

VOLUME LXXIV – N. 1

GENNAIO – MARZO 2020

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 QUADRO 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.
info@sieds.it, rivista@sieds.it

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

info@sieds.it

rivista@sieds.it

VOLUME FUORI COMMERCIO – DISTRIBUITO GRATUITAMENTE AI SOCI

INDICE

Margherita Gerolimetto, Stefano Magrini <i>Bootstrap methods for long-range dependence: Monte Carlo evidence</i>	5
Giuseppe Ricciardo Lamonica, Luca Salvati, Margherita Carlucci <i>Vertical specialization across the world: evidence from the world input-output table</i>	17
Marco Lattanzio <i>The new methodology for the STS indicator on the “number of persons employed”</i>	29
Sandro Stancampiano <i>Management of cultural heritage using Big Data Data</i>	41
Andrea Fasulo, Marco D. Terribili <i>International cooperation for redesigning Moldovan social surveys</i>	51
Claudia De Vitiis, Francesca Inglese <i>La valutazione dell’effetto tecnica nelle indagini sociali a tecnica mista</i>	65
Eugenia De Rosa, Claudia De Vitiis, Francesca Inglese, Alberto Vitalini <i>Il web-Respondent Driven Sampling per lo studio della popolazione LGBT+</i>	73
Antonella Bernardini, Daniela Bonardo, Alessandra Dentini, Maura Giacummo, Matteo Mazziotta, Alessandra Preti <i>The Italian continuous censuses: managing the complexity among registers, statistics and territories</i>	85
Patrizia Collesi, Alessandro Valentini <i>Le attività dell’Istat per la scuola e sul territorio</i>	97

BOOTSTRAP METHODS FOR LONG-RANGE DEPENDENCE: MONTE CARLO EVIDENCE

Margherita Gerolimetto, Stefano Magrini

1. Introduction

Bootstrap methods (Efron, 1979) are a computer intensive approach, based on resampling, to statistical inference issues without any statistical assumption on the underlying data generating process. The bootstrap is often adopted because it has much better performance than the conventional approaches (first order asymptotics) and it provides empirical and efficient inference for complicated problems. However, bootstrap methods cannot be equally applicable to all random processes.

In general, for data that are not independent identically distributed (iid), modifications of the original bootstrap set-up have been put forward in order to be able to resample without ignoring the dependence structure. For time series data, among the most common proposals are the block-bootstrap by Künsch (1989) that follows a “data-block” mechanism and the sieve bootstrap by Bühlman (1997) that shares the principle of data transformation. Both methods are time domain-based and have been developed for weakly dependent time series data (e.g. from short memory ARMA models) and. In a parallel line of research, there are also proposals in the frequency domain; in particular, Dahlhaus and Janas (1996) in their seminal paper established a bootstrap method for ratio statistics in case of weakly dependent data, where the Fourier transform is adopted in order to weaken the dependence structure.

In the last decades there has been an increasing interest in strongly dependent (or long-range dependent) time series processes, that can be in general categorized as those processes whose autocovariance decays slowly, in contrast to weakly dependent (or short-range dependent) processes whose covariance decay is fast. Statistical issues between strongly and weakly dependent data change dramatically and, unsurprisingly given the high persistence implied by a slowly decaying autocovariance, the use of resampling methods is more complicated for highly persistent data. The degree of complexity increases even more in case of the dependence structure is not just highly persistent but actually nonstationary. As we

will see in the following sections, this will require a specific additional step of first order differencing of the time series.

The aim of this work is two-fold. On the one hand, we study via Monte Carlo simulations the performance of block bootstrap and sieve bootstrap methods, originally developed for weakly dependent time series, in case of strongly dependent time series, both stationary and nonstationary. On the other hand, we propose an approach to improve the performance of those methods in case of nonstationarity and we will show its finite sample performance via Monte Carlo.

The strongly dependent data generating processes we consider is long memory processes, in particular FARIMA(p, d, q). For these processes, in their stationary version ($d < 1/2$), in literature there are contributes where it has been extended the validity of the block bootstrap (Kim and Nordman, 2011) and sieve bootstrap (Poskitt, 2008). To the best of our knowledge, what happens in the nonstationarity case ($d > 1/2$) is not clear yet and in this paper we intend to make a move in this direction by investigating the effective performance of these methods for data generating processes that are not just strongly dependent, but non-stationary. Results show that the performance of these bootstrap methods worsens with the increase of the persistence, however it improves if the bootstrap algorithms are augmented with the additional step depicted in this paper.

The structure of the paper is as follows. In the second section, we provide details on the long memory. In the third section, we describe the bootstrap methods under examination. In the fourth section, we present our proposal of a modification of the bootstrap algorithms to handle nonstationarity. In the fifth section, we present our Monte Carlo experiments and some conclusions.

2. Long-range dependence

A linear stationary process $\{X_t\}$ with mean $E(X_t) = \mu$ is defined as

$$X_t = \mu + \sum_{j \in \mathbb{Z}} b_j \epsilon_{t-j} \quad (1)$$

where $\{\epsilon_t\}$ are i.i.d. innovations with $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) < \infty$, $\sum_{j \in \mathbb{Z}} b_j^2 < \infty$. The process $\{X_t\}$ is characterized by long-range dependence (LRD) if the autocovariance function $r(k) = Cov(X_0, X_k)$ satisfies the following slow decay condition:

$$r(k) \sim \sigma^2 k^{-\theta}, k \rightarrow \infty \quad (2)$$

for some $\theta \in (0,1)$ and $\sigma^2 > 0$. If $\theta = 1$, the process is characterized by short-range dependence (SRD). Another condition for the process $\{X_t\}$ to be LRD is that the partial sums of the autocovariances $\sum_{k=1}^{\infty} |r(k)| = O(n^{1-\theta})$ diverges as $n \rightarrow \infty$ (Robinson, 1995). On the other hand, for short-range dependent time series the autocovariance decay rapidly to 0 as $k \rightarrow \infty$, so that $\sum_{k=1}^{\infty} |r(k)| < \infty$.

Since in this paper we consider inference about the sample mean $\bar{X}_n = n^{-1} \sum_{t=1}^n X_t$, we need to characterize the LRD property in terms of the behavior of the sample variance of the sample mean. If we denote $\sigma_{n,\theta}^2 = n^\theta \text{var}(\bar{X}_n)$, the slow decay condition implies that

$$\lim_{n \rightarrow \infty} \sigma_{n,\theta}^2 = \sigma_{n,\infty}^2 > 0 \quad (3)$$

holds for a constant $\sigma_{n,\infty}^2$, depending on $\theta \in (0,1)$. The implication behind (3) is important as it means that the sample variance estimator of the sample mean, $\text{var}(\bar{X}_n)$, decays under LRD at a slower rate $O(n^{-\theta})$ as $n \rightarrow \infty$ than the usual $O(n^{-1})$ under SRD (that indeed it is obtained if $\theta = 1$).

So, for a linear process with LRD, if condition (3) holds, we have the normal limit for the scaled sample mean (Davidov, 1970)

$$n^{\theta/2}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma_{n,\infty}^2) \quad \text{as } n \rightarrow \infty \quad (4)$$

In this paper, we will focus in particular on the class of long memory FARIMA(p,d,q) models

$$\phi(B)(1-B)^d X_t = \mu + \psi(B)\epsilon_t \quad (5)$$

where $\{\epsilon_t\}$ are white noise, as above, B is the lag operator and the long memory parameter d is $d = \frac{1-\theta}{2}$. So, if $\theta = 1$, we have that $d = 0$, *i.e.* ARMA(p,q) models. Stationarity holds if $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. The long memory parameter d can be estimated in a variety of ways, here we will adopt two well-known methods: the GHP method (Geweke and Porter-Hudack, 1983) and Whittle method (Fox and Taqqu, 1986; Dahlhaus, 1989).

3. Bootstrap for time series: an overview

The bootstrap is a method to estimate the distribution of an estimator or a test statistic by resampling one's data. Under iid conditions, the bootstrap provides

approximations that are at least as accurate first order asymptotics. Often the bootstrap provides approximations that are even more accurate, especially when the sample size is not large

When data are not iid, as in time series case, the bootstrap should be revisited in order to preserve the dependence structure in the data generating process (DGP). In the next subsections, we revise two methods developed with that purpose.

3.1. Block Bootstrap

The block bootstrap (Künsch, 1989) is based on the principle of capturing or preserving the dependence structure of the original time series. In practice, it means resampling data blocks, consisting of consecutive groups of data points.

Two types of block bootstrap have been proposed and for both validity has been originally proved under short-range dependence. One is the moving block bootstrap (MBB), when resampling is done with respect to overlapping blocks (Künsch, 1989) the other one (Carlstein, 1986) is when resampling is done with respect to non-overlapping blocks (NBB). The following is the algorithm:

- 1) Given a time series $X_t, t=1, \dots, n$, define block size $l < n$ and compute the number of blocks for the resampling procedure $b = \lfloor \frac{n}{l} \rfloor$
- 2) Construct data blocks (either overlapping, MBB, or not overlapping, NBB)
- 3) Generate bootstrap replicates from the data-block set as:
 - (a) for a MBB series X_1^*, \dots, X_N^* , where $N = bl$, we generate I_1^*, \dots, I_b^* from iid uniform random variables $\{I_1^*, \dots, I_{n-l+1}^*\}$
 - (b) for a NBB series X_1^*, \dots, X_N^* , where $N = bl$, we generate I_1^*, \dots, I_b^* from iid uniform random variables $\{I_1^*, \dots, I_{n+l(b-1)}^*\}$
- 4) Make a MBB/NBB series X_1^*, \dots, X_N^* , where $N = bl$

Approximating the distribution of $n^{\theta/2}(\bar{X}_n - \mu)$ with the bootstrap counterpart $N^{\theta/2}(\bar{X}_n - E_*\bar{X}_n^*)$, that is valid in case of short-range dependence, appears to be a natural choice also in case of LRD. However, in this latter case, it is actually wrong as pointed out by Künsch (1989) and Lahiri (1993). More recently, Kim and Nordman (2011) proved that for long memory time series, the sample mean should be inflated by adjustment factor of $b^{1-\theta/2}$ and the distribution of $n^{\theta/2}(\bar{X}_n - \mu)$ will be approximated as follows:

$$\sup_{x \in \mathbb{R}} \left| P_* \left(b^{\frac{1}{2}} l^{\frac{\theta}{2}} (\bar{X}_n^* - E_* \bar{X}_n^*) \leq x \right) - P \left(n^{\frac{\theta}{2}} (\bar{X}_n - \mu) \leq x \right) \right| \xrightarrow{p} 0 \quad (6)$$

In other words, the inflating factors ensures that the MBB version $b^{\frac{1}{2}}l^{\frac{\theta}{2}}\bar{X}_n^*$ has the “right” variance for approximating the distribution of $n^{\theta/2}\bar{X}_n^*$ under LRD (the same hold for NBB).¹

Note that the performance of the BB depends on a block choice and optimal blocks for variance estimation are shown to decrease as the strength of the underlying process increases. Blocks of size $O(n^{1/2})$ may be a compromise for use in practice.

3.2. Sieve Bootstrap

The sieve bootstrap (Bühlmann, 1997) approximates a general linear invertible process by a finite autoregressive model of order $p = p(n)$, where $p(n) \rightarrow \infty$, then it resamples from the approximated autoregression. This method takes up the older idea of fitting parametric models first and then resampling from the residuals, but instead of considering a fixed finite-dimensional model, an infinite-dimensional, non-parametric model is approximated by a sequence of finite-dimensional models. This method can be considered non parametric because it is model-free in the class of the linear invertible processes.

The properties of the sieve bootstrap have been rigorously investigated, among other, by Kreiss (1992), Paparoditis (1996), Bühlmann (1997), Bickel and Bühlmann (1999) who established its asymptotic validity for several statistics assuming that the data generating process is an infinite order autoregressive. Kapetanios and Psaradakis (2006) and Poskitt (2008) proved that under regularity conditions (satisfied by stationary long memory processes) the sieve bootstrap provides an asymptotically valid approximation to the distributions of several statistics. The following is the algorithm:

- 1) Given a time series X_t , $t = 1, \dots, n$, fit an AR(h) model $h > 0$. Obtain the residuals of the AR(h) model and standardize them, denote by $\hat{\zeta}_t$
- 2) Create a new randomly resampled residuals set, denoted by $\hat{\zeta}_t^*$
- 3) Generate the bootstrap time series X_t^* as

$$X_t^* = \hat{\alpha}_1 X_{t-1}^* + \dots + \hat{\alpha}_h X_{t-h}^* + \hat{\zeta}_t^* \quad (7)$$

- 4) By repeating the above procedure a number of times B we obtain a bootstrap approximation to the distribution of the desired statistic

¹ We emphasize that (6) holds for stationary long memory, *i.e.* $|d| < 1/2$, the nonstationary area ($d > 1/2$) has not been investigated yet.

For stationary long memory time series Poskitt (2008) proved that the distribution of $n^{\theta/2}(\bar{X}_n - \mu)$ will be approximated as follows:

$$\sup_{x \in \mathbb{R}} \left| P_* \left(n^{\frac{\theta}{2}} (\bar{X}_n^* - \bar{X}_n) \leq x \right) - P \left(n^{\frac{\theta}{2}} (\bar{X}_n - \mu) \leq x \right) \right| \xrightarrow{p} 0 \quad (8)$$

Note that the SB performance depends on the order h of the AR model fitted to the data. A reasonable choice is to adopt the Akaike criterion as suggested in Bühlmann (1997).

4. Bootstrapping nonstationary long memory time series

When, X_t is nonstationary, *i.e.* $d > 1/2$, bootstrapping is an even more complex issue. Inspired by the literature of bootstrap for unit root-test, the idea we propose here is to preliminarily first difference the nonstationary long memory time series (Palm *et al.* 2008), in order to bring it back to the stationary area; then we apply the chosen bootstrap to $Z_t = \Delta X_t$. This is also in line with Psaradakis (2001) and Chang and Park (2003) who proved that applying the sieve bootstrap to first difference is a valid bootstrap approach to nonstationary I(1) time series.²

So, the chosen bootstrap algorithm is augmented by two additional steps, before and after the implementation of the bootstrap itself:

[Add 1]: Take the first difference $Z_t = \Delta X_t$

[Boot]: Obtain a bootstrap sample Z_t^* with the chosen bootstrap

[Add 2]: Obtain $X_t^* = X_{t-1}^* + Z_t^*$

By preliminarily differentiate the time series, the conditions required to guarantee the validity of the bootstrap are respected. Indeed, the series which is effectively bootstrapped is $Z_t = \Delta X_t$, whose long memory parameter is $d_z = d - 1$, will be $|d_z| < 1/2$, so stationary long memory.

² Palm *et al.* (2008) also showed that for some data generating processes (not our case), residuals from a first order autoregression can lead to an even better performance of the sieve bootstrap in terms of asymptotic validity, compared to first difference.

5. Monte Carlo experiment

The aim of the experiment is to investigate the performance of block bootstrap and sieve bootstrap to obtain confidence intervals for μ the mean of the process for a variety of long memory DPGs, both stationary and nonstationary. The number of MC replications is $S=5000$. Both stationary ($d=0.25, 0.45$) and nonstationary ($d=0.65, 0.80$) FARIMA models have been considered, in four versions (that include or not the short-range components):

DGP1: FARIMA(1,d,1), $\phi=0.3, \psi=0.4$

DGP2: FARIMA(0,d,1), $\psi=0.4$

DGP3: FARIMA(1,d,0), $\phi=0.3$

DGP4: FARIMA(0,d,0)

A further DGP has been included for benchmarking with the short memory case.

DGP5: ARMA(1,0,1), $\phi=0.3, \psi=0.4$

The sample size is $n = 100, 250, 500$, innovations are $\epsilon_t \sim N(0,1)$. The considered methods are block bootstrap (BB) and sieve bootstrap (SB) both in the stationary and revised version for nonstationarity.

The BB is implemented in the overlapping version, the block length is $l = \lceil \sqrt{n} \rceil$, in line with, for instance Kim and Nordman (2011). The SB is implemented using Yule walker estimates, the AR order h is set adopting the Akaike criterion as in Bühlmann (1997). $B=500$ is the number of bootstrap samples.

The performance is expressed in terms of empirical coverage of (symmetric) bootstrap intervals at 90%: $\bar{X}_n^* \pm n^{-\theta/2} q_{0.90}^*$, where $q_{0.90}^*$ is defined as

$$P_* \left(\frac{1}{b^2 l^2} |\bar{X}_n^* - E_* \bar{X}_n^*| \leq q_{0.90}^* \right) = 0.9 \text{ for the block bootstrap}$$

$$P_* \left(\frac{\theta}{n^2} |\bar{X}_n^* - \bar{X}_n| \leq q_{0.90}^* \right) = 0.9 \text{ for the sieve bootstrap}$$

Note that $\theta = 1 - 2d$. Bootstrap interval are computed without knowledge on d and it is adopted the Whittle or GPH estimate, so that $\hat{\theta} = 1 - 2\hat{d}$.

Results are presented in the following Tables 1, 2, 3. The performance, *i.e.* the empirical coverage, should be read in the sense that the closer it is to 0.90, the more satisfactory it the behavior of the bootstrap.

Table 1 – Empirical coverage of bootstrap confidence intervals at 90% - stationarity.

DGP (ϕ, d, ψ)	Bootstrap	N=100		N=250		N=500	
		\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}
DGP 1, $d=0.25$ (0.3, d , 0.4)	BB	0.648	0.773	0.779	0.756	0.810	0.820
	SB	0.635	0.607	0.788	0.848	0.813	0.884
DGP 2, $d=0.25$ (0.0, d , 0.4)	BB	0.621	0.735	0.746	0.790	0.800	0.842
	SB	0.632	0.687	0.722	0.854	0.823	0.862
DGP 3, $d=0.25$ (0.3, d , 0.0)	BB	0.638	0.766	0.745	0.784	0.815	0.822
	SB	0.612	0.683	0.732	0.871	0.821	0.893
DGP 4, $d=0.25$ (0.0, d , 0.0)	BB	0.786	0.758	0.826	0.789	0.892	0.839
	SB	0.808	0.689	0.851	0.845	0.901	0.883
DGP 1, $d=0.45$ (0.3, d , 0.4)	BB	0.489	0.594	0.530	0.601	0.569	0.619
	SB	0.565	0.615	0.540	0.628	0.581	0.763
DGP 2, $d=0.45$ (0.0, d , 0.4)	BB	0.501	0.565	0.527	0.584	0.551	0.604
	SB	0.544	0.581	0.550	0.799	0.561	0.861
DGP 3, $d=0.45$ (0.3, d , 0.0)	BB	0.523	0.576	0.539	0.598	0.581	0.611
	SB	0.534	0.505	0.547	0.823	0.590	0.857
DGP 4, $d=0.45$ (0.0, d , 0.0)	BB	0.588	0.566	0.632	0.570	0.670	0.606
	SB	0.649	0.532	0.700	0.815	0.748	0.821
DGP 5, short mem. (0.3, 0, 0.4)	BB	0.833		0.843		0.874	
	SB	0.801		0.822		0.855	

Table 1, shows the performance of BB and SB in case of stationary long memory. We see that, as expected, the performance improves with the increase of n . Not surprisingly it tends to worsen when $d=0.45$, as a consequence of the closeness to the nonstationarity region. There is no large difference between the SB and the BB method, however, especially for $d=0.45$ it seems that the SB performs slightly better. As for the type of DGP, the presence of short memory components in DGP 1-3 slightly affects the performance of the method. Indeed, in case of fractional noise, DGP 4, BB and SB perform better, especially if the estimate of d is obtained via Whittle method. Nevertheless, the GPH method is still a very good option, especially in those case where there is no exact knowledge about the short memory components. The short memory DGP (DGP 5) has been included as a benchmark, to show how effectively this methods work in case of weak dependence, i.e. what they have been created for. The performance is on average better, but encouragingly not too far from that of the long memory DGP.

Table 2 and 3 focus on nonstationary versions of DGPs 1-4. In the table the original version of the SB and BB has been coupled with the revised versions (Rev-BB, Rev-SB) proposed in the previous section.

Table 2 – Empirical coverage of bootstrap confidence intervals at 90% - nonstationarity $d=0.65$

DGP (ϕ, d, ψ)	Bootstrap	N=100		N=250		N=500	
		\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}
DGP 1, $d=0.65$ (0.3, $d, 0.4$)	BB	0.499	0.537	0.600	0.622	0.679	0.744
	Rev-BB	0.548	0.638	0.676	0.682	0.700	0.781
	SB	0.500	0.589	0.619	0.711	0.758	0.806
	Rev-SB	0.589	0.658	0.700	0.803	0.850	0.886
DGP 2, $d=0.65$ (0.0, $d, 0.4$)	BB	0.500	0.578	0.600	0.634	0.659	0.754
	Rev-BB	0.623	0.703	0.639	0.689	0.700	0.980
	SB	0.526	0.586	0.623	0.708	0.717	0.808
	Rev-SB	0.684	0.754	0.678	0.779	0.779	0.891
DGP 3, $d=0.65$ (0.3, $d, 0.0$)	BB	0.528	0.588	0.598	0.681	0.676	0.671
	Rev-BB	0.589	0.685	0.739	0.709	0.719	0.773
	SB	0.550	0.595	0.628	0.739	0.709	0.821
	Rev-SB	0.723	0.644	0.747	0.777	0.801	0.847
DGP 4, $d=0.65$ (0.0, $d, 0.0$)	BB	0.686	0.593	0.706	0.601	0.714	0.694
	Rev-BB	0.729	0.637	0.829	0.717	0.841	0.780
	SB	0.734	0.622	0.801	0.770	0.856	0.807
	Rev-SB	0.778	0.648	0.852	0.810	0.890	0.880

Table 3 – Empirical coverage of bootstrap confidence intervals at 90% - nonstationarity $d=0.80$

DGP (ϕ, d, ψ)	Bootstrap	N=100		N=250		N=500	
		\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}	\hat{d}_{wh}	\hat{d}_{gph}
DGP 1, $d=0.80$ (0.3, $d, 0.4$)	BB	0.469	0.515	0.584	0.628	0.688	0.744
	Rev-BB	0.539	0.600	0.581	0.678	0.676	0.774
	SB	0.500	0.523	0.504	0.572	0.605	0.706
	Rev-SB	0.564	0.617	0.629	0.713	0.717	0.805
DGP 2, $d=0.80$ (0.0, $d, 0.4$)	BB	0.477	0.538	0.536	0.604	0.588	0.699
	Rev-BB	0.615	0.674	0.585	0.683	0.677	0.789
	SB	0.518	0.529	0.512	0.573	0.590	0.620
	Rev-SB	0.674	0.701	0.658	0.746	0.731	0.818
DGP 3, $d=0.80$ (0.3, $d, 0.0$)	BB	0.500	0.515	0.533	0.623	0.595	0.683
	Rev-BB	0.564	0.633	0.617	0.705	0.686	0.747
	SB	0.529	0.584	0.544	0.632	0.660	0.700
	Rev-SB	0.660	0.717	0.698	0.761	0.773	0.821
DGP 4, $d=0.80$ (0.0, $d, 0.0$)	BB	0.678	0.569	0.663	0.602	0.729	0.701
	Rev-BB	0.724	0.629	0.729	0.667	0.834	0.786
	SB	0.708	0.607	0.710	0.654	0.800	0.720
	Rev-SB	0.756	0.676	0.784	0.758	0.876	0.835

By reading Table 2 and 3, also in comparison with Table 1, we see that the performance of the BB and SB worsens, the more deeply we enter into the nonstationary areas. However, the revised versions (Rev-BB, Rev-SB) improve with respect to the unrevised ones, as their performance resembles in magnitude order the performance in case of stationary long memory. The improvement is more evident for the SB, that was better performing also in the original version.

Also in this nonstationary case, both Whittle and GPH are very reasonable tool to estimate d .

Future research lines are to consider the same issues with respect to resampling the periodogram ordinates for long memory, also outside the case when the analytical expression of the spectral density function is known.

References

- BERAN, J. 1994. *Statistics for long memory time processes*. New York: Chapman and Hall.
- BICKEL P.J, BÜHLMANN P. 1999. A new mixing notion and functional central limit theorems for a sieve bootstrap in time series, *Bernoulli*, Vol. 5, pp. 413-446.
- BÜHLMANN, P. 1997. Sieve bootstrap for time series, *Bernoulli*, Vol. 3, pp. 123-148.
- CARLSTEIN, E. 1986. The use of subseries methods for estimating the variance of a general statistic for a stationary time series, *Annals of Statistics*, Vol. 14, pp. 1171-1179.
- CHANG Y., PARK J.Y. 2003. A Sieve Bootstrap For The Test Of A Unit Root, *Journal of Time Series Analysis*, Vol. 24, pp. 379-400.
- DAHLHAUS, R. 1989. Efficient parameter estimation for self-similar processes, *The Annals of Statistics*, Vol. 17, pp. 1749-1766.
- DAHLHAUS R., JANAS D. 1996. A frequency domain bootstrap for ratio statistics in time series analysis, *The Annals of Statistics*, Vol. 24, pp. 1934-1963.
- DAVIDOV, Y.A. 1970. The invariance principle for stationary processes, *Theory of Probability & Its Applications*, Vol. 15, pp. 487-498.
- EFRON, B. 1979. Bootstrap methods: another look at the jackknife, *Annals of Statistics*, Vol. 7, pp. 1-26.
- FOX R., TAQQU M.S. 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *The Annals of Statistics*, Vol. 14, pp. 517-532.
- GEWEKE J., PORTER-HUDACK S. 1983. The Estimation and Application of Long-Memory Time Series Models, *Journal of Time Series Analysis*, Vol. 4, pp. 221-237.
- KAPETANIOS G., PSARADAKIS Z. 2006. Sieve Bootstrap for Strongly Dependent Stationary Processes. *Working Papers 552*, Queen Mary University of London, School of Economics and Finance.
- KIM Y.M., NORDMAN D.J. 2011. Properties of block bootstrap under long range dependence, *Sankhya: The Indian Journal of Statistics, series*, Vol. 73, pp. 79-109.

- KREISS, J.P. 1992. Bootstrap procedures for $AR(\infty)$ -processes. In *Lecture Notes in Economics and Mathematical Systems, Bootstrapping and Related Techniques*, Vol. 376, pp. 107-113.
- KÜNSCH, H.R. 1987. Statistical aspects of self-similar processes. In *Proceedings of the 1st World Congress of Bernoulli Society*, Vol. 1, Utrecht: VNU Science Press, pp.67-74.
- KÜNSCH, H.R. 1989. The jackknife and the bootstrap for general stationary observations, *The Annals of Statistics*, Vol. 17, pp. 1217-1241.
- LAHIRI, S.N. 1993. On the moving block bootstrap under long-range dependence, *Statistics and Probability Letters*, Vol. 18, pp. 405-413.
- PALM F.C., SMEEKES S., URBAIN J.P. 2008. Bootstrap Unit Root test: Comparison and Extensions, *Journal of time series analysis*, Vol. 29, pp. 371-400.
- PAPARODITIS, E. 1996. Bootstrapping autoregressive and moving average parameter estimates of infinite order vector autoregressive processes, *Journal of Multivariate Analysis*, Vol. 57, pp. 277-296.
- POSKITT, D.S. 2008. Properties of the Sieve Bootstrap for Fractionally Integrated and Non-Invertible Processes, *Journal of time series analysis*, Vol. 29, pp. 224-250.
- PSARADAKIS, Z. 2001. Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors, *Journal of Time Series Analysis*, Vol. 22, pp-577-594.

SUMMARY

Bootstrap methods for long-range dependence: Monte Carlo evidence

In this paper we present a review of some well-known bootstrap methods for time series data. We concentrate on block bootstrap and sieve bootstrap, whose validity has been proved to be extended to stationary long memory time series.

