5	0.064	0.079	0.063	0.051	0.050	0.045	0.041
6	0.061	0.078	0.060	0.050	0.050	0.048	0.047

The size part of the experiment is based on data generated from several Singh-Maddala distributions for which the Theil (and Gini) inequality index is the same and the tail index varies in [2.9,6.26], the value 2.9 corresponds to the most severe case of heavy tail. The sample size is small, $m = n = 50$. As a benchmark it is used the Singh-Maddala distribution with tail index equal to 4.76, the one used in DF07.

Results for the case of identical distributions are presented in tables 3 (Theil index) and 4 (Gini index).

Table 4 – Gini index – Empirical size (same distribution, (r) stands for rescaled sample)

α	asy	boot S_b	perm $T_*(r)$	perm $S_*(r)$	perm S_*	boot $S_\circ(r)$	boot S_\circ
2.6	0.125	0.072	0.095	0.065	0.052	0.055	0.047
2.9	0.117	0.071	0.087	0.064	0.051	0.054	0.048
3.2	0.106	0.071	0.074	0.060	0.050	0.0500	0.048
4	0.090	0.070	0.068	0.057	0.050	0.050	0.049
5	0.081	0.069	0.062	0.053	0.050	0.050	0.049
6	0.077	0.068	0.057	0.050	0.050	0.050	0.049

As we can see from tables 3 and 4, in case of identical distributions results show that asymptotic and standard bootstrap do not perform well for the heaviest tails cases, whilst permutation tests and bootstrap based on S statistic perform well. Test based on original samples provide exact inference, unlike test based on rescaled samples (this is because rescaling is done via the sample mean).

Table 5 – Theil index – Empirical size (different distribution, (r) stands for rescaled sample)

α	asy	boot S_b	perm $T_*(r)$	perm $S_*(r)$	perm S_*	boot $S_\circ(r)$	boot S_\circ
2.9	0.145	0.171	0.173	0.125	0.124	0.115	0.116
3.2	0.120	0.135	0.137	0.104	0.096	0.090	0.088
3.5	0.083	0.105	0.090	0.071	0.068	0.062	0.062
4	0.071	0.091	0.073	0.055	0.053	0.050	0.050
4.76	0.062	0.087	0.065	0.053	0.052	0.047	0.048
5.5	0.061	0.086	0.061	0.051	0.051	0.048	0.048
6	0.062	0.086	0.061	0.051	0.051	0.049	0.050

Table 6 – Gini index – Empirical size (different distribution, (r) stands for rescaled sample)

α	asy	boot S_b	perm T_* (r)	perm S_* (r)	perm S_*	boot S_* (r)	boot S_*
2.6	0.138	0.090	0.105	0.085	0.083	0.082	0.081
2.9	0.082	0.087	0.092	0.077	0.078	0.070	0.070
3.5	0.082	0.072	0.071	0.091	0.059	0.057	0.056
4	0.078	0.070	0.068	0.058	0.055	0.053	0.053
4.76	0.077	0.069	0.065	0.058	0.054	0.050	0.050
5	0.072	0.068	0.065	0.054	0.053	0.050	0.050

In case of different distributions (empirical size in tables 5 and 6), the distribution of X is fixed ($\alpha=4.76$) and the distribution of Y changes and its tail becomes heavier (the lower the value of α). The results show an overall performance that worsens the more F_Y is heavy tailed than F_X . Permutation tests and bootstrap based on S statistic perform similarly and better than other methods. DFK18 also consider the effects of the increase of the sample size (results not reported here). When $F_X = F_Y$, the rejection frequencies decrease slowly for asymptotic and standard bootstrap test. Instead, permutation and bootstrap under the null based on studentized statistic perform very well in all cases. When $F_X \neq F_Y$ the rejection frequencies decrease slowly for all methods, but permutation and bootstrap under the null based on studentized statistic outperform the other methods.

DFK18 also study the effects of the unequal sample sizes. In case $F_X = F_Y$, the more unequal are the sample sizes, the overrejections grow quickly for asymptotic and bootstrap test. Instead, permutation based on studentized statistic perform very well in all cases. In case $F_X \neq F_Y$ for all methods overrejections grow quickly as the sample sizes are more unequal, but permutation tests this happens more slowly.

As for the power side of the Monte Carlo experiment, DFK18 test the equality of an inequality measure between two samples, when the sample come from two distributions with different value of the inequality measure (F_X is fixed and F_Y varies). Power comparison of the considered permutation and bootstrap methods are valid since rejection probabilities under the null hypothesis $\theta(F_X) - \theta(F_Y) = 0$ are close to the nominal level. The permutation approach (rescaled and standard) is more powerful than the bootstrap under the null (rescaled and standard), the difference between the two approaches being resampling without replacement rather than with replacement. Studentized permutation test based on rescaled outperforms all other methods, especially when F_Y is heavier tailed than F_X .

All in all, results show that Monte Carlo methods outperform competitors both in terms of empirical size and power. Substantial reduction in size distortion is achieved more generally and studentized rescaled permutation tests outperforms the competing methods in terms of power.

5. Conclusions

In this work we present a review of some contributes of the econometric literature on comparing inequality measures. The main issue behind this bulk of recent literature is the heavy right tail of the income distribution, a condition under which standard methods of inference, both asymptotic and bootstrap are unreliable.

The papers we review are attempts to deal with this inference severe problem focusing on bootstrap and permutation, as well as on asymptotic expansion. Theoretical and simulations performance has been provided by all authors and, as far as we can tell by results, it is, in particular, the most recent proposal based on permutation methods by DFK18 the one that seems to better handle the issue.

References

- BICKEL P., GOTZE F., VAN ZWET W.R., 1997. Resampling fewer than n observations: gains, losses, and remedies for losses. *Statistica Sinica* Vol 7, pp. 1–32.
- BRACHMAN K., STICH A., TREDE M., 1996. Evaluating parametric income distribution models. *Allgemeines Statistisches Archiv*, Vol. 80, pp. 285–298.
- COLES, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.
- CHUNG E., ROMANO J. P., 2013. Exact and asymptotically robust permutation tests. *Annals of Statistics*, Vol. 41, pp. 484–507.
- DAVIDSON R., FLACHAIRE E., 2007. Asymptotic and bootstrap inference for inequality and poverty measures. *Journal of Econometrics*, Vol. 141, pp. 141–66.
- DONALD S.G., HSU Y.-C., BARRET G..F., 2012. Incorporating covariates in the measurement of welfare and inequality: methods and applications, *The Econometrics Journal*, Vol. 5, pp. C1-C30.
- DUFOUR J.-M., FLACHAIRE E., KHALAF L., 2018. Permutation tests for comparing inequality measures. *Journal of Business & Economic Statistics*, DOI: 10.1080/07350015.2017.1371027
- DWASS M., 1957. Modified randomization tests for nonparametric hypotheses. *Annals of Mathematical Statistics*, Vol. 28, pp. 181–187.

- POLITIS D.N., ROMANO J.P., 1994. The stationary bootstrap. *Journal of American Statistical Association*, Vol. 89, pp. 1303–1313.
- ROMANO J. P., 1990. On the behavior of randomized tests without a group invariance assumption. *Journal of the American Statistical Association*, Vol. 85, pp. 686–692.
- SCHLUTER C., VAN GARDEREN K. J., 2009. Edgeworth Expansions and Normalizing Transforms for Inequality Measures. *Journal of Econometrics*, Vol. 150, pp. 16–29.
- SCHLUTER C., TREDE M., 2002. Tails of Lorentz curves. *Journal of Econometrics*, Vol. 109, pp. 151–166.
- SCHLUTER C., 2012. On the problem of inference for inequality measures for heavy-tailed distributions. *The Econometrics Journal*, Vol. 15, pp. 125–153.

SUMMARY

Inference for inequality measures: a review

In this paper we present a review of the most recent contributes of the econometric literature on comparing inequality measures, focusing in particular on Theil index and Gini index.

We will start by discussing the main issue behind this bulk of literature, which is the heavy tail of the income distribution. Specifically, the severity of the inference problem responds to the exact nature of the right tail of the distribution. Attention in the literature has been given to determining the limits of conventional inference in the presence of heavy tails and, in particular, of bootstrap inference. Then we review a number of methods based on alternative parametric bootstrap and, more recently on permutations that heated in this debate in the last 10 years.

MARGHERITA GEROLIMETTO, Università Ca' Foscari Venezia, Dipartimento di Economia. E-mail: margherita.gerolimetto@unive.it

STEFANO MAGRINI, Università Ca' Foscari Venezia, Dipartimento di Economia. E-mail: stefano.magrini@unive.it

OPPORTUNITY AND DISCRIMINATION IN TERTIARY EDUCATION: A PROPOSAL OF AGGREGATION FOR SOME EUROPEAN COUNTRIES

Francesco M. Chelli, Mariateresa Ciommi, Francesca Mariani,
Maria Cristina Recchioni

1. Introduction

There is strict relationship between education achievement and socioeconomic outcomes. Usually, higher education is associated with improved health, higher incomes earned and greater possibilities of establishment in the labour market. On the other hand, education can also perpetuate since there exists a connection on the levels of educational attainment among generations. For instance, OECD (2017) data reveals that adults (30-59 years-old) from highly educated families more often reach a tertiary educational level than adults whose parents are not tertiary-educated (Figure 1).

In addition, according to Brown (2014), universities have an important role in ensuring real social mobility. Thus, government policies should be focused on encouraging and supporting universities to recruit more students from disadvantaged backgrounds with the aim of facilitating social mobility.

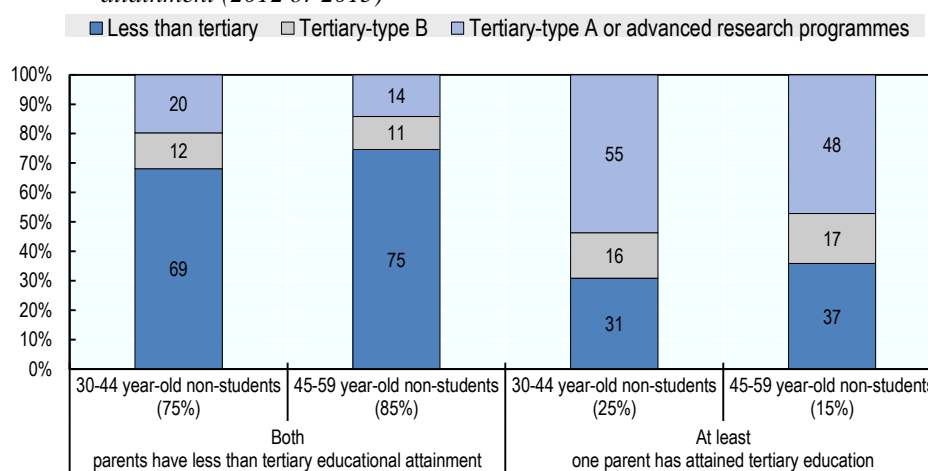
We are interested in analyzing (tertiary) education achievement and, in particular, if for a given individual (child), the family's educational level matters for obtaining the maximum educational title.

To achieve our aim, we analyze some OECD countries and we restrict to those countries belonging to EU15, namely Austria, Denmark, Finland, France, Germany, Greece, Italy, Netherlands, Spain and Sweden. As well it is known, among OECD countries, there is a great disparity in the education achievement, especially for what concern tertiary education. We use data from the OECD's publication "Education at a Glance 2017" (the reference year is 2015). We focus on individuals aged 30-44-year-old.

There are two distinct aspects of the educational achievement to be investigated: opportunity and discrimination. Opportunity refers to the probability for an adult to get a degree. Discrimination refers to the difference of opportunity among adults with different family status.

Thus, our main aim is twofold. Firstly, we are interested in analysing the ranking of the countries according to the two dimensions, opportunity and discrimination, separately; secondly, since the two aspects are the faces of the same coin, we investigate how to develop an index that takes into account the two dimensions, simultaneously.

Figure 1 – Educational attainment of 30-44 and 45-59 year-olds, by parents' educational attainment (2012 or 2015)



Notes: Percentage in parentheses represents the share of the population in each group. The values may not add up to 100% because of missing values in the source table. Data on educational attainment are based on ISCED-97

Source: OECD (2017), Tables A4.1 and A4.2. <http://dx.doi.org/10.1787/888933557147>

The rest of the paper is organised as follows: Section 2 briefly describes the data. In Section 3, after having discussed some mobility indices, we present the index of opportunity, the index of discrimination and a first proposal of aggregation of the two dimensions. Finally, Section 4 discusses further research and concludes.

2. Data

We use data (at country level) on education collected in 2015 (reference year) by the Survey of Adult Skills under the OECD Programme for the International Assessment of Adult Competencies (PIAAC) (OECD, 2017). The data regard 10 OECD countries.

We focus on individuals 30-44 years of age. Specifically, we consider four different percentages: 1) the percentage of adult that have attained tertiary

education given that both parents have less than tertiary education; 2) the percentage of adults that have attained tertiary education when at least one parent attained tertiary education; 3) the percentage of all adults whose parents (both) have less than tertiary education; 4) the percentage of all adults who have at least one parent that have attained tertiary education.

3. Opportunity and discrimination

There are several possible ways in which opportunity and discrimination can be measured. For us, opportunity refers to the probability, for an adult, in getting a degree. Discrimination refers to the difference of opportunity among adults with different family status.

Let us denote by E the double entry table summarizing all the information on education achievements of parents and children. The columns display the percentage of children that have achieved tertiary education (CL) or not (\overline{CL}). Similarly, rows indicate the percentage of parents (at least one) that have achieved tertiary education (PL) or not (\overline{PL}). Thus, the double entry E reads as:

	\overline{CL}	CL
\overline{PL}	p	$1 - p$
PL	q	$1 - q$

We associate with E the P matrix:

$$P = \begin{pmatrix} p & 1 - p \\ q & 1 - q \end{pmatrix}$$

Consequently, p is the probability for a child that has not achieved a tertiary education level to have parents that have not achieved the tertiary education level. Similarly, q is the probability for a child that has not achieved a tertiary education level to have at least one parent that has achieved a tertiary education level. Formally:

$$p = \Pr(\overline{CL} | \overline{PL}) \quad \text{and} \quad q = \Pr(\overline{CL} | PL)$$

and, without loss of generality, we assume $p > q$.¹

¹ This is not restrictive. In fact, the empirical analysis confirms that this holds for every OECD countries and in particular for the 10 selected countries.

Thus, P can be interpreted as a transition matrix. Bearing this in mind, the elements on the diagonal represent stayers, that is, individuals whose educational status does not change respect to their parents. By contrast, the off-diagonal elements are the movers. Of course, in the limit case in which everyone stays in the same class, the trace of matrix is 2.

Just to have an intuition, let P^{Italy} denote the transition matrix for Italy:

$$P^{Italy} = \begin{pmatrix} 0.86 & 0.14 \\ 0.32 & 0.68 \end{pmatrix}.$$

The first row indicates that, among all the adults from a family where both parents have not achieved tertiary education, 86% do not achieve tertiary education, while the remaining 14% achieve tertiary education. Similarly, the second row means that, among the adults from a family where at least one parent has achieved tertiary education, 32% do not achieve tertiary education, while the remaining 68% achieve tertiary education.

For each country, we also consider a parameter, k , representing the probability for a child to have at least one of the parents that have not a tertiary degree. That is,

$$k = \Pr(\overline{PL}).$$

Using this notation, the quantity $1 - k = 1 - \Pr(\overline{PL}) = \Pr(PL)$ is the probability for a child to belong to a family with at least one parents with the tertiary education level.

The comparison of countries or regions according to these matrices is not an easy task since it requires an ordering on these matrices. For this reason, scholars have introduced several indices, called mobility indices. Usually, the term mobility is used to denote changes in economic status of individuals due to an income variation. However, (most of) the measures introduced can be applied in different areas, for instance, to measure changing states due to education, occupation or marital status. Here, we refer to the former.

Mobility indexes are constructed such that zero mobility implies that the value of the index is 0, while perfect mobility implies that the value of the index is 1. Shorroks (1978) proposes two different indices, the so-called *Trace index* and the *Determinant measure*. The *Trace index* I is essentially based on the trace of P , that is the sum of the elements on the principal diagonal:

$$I' = \frac{n - Tr(P)}{n - 1},$$

where n denotes the number of different (economic or social) status. By definition, I' is a sort of measure of the concentration of the matrix around its diagonal.

The *Determinant measure* I'' , is defined as:

$$I'' = 1 - \frac{\det(P)}{n - 1}.$$

Similarly, Sommers and Conlisk (1979) propose to measure mobility through an index I''' that is a function of the eigenvalues of P :

$$I''' = 1 - |\lambda|,$$

where λ denotes the second largest eigenvalue.

Finally, the Bartholomew index (1973) I'''' weights transitions by the number of categories traversed. Thus, it is computed as a kind of average of the elements of the transition matrix (p_{ij}):

$$I'''' = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n p_{ij} |i - j|.$$

We observe that Shorrocks (1978), Sommers and Conlisk (1979) and Bartholomew (1982) defined mobility measures based on a quantile transition matrix. For a complete review of the literature about mobility indices see Dardanoni (1993) and Checchi and Dardanoni (2003).

In our work, we consider only two states, i.e. $n = 2$. It is easy to prove that, the above listed indices reduce to the same quantity², that is

$$I^1 = 1 - p + q.$$

3.1. Discrimination

We define the discrimination, Δ , associated with the matrix P as the difference in the probability of getting tertiary education given the family status, that is:

$$\Delta = \Pr(\text{CL}|\text{PL}) - \Pr(\text{CL}|\overline{\text{PL}}) = (1 - q) - (1 - p) = p - q.$$

² To be more precise, the indices reduce to the same quantity for $n = 2$, $I^1 = I' = I'' = I''' = I'''' = 2I''''$.

The discrimination index is the complement to 1 of the mobility index associated to the matrix P :

$$1 - \Delta = 1 - (p - q) = I^1.$$

Thus, the discrimination can be interpreted as the *opposite* of the mobility index.

A positive discrimination, $\Delta > 0$, implies that, for a given country, it is more likely to graduate if at least one of the parents is graduated. Furthermore, the greater the value, the lower the social mobility. When Δ approaches to 0, the degree of parents is irrelevant for the achievement of the title of the children. Since we assume $p > q$, we have that Δ is strictly positive.

The lowest discrimination, that in turn corresponds to the highest mobility, is achieved if the transition matrix has all rows identical, that is $p = q$. In this case, we have $\Delta = 0$ and the parent's education level has no effect on the child's education level (perfect mobility). In other words, we have equality of opportunity or null discrimination among the children coming from families with low or high educational level. For this reason, we refer to it as the "best case".

Table 1 – Ranking of some OECD countries according to discrimination index.

Ranking	Country	Value of the index
1	Finland	0.199
2	Austria	0.269
3	Sweden	0.274
4	Netherlands	0.295
5	Denmark	0.333
6	Germany	0.353
7	Spain	0.404
8	Greece	0.441
9	France	0.462
10	Italy	0.538

Our elaboration on OECD (2017) data.

The highest discrimination (or lowest mobility) is achieved if the transition matrix coincides with the identity matrix, that is $p = 1 - q = 1$. In this case we have no intergenerational transition between the educational levels and we refer to it as the "worst case"

Table 1 shows the country ranking according to the discrimination index.

3.2. Opportunity

We define the “opportunity loss index”, denoted by h , as the probability for a child of not attaining tertiary education independently from the education level attained by his parents. Formally:

$$h = kp + (1 - k)q$$

The h index ranges in the interval $[0,1]$. Specifically, $h = 0$ means that everyone has attained tertiary education, we refer to this as the “best case”. In contrast, the value $h = 1$ implies that nobody has attained tertiary education, that is the “worst case”.

Table 2 shows the country ranking according to the opportunity loss index.

Table 2 – Ranking of some OECD countries according to opportunity loss index.

Ranking	Country	Value of the index
1	Finland	0.482
2	Denmark	0.532
3	France	0.594
4	Netherlands	0.597
5	Sweden	0.607
6	Spain	0.627
7	Germany	0.629
8	Greece	0.704
9	Austria	0.790
10	Italy	0.834

Our elaboration on OECD (2017) data.

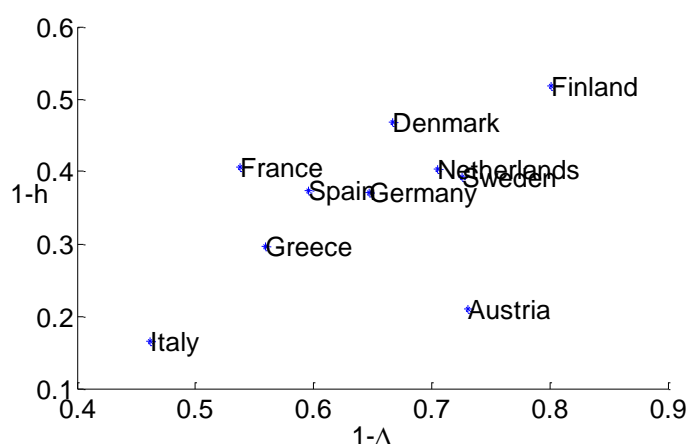
3.3. A proposal of aggregation

Comparing the ranking according to discrimination (Table 1) with loss of opportunity (Table 2), we find some conflicting results. Countries with higher performance in one dimension do not display the same performance according to the second dimension. For instance, Austria occupies the second position according to the discrimination, while is relegated to the penultimate position if we look at ranking generated computing the opportunity loss index.

Figure 2 displays the scatter plot of the opportunity loss index as function of the mobility index. Note that the scatter plot, though containing a global trend, is dominated by the scatter noise, that is, there is low positive correlation between the two. In fact, the two indices proposed do not display high level of correlation (we find a correlation coefficient equal to 0.5091).

To overcome this drawback, we introduce an index that takes into account both dimensions, $I(\Delta, h)$. We suppose that the two dimensions have the same weight and we adopt a non-compensative approach. This means that if a country achieves the minimum in one dimension, we do not desire that this component could be compensated by a high performance in the second dimension, as it happens for the arithmetic average. In addition, we desire that the index ranges in the interval $[0,1]$.

Figure 2 – Relationship between the two dimensions.



Source: Our elaboration on OECD (2017) data

Thus, we propose the following index:

$$I(\Delta, h) = \Delta h.$$

By definition, the smaller the value is the higher the performance of the country.

Table 3 reports the results according to $I(\Delta, h)$. Finland is the country with the best performance.

The ranking reveals that Scandinavian countries are the best performing countries, while Italy still occupies the last position.

Table 3 – Ranking of some OECD countries according to the new index.

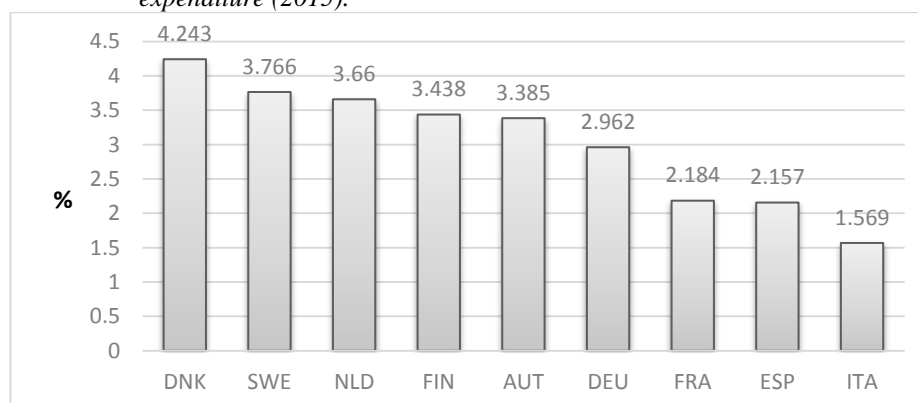
Ranking	Country	Value of the index
1	Finland	0.096
2	Sweden	0.166
3	Netherlands	0.176

4	Denmark	0.177
5	Austria	0.213
6	Germany	0.222
7	Spain	0.254
8	France	0.274
9	Greece	0.311
10	Italy	0.449

Our elaboration on OECD (2017) data.

