VOLUME LXXVI – N. 1 GENNAIO-MARZO 2022

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

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

SIEDS
SOCIETÀ ITALIANA
DI ECONOMIA DEMOGRAFIA E STATISTICA

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

IN MEMORY OF

PROF. GIOVANNI MARIA GIORGI

Macerata, 2 May 1947 – Rome, 10 October 2021

On 10 October 2021, Prof. Giovanni Maria Giorgi passed away.

He was the President of the Italian Society of Economics, Demography and Statistics (*Società Italiana di Economia Demografia e Statistica - SIEDS*) from 2008 to 2014 and again since 2015, as an honorary president.

Professor Giorgi originated from the Marche region, something which he was very proud of, he was born in Macerata on 2 May 1947. He then studied in Rome at the Faculty of Statistics, Demographic and Actuarial Sciences at the end of the 1960s. He later moved to Siena where he began his university career, first as an assistant and then as an associate professor. In 1994, he returned to Rome, as a full professor of Statistics, at the "La Sapienza" University of Rome, where, until 2015, he held the Multivariate Statistics and Complex Sample Design course. He had been a member of the Academic Board of the Doctorate in Methodological Statistics and, furthermore, he became - from 2003 to 2006 and from 2013 to 2017 – Editor in Chief of Metron (founded by Corrado Gini in 1920).

He has a deep bond with SIEDS. His presidency coincided with a period of great difficulty that affected Italian universities and research in general, as

well as having an impact on SIEDS. However, with his commitment, stubbornness (the "*tigna*" as he used to say) and the spirit with which he always characterized his actions, he was able to firmly guide the Society during these difficult years, laying the foundations for its success. During his presidency, Professor Giorgi dedicated particular efforts to the progress of the Italian Journal of Economic, Demographic and Statistical Studies (*Rivista Italiana di Economia Demografia e Statistica* – RIEDS), the official body of the SIEDS, aware of the importance of scientific production and dissemination at an international level. Even after the end of his presidency, he never left out his presence, his valuable contribution of knowledge, experience and support, not only to the board of directors, but also to the editorial board of the Journal and during the yearly scientific meetings of the Society. Professor Giorgi also represented an important bridge between the history of our Society (founded in 1937) and the Italian school of statistics of which he had a deep knowledge due to his studies and the numerous anecdotes collected during his friendship with Carlo Benedetti (1921-2015), pupil of Carlo Emilio Bonferroni and collaborator of Corrado Gini.

In our Society, he brought about his esteemed competence, his scientific rigour, as well as his work ethic. He was unanimously recognized as one of the most important worldwide experts on inequality measures and in particular on the Gini concentration index. The large number of scientific articles and contributions published in national and international journals, handbooks and encyclopaedias, written by him are proof of this. The last invited speech held by Professor Giorgi during the LVI SIEDS Scientific Meeting in Ascoli Piceno in 2019, entitled "The Gini concentration ratio: back to the future", provides an exhaustive and complete picture of the origin, current development and future perspective of the famous concentration index. In the SIEDS Scientific Meetings held under his presidency, special consideration was devoted to the relevant issues related to welfare. We would like to highlight, in particular, the session on "Poverty, inequality, well-being: methodological, historical and evolutionary aspects" during the XLVIII Scientific Meeting in 2011 in Rome, for which he personally laid down a lot and involved eminent speakers, both national and international.

The news of his death came as a great surprise, to us all. Up until his last day he continued to be interested in "his things" and to carry out his numerous

scientific projects. His passion for research and the warmth of his family filled his life and allowed him to overcome his health issues which had affected him in the last years.

Our thoughts and affection are with his wife, Liliana, who also accompanied him to the last Scientific Meetings of the SIEDS and to his two daughters, Giulia and Grazia, who were close to him until the end.

With him, SIEDS loses an important reference point, a prominent scholar but, furthermore, a strong, constant, discreet and polite presence, a guide, an authoritative and never trivial voice, an honest, clever, sarcastic and self-ironic person, an example for those who approach the world of research, but for us, more profoundly, a friend with whom we shared a part of our life, not just professionally.

Certainly, his absence will be felt, especially, in times of trouble and impasse moments, in which with his sarcasm he was able to face every problem and with his wisdom find the solution. His example and his teachings will forever help us.

Goodbye, dear Giovanni

SIEDS Executive Committee

IN RICORDO DEL

PROF. GIOVANNI MARIA GIORGI

Il 10 ottobre 2021 il prof. Giovanni Maria Giorgi ci ha lasciati.

È stato presidente della Società Italiana di Economia, Demografia e Statistica (SIEDS) dal 2008 al 2014 e, dal 2015, presidente onorario.

Il prof. Giorgi aveva origini marchigiane delle quali andava orgogliosamente e giustamente fiero: era nato a Macerata il 2 maggio del 1947. Aveva compiuto il suo percorso universitario a Roma, nella Facoltà di Scienze Statistiche, Demografiche ed Attuariali alla fine degli anni '60. Successivamente si era trasferito a Siena dove aveva iniziato la sua carriera universitaria, prima come assistente e poi come professore associato. Nel 1994 era stato chiamato, come professore ordinario di Statistica, all'Università "La Sapienza" di Roma, dove, fino al 2015, ha tenuto il corso di Statistica Multivariata e Disegni campionari complessi. Membro del Collegio Docenti del Dottorato in Statistica Metodologica, aveva ricoperto - dal 2003 al 2006 e dal 2013 al 2017 – la carica di direttore della Rivista Metron (fondata da Corrado Gini nel 1920).

Il suo legame con la SIEDS è stato profondissimo. La sua presidenza è coincisa con un periodo di grandi difficoltà che avevano investito il mondo dell'università italiana e della ricerca in generale e, di riflesso, anche la SIEDS. Tuttavia, con il suo impegno, la sua caparbietà (la "tigna" come la chiamava lui) e con lo spirito che hanno sempre contraddistinto il suo agire, ha saputo guidare saldamente, in quei difficili anni, la Società ponendo le basi per il suo rilancio. Durante la sua presidenza, il prof. Giorgi ha dedicato particolare impegno al progresso della Rivista Italiana di Economia, Demografia e Statistica (RIEDS), organo ufficiale della SIEDS, consapevole dell'importanza di una produzione e divulgazione scientifica di livello e di respiro internazionale. Anche una volta terminato il suo mandato da presidente, la sua presenza, il suo prezioso contributo di conoscenza, esperienza e supporto non sono mai venuti meno, non solo all'interno del consiglio direttivo, ma anche per il comitato di redazione della rivista e durante le riunioni scientifiche della Società. Il prof. Giorgi ha rappresentato

un importante ponte con la storia della nostra Società (fondata nel 1937) e della scuola italiana della statistica di cui era profondo conoscitore per via dei suoi studi e dei numerosi aneddoti raccolti durante l'amicizia intrattenuta con Carlo Benedetti (1921-2015), allievo di Carlo Emilio Bonferroni e collaboratore di Corrado Gini.

Nella Società ha portato la sua stimata competenza, il suo rigore scientifico, la sua etica del lavoro. Era riconosciuto come uno dei massimi esperti internazionali delle misure di disuguaglianza ed in particolar modo dell'indice di concentrazione di Gini. Ne sono prova i suoi numerosi articoli scientifici e contributi pubblicati su riviste, handbook ed enciclopedie nazionali ed internazionali. L'ultima relazione invitata che il Prof. Giorgi ha tenuto durante la LVI riunione scientifica SIEDS di Ascoli Piceno nel 2019, dal titolo "The Gini concentration ratio: back to the future", fornisce un quadro esaustivo e completo dell'origine, dell'attualità e delle prospettive future riguardo il famoso indice di concentrazione. Nelle riunioni scientifiche organizzate sotto la sua presidenza, uno spazio era spesso riservato a tematiche di forte attualità legate al welfare. Ricordiamo, in particolare, la sessione su "Povertà, disuguaglianza, benessere: aspetti metodologici, storici ed evolutivi" durante la XLVIII riunione scientifica nel 2011 a Roma per cui si era speso tantissimo in prima persona e che aveva visto il coinvolgimento di eminenti relatori sul piano nazionale ed internazionale.

La notizia della sua scomparsa è stata improvvisa. Fino al giorno precedente ha continuato ad interessarsi delle "sue cose" e a portare avanti i numerosi progetti scientifici che aveva ancora in mente. La passione per la ricerca ed il calore della famiglia hanno riempito la sua vita e gli hanno consentito di superare gli acciacchi che negli ultimi anni ne avevano minato la salute.

Ai suoi cari, la moglie Liliana, che lo ha anche accompagnato nelle ultime riunioni scientifiche della SIEDS, e le sue due figlie, Giulia e Grazia, che gli sono stati vicino fino alla fine, va tutto il nostro affetto.

Con lui la SIEDS perde un punto di riferimento importante, uno studioso di rilievo ma soprattutto una presenza forte, costante ma discreta e garbata, una guida, una voce autorevole e mai banale, una persona arguta, schietta, ironica ed autoironica, un esempio per chi si affacciava al mondo della

ricerca ma per noi soprattutto un amico con cui abbiamo condiviso una parte della nostra vita non solo professionale.
Sicuramente la sua mancanza si farà sentire, ancor di più nei momenti di difficoltà e di impasse in cui con la sua ironia riusciva a stemperare ogni problema e con il suo buon senso a trovarne la soluzione. Il suo esempio ed i suoi insegnamenti ci saranno d'aiuto per sempre.

Ciao caro Giovanni

Il Consiglio Direttivo della SIEDS

JOB INSTABILITY IN THE ITALIAN LABOUR MARKET: IS THERE AN ETHNIC BIAS?

Andrea Ciccarelli, Rinaldo Evangelista, Elena Fabrizi, Silvia Longhi

1. Preliminary remarks

The study of migratory flows is becoming increasingly central in political and economic analysis, not only for its direct connection with to the issues of social integration between migrants and natives, but also – and above all – for the great contribution that immigrant workers can give to the economy and welfare state.

Looking at Eurostat data the magnitude of net migration for Italy has changed significantly over the last years: the population share of migrants rose very rapidly, from 1.1 per cent in 1995 to 8.4 per cent in 2019 (about 5 million people); however, the number of immigrants participating as active population in the Italian labor market has risen appreciably over the past 10 years (Strozza and De Santis, 2017). Migrants tend to be concentrated among blue-collars (while Italians employed as blue-collars have decreased over time) and, at a sectoral level, in "Construction" and "Hotels and restaurants"; they are more willing to move within a territory, due to fewer family ties in the emigration country; in general, migrants tend to be younger, employed in low-wage sectors and in low-skilled professions (Cozzolino *et al.*, 2018).

Many studies have analyzed the dynamics of immigrants' integration on the basis of data not always able to outline – in a longitudinal perspective – the career in the long run (Cebolla-Boado and Finotelli, 2015; Venturini and Villosio, 2008). The aim of this is paper is to explore the existence, the level and nature of ethnic bias in the Italian labour market over the period 2004-2017.

More specifically, we look for the existence of significant differences between Italian and non-Italian workers in the type of labour contracts and in the level of job continuity among the two groups of workers. The empirical analysis is based on the AD-SILC database, a microdata panel containing INPS administrative data merged with very detailed data drawn from EU-SILC, the European survey on income and living conditions. The use of such rich dataset allows us to adopt a long run longitudinal perspective on the issue explored and overcome the limitations characterizing the majority of existing studies based on cross-sections short-term survey data.

2. Some introductory issues

A growing literature has put emphasis on investigating the role of ethnic bias in the inclusion and access of ethnic minorities in the labour market. Empirical research generally shows that migrants in general do not compete with natives in the labour market: foreign workers have more fragmented careers and less job stability compared to natives (more seasonal or short-term jobs; sometimes alternating between legal and illegal employment – Fullin, 2016; Dell'Aringa *et al.*, 2015).

Several studies show the existence of a gap between immigrant and native workers in terms of socio-economic status (Ballarino and Panichella, 2018; Ballarino and Panichella, 2015), consistency of employment with respect to qualification (Prokic-Breuer and McManus, 2016) and remuneration (Heath and Cheung, 2007). Part of these gaps depends on the "disadvantaged" composition of immigrant workers (educational qualifications, age, greater attitude for less skilled jobs). However almost all immigrant groups suffer a further disadvantage that is not dependent on individual characteristics: this residual disadvantage is called "ethnic penalty" (Heath and Cheung, 2007).

Immigrants are often less educated than the natives, especially if they come from less developed countries, and if they are highly educated, their qualifications are not easily recognized in many host countries. In addition, most immigrants have poor knowledge about the functioning of the labour market in the host country, so making difficult for them to find a job matching their skills and expectations (Kogan, 2007; Borjas, 1994).