We will start by reviewing briefly the peculiar features of the bootstrap methods and the issues raised in case of long range dependent data; then we present a Monte Carlo experiment to compare the performance of the methods for a variety of ARFIMA processes. Comments about the finite sample performance of the methods will be provided also in light of the established theoretical properties of the methods

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

VERTICAL SPECIALIZATION ACROSS THE WORLD: EVIDENCE FROM THE WORLD INPUT-OUTPUT TABLE

Giuseppe Ricciardo Lamonica, Luca Salvati, Margherita Carlucci

1. Introduction

The current world economy is characterized by a strong economic interdependence among countries which ties them together, so that the production of a finished good involves the participation of many countries specialised in different stages of production. Indeed, a country imports goods from other countries, uses them as inputs in the production of its own good, which is then exported to the countries specialised in the next stage of production. This sequence continues until the good reaches its final consumers.

In literature this phenomenon known as “vertical specialization”, “slicing the value chain”, “international production sharing” or “outsourcing” is widely used to describe this kind of production process and the resulting trade pattern.

The phenomenon of vertical specialization has been extensively studied in literature and there is an increasing empirical documentation of the rising role of vertical specialization of the country economies. See for example but not limited to: Bridgman (2012), Dean and Lovely (2010), Dixit and Grossman (1982), Feenstra (1998); Feenstra and Hanson (1997), Goh and Henry (2005), Leung (2016), López (2014), Jones and Kierzkowski (2001), Hogan et al. (2005), Sanyal (1983), Yücer et al. (2014), Vecchiu and Makhoul (2014).

There are three standard methods to measure vertical specialization. While the first one uses firm surveys, the second method uses a fine industrial classification of trade, as in Athukorala and Yamashita (2006). The third and most popular method (Hummels et al., 2001) considers national input-output table (NIOT).

The measure proposed by Hummels et al. (2001) and generally used in the empirical analysis is the amount of imported inputs in domestic production embodied in goods that are later exported to other countries, either as a final product or as a good-in-process.

The availability of new databases and in particular multi-country IOT requires a re-examination and adaptation of the original index proposed by Hummels et al. (2001). Indeed, on the one hand the multi-country IOT provides, unlike the NIOT, a more detailed description of the interdependence of industries (economic sectors) among countries. On the other hand, an important part of the exports are endogenous

(i.e. depend on the final demand of other countries) in the multi-country framework, whereas they are considered exogenous when using a national framework. Consequently, the index proposed by Hummels et al. (2001) is no longer sufficient to measure the degree of vertical specialization under multi-country IOT.

This paper, by means of the multi-country IOT derived from the World IOT (WIOT) database, furnishes a detailed analysis of the fragmentation degree of the world's economies applying a new measure, in the spirit of the Hummels et al. (2001) approach, that will be introduced in the next section.

The WIOT (www.wiod.org) is a multi-country IOT and is built in current U.S. dollars with a classification for 35 industries (sectors). The database covers 40 countries plus the Rest of the World and we are able to analyse the phenomena of the vertical specialization in three macro regions, the European Union (i.e. 27 countries that were members in 2011), NAFTA (Canada, Mexico and United States) and East Asia (China, Japan, South Korea and Taiwan) for the period from 1995 to 2011. The main results of this analysis show that the production chain of the countries has become increasingly fragmented since 1995. However, the degree and rates of change of international fragmentation vary considerably across the countries and economic sectors. Moreover, we find evidence that the global financial crises caused a structural break in the time evolution of the considered phenomena.

2. Methodology

Considering the following simplified pattern, in block matrix notation, of a NIOT:

Figure 1 – *Pattern of a NIOT.*

$$\begin{array}{c} \hline \mathbf{Z} \quad \mathbf{f} \quad \mathbf{e} \quad \mathbf{x} \\ \mathbf{M} \\ \mathbf{v}' \\ \mathbf{x}' \\ \hline \end{array}$$

where:

- \mathbf{Z} is a $n \times n$ matrix whose entries (z_{ij}) are the flows for intermediate use from the i -th sector to the j -th sector;
- \mathbf{f} is a $n \times 1$ vector whose entries are the flows from the i -th sector to the final sector;
- \mathbf{e} is a $n \times 1$ vector whose entries are the exportations of the i -th sector;

- \mathbf{M} is a nxn matrix whose entries (m_{ij}) are the imported flows, for intermediate use, from the i -th foreign sector to the j -th domestic sector;
- \mathbf{v}' is $1 \times n$ vector whose entries are the added value of the i -th sector (' is the transposition symbol);
- \mathbf{x} is a $nx1$ vector whose entries (x_j) are the total production (gross output) of the j -th domestic sector.

As proposed by Hummels et al. (2001) a measure of the vertical specialization of an economic sector of a country R is the value of directly imported intermediates embodied in goods that are exported. Formally:

$$DVS_{Rj} = a_{M,j} e_j (\mathbf{i}' \mathbf{e})^{-1} \quad \text{and} \quad a_{M,j} = \sum_{i=1}^k a_{Mij} \quad (1)$$

where, \mathbf{i} is a vector of 1's, $a_{Mij} = m_{ij}/x_j$ is the generic entry of the \mathbf{A}_M matrix of the direct imported coefficients. Thus, a_{Mij} is the total amount of i -th product imported and used as input for the production of one monetary unit of industry j 's output.

For the economic system of a country, vertical specialization is simply the sum of DVS_j across all j . In matrix notation:

$$DVS_R = \mathbf{i}' \mathbf{A}_M \mathbf{e} (\mathbf{i}' \mathbf{e})^{-1} \quad (2)$$

The DVS index is a weighted average of the direct import coefficients using the sectorial exports as weights.

A more detailed index of a sector's vertical specialization is the following (3). Indeed, an intermediate good can be initially imported as input of one domestic sector and then used as an intermediate good in a second domestic sector and so on, until the imported product is finally embodied in a good that is exported:

$$VS_{Rj} = A_{Mj}' \mathbf{L}_j \frac{e_j}{\mathbf{i}' \mathbf{e}} \quad (3)$$

Where A_{Mj}' is the j -th column of the \mathbf{A}_M matrix and \mathbf{L}_j the j -th column of the Leontief inverse matrix $\mathbf{L} = (\mathbf{I} - \mathbf{A}_Z)^{-1}$. This last index includes both the directly and indirectly imported input content in exports of a sector. Indeed, equation (3) gives the total amount of imports that is directly and indirectly required to satisfy one unit (1\$) of the j -th product exported. The vertical specialization of the economic system of a country is:

$$VS_R = \mathbf{i}' \mathbf{A}_M \mathbf{L} \mathbf{e} (\mathbf{i}' \mathbf{e})^{-1} \quad (4)$$

The VS_R index, similarly to the DVS_R , is a weighted average of the import multipliers, with the sectorial exports as weights.

In general, a high value of the VS_R index indicates that imported intermediate goods make up a large proportion of the value of an economic sector (or country) exports and consequently indicate a country's greater degree of involvement in global production chains.

The application of the indices discussed earlier for measuring vertical specialization when a multi-country IOT is used is not completely suitable. In this regard let us consider the simplified pattern, in block matrix notation, of the WIOT depicted in Figure 2, where:

- $Z_{R,R}$ ($R=1,\dots,41$) is a 35×35 matrix whose entries are the flows for intermediate use from the i -th sector of country R to the j -th sector of the same country;
- $Z_{R,K}$ ($R,K=1,\dots,41$ and $R \neq K$) is a 35×35 matrix whose entries are the exports for intermediate use from the i -th sector of country R to the j -th sector of country K ;
- $C_{R,R}$ ($R=1,\dots,41$) is a 35×5 matrix of the domestic final demand in country R ;
- $E_{R,K}$ ($R,K=1,\dots,41$ and $R \neq K$) is a 35×5 matrix of the exports of country R for final demand purposes in country K ;
- x_R ($R=1,\dots,41$) is a 35×1 vector whose entries are the sectorial output of country R ;
- v_R ($R=1,\dots,41$) is a 35×1 vectors whose entries are the sectorial added-value of country R .

Figure 2 – Pattern of a WIOT

$Z_{1,1}$	$Z_{1,2}$...	$Z_{1,40}$	$Z_{1,41}$	$C_{1,1}$	$E_{1,2}$...	$E_{1,41}$	x_1
$Z_{2,1}$	$Z_{2,2}$...	$Z_{2,40}$	$Z_{2,41}$	$E_{2,1}$	$C_{2,2}$...	$E_{2,41}$	x_2
...
$Z_{40,1}$	$Z_{40,2}$...	$Z_{40,40}$	$Z_{40,41}$	$E_{40,1}$...	$C_{40,40}$	$E_{40,41}$	x_{40}
$Z_{41,1}$	$Z_{41,2}$...	$Z_{41,40}$	$Z_{41,41}$	$E_{41,1}$...	$E_{41,40}$	$C_{41,41}$	x_{41}
v'_1	v'_2	...	v'_{40}	v'_{41}					
x'_1	x'_2	...	x'_{40}	x'_{41}					

Legend: 41 is the Rest of World

As it is possible to note, contrary to NIOT where all exports go into the final demand section and are assumed to be exogenous to the economic system of a country, in the WIOT they are subdivided into exports of final goods and services ($E_{R,K}$) and exports of intermediate goods and services ($Z_{R,K}$ for $R \neq K$). While the former are exogenous to the economic system, the latter, on the contrary, depend on

the output of other countries ($\sum_{K=1}^{41} \mathbf{Z}_{R,K} = \sum_{K=1}^{41} \mathbf{A}_{R,K} \mathbf{x}_K$ and $R \neq K$) and are thus endogenous to the economic system.

Consequently, in a multi-country input-output context the $VS_R (VS_{R_i})$ indices do not adequately measure the degree of specialization of a country because they consider all exports exogenous. The following fundamental relation is the starting point to solve this problem:

$$\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{41} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,41} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{41,1} & \cdots & \mathbf{A}_{41,41} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{41} \end{bmatrix} + \begin{bmatrix} \mathbf{c}_1 + \mathbf{e}_1 \\ \vdots \\ \mathbf{c}_{41} + \mathbf{e}_{41} \end{bmatrix} \quad (5)$$

Where: $\mathbf{A}_{R,K} = [Z_{ij,R,K} / x_{j,K}]$; $\mathbf{c}_R = \mathbf{C}_{R,R} \cdot \mathbf{i}$ ($R=1, \dots, 41$); $\mathbf{e}_R = (\sum_{K=1}^{41} \mathbf{E}_{R,K}) \cdot \mathbf{i}$ ($R=1, \dots, 41$).

Alternatively:

$$\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{41} \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \mathbf{A}_{1,1} & \cdots & -\mathbf{A}_{1,41} \\ \vdots & \ddots & \vdots \\ -\mathbf{A}_{41,1} & \cdots & \mathbf{I} - \mathbf{A}_{41,41} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c}_1 + \mathbf{e}_1 \\ \vdots \\ \mathbf{c}_{41} + \mathbf{e}_{41} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{1,1} & \cdots & \mathbf{L}_{1,41} \\ \vdots & \ddots & \vdots \\ \mathbf{L}_{41,1} & \cdots & \mathbf{L}_{41,41} \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 + \mathbf{e}_1 \\ \vdots \\ \mathbf{c}_{41} + \mathbf{e}_{41} \end{bmatrix} \quad (6)$$

To determine the import content of exports of a generic country R , let us first consider the matrix of the total imports ($\mathbf{M}_{\bullet R}$):

$$\mathbf{M}_{\bullet R} = \sum_{K=1}^{41} \mathbf{Z}_{K,R} = (\sum_{K=1}^{41} \mathbf{A}_{K,R}) \mathbf{D}(\mathbf{x}_R) = \mathbf{A}_{\bullet R} \mathbf{D}(\mathbf{x}_R) \quad \text{for } R \neq K \quad (7)$$

Where $\mathbf{D}(\mathbf{x}_R)$ is the diagonal matrix whose entries, $d_{jj}(\mathbf{x}_R) = x_{jR}$, correspond to the total production of economic sectors in country R . Post-multiplying equation (7) by \mathbf{i} , the vector of the total sectorial imports (\mathbf{m}_R) is obtained:

$$\mathbf{M}_{\bullet R} \mathbf{i} = \mathbf{m}_R = \mathbf{A}_{\bullet R} \mathbf{D}(\mathbf{x}_R) \mathbf{i} = \mathbf{A}_{\bullet R} \mathbf{x}_R \quad (8)$$

Since (see Figure 2):

$$\mathbf{x}_R = \mathbf{c}_R + \mathbf{e}_R + \mathbf{Z}_{R,R} \mathbf{i} + \sum_{K=1}^{41} \mathbf{Z}_{R,K} \mathbf{i} \quad (9)$$

Then substituting (9) in (8) and in consideration that $\mathbf{Z}_{R,K} = \mathbf{A}_{R,K} \mathbf{x}_K$ for $R \neq K$, we have:

$$\mathbf{m}_R = \mathbf{A}_{\bullet R} \left(\mathbf{c}_R + \mathbf{e}_R + \mathbf{Z}_{R,R} \mathbf{i} + \sum_{\substack{K=1 \\ K \neq R}}^{41} \mathbf{A}_{R,K} \mathbf{x}_K \mathbf{i} \right) \quad (10)$$

By means of equation (6) the vector of the sectorial production (\mathbf{x}_K) of country K can be rewritten in terms of production generated by the final demands of the other countries:

$$\mathbf{x}_K = \sum_{j=1}^{41} \mathbf{L}_{K,j} (\mathbf{c}_j + \mathbf{e}_j) \quad (11)$$

Finally, substituting the equation (11) into (10), we have:

$$\mathbf{m}_R = \mathbf{A}_{\bullet R} \left[\mathbf{c}_R + \mathbf{e}_R + \mathbf{Z}_{R,R} \mathbf{i} + \sum_{K=1}^{41} \mathbf{A}_{R,K} \left(\sum_{j=1}^{41} \mathbf{L}_{K,j} (\mathbf{c}_j + \mathbf{e}_j) \right) \mathbf{i} \right] \quad \text{for } R \neq K \quad (12)$$

At this point, consider the following quantities:

–The *exogenous exports* of country R (i.e. final exports): the j-th entry of this vector gives the exogenous exports of the j-th economic sector: $\mathbf{exp}_1 = \mathbf{e}_R$;

–The *imports content of the exogenous export* of country R: from equation (11), the vector $\mathbf{L}_{R,R} \mathbf{e}_R$, gives the domestic output necessary for the exogenous exports. Consequently, the generic entry of following vector gives the imports of sector j embodied directly and indirectly in the exogenous exports of the same sector: $\mathbf{imp}_1 = \mathbf{A}_{\bullet R} \mathbf{L}_{R,R} \mathbf{e}_R$;

–The *endogenous exports* of country R (i.e. exports that are used as inputs in the production process of the other countries): the j-th entry of the following vector gives the endogenous exports of the j-th economic sector: $\mathbf{exp}_2 = \sum_{K=1}^{41} \mathbf{Z}_{R,K} \mathbf{i} = \sum_{K=1}^{41} \mathbf{A}_{R,K} \mathbf{x}_K = \sum_{K=1}^{41} \mathbf{A}_{R,K} \left(\sum_{j=1}^{41} \mathbf{L}_{K,j} (\mathbf{c}_j + \mathbf{e}_j) \right)$ for $R \neq K$;

–The *imports content of endogenous exports* of country R: this quantity is given by the vector $\mathbf{imp}_2 = \mathbf{A}_{\bullet R} \sum_{K=1}^{41} \mathbf{A}_{R,K} \left(\sum_{j=1}^{41} \mathbf{L}_{K,j} (\mathbf{c}_j + \mathbf{e}_j) \right)$ for $R \neq K$ whose entries give the imports of sector j embodied directly and indirectly in the endogenous exports of the same sector.

Consequently, the following index is adopted to measure the degree of vertical specialization of the j-th economic sector:

$$VS1_{Rj} = \frac{\mathbf{imp}_{1j} + \mathbf{imp}_{2j}}{\mathbf{exp}_{1j} + \mathbf{exp}_{2j}} \quad (13)$$

Accordingly, the overall degree of vertical specialization at a national level is given by:

$$VS1_R = \frac{i'(\text{imp}_1 + \text{imp}_2)}{i'(\text{exp}_1 + \text{exp}_2)} = \frac{i'(\mathbf{A}_{\bullet R} \mathbf{L}_{R,R} \mathbf{e}_R + \mathbf{A}_{\bullet R} \sum_{K=1}^{41} \mathbf{A}_{R,K} (\sum_{j=1}^{41} \mathbf{L}_{K,j} (c_j + e_j)))}{i'(\mathbf{e}_R + \sum_{K=1}^{41} \mathbf{A}_{R,K} (\sum_{j=1}^{41} \mathbf{L}_{K,j} (c_j + e_j)))} \text{ for } R \neq K \quad (14)$$

If only NIOT is available, the endogenous exports, $\sum_{K=1}^{41} \mathbf{Z}_{R,K}$ (for $R \neq K$), belong to the final exportations of the country R and are considered exogenous. Thus, equations (11) and (12) become respectively:

$$\mathbf{x}_R = \mathbf{L}_{R,R} (\mathbf{c}_R + \mathbf{e}_R + \sum_{K=1}^{41} \mathbf{Z}_{R,K}) \text{ for } R \neq K \quad (15)$$

and:

$$\mathbf{m}_R = \mathbf{A}_{\bullet R} [\mathbf{L}_{R,R} (\mathbf{c}_R + \mathbf{e}_R + \sum_{K=1}^{41} \mathbf{Z}_{R,K})] \text{ for } R \neq K \quad (16)$$

Consequently, the $VS1_R$ index becomes:

$$\overline{VS1}_R = \frac{i'(\mathbf{A}_{\bullet R} \mathbf{L}_{R,R} (\mathbf{e}_R + \sum_{K=1}^{41} \mathbf{Z}_{R,K}))}{i'(\mathbf{e}_R + \sum_{K=1}^{41} \mathbf{Z}_{R,K})} \quad (17)$$

that is the original index proposed by Hummels et al. (2001).

3. Results and discussion

In this section, the vertical specialization of the world economy between 1995 and 2011 is examined by means of the $VS1_R$ index.

Table 1 depicts the values of the $VS1_R$ index by countries and for same years. The detailed values are available from the authors upon request.

Analysis of this index over time highlights that vertical specialization increases in all countries with two exceptions: (i) Bulgaria, Cyprus and Indonesia, where the values remain almost constant over time, and (ii) Canada, Estonia, Malta, Portugal and Russia, where the vertical specialization decreases.

Japan, India and Poland realized a per-year growth rate by 7.5%, 6% and 5.2% respectively. These are followed by Brazil (4.7%), Hungary (4.2%) and Czech Republic (3.9%). The remaining countries register an annual growth rate ranging between 0.1% (Cyprus) and 3.3% (Korea). Moreover, at the beginning of the financial crisis in 2009 all countries experienced a turnaround estimated about 0.04\$. In particular, with the exception of Ireland, Luxemburg and Mexico, where the index

value in the 2009 is identical to the one of 2008, the rest of the countries show an average decrease amounting to 5.5%. Bulgaria (-29%), India (-6%) and Taiwan (-6%) realized the highest decrease in the sample.

Table 1 – $VS1_R$ index by countries and same years.

	1995	2000	2005	2010	2011		1995	2000	2005	2010	2011
AUS	0.08	0.09	0.10	0.09	0.10	IRL	0.35	0.41	0.37	0.37	0.38
AUT	0.20	0.24	0.27	0.27	0.29	ITA	0.15	0.17	0.18	0.21	0.22
BEL	0.34	0.36	0.36	0.38	0.41	JPN	0.04	0.06	0.08	0.10	0.11
BGR	0.25	0.29	0.25	0.25	0.25	KOR	0.17	0.22	0.23	0.25	0.27
BRA	0.05	0.09	0.08	0.07	0.09	LTU	0.26	0.29	0.32	0.30	0.30
CAN	0.22	0.24	0.20	0.16	0.16	LUX	0.39	0.47	0.47	0.50	0.50
CHN	0.12	0.13	0.20	0.15	0.15	LVA	0.20	0.20	0.24	0.20	0.19
CYP	0.24	0.28	0.20	0.23	0.23	MEX	0.23	0.28	0.27	0.28	0.28
CZE	0.22	0.30	0.35	0.38	0.39	MLT	0.40	0.44	0.35	0.33	0.33
DEU	0.15	0.20	0.22	0.23	0.24	NLD	0.28	0.31	0.30	0.34	0.35
DNK	0.24	0.27	0.29	0.32	0.33	POL	0.13	0.21	0.25	0.27	0.28
ESP	0.17	0.23	0.21	0.22	0.24	PRT	0.24	0.25	0.26	0.22	0.22
EST	0.29	0.35	0.31	0.26	0.26	ROM	0.16	0.21	0.23	0.19	0.19
FIN	0.18	0.21	0.24	0.25	0.26	RUS	0.05	0.07	0.05	0.03	0.04
FRA	0.16	0.21	0.21	0.22	0.24	SVK	0.23	0.35	0.40	0.36	0.36
GBR	0.16	0.16	0.15	0.18	0.19	SVN	0.28	0.31	0.35	0.31	0.31
GRC	0.15	0.26	0.22	0.21	0.22	SWE	0.21	0.25	0.25	0.26	0.26
HUN	0.22	0.42	0.39	0.40	0.40	TUR	0.12	0.16	0.22	0.16	0.18
IDN	0.12	0.16	0.15	0.11	0.12	TWN	0.27	0.30	0.35	0.36	0.37
IND	0.08	0.11	0.16	0.18	0.19	USA	0.08	0.09	0.10	0.11	0.12

The mean value of the $VS1_R$ index varies across countries from 0.05\$ (Russia) to 0.47\$ (Luxemburg). Accordingly, countries were classified into four groups based on the mean level of the index:

- 1) Countries with low value (i.e. $VS1_R < 0.10$): Australia, Brazil, Japan, Russia, and USA;
- 2) Countries with mean-low vertical specialization (i.e. $0.10 < VS1_R \leq 0.20$): China, Germany, France, Great Britain, India, Indonesia, Italy, and Turkey;
- 3) Countries with mean-high (i.e. $0.20 < VS1_R \leq 0.30$): Austria, Bulgaria, Canada, Cyprus, Denmark, Estonia, Finland, Greece, Korea, Lithuania, Latvia, Mexico, Holland, Poland, Portugal, Romania and Sweden;
- 4) Countries with high values (i.e. $0.30 < VS1_R \leq 0.53$): Belgium, Czech Republic, Estonia Hungary, Ireland, Luxembourg, Malta, Slovenia, Slovakia and Taiwan.

Table 2 reports descriptive statistics of the $VS1_R$ index by year. An average degree of outsourcing by 0.20\$ was observed in 1995 and one half of the units

presented a value not exceeding 0.20\$. The lowest value is recorded for Japan (0.05\$) and the highest value for Malta (0.40\$), as illustrated in Table A5.

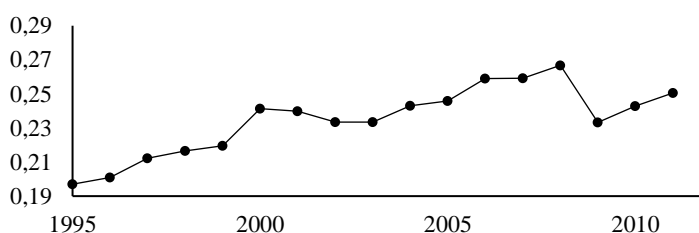
The mean value increased to 0.25\$ in 2011, when Luxemburg was the country with the highest value (0.50\$), while Russia had the lowest value (0.04\$). The mean value of the index increased by 0.05\$ during 1995-2011 (Figure 3).

The first quartile grows from 0.15\$ to 0.19\$. The median changes from 0.20\$ to 0.25\$ and the third quartile from 0.24\$ to 0.32\$. This confirms that the degree of specialization of the world economy has grown markedly during 1995-2011.

Table 2 – Descriptive statistics of the VSI_R index by year.

Year	Mean	Std. Dev.	1° Quart.	Median	3° Quart.	Min	Max
1995	0.20	0.09	0.15	0.20	0.24	0.04	0.40
1996	0.20	0.09	0.14	0.21	0.26	0.04	0.40
1997	0.21	0.09	0.15	0.22	0.28	0.05	0.44
1998	0.22	0.09	0.17	0.23	0.28	0.06	0.45
1999	0.22	0.09	0.15	0.22	0.28	0.05	0.45
2000	0.24	0.10	0.17	0.24	0.30	0.06	0.47
2001	0.24	0.10	0.18	0.24	0.30	0.06	0.47
2002	0.23	0.10	0.18	0.24	0.28	0.06	0.48
2003	0.23	0.09	0.18	0.24	0.30	0.06	0.46
2004	0.24	0.10	0.20	0.24	0.30	0.05	0.48
2005	0.25	0.09	0.20	0.24	0.31	0.05	0.47
2006	0.26	0.10	0.21	0.26	0.33	0.05	0.48
2007	0.26	0.10	0.20	0.26	0.34	0.04	0.48
2008	0.27	0.10	0.21	0.27	0.35	0.04	0.50
2009	0.23	0.10	0.18	0.23	0.30	0.03	0.50
2010	0.24	0.10	0.18	0.24	0.31	0.03	0.50
2011	0.25	0.10	0.19	0.25	0.32	0.04	0.50

Figura 3 – Time series of the mean value of VSI_R index between 1995 and 2011.



On average, small countries, especially in Europe, realized the highest degree of vertical specialization. It is not surprising because smaller countries are typically

more dependent on imported inputs for their production than larger countries, which are more self-supporting.

4. Conclusions

With the global economic systems becoming more integrated, production processes, previously carried out in one site, are now split up into stages (or fragments), each to be performed in a different location and often beyond national boundaries.

These changes have been extensively studied under different names, such as “disintegration of production”, “vertical specialization”, “fragmentation”, “outsourcing”, “offshoring”, among others. Thus, one important feature of the new international and globalized economy is the significant increase of intermediate goods trade, crossing several borders along the supply chains.

At the same time, fragmentation of production processes plays a role in societal well-being, improving living conditions and wealth of resident population.

Using the measure proposed by Hummels *et al.* (2001) and modified to be used in a multi-country IOT, this study has analysed the degree and evolution of vertical specialization of the world economy in the period 1995-2011.

Our main data source has been the World Input Output Table, which covers 40 countries with a classification for 35 economic sectors.

Empirical results confirm the growth of foreign intermediate goods content of exports in all economies. The positive trend declined slightly at the beginning of the financial crisis.

The degree of specialization is greatly diversified across countries. In this regard, four distinct groups of countries were identified.

Moreover, the analysis identified a negative correlation between the vertical specialization index and the size of country’s economies: small countries have the highest degree of specialization and are concentrated in Europe.

References

- ATHUKORALA P.C., YAMASHITA N. 2006. Production fragmentation and trade integration: East Asia in a global context, *North American Journal of Economics and Finance*, Vol.17, No. 3, pp. 233-256.
- BRIDGMAN, B. 2012. The rise of vertical specialization trade. *Journal of International Economics*, Vol. 86, pp. 133–140.