Finally, we analyse the percentage of public spending on tertiary education respect to the totality of public expenditure. Figure 2 shows these percentages.³ According to OECD (2017), the “percentage of public spending on tertiary education” indicator reveals the priority given by governments to education in comparison with other public area of investment such as health care, social security or defence.

Figure 3 – *Public spending on tertiary education as percentage of the total government expenditure (2015).*



Source: Our elaboration on OECD (2017) data on public spending on education (indicator). doi: 10.1787/f99b45d0-en (Accessed on 13 July 2018).

It is interesting to note that Italy, which is ranked in the last position according to our index, is the country with the lowest level of public expenditure on tertiary education. Similarly, Denmark, Sweden, Netherland and Finland are ranked in the top of the ranking according to our index and also according to the expenditure. This suggests that there is a relationship between educational attainment and the amount of public expenditure that need to be more investigated.

³ The data refer to the last available year, that is 2014. For Greece, there is no data available.

4. Concluding Remarks

The paper deals with two aspects related to the possibility of reaching a given educational level conditioned on the educational level of the family: discrimination and opportunity. Ten European countries are ranked according to these indices, showing a different pattern. Consequently, a new index that takes into account the two aspects is proposed.

The index proposed here can be used by policy makers to monitor and evaluate the effectiveness of public policies. In fact, in contrast to the unidimensional mobility indices, that take into account only one dimension, the new index is able to capture simultaneously two different information: on one hand the difference in the probability of getting tertiary education given the family status (i.e., the discrimination index, Δ) and, on the other hand, the probability for a child of not attaining tertiary education independently from the education level attained by his/her parents (i.e., the opportunity loss index, h). Thus, $I(\Delta, h)$ is more informative than the unidimensional indices while preserving a simple functional form that requires an elementary computation. These are two key-properties necessary to define good composite indicators (OECD, 2008).

However, our index is defined by assuming that the two dimensions are equally weighted. Therefore, further researches will be devoted to the analysis of how combining these two indices in a different way. For instance, a different system of weights or a different functional form can be used.

References

- BARTHOLOMEW D. J. 1973. *Stochastic Models of Social Processes*, Second Edition London, Wiley.
- BROWN M. 2014. Higher Education as a tool of social mobility: Reforming the delivery of HE and measuring professional graduate output success. In: *London: Centre Forum*. Available at: <https://www.centreforum.org/assets/pubs/he-as-a-tool-of-social-mobility.pdf>
- CHECCHI D., DARDANONI V. 2003. Mobility comparisons: Does using different measures matter?, *Research in Economic and Inequality*, Vol.9, pp.113–145.
- DARDANONI V. 1993. On measuring social mobility, *Journal of Economic Theory*, Vol.61, pp. 372–394.
- OECD. (2008). Handbook on constructing composite indicators: Methodology and user guide. OECD Publishing, Paris. <https://www.oecd.org/sdd/42495745.pdf>
- OECD 2017. Education at a Glance 2017: OECD Indicators, OECD Publishing, Paris. <http://dx.doi.org/10.1787/eag-2017-en>

- SHORROCKS A. F. 1978. The Measurement of Mobility, *Econometrica*, Vol.46, No.5, pp. 1013-1024.
- SOMMERS P. M., CONLISK J. 1979. Eigenvalue immobility measures for Markov chains. *Journal of Mathematical Sociology*, Vol.6, No.2, pp. 253-276.

SUMMARY

Opportunity and Discrimination in Tertiary Education: a proposal of aggregation for some European Countries

Several studies underline the relationship between education achievement and socioeconomic outcomes. Higher education is associated with improved health, higher incomes earned and greater possibilities of establishment in the labour market. For these reasons, there is an increasing interest in analysing education achievement. Among OECD countries, there is a great disparity in the education achievement, in particular for what concern tertiary education.

Looking at individuals aged 30-44-year-olds, adults from highly educated families more often complete tertiary-education compared with adults whose parents have not a tertiary education. There are two distinct aspects to be investigated: opportunity and discrimination. Opportunity refers to the probability, for an adult, in getting degree. Discrimination refers to the difference of opportunity between adult with different family status.

We use data from the OECD's publication "Education at a Glance 2017" (the reference year is 2015) and, among the OECD countries, we focus on the EU15 countries.

The selected countries are compared according to an index of discrimination and an index of opportunity, respectively. In addition, we propose a new index that accounts simultaneously for the two dimensions.

Francesco M. CHELLI, Università Politecnica delle Marche, f.chelli@univpm.it
Mariateresa CIOMMI, Università Politecnica delle Marche, m.ciommi@univpm.it
Francesca MARIANI, Università Politecnica delle Marche, f.mariani@univpm.it
Maria Cristina RECCHIONI, Università Politecnica delle Marche, m.c.recchioni@univpm.it

IMPATTI DELL'AUTOMAZIONE SUL MERCATO DEL LAVORO. NUOVE STIME PER IL CASO ITALIANO

Massimiliano Bruno, Alessandro Polli

1. Introduzione

Obiettivo del presente contributo è stimare gli effetti della diffusione di processi produttivi innovativi (quali quelli connessi all'introduzione dell'intelligenza artificiale in una o più fasi della lavorazione e quelli associati all'automazione) sul mercato del lavoro, affinando ed estendendo la metodologia già presentata dagli Autori in un precedente contributo (Bruno e Polli, 2017).

Nel corrente dibattito scientifico, sebbene non si sia ancora giunti ad un accordo circa segno e portata dell'impatto provocato dalla diffusione dell'innovazione sul mercato del lavoro, numerosi autori, pur sottolineando la difficoltà di separare l'effetto specifico della tecnologia da quelli più generali legati alla globalizzazione (Rotman, 2015), tenderebbero a evidenziare una portata sostanzialmente negativa delle recenti evoluzioni.

In particolare, il Forrester Institute (2016) stima in circa il 6% la perdita netta di posti di lavoro negli Stati Uniti che si concretizzerebbe entro il 2021, mentre nel lungo periodo Bowles (2014) afferma che il 54% della forza lavoro europea rischia nei prossimi decenni di essere sostituita da processi automatizzati.

Di segno opposto le considerazioni di Graetz e Michaels (2015), che preconizzano il declino dei lavori a bassa specializzazione, sostituiti dall'automazione, ma aggiungono che il progresso tecnologico, incrementando la produttività, determinerà al contrario un incremento nella domanda di posti di lavoro a elevata specializzazione.

Dello stesso segno le valutazioni del Boston Consulting Group (2015) che, con riferimento al sistema produttivo tedesco, afferma che «i produttori sono in grado di aumentare la competitività e dunque di espandere la forza lavoro interna, sia perché la produzione capital-intensiva rende meno vantaggiosa la delocalizzazione del lavoro, sia per soddisfare la domanda che deriva da nuovi prodotti e servizi».

Centrale nella discussione appare la velocità di diffusione delle nuove tecnologie. Secondo un report recentemente diffuso dal Boston Consulting Group e dalla MIT Sloan Management Review (2017) – «Is Your Business Ready for Artificial Intelligence?» – a livello mondiale solo un'impresa su 5 ha implementato

l'IA in qualche prodotto o processo, mentre appena una su 20 ha introdotto in maniera massiccia strumenti di IA. La ricerca rivela inoltre l'esistenza di un ampio divario tra imprese leader – quelle che hanno già maturato una significativa esperienza nell'adozione di strumenti di IA – e le restanti imprese, che ancora sono incerte sulla strada da percorrere.

L'obiettivo dichiarato del report è smitizzare la narrativa sull'IA e fornire una base realistica di discussione. La ricerca, che ha coinvolto oltre tremila executive, manager ed esperti di 21 aree produttive provenienti da 112 paesi, ha evidenziato che sebbene più di tre quarti degli intervistati concordino nel ritenere che nei prossimi anni tale tecnologia sarà fondamentale per la crescita economica e la competitività, vi è una sostanziale differenza tra la consapevolezza che i manager hanno delle problematiche associate al reperimento di dati e all'addestramento di una IA e l'approccio operativo che effettivamente seguono.

Inoltre, per quanto molti degli intervistati ritengano che il divario tra imprese leader e «ritardatari» sia destinato a crescere nei prossimi 5 anni, quasi nessuna di queste ultime ha le competenze per creare un vantaggio competitivo e generare valore implementando soluzioni di IA nei rispettivi settori di competenza.

2. Intelligenza artificiale. Mito e realtà

Ogni seria riflessione sui rapporti tra automazione e mercato del lavoro non può prescindere da una considerazione realistica delle potenzialità che caratterizzano le attuali innovazioni, in particolare l'intelligenza artificiale¹.

Le più avanzate architetture multi-layer oggi in fase di sperimentazione si compongono di 16 milioni di neuroni, ben lontane quindi dagli 80 miliardi di neuroni di un cervello umano (Herculano-Houzel, 2009).

Uno studio pubblicato lo scorso anno da tre ricercatori della Chinese Academy of Sciences (Feng *et al.*, 2017) ha proposto una metodologia per la misurazione del quoziente intellettivo dei più avanzati sistemi di intelligenza artificiale. Lo studio dimostrerebbe come l'IA di Google (AlphaGo) sia leader con un QI di 47,29, il doppio di quello rilevato sul sistema Siri di Apple, ma che entrambi siano meno sviluppati rispetto al quoziente intellettivo medio di un bambino di sei anni (55,5) e ciò nonostante gli evidenti progressi compiuti nell'ultimo biennio in termini di comprensione del linguaggio naturale e di riconoscimento delle immagini.

¹ Da adesso in poi, intenderemo per intelligenza artificiale «una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana» (Marco Somalvico).

Ovviamente non è solo un problema di dimensioni, ma anche di efficienza computazionale: infatti le moderne reti neurali artificiali operano su supercomputer, mentre è ancora proibitivo utilizzare i normali desktop, per quanto potenti, considerato che sono necessari mediamente 12 giorni per «addestrare» una rete neurale composta da appena 100 mila neuroni.

È per tale motivo che iniziano ad affiorare forti perplessità circa le effettive possibilità di sviluppo in un immediato futuro dell'intelligenza artificiale. In particolare, alcuni scienziati iniziano a domandarsi se il deep learning sia realmente la soluzione, come nota Steve Lohr in un articolo pubblicato sul New York Times (2018).

Non a caso, il deep learning arranca quando si avventura in terreno aperto, e cioè negli ambiti più propri dell'intelligenza umana, come ricerca di significato, ragionamento e conoscenza basata sul senso comune. In altri termini, ciò che manca al deep learning è la capacità di generalizzazione, il che ne limita pesantemente le possibilità di introduzione nei processi produttivi.

«Gli algoritmi di deep learning», osserva Steve Lohr, «si addestrano su insiemi di dati correlati – come le immagini di volti umani – alimentati da masse crescenti di dati, che ne migliorano l'accuratezza». In questa semplice osservazione c'è la chiave che determina il successo di un'architettura di deep learning, che ottiene risultati se utilizzata per compiti specifici (come il riconoscimento di oggetti o le applicazioni di speech-to-text) e a condizione che siano disponibili enormi quantità di dati.

La precedente considerazione, detto *en passant*, spiegherebbe anche i motivi dello scarso interesse mostrato per l'intelligenza artificiale dalla piccola e media impresa, così come evidenziato da numerosi studi: difficoltà nell'individuazione di ambiti specifici di applicazione all'interno di un processo produttivo di solito relativamente poco complesso e difficoltà di raccolta della base dati necessaria all'addestramento di un algoritmo di intelligenza artificiale.

A ulteriore conferma che l'automazione è uno strumento sostanzialmente adottato solo nelle grandi imprese i risultati di una ricerca condotta dalla Lawless Research per conto della Service Now (2017), che evidenziano come il 50% delle aziende intervistate con più di cinquecento dipendenti abbia iniziato ad utilizzare l'automazione intelligente in uno o più processi di business, mentre il 75% afferma che la massa di dati raccolti è in crescita esponenziale grazie al ricorso a dispositivi mobili e Internet of Things.

Quindi, in sintesi, se da un lato risultano chiari gli attuali limiti di applicazione di processi di automazione intelligente, il mondo della ricerca mostra ancora notevoli difficoltà nel chiarire la tempistica con cui tali processi si mostreranno pienamente operativi e in quali anelli della *value chain* l'intelligenza artificiale

esplicherà i maggiori impatti, rendendo quindi arduo valutarne la portata in termini occupazionali.

3. Tecnologia e mercato del lavoro. Nuove valutazioni per il caso italiano

La metodologia proposta dagli Autori lo scorso anno (Bruno e Polli, 2017) intende fornire una risposta pragmatica all'obiezione avanzata da Rotman (2015) richiamata all'inizio della presente trattazione, che tende a evidenziare la complessità nella separazione degli effetti specifici derivanti dalla diffusione della tecnologia sul mercato del lavoro da altri fattori, quali i processi di delocalizzazione.

Nella metodologia sviluppata dai due Autori si capovolge la strategia di soluzione del problema: infatti, si procede preliminarmente all'individuazione di una misura di probabilità dell'evento «perdita del posto di lavoro a seguito della diffusione di nuove tecnologie nel processo produttivo», distinta per tipologia occupazionale, utilizzata per valutare gli impatti che si verificherebbero nei vari comparti produttivi se tali tecnologie fossero già disponibili e adottate oggi.

La predetta misura di probabilità è desunta da uno studio di Frey e Osborne (2013), in cui si stima la probabilità che un certo insieme di figure professionali (702 nello studio citato) siano «potenzialmente automatizzabili». Tali probabilità, applicate al mercato del lavoro statunitense, mostrerebbero che il 47% dei posti di lavoro risulterebbe potenzialmente vulnerabile alla tecnologia.

Gli Autori riconducono tale elenco alla classificazione CP2011 adottata dall'Istituto nazionale di statistica (Istat, 2013) per il mercato del lavoro italiano e che, sostanzialmente, recepisce le innovazioni introdotte dalla nomenclatura ISCO 2008 del Bureau International du Travail.

L'armonizzazione delle due classificazioni ha consentito di individuare 663 figure professionali, per le quali sono note, dallo studio di Frey e Osborne, le misure di probabilità descritte in precedenza, mentre sono stati desunti dal Registro statistico dell'occupazione delle imprese (ASIA Occupazione) i livelli occupazionali medi per codice ATECO e per macro-ripartizione geografica al 2016.

Preliminarmente, le 663 figure professionali sono state attribuite alle sezioni di classificazione ATECO. L'attribuzione si rende necessaria in quanto l'archivio ASIA Occupazione contiene informazioni sulla tipologia occupazionale (distinta in interna – dipendente e indipendente – ed esterna, quando il lavoratore ha un contratto di collaborazione o somministrazione) e sulla qualifica (dirigente, quadro, impiegato, operaio, ecc.), ma non sulla figura professionale.

Tale attribuzione è nella gran parte dei casi univoca, poiché la descrizione della figura professionale della nomenclatura ISCO è molto dettagliata anche con riferimento al settore di attività: valga per tutti il caso degli «Imprenditori e amministratori di grandi aziende nei servizi di alloggio e ristorazione», che ha univoca collocazione nelle sezioni G-I della classificazione ATECO. Solo in un numero estremamente limitato di casi l'attribuzione è stata svolta applicando un criterio di prevalenza.

Una volta attribuite le figure professionali alle sezioni di classificazione ATECO, si procede al calcolo della probabilità media di perdita del posto di lavoro a seguito della diffusione delle nuove tecnologie come media semplice delle misure di probabilità associate alle figure professionali attribuite a ciascuna sezione di classificazione. Tali probabilità medie sono riportate nella Tabella 1.

Tabella 1 – *Distribuzione delle figure professionali per sezioni di classificazione ATECO 2007, probabilità media di perdita del posto di lavoro a seguito della diffusione delle nuove tecnologie (\bar{p}_u), probabilità dell'evento complemento ($1 - \bar{p}_u$).*

ATECO 2007	Figure professionali censite	\bar{p}_u	$1 - \bar{p}_u$
b-e	179	0,664	0,336
f	36	0,709	0,291
g-i	133	0,665	0,335
j-s ⁽¹⁾	315	0,397	0,603
Totale	663	0,540	0,460

(1) *Escluse le sezioni di classificazione O (Amministrazione pubblica e difesa; assicurazione sociale obbligatoria) e S, divisione 94 (attività di organizzazioni associative)*

Da un sommario esame dei dati riportati nella precedente tabella, già nel lavoro pubblicato lo scorso anno si evidenziava la sostanziale concordanza tra i risultati dello studio e quelli ottenuti da Bowles (2014) con riferimento ai 28 paesi dell'UE. In sintesi, con riferimento alle 663 figure professionali individuate, si registrerebbe un rischio medio di perdita del posto di lavoro a seguito della diffusione delle nuove tecnologie pari al 54%.

Ovviamente, la reale perdita di posti di lavoro imputabile all'automazione dipenderà, a livello territoriale, dalla loro distribuzione tra settori produttivi e dalle figure professionali prevalenti in ciascun comparto di attività.

La nuova metodologia, ad ogni modo, tiene espressamente conto del dibattito sinteticamente riportato nei primi due paragrafi e applicherà tali probabilità solo alle imprese con un numero di addetti maggiore o uguale alle 50 unità, nell'ipotesi implicita, suffragata dagli studi citati, che al di sotto di quella classe dimensionale non vi sia alcun incentivo all'introduzione delle nuove tecnologie.

Nella Tabella 2 riportiamo l'occupazione registrata nell'anno 2016 nelle imprese appartenenti alle sezioni di classificazione ATECO 2007 del registro ASIA Occupazione, distinte per macro-ripartizioni geografiche.

Tabella 2 – *Occupati per sezioni di classificazione ATECO 2007 e macro-ripartizioni geografiche. Italia 2016.*

Ripartizione geografica	ATECO 2007				Totale
	b-e	f	g-i	j-s ⁽¹⁾	
Nord-Ovest	1.552.254	414.862	1.800.626	2.050.289	5.818.031
Nord-Est	1.215.304	307.795	1.317.218	1.186.178	4.026.495
Centro	758.774	267.396	1.348.466	1.338.299	3.712.934
Sud	468.522	250.929	1.042.001	761.038	2.522.490
Isole	139.467	100.698	449.566	337.613	1.027.344
Totale	4.134.321	1.341.680	5.957.878	5.673.417	17.107.295

(1) Vedi Nota Tabella 1

Fonte: Ns elaborazioni su Istat, ASIA Occupazione 2016

Dall'esame della precedente tabella è agevole evincere che gli occupati censiti ammontavano nel 2016 a 17,1 milioni di unità, di cui il 24,2% nei settori estrattivo, manifatturiero e *public utilities* (sezioni B-E della classificazione ATECO 2007), il 7,8% nel settore delle costruzioni (sezione E), il 34,8% nel commercio, trasporto, magazzinaggio, alloggio e ristorazione (sezioni G-I) e il restante 33,2% negli altri servizi (sezioni J-S).

Con riferimento alle macro-ripartizioni, la distribuzione dell'occupazione fra settori produttivi si differenzia in maniera relativamente netta, con un'occupazione nei comparti estrattivo, manifatturiero e *public utilities* superiore alla media nazionale nel Nord-Ovest e nel Nord-Est (pari rispettivamente al 26,7% e al 30,2%), in linea con la media nazionale nel Centro e al di sotto della media nelle restanti macro-ripartizioni (18,7% al Sud e 13,6% nelle Isole).

Una situazione speculare si registra nel comparto delle costruzioni, con un livello di occupazione superiore a quello medio nazionale nelle macro-ripartizioni

meridionali – 9,9% al Sud e 9,8% nelle Isole – e inferiore nelle restanti – 7,1% nel Nord-Ovest, 7,6% nel Nord-Est e 7,2% al Centro.

Una situazione sostanzialmente analoga caratterizza la sezione di classificazione riferita al commercio, trasporti, magazzinaggio, alloggio e ristorazione, con le regioni centro-meridionali che presentano un'occupazione superiore a quella media nazionale – 36,3% al Centro, 41,3% al Sud e 43,8% nelle Isole – mentre una situazione opposta caratterizza le restanti macro-ripartizioni – 30,9% nel Nord-Ovest e 32,7% nel Nord-Est.

In ultimo, con riferimento agli altri servizi, il Nord-Ovest (35,2%) e il Centro (36%) presentano un'occupazione superiore a quella media nazionale, mentre un livello inferiore contraddistingue le restanti macro-ripartizioni (29,5% nel Nord-Est, 30,2% al Sud e 32,9% nelle Isole).

Tabella 3 – Occupati per sezioni di classificazione ATECO 2007 e macro-ripartizioni geografiche. Medie e grandi imprese. Italia 2016.

Ripartizione geografica	ATECO 2007				Totale
	b-e	f	g-i	j-s ⁽¹⁾	
Nord-Ovest	808.917	55.269	655.183	899.243	2.418.611
Nord-Est	589.674	34.225	362.090	430.668	1.416.657
Centro	321.204	27.050	451.832	527.931	1.328.018
Sud	157.224	22.518	144.080	174.802	498.625
Isole	36.237	7.497	57.081	72.402	173.217
Totale	1.913.256	146.560	1.670.266	2.105.046	5.835.128

(1) Vedi Nota Tabella 1

Fonte: Ns elaborazioni su Istat, ASIA Occupazione 2016

Sottolineata la rilevanza della dimensione di impresa nel determinare la probabilità di implementazione delle nuove tecnologie nel processo produttivo, appare utile riportare (Tabella 3) la distribuzione degli addetti nelle medie e grandi imprese, cioè quelle caratterizzate da almeno 50 addetti, in cui gli occupati si attestavano nel 2016 a 5,8 milioni di unità, di cui il 32,8% nei settori estrattivo, manifatturiero e *public utilities*, il 2,5% nel settore delle costruzioni, il 28,6% nel commercio, trasporto, magazzinaggio, alloggio e ristorazione e il restante 36,1% negli altri servizi.

Con riferimento alle macro-ripartizioni, l'occupazione nei comparti estrattivo, manifatturiero e *public utilities* è superiore alla media nazionale nel Nord-Ovest e nel Nord-Est (pari rispettivamente al 33,4% e al 41,6%), in linea con la media

nazionale nel Sud e al di sotto della media nelle restanti macro-ripartizioni (24,2% al Centro e 20,9% nelle Isole).

Una situazione speculare si registra nel comparto delle costruzioni, con un livello di occupazione superiore a quello medio nazionale nelle macro-ripartizioni meridionali – 4,5% al Sud e 4,3% nelle Isole – e inferiore nelle restanti – 2,3% nel Nord-Ovest, 2,4% nel Nord-Est e 2% al Centro.

Una situazione sostanzialmente analoga caratterizza la sezione di classificazione riferita al commercio, trasporti, magazzinaggio, alloggio e ristorazione, con le regioni centro-meridionali che presentano un'occupazione superiore a quella media nazionale – 34% al Centro, 28,9% al Sud e 32,9% nelle Isole – mentre una situazione opposta caratterizza le restanti macro-ripartizioni – 27,1% nel Nord-Ovest e 25,6% nel Nord-Est.

In ultimo, con riferimento agli altri servizi, il Nord-Ovest (37,2%), il Centro (39,7%) e le Isole (41,8%) presentano un'occupazione superiore a quella media nazionale, mentre un livello inferiore contraddistingue le restanti macro-ripartizioni (30,4% nel Nord-Est, 35,1% al Sud).

L'applicazione del vettore di probabilità stimato sulla base dello studio di Frey e Osborne alla distribuzione delle figure professionali per sezioni di classificazione ATECO 2007 nelle macro-ripartizioni geografiche, con riferimento alle medie e grandi imprese, consente quindi di quantificare l'impatto sul mercato del lavoro che si verificherebbe se le nuove tecnologie fossero già applicate nei comparti produttivi.