The focus has mainly been on the analysis of the existing difference in earnings between stayers and immigrants, examining those factors that can contribute to the presence of significant ethnic penalties. As regards to the level of wages there is an "unadjusted" differential of about 30-40% between natives and immigrants (in favour of the former, of course); this is because immigrants tend to be younger, in low-wage sectors and less skilled professions; if you take into account socio economic characteristics (ie for the same sectors, professions, gender, age, contract, etc.) the differential is significantly lower (about 13% - Cozzolino *et al.*, 2018). Furthermore, the wage gap seems to be greater for graduates: in fact, while for natives the degree involves an increase in salary levels, this is not the case for immigrants (Tosi *et al*., 2020)

A strand of literature (Dell'Aringa *et al.*, 2015 and Cappellari *et al*., 2012) also explores the return of human capital (HC), considering factors such as education level and work experience, based on Chiswick's assimilation theory (Chiswick, 1978). Evidence shows how education and labour market experience attained in a foreign country is often significantly less evaluated than human capital obtained in the domestic country, suggesting the existence of the so called "imperfect

transferability" of human capital. This imperfect transferability of human capital across countries has been modelled in several studies and findings suggest that it can contribute to explain immigrant-native wage gap (see, among others, Basilio *et al.*, 2017).

A line of research investigates the impact of country-specific labour demand on immigrants' integration in the labour market (LM) (Fullin *et al.*, 2011). LM structural characteristics outline the pattern of immigrants' participation; a higher level of segmentation clusters immigrants in jobs with less skill and lower opportunities for advancement or raises. Moreover, occupational segregation in low-skilled jobs is shown to be associated with wage penalties.

Both approaches HC and LM segmentation are useful in the analyses of ethnic penalties in labour force participation and wages, exploring the contribution of different components in driving the existing bias.

However, little attention has been paid to qualitative aspects of immigrant's participation in the labour market in terms of fragmentation of career paths.

The aim of this paper is to contribute to the debate investigating the existence of an "ethnic bias" in the Italian labour market expanding the analysis to measure the level of fragmentation degree in career paths of immigrants focusing on critical dimensions such as (1) type of contract, (2) duration of employment, (3) time of stabilization and (4) contract transitions.

3. The empirical evidence

As anticipated in the introduction, the empirical analysis presented in this paper aims at exploring the existence of systematic differences in the way native and non-native workers participate in the labour market. The time span covered by our analysis is from 2004 to 2017, long enough to highlight trends and long-term changes in some structural features of working conditions and on the specific role played by citizenship. We are interested in verifying whether the working conditions and the careers differ in some key elements both in the short and in the long run. More specifically, our focus is on the contractual arrangements and the average duration of the contracts (in the short run, number of days in a year), the transition between different working status (looking at a 5-years horizon) and the overall stability of careers measured by the length of standard contracts over the period 2004-2017.

The analysis is based on an innovative database named AD-SILC. This is built by merging data from the Italian version of the EU-SILC survey with longitudinal information provided by administrative archives managed by the National Institute of Social Security (henceforth, INPS). From INPS administrative archives it is possible to obtain information on the working history of individuals. From IT-SILC

survey, the Italian database of the European union survey on income and living conditions (EU-SILC), collected by Istat, are collected information on the socio-demographic and economic status of interviewed individuals. This allows to reconstruct individuals' work and life patterns by controlling for a much higher number of variables than those included in both the original INPS and SILC datasets (for a detailed information on AD-SILC data see INAPP, 2020[1]).

The first and very basic qualitative dimension we take into account to assess the level of instability of work history patterns is the type of contract. Figures 1 and 2 show the relevance of different types of contractual arrangements among respectively immigrant and native workers aged 25-64. The AD-SILC database allows us to distinguish between the following labour contracts and professional categories: open-ended contracts, fixed-term contracts, professionals, self-employed work, atypical worker. Non-standard contractual arrangements such as casual workers, voucher-based workers are not covered by AD-SILC and are not included in our analysis.

For both native and non-native groups of workers the most widespread types of working arrangements are those taking either the "permanent" or "fixed term" contractual form. The most relevant difference between the two groups of workers has to do with the relative weight assumed by these two contractual arrangements. Permanent jobs are much more diffused among native workers than among immigrants. For workers with an Italian citizenship permanent contracts account for around 60% of total jobs, whereas among the immigrant component of the labour force this type of contractual arrangement accounts only for around 50% of total jobs. In the case of the immigrant component of the labour force it is also worth noticing the significant increase over time of the share of fixed term contracts which at the end of the period account for more than 40% of total jobs. All in all, Figures 1 and 2, provide a first, although rather clear-cut, indication of the existence of higher level of instability of the working status and conditions of immigrants in comparison to what is found in the case of an Italian citizens.

[1]https://inapp.org/sites/default/files/progetticompetitivi/mospi/documenti/T_Dymm%203.0%20Forecast%20model%20report.pdf

Figure 1 – *Distribution by employment status of the workforce. Non-native workers.*

Source: Our elaboration on AD-SILC data.

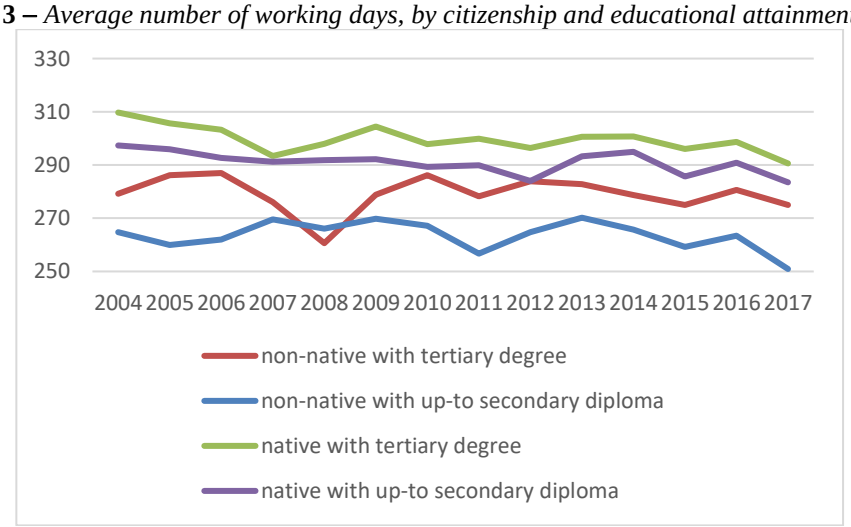
Figure 2 – *Distribution by employment status of the workforce. Native workers.*

Source: Our elaboration on AD-SILC data.

The second dimension used to measure the instability of work history patterns is the length of the contracts. Figure 3 shows the average number of working days – computed on a yearly basis – over the period 2004-2017, respectively for immigrants and native workers. The role played by the educational level is also introduced to include the distinction between graduate workers from those who only have an up-to-secondary diploma. The figure clearly shows that the average length of the

contracts is strictly correlated both to the level of education and the citizenship of the workers. The longest contacts are appanage of Italian workers having a tertiary degree. For this group of workers, on average, the contract duration is around 300 days per year although the time length of the contracts has considerably shrunk over the 2004-17 period. Immigrant workers having only a secondary diploma show, on the contrary, the shortest contracts. Also, for this latter group of workers the duration has progressively declined reaching, at the end of the period, on average, the level of 250 days per year. In summary, for both natives and immigrant workers, the educational level affects the duration of the contracts although, in the case of immigrant workers, the tertiary degree is not a guarantee of long-lasting jobs. In fact, immigrants holding a tertiary education work on average a higher number of days than less educated non-native workers but the length of their contract still remains below the average length of native workers with a secondary diploma. Summing up, the rather limited number of days worked during the year by immigrant workers provides further evidence on the existence of a high level of precariousness and fragility of the working status of this component of the Italian labour force.

Figure 3 – *Average number of working days, by citizenship and educational attainment.*



Source: Our elaboration on AD-SILC data.

The analysis of the job transitions covers a 5-year period and allows us to provide a more in depth and dynamic picture of the level of instability of the work histories of the two groups of workers considered in this study (non-native and native workers). Through the transition matrix it is possible to describe the probability of a change from one working status to another one (net to job-to-job transitions). In

order to compare the dynamic picture of the careers of native and non-native workers, the initial period (t_0) has been associated with the state observed after five years (t_5). Each value in the transition matrix should be read as the (average) probability of changing the status over the period 2004-2013.

Table 1 – *Work-status transition matrix from time (t_0) to (t_5). Immigrant and Native workforce.*

		Immigrant					
		(t_5) time					
		Permanent	Fixed	Professional	Self-employed	Atypical	Out of work
(t_0)time	Permanent	58,2	11,2	0,4	4,5	0,8	24,9
	Fixed	27,1	32,4	0,2	1,3	1,0	38,1
	Professional	7,4	2,0	77,7	2,7	4,1	6,1
	Self-empl.	7,2	3,9	0,0	71,7	2,5	14,7
	Atypical	35,4	15,1	3,6	5,6	17,4	23,0
		Native					
		(t_5) time					
		Permanent	Fixed	Professional	Self-employed	Atypical	Out of work
(t_0) time	Permanent	74,4	6,7	0,7	2,6	1,3	14,2
	Fixed	34,4	18,3	1,0	2,2	1,7	42,5
	Professional	8,5	2,0	82,9	1,6	1,4	3,6
	Self-empl.	5,5	3,1	0,3	81,2	1,8	8,0
	Atypical	30,0	11,6	6,3	5,1	29,7	17,2

Source: Our elaboration on AD-SILC data.

Table 1 and 2 show the transitions respectively for all native and non-native workers and by level of educational attainment. The main results emerging from this comparison are the following: when compared to immigrants, native workers have a higher probability to maintain over time a permanent working status. Almost three out of four native workers with a permanent contract at time t_0 are found having the same contractual arrangement after five years. In the case of non-native such percentage value drops to less than 60%. Furthermore, the probability of moving from a fixed-term contract to a permanent position is low for non-native (vis-à-vis native workers) indicating a greater difficulty for those workers to reduce the instability over their careers. The high level of instability of the work history patterns of migrant workers is also confirmed by the rather high percentage of them working at time t_0 but being "out of work" at the end of the 5-years period and this is hold

true irrespective of the working status observed at time t_0. The same type of ethnic gaps in the work histories emerge looking at graduated workers. Table 2, replicating the transition matrixes for the more educated components of the labour force, shows, for native workers, a higher probability of maintaining a permanent job, a greater possibility of moving towards more stable jobs, and the lower risk of becoming unemployed.

As shown in the transition matrix, having a permanent contract does not guaranty a real stability over the career in the long run, given the possibility, especially in the case of small firms, in particular sectors and after the recent reform (i.e., Jobs Act, Legislative Decree 81/2015) to be laid off and a "forced end" of these types of contracts. The real level stability (instability) of work history patterns can be assessed taking into account the individual work histories in a sufficiently long period and measuring (ex-post) the actual duration of stable contracts.

Table 2 – *Work-status transition matrix from time (t_0) to (t_5). Immigrant and Native workforce with tertiary degree.*

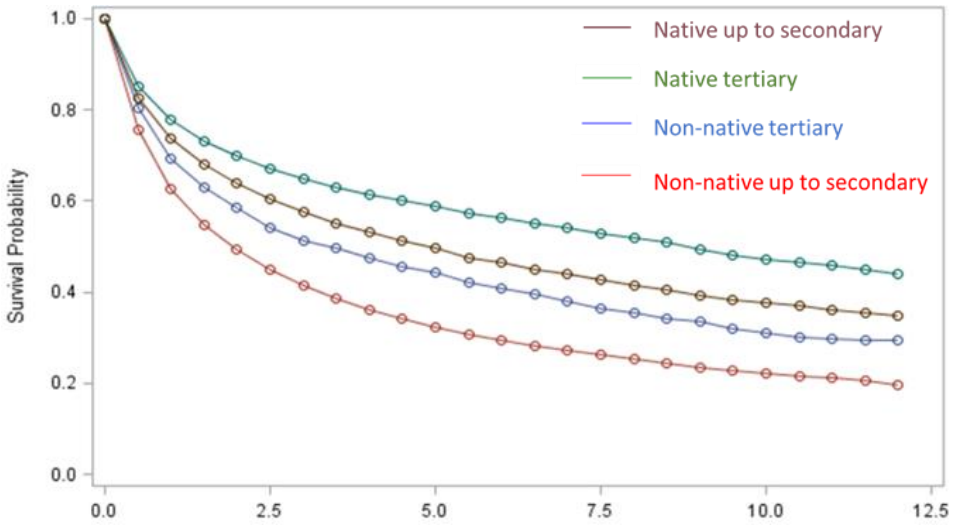
		Immigrant					
		(t_5) time					
		Permanent	Fixed	Professional	Self-employed	Atypical	Out of work
(t_0)time	Permanent	63,7	10,9	2,1	2,5	1,2	19,7
	Fixed	29,7	28,1	0,7	1,7	2,8	37,1
	Professional	6,3	1,0	85,4	2,1	2,1	3,1
	Self-empl.	7,0	3,5	0,0	70,2	3,5	15,8
	Atypical	35,1	13,4	7,2	1,0	22,7	20,6
		Native					
		(t_5) time					
		Permanent	Fixed	Professional	Self-employed	Atypical	Out of work
(t_0) time	Permanent	79,4	4,8	2,8	1,2	2,1	9,7
	Fixed	39,6	13,7	3,6	1,4	3,2	38,5
	Professional	8,2	2,3	85,1	0,9	1,1	2,5
	Self-empl.	9,6	5,6	2,4	69,4	3,2	9,9
	Atypical	32,0	12,9	10,7	2,6	25,8	16,0

Source: Our elaboration on AD-SILC data.