- DEAN J.M., FUNG K.C., WANG Z. 2011. Measuring vertical specialization: The case of China, *Review of International Economics*, Vol.19, No. 4, pp. 609–625.
- DEAN J.M., LOVELY M.E. 2010. Trade growth, production fragmentation, and China's environment. In Feenstra R.C. and Shang-jin W. (Eds.) *China's Growing Role in World Trade*, The University of Chicago Press, pp. 429-469.
- DIXIT A.K., GROSSMAN G.M. 1982. Trade and protection with multistage production, *Review of Economic Studies*, Vol. 49, pp. 583–594.
- FEENSTRA, R.C. 1998. Integration of trade and disintegration of production in the global economy, *Journal of Economic Perspective*, Vol. 12, No.4, pp. 31-50.
- FEENSTRA R.C., HANSON G.H. 1997. Foreign direct investment and relative wages: evidence from Mexico's maquiladoras, *Journal of International Economics*, Vol. 42, pp. 371–394.
- GOH A.T., HENRY Y.W. 2005. Fragmentation, Engel's law, and learning, *Review of International Economics*, Vol. 13, pp. 518–528.
- HOGAN C., KONDRATOWICZ M., YI K.M. 2005. Vertical specialization and three facts about U.S. international trade, *North American Journal of Economics and Finance*, Vol. 16, No. 1, pp. 35–59.
- HUMMELS D., ISHII J., YI K.M. 2001. The nature and growth of vertical specialization in world trade, *Journal of International Economics*, Vol. 54, No. 1, pp. 75–96.
- LEUNG, J.Y. 2016. Bilateral vertical specialization between the U.S. and its trade partners - before and after the free trade agreements, *International Review of Economics and Finance*, Vol. 45, pp. 177-196.
- LÓPEZ, A. 2014. Outsourcing and firm productivity: a production function approach, *Empirical Economics*, Vol. 47, pp. 977–998.
- JONES R.W., KIERZKOWSKI H., CHEN L. 2005. What does evidence tell us about fragmentation and outsourcing? *International Review of Economics & Finance*, Vol. 14, pp. 305–316.
- SANYAL, K.K. 1983. Vertical specialization in a Ricardian model with a continuum of stages of production, *Economica*, Vol. 50, pp. 71–78.
- YÜCER A., GUILHOTO J., SIROËN J.M. 2014. Internal and international vertical specialization of brazilian states. An input-output analysis, *Revue d'Economie Politique*, Vol. 124, No. 4, pp. 598-624.
- VECHIU N., MAKHLOUF F. 2014. Economic integration and specialization in production in the EU27: does FDI influence countries' specialization? *Empirical Economics*, Vol.46, pp. 543–572.

SUMMARY

Vertical specialization across the world: evidence from the world input-output table

Fragmentation of vertically integrated production processes (known also as vertical specialization) implies that production blocks are located in different countries and geographical areas across the world countries. This study extends the quantitative measure of vertical specialization, originally proposed by Hummels et al. (2001), into a multi-country framework. Based on this approach, vertical specialization in the world economy was studied between 1995 and 2011. Empirical evidence supports a significant increase in the vertical specialization of the world economy during the study period.

Giuseppe RICCIARDO LAMONICA, Università Politecnica delle Marche,
g.ricciardo@staff.univpm.it

Luca SALVATI, Council for Agricultural Research and Economics (CREA),
luca.salvati@uniroma1.it

Margherita CARLUCCI, Sapienza Università di Roma,
margherita.carlucci@uniroma1.it

THE NEW METHODOLOGY FOR THE STS INDICATOR ON THE “NUMBER OF PERSONS EMPLOYED”

Marco Lattanzio

1. Introduction

Starting from May 2018 Istat produces the quarterly indicator on the Number of Persons Employed required by the Eurostat-STS regulation (Reg. EU n.1165/98) according to a new compilation methodology. The regulation concerns the production of harmonized short-term statistics on businesses and this is the unique indicator on employment on the business side. The first innovation of this paper consists in the new exploitation of a short-term quarterly variable, the number of self-employees calculated in the context of the Labour Force Survey (LFS), used as indicator in the temporal disaggregation of the annual benchmark, represented by the number of self-employees from National Accounts (NA). This latter source of information better matches the definition requirements. Furthermore, the second aspect of the innovation of this paper is the new method for quarterly temporal and sectorial disaggregation of the NA source used. The LFS source is mainly used to give information about the short-term dynamic that with this new method is considerably improved, considering also that, as it will be described, the previous temporal disaggregation method used as indicator the time series of the employees, that is a source very poor in order to give information about the self-employee's short-term dynamic. Nevertheless, improvements are balanced with the unavoidable increasing of the variability in the series, consequence of the use of sample estimates for the self-employee component.

The document is organized as follows: paragraph 2 recaps some information about the objective of the indicator, the regulation framework and the STS requirements. Paragraph 3 describes briefly the source of information and the methodology used for estimating of the employee component. Paragraph 4 highlights on the methodological differences between the new and the old sources and new and old methods. Paragraph 5 shows some graphical comparisons between old and new time series, highlighting the improvements reached by the new method and reflecting what expected from the methodological point of view. Finally, paragraph 6 gives conclusions, describes critical aspects and suggests enhancements.

2. Requirements and definitions for STS variable 210 - Number of persons employed.

As written on the Eurostat web site, “...*STS indicators are tools for formulating and monitoring the economic and monetary policy of the European Union and the euro area... STS data are in great demand for economic analysis in the European Commission and European Central Bank (ECB), national governments and central banks, companies and financial market*”¹. The importance of these indicators at European level has encouraged Istat to improve the methodology and remove the confidentiality status, still characterizing the delivery of the Number of persons employed indicator. Until February 2019, even if the Italian indicator was used by Eurostat to calculate the European countries’ aggregated indicators for all the levels of detail of the economic activity classification required, it was not published. The indicator is now available in the EUROIND database². The STS regulation defines the statistical requirements of the variable 210: Number of persons employed, not only in terms of coverage of the population observed but also in terms of levels of detail by economic activity classification, timeliness and time interval of the series. The classifications used are the Italian version of the European Nace rev.2 nomenclature, that is the Ateco 2007 classification³, and the MIGs (Main Industrial Groupings) classification. In particular indicators are organized in four annexes (Industry without construction, Construction, Retail Trade, Services) and for each one different level of details are required, broken down until 2-digit Nace level. The target population are the persons employed, made by employees and self-employees, for which a regular contract exists between worker and employer. Fixed base indices are produced, on a quarterly basis, starting from the 1st quarter of 2000, and released 60 days after the end of the reference period. For each release, also revised data on the previous quarters are given following a defined revision policy.

3. Employees estimates: source and methodology

The employee component of the indicator is approximated by the Number of jobs produced in the context of the Oros⁴ survey, a process where administrative data

¹ <https://ec.europa.eu/eurostat/web/short-term-business-statistics/overview/sts-in-brief>

² at the link <https://ec.europa.eu/eurostat/web/euro-indicators/> in the *industry, trade and services/labour input* section

³ https://www.istat.it/it/files//2011/03/metenorme09_40classificazione_attivita_economiche_2007.pdf

⁴ Oros is the acronym for Occupazione, Retribuzioni e Oneri Sociali. For more information about the survey see Istat 2019, and for what concerns the number of jobs’ estimation see the work LATTANZIO et al 2019

related to the population of Italian firms with employees, acquired by the Italian National Social Security Institute (Inps), are integrated with survey data related to the Large Enterprises subpopulation (500+ employees) that come from an ad-hoc survey. The Oros process has also in charge the compilation of the STS indicator on the total wages and salaries.

4. Self-Employees estimates: old and new sources and methodology

4.1. The self-employee annual source from the National Accounts

The main source used to estimate the self-employee component is the annual number of self-employees calculated in the context of the National Accounts (NA). An ad-hoc elaboration prepared by NA experts is used in order to extract only the regular component of the self-employment, as required. These data, calculated only for the STS purposes, are used as a benchmark, as in the past, since they represent the best match with the statistical and timeliness STS requirements⁵. Actually, the main source used in NA for the calculation of the number of self-employees is the Business Register (BR), from which the Nace classification is also derived. In particular, say a the year of the release, the BR is used to estimate year $a-2$, while data for $a-1$ are estimated using other sources, as LFS data and administrative data from Inps, used also in order to integrate little typologies not covered by the BR. Data are available for the 2-digit level of detail of the Nace rev.2 classification, for the B to N sectors. At this level of detail, however, only annual data are available, so a procedure of quarterly disaggregation has to be performed.

4.2. The old temporal and sectorial disaggregation method

In the previous method used to estimate the number of self-employees, the sectorial disaggregation was not necessary and only a quarterly temporal disaggregation of the annual time series was performed, since data on self-employees were calculated on a 2-digit Nace level base. Until the delivery of February 2018, the procedure for the temporal disaggregation applied was very simple, aimed at preserving the quarterly dynamic of the indicator series chosen, that was the number of jobs. Let NPE denote the total number of employees, E denote the number of employees and SE denote the number of self-employees. Say $SE_{j,a}^{NA}$ the NA annual estimates of self-employees related to year a and sector j , say E_{j,t_a} the Oros quarterly estimates of Jobs related to the quarter t_a of year a and sector j . The quarterly estimates of the self-employees, \hat{SE}_{j,t_a} , was obtained in the following way:

⁵ Data on regular self-employment are available also from the Business Register but with a higher time delay.

$$\hat{S}E_{j,t_a} = SE_{j,a}^{NA} \frac{E_{j,t_a}}{\frac{\sum_{t_a=1}^4 E_{j,t_a}}{4}} = \alpha_{j,a} \cdot E_{j,t_a} \cdot SE_{j,a}^{NA}$$

where $\alpha_{j,a} = \frac{4}{\sum_{t_a=1}^4 E_{j,t_a}}$. Finally, estimates on the number of persons employed

$N\hat{P}E_{j,t}$ were obtained simply adding the number of jobs to the estimate of the number of self-employees. The most critical aspect of this method is that the quarterly dynamic of self-employees, and also that of the total number of persons employed, were strongly affected by that of the employees. This fact can be easily proved calculating analytically the dependence between the number of persons employed year-on-year growth rate ($V(NPE)_{t,t-4}$) and quarter-on-quarter growth rate ($V(NPE)_{t,t-1}$) and the self-employees and employees growth rate ($V(E)$). This can be done using the same analytical tools used in LATTANZIO 2019⁶:

$$V(NPE)_{t,t-1} = V(E)_{t,t-1} \quad (\text{for } t = Q2, Q3, Q4) \quad (1)$$

$$V(NPE)_{t,t-4} \approx V(E)_{t,t-4} + V(1 + \alpha_a \cdot SE_a^{NA})_{a,a-1} \quad (2)$$

Considering the results of the decomposition formula, the dynamic of the employee part on the total number of persons employed is of a great impact. This was the first reason that induced Istat to study a new method and exploit a new source in order to acquire information about the dynamic of the self-employee component.

4.3. The new additional source for the temporal disaggregation

Data related to the quarterly number of self-employees are calculated in the context of the Labour Force Survey (LFS), according to a two-stage sampling (first municipalities, then households) survey conducted on households. The sample is rotated every week and the calibration estimator is used to calculate quarterly estimates, available at 2-digit Nace level⁷. The main difference with the NA self-employees is that LFS register *heads* and not *jobs*. Furthermore, it is important to say that the Nace code is communicated by the person interviewed and translated by the interviewer, so an acceptable quality of this variable is achieved only at a high

⁶For the sake of simplicity the proof of these results is not presented but can be found in the reference paper.

⁷ For more information about the survey and this source see Istat 2006.

⁸ Note that, although LFS data are available at 2-digit NACE level, they were not used to estimate directly the quarterly number of self-employees due to the issues described in paragraph 4.3 and the difference with respect to the NA source.

aggregated level. Last, but not for importance, LFS includes also the non-regular employment.

4.4. The new temporal and sectorial disaggregation method

The exploitation of the time series of the Number of self-employees calculated in the context of the LFS survey, allows to consider information about the quarterly dynamic of this population. The new methodology used to estimate the self-employee part of the indicator can be summarized in two steps:

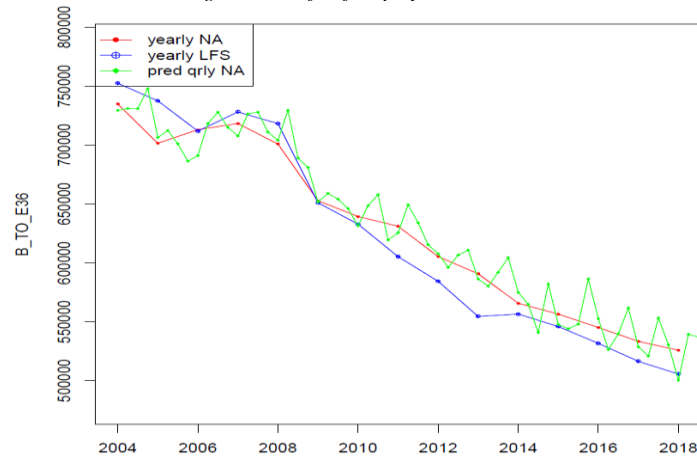
- 1st step - Temporal disaggregation

The first step is the temporal disaggregation of the NA annual time series of the Number of self-employees. For this purpose the Chow-Lin regression method (see the Appendix) is applied using the generalized R-package for temporal disaggregation “tempdisagg”⁹. The methodology, ordinarily used also at Istat in the context of the Quarterly National Accounts, is performed on data aggregated at a macro Nace level of detail, in order to ensure good estimates of the parameters and for which a linear relation can be resumed between the annual time series compared (NA and annualized LFS). More methodological details on the method are presented in the appendix of this work. The method is applied on data aggregated in 5 macro sectors. For the sake of simplicity, results regarding only the macro sector Industry without construction (indicated with the label B_TO_E36 in the graphics) are presented, showing a graphical comparison between the annual series and the predicted one (Figure 1), and output statistics obtained with the R package (Figure 2).

⁹ <https://cran.r-project.org/web/packages/tempdisagg/tempdisagg.pdf>.

Figure 1 – Comparison between NA annual number of self-employees (red line), LFS annualized number of self-employees (blue line) and the results of the quarterly disaggregation (green line) for the macro sector Industry without construction (B_TO_E36).

Period 2004-2018, average number of self-employees-



Source: Istat – Elaboration of LFS survey data and NA data.

Figure 2 – Output statistics of the R “tempdisagg” package for the application of the Chow-Lin regression method - sector B_TO_E36

```
fit statistics ~ se.lfs aggregate B_TO_E36
Call:
td(formula = NA[, k] ~ 1 + LFS[, k], conversion = average)

Residuals:
    Min       1Q   Median       3Q      Max
-13306.7  -7216.8  -654.9   8653.1  15632.0

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.230e+05  2.700e+04  4.558 0.000658 ***
LFS[, k]    8.173e-01  4.272e-02  19.133 2.34e-10 ***

'chow-lin-maxlog' disaggregation with 'average' conversion
14 low-freq. obs. converted to 57 high-freq. obs.
Adjusted R-squared: 0.9656  AR1-Parameter: 0.7081
```

Source: Istat – Elaboration of LFS survey data and NA data.

In the first figure a high correlation between the annual series of the LFS and NA number of self-employees can be seen, same results are obtained from the output statistics looking at the goodness of fit for the parameters of the regression, the high adjusted R-squared and the estimate of the autoregressive parameter.

- 2nd step - Sectorial disaggregation

The second step consists in splitting the estimates of the number of self-employees from the macro aggregation to the 2-digit Nace level detail, as required by the regulation. A procedure of sectorial disaggregation is then performed, that can be formulated as follows. Say $SE_{j,t}^{LFS}$ the LFS number of self-employees observed at quarter t in the Nace division j , $SE_{S,t}^{LFS}$ the same number for the macro sector S , then the final quarterly estimate of the number of self-employees, $\hat{SE}_{j,t}$, is obtained as:

$$\hat{SE}_{j,t} = \hat{SE}_{S,t} \frac{SE_{j,t}^{LFS}}{SE_{S,t}^{LFS}} \quad (3)$$

This method is aimed to preserve the LFS quarterly sectorial structure, represented by the ratio estimate $\frac{SE_{j,t}^{LFS}}{SE_{S,t}^{LFS}}$, that is applied to the quarterly estimate derived from the regression model, related to the macro sector S . It worth to note that the LFS time series on self-employees were available only starting from the 1st quarter of 2004. For the time interval ranging from the 1st quarter of 2000 to the 4th quarter of 2003 only data on the total number of persons employed and at a higher aggregated level with respect to the 2-digit Nace detail were available from LFS. For this reasons a different disaggregation level of detail was used for the temporal disaggregation and a process of reconciliation with the time series ranging from the 1st quarter of 2004 to the 4th quarter of 2017 applied. Apart from the reconciliation procedure, the temporal and sectorial disaggregation method used for the first time window reflect entirely that used for the other part of the time series and described here. In the following paragraph final results on the number of persons employed will be presented, in comparison with the results of the old method.

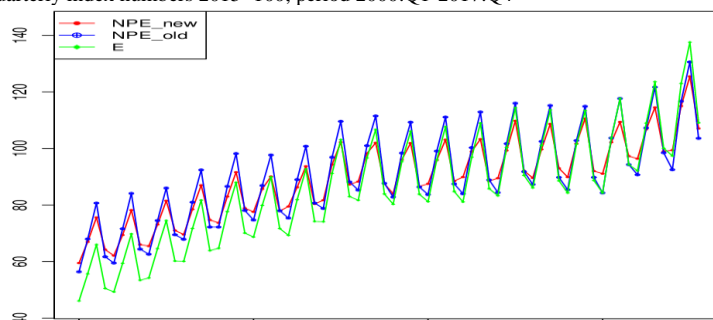
5. Results comparison

In order to show the improvement reached in the short-term dynamic of the indicator, indices and growth rates of the number of persons employed calculated in new (red line) and old (blue line) method were compared with that calculated on the number of employees (green line). In particular in figures 3 and 4 the series for the section I are compared, sector for which the self-employees represent a share of almost 38% on the total number of persons employed, in average on the entire time series, and for which can be observed a strong seasonal pattern for the number of employees. Looking at figure 4, apart from all the first quarters, the pattern is entirely reflected by the total number of persons employed calculated according to the old

method, consequence described in the quarter-on-quarter growth rate formulation in (1). Different is instead the dynamic for the series calculated with the new methodology.

Figure 3 – New (*NPE_new*) and old (*NPE_old*) number of persons employed and number of employees (*E*) in the Nace section I

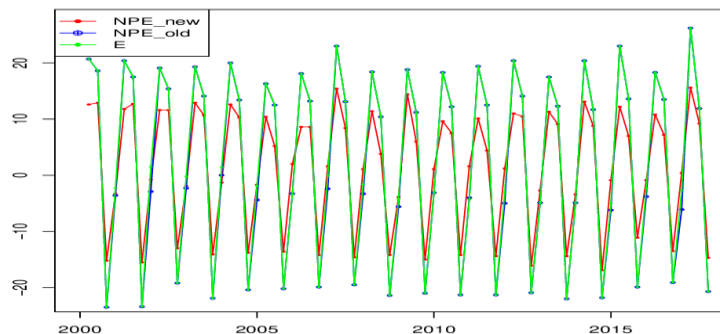
Quarterly index numbers 2015=100, period 2000:Q1-2017:Q4



Source: Istat – Elaboration of Oros and LFS survey data and NA data.

Figure 4 – New (*NPE_new*) and old (*NPE_old*) number of persons employed and number of employees (*E*) in the Nace section I.

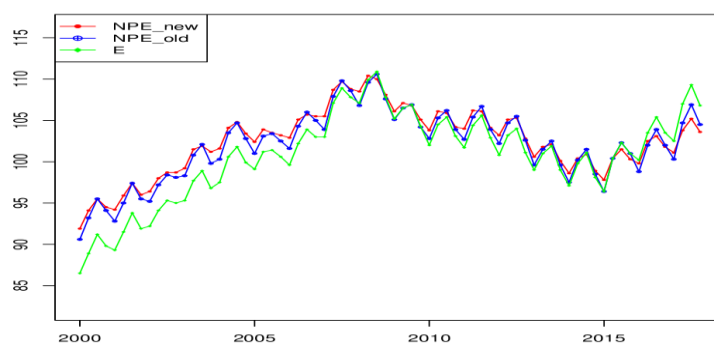
q-o-q growth rates, period 2001:Q1-2017:Q4.



Source: Istat – Elaboration of Oros and LFS survey data and NA data.

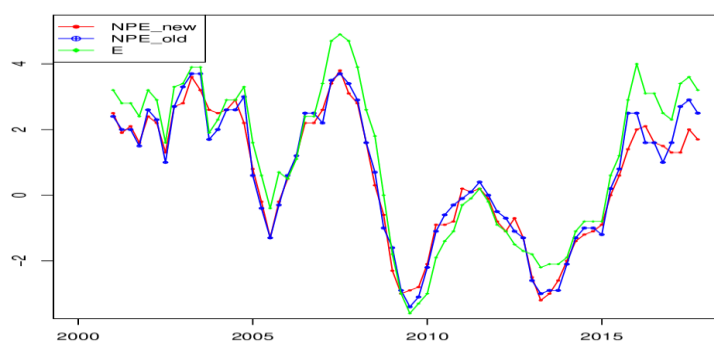
In figure 5 and 6 are shown series of the indices and the related year-on-year growth rates for the total industry and market services economy (indicated with the label B_TO_N in the graphics), for which the share of the self-employees is about 34% on the total number of persons employed, in order to show the improvement reached.

Figure 5 – New (NPE_new) and old (NPE_old) number of persons employed and number of employees (E) in the total economy B_TO_N for STS. Quarterly index numbers 2015=100, period 2000:Q1-2017:Q4.



Source: Istat – Elaboration of Oros and LFS survey data and NA data.

Figure 6 – New (NPE_new) and old (NPE_old) number of persons employed and number of employees (E) in the total economy B_TO_N for STS. y-o-y growth rates, period 2001:Q1-2017:Q4.



Source: Istat – Elaboration of Oros and LFS survey data and NA data.

6. Conclusions

The methodological improvement in the estimation of the number of persons employed has effects not only in terms of accuracy of the indicator produced but also in the use of a more statistically rigorous approach. Furthermore, it represents a case of exploitation of different administrative and statistical sources and an attempt of their integration in order to obtain new statistical information. An accurate analysis of the revisions implied by the new methodology and by the use of sources subjected

to revisions, such as NA data and employee Oros data, was subsequently performed and results compared with the same obtained analyzing Euro aggregate time series and those of the European countries. The results encouraged the release of the indicator on the number of persons employed as free instead of confidential, as flagged until the delivery of February 2019¹⁰. To conclude, some still critical aspects in the method remain and are object of study for possible future enhancements. In particular, the main aspects to take a closer look at are:

- NA data are subject to important revision, at Nace level, observed once a year and that affects totally the estimation process causing revision on the entire time series, through the modification of the temporal disaggregation's parameters. An attempt in order to have a lower impact could be to aggregate the data at a different level of Nace detail for which good estimate of the parameters in the regression model and lower revision of NA data are observed.
- An important aspect is the management of the accuracy, accounting also for the variability induced by the 2-digit Nace level LFS data in the sectorial disaggregation method described in (3). Another improvement could be the formulation and estimation of the dependence of the total variability with that of the LFS data, for which can be used the same approach described in LATTANZIO et al. 2017, including that induced by the temporal disaggregation.
- LFS data on self-employees contain also the non-regular component. An evaluation of the impact of this component should be done, in order to support an eventual procedure of decomposition.
- The NA annual number of self-employees covers also the outworker component of the self-employment, that, in order to satisfy the STS definition, should be excluded.

¹⁰ Results of this revisions analysis are not described in this work but a documentation is available at request.

Appendix

The Chow-Lin regression approach used for temporal disaggregation

The method (see Chow and Lin - 1971) consists in fitting a linear regression model between the aggregated quarterly series y , to be estimated, and the indicator series x , available, supposing the residuals to follow an autoregressive process at lag 1¹¹

$$y_t = \alpha + \beta x_t + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

Since the quarterly values for the aggregated series are not available, the model is fit on the annualized data that are, instead, available both for the aggregated and the indicator series. This implies that the linear relation observed on the annualized data is reproduced on the quarterly disaggregated data. The regression is made using the GLS (Generalized Least Squared) regression approach. The parameters to be estimated are α, β and the auto-correlation parameter ρ . In fact the variance-covariance matrix has a known structure and depends only on this value, that is calculated using the maximum log-likelihood of the GLS regression method (see BOURNAY and LAROQUE 1979). This is the form of the variance-covariance matrix.

$$\Sigma_{CL}(\rho) = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \begin{bmatrix} 1 & \dots & \rho^{n-1} \\ \vdots & \ddots & \vdots \\ \rho^{n-1} & \dots & 1 \end{bmatrix}$$

Riferimenti bibliografici

CHOW G., LIN A. 1971. Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series, *The Review of Economics and Statistics*, No. 53, pp. 372-375.

BOURNAY J., LAROQUE G. 1979. Reflexions sur le methode d'elaboration des comptes trimestriels, *Annales de L'Insee*, Vol. 36, pp. 3-29.

ISTAT 2006. *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*. Roma: Collana Metodi e Norme, Vol. 32.

SAX C., STEINER P. 2013. Temporal Disaggregation of Time Series, *The R Journal*, Vol. 5(2), pp. 80-87.

¹¹ Other versions of the Chow-Lin regression approach, such as the version by Litterman and that by Fernandez, propose different hypothesis on the distribution of the residuals, that is respectively a non-stationary process and a random walk.

ISTAT. AA.VV. 2015. Soluzioni metodologiche per l'utilizzo integrato delle fonti statistiche per le stime dell'occupazione, *Istat Working Papers*, No. 19.

LATTANZIO M. et al. 2017. Quality improvements in variance estimation for the labour force survey, *Rivista SIEDS*, Vol. 71, No. 4.

ISTAT. 2019. *La rilevazione trimestrale Oros su occupazione e costo del lavoro: indicatori e metodologie*. Roma: Collana Metodi.

LATTANZIO M., POGELLI F.R., TUZI D. 2019. L'imputazione delle mancate risposte sulle posizioni lavorative dipendenti nei dati provvisori di fonte Inps, *Istat Working Papers*, No. 4.

LATTANZIO, M. 2019. Indici semplici di variabili rapporto: analisi delle proprietà, scomposizione e caso di studio nella rilevazione Oros, *Istat Working Papers* (submitted to the review).

SUMMARY

The new methodology for the Italian STS indicator on the “Number of Persons Employed”

Since May 2018 Istat delivers to Eurostat quarterly business based indicators on the Number of Persons Employed, required by the EC-Regulation on Short-Term Statistics (STS), according to a new compilation methodology. The innovation concerns the calculation of the self-employed component. While for the employee part quarterly figures are calculated using data from the Istat Oros survey, for the self-employed part official statistics do not provide estimates at the frequency and business detail level required by the Regulation. Actually, figures on self-employed persons are available only at annual level from the National Accounts (NA) or from the Business Register (BR). For timeliness reasons, NA data were preferred. The approach at the basis consists in the quarterly disaggregation of the annual source. A simple disaggregation method was used also in the past methodology, as described in the paper, but the results were poorly meaningful, particularly in sectors with a significant self-employment influence. The new methodology includes two major improvements: 1) the exploitation of a new source for the estimation of the quarterly dynamic of self-employees, that is the Labour Force Survey (LFS) number of self-employed persons, used to temporally disaggregate the NA annual number of self-employees; 2) a new method for the temporal disaggregation, based on a mixed use of the Chow-Lin regression model, applied at an aggregated level, and a method aimed at preserving the LFS quarterly structure, for the Nace disaggregation (Eurostat requires data at 2-digit Nace level for industry and services). This work focuses on the innovations introduced, stressing on the differences with the old method showing the improvement in the short-term dynamic. Finally, some still critical aspects are described and possible enhancements suggested.

Marco LATTANZIO, Istat, lattanizio@istat.it

MANAGEMENT OF CULTURAL HERITAGE USING BIG DATA

Sandro Stancampiano

1. Introduction¹

The focus of this paper is about management of cultural heritage. We want to show the huge potential of data on the web to produce statistics in order to optimize decision-making processes.