I principali risultati sono riportati nella Tabella 4. Se le nuove tecnologie fossero state già disponibili nel 2016, si sarebbero registrate rilevanti perdite occupazionali in tutti i comparti, ad eccezione di quello degli altri servizi nel Nord-Ovest. In particolare, a livello nazionale, gli occupati si sarebbero attestati a poco meno di 15 milioni di unità, di cui il 21,5% nei settori estrattivo, manifatturiero e *public utilities*, l'8,3% nel settore delle costruzioni, il 33,5% nel commercio, trasporto, magazzinaggio, alloggio e ristorazione e il restante 36,7% negli altri servizi.

Diversamente dai risultati ottenuti lo scorso anno, le perdite occupazionali, una volta confinati gli effetti delle innovazioni tecnologiche alle sole medie e grandi imprese, risultano relativamente mitigati, colpendo con maggiore severità il Nord-est (con una perdita occupazionale pari al 18,1%) e il Centro (-18,6%), come appare immediato desumere esaminando i tassi di variazione percentuale riportati nella successiva Tabella 5.

In un quadro di generalizzato calo dell'occupazione, l'impatto delle innovazioni tecnologiche si rivelerebbe particolarmente negativo soprattutto nei comparti estrattivo, manifatturiero e *public utilities* in tutte le macro-ripartizioni geografiche, soprattutto nelle regioni del Centro e del Sud, che registrerebbero contrazioni

comprese tra il 28,8 e il 31% degli addetti totali. Nel complesso, la perdita di posti di lavoro si attesterebbe, in questo scenario di diffusione dell'automazione, al 12,5%, una valutazione nettamente inferiore rispetto a quelle ottenute in altri studi.

Come si diceva, l'unico comparto che registrerebbe un notevole aumento dell'occupazione, nella misura del 17,1%, è quello degli altri servizi nella macro-ripartizione del Nord-Ovest. Come si osservava in un precedente lavoro dei due Autori (Bruno e Polli, 2017), è in tale ripartizione che si concentrano le figure professionali meno suscettibili di essere automatizzate, in quanto maggiormente caratterizzate da capacità analitiche e creative in senso lato.

Tabella 4 – *Stima dell'occupazione in uno scenario di piena operatività delle nuove tecnologie per sezioni di classificazione ATECO 2007 e macro-ripartizioni geografiche. Base Italia 2016.*

Ripartizione geografica	ATECO 2007				Totale
	b-e	f	g-i	j-s ⁽¹⁾	
Nord-Ovest	1.331.819	387.634	1.521.677	2.400.114	5.641.244
Nord-Est	923.831	285.127	1.089.707	996.970	3.295.636
Centro	523.443	247.847	1.068.484	1.181.900	3.021.675
Sud	333.484	234.013	935.101	634.206	2.136.803
Isole	104.539	93.846	397.718	271.571	867.673
Totale	3.217.115	1.248.467	5.012.686	5.484.762	14.963.031

(1) Vedi Nota Tabella 1

Fonte: Ns elaborazioni su Frey e Osborne (2013) e Istat, ASIA Occupazione 2016

Tabella 5 – Tassi di variazione percentuale dell'occupazione in uno scenario di piena operatività delle nuove tecnologie per sezioni di classificazione ATECO 2007 e macro-ripartizioni geografiche. Confronto con scenario base Italia 2016.

Ripartizione geografica	ATECO 2007				Totale
	b-e	f	g-i	j-s ⁽¹⁾	
Nord-Ovest	-14,20	-6,56	-15,49	17,06	-3,04
Nord-Est	-23,98	-7,36	-17,27	-15,95	-18,15
Centro	-31,01	-7,31	-20,76	-11,69	-18,62
Sud	-28,82	-6,74	-10,26	-16,67	-15,29
Isole	-25,04	-6,80	-11,53	-19,56	-15,54
Totale	-22,19	-6,95	-15,86	-3,33	-12,53

(1) Vedi Nota Tabella 1

Fonte: Ns elaborazioni su Frey e Osborne (2013) e Istat, ASIA Occupazione 2016

4. Conclusioni

Esaurita la fase esplorativa dell'indagine, appare chiaro che la valutazione delle ricadute occupazionali dell'introduzione dell'automazione intelligente nei processi produttivi debba inquadrarsi all'interno di un solido *framework* teorico, un aspetto generalmente trascurato nella letteratura esaminata dai due Autori.

A questo scopo, nelle future fasi della ricerca si procederà ad una *review* della letteratura economica, soprattutto allo scopo di accertare quali siano gli impatti occupazionali attesi, nel breve, medio e lungo termine, determinati dalla diffusione di tali innovazioni nei processi produttivi, che in prima approssimazione possiamo etichettare come *labour-saving*.

Parallelamente a questo filone d'indagine, per consolidare l'analisi già svolta appare necessario procedere ad una valutazione realistica dello stato dell'arte in tema di applicazioni industriali dell'intelligenza artificiale, allo scopo di accertare in quali fasi dei processi produttivi l'intelligenza artificiale sia già applicata con successo, quali siano le future prospettive di sviluppo di tali architetture e quali le effettive ricadute occupazionali per le figure professionali coinvolte, anche al fine di verificare l'attendibilità dei risultati ottenuti da Frey e Osborne (2013), che sono alla base della presente ricerca.

Riferimenti bibliografici

- BOSTON CONSULTING GROUP e MIT SLOAN MANAGEMENT REVIEW. 2017. *Is Your Business Ready for Artificial Intelligence?*.
- BOWLES J. 2014. *The Computerization of European Jobs*, Bruegel, Bruxelles. Disponibile online all'indirizzo <http://bruegel.org/2014/07/the-computerisation-of-european-jobs/>.
- BRUNO M., POLLI A. 2017. Impatti dell'automazione sul mercato del lavoro. Prime stime per il caso italiano. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXI, No. 3, pp. 53-64.
- FENG L, YING L., YONG S. 2017. *Intelligence Quotient and Intelligence Grade of Artificial Intelligence*. Chinese Academy of Sciences.
- FORRESTER RESEARCH. 2016. *The Future of White-Collar Work: Sharing Your Cubicle With Robots*. Cambridge MA: Forrester Research Inc.
- FREY C.B., OSBORNE M.A. 2013. *The Future of employment. How susceptible are Jobs to Computerisation?* Oxford Martin School, Working Paper. Disponibile online all'indirizzo <http://www.oxfordmartin.ox.ac.uk/downloads/academic/future-of-employment.pdf>
- GRAETZ G., MICHAELS G. 2015. *Robots at Work: the impact on productivity and jobs*, Centre for Economic Performance, London: London School of Economics.
- HERCULANO-HOUZEL S. 2009. *The human brain in numbers: a linearly scaled-up primate brain*. *Frontiers in Human Neuroscience*. Vol. 3, November 2009, pp. 1-11.
- ISTAT. 2013. *La classificazione delle professioni*. Roma: Istat.
- LOHR S. 2018. *Researchers Seek Smarter Paths to A.I.* The New York Times. June 21, p. 81. Disponibile online all'indirizzo <https://www.nytimes.com/2018/06/20/technology/deep-learning-artificial-intelligence.html>.
- ROTMAN D. 2015. *Who Will Own the Robots?*. Cambridge MA: MIT Technology Review.
- SERVICE NOW. 2017. *Research - Today's state of work: at the breaking point*.

SUMMARY

Automation and Labor Market. New estimates for Italy

The impact of automation on the labor market is one of the most debated issues in recent years, partly because of the remarkable achievements in the field of robotics and artificial intelligence. It is quite clear that a scientifically grounded discussion on these issues – above all in terms of lost jobs and the related impacts on welfare - cannot be separated from the theoretical context and the analysis of concrete situations.

Last year already, during the Wharton Global Forum 2017 held in Hong Kong, a suspicion arose among a panel of experts that the «stunning» developments in the domain of artificial intelligence were mostly a media creation and that this hype entailed a full misrepresentation of reality.

On the other hand, at a global level, some surveys show that the diffusion of artificial intelligence is rather weak and confined to a narrow range of products and production processes (BOSTON CONSULTING GROUP, 2017), while at national level the results are quite contradictory: our country ranks at the top for use of Internet of Things, in the face of a large majority of micro-enterprises relatively closed to any form of innovation.

Therefore, this paper aims to achieve new estimates of the impact of automation on the Italian labor market, differentiated by sector and macro-geographical areas, taking into account the size of enterprises that affects the propensity in implementing the new technologies, as confirmed by many studies.

IL LAVORO IDEALE

Nidia Batic

1. Premessa

Il lavoro è indubbiamente un elemento centrale nella vita di ognuno, ma per diverse ragioni non sempre è portatore di soddisfazioni e gratificazioni. È diventata prassi diffusa lamentarsi, ma perché solo recriminare? Quale sarebbe allora il lavoro ideale? Quello che non produrrebbe più lamentele? Quello che cambierebbe modo di rapportarsi al lavoro di così tante persone?

2. Obiettivi, strumento e metodo

L'obiettivo di questa ricerca è stato quello di sondare i desideri delle persone che lavorano e di quelle che non sono occupate, per verificare non solo quale potrebbe essere il loro lavoro ideale ma soprattutto quali sono le condizioni che lo rendono tale e quali sacrifici le persone sarebbero disposte ad affrontare pur di ottenerlo.

A tal fine è stato predisposto un questionario di 25 domande e le ipotesi di lavoro che ne hanno guidato la strutturazione intendono verificare se ci sono differenze generazionali e di genere nel modo di concepire il lavoro ideale. Il questionario esplora da un lato il lavoro svolto e dall'altro il lavoro ideale. Oltre ai dati fattuali si è chiesta l'attuale condizione occupazionale (o quella precedente al pensionamento), in termini di tipologia di rapporto e durata del contratto, il livello di soddisfazione per tale lavoro, i fattori che ne hanno condizionato la scelta e la chiarezza di idee già durante gli studi; la corrispondenza tra il lavoro svolto e i propri valori e la condizione psicologica in cui uno si sente più propenso a "vivere per lavorare" o a "lavorare per vivere". Per capire quali requisiti dovrebbe avere il "lavoro ideale", si è chiesto come dovrebbe essere (tipologia di rapporto, durata, distribuzione nella giornata e nella settimana, nel settore pubblico o privato). Per definire le caratteristiche del lavoro ideale sono state previste 25 situazioni/condizioni valutate su una scala di importanza da 1 a 5. Si è chiesto inoltre di indicare la propria predisposizione al rischio e una riflessione sul proprio

carattere in relazione con le scelte lavorative. A quanti non svolgono il lavoro ideale è stato chiesto di indicarne le ragioni, esplorando quindi la disponibilità a fare delle rinunce per poterlo ottenere.

Il questionario è stato somministrato¹ ad un campione ragionato, quindi non probabilistico, di persone² scelte secondo i seguenti vincoli: 1) persone con almeno 19 anni compiuti nel 2017; 2) 15 maschi e 15 femmine per ciascun intervistatore; 3) cinque soggetti per classe decennale di età (in tutto 6 classi), a partire da 19-28 anni, con classe aperta dai 69 anni in su.

Dopo aver sottoposto a controllo i questionari, ed aver eliminato quelli con risposte non coerenti, ne sono risultati validi 4.340. I missing values sono risultati complessivamente non superiori all'1% e distribuiti in maniera non sistematica, per cui non sono stati sostituiti e nelle statistiche descrittive le percentuali sono state calcolate sulle risposte valide.

Pur con i noti limiti di una indagine non probabilistica, si ritiene che i risultati ottenuti possano fornire interessanti spunti di riflessione sul fenomeno studiato.

3. Analisi dei dati e verifica d'ipotesi

3.1 Descrizione del campione

Il campione è risultato formato per il 49,9% da maschi e per il 50,1% da femmine, di fatto equiripartito tra le 6 classi d'età, con un'età media pari a 47 anni e 7 mesi per le femmine e 48 anni e 11 mesi per i maschi, compiuti nel 2017.

Il livello di scolarizzazione modale è rappresentato dal diploma di scuola secondaria di secondo grado (41,9%), sia per i maschi che per le femmine, e si registra una differenza statisticamente significativa nella distribuzione dei titoli di studio per genere³: più femmine con la licenza elementare (11,0% contro il 7,2% dei maschi) e con la laurea triennale (8,0% contro il 6,2% dei maschi); più maschi con la licenza media (22,5% contro il 18,1% delle femmine) e uguale percentuale di laureati magistrali tra maschi e femmine (14,2%).

Al momento dell'intervista il 47,3% era lavoratore dipendente e l'11,3% autonomo, circa un quinto degli intervistati era pensionato/a (20,1%) e il 21,3%

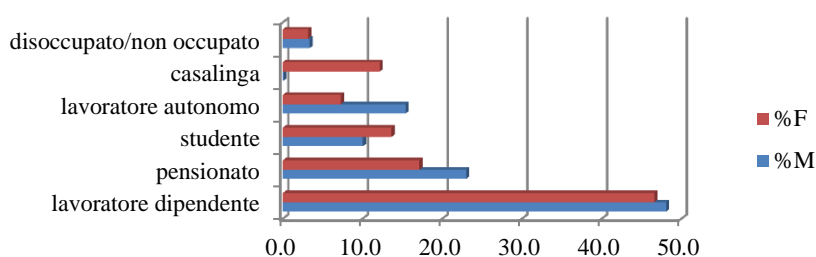
¹ Il questionario è stato consegnato alle persone che lo hanno compilato in presenza dell'intervistatore che ha letto via via le domande ad alta voce.

² Le interviste sono state effettuate da studenti del Corso di studi in Scienze della Formazione primaria dell'Università di Udine durante il mese di dicembre 2017 nelle province del Friuli Venezia Giulia e del Veneto. Gli intervistati sono stati scelti in modo da saturare le quote assegnate.

³ Test del $\chi^2 = 33,237$; il valore critico, con $\alpha = 0,01$, g.l. = 5, è pari a 15,086. Il livello di significatività utilizzato in tutti i test applicati nel presente lavoro è pari a 0,01 (per i test del χ^2 si rinvia a Ian Diamond e Julie Jefferies, 2006).

non era occupato (studenti (11,3%), casalinghe (6,1%), disoccupati o in cerca di occupazione (3,3%)), ed emerge una differenza di genere statisticamente significative⁴ nella composizione del campione (fig.1).

Figura 1 – Condizione occupazionale del campione.



Statisticamente significativa è la differenza nella tipologia contrattuale degli occupati a seconda del genere⁵: il 78,0% dei maschi e il 60,8% delle femmine lavorano a tempo pieno; il 27,7% delle femmine contro il 6,2% dei maschi ha un part-time; il 12,7% dei maschi e il 9,2% delle femmine è occupato a tempo continuato e infine il 3,1% dei maschi e il 2,3% delle femmine lavora a progetto. Non c'è invece differenza di genere⁶ nella durata del rapporto di lavoro, per l'84,5% a tempo indeterminato e il restante 15,5% a tempo determinato.

Oltre ai dati fattuali, è stato chiesto agli intervistati di fornire alcune informazioni di natura caratteriale che potessero essere correlate con le opinioni e le valutazioni espresse sul tema del lavoro. Essi si descrivono prevalentemente come ottimisti (78,0%), creativi (72,8%), riflessivi (63,6%) ed estroversi (60,6%) e si notano differenze di genere statisticamente significative⁷: i maschi risultano più ottimisti (79,9% contro il 76,1% delle femmine) mentre le femmine sono più estroverse (63,4% contro il 57,7% dei maschi).

Le persone sono convinte e consapevoli che il loro carattere ha influito "abbastanza" o "molto" sulle loro scelte lavorative nell'85,8% dei casi, e non sono emerse differenze statisticamente significative riconducibili al genere⁸.

Per completare le note caratteriali, scopriamo che l'88,3% "lavora per vivere", mentre il restante 11,7% dichiara di "vive per lavorare", e sono più gli uomini (12,9%) rispetto alle donne (10,3%) a mettere il lavoro al centro delle loro vite⁹.

⁴ Test del $\chi^2 = 351,026$; il valore critico, con $\alpha = 0,01$, g.l. = 5, è pari a 15,086.

⁵ Test del $\chi^2 = 265,268$; il valore critico, con $\alpha = 0,01$, g.l. = 3, è pari a 11,345.

⁶ Test del $\chi^2 = 5,544$; il valore critico, con $\alpha = 0,01$, g.l. = 1, è pari a 6,635.

⁷ χ^2 ottimisti/pessimisti per genere = 9,2 chi quadro introversi/estroversi per genere = 14,66, per ambedue il valore critico con $\alpha = 0,01$ e g.l. = 1, è pari a 6,635.

⁸ Test del $\chi^2 = 4,1$; il valore critico, con $\alpha = 0,01$, g.l. = 3, è pari a 11,345.

3.2 Rapporto con il proprio lavoro

Tra le persone che lavorano o che sono in pensione, il 45,3% è riuscito a fare il lavoro che pensava di fare alla fine del percorso di studi, senza distinzione di genere.

Alla domanda “da quali fattori esterni è condizionato il lavoro svolto” (attuale o prima del pensionamento), il 51,9% ha risposto che è stata una libera scelta, ma con maggiore libertà per gli uomini (54,8%) rispetto alle donne (48,4%), mentre il rimanente 48,1% ha dovuto assecondare altre esigenze, e in questo le donne ne hanno risentito di più rispetto agli uomini, su tutti i fronti (tab.1).

Tabella 1 – *Fattori che hanno condizionato il lavoro svolto attualmente o prima del pensionamento, per genere (valori percentuali).*

	Maschi	Femmine	Totale
Nessuno, è stata una libera scelta	54,8	48,4	51,9
Necessità economiche	27,1	32,9	29,7
Famiglia	16,7	23,1	19,6
Mancanza di opportunità diverse	16,7	19,1	17,8
Titolo di studio	16,1	19,3	17,6
Sede facilmente raggiungibile	10,3	11,5	10,8
Motivi logistici	3,9	4,0	4,0
Eredità	3,7	1,7	2,8
Motivi di salute	1,0	0,9	0,9
Condizioni estetiche	0,4	0,2	0,3

Prima di parlare del lavoro ideale è interessante scoprire se il campione intervistato è soddisfatto di quello che sta facendo o di quello che faceva prima del pensionamento. A tal fine è stata utilizzata una scala a 10 posizioni (da 1 a 10) e il livello medio di soddisfazione si attesta su 7,7 con un range che oscilla tra 7,0 e 8,5, quindi tendenzialmente elevato e con una variabilità molto contenuta (scarto quadratico medio pari a 1,743), che tuttavia fa emergere due distinti profili, sopra e sotto la media (tab. 2). I più soddisfatti sono quanti dichiarano di “vivere per lavorare” e chi fa il lavoro che pensava di fare alla fine degli studi, mentre i meno soddisfatti sono i pessimisti e quanti non fanno il lavoro che avrebbero voluto quando hanno terminato gli studi.

⁹ Il test del χ^2 fornisce un valore pari a 5, statisticamente non significativo per un livello di significatività pari ad $\alpha=0,01$, g.l.=1 e valore critico pari a 6,635.

Tabella 2 – Livello medio di soddisfazione per il proprio lavoro per condizione.

Maschio	7,7	Lav. dipendente	7,4	Ottimista	7,9
Femmina	7,6	Lav. autonomo	8,1	Pessimista	7,0
19-28	7,5	Pensionato/a	8,1	Introverso	7,5
29-38	7,5	A t. determinato	7,4	Estroverso	7,8
39-48	7,6	A t. indeterminato	7,7	Creativo	7,8
49-58	7,5	A progetto	7,9	Non creativo	7,5
59-68	7,8	A tempo pieno	7,7	Impulsivo	7,6
69 e oltre	8,1	Part-time	7,4	Riflessivo	7,8
Elementari	7,9	Tempo continuato	8,0	Fa il lavoro che pensava di fare alla fine degli studi	8,2
Medie	7,6	Vivo per lavorare	8,5	Non fa il lavoro che pensava di fare alla fine degli studi	7,2
Superiori	7,5	Lavoro per vivere	7,6		
Laurea triennale	7,7	Rispecchia i propri valori	8,1		
Laurea magistrale	7,9				

Applicando il test del chi quadrato è emerso che le distribuzioni dei giudizi, all'interno della scala di valutazione, sono statisticamente diverse tra tutte le modalità di classificazione del campione, con la sola eccezione del genere¹⁰. È stata inoltre individuata la valutazione mediana che è pari a 8 per ciascuna mutabile, che assume valori diversi solo per i pessimisti (7), coloro che svolgono il lavoro che avevano progettato alla fine degli studi (7,5) e quelli che “vivono per lavorare” (9).

Per verificare se la probabilità di essere soddisfatti per il proprio lavoro è riconducibile ad un insieme di variabili esplicative osservate, è stato applicato un modello logit (Delvecchio, 2002) in cui la variabile dipendente “livello di soddisfazione” è stata dicotomizzata (giudizi da 1 a 5 = 0 insoddisfatto, giudizi da 6 a 10 = 1 soddisfatto) (tab. 3). Nella tabella 3 sono presentate le variabili considerate e tra parentesi i riferimenti (*baseline*) utilizzati nel modello. Considerando i livelli di significatività minori o uguali a 0,05, emerge che la probabilità di essere più soddisfatti del proprio lavoro non è riconducibile al genere né all'età né al titolo di studio, mentre per chi fa un lavoro in armonia con i propri valori e ideali è 8 volte più elevata rispetto a chi non lavora in tali condizioni (*odds ratio*=8,365). La propensione alla soddisfazione è maggiore anche per: i pensionati (*odds ratio*=2,895); chi vive per lavorare rispetto a chi lavora per vivere (*odds ratio*

¹⁰ I test del χ^2 applicati alle diverse modalità dei caratteri osservati hanno fatto emergere risultati statisticamente significativi per $\alpha = 0,01$ e i gradi di libertà riferiti alle diverse modalità. Unica eccezione il genere, con un $\chi^2 = 12,384$, per cui il valore critico per $\alpha = 0,01$ e g.l. = 9 è pari a 21,666 e il caso border-line di chi lavora tempo determinato o indeterminato per cui il $\chi^2 = 20,849$ è non significativo per $\alpha = 0,01$ (valore critico 21,666).

Tabella 3 – Modello Logit, livello di soddisfazione per il proprio lavoro (baseline).