The Kaplan-Meier reported in Figure 4 provides an estimate of the duration of stable contracts. The event estimated is observed whereas a stable contract is interrupted (therefore net to job-to-job transition among stable jobs) and a transition

toward a fixed term contract, or an atypical contract, or a job interruption is observed. Steep and low curve in the Figure 4 represent the worst-case scenario: for every year in the x axis, each point of the curve represents the number of workers, conditional on keeping the stable contract until that moment. In Figure 4 the median duration of stable contracts is shown respectively for Italian and migrant workers, according to their educational attainment. The median duration of stable contracts is 9 years for graduated natives, 5 years for natives with an up to secondary diploma, 3.5 years for non-natives with a tertiary degree and 2 years for non-natives with an up to secondary diploma.

Figure 4 – *Duration in years of stable contracts for native and non-native workers: Kaplan-Meier estimate.*



Source: Our elaboration on AD-SILC data.

4. Some concluding remarks

The existence and relevance of an ethnic bias in the labour market of advanced economies is a theme of policy and social concern and one attracting an increasing amount of empirical research. So far, the bulk of the literature on this topic, including the one looking at the Italian case, has been largely focussed at assessing and quantifying such a bias looking at wages and at exploring the possible determinants of a wage gap between native and immigrant workers. Other more qualitative aspects differentiating the modalities through which immigrant and native workers participate to the labour market have remained under investigated. With this

contribution we have provided fresh descriptive evidence on some of these qualitative aspects and in particular on those able to quantify and qualify the level of stability (instability) of jobs and the work histories of immigrant and native workers. We have focussed our empirical analysis on the contractual forms under which work activities take place and the duration of such contracts, analysing these two fundamental aspects in a dynamic longitudinal perspective and taking into account a rather long-time span (2004-2017). This has been possible thank to the matching of the longitudinal microlevel data from the National Institute of Social Security (INPS) with data collected by the EU harmonised survey on "Income and Living Conditions (SILC)" carried out by the Italian Institute of Statistics (Istat). Despite the evidence presented in this contribution is purely descriptive it has provided relevant insights confirming the presence of strong and persistent differences between immigrant and native workers in the ways in which these two groups of works take part in the Italian labour market. The paper has shown that the work histories of immigrants when compared to native ones, are characterized by a higher level of precariousness and fragility and in particular by a higher relevance of short-term contracts, a shorter average job duration, a higher difficulty of moving towards permanent and long-lasting contracts. Two main conceptual and methodological implications can be drawn from the evidence presented: a) the very high level of fragmentation of the working conditions of immigrants is likely to be a very important factor behind the wage gap between native and non-native workers found by existing empirical studies. This implies that any empirical exercise aiming at assessing such a gap should take this into account; b) the variables and dimensions taken into account in this study (i.e. the type and duration of contracts, the level of saturation of the potential working time, the specific characteristics of the individual work histories) are even more crucial when we conceptualize and measure the wage gap issue (as should be done) in a more dynamic and long term perspective, that is taking into differences in the cumulative capacity to generate incomes.

References

BASILIO T.L, BAUER T. 2017. The Transferability of Human Capital and immigrant Assimilation: An Analysis for Germany, *Labour*, Vol. 31, No. 3, pp. 245-264.

BALLARINO G., PANICHELLA N. 2015. The Occupational Integration of Male Migrants in Western European Countries: Assimilation or Persistent Disadvantage? *International Migration*, Vol. 53, No 2, pp. 338-352.

BALLARINO G., PANICHELLA N. 2018. The Occupational Integration of Migrant Women in Western European Labour Markets, *Acta Sociologica*, Vol. 61, No. 2, pp. 126-142.

BORJAS G.J. 1994. The Economics of Immigration, *Journal of Economic Literature*, Vol. 32, pp. 1667-1717.

CEBOLLA-BOADO H., FINOTELLI C. 2015. Is There a North–South Divide in Integration Outcomes? A Comparison of the Integration Outcomes of Immigrants in Southern and Northern Europe, *European Journal of Population*, Vol. 31, No. 1, pp. 77-102.

CHISWICK B. R. 1978. The Effect of Americanization on the Earnings of Foreign-Born Men, *Journal of Political Economy*, Vol. 86, No. 5, pp. 897-921.

CAPPELLARI L., DELL'ARINGA C., LEONARDI M. 2012. Temporary Employment, Job Flows and Productivity: A Tale of two Reforms, *The Economic Journal*, Vol. 122, No. 562, pp. F188-F215.

COZZOLINO M., DI PORTO E., MARTINO E.M., NATICCHIONI P. 2018. Gli immigrati nel mercato del lavoro italiano: uno sguardo all'universo dei lavoratori dipendenti 1995-2015. In De Arcangelis G., Giovannini E., and Schivardi F. (Eds.) *Economia Italiana*, Vol. 1., pp. 35-72.

DELL'ARINGA C., LUCIFORA C., PAGANI L. 2015. Earnings Differentials Between Immigrants and Natives: The Role of Occupational Attainment, *IZA Journal of Migration*, Vol. 4, No. 8, pp. 1-18.

FULLIN G., REYNERI E. 2011. Low Unemployment and Bad Jobs for New Immigrants in Italy, *International Migration*, Vol. 49, No. 1, pp. 118-147.

FULLIN G. 2016. Labour Market Outcomes of Immigrants in a South European Country: Do Race and Religion Matter? *Work, Employment & Society*, Vol. 30, No. 3, pp. 391-409.

HEATH A., CHEUNG S.Y. 2007. Unequal chances: ethnic minorities in western labour markets. In *Proceedings of the British Academy,* Vol. 137, Oxford: Oxford University Press.

KOGAN I. 2007. *Working through barriers: host countries' institutions and immigrant labour market performance in Europe*. Dordrecht: Springer.

PROKIC-BREUER T., MCMANUS P.A. 2016. Immigrant Educational Mismatch in Western Europe, Apparent or Real? *European Sociological Review*, Vol. 32, No. 3, pp. 411-438.

STROZZA S., DE SANTIS S. 2017. *Rapporto sulla popolazione: le molte facce della presenza straniera in Italia*. Bologna: Il Mulino.

TOSI M., BORGNA C., BELLONI M. 2020. Quando lo studio non paga. Un'analisi sui differenziali salariali fra stranieri e italiani, *Economia e Società Regionale*, Vol. 1., pp. 42-59.

VENTURINI A., VILLOSIO C. (2008), Labour-Market Assimilation of Foreign Workers in Italy, *Oxford Review of Economic Policy*, Vol. 24, No. 3, pp.517-541.

SUMMARY

Job instability in the Italian labour market: is there an ethnic bias?

By using a unique dataset (AD-SILC) merging the longitudinal microlevel data from the National Institute of Social Security (INPS) with data collected by the EU harmonised survey on "Income and Living Conditions (SILC)" carried out by the Italian Institute of Statistics (Istat), this contribution aims at investigating the existence of an "ethnic bias" in the Italian labour market. More specifically, the paper explores the existence of structural differences in the way immigrant and native workers participate in the labour market taking into account: a) the type of contract (permanent versus fixed term); b) the duration of the contracts; c) the patterns of transition among different contractual arrangements; the overall long term level of stability of the work history trajectories. The evidence presented shows that the work histories of immigrants when compared to natives' ones, are characterized by a higher level of precariousness and fragility and in particular by a higher relevance of short-term contracts, a shorter average job duration, a higher difficulty of moving towards permanent and long-lasting contracts.

Andrea CICCARELLI, University of Teramo, aciccarelli@unite.it
Rinaldo EVANGELISTA, University of Camerino, rinaldo.evangelista@unicam.it
Elena FABRIZI, University of Teramo, efabrizi@unite.it
Silvia LONGHI, University of Teramo, slonghi@unite.it

LOCATION QUOTIENT AS A LOCAL INDEX OF RESIDENTIAL SEGREGATION. THEORETICAL AND APPLIED ASPECTS[1]

Federico Benassi, Massimiliano Crisci, Stefania M.L. Rimoldi

1. Introduction

The location quotient (LQ), a ratio of ratios, is a widely known geographic index (Isard, 1960). It is used to measure and map relative distributions or relative concentrations of a character in a subarea compared to the area as a whole (Wheeler, 2005). Originally adopted in regional economic studies (Crawley *et al.*, 2013), recently it has been proposed as a local index in studies of residential segregation of foreign population (Apparicio *et al.*, 2008; Iglesias-Pascual *et al.*, 2019). In this field of studies, the LQs are particularly useful when applied to a metropolitan area where they allow to identify the spatial units in which a population group is under-represented (LQ<1) or conversely, over-represented (LQ>1). Despite their apparent simplicity, the location quotients present some slippery features this contribution tries to reflect on. The caveats refer principally to the choice of the reference population, the distribution cut-offs for maps and comments, and their aspatial nature.

This paper examines some different ways to compute the LQs and their results, by using data on residents by citizenship in the metropolitan area of Milan. Stocks of the annual resident population on 1st January 2020, provided by the Statistical Municipality Office of the city and by Istat, are broken down by municipality and sub-municipal areas. Therefore, we analyse and compare the statistical distributions of the LQs of selected foreign population's groups, obtained with the aforementioned approaches. Finally, we propose an alternative measure, the Locational Differential Index (LDI), to take account of density, conveying the spatial component, absent in location quotients (Bressan *et al.*, 2008).

[1] All authors contributed to conceive the idea and develop the application. In particular, F. Benassi wrote Sections 1 and 2, S.M.L. Rimoldi wrote Sections 3 and 4, and M. Crisci wrote Sections 5 and 6.

2. Residential segregation: concept and measures

The concept of residential segregation generally is used to indicate the spatial separation of two or more groups of a population within a given spatial environment, most frequently an urban context (Feitosa *et al.*, 2007).

This general definition evokes a multidimensional process that found in literature various attempt to measure it, applied to various phenomena. It should be noticed that no consensus has been found among scholars about which is the optimal index, although some requisites seem to be necessary: 1) it must not be distorted by the relative size of the minority group in the population as a whole; 2) it must not depend on the overall size of the population and of the area; 3) it must not depend on the number of sub-areas into which the overall area is divided; 4) it must be standardizable, so as to vary between 0 and 1, where 0 indicates the situation in which in each sub-area the ratio between the groups is the same as that observed for the whole region, and 1 corresponds to the situation in which the groups are clearly separated in the sub-areas; 5) it must be sensible to the movement of one or more units from one sub-area to another; 6) it must be invariant to scale transformations in the composition (i.e. either an increase in the absolute level of a particular group in all sub-areas, or an increase in the absolute level of all groups in a particular sub-area).

According to Massey and Denton (1988), the segregation indices can be classified into the categories illustrated in Table 1, according to the dimension they aim to capture.

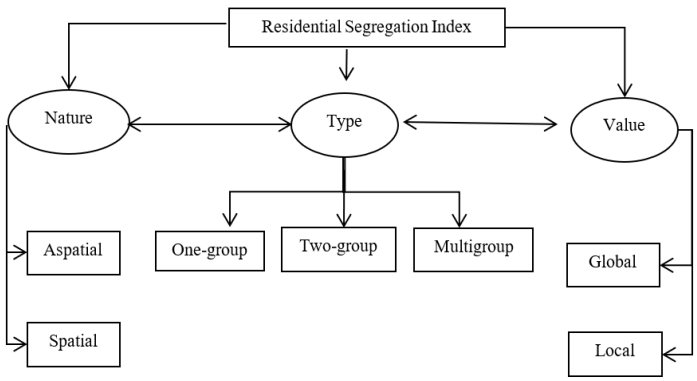
Then, segregation indices can be classified by their essential features (Figure 1): type (one-group, two-group, multigroup), nature (either spatial or aspatial), and value (local or global). According to their type, the one-group indices measure the distribution of a population group compared to the total population, while the two-group indices compare the distribution of two different population groups, and finally, the multigroup indices analyse the distribution of several population groups simultaneously. Considering their nature, aspatial indices are independent from the spatial information about location (shape and/or size of spatial sub-units, and their contiguity), while, on the opposite, spatial indices are based on the spatial setting of sub-units. Finally, as far as their value is concerned, global measures provide a summary value for the area as a whole, while the local measures provide one value for each of the spatial sub-units of the area.