We have chosen three of the most important Italian cities of art, characterised by stunning works of art but also by many administrative difficulties.

We build a corpus - in Latin language - composed of recent reviews about St Mark's Basilica (Venice), Colosseum (Rome) and Norman Palace (Palermo).

We want to discover regularity in the text examined by cluster analysis (BOLASCO, 2014).

The tourism industry represents an important source of income for the Italy, as a matter of fact italian state museums registered an increased revenue between 2013 and 2018, according to the statistics released by Mibact Statistic Office (Table 1).

Table 1 – *Visitors and income by year.*

Year	visitors	revenue (€)
2014	40.744.763	135.510.702
2015	43.288.366	155.494.415
2016	45.383.873	173.383.941
2017	50.169.310	193.915.765
2018	55.504.372	229.360.234

Source: Mibact - Statistic Office, 2019.

¹ The views expressed in this paper are those of the author and do not necessarily reflect the policies of ISTAT (Italian National Institute of Statistics).

In Table 2, we show data provided by “Trips and Holidays”, a focus included in the Istat Household Budget Survey, certify the wide diffusion of Internet as medium of booking travel (ISTAT, 2019). Reservation made using Internet have increased by 14.2% (31.8% in 2014 compared with 46% in 2018)².

Table 2 – *Trips by type.*

Year	holidays	business	total
2014	30.2	42.8	31.8
2018	45.5	50.3	46.0

Source: Istat - Trips and holidays in Italy and abroad, 2019

Many visitors (including those visiting museums in their city of residence) assign ratings to places, adding considerations on the state of conservation of the monuments, services and disservices they have noticed. We believe that by analysing these comments, it is possible to deduce valuable information (STANCAMPIANO, 2018).

This paper is structured as follows: in section 2 we present the theoretical framework; in section 3 we describe the corpus and the methods; in section 4 we illustrate the main results; in section 5 we delineate the conclusions and the future steps.

2. Theoretical framework

The texts we are analysing in this study have been collected using the Diogenes software³.

The Diogenes software is developed by the author in the Java language: allows to store on a relational Data Base texts extracted from the web using web scraping techniques⁴.

Within the activities related to this study were automatically collected and stored more than 6000 documents (each document is a review).

² Tourism, according to Istat definition, is the activity of travelling made by visitors to a main destination outside their usual environment.

³ <http://diogenes.statsapp.it/> [Accessed 26 Oct, 2019]

⁴ https://en.wikipedia.org/wiki/Web_scraping [Accessed 26 Oct, 2019]

Tripadvisor was chosen among the many websites used by users to produce content. Registered users use the site to write their reviews on the places they went to sharing their experiences and exchanging advice.

The reasons why people publish content on the web are partly unknown; although some potential reasons have been investigated (HO and DEMPSEY, 2008).

We believe that, regardless of the reasons why this happens, studying and analysing this content is an opportunity to be adequately valued (CERON *et al.*, 2014).

In this study we applied the methodology proposed by Reinert: the speaker exposes a series of mental places in sequence that impose their vocabulary. The statistical examination of the distribution of this vocabulary makes it possible to identify the "mental rooms" subsequently inhabited by the speaker (REINERT, 1995).

We can identify the main topics of the corpus highlighting the "lexical worlds": this happens because the recurrence of certain words or groups of words (in the same discursive contexts) is not a random fact.

To understand the content of a text, without reading it, Reinert proposes the study of co-occurrences.

We use the software IRaMuTeQ, created by Pierre Ratinaud, in order to apply the ALCESTE (Analyse Lexicale para Context d'un Ensemble de Segments de Texte) method, that allows us to classify the contents of the corpus analysing the words full (nouns, verbs, adjectives and some adverbs) reduced to their lexeme (GRECO, 2016).

IRaMuTeQ is a graphical interface of the statistical software R developed in the Python language (SOUZA *et al.*, 2018).

Applying the method implemented in IRaMuTeQ we obtain the dendrogram in which each class represents a theme of the corpus. The analyzed corpus is composed of a set of reviews dealing with similar topics, in this case the account of a cultural experience.

Descending Hierarchical Classification (DHC) allows you to switch from Initial Context Unit (each review is an ICU) to Elementary Context Unit (ECU) or text segments. Text segments (TS) are the words context; the corpus division in TS is automatically done.

3. Corpus and Methods

The 6569 reviews collected, published between 2015 and 2019, are divided as follows: Colosseum 2421 (36.8%), Norman Palace 2049 (31.1%), and St Mark's Basilica 2099 (31.9%). The corpus analyzed with the purpose of identifying different arguments about the

same topic consists of 7125 texts segments, 231229 occurrences and 13122 forms. The number of active forms with a frequency greater than 3 is 2868.

DHC was performed taking into account only full words (adjectives, adverbs, nouns and verbs) and produced a division in four clusters of text segments. Each cluster represents a recurring topic within the texts. The algorithm maximizes similarity between statements in the same class. IRaMuTeQ divided the corpus into two subcorpus classifying ranking 6416 out of 7125, more than 90% a very efficient result (CAMARGO and JUSTO, 2018).

The first subcorpus consists of 2006 ECU that correspond to 31.3% of the total while the second subcorpus had a further division in three classes: class 1 with 1739 ECU (27.1%), class 2 with 1062 ECU (16.6%) and class 3 with 1609 ECU_s (25.1%) (Figure 1).

Figure 1 – Dendrogram of the classes.



Source: Iramuteq data.

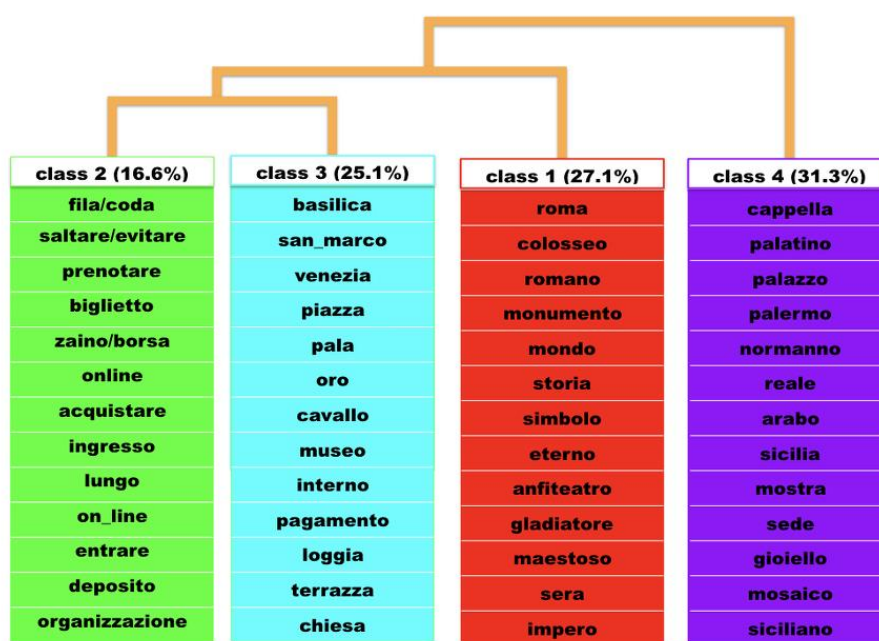
The analysis made it possible to identify groups of segments that are homogeneous within them and heterogeneous among themselves regarding the “concepts” expressed in the whole corpus.

The classes, identified by hierarchical clustering, divide the vocabulary and allow the definition of the lexical worlds. Alceste textual analysis methodology allows to extract the most important contents because repeated by many users in their stories. Users by observing what happens and telling their experience on social network can become sentinels and facilitate the task of those who must manage cultural heritage.

4. Main results and Discussion

Cluster analysis allows us to group statistical units by maximizing the cohesion and homogeneity of the words included in each group and at the same time minimizing the logical link between those assigned to different groups / classes.

Figure 2: Dendrogram of most representative words



Source: Own processing on Iramuteq data

The dendrogram (Figure 2) shows the division of the corpus into 4 classes. The words contained in each class make it possible to identify the types of topics covered in the corpus, applying the Alceste methodology proposed by Reinert and

implemented in the IRaMuTeQ software (REINERT, 1995). In Figure 2 we observe the words belonging to the four groups and how they are related to the three point of interest. These word clusters integrate several segments according to the vocabulary distribution.

Clusters 1, 3 and 4 concern respectively the visit of the Colosseum, St Mark's Basilica and Norman Palace as we can imagine from the characteristic words. The words are ordered by the χ^2 of association between word and cluster: coefficient calculated with one degree of freedom on the contingency table that crosses the presence/absence of the word in an ECU with the fact whether or not this ECU belongs to the class considered (REINERT, 1995). Obviously we find words that are closely related to the place visited in the first positions. There are theme words like *roma*, *colosseo*, *simbolo*, *anfiteatro* and *gladiatore* concerning cluster 1, *basilica*, *san_marco*, *venezia*, *pala* and *oro* regarding cluster 3 and words as *cappella*, *palatino*, *palazzo*, *palermo* and *normanno* in cluster 4; moreover in these three clusters there are words that represents TS that express positivity and actions related to the visit.

Table 3 – Typical ECU of cluster 2.

prima domenica del mese ingresso gratis la biglietteria apre alle 8 30 ma già alle 8 c'è fila. Tutt'intorno è pieno di venditori più o meno abusivi di visite guidate con opportunità di saltare la fila quella all'ingresso non quella alla biglietteria
colosseo fantastico, da rivedere l'organizzazione, monumento di bellezza più unica che rara merita assolutamente la visita, difetta di organizzazione nella vendita dei biglietti visto che polizia municipale ed un addetto mi hanno fatto fare la fila al botteghino salvo poi scoprire che il biglietto si comprava direttamente all'ingresso.
caotico per accedere a questa basilica è obbligatorio fare code chilometriche e bisogna informarsi molto bene degli orari, noi l'abbiamo fatta per poi scoprire che era chiusa, inoltre occhio a non avere zaini o borse grandi altrimenti verrete respinti sinceramente mi aspettavo migliore organizzazione
la basilica merita sicuramente qualche minuto di coda ed è gratis non ci si può accedere con gli zaini ma viene offerto un deposito gratuito presso il quale ti danno un pass che dura un'ora e che ti permette anche di saltare la fila
nonostante gli sforzi per permettere l'accesso ai diversamente abili motori, resta molto da fare. La "pala" è visibile solo con accompagnatori così come il "tesoro"
Belle le stanze e davvero da non perdere la cappella palatini. Accessibile anche a chi ha difficoltà motorie, è davvero un monumento da vedere.

Source: Own processing on IRaMuTeQ data.

From our point of view the most interesting group is the cluster 2. We observe that in this cluster we find words like *fila*, *saltare*, *prenotare*, *biglietto*, *deposito* and *organizzazione*. The text segments related to these words represents economic and practical aspects that in some cases may cause discomfort during the visit.

The underlying topics are related to the cost of the ticket, waiting list and the manner of the visit with both positive and negative connotations depending on the particular situation described by the user.

In Table 3 we present some of the typical text segments of the cluster. The segments cover all three attractions examined.

As we can read these reviews are full of useful and precise indications regarding times and visit arrangements as well as regarding the usability of the monuments.

5. Conclusion

The issues highlighted are of interest to public administrators, who can hear directly from the voice of citizens what are the main problems from the point of view of users. Based on this type of analysis, the decision maker can evaluate if and how to intervene to improve the management of cultural sites and heritage. The flow of information starts from the citizen who at the end of the process can obtain tangible benefits thanks to the data that he has put on the web. The process described in this paper shows a classic use of Big Data: data produced with a specific purpose are subsequently used to achieve other objectives, bringing an undeniable added value (RUDDER, 2015). The applied text mining techniques allowed to enhance information that otherwise would have remained unused. Further and more detailed analysis can be carried out using the same methodology and the same software used in this work. You can continue monitoring, increasing the corpus to conduct further analysis on these same monuments or study other cities and other cultural assets in order to improve management policies and optimize decision-making processes.

References

- BOLASCO, S. 2014. *Analisi Multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*. Roma: Carocci editore.
- CAMARGO B.V., JUSTO A.M. 2018. Tutorial para uso do software de análise textual IRAMUTEQ. Universidade Federal de Santa Catarina, Available at: <http://www.iramuteq.org/documentation/fichiers/tutoriel-portugais-22-11-2018> [Accessed 27 Jun, 2019].
- CERON A., CURINI L., IACUS S. M. 2014. *Social Media e Sentiment Analysis. L'evoluzione dei fenomeni sociali attraverso la Rete*. Springer Italia.
- GRECO, F. 2014. *Integrare la disabilità. Una metodologia interdisciplinare per leggere il cambiamento culturale*. Milano: FrancoAngeli.
- HO J.Y.C., DEMPSEY M. 2008. Viral marketing: Motivations to forward online content, *Journal of Business Research*, Vol. 63, pp. 1000-1006, <https://doi.org/10.1016/j.jbusres.2008.08.010> [Accessed 27 Jun, 2019].
- ISTAT 2019. Trips and holidays in Italy and abroad. <https://www.istat.it/en/archivio/22702>.
- REINERT, M. 1995. I mondi lessicali di un corpus di 304 racconti di incubi attraverso il metodo "Alceste". In CIPRIANI R. e BOLASCO S. (Eds), *Ricerca Qualitativa e Computer. Teorie, metodi e applicazioni*. Milano: FrancoAngeli, pp. 203 - 223.
- RUDDER, C. 2015. *Dataclisma. Chi siamo quando pensiamo che nessuno ci stia guardando*. Milano: Mondadori.
- SOUZA M.A.R., WALL M.L., THULER A.C.M.C., LOWEN I.M.V., PERES A.M. 2018. O uso do software IRAMUTEQ na análise de dados em pesquisas qualitativas, *Revista da Escola de Enfermagem da USP*, Vol. 52, <http://dx.doi.org/10.1590/S1980-220X2017015003353> [Accessed 27 Jun, 2019].
- STANCAMPIANO, S. 2018. Misurare, monitorare e governare le città con i Big Data. In IEZZI DOMENICA F., CELARDO L., MISURACA M. (Eds), *JADT' 18. Proceedings of the 14th International Conference on the Statistical Analysis of Textual Data*, Roma: UniversItalia, pp. 748 - 754.

SUMMARY

Management of cultural heritage using Big Data

The results of this research can support the administrators in the management of services dedicated to users of cultural heritage in the area. The experiment uses text analysis to extract information from reviews downloaded from the web using web scraping techniques. The flow of information starts from the citizen who at the end of the process can obtain tangible benefits thanks to the data put on the web.

The process described in this paper shows a classic use of Big Data: data produced with a specific purpose are subsequently used to achieve other objectives, bringing an undeniable added value.

The applied text mining techniques allowed to enhance information that otherwise would have remained unused.

INTERNATIONAL COOPERATION FOR REDESIGNING MOLDOVAN SOCIAL SURVEYS

Andrea Fasulo, Marco D. Terribili

1. Collaboration goals

Collaboration between National Statistical Institutes (NSIs) for sharing knowledge and expertise is a good practice which has proven to be successful in many international experiences, as technical reports and scientific paper show.

In 2018 we started to collaborate with the colleagues of the Moldovan National Bureau of Statistics (NBS), with the sponsorship of the World Bank, with the aim of providing methodological support to design the new sampling strategies for two important social surveys: Labour Force Survey (LFS) and Household Budget Survey (HBS).

These strategies could allow NBS to reach the European standard estimation accuracy levels -although Republic of Moldova is not obligate to respect them, not being a European country- and to produce better indicators about occupational status, household expenditures and other important social phenomena.

1.1. Collaboration phases

Collaboration lasted formally one month, with two one-week missions in Chişinău, and two home-based weeks in the middle.

Initially we acquired information on the Moldovan socio-demographic characteristics, mainly about emigrations, difference between urban and rural areas, informatization of the population, data source history, etc. This phase was really important to focus on the possibility of improving statistics produced, according to the socio-demographic characteristics of the country and of its populations.

Then we start to study the actual sampling design of both the surveys, mainly analysing the new data sources acquired by NBS (geolocalized census sectors, electricity register, 2014 Population and Housing Census data, etc.). The exploitation of these new data could increase the statistics quality, but a preliminary data quality analysis is needed.

During the two home-based weeks, that we spent in Rome, we try to find methodological solutions suitable with the Moldovan statistical peculiarities. Different scenarios have been studied, from every methodological point of view, such as the use of different sampling frames and their combination, data collection on the territory, estimates and expected errors computation.

Finally, we proposed to the NBS top management the scenarios and we discussed them, perfecting our proposal according to budget, logistic and precision constraints given by NBS.

1.2. Moldovan socio-demografic characteristics

Republic of Moldova, is a landlocked country bordered by Romania and Ukraine. The capital city is Chişinău. The country is divided into about 900 municipalities (12000 census enumeration areas) belonging to 32 districts, three main municipalities and two autonomous regions, Gagauzia and Transnistria. Gagauzia autonomy is ethnically motivated by the predominance of Orthodox-Turkish-speaking people. Transnistria situation is more complex because the central government do not control the area, which is controlled by the unrecognized Pridnestrovian Moldovan Republic, since USSR dissolution, in 1991. Statistics survey in these two areas are really difficult, and logistic difficulties persist also for collecting data in rural area, where villages are really small, often administratively part of either communes and where house numbering scheme is not always adopted.

Figure 1 – Administrative districts division of Moldova.

Over the last few years a huge amount of Moldovan citizens have emigrated and started to work abroad, mainly in Russia, Europe and Turkey. Estimation of this phenomenon magnitude is quite hard to estimate but we can certainly talk about a migration rate over the 30% of the whole population during the last twenty years. Usually just some members of an household emigrate (mainly women who can easily find in-home nursing jobs) and sometimes the family join the emigrated member after a stabilization period.

Moldovan migration entails two important side effects from a statistical point of view: many vacant houses, which make difficult the door-to-door data collection using PAPI or CAPI methods and a strong internet-use propensity of the population, accustomed to use PCs and smartphone to communicate with relatives and friends who live abroad. These population characteristics allow the NBS to rethink the data-collection modes adopting web interviews: their lower cost can

lead to a larger sample size too without introducing a relevant response bias in the survey.

Other important peculiarity of the Moldovan population is the remarkable number (almost a third) of citizens with double citizenship: Romanian speaking people can easily access for ethnic reasons to Romanian citizenship and passport, which guarantee to easily travel and move through European Union countries, without any visa. Moreover, in the last few years Romanian government have encouraged economically demographic growth: birth incentives have pushed many Moldovan parents with double citizenship to directly give Romanian citizenship to their children, also if they will effectively grow up in Moldova. Obviously this phenomenon leads to an underestimation of the whole population to consider both in design and in estimation phases.

2. Actual sampling design of the Moldovan social surveys

In 2003 – 2006, NBS developed a Multifunctional Sample for Surveys from Social Field, so-called EMDOS, which is used for carrying out HBS and LFS, as well as other household surveys from social field. Thus, in 2006 major changes took place in sampling field, and up to 2018 LFS and HBS were carried out, as two integrated surveys at the level of the single territorial sample – EMDOS. In order to increase the efficiency of available resources one common network of interviewers has been created in the frame of integrated master sample – EMDOS.

Specifically, LFS and HBS samples were probabilistic samples ensuring computation of the main indicators at national level, by areas of residence and at the level of 4 statistical zones – North, Center, South and Chisinau. LFS and HBS samples represent independent samples within each stratus. The different design of the samples was registered in Chisinau and Balti municipality, where their territory was divided into 20 and respectively 5 sub-sectors, as well as in Cahul and Ungheni cities, where the territory was divided into 2 sub-sectors. The first sampling stage, that was EMDOS extraction consists in dividing the territory of the Republic of Moldova in PSUs – these being composed generally from level-two-administrative-territorial units (cities or communes), but also from smaller parts of a locality from urban area. Afterwards the PSUs were stratified depending on their geographical location, area of residence and their size. The strata were determined with the purpose to differentiate the population by level of its employment on labor market and implicitly to homogenize the population. The number of PSUs extracted from every stratus was proportional to the stratus size. The probability of

extracting a PSU was proportional to the number of its population, according to the data of the Population Census from 2004. At the second stage, households were selected from the lists established based on electricity consumers data offered by Union Fenosa, RED Nord and RED Nord-Vest, and updated in rural areas via a special listing procedure.

According to the National Strategy for Development of Statistical System in 2016-2020 period, NBS has to “improve the sampling methods, including sample framework for households surveys in social area in 2018” (activity 1.2.8). The methodological innovations presented in the next chapter regards in particular the improvement of the sample design and so the PSUs to be selected in order to ensure better quality estimates that can respect the European efficient standards.

3. Finding methodological solutions

According to actual sampling design and to socio-demographic characteristics of the countries several scenarios have been proposed to discuss. In the table below the pros and cons of each of scenario are summarized.

Table 1 - *Redesigning scenarios proposed with pros and cons*

	Scenario	Pros	Cons
1.	Area sampling survey of enumeration areas:	no households list needed;	no control of sample size.
2.	Systematic sampling survey:	no households list needed, control of sample size;	households amount needed, no proper for a NSI.
3.	Combined two stages sampling design: Area + Households	households list used just in some PSUs (when available);	control of sample size just in some PSUs.
4.	Multivariate optimal allocation for different domains in a two stages stratified sampling design	control, since design phase, the sample size and the precision of interest estimations of several variables, both at national and at regional level;	households list needed, more complex allocation process.

Scenarios are increasingly more complex in terms of information needed in the design phase and control of sample size and estimates attended accuracy.

The first scenario proposed based on an area sampling survey of EAs. This scenario solved the lack of a complete household list but, on the other hand, didn't allow to set and control from the design phase the survey sample size. The impossibility to control the sample size has led NBI managers to not choosing this scenario.

The second scenario has characterized by a systematic sampling survey. This sample design allowed to control the sample size but has been not chosen because the NBI manager did not consider appropriate to use this sample design for the two main Moldovan surveys.

So, in the third scenario a combination of area sample design and household list sample design has been proposed. The use of two different sample design allowed to maximize the use of the information available and minimize the non-control of the sample size just for some PSUs. Also this scenario has been rejected for the impossibility to totally control the size of the samples.

The final solution on which all of us converged was the multivariate optimal allocation for different domains in a two-stages stratified sample design: this solution allows to control, since design phase, the precision of interest estimations of several variables, both at national and at regional level. Strata are defined by crossing 5 areas (North, Center, South, Chisinau and Gagauzia) with 2 area type (rural/urban), obtaining 10 strata; three estimation domains have been considered to give the estimates at national, area and area-type level.

3.1. Rotation and coordination scheme

Quarterly rotation scheme of the primary stage units (PSUs) has been studied too, to make the surveys longitudinally similar to the main Eurostat surveys on the fields of labour force and household budget.

In detail, the PSUs for the LFS follow the rotation scheme **1 (2) 1 (8) 1**, which indicates that PSUs included in the sample in quarter t are included also in $t+4$ and $t+13$ quarters. This scheme guarantees a quarterly PSU overlap of the 33%. For the HBS survey the rotation scheme adopted is defined as **1 (11) 1 (11) 1**, so PSUs included in the sample in t month are considered also in $t+13$ and $t+25$ months. This scheme guarantees a monthly PSU overlap of the 25%.

Furthermore, a positive coordination scheme has been adopted in order to maximize the PSU overlap between the LF and HBS surveys.

4. Results

Desired sample size is defined according to budget constraints, such as data-collection mode, cost of a single interview and number of cities involved in the survey. To reach the desired sample in both surveys some allocation procedure parameters have been tuned. These parameters are the minimum number of households interviewed in every single PSU and minimum number of PSUs in not-self-representative strata (NSR). Tuning these allocation parameters influence the the number of PSUs involved in the survey and the total number of sampled households.

In the following table the values assigned to parameters in the end of the tuning phase.

Table 2 – *New samples characteristics.*

Survey	Minimum n° hhs per EA (expected)	Minimum n° hhs per EA (effective)	Maximum n° hhs per EA (effective)	Minimum n° EA per NSR stratum	Enumeration Areas	Municip.	HHS
LFS	30	13	53	20	300+5	204	7565
HBS	24	6	42	20	220	118	3411

Number of municipalities involved in the surveys change with coordination: for HBS number of EA decrease from 220 to 200, analogously the municipalities numbers decrease from 118 to 79, increasing overlap rate of 1.5 %.

Table 3 – *Coordination between HBS and LFS survey results.*

EA coordination	after	Municipalities overlap before coordination	Municipalities after coordination	Municipalities overlap after coordination	HHS after coordination
200		61	79	42	3400

Proposed sampling design allow to produce more efficient estimates, both for HBS and LFS, maintaining the previous sample size. The use of enumeration areas guarantees cluster effect reduction: smaller PSUs make the sample more spread on the territory and more heterogeneous under the same SSUs sample size.

5. Further steps

Right now the methodological innovation proposed to the Moldovan NBI has been applied on the capital urban area of Chisinau, but it could be enlarged to the whole national territory.

Multivariate optimal allocation for different domains in a two stages stratified sampling design will be implemented in a R-package called `R2BEAT`, that will be soon available on CRAN. This package places side by side two others R Istat package, called `FS4` and `ReGenesees`, which carry out both selection and estimation process, respectively. Packages synergy create a practical framework which allow NSIs to control the redesign survey in every step of the process, from allocation to estimation.

The new designs proposed will guarantee to the Moldovan National Bureau of Statistics to provide more efficient estimates, regarding Labour Force and households expenditures, and to adjust these surveys to the European standard, adopting coordination and rotation scheme usually used by the main European NSIs.

Acknowledgements

Our heartfelt thanks to Lilian Galer, chief of the Statistical Methods Division of the Moldovan National Bureau of Statistics, for the proficient collaboration and to Alexandru Cojocaru, Senior Economist at World Bank, for supporting us during all the cooperation.

References

- BETHEL, J. 1989. Sample allocation in multivariate surveys, *Survey methodology*, Vol. 15(1), pp. 47-57.
- COCHRAN, W. G. 2007. Sampling techniques. John Wiley & Sons.
- COROBICA L., GALLAGER L. 2008. Sampling and Statistical Inference for the Surveys carried out based on EMDOS Master Sample, *Technical Report*.
- FALORSI S., RUSSO A. 2001. Il disegno di rilevazione per indagini Panel sulle famiglie, *Quaderni di Ricerca*, No. 3, Franco Angeli, pp. 55-90.
- ISTAT, MAUSS, R. 2013. Multivariate Allocation of Units in Sampling Surveys, https://www.istat.it/it/files/2011/02/user_and_methodological_manual.pdf (Web, 22 March 2019)..
- NEYMAN, J. 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection, *Journal of the Royal Statistical Society*, Vol. 97(4), pp. 558-625.
- TSCHUPROW, AL. A. 1923. On the mathematical expectation of the moments of frequency distributions in the case of correlated observations (Chapters 4-6). *Metron*, Vol.2, pp. 646-683.

SUMMARY

International cooperation for redesigning Moldovan social surveys

Since 1997, National Bureau of Statistics (NBS) of the Republic of Moldova is carrying out two important household surveys: Labour Force Survey (LFS) and Household Budget Survey (HBS). They are the main data sources on economic and well-being status of Moldovan country, at the national and regional level.

In 2018 a collaboration between statisticians from the Italian National Institute of Statistics and the NBS Sampling Survey Unit started, with the sponsorship of the World Bank. The aim of this collaboration was to provide support in defining the new sampling strategies for LFS and HBS: these strategies could allow NBS to reach the estimation accuracy level of the most common European standards, although Republic of Moldova is not obligated to respect them, and to produce better estimations about occupational status, expenditures and other important social phenomena.

The collaboration has been sketched out studying the actual sampling design of both the surveys, analysing the new data sources acquired by NBS (geolocalized census sectors, electricity register, 2014 Population and Housing Census data, etc.) and finding methodological solutions suitable with the Moldovan statistical peculiarities. Different scenarios have been proposed to the NBS top management to be discussed, according to budget, logistic and precision constraints.