	Stime dei parametri	Errori standard	Valore z	Pr(> z)	
Intercetta	5,218	1,399	3,729	0,000	***
Femmina (Maschio)	0,023	0,161	0,146	0,884	
ETÀ (età 19-28)					
Età 29-38	-0,181	0,327	-0,554	0,580	
Età 39-48	0,172	0,347	0,495	0,620	
Età 49-58	-0,329	0,340	-0,968	0,333	
Età 59-68	-0,618	0,361	-1,713	0,087	
Età 69 e oltre	-0,923	0,525	-1,757	0,079	
TITOLO DI STUDIO (Altri titoli)					
Licenza elementare	0,145	0,446	0,324	0,746	
Licenza media	0,023	0,322	0,071	0,943	
Diploma superiore	-0,136	0,303	-0,450	0,652	
Laurea triennale	0,240	0,425	0,564	0,573	
Laurea magistrale	0,298	0,361	0,825	0,410	
RAPPORTO CON IL LAVORO					
Lavoro in armonia con i propri valori (Non in armonia)	2,124	0,159	13,326	0,000	***
È il lavoro che pensava di fare alla fine degli studi (Non è il lavoro che pensava di fare alla fine degli studi)	0,503	0,194	2,589	0,010	**
Vivo per lavorare (Lavoro per vivere)	0,890	0,374	2,381	0,017	*
CONDIZIONE ATTUALE (Lavoratore dipendente)					
Lavoratore autonomo	-0,173	0,266	-0,651	0,515	
Pensionato	1,063	0,380	2,802	0,005	**
LAVORO IDEALE					
Lavoro a tempo indeterminato (Tempo determinato)	0,408	0,196	2,084	0,037	*
Lavoro a tempo pieno (Lavoro a progetto)	-1,100	0,666	-1,650	0,099	
Lavoro part-time (Lavoro a progetto)	-1,369	0,681	-2,011	0,044	*
Lavoro a tempo continuato (Lavoro a progetto)	-0,991	0,708	-1,398	0,162	
FATTORI CHE HANNO CONDIZIONATO LA SCELTA LAVORATIVA (dicotomico)					
È stata una libera scelta	0,532	0,216	2,464	0,014	*
È stato influenzato dalla famiglia	-0,006	0,185	-0,034	0,973	
Necessità economiche	-0,310	0,167	-1,855	0,064	
Mancanza di opportunità diverse	-0,286	0,170	-1,687	0,092	
Titolo di studio	-0,035	0,190	-0,182	0,856	
Eredità	0,968	0,779	1,244	0,214	
Sede raggiungibile	0,256	0,216	1,182	0,237	
Motivi logistici	-0,001	0,334	-0,002	0,998	
Problemi di salute	-0,681	0,570	-1,196	0,232	
PREDISPOSIZIONE AL RISCHIO (Niente)					
Poco	0,102	0,354	0,289	0,772	
Abbastanza	0,190	0,344	0,552	0,581	
Molto	0,339	0,374	0,905	0,366	
CARATTERE					
Ottimista (Pessimista)	0,599	0,162	3,690	0,000	***
Estroverso (Introverso)	0,077	0,159	0,486	0,627	
Creativo (Non creativo)	0,145	0,163	0,890	0,373	
Riflessivo (Impulsivo)	0,466	0,153	3,035	0,002	**

Livello di significatività: ***=0,001; **=0,01; *=0,05

=2,495); gli ottimisti (*odds ratio*=1,820); quelli che hanno scelto liberamente il lavoro da fare (*odds ratio*=1,702); chi fa il lavoro che pensava di fare quando ha finito gli studi (*odds ratio*=1,654); chi è riflessivo (*odds ratio*=1,594) e chi ha un lavoro a tempo indeterminato (*odds ratio*=1,504). Chi invece ha un lavoro a part-time ha una probabilità di essere insoddisfatto circa 4 volte maggiore di chi lavora a progetto (*odds ratio*=3,931).

3.3 Il lavoro ideale

Tutti, occupati, pensionati e non occupati, hanno descritto quello che per loro sarebbe o sarebbe stato il lavoro ideale. Le scelte espresse dagli intervistati sono state confrontate con il genere, l'età, il titolo di studio, la condizione occupazionale e l'indicazione se la persona vive per lavorare o lavora per vivere. Un primo risultato che colpisce è la presenza di molti "indifferenti", per cui nella tabella 4 si è scelto di ordinare secondo questo criterio la presentazione dei dati generali, mentre nel testo, seguendo tale ordine, si forniscono ulteriori dati ritenuti più significativi. In generale il settore privato è quello preferito, fatta eccezione per i lavoratori dipendenti (25,1%), le casalinghe (26,6%) e le donne (28,5%), che sceglierebbero un impiego pubblico. Il 54,3% reputa importante la presenza di particolari agevolazioni per motivi familiari; in questo le donne (65,9%) sono più sensibili degli uomini (42,7%) e in particolare il 71,5% di tutte le casalinghe.

Tabella 4 – Caratteristiche del lavoro ideale (valori percentuali).

Privato 26,5	Pubblico 23,4				Indifferente 50,1
Con agevolazioni per motivi familiari					Indifferente 45,7
Settimana corta 52,5	Settimana lunga 10,4				Indifferente 37,0
Dipendente 35,9	Non dipendente 28,1				Indifferente 36,0
Mattina 59,3	Pomeriggio 33,4	A turni 10,4	Sera 4,2	Notte 1,2	Indifferente 30,7
Tempo pieno 44,9	Part-time 26,1	Tempo continuato 6,4			Indifferente 22,6
Orario flessibile 47,6	Orario fisso 30,3				Indifferente 22,1
T. indeterminato 79,7	T. determinato 4,0	A progetto 3,9			Indifferente 12,4

In merito alla distribuzione delle giornate lavorative nella settimana c'è una tendenziale preferenza delle donne per la settimana corta (58,4% contro il 46,7% dei maschi), e con riferimento all'età, man mano che si avanza negli anni si registra un aumento di interesse per la settimana lunga dal 9,9% al 14,7%.

Per quanto riguarda il rapporto di lavoro, si nota che le donne sono più propense ad un lavoro dipendente (42,0%) rispetto agli uomini (29,8%), mentre tra le diverse età l'interesse per il lavoro autonomo passa dal 22,9% dei soggetti con 19-28 anni al 32,5% degli ultra 68enni. Con riferimento alla condizione occupazionale, il 57,9% dei lavoratori autonomi continua a considerare questa la condizione di lavoro ideale, e lo stesso vale per il 40,8% dei dipendenti che preferirebbero continuare ad esserlo. Chi "lavora per vivere" si orienta di più verso un lavoro dipendente (37,4%), chi "vive per lavorare" nel 39,4% dei casi preferisce un lavoro non dipendente.

La fascia oraria preferita dalle donne è al mattino (68,0%) contro il 50,9% degli uomini e nella scelta della durata del lavoro, il tempo pieno è sostenuto dal 56,2% dei maschi e dal 33,6% delle femmine, che sono invece più orientate ad un lavoro part-time (40,3%) rispetto all'11,9% dei maschi. In particolare il 55,1% delle casalinghe immaginerebbe un'occupazione part-time mentre il 56,1% dei pensionati pensa al lavoro ideale a tempo pieno.

Riguarda l'orario di lavoro, prevale l'interesse per un orario flessibile, anche se e a livello anagrafico nel tempo questa esigenza diminuisce progressivamente, passando dal 53,7% dei più giovani (età 19-28) al 35,9% dei più grandi (69 e oltre). La preferenza per l'orario flessibile è inoltre associata al livello di istruzione e diminuisce passando dai laureati magistrali (57,5%) a chi possiede la licenza elementare (34,1%). I più contrari all'orario fisso sono i lavoratori autonomi (solo il 19,0% lo sceglie) mentre ne sono più convinte le casalinghe (39,5%).

Per concludere, molto più chiare sono le idee riguardanti la durata del rapporto di lavoro, con solo il 12,4% di indifferenti e il 79,7% di scelte per il tempo indeterminato.

Dopo aver delineato il profilo del proprio lavoro ideale, è stato chiesto quanto siano importanti, nella ricerca di un lavoro ideale, determinati requisiti di natura sia relazionale che logistico-organizzativa. È stata utilizzata una scala di importanza tipo Likert a cinque posizioni (da 1 a 5) e sono state calcolate le risposte mediane per ciascun item (tab. 5). Come si può notare, i primi requisiti per un lavoro ideale non riguardano il compenso o la carriera, la comodità di averlo sotto casa o il prestigio. Ciò che conta ($M_e=5$) è in primo luogo il benessere personale che deriva dalla soddisfazione per quello che si fa, dalla condivisione di valori quali il rispetto, l'onestà e la trasparenza, e dalla presenza di buoni rapporti con i colleghi. Inoltre molto importante per definire un lavoro come ideale è la presenza di superiori competenti. Decisamente meno gradite di tutte sono le rinunce personali ($M_e=2$).

Tabella 5 – *Importanza di talune condizioni nella definizione del Lavoro ideale.*

Mediana	Condizioni nella definizione del Lavoro ideale
5	La soddisfazione per il proprio lavoro - Ambiente dove si ritrovano valori come rispetto, onestà, trasparenza - Buoni rapporti con i colleghi - Avere superiori competenti
4	Il compenso - Ambiente fisico accogliente - Lavoro dinamico - A contatto con le persone/utenti/clienti - Lavoro che consente di mettersi alla prova - Vicino a casa - Lavoro creativo - Avere responsabilità - Lavoro non stressante - Avere spazio fisico per sé - Lavorare in gruppo - Lavoro socialmente utile - Avere la possibilità di dirigere/coordinare
3	Lavoro intellettuale - Fare carriera - Lavoro manuale - Lavoro che dà prestigio - Al chiuso - All'aperto - Avere rapporti con l'estero
2	Lavoro in cui si devono fare delle rinunce personali

Nella loro struttura, le domande relative all'importanza di una batteria di fattori nella definizione del lavoro ideale, sono state assimilate alle scale aggettivali bipolari utilizzate per il differenziale semantico (Osgood, 1957) in cui il concetto da valutare è stato "il lavoro ideale", solo che in questo caso non sono state esplicitate le condizioni "opposte" (ad esempio per l'opzione "vicino a casa" l'estremo opposto della scala è "lontano da casa"). Per tale motivo le stesse scale di valutazione sono state trattate come variabili numeriche discrete e sottoposte ad analisi fattoriale¹¹ e sono stati estratti 7 fattori che spiegano il 56,31% della variabilità osservata. Il primo può essere definito il fattore della *comodità*, rappresentato dal desiderio di stare vicino a casa, stare in un ambiente accogliente, non stressante ed avere uno spazio lavorativo per sé. Il secondo è un fattore *dinamico/gratificante*, descritto dalla soddisfazione per il proprio lavoro, che deve essere anche dinamico, dove si può mettersi alla prova, creativo, in cui si hanno delle responsabilità. Il terzo è il fattore *funzionale/gestionale*, dove si colloca l'esigenza di poter lavorare in gruppo, di avere superiori competenti, dove si ha la possibilità di dirigere e coordinare e avere buoni rapporti con i colleghi. Il quarto fattore riguarda la *carriera*: possibilità di ricevere un buon compenso, fare carriera, avere prestigio. Segue un fattore che guarda alla *tipologia di lavoro*: intellettuale e in cui si hanno anche rapporti con l'estero. Il sesto fattore è rappresentato dalla *disponibilità a fare delle rinunce* e il settimo dalla possibilità di *lavorare all'aperto*.

Se questi sono i requisiti del lavoro ideale, le persone sono riuscite a realizzarlo? ed eventualmente quali sono le ragioni per cui non hanno potuto farlo? Il 41,0% degli intervistati (42,7% uomini e 38,9% donne) dichiara una elevatissima corrispondenza tra lavoro attuale (o precedente al pensionamento) e il lavoro ideale

¹¹ È stata applicata l'analisi delle componenti principali e la matrice dei fattori è stata ruotata con il metodo Varimax al fine di rendere perpendicolari le scale ai fattori. Sono stati estratti i fattori con auto valore maggiore o uguale a 1 (Fabbris, 1990).

(o addirittura lo sta già facendo). Queste persone non hanno continuato il questionario, mentre il rimanente 59.1% ha spiegato che la ragione principale per cui il proprio sogno lavorativo non si è realizzato è da imputare ad un titolo di studio inadeguato per il 41,4% delle donne e il 34,6% degli uomini e da problemi di natura familiare per il 42,5% delle donne contro il 28,2% degli uomini. A scalare ci sono stati motivi logistici (20,4%) o la mancanza di motivazione (11,5%). Pochi hanno rinunciato al lavoro ideale perché sottopagato (8,3%) e pochissimi hanno addotto motivi di salute (2,8%) o ragioni estetiche (0,5%, pari a 10 persone).

Si è visto che nella classifica dei requisiti del lavoro ideale, all'ultimo posto si colloca la disponibilità a pensare ad un lavoro in cui sono necessarie delle rinunce personali, e in una domanda di controllo finale si è chiesto a cosa sarebbero disposti a rinunciare pur di poter realizzare il sogno del lavoro ideale. Utilizzando una scala qualitativa (da 0=per niente a 3=molto) e calcolando al mediana delle risposte, emerge che sarebbero "abbastanza" disponibili a rinunciare all'orario fisso, alla posizione nel lavoro, alla responsabilità, al tempo libero e alla vicinanza a casa. "Poco" disponibili invece a rinunciare alla vicinanza alla famiglia, ad una buona retribuzione e alla sicurezza economica, alla sicurezza lavorativa e alle relazioni sociali. "Per niente" disponibili a rinunciare a formarsi una famiglia in cambio del lavoro ideale.

Diverso è l'atteggiamento verso la possibilità di occupare una posizione subordinata pur di fare il lavoro ideale: solo il 9,3% non si adatterebbe (i più decisi sono l'11,6% dei maschi contro il 6,8% delle femmine), mentre il 44,4% risponde con un sì incondizionato ed il restante 46,3% accetterebbe ma solo per il tempo necessario per imparare il nuovo lavoro.

Ciò che attirerebbe "abbastanza" o "molto" gli insoddisfatti sarebbe un lavoro con maggiori responsabilità (82,8%), con una differenza di genere statisticamente significativa¹²: sono più propensi gli uomini (85,8%) che le donne (79,3%).

Molto elevata è la disponibilità a partecipare a corsi di aggiornamento (92,8%), molti di meno sono disposti a spostarsi da casa (53,8%). Tra questi, il 44,0% si sposterebbe al massimo entro il territorio della propria regione e poco più di un terzo andrebbe anche all'estero, con una netta differenza di genere (il 61,4% dei maschi e il 45,4% delle donne). Evidenti differenze sono anche ascrivibili all'età: se i giovani tra i 19 e i 28 anni nell'82,5% dei casi sono disposti alla mobilità, nel tempo si contrae tale disponibilità fino ad arrivare al 41,7 degli ultra 68enni.

¹² Test del $\chi^2 = 35,310$; il valore critico, con $\alpha=0,01$, g.l.=3 è pari a 11,345.

4. Conclusioni

Qual è il lavoro ideale? Quali condizioni lo rendono tale? Quali sacrifici sarebbero disposti ad affrontare le persone per poterlo fare?

I 4.340 intervistati hanno manifestato grande indifferenza nel momento in cui dovevano esprimersi sul loro lavoro ideale e resta il dubbio se interpretare questo dato come mancanza di idee o grande capacità di adattamento, ma i dati disponibili non consentono di verificare tali ipotesi.

I requisiti che rendono il lavoro “ideale” hanno registrato una forte convergenza di opinioni, pur all’interno di una certa variabilità. A livello dichiarato è per tutti prioritaria la soddisfazione per il proprio lavoro seguita a ruota dalla componente valoriale e relazionale. Solo in seconda battuta, sebbene non poco importante, si pensa alla competenza gestionale e al compenso. A livello emozionale/inconscio, il primo fattore estratto applicando l’analisi fattoriale rappresenta invece la *comodità*.

Il pensiero al lavoro ideale non fa scattare una elevata disponibilità rinunciare a qualcosa che si ha già, ma porta a considerare la possibilità di avere maggiori responsabilità o partecipare a iniziative di aggiornamento.

Potendo scegliere, solo poco più della metà accetterebbe di percorrere grandi distanze per raggiungere il lavoro ideale ma, dovendo scegliere, poco meno della metà non sarebbe disposto a uscire dai confini regionali!

E allora? Si può dire che il lavoro ideale dev’essere dipendente e con contratto a tempo indeterminato, a tempo pieno e più nel settore privato che nel pubblico; con orario flessibile e preferibilmente al mattino, in settimana corta e con agevolazioni per i lavoratori per motivi familiari. Un lavoro gratificante, in un ambiente in cui sono importanti i valori e le relazioni umane; un lavoro che non comporti sacrifici ma per cui si è disposti ad investire per imparare, e ... preferibilmente sotto casa!

Riferimenti bibliografici

- DELVECCHIO F. 2002. *Statistica per la ricerca sociale*. Bari: Cacucci Editore.
DIAMOND I., JEFFERIES J. 2006. *Introduzione alla statistica per le scienze sociali*. Milano: McGraw-Hill.
FABBRIS L. 1990. *Analisi esplorative di dati multidimensionali*. Padova: CLEUP.
OSGOOD C.E., SUCI G.J., TANNENBAUM P.H. 1957. *The Measurement of Meaning*. University of Illinois: Urbana.

SUMMARY

The perfect job

4.340 people from Friuli Venezia Giulia and Veneto (in northeast Italy) were interviewed to verify what is the ideal job, its characteristics and what people are willing to give up so they can do it. The sample for quotas is non-probabilistic, however the results provide interesting food for thought. A structured questionnaire was used with closed or structured questions and some evaluation scales. The data were submitted to descriptive analysis and on the evaluation scales of the elements that contribute to defining the work as "ideal", factorial analysis was applied, while the level of satisfaction for one's work was related to the independent variables by applying a Logit model. It emerges that the respondents are quite satisfied with their work, however "only" two fifths find a very high correspondence between the job done and the ideal one. But what is the ideal job? Men and women, young and old, people with different qualifications and different working conditions, have provided answers that are sometimes very different, but basically converging on the following profile: the ideal job must be dependent rather than autonomous and with a permanent contract; full-time and in the private sector rather than in the public; with flexible hours and preferably in the morning; in a short week and where facilities are provided for workers with family problems. The requirements that make a job "ideal", have seen a strong convergence of opinions, within a certain variability. At the declared level, the most important requirement of the ideal job is, for everyone, to do a rewarding job, to work in an environment where there are values such as respect, honesty and transparency and there are good human relationships; then after people think about management competence and salary. Factor analysis was applied to bring out the emotional dimension and the first factor extracted is identified as a comfort factor, followed by the factor of dynamism and the functional/managerial factor. To conclude, the ideal job should not involve sacrifices but for it people are willing to invest to learn, and in any case it must be ... preferably near home!

THE LABOUR REGISTER IN ITALY: THE NEW HEART OF THE SYSTEM OF LABOUR STATISTICS

Ciro Baldi, Claudio Ceccarelli, Sara Gigante, Silvia Pacini, Francesca Rossetti

1. The role of the Labour Register in the labour statistics

Modern statistical production is more and more based on registers, micro level databases that comprise all units belonging to the target population. This is the way many NSIs are facing the challenge of improving the efficiency of the production processes in order to release integrated set of indicators, in a context of rapid growth of demand of statistical information. These are, in fact, the key points of the strategy of the European Statistical System for the next decade (ESS, (2014)).

In Italy, the Istat modernization process (Istat, 2016) is based on a new Integrated System of Register. Among them the Labour Register (LR) is a link between the Base Register of the Productive Units and that of Population and the new heart of the labour statistics. It is a micro level database with a Linked Employer Employee Database (LEED) structure with multiple information on labour activities.

With a very complex and rich structure of information, it aims to be the basis for most labour market indicators and analysis, requested by international regulations and national needs, either directly using its information for the output, or indirectly, coupled with surveys, to estimate variables not present in it or that do not perfectly fit the statistical definitions.

More in detail, the LR can provide information to calculate a considerable number of labour indicators such as those requested by labour statistics regulations both on business side (LCS, SES, LCI, GPG etc..) and on social one (LFS, Eu-Silc) and by other household (e.g the population census) and business statistics (such as SBS, STS...). In addition to providing variables that can be directly used (that is register based statistics), this register can provide auxiliary information to greatly enhance various phases of already in place survey based processes such as the sampling design, the E&I, the estimation. The integration will naturally be developed into mixed survey register processes by redesigning the first in order to complement the second, thus reaching the considerable advantage of "slimming down" the surveys and reducing the response burden, especially of enterprises. Besides being used for the production of aggregate indicators, the LR can become

the basis for the release of longitudinal samples of work relationships and/or workers from which a wealth of econometric studies can spur and an important database for the evaluation of labour policy programs.

To the extent of a deeper integration within the system, placing a Labour Register at the heart of the domain of labour statistics can bring about what is probably its most remarkable advantage: the capacity of generating an augmented degree of coordination of the different processes and consistency among the various output of the system.

2. The Labour Register: characteristics and its main pillars

2.1 Structure and variables

Being a theme focused register, the main aim of the LR is to provide as much information as possible at the level of each work relationship that links an employer to an employee. This fundamental statistical unit is then the basis to reconstruct the information at the level of employee or at the level of employer. It aims to cover the whole population of work relationships, in the entire national territory, in every sector of economic activity and all types of work relationships either characterizing the dependent work or the not dependent one (even when the employer and the employee refer to the same person). Its main variables are the type of work relationship, the type of contract, the type of working time, measure(s) of employment, the number of hours paid/and or worked, the wages (gross and possibly net) for each pay period (typically the month); the social contributions (in the next future). The structure of the database is longitudinal that is the same work relationship are followed over time; this implies to establish the moment the work relationship start and the moment it ends. Finally, the register has to be easily linked with the other registers of the system in particular the Business Register to retrieve employer level information (such as the NACE sector, the localization of the economic units or the local unit where the worker works at, etc..), with the Population Register to retrieve worker level information (sex, age, education, etc..) and also with the Income Register and so on.

2.2 The metadata system

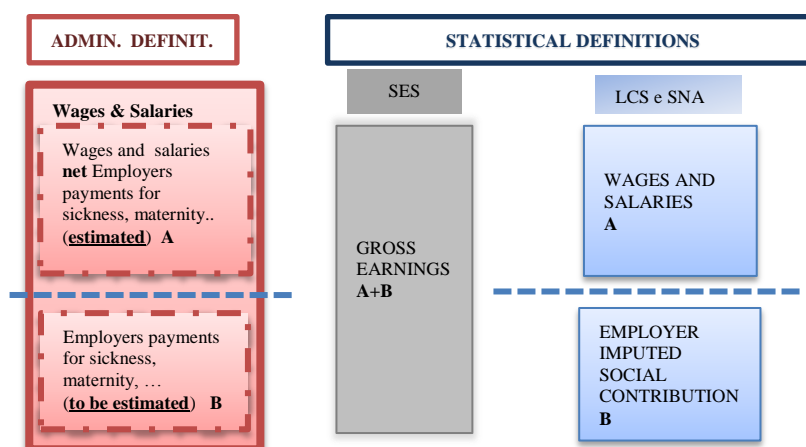
The centrality of LR in the Labour statistics is based on the possibility to satisfy different information needs from a single integrated database. The process that, starting from the information contained in the various administrative sources, allows the estimation of the variables and indicators is entirely guided by a complex system of metadata.

The register should be able to produce statistics that meet different regulations, with different classifications, indicators, variables and, sometimes, different definitions for similar concepts.

The metadata system needs first to evaluate for each atomic statistical item needed the availability in the administrative data or the possibility to calculate or estimate it, also via other administrative and statistical sources. Secondly, it has to map the administrative input data into the statistical output figures.

The variables “wages and salaries” and “earnings”, required respectively by the LCS and ESA regulations on one side and SES on the other are an example: they refer to very similar concepts with a slightly different definition. The main differences are related to the payments by the employer to the employee in the face of sickness, maternity, accidents etc., which are included in the definition of “Earnings” but excluded from the concept of “Wages and salaries”. The “Employers payments for sickness, maternity, accidents, etc..” represents an atomic statistical item in the metadata system. It is combined with “wages and salaries net employers payments for sickness, maternity” in order to satisfy the SES definition of “earnings” while it is excluded from “wages and salaries” required by LCS regulations whereas it is one of the elements of “employer imputed social contribution” (Figure 1).

Figure 1 – An example of the atomic items necessary to satisfy different statistical definitions: a synthetic description.



Another example of the required flexibility to cope with difference in classification across statistics is related to the status of employment. The project workers are classified as self-employed in the context of NA, as outworkers with respect to the economic units they work for in the BR and dependent contractors

within the more general category of employees, in the incoming International classification of status in employment (ICSE 2018). Therefore, the same input category has to be mapped differently in the different outputs in order to fulfil the different standards.

The metadata system construction can thus be described as a three-step process: a) a recognition of all the statistical requirements to produce an output metadata system; b) the collection of the metadata of various input sources and their integration in a input metadata system that represent the tax-social security reality in a unified way; c) the mapping of the input metadata into the output one.

2.3 *The basic statistical unit and measurement*

In order to satisfy the variety of needs stemming from the social and economic studies, the LR will allow analysis based on three main statistical units: the individual, the economic unit, and the job position or work relationship. This last one, very specific of this register, is the position involving one or two subjects to perform a labour activity.