Table 1 – *Dimensions of segregation.*

Dimension	Description
Evenness	Evenness represents the distribution of groups across the spatial sub-units of an area. Evenness indices measure a group's over or under-representation in each spatial sub-unit: the more unevenly a group is distributed across these spatial sub-units, the more segregated it is.
Exposure	Exposure is the degree of potential contact between members of the same group or between members of two groups inside the spatial sub-units. It measures the probability that members of one group will encounter members of their own group (isolation) or another group (interaction) in their spatial unit. Isolation and interaction are complementary: the least isolation (i.e., the highest interaction) means the least segregation.
Concentration	Concentration refers to the physical space occupied by a group. The less of the area a group occupies, the more concentrated it is. According to Massey and Denton (1988), segregated minorities generally occupy a small portion of the area.
Clustering	Clustering refers to another spatial feature, contiguity. The more contiguous spatial sub-units a group occupies -thereby forming an enclave within the overall area- the more clustered and therefore segregated it is.
Centralization	Centralization indices measure the degree to which a group is located in or near the center of the area, which is typically the central business district in a metropolitan area. The closer a group is to the city center, the more centralized and thus segregated it is.

Source: adapted from Martori and Apparicio (2011).

Figure 1 – *Segregation measures by their essential features.*



Source: adapted from Benassi et al. (2016).

Then, by considering in the meantime the two classifications introduced above (by dimensions and by nature), a tentative (not exhaustive) list of the most used indices of segregation has been suggested by Apparicio *et al.* (2008). Here, the

Location Quotients are classified as 'local' indices of segregation but they are not associated to a specific dimension of residential segregation.

3. Location Quotient and residential segregation

Local Quotient (LQ) is a widely known geographic index founded on the work of Walter Isard (1960) that is considered the father of regional science. It is used to measure and map relative distributions or relative concentrations of a character in a subarea compared to the area as a whole (Wheeler, 2005). Originally adopted in regional economic studies (Isard, 1960; Crawley *et al.*, 2013), more recently it has been proposed as a local index in the study of residential segregation of foreign population (Apparicio *et al.*, 2008; Brown and Chung, 2006).

There are different ways to compute LQ. With reference to residential segregation, we initially refer to the definition proposed by Apparicio *et al.* (2014):

$$LQ_i = \frac{x_i/t_i}{X/T} \qquad i = 1, \dots, n \tag{1}$$

given i the spatial sub-unit of the area, x_i the population of group X in spatial sub-unit i, t_i the total population in spatial sub-unit i, X the population of group X in the overall area, T the total population in the overall area. According to the formula (1), LQ points to evaluate the geographic dimension of the index, answering the question "where group X is over/under-represented compared to the average".

However, the (1) can be simply transformed into:

$$LQ_i = \frac{x_i/t_i}{X/T} = \frac{x_i}{t_i} \cdot \frac{T}{X} = \frac{x_i/X}{t_i/T} \tag{2}$$

which aspires to evaluate the pattern of group X, answering the question "how group X behaves compared to the overall population". According to both (1) and (2), LQ appears to measure the dimension of evenness, in segregation.

Indeed, despite their apparent simplicity, the location quotients present some other slippery features. These caveats concern principally the identification of the reference population, their distributions' cut-offs (paying attention to the absence of an upper limit of the LQ) and, accordingly, the way of mapping and commenting on them, all aspects that become crucial in studying residential segregation.

As far as the first caveat is concerned, that is the reference population, it should be observed that traditionally, segregation indices are based on the "spatial assimilation theory" (Massey and Denton, 1988; Massey, 1985) , which extends the concept of assimilation to spatial behaviour (i.e., spatially segregated ethnic groups as far as they become culturally, socially and economically assimilated to the majority group, they tend to spread out in the city).

In the case of LQ we can follow several approaches with fascinating implications. First, by assuming that the population is divided into two groups, X (foreigners), and

Y (Italians), the LQ aims at evaluating the pattern of *X* (minority group) in comparison to *Y* (majority group); it becomes a two-groups index in the dimension of evenness. Second, when groups are more than two, considering the foreign population as a pseudo-total population, the LQ of a particular subgroup *X* of foreigners is a one-group index that tell us which spatial units *i* the group *X* is over/under-represented compared to total foreign population. Finally, we can compare group *X* to total population but *X*; this way to calculate the LQ removes the redundancy of group *X* in calculating the average pattern of total population, and this is particularly important when dealing with sizeable groups. To summarise, LQ can be viewed as a mono or two -groups measure of residential segregation; it is a local measure, but aspatial; LQ has no upper limit in each kind of computation; in the study of residential segregation, LQ's interpretation (referring to spatial assimilation theory) can be ambiguous.

To illustrate the caveats partially faced by some scholars (Crawley *et al*., 2013), and to deepen the LQ distributions' cut-offs and the way of mapping and commenting on them, we use the case study of the Metropolitan area of Milan, in the next section.

4. Case study: The Milan Metropolitan Area

The geographical data refer to the Milan Metropolitan Area (MMA), including the city of Milan and the municipalities of the provinces of Milan and Monza-Brianza: the city of Milan is divided in 88 sub-areas, the NILs (*Nuclei di Identità Locale*), while the rest of the area is considered at municipality level, for a total of 187 units. To highlight differences in the spatial patterns, the MMA is split into five areas, as a result of an aggregation of NIL and municipalities: two areas inside the municipality of Milan (the City center and the Rest of the city); and three areas outside the city of Milan (the hinterland), including the other municipalities of the MMA (the First belt, the Second belt and the Rest of the Metropolitan area). We consider the resident population by country of origin at 1.1.2020; data come from the population register (*Anagrafe*) and Istat. To illustrate the LQs, we identified the two sub-population of Romanians (69,051 individuals) and Chinese (49,006 individuals), the first showing a dispersal pattern outside the city center, the second showing a sparsely clustered pattern.

Figure 2 – *Location quotients of Romanians in the MMA (1.1.2020), by different reference population.*

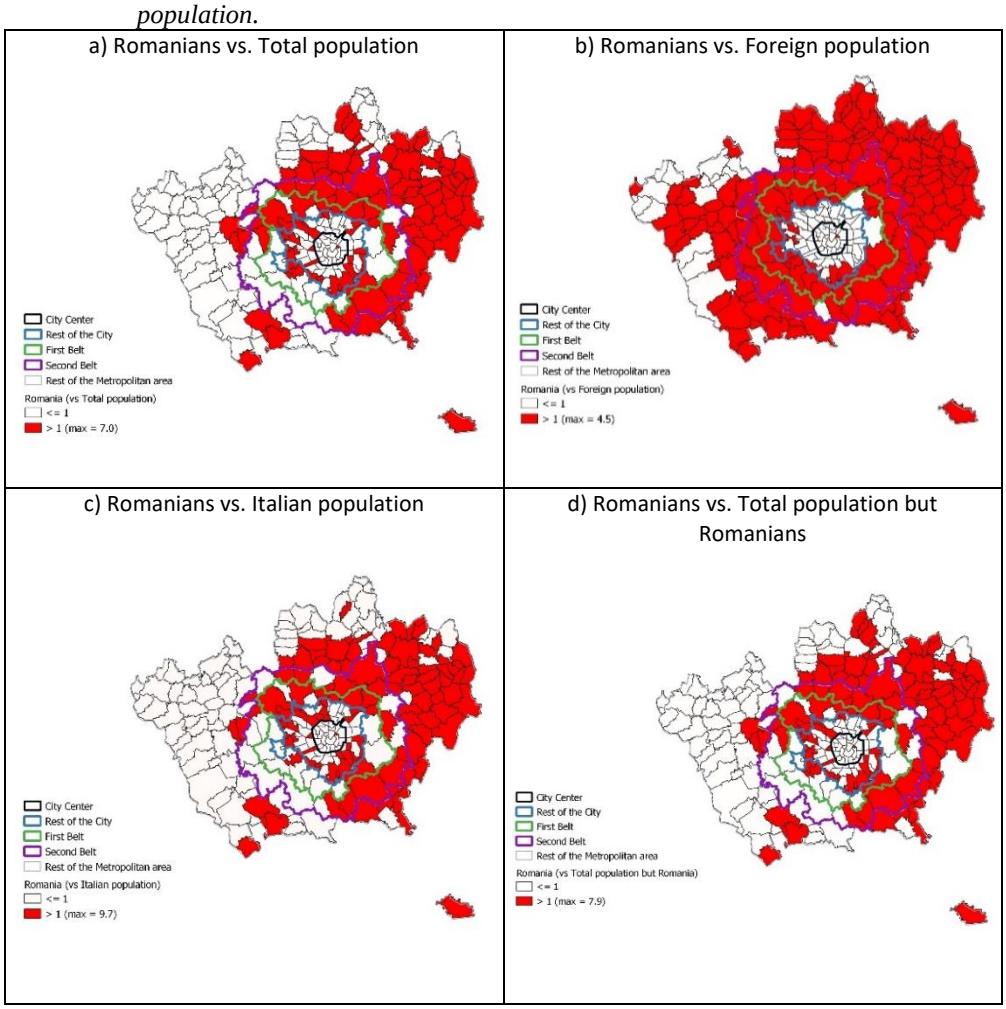
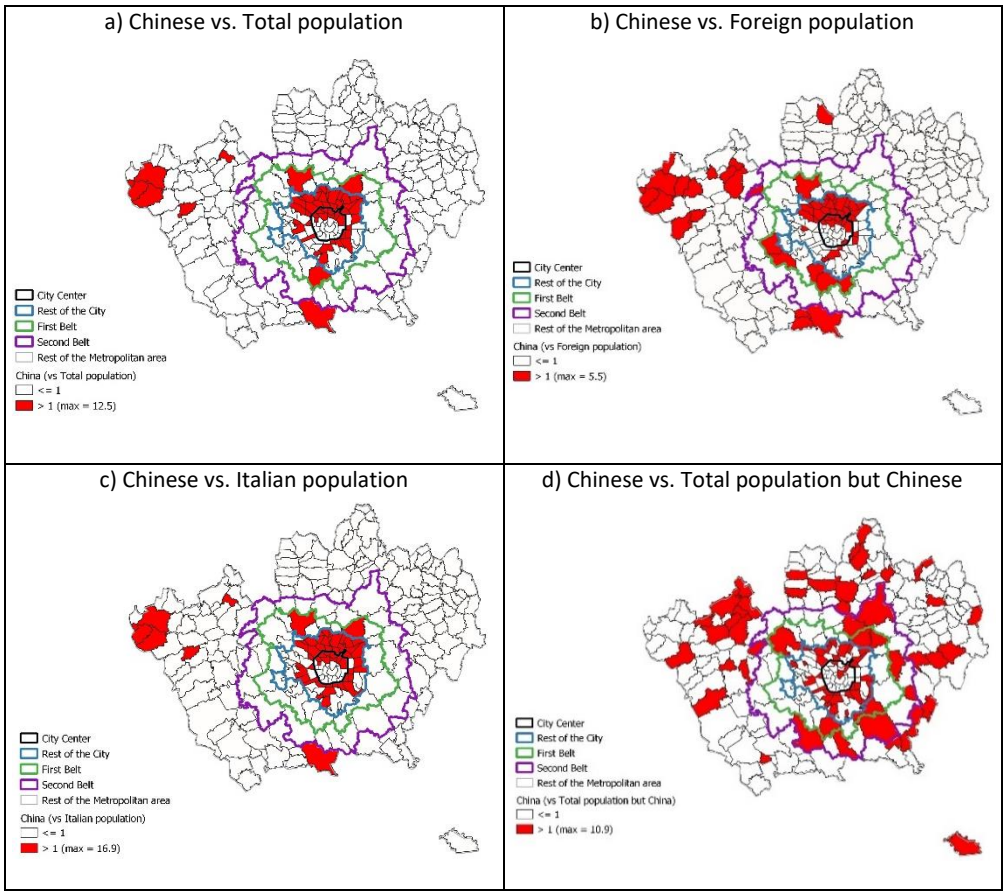


Figure 2 clearly shows how changing the reference population leads to different interpretations of the phenomenon. Notice the negligible differences between figures (a), (c), and (d), due to the relevant size of the majority group (Italians) compared to the size of Romanians; this circumstance is not very frequent, see, for example, the traditional segregation studies concerning the US population that analyse ethnic segregation of large groups like Blacks or Hispanics versus Whites. Therefore, limiting the attention to Figure 2 (a) and (b), we can conclude that Romanians' residential pattern concentrates outside the city centre and in the north-eastern and southern parts of the metropolitan area, when compared to total population, while it