The final solution on which all of us converged was the multivariate optimal allocation for different domains in a two stages stratified sample design: this solution could allow to control, since design phase, the precision of interest estimations of several variables, both at national and at regional level. A quarterly rotation scheme of the primary stage units (PSUs) has been studied too, to make the surveys longitudinally similar to the main Eurostat surveys on the fields of labour force and household budget.

This redesign is still a work in progress. Right now it has been applied on the capital urban area of Chisinau, but it could be enlarged. The redesign process could concern also the estimation phase and the development of an R package for the multivariate optimal allocation for different domains in two stage stratified sample design, proposed in Moldovan experience.

Andrea FASULO, Istat, fasulo@istat.it

Marco D. TERRIBILI, Istat, terribili@istat.it

LA VALUTAZIONE DELL'EFFETTO TECNICA NELLE INDAGINI SOCIALI A TECNICA MISTA

Claudia De Vitiis, Francesca Inglese

1. Introduzione

L'Istituto Nazionale di Statistica da alcuni anni ha adottato l'uso congiunto di più tecniche di rilevazione (mixed-mode) per un numero crescente di indagini sociali. Questa scelta deriva sia dall'esigenza di contenere i costi introducendo la tecnica web, sia, soprattutto, dalla necessità di migliorare i tassi di risposta e la copertura della popolazione, in quanto l'utilizzo di più tecniche di rilevazione consente di raggiungere differenti target di popolazione.

L'uso della tecnica mista può introdurre, tuttavia, un effetto distorsivo nei dati rilevati, comunemente denominato *effetto tecnica*, che determina differenze sistematiche nelle stime finali. Diventa pertanto opportuno verificare che le stime prodotte dall'indagine siano coerenti e confrontabili con quelle analoghe ottenute nelle edizioni precedenti e che le variazioni nelle serie storiche siano dovute a reali cambiamenti del fenomeno osservato.

L'obiettivo del presente lavoro è illustrare l'analisi svolta per valutare l'impatto dell'introduzione del disegno mixed-mode (MM) rispetto al disegno unico o single-mode (SM) sulle stime dell'indagine ISTAT "Aspetti della vita quotidiana" (AVQ), che nell'edizione 2017 è stata realizzata per la prima volta con tecnica mista. Per analizzare tale impatto è stato realizzato un disegno sperimentale basato su due campioni indipendenti, a cui sono stati somministrati la doppia tecnica web/PAPI o la sola tecnica PAPI (campione di controllo). Le analisi condotte in tale contesto sperimentale hanno lo scopo di evidenziare e capire le ragioni che determinano differenze significative nelle stime ottenute con i due disegni. In particolare le analisi per lo studio della mancata risposta totale sono finalizzate alla valutazione della qualità complessiva dei campioni di rispondenti delle due indagini. I metodi utilizzati per la stima dell'effetto tecnica, poi, hanno l'obiettivo di evidenziare la presenza di errori di misura.

Il lavoro illustra nella sezione 2 il quadro teorico di riferimento e i principali approcci presenti in letteratura per l'analisi dell'effetto tecnica. Nella sezione 3 si descrivono le indagini sociali realizzate dall'ISTAT con tecnica mista, alcune caratteristiche dell'indagine AVQ e il disegno sperimentale adottato in occasione

dell'introduzione della doppia tecnica. La sezione 4 illustra l'impianto generale delle analisi condotte, mentre le sezioni 5, 6 e 7 presentano rispettivamente: l'analisi sulla mancata risposta totale; la stima delle componenti dell'effetto tecnica (selezione e misurazione); un confronto tra stime ottenute applicando diversi metodi di aggiustamento dell'effetto selezione. Infine, nella sezione 8, sono tratte alcune conclusioni.

2. Quadro teorico di riferimento per l'analisi dell'effetto tecnica

L'*effetto tecnica* introdotto dalla tecnica mista consiste in un effetto distorsivo che può determinare differenze sistematiche nelle stime finali. Tale effetto è dovuto sia a differenze negli errori di misura (*effetto misurazione*), causati dall'uso di strumenti di rilevazione differenti, sia alla diversa composizione dei campioni di rispondenti alle diverse tecniche di raccolta dei dati (*effetto selezione*).

Le tecniche di rilevazione possono essere combinate e assegnate alle unità campionarie secondo differenti disegni: concorrente quando la tecnica di rilevazione è assegnata ex-ante alle unità del campione in modo casuale o sulla base di caratteristiche note a priori (variabili socio-demografiche o *preference mode*); sequenziale quando si propone prima la medesima tecnica a tutte le unità e successivamente si torna sui non rispondenti con una seconda o terza tecnica. Nei disegni di tipo concorrente, la diversa composizione dei campioni di rispondenti è legata proprio a differenti popolazioni di riferimento. Infatti, per garantire la rappresentatività dell'intera popolazione, campioni indipendenti sono selezionati da popolazioni raggiungibili con diverse modalità di rilevazione (ad es. la tecnica CATI per le famiglie con telefono fisso). Nei disegni di tipo sequenziale, la diversa composizione dei campioni è determinata dall'autoselezione dei rispondenti che scelgono modalità di risposta differenti (de Leeuw, 2005; Luzi et al., 2017).

L'effetto selezione rappresenta una conseguenza naturale e auspicata di questo tipo di disegni ma introduce una complicazione poiché si confonde con l'effetto misurazione (de Leeuw, 2005). Generalmente i due effetti sono contenuti nella relazione tra variabile d'indagine e tecnica e possono avere impatti differenti sulla stima di parametri della popolazione o sulla stima di modelli: l'errore di misura produce, infatti, differenze tra le medie nei diversi campioni di rispondenti, mentre l'effetto selezione influenza la struttura di correlazione dei dati.

Da un punto di vista inferenziale, gli effetti selezione e misurazione devono essere analizzati separatamente, sia per ottenere una corretta formulazione dell'errore non campionario totale, sia per l'applicazione di metodi che correggano le stime dei parametri di interesse dell'indagine dagli effetti distorsivi determinati da una molteplicità di cause. Nelle indagini a tecnica mista, infatti, il processo

inferenziale è accurato soltanto se la misurazione effettuata con più tecniche è equivalente (Hox et al., 2015).

Il problema della confusione tra i due effetti (Morgan e Winship, 2009) è il tema centrale della teoria dell'inferenza causale dove l'errore di misurazione è concettualizzato come effetto causale della tecnica sulla variabile d'indagine, mentre l'effetto selezione è visto come una correlazione spuria tra la variabile e la tecnica. Negli studi osservazionali, per la stima di entrambi gli effetti, l'inferenza causale viene utilizzata secondo una prospettiva controfattuale ipotizzando, cioè, l'esistenza di un risultato potenziale non realmente osservato (il valore che l'intervistato avrebbe fornito con l'altra modalità). Approcci alternativi, tipici di disegni sperimentali si basano sull'uso di variabili strumentali, quando è disponibile un'indagine di riferimento (Vannieuwenhuyze et al., 2010), o sulle re-interviste realizzate con una seconda tecnica per un sottoinsieme di intervistati che hanno già risposto attraverso una prima tecnica (Klausch et al., 2018).

I risultati delle analisi dell'effetto tecnica, sono condizionati dalla piena validità dei metodi applicati e delle ipotesi sottostanti. Si può scegliere di utilizzare metodi di aggiustamento del solo effetto selezione oppure metodi di correzione degli errori di misurazione generati dalle tecniche. L'effetto selezione può essere corretto con metodi di aggiustamento dei pesi campionari (post-stratificazione, calibrazione, propensity score) che hanno l'obiettivo di bilanciare le distribuzioni dei campioni di rispondenti a tecniche differenti rispetto ad un set di variabili ausiliarie esplicative dell'effetto selezione e *mode-insensitive*. I metodi di aggiustamento dell'effetto misurazione sono invece finalizzati alla correzione delle stime di singole variabili e non sono facilmente applicabili a indagini il cui obiettivo sia produrre una notevole molteplicità di stime. Si tratta di metodi che assumono, generalmente, l'ignorabilità dell'effetto selezione e una tecnica come riferimento (benchmark, ad esempio la tecnica ritenuta più affidabile, o quella storicamente utilizzata dall'indagine). Alcuni metodi possono essere utilizzati per produrre direttamente stime aggregate (modelli di regressione generalizzati), altri, come quelli basati su modelli di imputazione multipla, per modificare i microdati convertendo le risposte fornite con una specifica tecnica in risposte controfattuali. In quest'ultimo approccio l'errore di misura è concettualizzato come un problema di dati mancanti che può essere trattato con metodi di imputazione standard, oppure con il metodo di imputazione con pesi frazionati (Park et al., 2016) o il metodo di imputazione con modello di risposta e modello di selezione (Suzer-Gurtekin et al., 2012). Tali metodi si differenziano per le assunzioni circa il meccanismo che genera i dati missing (MAR, NMAR) e per l'utilizzo o meno di una tecnica di riferimento (Luzi et al., 2017).

3. Indagini sociali a tecnica mista in Istat e il contesto sperimentale dell'indagine *Aspetti della vita quotidiana*

L'indagine longitudinale sulle Forze di lavoro è la prima indagine sociale che in Istat ha utilizzato il mixed mode, un disegno d'indagine di tipo concorrente con tecnica CAPI alla prima occasione e CATI alle successive tre. Un disegno d'indagine analogo (CAPI-CATI concorrente) viene utilizzato anche nell'indagine longitudinale Eusilc (European Survey on living condition). Altre indagini su famiglie o individui tradizionalmente realizzate con tecnica CATI sono passate a tecnica mista per problemi di copertura della rete di telefonia fissa, quali: *Sicurezza dei cittadini e violenza sulle donne* da CATI a CAPI/CATI concorrenti; *Percorsi di studio e lavoro di Laureati e Diplomati*, da CATI a web/CATI sequenziale. Negli ultimi anni, le indagini su famiglie tradizionalmente realizzate con tecnica PAPI, come l'Indagine su *Cittadini e tempo libero* 2015 e l'Indagine annuale *Aspetti della vita quotidiana*, sono passate dalla tecnica PAPI alla tecnica mista (web/PAPI) e hanno adottato un disegno di tipo sequenziale.

L'Indagine multiscopo sulle famiglie "Aspetti della vita quotidiana" è basata su un campione di circa 24.000 famiglie (di cui un gruppo di circa 18.000 famiglie intervistate, 38.000 individui), selezionate con disegno a due stadi (comuni/famiglie) dal registro comunale centralizzato (LAC). Nell'indagine sono coinvolti quasi 850 comuni italiani. L'indagine raccoglie informazioni sulle attività ricreative e culturali nel tempo libero, come sport, lettura, cinema, musica, Internet, relazioni sociali, problemi per la qualità della vita delle persone e fornisce all'Eurostat ogni anno alcuni importanti indicatori sul benessere e sull'utilizzo delle tecnologie informatiche (ICT).

Nell'edizione 2017 dell'indagine AVQ è stata introdotta per la prima volta una tecnica mista di rilevazione con disegno sequenziale web/PAPI. Dapprima, a tutto il campione è stata proposta, tramite lettera di invito inviata dall'ISTAT, la compilazione di un questionario in modalità web e successivamente alle famiglie non rispondenti è stata sottoposta l'intervista diretta con tecnica PAPI. Il disegno sperimentale progettato ha previsto la realizzazione di due indagini parallele su campioni indipendenti: un'indagine con tecnica mista (MM) e un'indagine con tecnica singola (SM). Il campione di ciascun comune è stato diviso casualmente in due sottocampioni: al primo, di dimensioni maggiori, sono state assegnate le tecniche web e PAPI in modo sequenziale (MM); al secondo, è stata assegnata la sola intervista PAPI (SM, campione di controllo).

Il tasso di risposta alle due indagini è stato leggermente più alto per il campione MM (74% contro il 71% del campione SM), mentre, il tasso di risposta web è risultato fortemente dissimile tra le aree geografiche (dal 36% nel Nord Est al 17% nel Sud e isole).

4. Schema delle analisi

Lo studio dell'impatto sulle stime dell'indagine AVQ derivante dall'introduzione del disegno mixed-mode rispetto al disegno single-mode è stato sviluppato su più livelli di analisi: il primo basato sul confronto tra i due disegni SM e MM per una valutazione della qualità dei risultati delle due indagini; il secondo finalizzato alla valutazione dell'effetto tecnica (selezione e misurazione).

Nel primo livello di analisi sono stati effettuati test statistici sulle differenze delle stime calcolate sui campioni SM e MM. Le analisi sono state finalizzate allo studio della distorsione dovuta alla mancata risposta totale nei due campioni. A tal fine sono stati definiti due distinti modelli di risposta e a partire dalle probabilità predette sono stati calcolati gli indicatori di rappresentatività della risposta per individuare difformità che potessero spiegare, in parte, differenze nelle stime dell'indagine. La diversa composizione dei campioni di rispondenti dovuta a differenze nella mancata risposta potrebbe, infatti, contribuire a generare differenze nelle stime.

Il secondo livello di analisi è stato incentrato sulla stima dell'effetto tecnica (selezione e misurazione) e quindi sull'applicazione di metodi che permettessero di separare i due effetti e di stimare, per alcune variabili target dell'indagine AVQ, gli errori di misura. I metodi di stima applicati sono: l'*approccio con variabile strumentale* (Vannieuwenhuyze et al., 2010), che presuppone l'esistenza di una tecnica benchmark, è stato applicato nell'ottica di comparare i campioni di rispondenti con tecnica mista e con la tecnica PAPI (tecnica di riferimento); il *propensity score*, che richiede la disponibilità di variabili ausiliarie *mode-insensitive* (Rosenbaum e Rubin, 1983), è stato applicato, invece, al solo disegno d'indagine con tecnica mista al fine di rendere comparabili i campioni di rispondenti web e PAPI. Il metodo, che assume che le differenze che rimangono tra gruppi comparabili siano causate soltanto da errori di misura, è stato utilizzato sia per stimare gli effetti separatamente sia come *metodo di aggiustamento* dell'effetto selezione.

Le variabili target prese in considerazione sono quelle i cui test sulle differenze tra stime sono risultati significativi (Martin e Lynn, 2011), come "Lettura di libri negli ultimi 12 mesi", "Uso di internet", "Uso del PC" e "Frequenza con cui vedi gli amici". Le variabili ausiliarie utilizzate nei modelli di risposta e nel propensity score, oltre a quelle di tipo territoriale (regione e tipologia comunale) sono state acquisite a livello individuale dal DB Archimede - Archivio integrato di micro dati economici e demografici - e poi ricondotte a un livello familiare definendo delle tipologie familiari per numero di componenti e età, per classi di reddito (quintili), per titolo di studio più alto e cittadinanza. Tali variabili sono risultate essere i predittori più significativi di entrambi i modelli (risposta e propensity score).

5. Indicatori di rappresentatività della risposta

Per valutare la qualità complessiva dei campioni di rispondenti in termini di distorsione è stato utilizzato un indicatore di rappresentatività (*R*-indicator) che è basato su una misura di variabilità della propensione alla risposta. L'indicatore descrive come il campione dei rispondenti ad un'indagine riflette la popolazione di interesse rispetto a un insieme di variabili caratterizzanti la popolazione stessa e misura di quanto il campione di rispondenti si discosta dalla risposta rappresentativa. La stima dell'indicatore *R* è esprimibile come

$$\hat{R}(\hat{\rho}_X) = 1 - 2\hat{S}(\hat{\rho}_X), \quad (1)$$

dove $\hat{\rho}_X$ è la propensione alla risposta stimata rispetto alle variabili ausiliarie \mathbf{X} , disponibili per le unità del campione teorico, e $\hat{S}(\hat{\rho}_X)$ è la stima della standard deviation di $\hat{\rho}_X$ (Schouten et al., 2009, 2011).

Nella tabella 1 sono riportati i valori dell'indicatore *R* calcolato utilizzando le probabilità predette sulla base dei modelli di risposta stimati sia a livello nazionale e, considerando l'elevata variabilità dei tassi di risposta web tra le aree geografiche, per ciascuna ripartizione geografica. Dal confronto tra i valori ottenuti per i due campioni (Tab. 1) emerge che a livello Italia il campione di rispondenti MM si discosta meno dalla risposta rappresentativa rispetto al campione di rispondenti SM; infatti la stima dell'indicatore *R* è pari a 0,852 nel primo e a 0,812 nel secondo. I valori dell'indicatore sono invece pressoché simili per i due campioni nella ripartizione Nord, mentre sono molto diversi nelle altre ripartizioni: la risposta è più rappresentativa nelle ripartizioni Centro e Sud e Isole quando si adotta il disegno di indagine MM. Pur essendo i tassi di risposta web molto più bassi al Sud e Isole, nel disegno misto il campione di rispondenti riflette meglio la popolazione di interesse rispetto alle variabili che la caratterizzano.

Tabella 1 – Indicatori *R* nei campioni SM and MM.

Area geografica	Disegno SM	Disegno MM
Italia	0.812	0.852
Nord	0.847	0.840
Centro	0.752	0.842
Sud e isole	0.840	0.907

6. Stima degli effetti selezione e misurazione

Al fine di ottenere una stima dell'effetto tecnica separando gli effetti selezione e misurazione, sono stati implementati i due metodi menzionati nel paragrafo 4.

Il metodo con variabile strumentale (Vannieuwenhuyze et al. 2010) richiede che siano soddisfatte alcune assunzioni: l'equivalenza della misurazione e la rappresentatività dei campioni. Per rendere comparabili i campioni di rispondenti SM e MM è stato adottato un sistema di calibrazione con vincoli sulle distribuzioni delle variabili classe d'età, sesso, cittadinanza, regione e tipologia comunale. La variabile strumentale di tipo binario è stata utilizzata per suddividere il campione dei rispondenti all'indagine in due gruppi a seconda della tecnica, singola o mista, con cui avevano risposto. La stima degli effetti selezione e misurazione è stata determinata, a partire dalle distribuzioni di probabilità delle risposte del disegno SM (tecnica PAPI come benchmark) e dalle distribuzioni di probabilità condizionate delle risposte alle tecniche web e PAPI del disegno MM, per le variabili "lettura di libri negli ultimi 12 mesi" (Tab. 2) e "uso di internet" (Tab. 3). La tabella 2 mostra la presenza di entrambi gli effetti: l'effetto selezione positivo per la modalità NO può essere indicativo del fatto che i rispondenti PAPI nel campione MM hanno più probabilità di rispondere NO rispetto ai rispondenti web. Nella tabella 3, il segno positivo dell'effetto selezione per tutte le modalità, ad eccezione della prima, indica che i rispondenti PAPI nel disegno MM hanno maggiori probabilità di fornire tali risposte rispetto ai rispondenti web.

Tabella 2 – *Stima degli effetti selezione e misurazione per la variabile "lettura di libri negli ultimi 12 mesi"*.

Variabile	Modalità	Effetto selezione	Effetto misurazione
Lettura di libri negli ultimi 12 mesi	No	0,148	-0,073
	Si	-0,177	0,042
	N.r.	0,028	0,031

Tabella 3 – *Stima degli effetti selezione e misurazione per la variabile "uso di internet"*.

Variabile	Modalità	Effetto selezione	Effetto misurazione
Uso di internet	Si, negli ultimi tre mesi	-0,261	-0,052
	Si, da 3 mesi a 1 anno fa	0,001	0,007
	Si, più di un anno fa	0,005	0,014
	Mai	0,227	0,001
	N.r.	0,028	0,030

Il propensity score è un metodo dell'inferenza causale in ottica controfattuale che nel contesto delle indagini a tecnica mista e sotto l'ipotesi di ignorabilità dell'effetto selezione della tecnica, può essere interpretato come la probabilità di assegnazione (o scelta) di una tecnica condizionatamente alle covariate osservate (Rosembaum e

Rubin, 1983). Il propensity score di ciascun rispondente è definito come la probabilità di partecipare alla tecnica web stimata mediante un modello binomiale logistico su un insieme di variabili ausiliarie: $\text{logit}[p(x)] = \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j$. Il metodo consiste nel creare classi o gruppi omogenei di unità (famiglie) rispetto ai predittori del modello risultati significativi.

A partire dalla distribuzione delle probabilità predette dal modello di risposta web sono stati definiti i decili, e per ciascun gruppo generato è stata verificata l'ipotesi di bilanciamento (test d'ipotesi del Chi-quadrato), ossia che ognuna delle variabili ausiliarie utilizzate nel modello sia indipendente dalla modalità di rilevazione web/PAPI. L'ipotesi è stata verificata in otto dei dieci gruppi individuati, dove l'ipotesi di indipendenza veniva rifiutata per tutte le variabili. I gruppi sono stati creati correttamente, garantendo che le famiglie che si trovano all'interno dello stesso decile hanno una propensione alla risposta web simile.

Per ciascun gruppo bilanciato k , è stato calcolato il peso che rende equivalenti la proporzione pesata dei rispondenti alla tecnica web (m_1) alla proporzione dei rispondenti alla tecnica PAPI (m_2) nella stessa classe

$$w_k = \frac{n_{k,m_2}/n_{m_2}}{n_{k,m_1}/n_{m_1}} \dots \dots \dots (2)$$

Tramite i correttori definiti per le classi bilanciate sono stati stimati i due effetti (Vandenplas et al, 2016): l'effetto selezione, come differenza tra le stime pesate e non pesate dei rispondenti alla tecnica m_1

$$SE_{m_1}(\bar{y}) = \frac{\sum_{i=1}^{n_{m_1}} y_{i,m_1}}{n_{m_1}} - \frac{\sum_{i=1}^{n_{m_1}} w_{k,i} y_{i,m_1}}{n_{m_1}}; \dots \dots \dots (3)$$

l'effetto misurazione, come differenza tra la stima pesata dei rispondenti alla tecnica m_1 e la stima non pesata dei rispondenti a m_2

$$ME_{m_1}(\bar{y}) = \frac{\sum_{i=1}^{n_{m_1}} w_{k,i} y_{i,m_1}}{n_{m_1}} - \frac{\sum_{i=1}^{n_{m_2}} y_{i,m_2}}{n_{m_2}}. \dots \dots \dots (4)$$

Nella tabella 4 sono riportate le stime degli effetti per diverse variabili target dell'indagine AVQ. Nella maggior parte delle variabili riportate sembra esserci un impatto di entrambi gli effetti selezione e misurazione. Tuttavia per diverse modalità delle variabili "lettura di libri negli ultimi 12 mesi" e "frequenza con cui vede gli amici" è presente maggiormente un effetto misurazione.

Tabella 4 – *Stima degli effetti selezione e misurazione per alcune variabili target dell'indagine.*

Modalità	Stime			Effetti	
	Media web	Media web pesata	Media PAPI	Selezione	Misurazione
<i>Letture di libri negli ultimi 12 mesi</i>					
No	0.451	0.485	0.618	0.034	-0.132
Si	0.508	0.432	0.347	-0.075	0.085
N.r.	0.041	0.043	0.035	0.002	0.007
<i>Uso di internet</i>					
Si, negli ultimi tre mesi	0.698	0.662	0.548	0.042	0.077
Si, da 3 mesi a 1 anno fa	0.026	0.026	0.020	-0.001	0.008
Si, più di un anno fa	0.049	0.049	0.030	0.000	0.018
Mai	0.190	0.221	0.372	-0.040	-0.109
N.r.	0.037	0.042	0.031	-0.006	0.013
<i>Uso del PC</i>					
Si, negli ultimi tre mesi	0.623	0.577	0.423	0.055	0.109
Si, da 3 mesi a 1 anno fa	0.032	0.034	0.025	-0.002	0.010
Si, più di un anno fa	0.059	0.058	0.049	0.001	0.010
Mai	0.249	0.291	0.471	-0.055	-0.134
N.r.	0.037	0.041	0.032	-0.005	0.011
<i>Frequenza con cui vedi gli amici</i>					
Tutti i giorni	0.086	0.097	0.200	-0.015	-0.086
Qualche volta a settimana	0.254	0.250	0.267	0.004	-0.019
Una volta a settimana	0.212	0.208	0.188	0.003	0.017
Qualche volta al mese	0.219	0.211	0.174	0.009	0.028
Qualche volta l'anno	0.146	0.143	0.085	0.002	0.058
Mai	0.044	0.048	0.056	-0.005	-0.006
No ho amici	0.017	0.019	0.019	-0.003	-0.001
N.r.	0.022	0.024	0.011	-0.002	0.015

7. Confronto tra metodi di aggiustamento dell'effetto selezione

I metodi di aggiustamento dell'effetto selezione assumono l'ignorabilità dell'effetto selezione e l'invarianza dell'errore di misura determinato dal mix di tecniche nel tempo. Tale assunzione è, tuttavia, poco sostenibile per le indagini ripetute nel tempo con tecniche miste/sequenziali, dove la composizione dei rispondenti per tecnica può variare nelle successive edizioni dell'indagine, comportando variazioni nell'errore totale di misura. Per stabilizzare l'errore di misura e evitare errate valutazioni sulle variazioni dei fenomeni nel tempo, Buelens e Van den Brakel (2011) propongono un metodo basato su due calibrazioni simultanee, una rispetto a variabili ausiliarie per correggere l'effetto selezione e l'altra rispetto a livelli fissi della composizione dei campioni per tecnica.

In questa sezione sono presentati i risultati dell'applicazione di diversi metodi di aggiustamento dell'effetto selezione sul campione MM, tutti basati su calibrazioni rispetto a totali di variabili socio-demografiche (classi di età, sesso, istruzione) a

livello di area geografica. I metodi si differenziano per altri aspetti e sono: A) Calibrazione solo su variabili socio-demografiche; B) Calibrazione su variabili socio-demografiche e livelli di composizione per tecnica e tipologia comunale (Buelens e Van den Brakel, 2011) osservati (B1) e ipotetici (B2); C) Calibrazione su variabili socio-demografiche e pesi diretti corretti per l'effetto selezione con i fattori correttivi w_k ottenuti con il propensity score. I risultati dei metodi applicati al campione MM sono confrontati con le stime calibrate sugli stessi totali noti ottenute dal campione SM. Dalla tabella 5 emergono differenze nelle stime delle due indagini ma non tra quelle ottenute con diversi metodi di calibrazione.

Tabella 5 – *Stima delle distribuzioni di “Lettura dei libri negli ultimi 12 mesi” con differenti metodi (valori percentuali).*

Modalità	SM	Metodo A	Metodo B1	Metodo B2	Metodo C
No	57,81	59,92	59,00	58,66	59,92
Si	39,68	36,51	37,43	37,73	36,33
N.r	2,49	3,58	3,56	3,61	3,75

8. Conclusioni

Il disegno sperimentale che ha coinvolto l'indagine AVQ del 2017 ha consentito di effettuare una serie di analisi che hanno messo in luce l'impatto dell'introduzione della tecnica mista sulla qualità delle stime. Il principale vantaggio del disegno a tecnica mista è che risultano attenuati alcuni effetti distorsivi determinati dalla mancata risposta totale, indicando che il campione dei rispondenti rappresenta meglio la popolazione di interesse rispetto alle principali caratteristiche socio-demografiche. Tuttavia la qualità delle stime, soprattutto per alcuni indicatori, sembra essere influenzata dalla presenza di un effetto di misurazione. I risultati delle analisi, solo in parte riportati nel presente articolo, possono comunque fornire indicazioni utili alla progettazione delle successive edizioni dell'indagine, sia per la definizione dei questionari, sia per la raccolta dei dati. In tali fasi, infatti, devono essere messi in campo strumenti per limitare l'effetto misurazione, che è il principale svantaggio della tecnica mista. Le evidenze emerse in questo lavoro sono estendibili ad altre indagini sociali realizzate con tecnica web insieme a tecniche caratterizzate dalla presenza del rilevatore. In tali situazioni, azioni specifiche e mirate nella fase di raccolta dei dati sono ancor più necessarie se si tiene conto dell'eccezionalità della sperimentazione condotta su AVQ 2017.