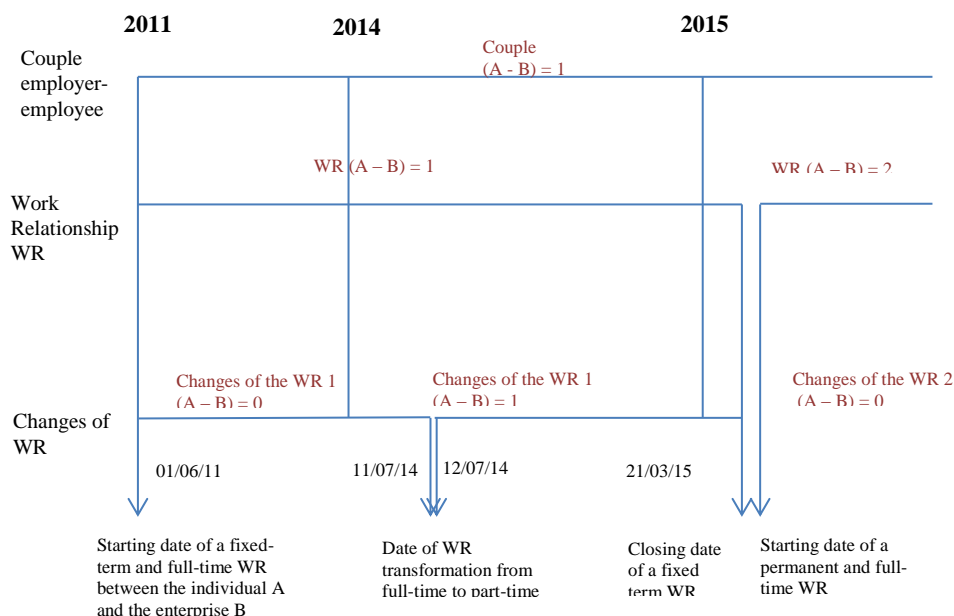
The job position is defined as the work relationship established between an economic unit and an individual and defined by a starting date. An important attribute of the job position is the ending date, which will be lacking up to the moment of the cessation of the relationship. In the context of dependent (or assimilated) employment, this statistical unit corresponds to the employment contract between an employer and an employee where the starting date is the date of activation of the contract and the ending date its cessation. Moreover, each job position is characterized by a wealth of other attributes such as the type of contract, the working time, the contractual occupation, the workplace, etc..¹

The basic unity of the register is the result of a process of control and correction that has the objective to resolve cases of:

- a. incoherence dates assumption and date cessation
- b. duplicated declarations
- c. juridical transformations
- d. more events of assumption e/o cessation inside every month
- e. dependent job positions totally or partially overlapped between the same enterprise and the same worker

Figure 2 represents, for instance, the evolution of an hypothetical relationship among the individual A and the enterprise B.

¹ Since some of these attributes can change within the duration of a job relationship, the LR will also allow analysis on the temporal portion of the job homogeneous for a number of these attributes: type of working time, type of contract, and occupational qualification. This might be very important to evaluate the contract transformations e.g. from fixed term to permanent, that do not constitute an interruption of the same position.

Figure 2 – Schematic description of the LR base statistical unit.

The rationale for defining this statistical unit in this way is multiple. First, it is the basis for analysis of all labour market policies that have an impact on the definition of contracts or birth of new ones. Second, being the natural link with the Compulsory communication (Cc) unit, it is the basis of the extension of the register to the information contained in this source especially useful for short-term estimates of employment dynamic (Rapiti et al, 2018). Third, the duration of the job position is an essential policy parameter as it makes possible to better evaluate the fragmentation of the employment careers with all its implications.

All in all, this statistical unit permits a very rich stock-flow accounting connecting on one side levels of employment, gross and net changes and on the other workers and jobs. Fourth, the business based labour statistics measure parameters referred to job position.

The passage from a statistical unit to another is governed by a set of rules. For instance to characterize, in a given period - say a year, an employee in terms of the type of contract or type of working time it is needed to decide which statistical rules to apply when an employee has had more than one position in the year.

The multiplicity of statistical units and these rules allows providing alternative representations of the same or related phenomena projected on populations of different units. For instance, in conjunction with the metadata, it permits to analyse

the relationships between the inequalities of price of labour, defined at the level of job position, to the inequality of earnings among individuals. Moreover, it allows switching from measures that make sense only for some units (the wages can be defined both at enterprises and at job level) to other that makes sense only for another unit (the income that is related to the individual level).

In the following, some figures are provided to assess the cardinality of the LR in terms of different statistical units. In the year 2015 (Table 1), the segment of the LR referred to the dependent employment in the private sector, is composed of 1.6 millions of employers, 14 millions of employees that generate 17.5 millions of work positions.

Table 1 – *Census of the dependent work relationships in the private sector extra-agriculture and different unit of analyses. Year 2015.*

Units	Measurements
Employers	1,573,529
Employees	14,004,534
Couples employers / employees	16,075,725
Work relationships	17,495,470
Work relationships with changes	18,752,967

From the individual point of view, among the 14.004.534 individuals in RACLI in the years 2014-2015, the 78,86% worked only for an enterprise, the 15,59% for two enterprises and the 3,93% for three. The different individuals have activated for the 72% of the cases only a work relationship, for the 16,92% two relationships and for the 1,45% more than five. Besides, the 72% entertained only a relationship with only an enterprise during the period, while the 4,79% entertained two relationships with the same enterprise. The 1,11% had three relationships with only an enterprise and the 2,64% had an unique relationships in three different enterprises (Table 2).

Different metrics of labour input are necessary to fulfil different statistical scopes. The number of work relationship formally active in specific dates, along with those activated and closed in an interval, provides the necessary data for the stock-flow accounting that is the number of days each contract has been formally active during a period (e.g. a year). A natural extension of these measures is the average number of formally active job positions in an interval where the average is calculated using as weights the number of days of formal duration in the interval.

Those measures are important also because they are the same used within the Cc system, with the important value added that the LR provides not only the magnitude of the gross and net job flows but also the level of the job stocks.

Table 2 – Percentage of individuals for number of employers and number of relationships (only dependent work relationships in the private extra-agriculture sectors). Years 2014 – 2015.

Number of employers	Number of work relationship						Total
	1	2	3	4	5	>5	
1	72.00	4.79	1.11	0.43	0.24	0.29	78.86
2		12.13	2.24	0.64	0.27	0.31	15.59
3			2.64	0.76	0.27	0.26	3.93
4				0.63	0.24	0.21	1.08
5					0.16	0.16	0.32
>5						0.22	0.22
Total	72.00	16.92	5.99	2.46	1.18	1.45	100,00

A set of measures on labour input: stock-flow accounting, hours paid and hours worked

Table 3 – Stock-flow accounting in the universe of the dependent jobs in the private sector. Year 2015

Concepts	Measures
Number of distinct job position active for at least one day in the period	17,495,470
Average number of effectively active Job positions (based on days covered by contributions)	11,381,696
Average number of formally active Job positions (based on days covered by contract)	12,018,315
Number of job position formally active in a determined date e.g. January 1 st	11,231,636
Number of jobs activated (activations) in an interval e.g. the first 6 months of the year	3,281,377
Number of jobs terminated (cessations) in an interval e.g. the first 6 months of the year	2,962,707

However, the formal duration of the job may provide wrong signals if one is interested in the effective work input especially for some types of contracts (e.g. job on call). In fact in traditional business statistics and the NA the measure used to calculate the average number of jobs is weighted with a quantity that expresses the effective activity (such as the number of days for which the job has been covered by social contributions). Compared with the former measure it reduces dramatically the weights of the contracts such as the jobs on call and, at the same time, corrects the importance of those cases, that might be affected by errors in the

dates of activations and cessations. Besides the measures of jobs stocks and flows, either pointwise or over intervals, the LR provides variables to evaluate the time paid by the employer and time worked (Table 4).

Table 4 – Hours paid and hours worked per-capita in the universe of the dependent jobs in the private sector. Year 2015- 2016

Size class	Economic activity	Hours paid		Hours worked	
		Year		Year	
		2016	2015	2016	2015
Total	Total	1778,82	1758,89	1396.87	1410.89
Total	T2-Ind	1943,70	1914,18	1541.78	1554.38
Total	T3-Ser	1685,37	1668,15	1312.21	1329.57
>100	Total	1834,92	1818,83	1405.94	1421.16
>100	T2-Ind	1997,57	1962,17	1549.23	1565.61
>100	T3-Ser	1754,38	1746,42	1333.56	1349.63
1-100	Total	1740,49	1717,63	1390.63	1403.88
1-100	T2-Ind	1911,85	1885,88	1547,74	1537,38
1-100	T3-Ser	1634,31	1609,33	1314,73	1296,17

The number of hours paid, used primarily as denominator of wages to capture the effective price of labour (e.g. the wage rates), reduces the weight of part-time workers and excludes from the computation the time paid by the social security for events of sicknesses, maternity, accidents, short time allowances, etc.. Moreover, the time for paid overtime is accounted. Finally, the number of hours worked, excluding the time spent on holidays and time in the above events paid by the employer, provides the foremost metric to measure the labour input, and as such, is used as denominator in important indicators (e.g. hourly labour cost, labour productivity).

3. New analyses of the labour market

Besides the renewal of the statistical processes and the role of the surveys, the availability of the LR allows new analyses on the labour market and on the impact of labour policies (ILO, 2017).

In particular, the LR immediately opens new possibilities for the study of wage inequality both alone and together with other sources either sample surveys or registers. Up to now the research on wage inequality in Italy has been based mostly on household sample surveys but the advantages of this new source over the classical sources is multifold. Since the register covers the whole population of employees, it overtakes the limits posed by the classical sample surveys. In fact, it allows an in depth study even of small subpopulations of workers or of firms (for

instance on the geographical or economic activity dimension) without incurring in sampling errors. Considering the territory, as an example, the register may shed light on the inequality between and within very small portion of territory like the Municipalities or Local system of labours. Moreover, the measure of wage is usually considered much more precise of whatever information can be obtained from the side of the worker that are likely to be affected by phenomena like proxy responses, memory biases, rounding biases, underreporting etc..

A similar advantage of the measure over the household surveys regards the labour input measure like the number of hours paid. Although, in the Register, it is subject to some estimation components to correct the definition of the information available from the administrative source, the measure is likely to be much more accurate than the survey one. A possible shortcoming of the wage figure, of course, is that, by definition, it does not include any irregular sum paid by the employer.

The following figure 3 shows the share of low paid job, at municipal level, identified as the dependent jobs of the private extra-agricultural sectors with an hourly earnings lower than $2/3$ of the national median value.

Figure 3 – *Low paid jobs incidence at municipal level of detail. Year 2016.*



Moreover, the work relationship which is the basic unit of the register, beyond relating two other important unit, the employer and the employees, is further connoted by the activation date of the contract and can be considered a stepping stone of more in depth analysis. In fact the wage of an employee, say in a year, can be reduced to the wages earned in each of the work relationships he has occupied in the year either with the same employer or with different employers, along with the duration in hours of each job. The inequality among workers, thus, can be tracked back to the wage differentials between job relationships. This can enrich remarkably the inequality analysis especially because in recent years the labour market has been increasingly characterized by a fragmentation of working relationships. A considerable number of short and very short jobs spells are enumerated and the annual wages of a sizeable fraction of workers is composed by what is earned in all these jobs, either with the same employer or with different employers. Together with length of time during the year of each job, the type of working time is showed to contribute a lot to wage inequality since approximately a quarter of employees are employed in part time jobs.

Another, very important, advantage of this register is its extensibility. Representing the universe of the worker it can be matched with any sample survey, either business or household, possessing the right identifiers, at micro level with the possibility of enriching the analysis with the specific information required by the survey or the subject analysis the other way round.

For example, observing the incidence of the couples with only a work relationship, for age and geographical area, the evidence is that it is less consistent for the young people (less than 29 years) and in the South and Islands (Figure 4).

Figure 4 – *Percentage incidence of the couples employers - employee with only one relationship for class of age and geographical division (Years 2014 - 2015)*



A further strength of the new data source is that, being an employer-employee database, it allows taking in account very accurately the characteristics of the firm/institutions. Moreover the information available on each firm are much richer

either because it is possible to draw information on related databases like the FRAME-SBS or because the same employee level information allows to calculate firm level indicators on the composition of the workforce. Therefore, it provides measures of inequality and their link to the economic structure and, in particular, the role of between versus within enterprises wage inequality and its correlation with indicators of firm performance and structure (Baldi et al., 2017).

4. Final remarks

Within the more general Integrated System of Register (SIR) that Istat is producing, the potentialities of the Labour Register makes it a powerful instrument for analyses on the characteristics of the work positions, both dependent and not dependent, at micro level.

Moreover its richness consists also in the possibility of changing the unit of analyses from the work position to the individual until the economic unit and with a longitudinal dimension which represents an important value added also for the economic policies evaluations on the labour market.

References

- BALDI C., COSTA S., PACINI S., ZURLO D. 2017. Looking at wage inequality through the lenses of economic structure. A first exploration of a new data source, Trends in inequality - International Conference - Bologna, 2-4 November 2017.
- ESS Vision 2020, Maggio, 2014 <http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020>.
- ILO 2017. Quick guide on sources and uses of labour. Geneva. Switzerland.
ISBN: 978-92-2-130119-6 statistics.
- ISTAT 2016. Istat's Modernisation Programme.
https://www.istat.it/en/files/2011/04/IstatsModernisationProgramme_EN.pdf
- RAPITI F., BALDI C., ICHIM D., PINTALDI F., PONTECORVO M.E., RIZZI R. 2018. Digging into labour market dynamics: toward a reconciliation of stock and flows short term indicators, SIS, Palermo, 19-22 Giugno 2018.

SUMMARY

The Labour Register in Italy: the new heart of the system of labour statistics

The ongoing construction of a Labour register is greatly enhancing the labour statistics either due to the possibility of producing output directly or by constituting a coordination framework able to provide a much greater alignment and coherence to the entire system for labour statistics, in the spirit of the ESS Vision 2020. Three pillars ensure this result. First, a set of clearly defined statistical units, the main one being: the employer, the worker and the core of the system, that is the work relationship. The data representations based on each unit feed respectively the business statistics (such as SBS), the social statistics (such as the census and LFS) and more specific labour market statistics. In this regard, the system allows producing an output of fluxes indicators (such as job creation and destruction) consistent with more usual stock statistics (number of jobs). Second, a set of classifications and coding rules able to map precisely the juridical/fiscal/contributive reality of the labour market into a statistical representation according to the international classifications, including the incoming versions of ILO ICSE. Third, a set of measures on labour input which in most cases are able to fulfill directly the statistical standards or can be used as intermediate input or auxiliary variables. This set of measures span from formal duration of the contract, to the number of hours actually worked, passing through measures of effective activity, remunerated activity and hours paid. The paper, beyond describing these measures and their relationships, illustrates their capacity of illuminating old and new labour market issues such as the fragmentation of the work activity.

Ciro BALDI, Italian Institute of Statistics, baldi@istat.it
Claudio CECCARELLI, Italian Institute of Statistics, cceccar@istat.it
Sara GIGANTE, Italian Institute of Statistics, gigante@istat.it
Silvia PACINI, Italian Institute of Statistics, pacini@istat.it
Francesca ROSSETTI, Italian Institute of Statistics, frosset@istat.it

IL CASO STUDIO DELLE ORE LAVORATE IN UN'OTTICA DI DESTAGIONALIZZAZIONE E REVISIONI

Ilaria Franco¹, Silvia Loriga e Andrea Spizzichino

1. Introduzione

Le ore lavorate rappresentano un indicatore di ampio interesse in quanto sono legate a temi cruciali e attuali sia in ambito economico sia sociale. Dal punto di vista economico possono essere considerate un indicatore dell'input di lavoro di un paese, visto anche ai fini dell'analisi della produttività del lavoro mentre in un'ottica sociale costituiscono un'informazione rilevante ai fini dell'analisi dei tempi di vita e della conciliazione tra lavoro e vita personale e familiare.

Negli ultimi anni, caratterizzati dalla ripresa economica e del mercato del lavoro, dopo la crisi iniziata nel 2008 e proseguita anche nel 2013, le ore lavorate dagli occupati ricoprono un nuovo interesse, legato al fatto che questa ripresa del mercato del lavoro mostra un ritmo di crescita degli occupati più forte di quello del monte ore lavorate, determinato dalla riduzione del numero di ore pro capite (Guzzi, Liasciandro 2017).

In questo contesto, l'indicatore delle ore lavorate è interessante anche in un'ottica di analisi di breve periodo, al fine di affiancare all'andamento del numero di persone occupate anche quello delle ore lavorate e valutare quindi in modo completo l'evoluzione dell'input di lavoro. A tal fine si è determinata la necessità di disporre di serie storiche aggiornate e tempestive sulle ore lavorate destagionalizzate al fine di analizzare correttamente le tendenze nel breve periodo.

Questo lavoro risulta innovativo sia per la produzione di serie destagionalizzate sulle ore effettivamente lavorate a partire dai dati della rilevazione sulle forze lavoro sia per il procedimento utilizzato. Indicatori sulle ore lavorate vengono tradizionalmente prodotti a partire da rilevazioni sulle imprese e nell'ambito della contabilità nazionale, le serie prodotte in questo lavoro mettono in più in risalto le relazioni tra le principali caratteristiche dell'attività lavorativa e il numero di ore effettivamente lavorate

Nell'esercizio sono state affrontate tutte le principali tematiche legate alla destagionalizzazione tra cui l'opportunità di inserire effetti di calendario, la scelta

¹ Il presente lavoro riporta alcune parti della tesi di laurea di Ilaria Franco, condotta nell'ambito del progetto EMOS che ha previsto un tirocinio presso l'Istat, nel quale l'autrice ha collaborato con Andrea Spizzichino (tutor del progetto) e Silvia Loriga.

tra tecniche dirette o indirette e la verifica dell'affidabilità dei risultati attraverso l'analisi delle revisioni.

Il lavoro si conclude con l'analisi dei risultati mettendo in evidenza come la quantità di ore effettivamente lavorate abbia risentito degli effetti della crisi economica iniziata nel 2008.

2. Il Quadro d'insieme

Nello scenario italiano le ore effettivamente lavorate sono entrate nel dibattito sull'analisi del mercato del lavoro a partire dal 2014, quando si sono cominciati a vedere i primi segni di ripresa dopo la crisi iniziata nel 2008 che ha colpito l'economia italiana e il mercato del lavoro. Il dibattito nasce dal fatto che questa ripresa riguarda solamente il numero di occupati, che è effettivamente ritornato ai livelli del 2008, ma non la quantità di ore lavorate, che risulta ancora molto lontana dai livelli raggiunti prima della crisi.

Tale situazione si è verificata come conseguenza della diffusione di forme di lavoro, come il part-time e i lavori occasionali, caratterizzate dal numero di occupati che cresce in misura più consistente rispetto al numero totale di ore lavorate.

Queste tendenze nel mondo del lavoro hanno accresciuto l'interesse per lo studio delle serie storiche delle ore effettivamente lavorate.

Come per tutte le principali dimensioni del mercato del lavoro, e non solo, gli indicatori fondamentali utilizzati per analizzare l'andamento nel tempo sono le variazioni tendenziali e congiunturali.

La prima prende in considerazione le differenze tra dati osservati nello stesso periodo di tempo a un anno di differenza, mentre la seconda considera il mutamento del fenomeno rispetto a misurazioni ottenute in periodi consecutivi.

La differenza tra questi due tipi di analisi risiede nel fatto che la variazione tendenziale può essere calcolata anche su dati "grezzi", o non trattati, quella congiunturale necessita invece che i dati vengano prima depurati dalle fluttuazioni dovute al fatto che il fenomeno di interesse è caratterizzato da stagionalità. Tale componente, che influenza un qualsiasi fenomeno nel tempo, è considerata un fattore di disturbo poiché provoca un oscuramento del segnale congiunturale che complica l'analisi e l'interpretazione delle variazioni.

Alla base di questo lavoro, c'è proprio la volontà di affiancare analisi congiunturali a quelle tendenziali già disponibili, riguardanti il numero di ore lavorate.

Nei paragrafi successivi vengono dati dei cenni metodologici sui principali aspetti del processo di destagionalizzazione, sia con riguardo alle fasi preliminari sia a quelle di valutazione dei risultati; vengono poi presentati i principali risultati

ottenuti sottolineando le motivazioni che hanno spinto a prendere determinate scelte; il lavoro si chiude con la valutazione in un'ottica di diffusione futura.

Come accennato nell'introduzione, i dati su cui si basa questo lavoro sono quelli raccolti tramite l'indagine sulle forze lavoro tra il primo trimestre del 2004 e il quarto trimestre del 2017 e i principali aggregati da destagionalizzare sono gli occupati, il monte ore lavorate e le ore pro capite effettivamente lavorate.

3. La Metodologia

Il processo di depurazione di una serie dalla stagionalità prende appunto il nome di destagionalizzazione; come noto un fenomeno dinamico può essere decomposto in tre componenti principali che ne spiegano la variabilità nel tempo: il ciclo-trend, la stagionalità e la componente irregolare (BEE DAGUM 2001).

La componente di ciclo-trend (CT) rappresenta la tendenza di medio-lungo periodo e quindi l'andamento tendenziale di fondo, crescente o decrescente, di un fenomeno.

La stagionalità (S) viene definita come la variazione periodica di una serie storica; si presenta a intervalli regolari ogni anno, con intensità simile nello stesso periodo di anni differenti (mese, trimestre) e diversa nei differenti periodi dello stesso anno.

Infine, la componente irregolare (I), rappresenta la parte residuale ottenuta eliminando le componenti precedenti ed è causata dai movimenti erratici, ossia fluttuazioni di breve periodo non sistematiche e non prevedibili.

La procedura di destagionalizzazione è composta da due fasi: preliminare o di pre-trattamento e individuazione e rimozione della componente stagionale.

La fase preliminare consiste in primo luogo nella scelta dello schema di decomposizione della serie storica. Le tre componenti che caratterizzano una serie possono essere aggregate secondo diversi modelli: additivo, moltiplicativo, logaritmico e misto. In particolare i modelli principalmente usati risultano quello additivo e moltiplicativo.

Il modello additivo assume che la serie storica equivalga alla somma delle componenti di ciclo-trend, stagionale e irregolare quindi può essere definito come segue:

$$Y_t = CT_t + S_t + I_t \quad (1)$$

L'ipotesi alla base di questo modello è che le tre componenti siano tra loro indipendenti, ovvero l'ampiezza delle oscillazioni della componente stagionale non dipende dalla dinamica del ciclo-trend.

Nel modello moltiplicativo la serie storica considerata è ottenibile tramite il prodotto delle tre componenti:

$$Y_t = CT_t \times S_t \times I_t \quad (2)$$

Questo tipo di scomposizione viene applicata quando le componenti non sono tra loro indipendenti e quindi l'ampiezza delle oscillazioni dovute alla componente stagionale è proporzionale al livello della serie.

Una volta scelto lo schema di decomposizione, la fase preliminare prosegue con la depurazione dei dati che saranno poi destagionalizzati. Questa fase è necessaria poiché solitamente i dati derivano da rilevazioni campionarie quindi, soprattutto nel caso in cui le rilevazioni abbiano una lunga durata nel tempo, risulta complicato il controllo dell'effetto dei fattori esterni. Una distinzione necessaria per trattare i fenomeni esterni, riguarda la conoscenza a priori o meno dell'istante in cui si esplicano. Per questo motivo sono presi in considerazione due diversi tipi di fattori esterni che possono provocare delle modificazioni della serie: effetti di calendario e valori anomali (outliers) (Battaglia, Bonacci, Calzini 1996).

Gli effetti di calendario sono considerati modificazioni con origine e cause note che, a differenza degli effetti stagionali, non si presentano sempre negli stessi periodi dell'anno, ma comportano comunque un cambiamento dell'andamento di una serie nei diversi giorni della settimana. I valori anomali sono dei valori numericamente distanti dal resto dei dati raccolti, dovuti principalmente a due diverse cause: errori generati durante le diverse fasi dell'indagine tra cui si distinguono errori campionari e non campionari, oppure fattori esterni di natura eccezionale come calamità naturali, cambiamenti legislativi, scioperi e fenomeni di natura economica.