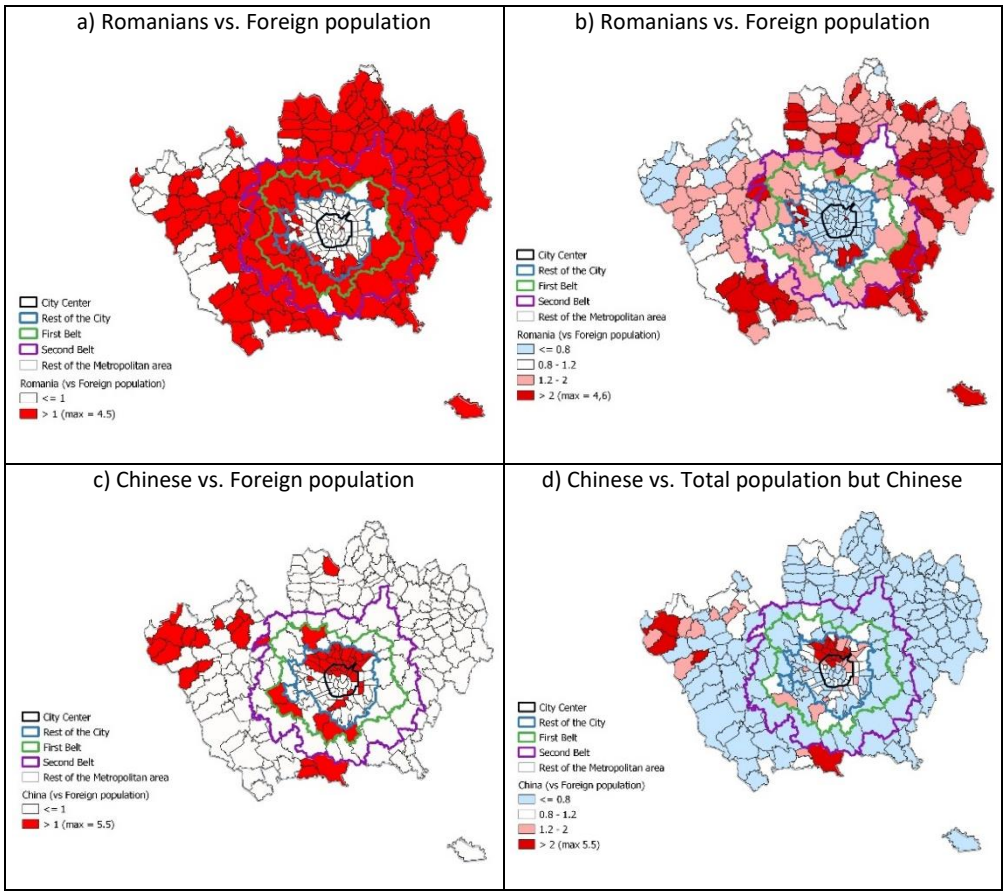
emerges that, compared to foreign population, Romanians show a diffusive residential behaviour outside the city centre.

Figure 3 – *Location quotients of Chinese in the MMA (1.1.2020), by different reference population.*



From Figure 3, differences in interpretation of LQs are more expressive. In particular, the residential pattern of Chinese compared to the total population but China (d), lets emerge a more diffusive behaviour, although visibly clustered (in part also in the city centre), and shows that what is relevant is the difference with the Italians (c).

Figure 4 – *Location quotients of Romanians and Chinese in the MMA (1.1.2020), by different cut-offs.*



Up to now we have just scraped the slipperiness of the information potential of LQs. We have, in fact, only discriminated between the under-representation (LQ < 1) and the over-representation (LQ > 1), where LQ = 1 represents the condition of evenness. Indeed, we can assume that evenness is achieved in a range of values "fairly" close to 1, and we can be interested to let emerge those situations where the group appears remarkably over-represented. Notice that here the interest is more focused on "where".

Taking as an example the LQs of Figures 2 (b) and 3 (b) (reference: foreign population), Figure 4 illustrates how the picture (and interpretation) changes, by adopting different cut-offs of the LQ's distribution. More specifically, it has been compared a binary LQs representation (LQ <= 1 or LQ > 1) and a LQs representation

with four cut-offs: <= 0.8, indicating under-representation; 0.8-1.2, indicating evenness; 1.2-2, indicating over-representation; >2, indicating relevant over-representation. For example, Figure 4 (b) reveals that Romanians tend to concentrate at the borders of the Metropolitan area, especially in a cluster of municipalities in the eastern part of the metropolitan area, while the under-representation of Chinese is largely diffused in the three areas of the MMA hinterland and the Southern part of Milan municipality. Therefore, the adoption of a cut-offs scale that remarks the groups' over-representation suggests the interpretation of the LQ as a concentration index (although aspatial), rather than an evenness index. Remarkable groups' over-representation in some sub-units suggests the interpretation of LQ as concentration index, although in an aspatial form.

5. A proposal to account for population density

Local measures are always affected by problems of robustness, that is the extreme variability due to small or even rare populations in local contexts. In fact, a low density in a certain sub-unit of an area can amplify the relative presence of a certain group of a population, even though the size of the group in that sub-unit is modest or even irrelevant. Therefore, to control for density, we advance the proposal of the Locational Differential Index (LDI) (Bressan *et al.*, 2008), as follows.

Given P_i^K the population of group *K* in the sub-unit *i*, $P_i^{\bar{K}}$ the whole population but *K* in sub-unit *i*, P^K the population of *K* in the whole area, $P^{\bar{K}}$ the whole population but *K* in the whole area, and the d_i the total population density ($d_i = P_i/A_i$, with A_i the surface area of *i*), then:

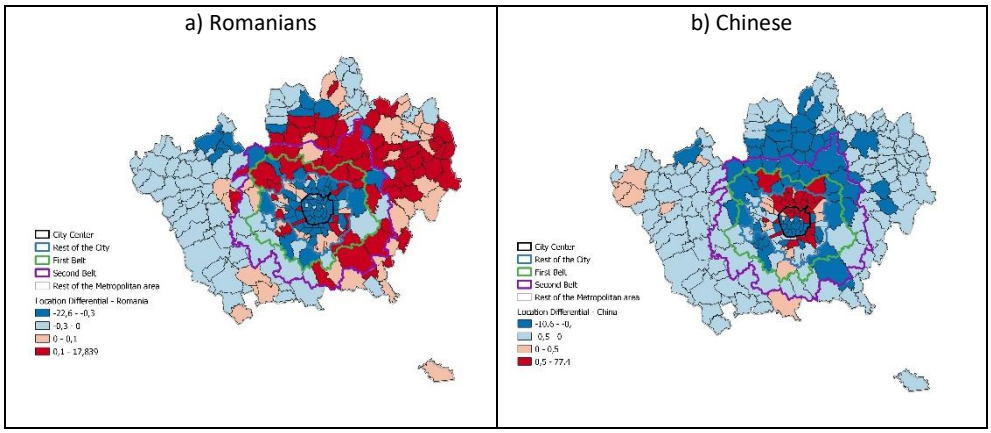
$$LD_i^K = 100 \cdot \left(\frac{P_i^K}{P^K} - \frac{P_i^{\bar{K}}}{P^{\bar{K}}}\right) \cdot d_i \qquad (3)$$

with $-\infty \leq LD_i^K \leq +\infty$.

Notice that LDI compares the share of each group in sub-unit *i* to the share of the rest of the population, weighting for density.

The introduction of density can change a lot the picture of a group distribution over an area. To visualise the effect, we recall the example presented in previous section (Figure 5).

Figure 5 – *Locational differentials of Romanians and Chinese in the MMA (1.1.2020).*



Notice that, accounting for population density, Romanians result more concentrated in some north-eastern municipalities (namely Monza and nearby) where the LQ indicated under-representation. Chinese result strongly over-represented only in the municipality of Milan, above all in some Northern and Southern neighbourhoods.

6. Conclusions

According to literature (Massey and Denton, 1988; Apparicio *et al.*, 2008), LQ can be viewed as a one or two -groups measure of residential segregation; it is a local measure, but aspatial; LQ has no upper limit in each kind of computation.

Moreover, LQ is a local but aspatial index, computed as a ratio of ratios; in the research strand of residential segregation, it can be related to different dimensions (evenness and concentration), producing a different kind of index (one or two groups).

LQ can be very difficult to map and interpret because it has no upper limit; empirical applications conducted on the Milan Metropolitan Area have shown how, in the field of (urban) residential segregation, the interpretation of maps ambiguously depends on the cut-offs scale adopted.

Finally, the absence of a spatial feature enhances ambiguous interpretation.

Provisional results of our ongoing research on the LQ's upper limit make us trust in the possibility of identifying an empirical maximum; this will be the object of our next study.

Acknowledgements

The analyses here presented are part of the project *Immigration, integration, settlement. Italian-Style,* financed by the Italian Ministry for Education, Universities, and Research.

References

APPARICIO P., MARTORI J.X., PEARSON A.L., FOURNIER È., APPARICIO D. 2014. An Open-Source Software for Calculating Indices of Urban Residential Segregation, *Social Science Computer Review*, Vol. 32, No. 1, pp. 117-128.

APPARICIO P., PETKEVITH V., CHARRON M. 2008. Segregation Analyzer. A C#. Net application for Calculating Residential Segregation Indices, *Cybergeo: European Journal of Geography*, Vol. 414, pp. 1-27.

BENASSI F., HEINS F., LIPIZZI F., PALUZZI E. 2016. Measuring residential segregation of selected foreign groups with aspatial and spatial evenness indices. A case study. In PERNA C., PRATESI M. and RUIZ-GAZEN A. (Eds.) *Studies in Theoretical and Applied Statistics*, Vol. 227, Springer Proceedings in Mathematics & Statistics, Springer, pp. 189-199.

BRESSAN M., FANFANI D., RADINI M. 2008. Prato: dalla città fabbrica alla società multiculturale. In FREGOLENT L. (Ed.) *Periferia e Periferie*, Roma: Aracne, pp. 334-379.

BROWN L.A., CHUNG S.Y. 2006. Spatial Segregation, Segregation Indices and the Geographical Perspective, *Population, Space and Place*, Vol. 12, No. 2, pp. 125-143.

CRAWLEY A., BEYNON M., MUNDAY M. 2013. Making Location Quotients More Relevant as a Policy Aid in Regional Spatial Analysis, *Urban Studies*, Vol. 50, No. 9, pp. 1854-1869.

FEITOSA F.F., CÂMARA G., MONTIERO A.M.V., KOSCHITKI T., SILVA M.P.S. 2007. Global and Local Spatial Indices of Urban Segregation, *International Journal* of *Geographical Information Science*, Vol. 21, No.3, pp. 299-323.

IGLESIAS-PASCUAL R., PALOMA V., GARCÍA-RAMÍREZ M. 2019. Contextual Factors and Prejudice at the Beginning of the Migrant Influx: The Moroccan Case in Seville, Spain, *Population, Space and Place*, Vol. 25, No. 3, e2204.

ISARD W. 1960. Methods of regional analysis: an introduction to regional science. Cambridge: The MIT Press.

MARTORI J. C., APPARICIO P. 2011. Changes in Spatial Patterns of the Immigrant Population of a Southern European Metropolis: The Case of the

Barcelona Metropolitan Area (2001–2008), *Tijdschrift Voor Economische En Sociale Geografie*, Vol. 102, No. 5, pp. 562-581.
MASSEY D.S., DENTON N.A. 1988. The Dimension of Residential Segregation, *Social Forces*, Vol. 67, No. 2, pp. 281-315.
MASSEY D.S. 1985. Ethnic Residential Segregation: A Theoretical Synthesis and Empirical Review, *Sociology and Social Research*, Vol. 69, pp. 315-350.
WHEELER J.O. 2005. Geography. In Kimberly K.L. (Ed.) *Encyclopedia of Social Measurement*, Elsevier, pp. 115-123.

SUMMARY

Location quotient as a local index of residential segregation. Theoretical and applied aspects

The location quotient (LQ), a ratio of ratios, is a widely known geographic index (Isard, 1960). It is used to measure and map relative distributions or relative concentrations of a character in a sub-area compared to the area as a whole (Wheeler, 2005). It has been recently proposed as a local index in the study of residential segregation of foreign population (Apparicio *et al.*, 2008; Benassi *et al.*, 2016). The LQs are particularly useful when applied to a metropolitan area where they allow to identify the spatial units in which a population group is under-represented (LQ<11) or conversely, over-represented (LQ>1). Despite their apparent simplicity, the location quotients present some slippery features this contribution tries to reflect on. This paper examines some different ways to compute the LQs and their results, by using data on residents by citizenship in the metropolitan area of Milan. We analyse and compare the statistical distributions of the LQs of selected foreign population's groups. Finally, we propose an alternative measure, the Locational Differential Index, to take account of density, conveying the spatial component, absent in location quotients.