Riferimenti bibliografici

- BUELENS B., VAN DEN BRAKEL J. A. 2015. Measurement error calibration in mixed-mode, *Sociological methods & Research*, Vol.4483, pp. 391-426.
- DE LEEUW, E. D. 2005. To mix or not to mix data collection modes in surveys, *Journal of Official Statistics*, Vol. 21(2), pp. 233–255.
- HOX J., DE LEEUW E. D., KLAUSCH T. 2015. Mixed Mode Research: Issues in Design and Analysis, *International Conference on Total survey error: improving quality in the era of big data*, Baltimore, 19-22 September.
- KLAUSCH T., SCHOUTEN B., BUELENS B., VAN DEN BRAKEL J. 2018. Adjusting measurement bias in sequential mixed-mode surveys using re-interview data, *Journal of Survey Statistics and Methodology*, Vol. 5(4), pp. 409-432.
- LUZI, O. et al. 2017. L'effetto tecnica nelle indagini mixed-mode, ISTAT, Metodi.
- MARTIN P., LYNN P. 2011. The effects of mixed mode survey designs on simple and complex analyses. *Working Paper Series*, No. 04.
- MORGAN S. L., WINSHIP C. 2009. *Counterfactuals and Causal Inference: Methods and Principles for Social Research. Analytical Methods for Social Research*, New York: Cambridge University Press.
- PARK S., KIM J. K., PARK S. 2016. An imputation approach for handling mixed-mode surveys, *The annals of Applied Statistics*, Vol. 10, No. 2, pp. 1063-1085.
- SCHOUTEN B., COBBEN F., BETHLEHEM J. 2009. Indicators for the Representativeness of Survey Response, *Survey Methodology*, 35, pp. 101–113.
- SCHOUTEN B., SHLOMO N., SKINNER C. 2011. Indicators for Monitoring and Improving Representativity of Response, *Journal of Official Statistics*, No. 27, pp. 231–253.
- SUZER-GURTEKIN Z. T., HEERINGA S. G., VALLIANT R. 2012. Investigating the bias of alternative statistical inference methods in sequential mixed-mode surveys, *Section on Survey Research Methods*, JSM 2012.
- ROSENBAUM P. R., RUBIN D. B. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, No. 70, pp. 41-55.
- VANDENPLAS C., LOOSVELDT G., VANNIEUWENHUYZE J. T. A. 2016. Assessing the use of mode preference as a covariate for the estimation of measurement effects between modes. A sequential mixed mode experiment, *Method, data, Analyses*, Vol. 10(2), pp. 119-142.
- VANNIEUWENHUYZE J. T. A., LOOSVELDT G., MOLENBERGHS G. 2010. A Method for Evaluating Mode Effects in Mixed-mode Surveys, *Public Opinion*.

SUMMARY

Assessment of mode effect in social surveys with mixed mode

For several years, the National Statistics Institute (ISTAT) has adopted several mixed-mode (MM) designs for a growing number of social surveys. This choice derives both from the need to contain costs by introducing the web technique, and, above all, from the need to improve response rates and population coverage, as the use of multiple techniques allows reaching different population targets.

However, mixed-mode introduces several issues that must be addressed both at the design phase, by defining the best collection instruments to contain the measurement error, and at the estimation phase, by assessing and adjusting the bias effects (mode effect). The surveys based on MM must be designed, in fact, considering the accuracy of the estimates that must be consistent and comparable with the analogue ones obtained in the previous survey editions. In this context the mode effect is determined by differences in measurement errors (measurement effect) due to the use of different techniques and by the different composition of the samples of respondents (selection effect).

Generally, these effects are confused and contained in the relationship between survey variable and technique: the measurement error produces differences between the averages in the different samples of respondents, while the selection effect influences the data correlation structure. The selection effect is not in itself a limit but, from an inferential point of view, it must be separated from the measurement effect, both to verify the real presence of measurement errors, and to assess whether the changes in estimates over time are real or determined by variations in the composition of the samples of respondents.

The focus of this work is the experience in the evaluation of mode effect in the experimental situation of ISTAT survey “Aspects of daily life - 2017”, a sequential web/PAPI survey for which an independent control single mode (SM) sample PAPI was planned to make an assessment of the introduction of mixed mode. The aim of the analyses presented in this work is to evaluate the impact on the survey estimates of the introduction of mixed-mode design, with respect to the previous single mode design. To this end, analyzes were carried out on the representativeness of the response of the two samples (SM and MM) and on the mode effect. The methods used to estimate the mode effect follow two different approaches, the first based on causal inference which requires the availability of auxiliary variables and the second based on instrumental variable which assumes a benchmark survey.

IL WEB-RESPONDENT DRIVEN SAMPLING PER LO STUDIO DELLA POPOLAZIONE LGBT+

Eugenia De Rosa, Claudia De Vitiis, Francesca Inglese, Alberto Vitalini

1. Introduzione

La realizzazione di indagini quantitative sulla popolazione LGBT+ (lesbiche, gay, bisessuali, trans, e tutte le altre diversità di orientamento sessuale e/o identità di genere) è fortemente condizionata dall'assenza di liste di persone di cui sono noti orientamento sessuale e/o identità di genere. Per questo tipo di popolazione, configurabile come "elusiva e rara", risulta difficile la definizione di disegni di campionamento probabilistici ottimali basati sulle liste della popolazione generale. La sensibilità del quesito sull'identità e/o orientamento sessuale caratterizza la popolazione LGBT+ come elusiva, mentre la bassa incidenza, soprattutto delle singole componenti, rispetto alla popolazione generale, ne connota la rarità. Un'altra peculiarità della popolazione LGBT+ è la sua natura composita: nell'acronimo, infatti, rientrano una pluralità di soggettività e gruppi sociali diversi caratterizzati da differenti livelli di visibilità e riconoscimento sociale.

Indagare una popolazione così complessa pone delle sfide di carattere metodologico e di campionamento. Gli Istituti Nazionali di Statistica conducono studi prevalentemente in maniera indiretta confrontando la popolazione LGB e eterosessuale su alcuni indicatori riferiti a diversi aspetti come la salute, il lavoro, gli stili di vita, ecc. Pochissimi sono gli esempi di indagini che, avendo lo scopo di assicurare la rappresentatività del campione rispetto alla popolazione generale, riescono a garantire anche la rappresentatività del sotto-campione LGBT+. A livello europeo è da segnalare il percorso avviato nel 2006 dall'Istituto UK (ONS) per la costruzione di indicatori utili all'analisi delle disuguaglianze legate all'orientamento sessuale con il «Sexual Identity Project». Il progetto ha sperimentato l'introduzione di quesiti sull'orientamento sessuale e l'identità di genere in un sistema di indagini sociali armonizzate (Integrated Household Survey) e dal 2009 fornisce dati disaggregati per la popolazione LGB.

Nella maggior parte dei casi, per studiare questo tipo di popolazione, si ricorre a tecniche di campionamento non probabilistico basate sui legami sociali, come lo snowball, oppure sui luoghi di frequentazione degli individui, come il campionamento nei luoghi di incontro della comunità LGBT+: tutte strategie che

hanno l'obiettivo di ridurre le distorsioni determinate dall'autoselezione di campioni realizzati sulla base della partecipazione volontaria.

Una tecnica di campionamento basata sui legami sociali ma di tipo probabilistico utilizzata nell'ultimo decennio è il *Respondent Driven Sampling* (RDS). A tal proposito si ricordano le indagini web realizzate in Vietnam e Svezia sui gay per studiare il rischio di trasmissione dell'HIV (Bengtsson et al. 2012; Strömdahl et al. 2015), l'indagine sulle donne transgender adulte (Wesson et al., 2017) a San Francisco e lo studio pilota condotto in Nuova Zelanda su gay e bisessuali per indagare le malattie trasmissibili sessualmente (Ludlam et al., 2015).

La natura probabilistica dei campioni RDS deriva dal fatto che la tecnica snowball è combinata con un modello matematico in cui il processo di reclutamento è formalizzato come una catena markoviana (Heckathorn, 1997, 2002). Sotto ipotesi molto stringenti, il campionamento RDS consente di fare inferenza sulla struttura della rete e da questa ottenere stime non distorte sulla popolazione di interesse.

In questo lavoro si esplorano potenzialità e criticità dell'applicazione del RDS nello studio di carattere sperimentale che l'ISTAT intende effettuare nell'ambito del progetto di ricerca su *Discriminazioni sul lavoro per orientamento sessuale e identità di genere e diversity policies nelle imprese* realizzato in collaborazione con l'Ufficio Nazionale Antidiscriminazioni Razziali (UNAR). Il progetto intende indagare il fenomeno discriminatorio su scala nazionale da più prospettive e livelli con una serie di indagini rivolte a differenti popolazioni target: le persone LGBT+, i datori di lavoro (le imprese) e i principali stakeholder nazionali. Con riferimento specifico alla parte del progetto riguardante l'esperienza delle persone LGBT+ rispetto al mondo del lavoro si è voluto tener conto dell'eterogeneità e pluralità di soggettività LGBT+ prevedendo la realizzazione di un'indagine sulle persone in unione civile selezionate dalla lista anagrafica comunale (LAC), un'indagine sugli utilizzatori trans di servizi/sportelli dedicati e membri di associazioni trans, e un'indagine sperimentale, con tecnica WebRDS (Wejnert e Heckathorn, 2008; Scheim et al., 2016; Vitalini, 2012), indirizzata a un segmento di persone LGBT+ non raggiunte tramite le due indagini precedenti (De Rosa e Inglese, 2018). Tale sperimentazione si pone l'obiettivo di indagare segmenti della popolazione di interesse con caratteristiche diverse da quelle delle persone in unione civile e degli utilizzatori di servizi/sportelli trans o associati, che possono essere poco rappresentati dalle altre due indagini.

Il presente articolo si struttura nel seguente modo. Si illustra la tecnica di campionamento RDS (§ 2). Si presenta successivamente come l'ISTAT intende implementare il metodo nell'ambito del progetto sulle discriminazioni lavorative verso le persone LGBT+ evidenziando una serie di questioni e sfide legate a tale implementazione (§ 3). Tali sfide rinviano alla complessità della popolazione oggetto di indagine e all'assunto base del RDS riguardante la connessione della

popolazione indagata (§ 3.1), alla definizione della popolazione target (§ 3.2) e del disegno WebRDS dove una particolare attenzione è posta alla garanzia della riservatezza delle persone coinvolte nell'indagine (§ 3.3). Nelle conclusioni, infine, vengono evidenziati i vantaggi dell'applicazione del RDS come metodo di campionamento basato sul web e le criticità legate, più in senso stretto, all'impianto teorico del RDS (§ 4).

2. Il Respondent Driven Sampling

2.1. Elementi base e ipotesi del RDS

La particolarità del respondent driven sampling (RDS) consiste nel combinare la tecnica snowball, in cui il campione si costruisce utilizzando i nominativi forniti da reclutatori iniziali (semi) e dai reclutati/reclutatori successivi (definiti nodi), con un modello matematico che formalizza il processo di reclutamento. Sotto certe condizioni il processo di reclutamento, infatti, può essere formalizzato come una catena markoviana, ovvero come un processo probabilistico (Heckathorn, 1997, 2002). Il processo di reclutamento si sviluppa in ondate generate a partire dai reclutatori iniziali, fino a raggiungere una condizione di equilibrio in cui le probabilità di inclusione delle unità campionarie si stabilizzano. I semi sono scelti in modo non casuale sulla base del criterio della differenziazione e dell'abilità di reclutare.

I dati raccolti durante il processo di campionamento sono utilizzati per fare inferenza sulla struttura della rete sociale e per ottenere stime non distorte sulla popolazione di interesse. Le proprietà dei nodi (rispondenti), le informazioni su chi recluta chi (matrice di reclutamento) e le dimensioni della rete personale degli intervistati (grado/numero di legami) costituiscono la base per generare inferenze sulle caratteristiche della popolazione. Sotto ipotesi molto stringenti, necessarie al trattamento probabilistico del campione, gli stimatori RDS sono asintoticamente non distorti (Salganik e Heckathorn, 2004).

Le ipotesi sottostanti il modello matematico riguardano sia la struttura della rete sia il campionamento. La rete della popolazione target deve essere abbastanza densa e connessa in modo che ogni nodo sia raggiungibile da ogni altro nodo. Inoltre, la popolazione non deve essere segmentata a tal punto che le catene di reclutamento rimangano intrappolate entro specifici sottogruppi altrimenti l'equilibrio non si raggiunge. I partecipanti all'indagine devono intrattenere relazioni di tipo simmetrico e devono riconoscersi l'un l'altro come membri della popolazione di riferimento (network non orientato). Il numero di legami fra i membri deve essere sufficientemente elevato da sostenere i processi di reclutamento (lunghe catene di reclutamento che attraversano più ondate) per assicurare che ogni componente della popolazione abbia una probabilità diversa da zero di entrare a far parte del campione.

Le ipotesi sul campionamento riguardano la selezione con reimmissione delle unità, l'accuratezza della stima dei legami, la casualità del reclutamento. Nello schema 1 sono riportate le assunzioni sulla struttura della rete e sul campionamento ponendole in relazione alla loro specifica funzione (Gile e Handcock, 2010).

Schema 1 – Sintesi delle assunzioni.

Funzione	Assunzioni sulla struttura della rete	Assunzioni sul campionamento
1. Modello Random Walk	Rete sociale ampia	<i>Campionamento con reimmissione (SWR)</i> Catena singola non ramificata
2. Rimuovere la dipendenza dai semi iniziali	<i>Network (grafo) connesso</i> Omofilia debole	Sufficiente numero di ondate
3. Caratteristiche dei rispondenti	<i>Reciprocità dei legami</i>	<i>Stima accurata dei legami</i> <i>Reclutamento casuale</i>

Lo schema riporta sulla riga del punto 1 le ipotesi che sono necessarie per l'approssimazione del processo di campionamento con una passeggiata aleatoria (random walk model). In corrispondenza del punto 2 si trovano invece le assunzioni necessarie a ridurre l'influenza dei semi del campione iniziale sulle stime finali. Sulla riga del punto 3 sono riportate le ipotesi necessarie per la stima delle probabilità di inclusione, legate al comportamento dei rispondenti, e, quindi alla precisione delle risposte fornite sulla dimensione della loro rete sociale. Le ipotesi indicate in corsivo sono necessarie ad ottenere stimatori RDS asintoticamente non distorti in un modello semplificato. La violazione delle ipotesi comporta comunque il rischio di produrre stime distorte e con elevata variabilità.

2.2. Probabilità di inclusione e stimatore

Per la definizione delle probabilità di inclusione delle unità del campione si deve tener conto della dimensione del nodo in termini di legami, delle ondate di reclutamento e dello stato di equilibrio del processo di reclutamento (Xin, 2013).

Se l'individuo i è selezionato nell'ondata campione t , la probabilità che ciascun nodo j (individuo) sia selezionato nell'ondata $t + 1$ è

$$Pr_{i \rightarrow j} = \begin{cases} \frac{1}{d_i} & \text{se esiste un legame tra } i \text{ e } j \\ 0 & \text{altrimenti} \end{cases} \quad (1)$$

dove d_i è il grado dell'unità i (dimensione del nodo in termini di legami). Il RDS può essere modellato come un processo di Markov con matrice di transizione

$$T = \begin{pmatrix} 0 & e_{12}/d_1 & \dots & e_{1N}/d_1 \\ e_{21}/d_2 & 0 & \dots & e_{2N}/d_2 \\ \vdots & \vdots & \ddots & \vdots \\ e_{N1}/d_N & e_{N2}/d_N & \dots & 0 \end{pmatrix}, \quad (2)$$

con $e_{ij} = 1$ se esiste un legame tra l'individuo i e l'individuo j , e $e_{ij} = 0$ altrimenti.

Si può dimostrare (Xin, 2013) che in un grafo connesso e non orientato la catena markoviana è irriducibile e la probabilità di inclusione di ciascun nodo, una volta raggiunto la condizione di equilibrio, è proporzionale al corrispondente grado,

$$Pr_i = \frac{d_i}{\sum_{j=1}^N d_j}. \quad (3)$$

Le probabilità di inclusione delle unità campionarie (nodi) sono usate per ottenere lo stimatore di Hansen-Hurwitz in cui le osservazioni sono ponderate con l'inverso delle probabilità di inclusione. La stima della proporzione di individui appartenenti al generico gruppo A della popolazione U può essere ottenuta applicando lo stimatore Volz-Heckathorn, noto come VH o RDSII (Volz e Heckathorn, 2008),

$$\hat{p}_A = \frac{\sum_{i \in A \cap U} d_i^{-1}}{\sum_{i \in U} d_i^{-1}}. \quad (4)$$

3. L'implementazione del WebRDS nell'indagine ISTAT sulle discriminazioni lavorative nei confronti delle persone LGBT

3.1. Connessione e complessità della popolazione LGBT+ in Italia

L'applicazione del RDS richiede come presupposto fondamentale che la popolazione oggetto di indagine abbia una struttura a rete chiaramente identificabile. L'eterogeneità e la complessità della popolazione LGBT+, in cui rientrano persone con differenti orientamenti sessuali e affettivi, identità di genere (binarie e non binarie), persone queer, intersessuali, rendono particolarmente delicata la fase di definizione teorica e operativa della popolazione di indagine. Tale fase implica una scelta riguardante il tipo di reti o legami sociali che si intendono considerare per operationalizzare la "connessione o connettività" della popolazione di interesse e, in primo luogo, la scelta da parte del ricercatore di considerare o meno tale popolazione come un unico aggregato o piuttosto composto da più aggregati/popolazioni.

Un primo strumento analitico utile è il concetto sociologico e contemporaneo di comunità che, applicato alla popolazione LGBT+, indica un aggregato che si costruisce "sulla base di dimensioni che vanno oltre il mero radicamento territoriale" (Corbisiero, 2013, 2016) laddove il principale fattore di coesione sociale è

rappresentato dal rifiuto dell'eteronormatività, cisnormatività e dei vari binarismi di genere. Il tratto principale di tale “sentire comune o sentimento di appartenenza” può essere rintracciato in un ampio progetto politico e culturale che si costituisce per differenza rispetto al mondo eteronormativo e cisnormativo, riconducibile, seppure non completamente, alle dimensioni dell'attivismo e della rivendicazione di diritti di LGBT+. Tale concetto di comunità ha quindi la capacità di includere alcuni, coloro che sentono l'appartenenza all'universo LGBT+, e di escluderne altri, ovvero coloro che, pur essendo LGBT+, non aderiscono a un generico sentire comune.

Nell'acronimo LGBT+ tuttavia confluiscono differenti gruppi sociali. Una prima distinzione è tra persone LGB+ e persone trans; due gruppi solo in parte sovrapponibili (trans non eterosessuali) laddove il primo si costruisce sulla base dell'orientamento sessuale e il secondo sul fattore dell'identità di genere. Tale distinzione non esaurisce la molteplicità interna al mondo LGBT+ dove coesistono gruppi con rivendicazioni identitarie, istanze, interessi e risorse differenti, talvolta in conflitto tra loro, tanto che è opportuno parlare di sotto-comunità o più comunità (es. comunità lesbica, comunità bisessuale, comunità trans, eccetera). Se per alcuni gruppi la presenza e visibilità di una comunità di riferimento è più forte (es. gay), per altri lo è meno come per la comunità bisessuale o transgender non med. Alcune comunità poi si caratterizzano per la conquista di spazi pubblici, trovando nella città livelli più alti di visibilità, prestigio e riconoscimento sociale (ibidem); di contro altre hanno costruito il proprio senso di appartenenza prevalentemente sul web tanto che la comunità virtuale sembra strutturare le relazioni maggiormente di quanto accade nella comunità reale. È questo il caso della comunità bisessuale e delle persone con identità di genere fluide e non binarie.

Oltre all'aggregato di persone il cui legame risiede nell'appartenenza ad una o più comunità LGBT+, per l'implementazione del RDS è possibile far riferimento anche ad altre reti e relazioni sociali, più circoscritte. Si pensi, ad esempio, all'esistenza di gruppi laddove le relazioni si costruiscono intorno ad attività comuni, di tipo culturale, e sportivo (es. Libera rugby, Gruppo pesce nuoto) o gruppi legati al turismo. Da considerare inoltre che la vita digitale assume un'importanza crescente con la diffusione di social network e app dedicate alla conoscenza reciproca, allo scambio di informazioni su temi specifici e su servizi LGBT+ (es. GayFriendlyItaly, piattaforma italiana dedicata al turismo LGBT+).

Nel complesso, ai fini dell'applicazione del RDS per lo studio delle popolazioni LGBT+ la fase di “formative study” acquista un ruolo fondamentale nel disegno di indagine.

3.2. *Il Formative study*

Nel progetto ISTAT la fase di formative study è indirizzata a dirimere alcune questioni preliminari all'implementazione del RDS, in particolare con riferimento:

alla definizione della popolazione (o delle popolazioni di interesse) che, come detto precedentemente, deve presentare una chiara struttura a rete; alla scelta dei reclutatori iniziali (semi) del campione RDS, soggetti conosciuti e visibili nella popolazione selezionata e che possono contare su numerosi legami sociali.

Per tali scelte, oltre all'analisi della letteratura accademica e grigia e dei dati quantitativi disponibili, si rende necessaria la raccolta diretta di informazioni qualitative attraverso colloqui e interviste in profondità a testimoni privilegiati ed esperti tematici (su LGBT+ e discriminazione) e consultazioni con le associazioni LGBT+, appartenenti e non al «Tavolo permanente delle associazioni LGBT per la promozione dei diritti e la tutela delle persone LGBT» istituito presso UNAR. Tali approfondimenti qualitativi sono supportati da strumenti di raccolta e analisi appositamente predisposti. Il coinvolgimento di esperti e associazioni è fondamentale per ridefinire la popolazione di riferimento e addivenire a una definizione operativa di comunità/rete LGBT+ o parti di essa (e specularmente individuare gruppi/soggetti che non appartengono a tale comunità/rete). In particolare, le associazioni, da una parte forniranno un quadro degli iscritti alle associazioni (caratteristiche socio-demografiche, composizione per orientamento sessuale e identità di genere) segnalando l'esistenza di gruppi sottorappresentati dall'associazionismo e dall'attivismo LGBT+, dall'altra daranno indicazioni su livello di connessione degli associati (interno all'associazione e tra associazioni), e più in generale della popolazione LGBT+, sulla diffusione di un senso di appartenenza ad una comunità o sotto-comunità LGBT+, sull'esistenza di altre attività comuni o reti caratterizzanti il mondo LGBT+ non riconducibile all'associazionismo che di fatto coglie una parte minoritaria di popolazione.

Il contributo di esperti e associazioni è fondamentale per la scelta dei semi o reclutatori iniziali del campione RDS. Partendo da alcuni criteri stabiliti dal team di ricerca per la scelta dei semi (es. avere un'ampia rete sociale, essere ben conosciuti dalla popolazione o popolazioni di riferimento, essere molto motivati alla ricerca, e altri criteri basati su caratteristiche socio-demografiche), associazioni ed esperti proporranno dei nominativi scelti tra soggetti particolarmente visibili e conosciuti nella/e reti individuate in fase di definizione della/e popolazione/i di interesse. Per superare i limiti degli studi che si concentrano sui soci di organizzazioni LGBT+ tali semi non dovranno appartenere esclusivamente alla realtà dell'associazionismo e attivismo (es. blogger e influencer in rete maggiormente seguiti, appartenenti ad associazioni sportive LGBT+).

Infine verranno predisposti quesiti ad hoc nel questionario rivolto agli uniti civilmente e nelle interviste agli stakeholder con l'intento di raccogliere informazioni, oltre che sulla vita associativa, sui tipi di relazioni sociali e sulla vita digitale delle persone LGBT+. Integrando tali informazioni con le altre fonti esistenti

l'obiettivo è ottenere una prima tipologia di gruppi appartenenti alla popolazione LGBT+.

3.3. *Disegno WebRDS e raccolta dati*

L'applicazione del WebRDS ha una serie di implicazioni di cui si deve tener conto, *in primis* la definizione di regole per garantire l'anonimato e la tutela della riservatezza delle persone coinvolte nell'indagine. Diversi sono gli strumenti e i dispositivi che devono essere messi in campo per garantirne il funzionamento, a partire dalla lista degli individui del campione RDS iniziale, dal questionario, fino ad arrivare ai programmi per la gestione del processo di reclutamento e della somministrazione del questionario on-line. La lista degli individui del campione RDS iniziale si avvarrà del contributo di esperti e associazioni per l'individuazione dei nominativi e sarà formata da coloro che accettano di partecipare all'indagine.

La formulazione del questionario, nella versione per il RDS, assume particolare importanza poiché, oltre alle sezioni tematiche, deve contenere domande relative al numero di legami che il rispondente ha nell'ambito della popolazione oggetto di interesse e alla conoscenza della persona che lo ha reclutato: informazioni utili alla stima della probabilità di inclusione dei nodi e al controllo della reciprocità delle relazioni tra le persone che reclutano e quelle che sono reclutate. Le pagine del questionario sono precedute da una pagina di consenso informato. All'inizio del questionario è valutata l'idoneità a partecipare all'indagine, le persone non eleggibili, ovvero non appartenenti alla popolazione di interesse, e le persone che hanno già risposto ad un invito precedente, non procedono con la compilazione del questionario. Alle persone risultate idonee sono fornite informazioni dettagliate sulla ricerca e, solo dopo aver fornito il consenso informato scritto, procedono alla compilazione del questionario. Se la persona risultata idonea non fornisce il consenso informato scritto, è indirizzata ad una apposita pagina in cui si chiede di fornire le ragioni della mancata partecipazione. Informazioni più dettagliate sul progetto di ricerca sono fornite in un link separato, accessibile in qualsiasi momento durante l'indagine. I loghi dell'ISTAT, dell'UNAR e delle associazioni che collaborano al progetto sono visualizzati su ogni pagina del questionario e ogni e-mail di invito, e collegati ai siti web associati.

L'intervistato che completa il questionario è indirizzato a una pagina in cui gli si chiede di invitare tre o quattro suoi contatti, persone che devono essere in possesso di specifiche caratteristiche affinché partecipino all'indagine. In questa pagina per ognuno dei possibili contatti è presentato il testo di un'e-mail di invito contenente un link per l'accesso al questionario. Il link, generato automaticamente dal sistema, è unico per ogni partecipante. Il codice univoco del link permette di controllare la partecipazione ripetuta dei rispondenti e di ricostruire le catene di reclutamento.

Sapere “chi ha reclutato chi” è, infatti, un’informazione indispensabile (§ 2.2) per il calcolo della situazione di equilibrio del processo di reclutamento.

L’indagine web sarà completamente anonima, a parte i reclutatori (semi) del campione RDS iniziale. Infatti, per garantire la riservatezza delle persone coinvolte nell’indagine, evitando la richiesta di indirizzi e-mail a cui inoltrare il questionario, il reclutatore stesso “taglia e incolla” il testo della e-mail e lo spedisce ai suoi contatti tramite whatsapp, messenger, e-mail, facebook. Il reclutatore può aggiungere un testo personale ad ogni messaggio. Lo svolgimento online dell’indagine consente di individuare eventuali duplicati (individui con più indirizzi e-mail) con controlli su indirizzi IP. Per garantire il rispetto delle norme di riservatezza, gli indirizzi IP sono convertiti in un codice anonimo univoco (uno per ogni indirizzo IP) utilizzando un algoritmo di crittografia unidirezionale. Gli indirizzi IP originali sono cancellati e tutte le comunicazioni tra l’utente e il server devono essere crittografate. Anche il database online dei dati è criptato e protetto da password. L’applicativo utilizzato per gestire il processo di reclutamento, la rilevazione online e il salvataggio dei dati registrati è LimeSurvey (www.limesurvey.org).