Il loro trattamento risulta molto più complesso in quanto oltre a stimare l'entità della variazione è necessario spesso anche determinare l'istante in cui si verificano. Al fine di stimare gli effetti di entrambe le modificazioni, si usa un modello di regressione ARIMA (reg-Arima). Questi modelli quindi migliorano le serie eliminando tutti i valori che influiscono negativamente sull'identificazione della componente stagionale, così che la successiva destagionalizzazione darà luogo a risultati di qualità più elevata.

Il modello Reg-Arima può essere descritto come una regressione lineare in cui la componente dell'errore segue un processo Arima e non un White Noise:

$$Y_t = f(\beta, w, t) + Z_t \quad (3)$$

e in cui $f(\beta, w, t)$ rappresenta gli effetti deterministici al tempo t (valori anomali e outliers) e Z_t rappresenta la componente stocastica.

Al termine della fase preliminare il risultato ottenuto è una nuova serie Z_t , derivante dalla depurazione della serie iniziale Y_t , sulla quale viene stimata la componente stagionale. A questo fine ci sono due approcci maggiormente diffusi nell'ambito della statistica ufficiale:

- Metodi Model Based che si basano sull'ipotesi che la struttura probabilistica del processo stocastico che genera la serie storica considerata sia adeguatamente spiegata da un modello statistico parametrico di tipo Arima.
- Metodi Filter based che, al contrario, stimano le componenti della serie storica considerata tramite l'applicazione successiva di filtri lineari costituiti da medie mobili centrate. Rappresentano quindi dei metodi non parametrici o semi-parametrici.

Oltre alla fase relativa alla disaggregazione delle serie nelle tre componenti, il processo di destagionalizzazione prevede una valutazione complessiva delle serie da destagionalizzare a monte e, dopo la destagionalizzazione, una fase di valutazione dei risultati ottenuti.

Alla base del processo di destagionalizzazione di serie storiche derivanti dall'aggregazione additiva o moltiplicativa di altre serie temporali, la pianificazione del procedimento non può prescindere dalla scelta tra un approccio diretto o indiretto alla destagionalizzazione.

La differenza principale tra i due approcci riguarda l'ordine con cui le serie vengono aggregate e destagionalizzate. La procedura diretta prevede che le serie siano in primo luogo aggregate e poi destagionalizzate; al contrario quella indiretta prevede che venga effettuata prima la destagionalizzazione delle serie disaggregate e solo dopo la loro aggregazione per ottenere serie aggregate. Generalmente il principio adottato è quello di usare la destagionalizzazione diretta nel caso in cui le serie componenti presentino caratteristiche simili mentre l'indiretta è preferibile nei casi in cui le serie siano caratterizzate da un andamento stagionale piuttosto differente o quando la componente stagionale sia molto debole e quindi difficile da identificare.

Come già accennato, un aspetto fondamentale legato alla destagionalizzazione, è la valutazione della qualità dei risultati ottenuti. In particolare in questo lavoro è stato approfondito un aspetto imprescindibile della qualità delle serie destagionalizzate rappresentato dall'analisi delle revisioni, ovvero delle differenze scaturite dal confronto tra serie destagionalizzate prodotte nelle release² successive (Eurostat 2013).

² Per release si intende la produzione di uno stesso dato rivista a seguito dell'aggiunta di un nuovo dato o di un evento che ne richieda un ricalcolo. Tipicamente la replicazione della procedura di destagionalizzazione, dovuta della disponibilità di un nuovo dato, genera revisioni nell'intera serie. Va sottolineato che l'entità delle revisioni è influenzato dalla procedura di destagionalizzazione scelta tra *concurrent approach* e *partial concurrent approach*.

Lo scopo dell'analisi delle revisioni è quello di capire se e quanto le stime iniziali di una serie destagionalizzata siano affidabili, ossia misurare la vicinanza tra la stima iniziale di un dato e le sue stime successive. Per valutare l'affidabilità delle stime, vengono usati degli indicatori di sintesi grazie ai quali si può capire se le revisioni sono o meno significative. Gli indicatori principalmente usati sono:

Revisione Media Assoluta (RMA)

$$RMA = \frac{1}{n} \sum_{t=1}^n |L_t - E_t| = \frac{1}{n} \sum_{t=1}^n |R_t| \quad (4)$$

Revisione Media (RM)

$$RM = \frac{1}{n} \sum_{t=1}^n (L_t - E_t) = \frac{1}{n} \sum_{t=1}^n R_t \quad (5)$$

dove n è il numero di stime per il periodo considerato, L_t la stima più recente, E_t la stima meno recente e R_t la loro differenza.

Per poter calcolare gli indicatori sintetici di revisione è necessario organizzare i dati in modo tale da creare, per ogni serie, un revision database. Quest'ultimo è composto da diversi fogli di calcolo che contengono due diverse forme di organizzazione dei dati: il triangolo di revisione e il prospetto di calcolo.

Il triangolo di revisione è un tipo di organizzazione dei dati tale che a seconda del tipo di analisi che si vuole condurre, le stime possono essere lette in senso orizzontale, verticale e diagonale.

Sulle righe del triangolo sono presenti le serie storiche o vintage rilasciate nel tempo, quindi la lettura orizzontale consente di studiare come si evolve il fenomeno di interesse. La lettura verticale permette invece di vedere come cambiano le stime pubblicate in riferimento a un singolo periodo temporale in modo tale da ottenere una misura delle revisioni e dunque un'analisi dell'affidabilità delle diverse pubblicazioni.

Infine, la lettura diagonale permette di individuare tutte le stime caratterizzate dalla stessa epoca di rilascio; per esempio le stime presenti nella diagonale principale più esterna, formano la serie costituita dalle prime stime rilasciate per ciascun periodo temporale.

Nel revision database sono contenuti tre triangoli di revisione: il primo contiene i livelli o indici della serie, il secondo le variazioni congiunturali delle stime e il terzo quelle tendenziali. I prospetti di calcolo contengono un riassunto di tutte le informazioni più importanti ai fini dell'analisi delle variazioni contenute nei fogli precedenti.

4. I principali risultati

Tutte le analisi prodotte in questo lavoro sono basate sulle stime prodotte dall'indagine sulle forze lavoro tra il primo trimestre del 2004 e il quarto trimestre del 2017. Lo studio si focalizza su tre aggregati: occupati, totale monte ore lavorate e ore pro capite effettivamente lavorate (con il terzo che si ottiene dal secondo fratto il primo).

I tre aggregati sono stati disaggregati rispetto alle principali variabili d'interesse: sesso, posizione lavorativa (dipendente o indipendente) e la tipologia di orario di lavoro (pieno o parziale), così da ottenere ventiquattro serie.

La scelta delle variabili di disaggregazione è ricaduta su quelle che influenzano maggiormente le ore effettivamente lavorate (Istat 2011) tenendo presente la necessità di avere serie di numerosità non troppo ridotta sulle quali non sarebbe possibile effettuare una destagionalizzazione.

La destagionalizzazione delle serie a nostra disposizione viene effettuata automaticamente tramite il software Demetra, un'interfaccia che consente di applicare la procedura scelta a livello europeo TRAMO- SEATS (Gomez e Maravall, 1996), in cui la fase di pretrattamento è svolta da TRAMO (Time series Regression with Arima noise, Missing observation and Outliers) mentre quella di destagionalizzazione da SEATS (Signal Extractio in Arima Time Series).

Inizialmente tutte le serie considerate sono state destagionalizzate prendendo in considerazione il periodo compreso tra il primo trimestre 2004 e il primo trimestre del 2015, per poi aggiungere uno alla volta i dati trimestrali successivi fino a arrivare all'ultimo dato disponibile relativo al quarto trimestre 2017.

Tutto ciò con l'obiettivo di studiare dopo quanto tempo la serie si stabilizza e di fornire un'informazione più accurata possibile che consenta di analizzare le revisioni che si producono con l'aggiunta del dato di ogni trimestre.

Per quanto riguarda le otto serie degli occupati fino al primo trimestre del 2015, l'identificazione dei modelli di destagionalizzazione svolta automaticamente dal software ha prodotto un ottimo risultato; solo le serie delle donne con un'occupazione dipendente a tempo pieno e quelle delle donne con un'occupazione indipendente sia a tempo pieno sia part-time hanno richiesto una identificazione manuale, a causa della presenza di outliers additivi di diverse entità.

Usando i modelli identificati per le serie fino al primo trimestre 2015, si è proceduto aggiungendo a ciascuna serie le osservazioni dei trimestri successivi fino ad arrivare al quarto trimestre del 2017.

Per la maggior parte delle serie considerate i modelli sono rimasti gli stessi fino all'aggiunta dell'ultima osservazione; le uniche eccezioni riguardano quelli delle serie degli occupati maschi indipendenti part-time e delle occupate femmine

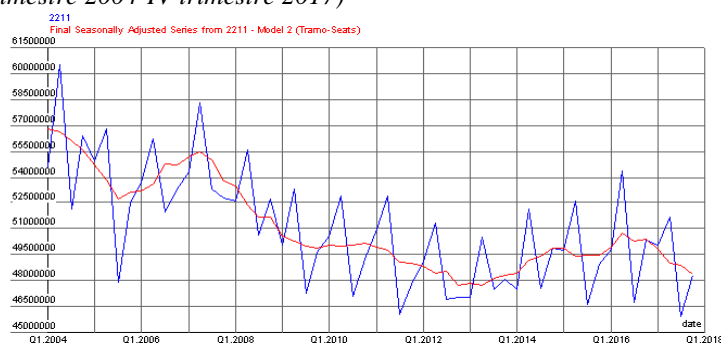
indipendenti a tempo pieno che nel secondo trimestre del 2016 sono stati identificati manualmente per la presenza di outliers additivi e transitory change³.

L'identificazione automatica dei modelli di destagionalizzazione delle otto serie del monte ore lavorate, che corrisponde all'ammontare complessivo delle ore effettivamente lavorate calcolato come la somma delle ore lavorate da tutti gli occupati, ha prodotto un risultato meno soddisfacente del precedente in quanto solo tre modelli identificati automaticamente sono stati ritenuti adeguati.

Le serie del monte ore svolto dagli uomini e dalle donne dipendenti part-time non presentavano grossi problemi ma è stato comunque scelto di identificare manualmente i modelli poichè, essendo la prima destagionalizzazione, si è cercato di ottenere i risultati migliori per evitare di incontrare problemi nelle release successive. Molto più problematiche risultavano la serie del monte ore svolte sia dagli uomini sia dalle donne indipendenti part-time e dalle donne indipendenti a tempo pieno, a causa della presenza di numerosi outliers additivi. I modelli identificati manualmente in questa prima fase di destagionalizzazione sono comunque rimasti gli stessi per tutti i trimestri successivi fino al quarto trimestre del 2017.

L'unico modello che è stato nuovamente identificato manualmente è quello della serie del monte ore svolte dalle donne con un impiego indipendente e part-time; una volta aggiunto il dato del secondo trimestre del 2016 infatti, questo modello è stato rigettato dal software in quanto non rispettava le condizioni poste sul test di normalità.

Figura 1 – Serie del monte ore lavorate dalle donne, indipendenti a tempo pieno(I trimestre 2004-IV trimestre 2017)



³ Generalmente non è consigliabile cambiare i modelli delle serie, soprattutto tra trimestri di uno stesso anno, in quanto si rischia di generare revisioni consistenti rispetto ai dati precedenti. Per queste ragioni prima di decidere di modificare i modelli è necessario accertarsi, facendo la differenza con i valori ottenuti nel trimestre precedente, che i cambiamenti apportati nelle serie non siano troppo consistenti.

Nella figura 1 è riportato a titolo di esempio l'output ottenuto tramite il software Demetra relativo alla serie del monte ore lavorate dalle donne indipendenti a tempo pieno. Nel grafico sono riportate la serie grezza originaria (spezzata) e quella destagionalizzata (continua). Già l'analisi grafica consente di valutare l'ottima qualità della destagionalizzazione della singola serie.

Un discorso particolare è necessario per le otto serie delle ore lavorate pro capite. L'indicatore sulle ore effettivamente lavorate non è ancora mai stato prodotto in forma destagionalizzata a partire da rilevazioni svolte sulle famiglie.

Questa variabile è il risultato del rapporto tra i due aggregati finora analizzati, quindi la destagionalizzazione può essere effettuata sia seguendo un approccio diretto sia indiretto. La destagionalizzazione diretta di questo gruppo di serie fino al primo trimestre del 2015, non ha prodotto dei buoni risultati in quanto per molti modelli c'è stato bisogno di interventi manuali. Queste correzioni sono dovute a ragioni differenti. In certi casi sono state necessarie per la presenza di outliers molto evidenti, in altri sono state apportate con lo scopo di migliorare la regolarità dell'andamento della serie destagionalizzata che seppur accettata dal software non risultava soddisfacente.

I risultati ottenuti con la destagionalizzazione diretta delle serie fino al primo trimestre del 2015, sono stati confrontati con quelli ottenuti nello stesso periodo con l'approccio indiretto, in modo da stabilire il miglior modo di procedere.

La scelta è ricaduta sulla tecnica indiretta, tenendo conto di due considerazioni principali: la prima riguarda il confronto tra i risultati ottenuti in modo diretto e indiretto, che sono risultati molto vicini tra di loro; la seconda si basa sul fatto che la ri-destagionalizzazione fatta sulle serie già destagionalizzate con l'obiettivo di verificare la presenza di stagionalità residua⁴ ha prodotto, per tutte le serie considerate, lo stesso risultato della destagionalizzazione indiretta, a prova del fatto che quest'ultima non necessita di ulteriori correzioni.

Tutte queste motivazioni hanno portato a utilizzare l'approccio indiretto e a limitare quindi le serie da destagionalizzare a quelle di occupati e monte ore lavorate (16 nel complesso) con cui ottenere per rapporto le ore effettivamente lavorate.

Al fine di verificare l'affidabilità delle stime ottenute si è proceduto con l'analisi delle revisioni dei risultati. A questo scopo i tre aggregati principali sono stati disaggregati nel seguente modo: totali, maschi, femmine, dipendenti, indipendenti, tempo pieno e part-time.

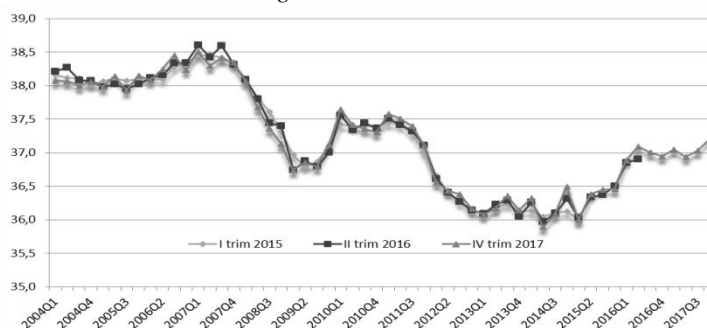
Mediante i triangoli di revisione è stata condotta un'analisi orizzontale, dalla quale emerge, nel periodo compreso tra il primo trimestre del 2015 e il quarto trimestre del 2017, per la maggior parte degli aggregati, una crescita caratterizzata

⁴ Per stagionalità residua si intende la manifestazione di stagionalità in dati già destagionalizzati.

da variazioni contenute. Partendo da queste informazioni sono state calcolate le revisioni tra le variazioni tendenziali e congiunturali di ogni indicatore a distanza di un trimestre e di un anno dalla prima uscita.

Tali revisioni sono state sintetizzate attraverso gli indicatori presentati nel paragrafo 3 (revisione media assoluta (RMA) e revisione media (RM)) che hanno evidenziato revisioni piuttosto contenute a dimostrazione della qualità delle destagionalizzazioni prodotte.

Figura 2 – Ore effettivamente lavorate dai maschi nelle release del I trimestre 2015, II 2016 e IV 2017. Dati destagionalizzati.



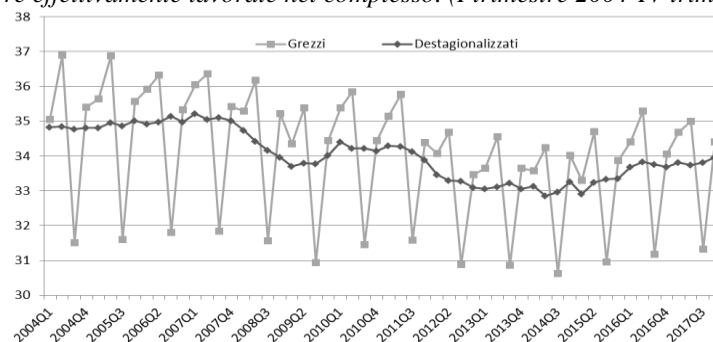
Anche rispetto alle revisioni il ricorso all'analisi grafica risulta utile: la figura 2 riporta l'esempio della serie storica destagionalizzata delle ore effettivamente lavorate dai maschi, stimate fino al I trimestre 2015, fino al II 2016 e fino al IV 2017. Il grafico mostra come cambia la serie destagionalizzata in seguito all'aggiunta dei nuovi trimestri, ed emerge chiaramente quanto i risultati ottenuti nella prima release vengano sostanzialmente confermati nelle successive.

Per concludere, nella figura 3 viene illustrato sia l'andamento dell'indicatore grezzo delle ore pro capite effettivamente lavorate dagli occupati sia quello destagionalizzato. La figura mostra come la serie destagionalizzata segua gli andamenti di quella grezza evidenziando il trend che ha caratterizzato le ore effettivamente lavorate tra il 2004 e il 2017 garantendo anche la possibilità di confronti tra trimestri diversi dello stesso anno o di anni differenti.

E' possibile notare come le ore effettivamente lavorate abbiano risentito degli effetti della crisi economica che ha investito il paese a partire dal 2007. Fino a questa data le ore lavorate erano caratterizzate da un andamento crescente per poi cambiare di tendenza e raggiungere dei livelli molto bassi nel 2009. Nel 2010 si registra una breve ripresa ma già dalla prima metà del 2011 i livelli ricominciano a diminuire fino al II trimestre del 2014 in cui si raggiungono i minimi storici. A partire dalla metà del 2015 inizia una nuova ripresa piuttosto consistente fino alla

metà del 2016. Ancora oggi i livelli di ore lavorate pro capite risultano comunque di oltre un'ora inferiori rispetto al periodo pre-crisi.

Figura 3 – Ore effettivamente lavorate nel complesso. (I trimestre 2004-IV trimestre 2017)



Si sottolinea infine, che quest'analisi, riportata a livello aggregato possa essere effettuata anche rispetto al sesso, la posizione lavorativa e la tipologia di orario di lavoro per vedere l'effetto composizione che determina il risultato nel complesso ed evidenziare come sia cambiata nel tempo la dinamica per i diversi aggregati.

5. Conclusioni

Il lavoro qui presentato fornisce i primi risultati di una importante innovazione che è stata avviata nell'ambito dell'analisi dei dati sul mercato del lavoro derivanti dalla rilevazione campionaria sulle forze di lavoro.

A partire dalle stime prodotte dalla rilevazione, per la prima volta, è stata prodotta la destagionalizzazione delle ore effettivamente lavorate dagli occupati consentendo di analizzare gli andamenti di questo indicatore che sempre di più, negli ultimi anni, risulta al centro delle attenzioni degli analisti.

Nel processo di produzione delle serie destagionalizzate sono state affrontate tutte le principali problematiche che si incontrano quando si destagionalizzano serie storiche.

In particolare oltre alla definizione dei modelli relativi alle diverse serie, l'attenzione è stata rivolta allo studio degli effetti di calendario (mai inseriti nelle precedenti destagionalizzazioni su dati di forze lavoro), alla scelta di un approccio diretto o indiretto e all'analisi della qualità dei risultati attraverso lo studio delle revisioni. Tutte le strategie adottate sono conformi agli standard di destagionalizzazione definiti in Istat e conformi alle indicazioni a livello europeo.

I risultati ottenuti dal processo di destagionalizzazione e l'ottima qualità degli indicatori di revisione, sia a livello aggregato sia disaggregato, rispetto alle principali variabili d'interesse, consentono di ipotizzare per il futuro la produzione

regolare e pubblicazione di indicatori destagionalizzati sulle ore lavorate di fonte forze lavoro da affiancare agli indicatori già prodotti dall'indagine stessa.

Riferimenti bibliografici

- BATTAGLIA F., BONACCI A., CALZINI C. 1996. Individuazione di dati anomali in serie storiche mediante interpolatori e previsori. *Statistica Applicata*, pp.8:577–594.
- BEE DAGUM E. 2001. Analisi delle serie storiche: modellistica, previsione e scomposizione. *Springer Science & Business Media*.
- EUROSTAT. 2013. ESS guidelines on revision policy for PEEIs. <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-13-016>
- GOMEZ V., MARAVALL A. (1996). Programs TRAMO (Time Series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instruction for the User. Working Paper 9628, Research Department, Bank of Spain.
- GUZZI G., LIASCIANDRO M. 2017. Gli occupati aumentano, le ore lavorate no. <http://www.lavoce.info/archives/49734/gli-occupati-risalgono-le-ore-lavorate-no/>
- ISTAT.2011. I tempi del lavoro. *Collana Argomenti n°40*, Istat

SUMMARY

Hours worked from the perspective of seasonal adjustment and revision

The interest in hours worked, which correspond to the number of hours actually worked by each employed person during the reference week, is related to the recent debate related to the recovery after the 2008 crisis that only concerns the level of employment, not the hours worked.

This work is focused on seasonal adjustment of hours worked time series, based on data collected by the Labour Force Survey during the period between the first quarter of 2004 to the fourth of 2017, which corresponds to the last data published.

A revision analysis was performed in order to verify the reliability of results of seasonal adjustment process. Revisions refer to changes of previously published data due to additional information available.

The results of the seasonal adjustment process and the high quality of the revision indicators make possible to assume a future publication of seasonal adjusted indicators of hours worked to be added to all the other indicators already published based on the Labour Force Survey.

Ilaria FRANCO, Università LA SAPIENZA, illyf93@gmail.com

Silvia LORIGA, Istat, siloriga@istat.it

Andrea SPIZZICHINO, Istat, spizzich@istat.it

ESSERE NEET IN ITALIA: I PRINCIPALI FATTORI DI RISCHIO

Irene Brunetti, Valentina Ferri

1. Introduzione

A partire dalla fine degli anni '90 nel Regno Unito si inizia a porre l'attenzione su quei giovani, con un'età compresa fra i 16 e i 18 anni, non inseriti in nessun percorso di istruzione o formazione né occupati, e quindi come tali considerati a rischio di esclusione sociale, definiti *Not in Education, Employment and Training* - NEET (SEU,1999).

La letteratura economica internazionale mostra da alcuni anni, in particolare dal 2008, molta attenzione all'indicatore del tasso di NEET (*Neet rate*) perché valori comparativamente più alti sono il segnale di difficoltà e ritardi nella transizione dallo studio al lavoro, di scarsa capacità del mercato di includere i giovani e di fenomeni di *skill mismatch* e scoraggiamento. La prolungata assenza dei Neet dal mercato del lavoro rischia, inoltre, di rendere più difficoltoso il loro inserimento o reinserimento aumentando quindi il rischio di disoccupazione di lungo periodo, salari più bassi, esclusione sociale e la dipendenza da ammortizzatori sociali (Balan, 2015).

Lo status di NEET a lungo termine produce effetti negativi non solo a livello individuale ma anche in termini di costi che un paese deve sostenere. Sebbene il calcolo del costo economico dei NEET sia un esercizio molto complesso, i ricercatori di Eurofound hanno stimato un costo molto elevato (Eurofound, 2012). Oltre alla perdita in termini di mancati guadagni, potenziale produzione e di produttività economica (SALTO-YOUTH Centro risorse per l'inclusione, 2015), Eurofound (2012) ha incluso nella stima anche l'impatto sulle finanze pubbliche in termini di regimi di welfare, nonché spese supplementari per la salute ed il benessere. Come evidenziato da Quintano *et al.* (2018), essere NEET significa essere improduttivi e perdere l'opportunità di migliorare il proprio capitale umano.