Federico BENASSI, Istat, benassi@istat.it
Massimiliano CRISCI, IRPPS-CNR, m.crisci@irpps.cnr.it
Stefania M.L. RIMOLDI, Università di Milano-Bicocca, stefania.rimoldi@unimib.it

SPATIAL DISTRIBUTION OF SERIOUS TRAFFIC ACCIDENTS AND ITS PERSISTENCE OVER TIME IN MILAN

Claudio Caterino, Luigi M. Solivetti

1. Introduction

Over the last decades, the road safety issue has drawn worldwide attention. Traffic accidents (hereafter TA) involve great costs both at an economic and at a human level and are a major cause of premature mortality, overly affecting young people. To tackle this problem, the EU created a Community Road Accident database meant to monitor TA, and in 2001 set for its member states the target of halving in ten years the number of TA victims, meeting, however, with only partial success. The United Nations, in turn, in 2015 incorporated road safety into its Sustainable Development Goals, aiming to halve the number of global victims from TA by 2030. Italy, too, paid attention to the TA issue at both the national and regional level. Istat (Italian National Statistical Office) took care of updating, together with Automobile Club d'Italia, its TA recording framework, and developed new TA indicators (Broccoli and Bruzzone, 2019).

According to the literature, TA causes can be divided into a few categories, as from the Haddon Matrix of accident factors: a) reckless driving (non-compliance with the minimum safe distance, speeding, using a mobile phone while driving etc.: Gariazzo *et al*., 2018), b) driver's poor conditions (drugs and alcohol abuse, tiredness, old age: Rolison *et al*., 2018), c) vehicle poor conditions (faulty braking etc.: Moodley and Allopi 2008), and, last, d) flaws in the road infrastructure (Demasi *et al*., 2018, Ishtiaque, 2013).

Often, TA are the consequence of more than one factor: e.g., the driver's poor conditions affect driving behaviour, causing a breach of the traffic rules (Ishtiaque, 2013); the road infrastructure flaws, when tackled without an increased caution, multiply the TA probabilities (Wang *et al*., 2009).

In an urban setting, where crossroads and intersections are plentiful, the non-compliance with the right of way – either intentional or due to inattention – is the primary cause of TA (Reason *et al*., 1990). Speeding is a further leading cause of serious TA. Proof of this is that traffic congestion – which affects the vehicle speed – mitigates the TA seriousness. However, congestion does not reduce the total number of TA significantly either in the urban setting or on high-speed roads

(Noland and Quddus, 2005). Since most of TA in the urban setting occur during rush hour, speed plays a less momentous role in urban TA.

Apart from the road infrastructure, all other TA factors are contingent, and they can be ascribed to a random combination of events. The TA physical context, and therefore the road infrastructure, represents the non-contingent component: it does not forgive reckless driving or poor conditions of both driver and vehicle. This non-random nature of the infrastructure goes beyond the effects of differences such as those between high- and low-traffic areas or few- and many-intersections zones.

This study intends to analyse the road infrastructure role in TA occurring in an urban setting. We hypothesise that the TA spatial distribution is highly differentiated and that such a distribution persists over time, owing to the infrastructure stability. Therefore, we intend not only to ascertain the presence of blackspots, namely places where TA cluster, but especially to verify their over-time persistence.

Several studies having been carried out using geographic information systems (GIS) to identify spatial models underlying TA in towns and cities (Erdogan, 2009). These studies are based on partitions of the territory by zones (administrative subdivisions, geometric grids: Le *et al*., 2020) or by features (crossroads, intersections, road segments and tunnels: Erdogan *et al*., 2008; Wang *et al*., 2009). A suitable combination of partition criteria and blackspot identification methods contributes to better results. Indeed, the territory partition represents a crucial problem for any spatial analysis of TA (Ghadi and Török, 2018).

It is noteworthy that both spatial smoothing techniques, such as the Kernel Density Estimation (KDE) and clusterisation techniques, such as the Getis-Ord Gi*, meant to calculate and plot TA dispersal, constitute planar methods. In contrast, the TA distribution does not have a two-dimensional nature. Indeed, an analysis aiming to identify places characterised by a *relatively high TA incidence* should consider additional information, such as traffic fluxes, congestions, and vehicles speed. To bypass this planar methods limitation, it is possible to adopt some expedient. In particular, it is possible to divide thoroughfares into segments and to calculate TA per segment (Bíl *et al.*, 2013). By doing this, one can normalise additional TA factors such as traffic flows – which are expected to be constant along the thoroughfare – and the average speed, which at times can be measured. Unfortunately, these expedients are not available in urban settings, owing to the intricacy of the road network. In urban settings, however, it is possible to identify places characterised by an *absolutely high TA incidence*. Such identification has an intrinsic utility as to interventions and policies meant to reduce TA incidence.

The specific issue of the persistence of TA spatial distribution patterns – the present study subject – has received little attention in current literature. Indeed, the persistence of TA distribution patterns has been mainly analysed as to *time* (Kingham *et al.*, 2011; Duarte Monedero *et al*., 2021) – also in order to boost the

data reliability – rather than as to *space*. A few studies have analysed the persistence of TA distribution patterns in terms of both space and time, but as a secondary issue (Cheng *et al*., 2018). Therefore, the persistence of TA spatial distribution patterns has not been satisfactorily tested with statistical tools.

2. Data and methods

Our dataset (from Istat 2021, Section *Incidenti Stradali*: http://dati.istat.it/) comprises accidents with casualties in Milan (the Italian city with the highest density of TA) in the five time-series waves from 2015 to 2019. Accidents georeferentiation was obtained from the TA street addresses by means of an experimental elaboration of geocoding (Cimbelli and Caterino, 2016). We divided the municipality territory (181.7 square km) into administrative and geometric zones and counted the TA number in those zones. We used QGIS, and OpenStreetMap (OSM) as background.

When dividing the territory, we moved from relatively larger units to smaller ones. In particular, from 88 administrative units (*Nuclei Identità Locale*, i.e., Local Identity Nuclei, hereafter NIL) with an average surface of 2.06 square km, and from a specially-made geometric grid of 88 squares (each with the same surface of the average NIL), to a second, specially-made grid comprising 5568 hexagons circumscribing 200-metre circles, and equivalent to a 34,641 square metres area.

The shift from administrative units to geometric grids was motivated by the fact that specially-made polygon grids, as opposed to administrative units, are of equal size and *do not* respond to administrative purposes. Since the present study goal is to test the hypotheses of a substantial difference in TA between the territorial units, and of the over-time persistence of this difference, geometric grids are better suited to the task. Their very construction would exclude the possibility that the abovementioned differential derives from the unequal size of the territorial units or from the administrative reasons behind a certain type of territorial partition (for instance, a partition into high- and low-urbanised sections etc.).

The shift from larger to smaller units, in turn, is motivated by the fact that it is advisable to check whether the differential in TA and its over-time persistence are the results of short-range TA blackspots, rather than spurious outcomes ascribable to average values of larger territorial units (Hashimoto *et al.*, 2016).

The comparison between the territorial units was carried out by calculating the TA number in each unit, the differential in TA between units, and the over-time consistency of the TA spatial distribution. As for the grid of smaller, and therefore more numerous, polygons (N = 5568), the comparison was conducted by means of a hotspot analysis Getis-Ord Gi* (Getis and Ord, 1992), which in turn used as weights a Queen's case contiguity matrix. The z scores obtained through this

analysis were then used to verify the difference in scores between the units and the over-time persistence of the spatial distribution of such scores.

In order to measure the over-time consistency of the TA spatial distribution, we made recourse to the Cronbach Alpha coefficient (Cronbach, 1951), and we calculated it for all three territorial partitions of the Milan municipality considered in this study, namely the NIL, the grid of squares and the grid of hexagons (1):

$$Cronbach\ \alpha_{[0,\ 1]} = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma^2 x}\right) \quad (1)$$

where k is the number of items (here, the 2015 to 2019 waves of TA counts – or of hotspot z scores – in the territorial units); σ^2_i is the variance of TA of the *i-th* wave; σ^2_x is the variance of the sum of all the waves counts for each observation unit.

The present study used factor analysis (principal factor method) as a further measure of the over-time consistency of the TA spatial distribution. The first factor would produce the same loadings (with value = 1) for all the waves, provided their TA exhibit correlations equal to 1. Therefore, the loadings std dev. would be inversely proportional to the over-time persistence of the TA spatial distribution. We added to the loadings a factor analysis post-estimation, namely the Kaiser-Meyer-Olkin (KMO) statistic. KMO measures the adequacy, for each variable in the model and for the complete model, of the data used for the factor analysis (2):

$$\mathrm{KMO}_{[0,\ 1]} = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} p_{ij}^2} \quad (2)$$

where r is the correlation between the variables i and j (here, the time-series waves of TA values – or of hotspot z scores – in the territorial units); and p is the partial correlation between the said variables.

The information provided by the KMO statistic is different from, and complementary to, the Alpha coefficient. While Alpha is a measure of the degree of closeness of the values for each observation unit over the various items, the KMO statistic is a summary of how small the partial correlations are relative to the original (zero-order) correlations between the items/variables. The partial correlation for each pair of variables is comprised of the correlation between those variables after partialling out the influence of all other variables. KMO, ultimately, is a measure of the proportion of variance among variables that might be ascribed to an underlying factor. In our case, we suppose that this factor is represented by an ecological component underlying the TA.

3. Results

Table 1 shows that the number of TA with casualties, of TA with pedestrian

casualties, of deaths and casualties due to TA in the Milan municipality remained rather stable over time, with a clear declining trend as to the number of deaths only.

Figure 1 shows that, in the 88 administrative NIL of Milan, the TA distribution is clearly differentiated. The over-time consistency of the TA spatial distribution can be visually appraised by means of the correlation matrix shown in Figure 1. Table 2 adds further information. The std dev. of the TA in the observation units is high. The Alpha coefficient, close to 1 (0.997), proves that the TA spatial distribution is highly persistent over time. The std dev. of the factor loadings is fractional (0.002), and the KMO measure (0.929) proves that almost all the variance between variables might be ascribed to an underlying factor.

Table 1 – *Traffic accidents with casualties. Milan municipality: 2015 to 2019.*

TA/Victims	2015	2016	2017	2018	2019
TA with casualties	8,729	8,935	8,559	8,523	8,263
TA with pedestrian casualties	1,259	1,368	1,317	1,365	1,305
TA deaths	53	50	53	49	34
TA casualties	11,465	11,905	11,123	11,112	10,743

Source: Istat 2021 – Yearly Survey on Road Accidents resulting in death or injury.

Figure 1 – *Choropleth map of traffic accidents with casualties. Milan municipality: 2015 to 2019. Partition: 88 NIL.*

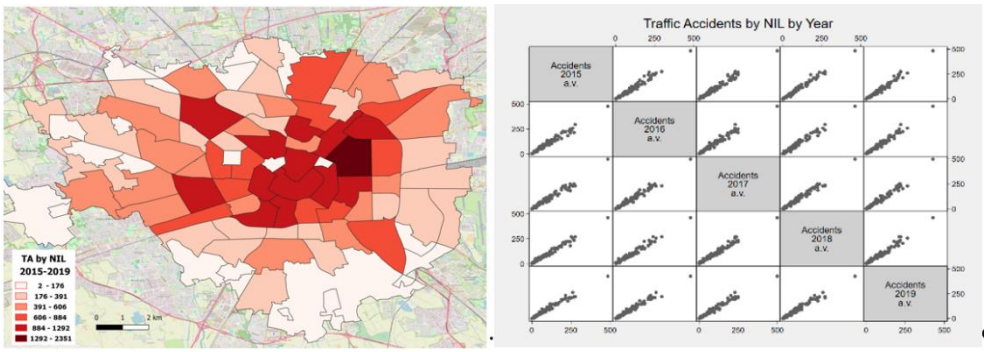


Figure 2 shows the TA distribution in the 88 square cells of the specially-made geometric grid covering the Milan territory. One can notice that in this case, too, the TA spatial distribution is uneven. Table 3 shows that the average std dev. of the TA over the five-year span is, in the case of the geometric grid, even higher than that for the NIL (Table 2). Concurrently, the Alpha coefficient for the 88 cells is higher than the NIL one. The loadings std dev. is slightly higher while the KMO measure is slightly lower. All in all, these findings prove that the differential in TA between the territorial units and the over-time consistency of the TA spatial distribution do not depend on the administrative nature of subdivisions such as the NIL.