4. Conclusioni

L’indagine sulle persone LGBT+ con Web Respondent Driven Sampling che si intende realizzare nell’ambito del progetto di ricerca ISTAT-UNAR su *Discriminazioni sul lavoro per orientamento sessuale e identità di genere e diversity policies nelle imprese* come detto si configura come sperimentale. Con tale sperimentazione si intende raggiungere, con un approccio quantitativo, una parte di popolazione non coperta (o non sufficientemente coperta) dalle due indagini che verranno condotte precedentemente e che sono rivolte a specifici segmenti della popolazione LGBT+ (le persone in unione civile individuate dalle liste anagrafiche comunali e le persone trans intercettate a partire da associazioni e sportelli/servizi dedicati a tale utenza).

Tale insieme di indagini è pensato in modo da garantire una buona rappresentazione dell’eterogenea popolazione LGBT+ e quindi offrire un quadro più fedele del fenomeno delle discriminazioni lavorative percepite e subite dalle persone LGBT+. Nello specifico la rilevazione con tecnica RDS mira a rappresentare e coinvolgere nella ricerca gruppi di popolazione sottorappresentati nell’indagine sugli uniti civilmente (come ad esempio le donne lesbiche e i più giovani dato che gli uniti per il 70% sono in coppie di uomini e nel complesso hanno un’età media di 49 anni) e più in generale gruppi di persone non già raggiunti (persone trans e unite civilmente) e che abbiano un qualche legame tra loro. L’esistenza di gruppi con un buon livello di connessione e una chiara struttura a rete fornisce una possibilità

concreta all'applicazione del campionamento WebRDS che ha il vantaggio di superare i limiti dei campioni di convenienza e rendere possibile delle stime.

La scelta di procedere con tecnica web circoscrive l'indagine sperimentale alla popolazione con competenze digitali, escludendo di fatto le persone che non conoscono/utilizzano i dispositivi informatici. Dall'altra parte l'acquisizione di informazioni tramite questionario autocompilato web meglio si adatta alla delicatezza dei quesiti sia sull'orientamento sessuale e l'identità di genere sia sulle discriminazioni in ambito lavorativo. La versione web del RDS risulta poi particolarmente conveniente sia in termini di costi che per il fatto che la popolazione raggiungibile è potenzialmente più ampia rispetto a quella che è possibile cogliere con altre tecniche di campionamento "non standard". Inoltre, da un punto di vista operativo, l'utilizzo del web permette di individuare eventuali duplicati, rischio alquanto contenuto quando non è previsto l'utilizzo di incentivi ai reclutatori.

A fronte dei vantaggi del campionamento RDS nella formulazione web precedentemente indicati, vi sono alcune criticità che si possono presentare già in fase di implementazione e che possono determinare la violazione di alcune ipotesi e l'allontanamento dalle condizioni ideali del modello sottostante tale tecnica. Una scelta oculata di semi iniziali capaci di generare lunghe catene di reclutamento e una formulazione accurata dei quesiti attraverso cui si rileva la reciprocità dei legami o la dimensione della rete sociale di ciascun rispondente (variabile importante per la stima delle probabilità di inclusione) possono contribuire al raggiungimento di risultati con una qualità accettabile. In fase di stima, inoltre, è opportuno analizzare la violazione delle ipotesi del modello RDS, applicare, attraverso simulazioni, metodi per identificare le ipotesi le cui violazioni hanno un'influenza significativa e, infine, sviluppare stimatori meno sensibili alla violazione delle ipotesi.

Da qui la consapevolezza della delicatezza sia della fase di implementazione del metodo sia della fase di stima per cui diventa fondamentale progettare e attuare azioni per assicurare la buona qualità dei risultati dell'indagine a partire dalla fase di progettazione degli strumenti di rilevazione.

Riferimenti bibliografici

- BENGTSSON L., LU X., NGUYEN Q. C., CAMITZ M., LE HOANG N., NGUYEN T. A., LILJEROS F., THORSON A. 2012. Implementation of Web-Based Respondent-Driven Sampling among Men Who Have Sex with Men in Vietnam. Published: <https://doi.org/10.1371/journal.pone.0049417>
- CORBISIERO, F. 2016. *Napoli DiverCity. Pratiche, prassi e metodi di ricerca sulla popolazione LGBT*. Roma: Carocci editore.

- CORBISIERO, F. 2013. Comunità omosessuali. Le scienze sociali sulla popolazione LGBT. Franco Angeli Editore.
- DE ROSA E., INGLESE F. 2018. Disuguaglianze e discriminazioni nei confronti delle persone LGBT: quale contributo della statistica ufficiale? *Rivista italiana di economia, demografia e statistica*, Vol. 72 , No. 4.
- GILE K. J., HANDCOCK M. S. 2010. Respondent-Driven Sampling: An assessing of current methodology, *Sociol Methodol*, Vol. 40(1), pp. 285-327.
- HECKATHORN, D. D. 1997. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations, *Social Problems*, Vol. 44, No. 2, pp. 174-199.
- HECKATHORN, D. D. 2002. Respondent-Driven Sampling II: Deriving Valid Population Estimates From Chain Referral Samples of Hidden Populations, *Social Problems*, Vol. 49, No. 1, pp.11-34.
- LUDLAM A., SAXTON P., DICKSON N. P., ADAMS J. 2015. Respondent-driven sampling among gay and bisexual men: experiences from a New Zealand pilot study, *BMC Res Notes*, Vol 8, No. 549.
- SALGANIK M. J., HECKATHORN D. D. 2004. Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling, *Sociological Methodology*, Vol. 34, pp. 193-239..
- SCHEIM A. I., BAUER G. R., COLEMAN T. A. 2016. Sociodemographic Differences by Survey Mode in a Respondent-Driven Sampling Study of Transgender People in Ontario, Canada, <https://doi.org/10.1089/lgbt.2015.0046>.
- STRÖMDAHL S., LU X., BENGTSSON L., LILJEROS F., THORSON A. 2015. Implementation of Web-Based Respondent Driven Sampling among Men Who Have Sex with Men in Sweden, <https://doi.org/10.1371/journal.pone.0049417>.
- VITALINI, A. 2012. *L'uso delle reti sociali per la costruzione di campioni probabilistici*. Roma: Aracne.
- VOLZ E., HECKATHORN D. D. 2008. Probability-Based Estimation Theory for Respondent-Driven Sampling, *Journal of Official Statistics*., Vol. 24, No. 1, pp. 79–97.
- WEJNERT C., HECKATHORN D. D. 2008. Web-Based Network Sampling Efficiency and Efficacy of Respondent-Driven Sampling for Online Research, *Sociological Methods & Research*, Vol. 37, No. 1, pp. 105-134.
- WESSON P., QABAZARD R. F., WILSON ERIN C., MCFARLAND W., FISHER R. H. 2013. Estimating the population size of transgender women in San Francisco using multiple methods, pp. 107-112, <https://doi.org/10.1080/15532739.2017.1376729>.
- XIN, L. 2013. Respondent-Driven Sampling: Theory, Limitations & Improvements. Karolinska Institute, Stockholm: US-AB.

SUMMARY

Web-Respondent Driven Sampling for the study of LGBT+ population

Quantitative surveys on the LGBT+ population (lesbian, gay, bisexual, transsexual, and all types of sexual orientation and/or gender identity) are difficult to carry out mainly due to the absence of lists of people whose sexual orientation and/or gender identity are known.

For this type of population, termed “elusive and rare”, it is difficult to define optimal probability sampling designs based on lists of the general population. Another peculiarity of the LGBT+ population is its composite nature, as it includes a plurality of subjectivities and different social groups characterized by different levels of visibility and social recognition. Investigating a complex population such as LGBT+ poses sampling challenges.

There are very few examples of surveys on the general population that can also guarantee representativeness of a target population sample.

A non-standard probability sampling technique based on the social network of individuals, Respondent Driven Sampling (RDS), has been used in the last decade in surveys on subgroups of the LGBT population. RDS consists of combining the snowball technique, in which the sample is constructed using the names provided by initial recruiters, with a mathematical model that formalizes, under certain conditions, the recruitment process as a Markov chain, or as a probabilistic process. The data collected during the sampling process are used to make inferences about the structure of the social network and obtain from this unbiased estimates on the population of interest.

In this work the potentialities and limits of the application of RDS are explored in the experimentation that ISTAT intends to carry out in the research project on “Discrimination at work for sexual orientation and gender identity and diversity policies in enterprises” in collaboration with the National Office for Racial Discrimination (UNAR). The survey on a large segment of the LGBT population aims to investigate sub-populations with characteristics different from those of persons in civil union and users of Transgender services, the subject of two other surveys envisaged by the project.

The article presents the challenges for the application of the RDS method arising from the complexity of the population and from the assumptions about links within the population. Moreover, the strategies for the implementation of the RDS as a web-based sampling method are outlined.

Eugenia DE ROSA, Istat, eugenia.derosa@istat.it

Claudia DE VITIIS, Istat, devitiis@istat.it

Francesca INGLESE, Istat, fringles@istat.it

Alberto VITALINI, Istat, vitalini@istat.it

THE ITALIAN CONTINUOUS CENSUSES: MANAGING THE COMPLEXITY AMONG REGISTERS, STATISTICS AND TERRITORIES¹

Antonella Bernardini, Daniela Bonardo, Alessandra Dentini, Maura Giacummo,
Matteo Mazziotta, Alessandra Preti

1. Introduction

The recent process of modernization of the Italian Institute of Statistics (Istat) is based on the increasing use of databases from administrative sources in order to support and integrate sample surveys; this reduces the statistical burden for respondents (citizens and economic units) and the costs borne by the community, guaranteeing the production of statistical information on an annual, biennial or three-year basis, depending on the census considered, and no longer as ten years as in the traditional census model.

The transition from the ten-year census to the continuous one represents an element of great innovation, with a huge reduction in the costs of the operation and an increase in the frequency of production and dissemination of data. The transition to a new census model is made possible by the acquisition, processing and use for statistical purposes of administrative sources that generate, through data validation processes, statistical registers updated with high time frequency.

The challenge is to maintain a high quality of the data by increasing the frequency of publications. This objective can be reached by breaking the classical scheme based on the exhaustive survey, the counting of population units, the estimation of the non-sampling error.

The use of administrative sources and sample surveys to support the correction of the sources themselves represents a real revolution in official statistics. The actors on stage increase and the functions differ, creating a new work paradigm that needs a different organizational structure than in the past. The simultaneous conduction of some censuses needs to share competences, roles and functions in order to create economies of scale that limit the waste of time and resources.

¹ Acknowledgments: The paper is the result of the common work of the authors. In particular, Antonella Bernardini has written sections 2.1 and 4.3, Daniela Bonardo has written sections 1 and 4.2, Alessandra Dentini has written sections 3 and 4.1, Maura Giacummo has written sections 2 and 4.4, Matteo Mazziotta has written section 2.3 and 4.5 and Alessandra Preti has written sections 2.2 and 4.

The paper aims to present some characteristics of the organizational management of the five Italian censuses (Population, Business, Agriculture, Public Institutions and Non-profit Institutions) in which a new organizational structure composed by the authors of this paper and called "Contact Group" aims to standardize common processes of the censuses allowing the connection between different functions and a large internal diffusion of strategic information.

2. New Strategies and Old Issues

The old censuses were characterized by an organizational scheme based on many actors that interacted with Istat: local authorities, trade associations, interviewers, universities, etc.

The survey covered all the units of the reference population, occurring at the same time and every 10 years. The statistical errors of a census were calculated at the end of the survey and they are called non-sampling errors, since they do not depend on the partial selection of units from the population (Istat, 2013; Istat, 2016).

The new censuses are very different: they do not necessarily cover all the units at the same time, they are also based on sample surveys and statistical registers and a frequency of results publication every year (or two or three, depending on the census) is provided. In fact, with the new continuous Censuses information on the main socio-economic characteristics of the country will always be more detailed and continuously updated. Thanks to the integration of administrative data with those provided by the surveys, costs will be significantly reduced compared to the past and more useful information will be obtained to plan more efficient interventions and services.

2.1. Households

In October 2018 the continuous Census of Population and Housing started. For the first time the surveys are annual and "on sample": they involve 1 million 400 thousand families and about 2,800 municipalities. By 2021, all the municipalities of Italy will have participated at least once in the surveys. The areal sample survey is useful to enumerate the individuals who really insist on the territory and, from methodological point of view, it represents the real value to compare with the individuals of the population register. In the past, in many developed countries,

areal surveys have measured the quality of the population census and this type of survey is called Post Enumeration Survey (PES). In Italy there is an excellent tradition of conducting these investigations: the experience gained can help to design and carry out new areal surveys to constantly improve the quality of the population register.

2.2. *Companies*

Starting from 2019 also the Business Census becomes continuous: every three years, to increase the quality of the information offer, the information of more than 250 thousand productive units will be collected. In this case, the reference register has a multi-year history; in fact, the statistical register of active companies (ASIA) has been perfected over time and today it represents an essential reality for all surveys on Italian companies. From this point of view, the “road” of the continuous census is simpler than the population one because the detection units are simpler to capture and to collect through administrative sources.

The first continuous Census of Non-Profit Institutions was already implemented in 2016, with the survey of about 43 thousand units. The next edition is scheduled for 2019 and every three years (2022, 2025 and so on) with a sample size of 70 thousands units. There are positive prospects for administrative data in support of the construction of sectorial registers (third sector reform toward a single register). The data collection technique is based on mixed mode in which the analytical information is structured by stable parts and variable modules.

The continuous Census of Public Institutions has a biennial frequency and is the only one, among the new censuses projected by Istat, not to be a sample; in fact it involves all the institutions present in the country (about 13 thousands institutions and related local units). The second survey of the new course ended at the end of 2018 and the next will be in 2020. The collection of analytical information on structure, operation, organizational, innovation, human capital, relations with the market is based on stable parts and variable modules. The integrated publication of the results with information from basic and extended registers will be progressively implemented. In fact, the extended register of Public Institutions (Frame PI) is in progress. The data collection technique is CAWI and then the cost is reduced.

The Agricultural Census will change starting in 2021: the last general census of agriculture with ten-year frequency, in which all Italian farms will participate, is in fact planned for 2020. Subsequently, the farm register will be corrected with sample surveys every three years (2023, 2026 and so on). The Integrated publication of the results with information from basic and extended registers will be progressively implemented on an annual basis. The size of 2020 survey is about

1.5 millions of farms. In the future, the burden and the cost reduction on respondents and the integrated use of multiple collection techniques will be the main characteristics of the new world of agricultural surveys.

2.3. Errors

In this context, the statistical error will be composed of a sample and a non-sample part and that the sources of the error are generated, in addition to the actors involved, also by the use of new administrative data. For this reason, it seems essential to focus attention on quality checks, some of which are included in the surveys themselves. Such a great revolution must be supported by a new management of skills and roles within the complex process of producing the final data, providing active management structures to solve problems and to facilitate dialogue between the actors involved.

3. Management of “Functional Processes”

The Contact Group, established in the spring of 2017, is made up of seven Istat researchers and technologists who present different skills in order to cover different scientific and organizational functions. In fact, skills range from management to information technology, from statistics to accounting; furthermore, the specificities of each member can adapt to the specificities of such diverse censuses. The main function is to provide a management approach to the most complex phases by finding a sustainable solution that can be used for all censuses.

Although the specificities of each census are very different, the goal of the Contact Group is to find common points in order to converge the pillars defining the general strategy. In particular, some urgent recommendations are proposed:

- outlining in a single Gantt the planning processes of the various censuses;
- developing shared guidelines for the General Census Plan (PGC) that take into account the continuity of the new censuses;
- synchronizing some tender’s procedures, in particular those concerning IT equipment or in any case provide for the optimization of the invested resources;
- encouraging the circulation of methodological notes and information materials on the processes that generate statistical registers;
- reasoning on the transversal and on the organizational optimization of the Advisory Committees;

- evaluating the resources to be committed on the various processes, in particular for the study of procedures for exploiting Big Data and data for the geolocation of the units;
- accurately defining the development of communication plans;
- deepening and coordinating the methods of data publication and more generally planning the IT management of the censuses.

For almost two years, the Contact Group has assisted the Istat structures to carry out their tasks, has organized technical meetings and organized seminars on frontier issues: the group has dealt with the general census budget, produced technical reports for specific commissions and provides methodological support in order to calculate the statistical error of the censuses. Today it has become a reference point for the exchange of relevant information towards a new era for official statistics.

4. Transversal Functions

We are aware of the different states of progress about the design of the five censuses, as well as of the different conditions in which we can realize them (specific sources, alliances and norms, experiences and internal resources etc.).

With different speed we must however establish a point in time in which to converge on the foundations that define the strategy of continuous censuses. The year on which to focus and succeed in aligning the founding elements of the strategy is 2021.

As mentioned in the previous sections, the decisive role of this great revolution in official statistics is the construction of statistical registers and the relations among themselves, samples and census surveys. The basic register is the result of a vast whole, to be continuously fed even with new acquisitions of administrative sources, together with periodic coverage surveys in order to evaluate the quality (no-sampling error). Since the target is to define the uniqueness of the size of the populations of interest in order to support the statistical Institute's production and the analysis of the processes. In fact, the identification of the fundamental administrative source (s), whose availability must be stable, foreseen by law, and whose correction on the basis of the basic Register may be of institutional interest (for example, municipal register, Chambers of Commerce, Land Register and so on).

In this context, some transversal functions are common to the five censuses and the work of the Contact Group is useful for sharing the different experiences in order to optimize phases and processes of the survey. The main transversal functions are listed below.

4.1. Administrative aspects

The Contact Group assisted the Istat organizational structures in order to achieve the planning objectives of some important administrative aspects concerning the censuses.

The first fundamental task was a survey of the status of administrative procedures such as the budget, the tenders for the awarding of services, the General Census Plans (PGC's) and other legal aspects concerning the obligation to reply to the questionnaire and the penalties for non-respondents. In the month of May 2017 the Contact Group has begun the recognition of the financial needs for the realization of the permanent censuses connecting with the different interested structures. The needs highlighted by the production structures and the transversal structures, also taking into account the scheduling of the tenders, were structured by macro item of expenditure and by single census. The expenditure forecasts have been analyzed for each year, starting from 2018, and reports have been prepared highlighting the expenses articulated for two years, distinguishing the period 2018/2019 and 2020/2021 from the following period (2022 - 2031). In order to allow the aggregation of the items and the comparison between the censuses, the subdivision of the expenses by macro-code has been prepared homogeneously for all the censuses. The expenses were divided into current expenses and capital expenditures, and further divided into macro-items of expenditure. The financial tables produced were the basis for writing the census law which was subsequently approved by the Italian Parliament.

Another task, developed in accordance between Contact Group and Istat thematic sectors, is the writing of the General Census Plans. The PGC is a planning act of a general nature that dictates the guidelines regarding the planning, organization and execution of operations carried out at each census. The General Census Plan, drawn up in accordance with the European regulation scheme, describes the legal foundations of census, the organization of the survey network and the tasks of the census bodies, the main technical and methodological aspects of the survey and the calendar of the operations. It represents a point of reference both for Istat and for all the organizations and actors involved in census operations. Moreover, in view of the permanent censuses, the PGCs assume even more

importance because they regulate many operations that are repeated over time with an annual frequency (or biennial or three-year) and not a decade.

4.2. Training

The training of the interviewers and of the various actors on the field is one of the main aspects for the success of such complex surveys as the censuses. The general principle underlying the design of a training process is to allow the detection networks to acquire the skills and abilities necessary to detect high quality data. In fact, training is one of the main strategies for preventing non-sampling errors. In detail, it is important to:

- provide all the tools necessary for the activity to be carried out;
- provide well-calibrated tools for recipients;
- ensure uniform training in the areas;
- provide adequate tools for continuous learning also through the support of the Moodle platform;
- maintain contact with the network for sharing critical issues and solutions;
- provide self-assessment tools and learning assessment;
- provide tools for process monitoring;
- provide tools for the evaluation of the training project.

4.3. Data Collection

The data collection phase is certainly among the most delicate because a bad management of the process can cause errors that cannot be corrected. In this regard, briefly, the main aspects to consider are the following:

- take into consideration the quality of the survey lists;
- adopt operational solutions for the implementation of a possible mix-mode;
- define the integrated tools for monitoring data collection;
- guarantee assistance to respondents and to the survey network.

In this context, it seems necessary to: anticipate any critical issues related to the conduct of data collection in the field, mainly related to any mix-mode; establish virtuous interrelationships between Istat actors and the survey network;

design of integrated and efficient management tools; give greater impetus to the use of administrative sources in a continuous census.

4.4. Survey Management System

The function of data collection is centralized in a single direction, giving the possibility to rationalize and integrate the data acquisition and monitoring phases of the operations. From a technological point of view (IT mainly) this has led to a standardization of the processing phases allowing the possibility to design generalized platforms consistent with well-defined processes. In particular, for the data collection and management phase, two web applications have been developed called the Survey Management System (SGI) and PANDA, as acquisition system, which allowed the first continuous census of population and housing to be carried out and they will also be reused for the census of agriculture and census of non-profit. The acquisition system allows the management of the multi-technique, it guarantees the management of the guided flows of the questions, it manages the consistency and the formal correctness rules. The Survey Management System follows all the collection phases in the field: establishment of the census offices, creation of its own survey network, assignment of work to surveyors, management of the field survey and monitoring of the system. From a technical point of view the new platforms are:

- designed according to the user centricity approach in order to facilitate their usability towards users;
- homogeneous from the point of view of the software infrastructure: this choice helps the IT management to manage any corrections and evolutionary maintenance in a more rational and immediate way;
- developed in compliance with the architectural principles of a service approach;
- integrated with the Institute's metadata systems to allow centralized management of all available information.

4.5. Data Quality

All register-based censuses cannot disregard the quality measurement and the correction of the starting list. The phases in which the statistical error can occur are many and it is necessary to apply different statistical methodologies in order to keep both the sampling part and the non-sampling part under control (Bernardini

and Mazziotta, 2013). The incidence of non-sampling errors, particularly in complex investigations such as census in which a considerable effort of memory and knowledge of the phenomenon are required, can seriously affect the reliability of final results. The quality takes on the meaning of precision that is expressed as an inverse function of the error statistics. The aim of Istat is to provide accurate estimates of the main non-sampling errors, particularly in complex investigations such as the Census (Grossi and Mazziotta, 2016). The researchers must take into consideration many phases, such as the construction of the census list (eligibility of the units, reduction of over/undercover errors and measurement of the error on the microdata), provisional data release, check and data correction and so on. Specifically, the measurement of coverage of the statistical register is a fundamental phase of the whole process of modernization of the official statistics in Italy. In order to evaluate the census coverage, an areal sample survey is carried out. In fact, the aim is to estimate the real number of units really existing on the Italian territory and the coverage rate is the ratio between the number of units pointed out in the register and the number of units really existing on the Italian territory at the same time of register construction. Particular estimation models allow for estimating the number of units for each reference territory of the census with very short time intervals (for example, every year for the population census. For more details, see section 2). In this phase, it is possible to correct the reference register and always have an updated situation of the statistical units present in the territory (individuals, farms, companies). Moreover, the register of places will allow to accurately localize the units providing an indispensable tool to the scientific community and to the public decision maker.

5. Conclusions

The era of big data and the use of administrative sources for statistical purposes is probably the biggest revolution in official statistics. Istat has initiated a process that through the construction of statistical registers completely changes the philosophy of data collection. The Contact Group represents a point of union of skills and professionalism available to the different sectors of the institute in order to standardize common processes towards a quality statistical product.

The Contact Group is a consistent choice with the modernization program, which is based on a common commitment to all the censuses in terms of the fundamentals of the production and output process.

In fact, the modernization is a great opportunity for:

- implementing the registers in an organic way with the different institutional processes;
- harmonizing the survey processes and implement unified solutions for problems common to various contexts;
- developing skills, advanced services, integrated products;
- strengthening the relevance of census products through a structured dialogue with stakeholders and other institutions (Ministries, Municipalities and Regions, Privacy Authority).

Moreover, some strengths are: usefulness of the operation; benefits in terms of cognitive content, relevance of the information produced; increase in timeliness, punctuality and accessibility of data; maintenance of data accuracy and comparability standards; high rate of innovation; reuse of the statistical information and consequent reduction of the burden; great cost reduction compared to traditional censuses.

New census translates into new products: supply of new and diversified annual products in new ways. In order to achieve this objective, the Contact Group has made available different skills in order to be able to cover numerous transversal functions in order to be able to co-deploy the Institute's resources in connections with internal and external structures. This task made it possible to standardize many operational functions by facilitating the exchange of information among the many actors involved in census operations. This is just the beginning of a long process of change of official statistics and censuses towards the joint use of direct surveys and statistical registers. This process also needs an adaptation of the organizational structure in order to optimize workflows and allow the different phases to become an efficient mechanism.

References

- BERNARDINI A., MAZZIOTTA M. 2013. Atti del 6° Censimento Generale dell'Agricoltura – La valutazione della qualità, Vol. 6, Roma: Istat. <https://www.istat.it/it/files/2014/02/5.La-valutazione-della-qualit%C3%A0.Atti-6%C2%B0-Cens-agr..pdf>
- GROSSI P., MAZZIOTTA M. 2016. Atti del 15° Censimento Generale della Popolazione e delle Abitazioni - La valutazione della qualità, Vol. 6., Roma: Istat. https://www.istat.it/it/files//2016/08/AttiPOP_Fascicolo-6-web.pdf

SUMMARY

The Italian Continuous Censuses: Managing the Complexity among Registers, Statistics and Territories

The tradition of censuses in Italy is very strong and, over the years, the experience has allowed us to arrive at exhaustive measurements that guarantee a high quality of the data. The challenge of the coming years will be to maintain the same level of quality for the same territorial domains by increasing the frequency of publications. This will be possible thanks to the combination of sample surveys and statistical registers; the use of administrative data, a new chain of information construction and a renewed organizational management will be the challenge that official statistics will have to win.

Antonella BERNARDINI, Istat, anbernar@istat.it

Daniela BONARDO, Istat, bonardo@istat.it

Alessandra DENTINI, Istat, aldentin@istat.it

Maura GIACUMMO, Istat, magiacum@istat.it

Matteo MAZZIOTTA, Istat, mazziott@istat.it

Alessandra PRETI, Istat, preti@istat.it

LE ATTIVITÀ DELL'ISTAT PER LA SCUOLA E SUL TERRITORIO¹

Patrizia Collesi, Alessandro Valentini

1. Introduzione

Come ben noto, la missione dell'Istat è quella di servire la collettività attraverso la produzione e la comunicazione di informazioni statistiche, analisi e previsioni di elevata qualità. In maniera consequenziale uno dei Programmi strategici di attuazione, il PG5 (*Migliore informazione e comunicazione*), ingloba le iniziative volte all'integrazione dei prodotti e servizi di diffusione e comunicazione.

In effetti l'obiettivo di sviluppo della cultura statistica è da molti anni centrale per l'Istituto (si vedano per esempio GIACCHE', 2008; GIOVANNINI, 2013). Come base delle numerose attività che si sviluppano e si applicano sia a livello centrale che capillarmente sul territorio è possibile considerare la seguente citazione, estratta dal sito dell'Istat²: “Sviluppare un'approfondita conoscenza della realtà ambientale, economica e sociale dell'Italia ai diversi livelli territoriali e favorire i processi decisionali di tutti i soggetti della società (cittadini, amministratori, ecc...)”.