Queste considerazioni giustificano quindi il notevole interesse mostrato sia in ambito accademico che politico e sociale. La Commissione Europea, nell'ambito dell'iniziativa *Youth on the Move* finalizzata a sostenere la nuova strategia per l'occupazione e lo sviluppo "Europa 2020", ha deciso un monitoraggio continuo

della situazione dei Neet, sulla base di dati comparabili a livello UE, che servirà all'elaborazione di politiche comuni e all'apprendimento reciproco in materia.

Molteplici sono infatti le cause alla base di questo fenomeno e solo in parte sono dovute alla recente crisi economica. L'obiettivo di questo paper è indagare le determinanti dello stato di NEET in Italia utilizzando un campione di osservazioni proveniente dalla Rilevazione Continua sulle Forze Lavoro (RCFL) per gli anni 2007 e 2017. In particolare, si stima la probabilità di diventare NEET controllando per una serie di caratteristiche individuali tra cui il sesso, la cittadinanza, la tipologia di scuola superiore in cui il soggetto si è diplomato ecc. La scelta degli anni considerati per l'analisi deriva dalla necessità di verificare se ci sia stato un cambiamento o meno nei fattori di rischio prima e dopo la crisi economica.

L'identificazione delle determinanti dello status di Neet può in qualche misura offrire suggerimenti ai *policy maker* relativamente alle azioni da intraprendere per contrastare l'aumento, e di conseguenza il rischio di esclusione sociale, dei Neet. Data l'importanza del rischio di disoccupazione giovanile o della disoccupazione stessa, con la nostra analisi diamo delle informazioni su come mitigare il problema.

Il paper è strutturato come segue: nella Sezione 2 passiamo brevemente in rassegna la recente letteratura riguardante il fenomeno dei Neet, la Sezione 3 descrive i dati e presenta alcune statistiche descrittive. La Sezione 4 presenta il modello stimato ed i principali risultati. Infine, la Sezione 5 conclude.

2. NEET: una *review* della letteratura

La definizione di NEET non è esattamente la stessa per tutti i paesi poiché gli istituti statistici internazionali utilizzano indicatori differenti in base all'età, alle condizioni nel mercato del lavoro e alla natura dei corsi di istruzione e formazione (Cicciomessere *et al.*, 2011). Nel Regno Unito il termine NEET è stato creato per affrontare un problema specifico, quello della difficile transizione tra scuola e lavoro (Robson, 2008). Dato che stiamo parlando dell'Italia in questo lavoro ci rifaremo alla definizione data da Istat e Banca d'Italia¹: "... [I NEET sono] la popolazione di età compresa tra 15 e 29 anni che non sono né occupati né in istruzione o formazione, il che significa qualsiasi tipo di istruzione scolastica / universitaria e qualsiasi tipo di attività di formazione..." (Istat, 2013).

Da un punto di vista puramente teorico, includere i giovani disoccupati in questa categoria non è corretto perché stanno attivamente cercando un lavoro. A

¹ La definizione data da Istat e Banca d'Italia si discosta dalla definizione data a livello europeo unicamente per la fascia d'età considerata. Eurostat definisce NEET i giovani di età compresa tra i 15 e i 24 anni che non sono né occupati, né in istruzione o formazione.

causa di errori di misurazione dovuti a risposte errate da parte degli intervistati, che a volte si auto definiscono disoccupati anziché inattivi, è difficile distinguere chi è effettivamente disoccupato da chi è invece inattivo (Robson, 2008; Marshall, 2012). Boeri e Garibaldi (2002) hanno stimato che in Italia circa il 45% di quelli classificati come disoccupati e il 10% di quelli classificati come inattivi lavorano in modo irregolare. Queste percentuali potrebbero essere ancora più elevate per i giovani dal momento che hanno meno potere contrattuale sul mercato del lavoro.

Il notevole aumento dei NEET avvenuto durante la crisi economica è certamente collegato alla crescita dei tassi di disoccupazione giovanile. Anzi, da una parte, in tempi di crisi, i giovani sono stati i soggetti più penalizzati rispetto ai lavoratori più anziani a causa della loro minor esperienza lavorativa in aggiunta a contratti di lavoro più deboli e più basse qualifiche (Marelli e Signorelli, 2015). Oltre a questi elementi, altri fattori possono essere considerati come determinanti dell'elevato numero di NEET. Le imposte sul lavoro, l'indennità di disoccupazione, i salari minimi, e politiche del mercato del lavoro a proteggere altre tipologie di persone dalle crisi, hanno di fatto contribuito all'aumento del tasso di disoccupazione giovanile. D'altro canto, lo status di NEET dipende anche da fattori non legati alla disoccupazione giovanile, come l'esclusione dall'istruzione e dalla formazione (Bell e Blanchflower, 2011). Nel 2016 in Italia sono circa 2,2 milioni i giovani tra i 15 e 29 anni che non sono in occupazione né in formazione, la quota più elevata tra i paesi dell'Unione Europea (24.3% vs valore medio del 14%) (Istat, 2017). Tuttavia, dopo il forte incremento registrato negli anni della crisi, il fenomeno sembra essere in lieve riduzione.

Altre ragioni che determinano lo scoraggiamento dei giovani possono essere ricondotte al ciclo economico, alla mancanza di conoscenza su come o dove cercare lavoro, all'incapacità di trovare un lavoro che corrisponda ai loro livelli di abilità o all'indisponibilità di allontanarsi quando non ci sono posti di lavoro nella loro area di residenza (ILO, 2015). Coppola e Di Laurea (2016) spiegano che la condizione di NEET potrebbe dipendere da un sistema di *welfare state* efficiente. In particolare, la percentuale di Neet potrebbe essere ridotta dalla disponibilità di specifici servizi di assistenza per malati e disabili che andrebbe ad aumentare la propensione al lavoro di chi dichiara di essere inattivo perché è malato o deve assistere un familiare.

L'essere NEET ha quindi cause molteplici e intrecciate. Questo stato è anche influenzato da fattori socio-economici individuali e familiari. Con riferimento ai fattori individuali, Ryan (2001) ha analizzato l'intero periodo tra la fine della scuola dell'obbligo e il raggiungimento di un'occupazione stabile a tempo pieno in alcuni paesi europei, in Giappone e negli Stati Uniti. Dalla sua analisi emerge che lo svantaggio socioeconomico e un basso livello di istruzione sono le principali forze trainanti per l'inattività della gioventù e la disoccupazione in tutti i paesi. Brunello

e De Paola (2014) e Furlong (2006) hanno evidenziato che lasciare presto la scuola è spesso associato allo stato di NEET.

Tuttavia, come messo in luce da Caroleo *et al.* (2018), anche tra i più istruiti, lo squilibrio tra le competenze acquisite e quelle richieste dal mercato del lavoro rappresenta uno dei fattori principali di rischio della condizione di NEET.

Infine, per quanto riguarda i fattori legati alla famiglia, una bassa estrazione sociale, un basso reddito familiare e bassi livelli di istruzione dei genitori sono fattori legati allo stato di Neet (Bynner e Parsons, 2002; Robson, 2008; Eurofound, 2012a).

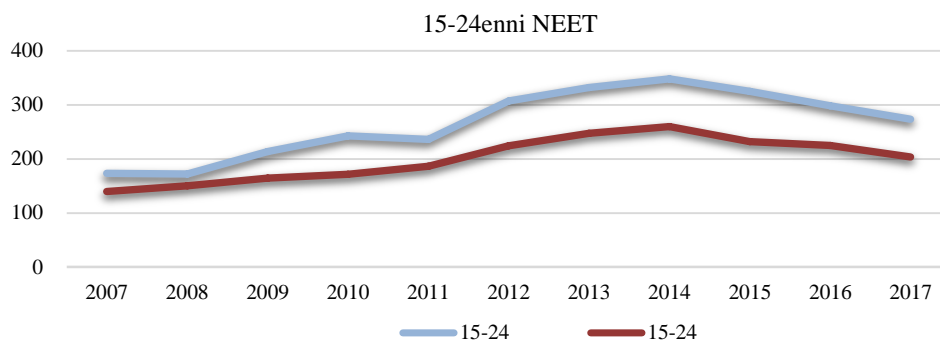
3. Dati e statistiche descrittive

L'indagine empirica si basa sui dati della *Rilevazione Continua sulle Forze di Lavoro* relativa agli anni 2007 e 2017 sviluppata dall'Istituto Nazionale di Statistica (ISTAT). Dall'indagine sulle forze di lavoro derivano le stime ufficiali degli individui occupati e in cerca di lavoro, nonché informazioni su aggregati dell'offerta di lavoro, professione, ramo di attività economica, ore lavorate, tipologia e durata dei contratti, formazione.

In questo paragrafo analizziamo il quadro nazionale del fenomeno dei NEET in Italia. Nonostante il punto più elevato sia stato nel 2014, sia per i 15-24enni che per i 25-34enni, l'aumento dei NEET nelle due unità di tempo considerate, 2007 e 2017, è molto rilevante. I NEET di sesso maschile nel 2017 aumentano del 91% rispetto al 2007 per ambo le fasce (Figura 1 e Figura 2). Per le donne invece la situazione risulta più complicata, pur attestandosi ad un livello inferiore tra i 15-24enni rispetto agli uomini, per la stessa fascia d'età, le donne NEET aumentano rispetto alla prima unità di tempo considerata, del 105%. Per le donne 25-34enni non in un percorso d'istruzione, formazione e lavoro, invece l'aumento è del 61%. Tali dati riguardano gli individui che secondo la classificazione europea sarebbero disoccupati. Gli stessi dati che includono invece il totale (compresi inattivi e indisponibili e non in cerca), mostrano un aumento del 30% e del 22,5% per gli uomini e per le donne 15-24enni. Del 30% e del 7% è invece l'aumento per uomini e donne NEET dai 25 ai 34 anni.

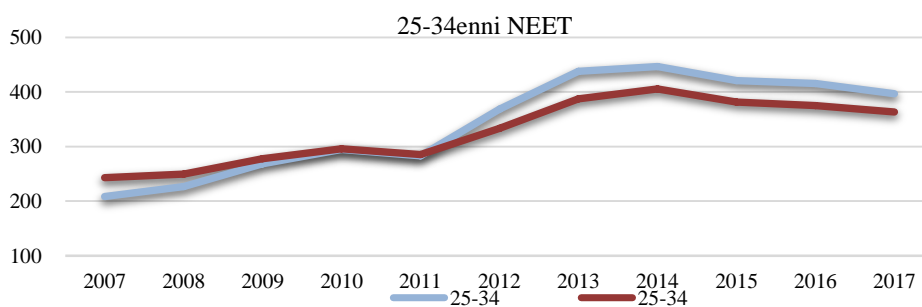
La Figura 3 mostra il numero assoluto NEET tra i 15 e i 34 anni. E' possibile osservare come per ognuno dei titoli di studio considerati il numero dei giovani che non sono inseriti in percorso di formazione, non vanno a scuola e non lavorano sia aumentato. La categoria maggiormente interessata al fenomeno è costituita dai diplomati.

Figura 1 - NEET 15-24enni a livello nazionale, categoria europea disoccupati, v.a.



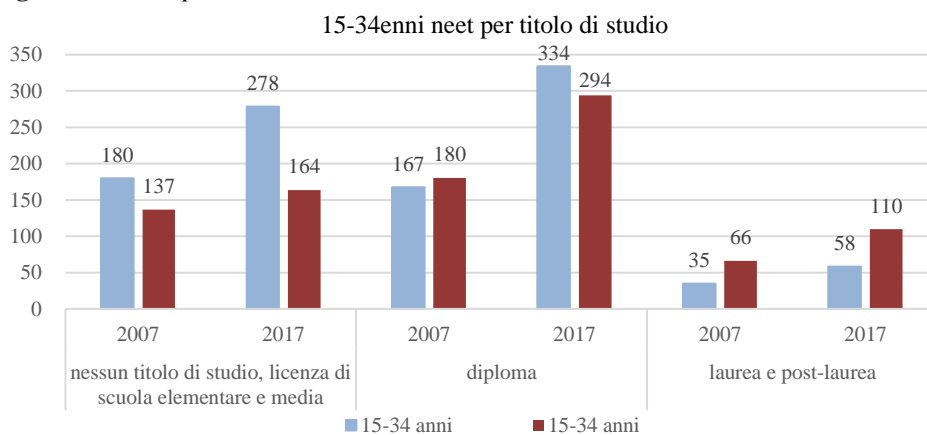
Note: Elaborazioni INAPP su dati ISTAT

Figura 2 - NEET 25-34enni a livello nazionale, categoria europea disoccupati, v.a.



Note: Elaborazioni INAPP su dati ISTAT

Figura 3 - NEET per titoli di studio 15-34enni, v.a.



Note: Elaborazioni INAPP su dati ISTAT.

4. Il modello per l'analisi empirica

L'analisi empirica si basa sulla stima di regressione Probit, al fine di calcolare la probabilità di un individuo di diventare NEET.

Dalle stime relative al 2007 emerge che le donne hanno un rischio di diventare NEET più elevato: tale rischio aumenta del 6% nel caso dei NEET da 15 a 24 anni, cresce invece dell'8,7% nel caso dei NEET dai 25 al 34 anni. Per i cittadini non italiani, aumenta la probabilità di diventare NEET del 5% sia che si consideri la prima fascia d'età, sia che si consideri la fascia d'età più elevata. Nella regressione Probit è stato incluso il tasso di disoccupazione dei quindici ventiquattrenni del 2007, per ogni regione. Questa variabile risulta correlata con un coefficiente dello 0,4% e dell'1,4% nel caso dei 25-34enni. Sui titoli di studio per i 15-24enni che non hanno alcun titolo, la probabilità di essere NEET, nel 2007, diminuisce del 10% per coloro i quali hanno il titolo di scuola media rispetto a coloro i quali hanno un titolo elementare o nessun titolo, le qualifiche diminuiscono circa del 6%. le probabilità di essere NEET in ambo le fasce d'età considerate. Il diploma diminuisce dell'11,8 e del 10,1 la probabilità di essere NEET. La laurea diminuisce tra i 15-24enni del 9% la probabilità rispetto a coloro che non hanno alcun titolo. Per la fascia d'età 25-34 invece diminuisce del 5%.

Essere coniugato ed essere divorziato diminuisce la probabilità di diventare NEET per i 25-34enni, rispettivamente del 4,7 e del 3,4. Tali coefficienti potrebbero essere collegati all'impegno familiare che porta alla necessità di lavorare.

Per la fascia d'età 15-24 invece lo stato civile non ha peso alcuno. Per i residenti del Sud e delle Isole, aumenta la probabilità di essere NEET dei 15-24enni del 2,8%. All'aumentare dell'età dei 15-24enni aumenta la probabilità di diventare NEET del 22%. Il contatto con i centri per l'impiego aumenta la probabilità di essere NEET del 16,4% e del 13,7%. Nonostante questo coefficiente possa essere letto in entrambe le direzioni, il segno positivo potrebbe significare che esisteva una correlazione nel 2007 tra i NEET e i centri per l'impiego.

Osservando le stime relative al 2017 la macro-area del Nord Est prevede un calo dell'1,8% rispetto al Nord Ovest per i 25-34enni. Risiedere al Centro sembra diminuire il rischio di essere NEET (-1,4%) del Nord Ovest per i 15-24enni.

Nel 2017, il problema di genere sembra non riguardare più i 15-24enni, aumenta tuttavia del 5,1% per le 25-34enni la probabilità di essere NEET. Essere cittadino straniero aumenta tale probabilità del 6,8% per i meno giovani delle fasce considerate e del 7,2% per i più giovani. Nelle stime relative al 2007, il valore relativo ai cittadini non italiani risulta più basso. L'aumento potrebbe rappresentare una maggiore difficoltà a trovare un'occupazione da parte degli stranieri.

Tabella 1 – Stime Probit relative al rischio di essere NEET per l'anno 2007 15-24enni e 25-34enni.

	2007pr_neet15	2007pr_neet25
	b/se	b/se
Nord Est	-0.008 (0.007)	-0.006 (0.007)
Centro	-0.007 (0.007)	0.006 (0.007)
Sud e Isole	0.028*** (0.009)	0.014 (0.012)
fem	0.006* (0.004)	0.087*** (0.004)
stran	0.050*** (0.009)	0.050*** (0.008)
Medie	-0.100*** (0.013)	-0.038*** (0.010)
Qualific	-0.065*** (0.015)	-0.066*** (0.012)
Diploma	-0.118*** (0.013)	-0.101*** (0.010)
Laurea	-0.094*** (0.016)	-0.050*** (0.010)
Coniugato	-0.006 (0.012)	-0.047*** (0.004)
Separato	0.074 (0.087)	-0.023 (0.019)
Divorziato	0.035 (0.074)	-0.034** (0.016)
Vedovo	0.000 (.)	-0.010 (0.028)
Unito civilemente	-0.083 (0.154)	-0.003 (0.044)
_15enni	0.004*** (0.000)	
A_eta	0.219*** (0.011)	-0.003 (0.018)
eta2	-0.005*** (0.000)	-0.000 (0.000)
cimpieg	0.164*** (0.004)	0.137*** (0.004)
_25enni		0.014*** (0.001)
Obs	41664	40975

Note: Elaborazioni INAPP su dati della Rilevazione Continua sulle Forze Lavoro (RCFL), Istat 2007.

Tabella 2 - Stime Probit relative al rischio di essere Neet per l'anno 2017 15-24enni e 25-34enni.

	2017pr_neet15	2017pr_neet25
	b/se	b/se
Nord Est	-0.001 (0.009)	-0.018* (0.010)
Centro	-0.014* (0.008)	0.007 (0.011)
Sud e Isole	0.003 (0.012)	0.022 (0.020)
fem	-0.007 (0.005)	0.051*** (0.006)
stran	0.072*** (0.009)	0.068*** (0.009)
Medie	-0.102*** (0.020)	-0.003 (0.019)
Qualific	-0.064*** (0.023)	-0.046** (0.023)
Diploma	-0.145*** (0.021)	-0.105*** (0.019)
Laurea	-0.131*** (0.024)	-0.092*** (0.020)
Coniugato	-0.011 (0.022)	-0.082*** (0.007)
Separato	-0.125 (0.081)	-0.062** (0.025)
Divorziato	0.000 (.)	0.010 (0.039)
Vedovo	-0.093 (0.177)	-0.015 (0.041)
Unito civilmente	0.000 (.)	0.006 (0.126)
15enni2017	0.004*** (0.000)	
A_eta	0.360*** (0.016)	-0.029 (0.028)
eta2	-0.008*** (0.000)	0.000 (0.000)
cimpieg	0.216*** (0.005)	0.201*** (0.005)
25enni 2017		0.008*** (0.001)
Obs	31586	26682

Note: Elaborazioni INAPP su dati della Rilevazione Continua sulle Forze Lavoro (RCFL), 2017.

I coefficienti che riguardano il titolo di studio diminuiscono sia per i diplomati che per i laureati. Un diplomato tra i 15 e i 24 anni ha il 14,5 in meno delle

probabilità e un 25-34enni ha il 10,5 in meno di probabilità di essere NEET. Un laureato ha il 13,1% in meno e il 9,2% in meno delle probabilità di essere Neet. Il fatto che si abbia avuto un contatto con un centro per l'impiego, aumenta del 21,6% e del 20,1% le probabilità di essere NEET. Questo coefficiente ancora una volta testimonia come nel caso dei NEET ci sia un contatto con i centri per l'impiego e che tale contatto si sia amplificato tra il 2007 e il 2017. Sicuramente l'intensità dei contatti con i centri per l'impiego potrebbe essere aumentata negli ultimi anni anche grazie alla importanza affidata agli stessi per l'intervento europeo Garanzia Giovani che contrasta il fenomeno NEET. Si controlla inoltre per i tassi di disoccupazione per i 15-24enni e per i 25-34enni.

5. Conclusioni

Il lavoro rappresenta una prima analisi sui fattori di rischio che incidono sulla probabilità di essere NEET.

Come abbiamo osservato dalle statistiche descrittive, dal 2007 al 2017 il fenomeno è aumentato, in particolare ha toccato la fase più acuta nel 2014. Dalle stime emerge inoltre una probabilità più elevata per le donne di non essere impegnate in alcun percorso formativo e lavorativo. E' nota in letteratura la difficoltà di partecipazione al mercato del lavoro delle donne e talvolta tale difficoltà può trasformarsi anche in demotivazione. Questo dato emerge in particolare per le fasce di donne 25-34enni, età in cui subentrano spesso gli impegni di cura della famiglia e dei figli.

In aggiunta, negli anni considerati, gli stranieri hanno una probabilità più elevata di non essere coinvolti in alcun percorso di istruzione, formazione e lavoro, probabilità che in uno scenario successivo agli anni di crisi più acuta, sembra aumentare. Tra i titoli di studio sono i diplomati a mostrare la probabilità meno elevate di diventare Neet, tuttavia anche la laurea mostra dei coefficienti negativi ma non troppo elevati. Ciò dimostra che probabilmente anche i titoli di studio più elevati aumentano il rischio di essere NEET.

Anche lo stato civile mostra alcune differenze: i coniugati e i separati hanno meno probabilità di diventare Neet, probabilmente questo si collega alla necessità di contribuire al reddito familiare.

Ulteriore elemento preso in considerazione è il contatto con i centri per l'impiego, cioè il contatto con cui, soprattutto a seguito dell'intervento di Garanzia Giovani, si cerca di contrastare il fenomeno. Si evidenzia infatti che i coefficienti relative al centro per l'impiego aumentano tra il 2007 e il 2017.

Il lavoro potrà essere arricchito successivamente da una serie di variabili esplicative che completano il quadro e con metodologie che tengano conto di problemi di selezione.

Riferimenti bibliografici

- BALAN M., 2015. Methods to estimate the structure and size of the “neet” youth, *Procedia Economics and Finance*, Vol. 32, pp: 119–124.
- BOERI T., GARIBALDI P., 2002. Shadow activity and unemployment in a depressed labour market, *CEPR Discussion paper*, No. 3433.
- BRUNELLO G., DE PAOLA M., 2014. The costs of early school leaving in Europe, *IZA Journal of Labour Policy*. <https://doi.org/10.1186/2193-9004-3-22>.
- BYNNER J., PARSONS S., 2002. Social exclusion and the transition from school to work: The case of young people not in education, employment or training (NEET), *Journal of Vocational Behaviour*, Vol. 60, No.2, pp.: 289–309.
- CAROLEO F. E., DEMIDOVA O., MARELLI E., SIGNORELLI M., (Eds.). 2018. *Young people and the labour market. A Comparative Perspective*, Oxford: Routledge.
- CICCIOMESSERE R., MONDAUTO M., MANIERI M., MARSALA A., CALABRESE S., 2011. NEET: i giovani che non studiano, non frequentano corsi di formazione e non lavorano. Caratteristiche e cause del fenomeno e analisi delle politiche per contenerlo e ridurlo, *Italia Lavoro*, disponibile su: http://www.robertocicciomessere.eu/Neet_ITALIA.pdf
- COPPOLA L., DI LAUREA D., 2016. Dynamics of persistent poverty in Italy at the beginning of the crisis, *Genus*, Vol. 72, No. 3.
- EUROFOUND, 2012. NEETs young people not in employment, education or training: Characteristics, costs and policy responses in Europe, *Luxembourg publications Office of the European Union*.
- FURLONG A., 2006. Not a very NEET solution: representing problematic labour market transitions among early school leavers, *Work. Employment and Society*, Vol. 20, No. 3, 553–569.
- INTERNATIONAL LABOUR OFFICE, 2015. *Global employment trends for youth*. Geneva: Scaling up investments in decent jobs for youth.
- ISTAT, 2013. *Noi Italia. 100 statistiche per capire il paese in cui viviamo*, Roma: Istat.