Table 2 – *Traffic accidents with casualties. Milan municipality: 2015 to 2019. Partition: 88 NIL.*

Variables	Territ. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA with casualties 2015	88	98.56	85.66	0.9955	0.9931	0.9141
TA with casualties 2016	88	100.69	83.80	0.9958	0.9909	0.9372
TA with casualties 2017	88	96.39	82.95	0.9960	0.9894	0.9489
TA with casualties 2018	88	95.77	80.93	0.9956	0.9927	0.9118
TA with casualties 2019	88	93.16	76.34	0.9959	0.9899	0.9362
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	88	96.91	81.94	-0.9966	[0.0017]	{0.9294}

Figure 2 – *Choropleth map of traffic accidents with casualties, and correlation matrix. Milan municipality: 2015 to 2019. Partition: 88 squares.*

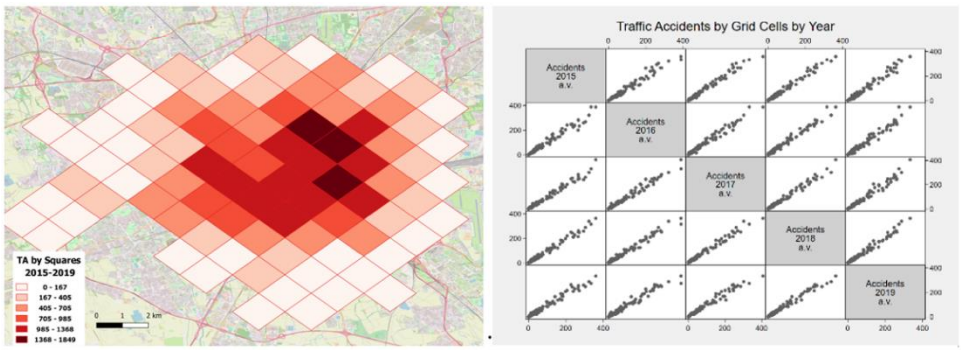


Table 4 tests the hypothesis that the differential in TA between territorial units is just the consequence of an uneven distribution of the road network extent over the same units. The results prove that – even when controlling for the road network extent – the differences between the territorial units remain vast (average std dev. 2.26 vs the five-year mean of 3.07). Concurrently, the Alpha coefficient changes only slightly (0.994), and the same occurs with the loadings std dev., while the KMO measure is even higher. The aforesaid hypothesis, therefore, should be rejected.

Table 5, in turn, intends to check whether the territorial units' differential in TA and the over-time consistency of the TA spatial distribution persist when selecting only a portion of the TA: that with pedestrian casualties. Table 5 shows that, even in the case of pedestrian casualties, the differences in the TA spatial distribution remain vast (with an average std dev. higher than the mean of the five-year period), while the consistency of the TA spatial distribution decreases fractionally (0.986). The loadings std dev. is higher than in the previous samples, but the KMO measure stays on values similar to the previous ones.

Table 3 – *Traffic accidents with casualties. Milan municipality: 2015 to 2019. Partition: 88 squares.*

Variables	Territ. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA with casualties 2015	88	96.73	94.75	0.9964	0.9937	0.9213
TA with casualties 2016	88	98.41	95.79	0.9969	0.9903	0.9411
TA with casualties 2017	88	94.69	93.28	0.9965	0.9928	0.9351
TA with casualties 2018	88	93.86	90.93	0.9960	0.9965	0.8918
TA with casualties 2019	88	91.44	86.06	0.9968	0.9908	0.9495
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	88	95.03	92.16	(0.9972)	[0.0025]	{0.9272}

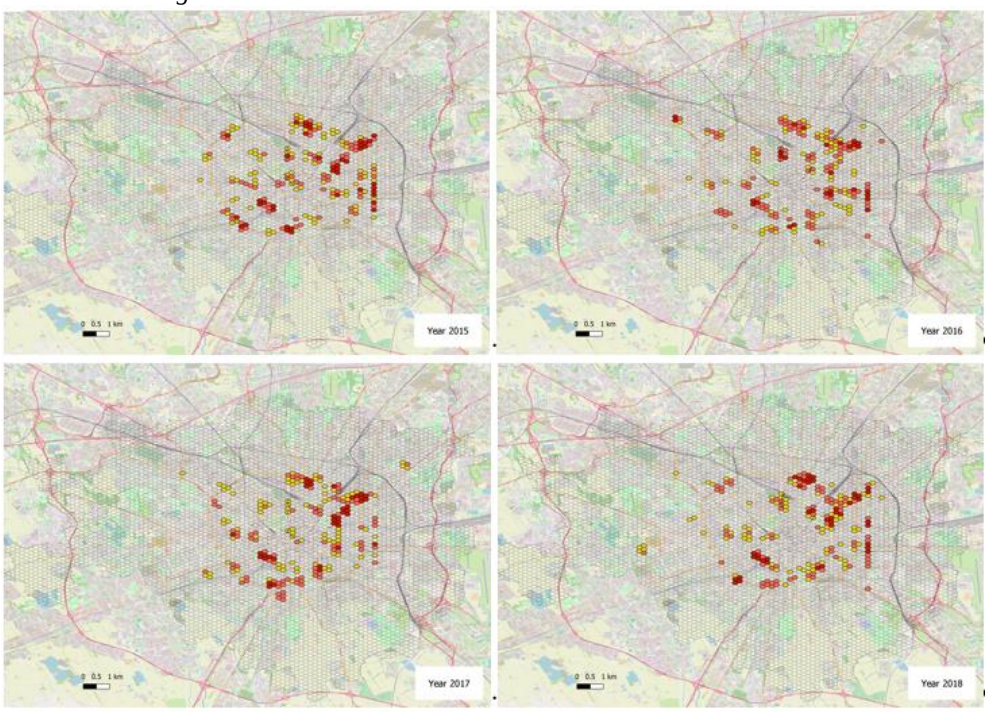
Table 4 – *Traffic accidents with casualties per kilometre of road. Milan municipality: 2015 to 2019. Partition: 88 squares.*

Variables	Terr. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA w. cas. per Km. 2015	88	3.12	2.34	0.9926	0.9882	0.9271
TA w. cas. per Km. 2016	88	3.18	2.36	0.9937	0.9805	0.9530
TA w. cas. per Km. 2017	88	3.05	2.27	0.9929	0.9860	0.9392
TA w. cas. per Km. 2018	88	3.05	2.20	0.9925	0.9896	0.9158
TA w. cas. per Km. 2019	88	2.98	2.11	0.9933	0.9835	0.9401
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	88	3.07	2.26	(0.9944)	[0.0037]	{0.9348}

Figure 3 and Table 6 check the over-time consistency of the TA spatial distribution when moving from larger to smaller territorial units (namely, the 5568 hexagons). The results show that the hotspot analysis z scores present vast differences in their spatial distribution (with an average std dev. much higher than the mean of the five years), an Alpha equal to 0.981, while both the loadings std dev. and the KMO measure hang around values not far from the previous ones.

Lastly, Table 7, based on the said hexagons, shows the over-time consistency of the z scores spatial distribution for those units regarded as real hotspots by the Getis-Ord Gi* procedure: namely, units exhibiting – in the first year of the time series – z scores with p value <0.10. By excluding units with low z scores, the sample shrank from 5568 to 203 units, generating, as expected, a fall in the std dev. when compared to the sample mean. However, Alpha remains high (0.892), and the variations registered by the loadings std dev. and by the KMO statistic are modest. These results suggest that the over-time consistency of the TA spatial distribution concerns the territorial units with higher TA incidence and not only the lower-incidence units.

Figure 3 – *Maps of hotspots of traffic accidents with casualties from Getis-Ord Gi*, and correlation matrix. Milan municipality: 2015 to 2019. Partition: 5568 hexagons. Yellow hexagons = 90% confidence level; orange hexagons = 95% c.l.; red hexagons = 99% c.l.*



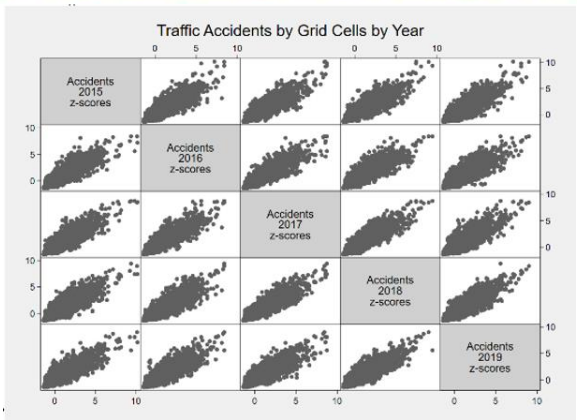


Table 5 – *Traffic accidents with pedestrian casualties. Milan municipality: 2015 to 2019. Partition: 88 squares.*

Variables	Terr. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA w. pedestrian cas. 2015	88	14.17	14.71	0.9834	0.9541	0.9472
TA w. pedestrian cas. 2016	88	15.30	16.05	0.9806	0.9744	0.9035
TA w. pedestrian cas. 2017	88	14.84	15.83	0.9813	0.9695	0.9146
TA w. pedestrian cas. 2018	88	15.30	15.92	0.9810	0.9713	0.9043
TA w. pedestrian cas. 2019	88	14.66	15.28	0.9839	0.9517	0.9280
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	88	14.85	15.56	(0.9856)	[0.0105]	{0.9190}

Table 6 – *Z scores of traffic accidents with casualties. Milan municipality: 2015 to 2019. Partition: 5568 hexagons.*

Variables	Terr. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA w. cas. (z scores) 2015	5568	−0.0012	0.685	0.9757	0.9516	0.9329
TA w. cas. (z scores) 2016	5568	−0.0011	0.677	0.9759	0.9510	0.9319
TA w. cas. (z scores) 2017	5568	−0.0019	0.682	0.9749	0.9578	0.9235
TA w. cas. (z scores) 2018	5568	−0.0009	0.682	0.9753	0.9546	0.9281
TA w. cas. (z scores) 2019	5568	−0.0019	0.662	0.9770	0.9427	0.9412
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	5568	−0.0014	0.678	(0.9805)	[0.0056]	{0.9315}

Table 7 – *Z scores of traffic accidents with casualties (p <0.10). Milan municipality: 2015 to 2019. Partition: 203 out of 5568 hexagons.*

Variables	Terr. Units	Mean	Std. Dev.	Alpha	Factor 1 loadings	KMO
TA w. cas. (z scores) 2015	203	2.23	0.560	0.8751	0.7460	0.9068
TA w. cas. (z scores) 2016	203	1.86	0.672	0.8760	0.7430	0.9030
TA w. cas. (z scores) 2017	203	1.95	0.691	0.8629	0.8036	0.8812
TA w. cas. (z scores) 2018	203	1.95	0.757	0.8585	0.8253	0.8648
TA w. cas. (z scores) 2019	203	1.85	0.721	0.8690	0.7791	0.8841
Av. val., (tot. Alpha), [std. dev.], {tot. KMO}	203	1.97	0.680	(0.8919)	[0.0358]	{0.8867}

4. Conclusions

This study analysed the problem of traffic accidents in the context of large cities. We hypothesised that TA-intensive areas retain this characteristic over time. To test this hypothesis, we used data concerning TA with casualties that occurred over a five-year period in the Milan municipality. We divided the territory into three types of partition in order to check whether the hypothesis held no matter the partition. We

found, in all three partitions, vast differences in their TA spatial distribution, even when controlling for the road network extent in the observation units. Moreover, we found a robust over-time persistence of the TA spatial distribution. In the case of TA with pedestrian casualties, the TA spatial distribution exhibited an over-time persistence similar to that of the generic case of TA with casualties. Further findings showed that the over-time persistence concerned the most TA-intensive territorial units, namely TA blackspots, and not only the TA-light units. The persistence of the TA spatial distribution was measured through multiple procedures: the Cronbach Alpha, the variance between the factor loadings, and the KMO statistic. Their results suggest that the high differential in TA spatial distribution and its persistence over time are not contingent and derive from tendentially stable features of the road infrastructure. All this allows the identification of areas in need of targeted interventions on the infrastructure, while optimising scarce resources in the perspective of a TA reduction in tune with national and supranational guidelines.

5. Limitations, and room for further research

This study analysed TA spatial distribution and its persistence by administrative units and various grid cells. Further territorial partitions could be profitably used: e.g., zones around road features such as intersections, crossways, and road segments. As for the methods, as well, more could be done. Our approach revolved around the over-time *internal consistency* of TA location data because we intended to investigate the persistence of traffic accidents in certain territorial units, rather than the accidents proximity or the impact of close-by zones on the accidents distribution. TA distribution, however, is suitable to spatial analyses using contiguity and distance matrices, kernel density estimations and random point process techniques, to mention just a few other geo-statistical approaches.

References

BÍL M., ANDRÁŠIK R., JANOŠKA Z. 2013. Identification of Hazardous Road Locations of Traffic Accidents by Means of Kernel Density Estimation and Cluster Significance Evaluation, *Accident Analysis and Prevention*, Vol. 55, pp. 265-273.

BROCCOLI M., BRUZZONE S. 2021. Utilizzo di Open Street Map per il calcolo di indicatori per l'incidentalità stradale sulla rete viaria italiana, *Statistica Sperimentale*, Istat, Roma.