Con questo obiettivo di fondo, negli ultimi anni sono state rinnovate le modalità di intervento basate su una strategia multi-dimensionale volta all'accrescimento della domanda e al conseguente incremento della rilevanza dell'informazione statistica, anche in coerenza con varie esperienze internazionali (UNECE, 2014; FERLIGOJ, 2015; DE SMEDT, 2016). Il nuovo approccio richiede un cambio di paradigma, volto in primo luogo alla fidelizzazione degli utenti e, in maniera simmetrica, alla specializzazione delle richieste da parte dell'utenza. Non è più sufficiente soddisfare la domanda esplicita ma è necessario andare oltre, abbinare ad ogni dato altre informazioni di pari livello informativo inerenti temi ad esse correlati. Al tempo stesso bisogna proporre, per lo stesso tema, informazioni più dettagliate, in maniera tale da soddisfare una platea più ampia e un'utenza più specialistica.

¹ La pubblicazione è frutto congiunto del lavoro dei due co-autori. Tuttavia Patrizia Collesi ha redatto i Paragrafi 1, 3, 5, 6, 8 e Alessandro Valentini i Paragrafi 1, 2, 4, 5, 7.

² <https://www.istat.it/it/organizzazione-e-attivita%C3%A0>

Questo è vero in particolare per il target degli studenti, ai quali bisogna dedicare un percorso di conoscenza che non può prescindere dalla scuola e in particolare dall'utilizzo della statistica nelle materie curriculari.

2. Le nuove linee guida dell'Istat per lo sviluppo della cultura statistica

L'attuale framework in cui opera l'Istat per sviluppare la cultura statistica è caratterizzato da molteplici ambiti. Tra questi è possibile citare: la promozione del paradigma della ricerca statistica; lo specifico orientamento all'utilizzazione e all'applicazione dei dati in vari contesti conoscitivi di decisione e valutazione; lo sviluppo di metodi per il processo decisionale, soprattutto in condizioni di incertezza, attraverso l'uso dei dati della statistica ufficiale; la conoscenza dei criteri che guidano l'attività di ricerca; la promozione della cultura del metadato; una specifica attenzione alla qualità e alla sua certificazione. La ormai sistematica scarsità di risorse a disposizione richiede inoltre di soffermarsi su obiettivi standardizzabili, sostenibili e replicabili.

In tutto ciò un fattore chiave è giocato da una specifica attenzione verso le scuole. Il sistema di educazione formale continua infatti, come da tradizione, a rivestire un ruolo cruciale per la disseminazione della cultura statistica. Ciò nonostante, anche in questo ambito l'Istituto sta apportando notevoli innovazioni, rivolgendo la propria attenzione non soltanto ai ragazzi ma anche agli insegnanti.

Il nuovo approccio verso gli studenti richiede la messa a punto di strumenti utili per la valorizzazione dei dati statistici come mezzo per governare l'incertezza ed incentivare l'uso di «evidenze quantitative». Richiede anche di attivare canali di interazione con i percorsi di *alternanza scuola-lavoro* previsti dalla normativa vigente.

Le azioni per gli insegnanti devono avere alcune caratteristiche: essere pianificate in maniera tale da non risultare atti di mera duplicazione rispetto ai servizi forniti dagli enti di formazione; riguardare l'approfondimento di aspetti applicativi dell'uso della statistica e non delle tecniche; essere incentrate ad incentivare l'utilizzo dei dati della statistica ufficiale.

L'organizzazione vigente dell'Istituto attribuisce la competenza dello sviluppo della cultura statistica alla Direzione centrale per lo sviluppo dell'informazione e della cultura statistica (DCSI), con il compito di progettare i prodotti e le iniziative in merito. Un contributo essenziale è poi quello che promana dalla rete territoriale dell'Istituto, che supporta questi processi attraverso la *rete territoriale per lo sviluppo della cultura statistica* (VALENTINI e CORTESE, 2014). Tale *rete* è attualmente composta da un rappresentante per ciascuna delle 18 regioni dove si trovano le sedi territoriali Istat, con specifiche competenze in materia.

I membri della *rete* in primo luogo assumono il ruolo di facilitatori per trasmettere a livello periferico le iniziative di tipo centrale (*top-down*). Hanno anche il compito di progettare iniziative per territorio (*bottom-up*), che possono essere standardizzabili (ovvero replicabili in altri contesti) oppure non standardizzabili.

Oggetto del presente paper sono le principali iniziative top-down rivolte principalmente alle scuole e agli studenti. Le linee di sviluppo proposte sono condivise e coerenti, inoltre sono spesso frutto di progetti di collaborazione con le istituzioni europee o sovranazionali.

3. Le competizioni

In collaborazione con Società scientifiche, come la Società italiana di statistica (Sis) e con Istituzioni comunitarie (Eurostat) o internazionali, l'Istat promuove anche la partecipazione a giochi e competizioni. In particolare le Olimpiadi di statistica e il concorso di poster statistici ISLP.

3.1. Le Olimpiadi di statistica

Nell'anno scolastico 2018/19 Istat e Società Italiana di statistica hanno organizzato la nona edizione delle Olimpiadi Italiane di Statistica, rivolte agli studenti delle prime quattro classi delle scuole superiori, composte di due fasi: una fase italiana e una europea, e complessivamente tre prove. Come avviene per le altre Olimpiadi scolastiche, anche quelle di statistica hanno una valorizzazione pubblica, infatti i vincitori della prima prova individuale sono inseriti nell'Albo nazionale delle eccellenze³. L'attività viene gestita dall'Istat e dalla Sis, che la aveva avviata 9 anni fa in Italia (POLLICE e BARBIERI, 2019)

L'obiettivo delle Olimpiadi è avvicinare gli studenti al ragionamento statistico, suscitando il loro interesse verso l'analisi dei dati e la probabilità, per metterli in condizione di intendere correttamente il significato delle informazioni quantitative. Mentre le prove individuali vertono più su conoscenze di argomenti di statistica; le prove a gruppi, sia quella che conclude la fase italiana che quella europea, mettono alla prova i ragazzi sulla conoscenza delle statistiche prodotte da Organismi ufficiali, quindi l'Istat e Eurostat. Dalla prima alla seconda fase è interessante questo passaggio dalla statistica teorica a statistiche prodotte, che possono impattare sulla vita quotidiana dei ragazzi coinvolti: l'obiettivo è quello di far comprendere come

³ L'Albo fa parte del [Programma annuale per la valorizzazione delle eccellenze](#) del Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) in base a quanto sarà stabilito nel relativo Decreto Ministeriale.

“buone statistiche” possano far conoscere meglio la realtà (European statistics competition).

L’ultima edizione (a.s. 2018/19) ha visto un notevole incremento nella partecipazione degli studenti, come evidenziato nella figura 1. Fondamentale per questo risultato è stata la collaborazione con la *rete territoriale per lo sviluppo della cultura statistica*, che ha contribuito alle attività sia nella fase di promozione che nella fase di tutoraggio degli studenti, insieme ai tutor della Sis. In valore assoluto si è passati da 2.377 studenti dell’a.s. 2017/18 a 2.810 partecipanti dell’a.s. 2018/19, con notevole incremento anche nel numero di Istituti partecipanti, passati da 76 a 120.

Anche nell’ultima occasione le Olimpiadi italiane di statistica si sono configurate come fase nazionale della European Statistics Competition (ESC). I vincitori delle prove a squadre hanno avuto accesso infatti alla Competizione europea, organizzata da Eurostat con la collaborazione degli Istituti nazionali di statistica. Tra gennaio e febbraio 2019 si è svolta la gara italiana, mentre il 15 maggio si è chiusa la fase europea delle Olimpiadi, coordinata da Eurostat, e inquadrata all’interno del progetto Digicom Package 4 (Communication and Dissemination). L’Istat ha partecipato insieme agli altri INS europei, con un video di due minuti sul tema “Europe reflected in statistics”. La squadra italiana di primo e secondo anno superiore si è classificata quinta a livello europeo.

Figura 1 – Numero di scuole (scala sinistra) e numero di studenti (scala destra) partecipanti alle Olimpiadi della Statistica. Anni scolastici 2014/15-2018/19.



3.2. L’International statistical literacy project (ISLP)

Il progetto ISLP è un concorso biennale internazionale per poster statistici organizzato nell’ambito dell’International Statistical Literacy Project, che cura le attività di cultura statistica per l’International Statistical organization (ISI). Si tratta di una iniziativa rivolta agli studenti delle scuole di tutto il mondo finalizzata ad

accrescere nei giovani la percezione della statistica ufficiale come strumento di conoscenza della realtà quotidiana. L'edizione 2018/2019 è stata aperta per la prima volta agli studenti universitari. I poster vincitori delle fasi nazionali hanno partecipato alla gara internazionale, in occasione del convegno biennale dell'ISI, ad agosto 2019 a Kuala Lumpur, e il poster italiano "Distanza tra realtà e percezione", realizzato dagli studenti del corso di Laurea triennale in economia e statistica per le organizzazioni dell'Università degli studi di Torino, ha vinto la gara internazionale⁴. Per ulteriori dettagli si veda ISLP, 2019.

4. Iniziative sul territorio: il contributo dell'Istat alla Notte europea dei Ricercatori

La *Notte europea dei ricercatori*, che di solito si svolge durante l'ultimo fine settimana del mese di settembre, ha l'obiettivo esplicito di avvicinare il grande pubblico al mondo della ricerca e in particolare alla figura del ricercatore, per accrescere nei cittadini la consapevolezza dell'importanza che oggi riveste la ricerca scientifica. La *Notte* è cofinanziata dalla Commissione Europea, all'interno del Programma Quadro europeo per la Ricerca e l'Innovazione HORIZON 2020. Per il biennio 2018-2019 in Italia partecipano varie associazioni (tra queste Bright, Closer, Luna2016, Made in Science, Sharper, Society, Frascati Scienza).

Da qualche anno i ricercatori dell'Istat contribuiscono agli eventi della *Notte*. Dal 2018, tuttavia, la partecipazione è divenuta ancor più strutturata grazie al rilascio a tutta la rete territoriale di un *Catalogo di iniziative* replicabili, in grado di lanciare in tutti i contesti geografici lo stesso messaggio e di fornire la medesima immagine dell'Istituto.

Le principali proposte del catalogo sono riportate nella Tabella 1. In coerenza con le linee generali di sviluppo della cultura statistica dell'Istituto, l'offerta è multi-target e spazia dai bambini molto piccoli ai ricercatori senior più esigenti.

Per i più piccoli, per esempio, è possibile ascoltare alcune fiabe basate su argomenti statistici (CAMPORESE e al., 2017). Inoltre si può partecipare in forma giocosa ad esperimenti vari inerenti il calcolo della probabilità, quale quello basato sulle carte da gioco, che ha l'obiettivo di distinguere un mazzo truccato da uno regolamentare.

Ai ragazzi delle scuole superiori vengono proposti esempi di narrazioni statistiche più avanzate, come la *Storia di Mattia* che racconta la giornata tipo di un ragazzo universitario basandola sui dati statistici ufficiali. Ma sono offerti anche giochi

⁴ Gli studenti vincitori dell'Università di Torino sono: Andrea Ughetti, Didier Crosetti, Maria Vittoria Chiaruttini, Matteo Canale, Marco Cencig, coordinatrice la professoressa Alessandra Durio (<https://www.istat.it/it/archivio/232829>).

statistici realizzati per lo smartphone, con l'obiettivo, per esempio, di identificare le *fake news*.

Il pubblico più specialista, come i ricercatori più raffinati, viene comunque incuriosito dalla visualizzazione dei vari strumenti di rappresentazione grafica dell'Istituto, nonché dai contenuti interattivi del sito.

Adatto a un pubblico generalista è poi il quiz (on line) denominato "Gioca con la statistica e scoprirai il tuo profilo", che consente sia di attirare l'attenzione del pubblico sia di misurare il livello di statistical literacy dei partecipanti sulla base di uno standard internazionale (VALENTINI, 2016).

Il bilancio del 2018 in termini di numero di visitatori è sicuramente positivo: i 13 eventi svolti sul territorio hanno consentito ad oltre 2.200 persone di visitare gli stand dell'Istat. Le attese per il 2019 sono di crescita, sia del numero di eventi che in particolare dei visitatori.

Tabella 1 – *Attività e giochi proposti nel Catalogo Istat per la Notte europea dei ricercatori". 28-30 settembre 2018. Le principali attività*

Denominazione	Descrizione	Target (età)
<i>Narrazioni: le fiabe</i>	Lettura del libro di fiabe statistiche "Le streghe di Bayes"	3 – 9
<i>Disegni statistici</i>	Predisposizione/realizzazione di disegni per i bambini a carattere statistico	3 – 9
<i>L'Italia in cifre si fa cruciverba</i>	Risoluzione di cruciverba a tema statistico con definizioni tratte da "L'Italia in cifre"	10 – 16
<i>Giochi con lo smartphone</i>	Esecuzione di quiz statistici basati sull'utilizzo dello smartphone	14 anni e oltre
<i>Narrazioni: la storia di Mattia</i>	Lettura di un racconto della vita quotidiana di un ragazzo universitario basato su dati statistici	16 - 25
<i>Giochi veloci con la probabilità</i>	Esecuzione di giochi manuali utili per introdurre concetti statistici basati sulla probabilità	Tutte le età
<i>Video e infografiche Istat</i>	Illustrazione/Proiezione di materiali a carattere istituzionale tratti dal sito dell'Istat	Tutte le età

5. Iniziative e azioni/materiali disponibili per la scuola e i docenti

Come già più volte citato, il mondo della scuola rappresenta un target essenziale per lo sviluppo della cultura statistica (GAL I. OGRAJENSEK, 2017), a cui è dedicata una intera sezione del sito istituzionale⁵. I ragazzi di oggi, infatti, rappresenteranno la classe dirigente di domani. L'idea base è quella di sviluppare un

⁵ <https://www.istat.it/it/informazioni-e-servizi/per-studenti-e-docenti>

insieme di azioni volte ad accrescere il senso critico degli interlocutori attraverso la corretta lettura di dati e informazioni statistiche. Di conseguenza la proposta dell'Istat si sviluppa nel diffondere l'uso pervasivo e ragionato delle statistiche all'interno delle discipline curricolari, orientate alle materie umanistiche e alle scienze sociali, oltre che alle discipline Stem (scienze, tecnologie, ingegneria, matematica).

Per questo motivo una attenzione specifica è rivolta sia al rilascio di materiali utili che ad attività specifiche per le scuole e per i docenti (COLLESI e RONCATI, 2019).

I componenti della rete territoriale per lo sviluppo della cultura statistica hanno messo a punto una serie di pacchetti didattici per le scuole (VALENTINI, CARBONARA, 2016). I pacchetti sono organizzati nella forma di dossier che includono:

- a. una presentazione in Power Point introduttiva del tema;
- b. esercizi interattivi in Excel (comprensivi delle relative risposte);
- c. una nota per l'insegnante in maniera tale da poter proporre i pacchetti nelle classi. Complessivamente i pacchetti rilasciati sono 30: 10 per la scuola primaria, 9 per la secondaria di primo grado e 11 per la secondaria di secondo grado. Le macrotematiche trattate (vedi Tabella 2 per il dettaglio) e il linguaggio adottato sono coerenti con il target di riferimento. I materiali per la scuola primaria, ad esempio, si caratterizzano per il linguaggio e gli esempi adatti ai ragazzi delle scuole elementari. Quelli per la scuola secondaria di secondo grado sono strutturati sotto forma di percorsi orientati ad una migliore comprensione del mondo.

6. A Scuola di Open Coesione (Asoc)

Tra le iniziative specifiche per le scuole un rilievo particolare è quello rivestito dalla collaborazione sviluppata con il progetto ASOC (A Scuola di Open Coesione), percorso di monitoraggio civico attuato nelle scuole superiori. Dopo la fase sperimentale svolta in Molise e Abruzzo per l'anno scolastico 2017/18 (BASILE e GROSSI, 2019), l'Istat ha stipulato un accordo con l'Agenzia per la Coesione territoriale della Presidenza del Consiglio dei Ministri per fornire il supporto dei dati statistici al progetto, divenendo ufficialmente partner di Asoc dal 2018 per tutte le regioni italiane. L'obiettivo è quello di favorire la diffusione della cultura statistica e di fornire un adeguato approfondimento sulle metodologie di ricerca e sulla costruzione di indici sintetici, arricchendo così la formazione degli studenti durante le loro attività di monitoraggio civico.

Il percorso di Asoc a partire dall'anno scolastico 2017/2018 è stato riconosciuto per la prima volta come percorso di Alternanza scuola lavoro e come attività valida per erogare crediti formativi per la formazione docenti.

Tabella 2 – I materiali dell'Istat per la scuola

Livello scolastico	Macro-tema	Pacchetti didattici
<i>Scuola primaria</i>	L'ABC della statistica Primi passi alla scoperta della statistica Giochiamo a fare lo statistico	10
<i>Scuola secondaria di primo grado</i>	Leggiamo la realtà attraverso la statistica Viaggio alla scoperta della statistica La probabilità: primi concetti	9
<i>Scuola secondaria di secondo grado</i>	Percorso introduttivo alla statistica ufficiale La statistica e il mondo del lavoro La rappresentazione dei dati La progettazione e l'implementazione del questionario	11

I team composti da studenti scelgono un progetto del loro territorio realizzato grazie ai fondi europei, ed effettuano il monitoraggio dello stato di avanzamento.

Il percorso didattico è articolato in 5 lezioni con consegne da effettuare, e inoltre da una visita di monitoraggio civico, la partecipazione attiva all'evento internazionale Open Data Day, un evento pubblico finale. La didattica è organizzata secondo un modello misto che prevede modalità di fruizione di contenuti online (tipo MOOC) e lavoro di gruppo project-based attraverso l'utilizzo di tecnologie, strumenti di condivisione online, blog, social network.

All'edizione 2018/19 hanno partecipato poco meno di 200 team, e due squadre supportate dagli esperti territoriali Istat sono arrivate sul podio; al primo posto la squadra ReStart dell'ITT "Mario Ciliberto" di Crotone, supportata dall'Ufficio territoriale della Calabria e al terzo posto quella sostenuta il team "Tangentalete 007" del Liceo "Giuseppe M. Galanti" di Campobasso, sostenuta dall'Ufficio del Molise.

Gli esperti Istat per questa prima edizione nazionale hanno dato la loro adesione su base volontaria. Si tratta di quasi 60 persone che, con il coordinamento centrale dello staff di Promozione e sviluppo di cultura statistica, hanno portato i dati e la statistica ufficiale tra le squadre di Asoc, attraverso assistenza sull'uso dei dati, lezioni frontali e webinar, partecipazione agli eventi della Settimana dell'Amministrazione Aperta.

7. L'esperienza dell'Istat nell'Alternanza scuola lavoro

L'alternanza scuola-lavoro è stata introdotta con la legge n. 107 del 2015 (c.d. legge sulla "buona scuola"). Si tratta di un periodo di tempo che gli studenti del triennio delle scuole superiori devono svolgere nel mondo del lavoro con l'obiettivo

di integrare l'educazione scolastica con una esperienza pratica. Il forte impatto iniziale, pari a 400 ore per gli studenti degli istituti tecnici e professionali, e a 200 ore per i licei, è stato ridotto in maniera significativa con legge di bilancio 2019, che ha anche modificato altri aspetti del percorso di alternanza (che si chiamerà "Percorsi per le competenze trasversali e per l'orientamento"). Ciò nonostante viene confermato sia l'impianto generale della legge che in particolare gli obiettivi di integrazione tra l'educazione scolastica e il mondo del lavoro.

Dopo una iniziale fase di attesa, l'Istat ha aperto le proprie porte ad un percorso di alternanza scuola-lavoro per la prima volta in assoluto durante l'anno scolastico 2018-2019 grazie ad una convenzione formalizzata dall'Ufficio territoriale per l'Emilia-Romagna, la Toscana e l'Umbria (sede per la Toscana) con l'Istituto "Russell-Newton" di Scandicci e ad un preliminare accordo con la Direzione scolastica della Toscana (VALENTINI e DA VALLE, 2019).

Tra i mesi di gennaio e maggio 2019 vari ricercatori dell'Istat hanno svolto complessivamente 60 ore di attività con un gruppo di 12 studenti delle classi terze e quarte dei licei e delle scuole tecniche. La prima parte del percorso è stata focalizzata sull'espletamento di moduli formativi sulla statistica ufficiale, sul processo d'indagine e sulla qualità dei dati. È seguita una parte a carattere prettamente pratico, che ha permesso ai ragazzi di apprendere come realizzare un questionario, implementarlo in versione cartacea e on line, somministrarlo alle unità di rilevazione, elaborare i risultati e presentarli con un evento a carattere pubblico dedicato ai non specialisti.

La realizzazione del questionario è stata anche l'occasione per far convergere più obiettivi tra loro distinti: quelli dei ragazzi impegnati nel percorso di alternanza, quelli del management scolastico e quelli dell'Istat. I ragazzi nel percorso di alternanza, grazie al contatto diretto con l'Istat, hanno acquisito la cassetta degli strumenti utile per gestire una rilevazione concreta. Il management scolastico ha potuto utilizzare i risultati del questionario, validi dal punto di vista statistico, al fine di valutare il gradimento delle strutture da parte degli studenti, nonché di valutare il rapporto degli studenti con lo studio, con le proprie scelte e con gli altri studenti. L'Istat, infine, ha potuto sfruttare il questionario per poter fornire un messaggio di tutti gli studenti della scuola.

L'indagine ha visto il coinvolgimento complessivo come unità di rilevazione di tutti gli studenti iscritti alle classi dalla seconda alla quinta per i vari corsi di studio (liceale e tecnico). Gli studenti delle classi prime sono stati esclusi in quanto la loro conoscenza della struttura scolastica è stata ritenuta ancora non adeguata per poter fornire risposte pertinenti. Agli studenti di seconda è stata somministrata una versione cartacea del questionario; quelli del triennio hanno invece risposto on line tramite il proprio smartphone.

Il questionario è stato somministrato complessivamente a 1.201 studenti, con un tasso di risposta medio dell'88 per cento, senza distinzione significativa per tipo di corso (liceale o tecnico) e per classe (dalla seconda alla quinta). Di fatto i non rispondenti corrispondono agli studenti assenti il giorno di somministrazione (marzo 2019). Non è possibile in questa sede entrare nel merito dei risultati: basti dire che gli stessi saranno oggetto di una attenta pianificazione da parte della dirigenza scolastica.

8. Prospettive di sviluppo

Il periodo 2016-2018 ha visto una riorganizzazione e un riorientamento delle attività di promozione e sviluppo della cultura statistica sia a livello centralizzato che a livello regionale. Questo ha portato, come descritto nei paragrafi precedenti e pure in un'ottica di economia, riuso e standardizzazione dei materiali, a fare sì che molte attività fossero realizzate e reingegnerizzate in un'ottica di maggiore efficienza.

Il percorso è ben lungi dall'essere terminato: quello che rimane da fare, certamente fondamentale per un Istituto di statistica, è la valutazione del percorso fatto, sulla base di evidenze quantitative, misurando cioè successi e insuccessi, impatto avuto e target di utenti raggiunti. Misurare quello che è stato fatto e avere dati certi per andare avanti con un modello di valutazione delle attività può costituire un obiettivo di breve-medio termine che porti a prospettive di sviluppo condivise, economiche, realizzabili, continuando a lavorare e ampliando le possibilità di partnership con Società scientifiche, altre Istituzioni nazionali, internazionali, realtà del territorio, affinché una regia condivisa permetta interventi mirati realizzabili, a più livelli e rivolti a target vari, e magari arrivi ad un percorso unico e certificato di apprendimento e competenze sulla statistica ufficiale.

Riferimenti bibliografici

- CAMPORESE R., DA VALLE S., LETARDI S., OSTI S., PEDDES S. 2017. Le streghe di Bayes e altre storie. Fiabe statistiche per bambine e bambini curiosi. Roma: Istat.
- COLLESI P., RONCATI R. 2019. Statistical literacy as a tool to attract audience starting from school: different audiences and different messages. Unece Workshop on the dissemination and communication of statistics. Gdansk, 12-14 June. https://statswiki.unece.org/pages/viewpage.action?pageId=212303959&preview=/212303959/255492310/Dico2019_5-1_Collesi_Roncati_Italy_AD.pdf
- DE SMEDT, M. 2016. European Statistics and Eurostat's contribution to improving statistical literacy. In J. Engel (Ed.) *Proceedings, IASE Roundtable on Promoting Understanding of Statistics About Society*, Berlin, http://iase-web.org/conference_Proceedings.php?p=Promoting_Understanding_of_Statistics_about_Society_2016
- FERLIGOJ, A. 2015. How to Improve Statistical Literacy? *Metodološki zvezki*, Vol. 12, No. 1, pp. 1-10.
- GAL I., OGRAJENSEK I. 2017. Official Statistics and Statistics Education: Bridging the Gap, *Journal of Official Statistics*, Vol. 33, No. 1, pp.79-100.
- GIACCHE', P. 2008. Italy National Institute of Statistics: emphasizing the metadata, in Sanchez, J. (Ed.), *Government statistical offices and statistical literacy*, ISLP Publications, cap. 6.
- GIOVANNINI, E. 2013. Istat and Statistical Literacy, *ISLPNewsletter*, Vol. 1(6), p. 11.
- BASILE C., GROSSI P. 2019. Open-Cohesion School: Istat regional office in Molise experience, *Rivista italiana di demografia e statistica*, Vol. 73, No. 1, pp. 123-134.
- ISLP 2018. https://iase-web.org/islp/Poster_Competition_2018-2019.php
- POLLICE A., BARBIERI M.M. 2019. Le Olimpiadi italiane di statistica, *Statistica e Società*, No. 2.
- UNECE 2014. Making data meaningful. Part 4: A guide for statistical organizations, Geneva: ed. United Nation Economic Commission for Europe.
- VALENTINI, A. 2016. Come è possibile misurare la cultura statistica? Resoconto di un'esperienza condotta in Toscana, *Induzioni*, Vol. 50(1), pp. 79-90.
- VALENTINI A., CARBONARA M. 2016. Italian NSI for the school. A new project to promote statistical literacy in Italian regions, *Electronic Journal of Applied Statistical Analysis*, Vol. 9, pp. 716-735.
- VALENTINI A., CORTESE P. F. 2014. Il nuovo approccio alla promozione della cultura statistica da parte della rete territoriale Istat, *Induzioni*, Vol. 48(1), pp. 79-93.

VALENTINI A., DA VALLE S. 2019. Un possibile modello di “alternanza scuola-lavoro” nell’ambito della statistica ufficiale: l’esperienza dell’Istat in Toscana, *Statistica e Società*, No. 2.

SUMMARY

Istat activities for schools and at territorial level

The paper focuses on explaining how statistical literacy is dealt with at Istat and presents an overview of practical examples of several projects directly carried on by Istat or with partnerships. Istat develops several project for schools and students in Italy. Some of these, namely the competitions, are carried out in collaboration with European or supranational institutions, while others, bigger projects which involve the targets for longer periods, are dealt centrally and then brought concretely on the territory in accordance with the Network of territorial offices, according to ways of development centrally agreed upon. They are: Italian and European Statistics Competition; ISLP (statistical posters from the ISI) European Researchers' Night (joint programme and re-use of materials). A thorough presentation of a relevant cooperation started in 2018-2019 school year as well as an experience of a project with secondary school students in Toscana are also reported. The final remarks give a suggestion of what could be done further: that is an assessment of previous activities so as to come to a well-gearred and common path for introducing statistical literacy.

Patrizia COLLESI, Istat, pcollesi@istat.it
Alessandro VALENTINI, Istat, alvalent@istat.it

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).
- LVI Benessere e Territorio: Metodi e Strategie (Università Politecnica delle Marche, Ascoli Piceno 23-24 maggio 2019).