- ISTAT, 2017. *Il Mercato del Lavoro: la lenta ripresa e le disparità nei gruppi sociali*, Roma: Istat.
- MARELLI, E., e SIGNORELLI, M. 2015. Convergence, crisis and unemployment in Europe: the need for innovative policies, *Croatian Economic Survey*, Vol. 17, No. 2, pp: 5–56.
- MARSHALL, K. 2012. Youth neither Enrolled nor Employed, Perspectives on Labour and Income, *Component of Statistics Canada Catalogue Statistics*, No. 75-001-X, Canada.
- QUINTANO, C., MAZZOCCHI, P., e ROCCA, A. 2018. The determinants of Italian NEETs and the effects of the economic crisis, *Genus*, Vol. 74, No 5.
- RYAN, P. 2001. The school-to-work transition: a cross-national perspective, *Journal of Economic Literature*, Vol. 39, No.1, pp: 34–92.
- ROBSON, K. 2008. Becoming NEET in Europe: A Comparison of Predictors and Later-life Outcomes, *Global Network on Inequality Mini-Conference*, New York City.
- SALTO-YOUTH Inclusion Resource Centre. 2015. On track—different youth work for different NEET situations. *Support & advanced Learning and training opportunities within the Erasmus+: Youth in action programme*.
- SEU (Social Exclusion Unit), 1999. *Bridging the gap: New opportunities for 16-18 years old Not in Education, Employment or Training*, London: Report.

SUMMARY

Being NEET in Italy: the determinants of this rising phenomenon

The [Not in Employment, Education or Training] NEET category, first adopted in England in the mid-90s, has generated a great interest in the last years' Italian research. The share of NEET is increasing in many European countries, such as Italy.

The wide diffusion of NEETs represents an alarming social issue, as being NEET predisposes young people to long-term unemployment and social exclusion. It also has a significant negative impact on the economic growth and welfare equilibrium of countries. The aim of this paper is to identify the determinants, or the risk factors, behind this phenomenon. This work puts together theory and empirical findings in order to carry out a complete statistical analysis using the most recent data from the Labour Force survey drawn by ISTAT. It also gives an insight about the effects of the 2008 economic crisis on young Italians' occupations and behaviors. The results confirm the crucial role of the education system, as well as the importance of the economic and social disparities between gender and the Italian territorial districts.

Irene BRUNETTI, INAPP – Istituto Nazionale per l'Analisi delle Politiche Pubbliche, i.brunetti@inapp.org

Valentina FERRI, INAPP – Istituto Nazionale per l'Analisi delle Politiche Pubbliche, v.ferri@inapp.org

MEDIATION: ECONOMIC CONCEPTS AND SOME EXAMPLES OF RATIONAL FRAMEWORK FOR LEGAL PROFESSIONALS

Antonio Cappiello

The skilful management of conflict is among the highest of human skills
S. Hampshire, *Justice is Conflict* (Princeton University Press, 2000)

1. Introduction

Mediation, as alternative dispute resolution, is an important option for policy makers to alleviate the judicial burden and allowing both parties more freedom in the agreement for the solution of the conflict. The customization of the output according to the claimants' specific needs and wills can therefore maximize the satisfaction of both parties beyond the legal constraints of a judicial decision. Since the focus is on settlement, parties may give up their legal rights and their constraints in the favour of a quick, peaceable and more convenient solution.

To reach this objective we should consider that one of the main issues in solving conflicts is the determination of rationales and interests. Both parties want to know how their antagonists arrive to the decisions. That is the reason we should concentrate more on rationales than on positions. We will underline the importance of establishing a common rational framework as a reference for analysing the positions taken by the parties. This approach is useful to solve litigation and/or facilitating mediation but is also a good way to argument a position when discussing some assumptions in any kind of reasoning.

2. Legal rational, economics and possible biases

It should be noted that the rational framework is a tool for the analysis once the legal analysis applied to the specific case has already been done.

Afterwards, we could associate some probability to possible scenarios. Every party is evaluating their positive outcome according to the probability of possible favourable judicial decision. Once probabilities have been assigned to each possible outcome, a decision tree can be a useful tool in order to correctly evaluate the outcome.

Suppose that the track to arrive to our outcome is composed of 4 stages, and the probability of positive outcome from each stage is 70%, a straightforward propensity to get over these 4 steps may be observed in the majority of parties

involved in a dispute. However, if you just write down a parametric scheme, you may reconsider your chances of success.

3. Writing down your intuitive evaluation and reconsidering the framework

If we write down all the stages and the track leading to the final output, we may have a clear vision on the probability connected to our successful possibilities.

The following is an example illustrating how writing down some intuitive reasoning may lead to a different vision of our possible choice. Suppose we have to complete 4 consecutive stages in order to achieve a final results. Each stage has 70% probability of success. Would you judge the reaching of the 4th final stage an event that will occur with a high probability? If we just do not consider the entire process we could intuitively say that this is very likely to happen. However, let us try to calculate:

The probability to reach a stage n (considering independent probabilities for each stage) is given by the following formula:

$$\prod_{i=1}^n p_i \quad (1)$$

Therefore, in the following table we can indicate the probability to successfully complete 1, 2, 3 or 4 stages.

	Stage 1	Stage 2	Stage 3	Stage 4
calculation		$\prod_{i=1}^2 p_i$	$\prod_{i=1}^3 p_i$	$\prod_{i=1}^4 p_i$
probability to successfully complete n stages	0,7	0,49	0,343	0,2401

Someone could perhaps find the result of the calculation rather different from what he was deducing from a pure intuitive reasoning.

Applying the same reasoning to a litigation case, we could visualize possible outcomes and avoid psychological biases (overconfidence, imperfect information, etc.).

Of course, we can not eliminate biases but at least we can be aware of them and try to consider all the aspects from the quantitative to the subjective evaluation.

For instance, the same person can have a different evaluation on the cost of a house (depending on whether he has to buy or sell it). Disputing parties will bring these same biases accordingly; they will concentrate more on their desired outcome than on their factual situation.

The mediator should agree with the parties to a shared rational grid in order to visualize together a common framework (and therefore shared possible outcome) generating the best solution for them.

It is important to explain how the outcome of their needs is positioned on the grid of rationales according to shared neutral parameters. The grid will underline inconsistencies among positions and needs (objectives).

Positions may be unconciliating but needs can be in a sharing area where both parties feel better off compared to a conflict situation where positions can in the paradox lead to a suboptimal condition also for the party which reached its desired position.

4. Parametric assessment and probability assessment

All the above said, it is now clearer that defining potential results is an essential step, but to each result we should try to assign a probability in order to be able to make a rational decision.

For instance, it seems great to expect infinite gain from a *Saint Petersburg lottery*, but we should assess the probability of higher gain, as well as the available time and resources to play this game.

It is also good to aim for the best lottery prize, but obviously the probability connected to it would discourage a rational agent.

All these assessments could be set parametrically and be represented graphically. There are many ways to represent probability framework for assessment.

We start describing in the next paragraph the decision tree; one of the most used graphic tools used by mediators, especially in North America. Therefore, we make some considerations on other alternative tools and reference scheme always connected to probability and decision-making.

5. Tree structure according to the graph theory and decision tree

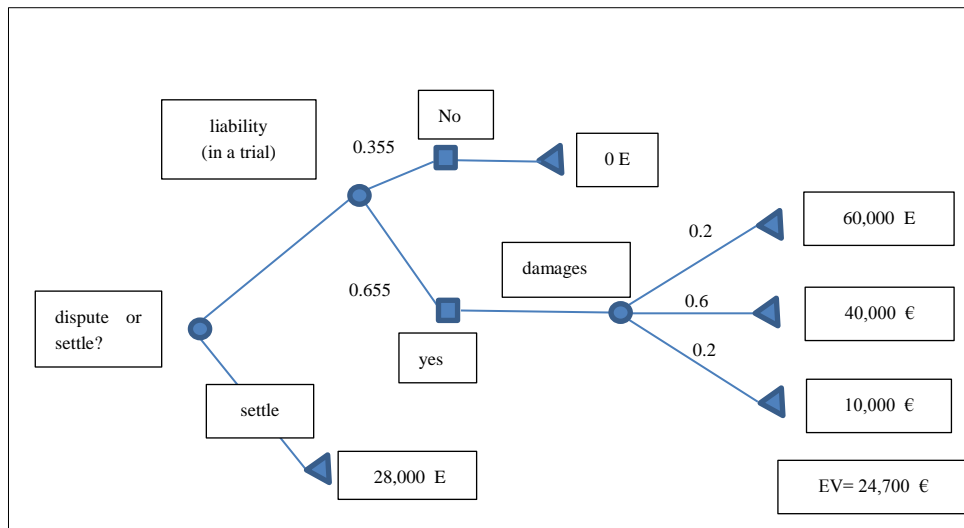
According to the graph theory, a *tree graph* Φ , with finitely n vertices (nodes), satisfies the following equivalent conditions:

- Φ has $n - 1$ edges and connected nodes
- Φ has $n - 1$ edges and has no simple cycles

We define: nodes” every tree element, “branches” (edges) the lines that connect nodes, “end nodes” nodes that do not produce other nodes or other patterns. In a finite tree graph there is always a member with no superior that is called “root node” (or simply “root”).

Applying the tree structure to a decisional process we would obtain a decision tree. This structure allows the visualisation of choices and their possible consequences according to a certain degree of probability

Figure 1 – *Simplified intuitive tree structure about 2 alternatives: settlement vs dispute.*



Each “node” of the tree is an event and each branch is the output of the event. Each “end node” is a class representing the final decision coming from its node path (the paths, going from the “node root” to the “end node” are intended as “classification rules”).

Table 1 – Components and rules of a decision tree.

Components and rules	description
Elements (nodes)	Nodes generating only splitting and no-converging paths (path direction: from left to right)
Rules (path)	Each path, from root to the end node, is a decision rule and the output is represented by the content of the end node e.g. if conditions (node 1, node 2 ...node n-1) are verified then the outcome “node n” is produced. For instance, an association rule to generate the path form a node to another node (always from left to right) can be based on causal or temporal relation.

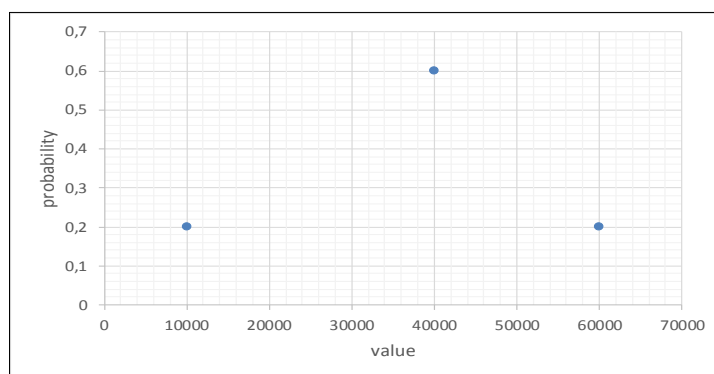
A tree structure for decision analysis allows the visualisation of the expected value (or expected utility) as consequences of alternatives.

Usually nodes can be of the following type: decision node (graphically represented by a square), chance node (graphically represented by a circle), end node (graphically represented by a triangle).

6. Expected value and risk assessment

In order to better evaluate the expected outcome of the trial, we could assess the risk associated with each possible outcome. In case of winning, we could graphically display the probability indicated in Figure 1.

Figure 2 – Probabilities of possible compensations for damages in case of winning.



However, taking into account the risk of losing, we should reconsider the EV of our outcome and the probabilities associated to each possible winning output would therefore be displayed as in Fig. 3. In figure 4, we can have a visual comparison on how, considering the possibility of losing, the outcome probabilities are proportionally reduced.

Figure 3 – Adjusted probability of possible compensation for damages.

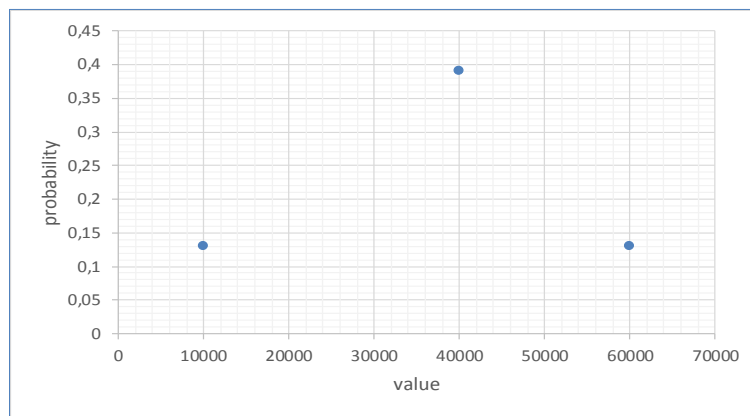
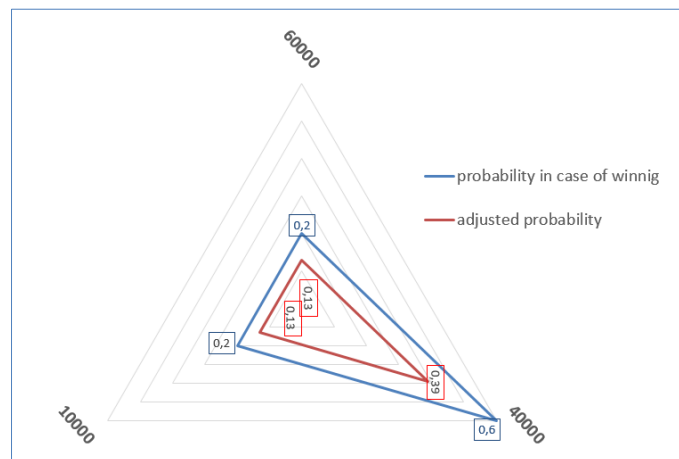


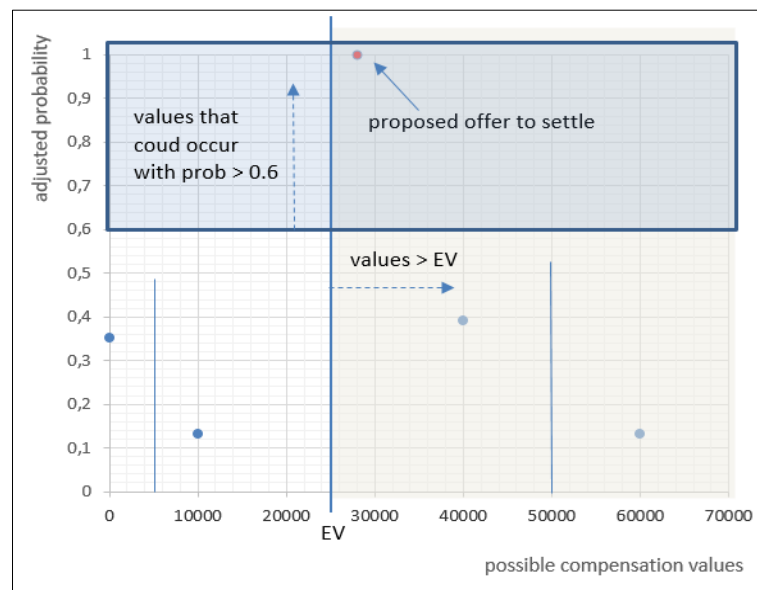
Figure 4 – Adjusted probabilities vs outcome probabilities in case of winning.



Finally, Figure 5 shows all the possible outcomes - including losing - and how an offer to settle in order to avoid the trial should be considered in the overall context.

The proposed sum to settle (28,000) is higher than the EV (24,700) and presents no risk. The temptation to get higher compensation related to the possible win 40,000 (39%) or 60,000 (13%) should be measured in a view of considering a fair possibility (48%) of getting smaller compensation of the proposed amount to settle (as indicated in Fig. 5: 0; 35% or 10,000;13%).

Figure 5 – Evaluating the overall risk: adjusted probability and EV.



7. Final considerations and the importance of the sensitivity analysis

In the previous paragraphs, we showed a decision process based on parameters and numbers but the synthesis of it is a very complex operation and requires wisdom. As Plato said a long time ago: “*a good decision is based on knowledge and not on numbers*” (Laches, 380 B.C.). In fact, we could find good examples in many academic studies about parametric frameworks that could help the legal professionals and their clients in making a cautious and efficient decision; actually, all proposed models in the literature seem reasonable. One of the main problems is, once defined the probable path of the events according to possible different choice, to assign the probability of each defined outcome. This would be usually done bearing in mind previous experiences and considering case law or any other empirical elements. In most of the cases, for several reasons (e.g. lack of previous judgements on similar issues, extreme complexity of the case involving elements

that could be hardly caged in common judgements and easy evaluation assessment), this is a very complicated exercise. If this could give you some comfort, at the beginning of the XX century de Finetti and Ramsey considered the probability as a manifestation of a personal interpretation on a certain observed phenomenon. The implication of this subjective approach is that the assessing of the probability will be intended as a process involving logical and psychological elements and producing personal views compatible with the available data. Therefore, contrary to the classical and frequentist theory, probability is not always an empirical process that leads to univocal considerations coming from data. At this point, the most important aspect for a mediator is to put order about needs, claims and desires of the parties and therefore sharing with them every single element in developing a coherent framework. As stated by de Finetti “*before approaching a problem to solve you need to see it*”, we should avoid to be trapped in formalism and use simpler and factual means (concrete tools) in order to reach our objectives. For instance, it will be crucial to write down as many elements as possible because only transforming the words in quantity we better understand the thought of every involved party. For instance, a “a good possibility for a certain event to happen” could be 60% according to a lawyer, but for his client could mean 80% and for their opponent only 55%. Therefore, the mediator will find it out only when writing down the decision framework after asking and listening all the opinions of the involved actors. The first mediator’s task would therefore be to start “mediating” the probabilities to assign to each possible outcome. At this point, as long as the parties agree on the proposed model and consider reasonable the assigned probability to each outcome, it will be easier to reach a satisfactory agreement. In fact, all the process leading to the settlement would be evaluated and visualised by both parties on the basis of clearer and shared scenarios that they themselves contribute to construct step by step.

References

- ANSCOMBE, F.J., AUMANN R.J. 1963. A definition of subjective probability. *Annals of Mathematical Statistics*, Vol. 34, pp. 199–205.
- AYRES I., NALEBU B.J. 1997. *Common Knowledge as a Barrier to Negotiation*, Yale ICF Working Paper No. 97-01.
- BROWN J.G., AYRES I. 1994. Economic Rationales for Mediation. *Virginia Law Review*, Vol. 80, pp. 323-402.
- CAPPIELLO A. 2016. Decision making and Saint Petersburg Paradox: focusing on heuristic parameters, considering the non-ergodic context and the gambling risks. *RIEDS Rivista Italiana di Economia, Demografia e Statistica - Italian Review of Economics, Demography and Statistics*, Vol. 70, No. 4, pp. 147-158.
- CAPPIELLO A. 2015. Underlined biases in the Doing Business Index: World Bank adjustments and passage from the percentile rank to the DTF rank. *Invited Paper at 6th International Academic Conference for Students and Graduate Students Statistical Methods for Analysis of the Economy and Society*, School of Statistics and Data Analysis, Faculty of Economics of the National Research University Higher School of Economics (HSE) (12-14 May)
- CAPPIELLO A. 2014. World Bank Doing Business Project and the statistical methods based on ranks: the paradox of the time indicator. *RIEDS - Rivista Italiana di Economia, Demografia e Statistica - Italian Review of Economics, Demography and Statistics*, Vol. 68, No. 1, pp. 79-86.
- CARNEVALE P.J., PRUITT D.G. 1992. Negotiation and Mediation. *Annual Review of Psychology*, Vol. 43, pp. 531-82.
- DE FINETTI B. 1937. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'institut Henri Poincaré*, Vol. 7. No. 1, pp. 1-68.
- DEBREU G. 1959. *Theory of Value*. New Haven, CT: Cowles Foundation.
- GERARDI D. 2004. Unmediated Communication in Games with Complete and Incomplete Information. *Journal of Economic Theory*, Vol. 114, pp. 104-131.
- KYDD A. 2006. When Can Mediators Build Trust. *American Political Science Review*. Vol. 100, pp. 449-462.
- PHILBIN D.R. 2008. The value of economic analysis in preparing for mediation. *Dispute Resolution Journal*, Vol. 3, No. 1, p. 49.
- RAIFFA H. 1982 *The Art and Science of Negotiation*. Harvard University Press.
- WALL J.A., STARK J.B., STANDIFER R.L. 2001. Mediation: A Current Review and Theory Development. *Journal of Conflict Resolution*, Vol. 45, pp. 370-391.

SUMMARY**Mediation: Economic Concepts
And Some Examples Of Rational Framework For Legal Professionals**

Mediation, as alternative dispute resolution, is an important option for policy makers to alleviate the judicial burden and allowing both parties more freedom in the agreement for the solution of the conflict. The customization of the output according to the claimants' specific needs and wills can therefore maximize the satisfaction of both parties beyond the legal constraints of a judicial decision.

We will underline the importance of establishing a common rational framework as a reference for analysing the positions taken by the parties. This approach is useful to solve litigation and/or facilitating mediation but is also a good way to argument a position when discussing some assumptions in any kind of reasoning.

Antonio CAPPIELLO, Economist, National Council of the Civil Law Notaries (Consiglio Nazionale del Notariato) and advisor at International Cooperation Center for Statistics ICstat. cnn.acappiello@notariato.it; economics@live.co.uk

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: Dott. CHIARA GIGLIARANO

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864



Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE

La copertina è stata ideata e realizzata da Pardini, Apostoli, Maggi p.a.m. @tin.it – Roma

Stampato da CLEUP sc
“Coop. Libreria Editrice Università di Padova”
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

ATTIVITÀ DELLA SOCIETÀ

A) RIUNIONI SCIENTIFICHE

- XXXVII La mobilità dei fattori produttivi nell'area del Mediterraneo (Palermo, 15-17 giugno 2000).
- XXXVIII Qualità dell'informazione statistica e strategie di programmazione a livello locale (Arcavacata di Rende, 10-12 maggio 2001).
- XXXIX L'Europa in trasformazione (Siena, 20-22 maggio 2002).
- XL Implicazioni demografiche, economiche e sociali dello sviluppo sostenibile (Bari, 15-17 maggio 2003).
- XLI Sviluppo economico e sociale e ulteriori ampliamenti dell'Unione Europea (Torino, 20-22 maggio 2004).
- XLII Sistemi urbani e riorganizzazione del territorio (Lucca, 19-21 maggio 2005).
- XLIII Mobilità delle risorse nel bacino del Mediterraneo e globalizzazione (Palermo, 25-27 maggio 2006).
- XLIV Impresa, lavoro e territorio nel quadro dei processi di localizzazione e trasformazione economica (Teramo 24-26 maggio 2007).
- XLV Geopolitica del Mediterraneo (Bari, 29-31 maggio 2008).
- XLVI Povertà ed esclusione sociale (Firenze 28-30 maggio 2009).
- XLVII Un mondo in movimento: approccio multidisciplinare ai fenomeni migratori (Milano 27-29 maggio 2010).
- XLVIII 150 anni di Statistica per lo sviluppo del territorio: 1861-2011. (Roma 26-28 maggio 2011).
- XLIX Mobilità e sviluppo: il ruolo del turismo. (San Benedetto del Tronto, 24-26 maggio 2012).
- L Trasformazioni economiche e sociali agli inizi del terzo millennio: analisi e prospettive (Università Europea di Roma, 29-31 maggio 2013).
- LI Popolazione, sviluppo e ambiente: il caso del Mediterraneo (Università Federico II di Napoli, 29-31 maggio 2014).
- LII Le dinamiche economiche e sociali in tempo di crisi (Università Politecnica delle Marche, 28-30 maggio 2015).
- LIII Mutamento economico e tendenze socio-demografiche tra sfide e opportunità (Università degli Studi Internazionali di Roma, 26-28 maggio 2016).
- LIV Mobilità territoriale, sociale ed economica: modelli e metodi di analisi (Università degli Studi Internazionali di Catania, 25-26 maggio 2017).
- LV Coesione sociale, welfare e sviluppo equo e sostenibile (Università degli Studi dell'Insubria, Varese 24-25 maggio 2018).