CHENG Z., ZU Z., LU J. 2018. Traffic Crash Evolution Characteristic Analysis and Spatiotemporal Hotspot Identification of Urban Road Intersections, *MDPI*

Sustainability, Vol. 11, No.1, pp. 1-17.
CIMBELLI A., CATERINO C. 2016. Georeferenziazione degli incidenti stradali e restituzione cartografica. In *Workshop Incidentalità Stradale,* Roma: Istat. https://www.slideshare.net/slideistat/alessandro-cimbelli-claudio-caterino-georefenziazione-degli-incidenti-stradali-e-restituzione-cartografica
CRONBACH, L. 1951. Coefficient Alpha and the Internal Structure of Tests, *Psychometrika*, Vol. 16, No. 3, pp. 297-334.
DEMASI F., LOPRENCIPE G., MORETTI L. 2018. Road Safety Analysis of Urban Roads, *MDPI Safety*, Vol. 4, No. 4, pp. 1-16.
DUARTE MONEDERO B., GIL-ALANA L.A., VALBUENA MARTÍNEZ M.C. 2021. Road Accidents in Spain: Are they Persistent? *IATTS Research*, Vol. 45, No. 3, pp. 317-325.
ERDOGAN S. 2009. Explorative Spatial Analysis of Traffic Accident Statistics and Road Mortality among the Provinces of Turkey, *Journal of Safety Research*, Vol. 40, No. 5, pp. 341-351.
ERDOGAN S., YLMAZ I., BAYBURA T., GULLU M. 2008. Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar, *Accident Analysis and Prevention*, Vol. 40, No. 1, pp.174-181.
GARIAZZO C., STAFOGGIA M., BRUZZONE S., PELLICCIONI A., FORASTIERE F. 2018. Association between Mobile Phone Traffic Volume and Road Crash Fatalities: A Population-Based Case-Crossover Study, *Accident Analysis & Prevention*, Vol. 115, pp. 25-33.
GETIS A., ORD, J.K. 1992. The Analysis of Spatial Association by Use of Distance Statistics, *Geographical Analysis*, Vol. 24, No. 3, pp. 189-206.
GHADI M., TÖRÖK Á. 2019. A Comparative Analysis of Blackspots Identification Methods and Road Accident Segmentation Methods, *Accident Analysis and Prevention*, Vol. 128, pp. 1-7.
HASHIMOTO S., YOSHIKI S., SAEKI R., MIMURA Y., ANDO R., NANBA S. 2016. Development and Application of Traffic Accident Density Estimation Models Using Kernel Density Estimation, *Journal of Traffic and Transportation Engineering*, Vol. 3, No. 3, pp. 262-270.
ISHTIAQUE A. 2013. Road Infrastructure and Road safety, *Transport and Communications Bulletin for Asia and the Pacific*, No. 83, pp. 19-25.
KINGHAM S., SABEL C.E., BARTIE P. 2011. The Impact of the 'School Run' on Road Traffic Accidents, *Journal of Transport Geography*, Vol. 19, No. 4, pp. 705-711.
LE K.G., LIU P., LIN T.L. 2020. Determining the Road Traffic Accident Hotspots Using GIS-Based Temporal-Spatial Statistical Analytic Techniques in Hanoi, Vietnam, *Geo-Spatial Information Science*, Vol. 23, No. 2, pp. 153-164.
MOODLEY S., ALLOPI D. 2008. An Analytical Study of Vehicle Defects and their

Contribution to Road Accidents. In *Proceedings of the 27th Southern African Transport Conference*, Pretoria: SATC, pp. 469-479.

NOLAND R.B., QUDDUS M.A. 2005. Congestion and Safety: A Spatial Analysis of London, *Transportation Research Part A: Policy and Practice*, Vol. 39, No. 7/9, pp. 737-754.

REASON J., MANSTEAD A., STRADLING S., BAXTER J., CAMPBELL K. 1990. Errors and Violations on the Roads: A Real Distinction? *Ergonomics*, Vol. 33, No. 10/11, pp. 1315-1332.

ROLISON J., REGEV S., MOUTARI S., FENEEY A. 2018. What Are the Factors that Contribute to Road Accidents? *Accident Analysis and Prevention*, Vol. 115, pp. 11-24.

WANG C., QUDDUS M.A., ISON S.G. 2009. The Effects of Area-Wide Road Speed and Curvature on Traffic Casualties in England, *Journal of Transport Geography*, Vol. 17, No. 5, pp. 385-395.

SUMMARY

Spatial distribution of serious traffic accidents and its persistence over time

This study intended to analyse the spatial distribution of serious road accidents in large cities to test the hypothesis of its over-time persistence.

Yearly road accidents with casualties in the Milan municipality from 2015 to 2019 constituted this study data. In order to analyse the accidents spatial distribution, we divided the municipality territory according to three different partitions: one of 88 administrative units, two specially-made grids of respectively 88 squares and 5568 hexagons. We evaluated the over-time persistence of the TA spatial distribution through correlations, Cronbach Alpha, factor loadings and the KMO statistic applied to the accidents densities and Getis-Ord Gi* z scores of the units within the three territorial partitions.

Repeated evidence emerged of a high and persistent differential in accidents between the territorial units, suggesting the existence of an infrastructure non-contingent factor underlying road accidents.

Claudio CATERINO, Istat, caterino@istat.it
Luigi M. SOLIVETTI, Sapienza University of Rome, luigi.solivetti@uniroma1.it

AN ANALYSIS ON CONSUMER PERCEPTION VIA "STEREOTYPE CONTENT" MODEL BETWEEN SUSTAINABLE BRAND AND COUNTRY OF ORIGIN

Gabriella Schoier, Giovanna Pegan, Stefano Fadini

1. Introduction

This paper aims to analyze the Country of Origin (COO) effect and the consumer's intention to purchase green brands, applying the revised Stereotypic Content Model, in which the competence dimension is combined with the warmth dimension divided into two stereotypic sub-dimensions of morality and sociality. While previous analyses have investigated the role of stereotypes of competence and warmth with regard to Country of Brand (COB) and green brands, so far the dimensions of morality and sociality have not been considered separately. This research aims to investigate how the competence and the two components of the warmth (morality and sociality) of the green brand and its country of origin can positively/negatively be related to the perceived quality and purchase intentions of a green brand.

Considering the importance of understanding if and how COB stereotypes can be related to those of the green brands and how this can be related to the perception of quality and the purchase intention of the brand, an exploratory quantitative research has been conducted on a sample of Italian consumers. In particular, the empirical research has been developed considering three COO: Italy, Germany and USA and two brands: green cars and green fashion clothing. The data, processed using SPSS software, have been analyzed through different statistical analysis.

The results suggest that the dimension of competence related to the green brand and its country of origin seems to affect the consumer's perception of quality and purchasing intentions; as well as morality and the combination of morality and competence, particularly for Germany and the United States. The same is not true in the case of Italy where the dimension of sociality becomes more relevant. Moreover, perceiving the moral and/or competent COO could help the green brand to be perceived positively. This research can provide useful suggestions to managers, committed to promoting green consumer styles, in terms of strategic positioning.

2. Theoretical framework

2.1 Brand origin and country of brand

In the global branding scenario, where manufacturers use multiple countries to find components and to produce and/or assemble goods (hybrid products) and often use brand naming that suggest linguistic origins other than the true origin of the brand (Samiee *et al.,* 2005), understanding how consumers correctly identify the origin of a brand is a very relevant issue. In fact, associating a brand with the wrong COO can affect the perception of its value and the consumer's willingness to buy it (Pegan *et al.,* 2020).

The country of brand, develops from the correspondence between the image of the country of origin and the product (Roth and Romeo, 1992), can be learned explicitly (by storing information) or implicitly (by classifying the brand in the country of origin through attributes). According to Hutchinson and Alba (1991), consumers in most cases don't learn intentionally. The predominance of implicit learning can be explained by the process of categorization, i.e., the tendency of consumers to structure their knowledge of product/brand alternatives into categories (Gutman, 1982). Han (1989) has shown that the COO can serve both as a stereotypical measure (coding function or summary construct) and as a proxy for other product attributes (inference function or halo effect) to simplify choice.

2.2 Stereotype Content Model and the country of brand

One way to explore how implicit and explicit country stereotypes can influence preferences for brands with different origins is using the Stereotype Content Model (SCM). This model is based on a construct able to effectively operationalize the stereotypes of the country not linked to any class of product or idiosyncratic category (Fiske *et al.,* 2002; Brambilla and Leach, 2014). Stereotypes are semantic associations about a particular group, activated, often automatically, with little cognitive activity in the presence of a member or symbolic equivalent of the stereotyped group (Blair, 2002). According to the SCM, each group can be described through its competence and warmth. In the context of the country of origin, the dimension of warmth should capture those aspects related to friendliness, availability, sincerity and trust, while competence would reflect aspects related to intelligence, capacity and effectiveness (Fiske, 2007). At the theoretical level, the more a country is characterised by a high level of warmth and/or competence, the more we expect a positive response to the products of that country (Chen *et al.,* 2014). The presence of information on the country of origin of the brand could

therefore automatically activate the stereotypes related to the country that will later influence the buying behaviour. Based on SCM (Fiske *et al.*, 2002), as well as on previous bi-dimensional models of the content of stereotypes, it can be found that warmth and competence are the basis of the content of most group stereotypes. Recent research has suggested that the dimension of warmth should be considered as divisible into two distinct constructs: sociality and morality (Leach *et al.*, 2007). While the construct of sociality is about being benevolent towards people in such a way as to facilitate affectionate relationships with such people (friendliness, kindness, pleasantness and sympathy), the construct of morality refers to being benevolent with people in such a way as to facilitate correct and principled relationships with them (honesty, reliability, credibility and sincerity). Morality traits may be more relevant to infer the intentions of others than those of sociality and competence (Brambilla, 2013).

2.3 Green brand and Stereotype content model

The response of firms to the global call of consumers to solve the most pressing societal problems (Pegan *et al.*, 2020) has led to a significant increase in the availability of green products/brand in the marketplace (Pinna, 2020). These products/brands are aimed to consumers who want to integrate ethical concerns into their consumption choices. Although consumers claim to be concerned about environmental well-being, the share of market of green products and brands is very low (Pinna, 2020). A critical aspect to assess is whether the association between a social stereotype and responsible consumers represents a psychological obstacle to buying green brands. Consumers are constantly engaged in maintaining or improving the concept of self through symbolic consumption that manifests itself in the purchase of brands that match a specific dimension of themselves (Sirgy, 1982). In short, a consumer with positive attitudes towards green products also has a positive image towards green brands in order to have consistency between personal goals and behaviour (Hogg *et al*., 2000).

The general socially widespread view of responsible consumers as stereotypical is frequent (Burgess *et al.*, 2013). Because brands are also used to influence consumers' perceptions of characteristics, brands can lead to bias. The SCM provides a useful framework for examining the stereotyping of "responsible" consumers. The key point is that personality traits attributed to a brand are also applied to its clients/users (Fennis and Pruyn, 2007). Brand identification research documents that users see brands as expressing their personal values and meanings (Stokburger-Sauer, 2012). It has also been seen that animosity towards a country can influence the perceptions of brands stereotypically associated with the country subject to

prejudice. From this perspective, if brands are perceived in terms of their warmth and competence, it is reasonable to expect that users of the same brands are perceived with the same type of social perception. In the way that buying a green brand is perceived as an ethical and/or altruistic action (Shang and Peloza, 2016) consumers of such alternatives are stereotyped as warm. According to the theory of the SCM, stereotyping is driven by the ability of consumers to recognize the ethical and altruistic nature of brands that convey an interest in society or the environment. Responsible brands tend to underline their "green" or "ethical" credentials and this, probably, influences the social perception of those who buy them. Luchs *et al.* (2010) saw that the perceived ethicality of a product or brand can be positively associated with attributes related to kindness, a trait in contrast to perceptions of effectiveness and competence. Ethical consumption choices are probably subject to less imitation as the stereotype of warmth seems to convey a relatively low social status. This effect is explained by the influence that warmth exerts on feelings of admiration and envy. Admiration is due to the appreciation of something praiseworthy that others have done or achieved (Smith and Decoster, 2000). Admiration is an emotion involved in learning as the individuals being admired become models and providers of information on the form of behaviour to be adopted. In a consumer context, the behaviour of others can easily be imitated: by buying the same brands, a consumer can share with his model the same symbols and meanings that generate admiration. Although research into admiration is limited, the evidence in international marketing shows that admiration for a country can lead to a more positive attitude towards products from that place. The issue is that stereotypes of warmth have a negative influence on envy (Cuddy *et al.*, 2008), as a result, the ethical characteristics of a brand can indirectly reduce the envy felt towards consumers who bought them. From this point of view, the stereotype of warmth can damage green brands because it makes them less likely for improving the social status of consumers. Lastly, the stereotypes of warmth attributed to brand green users could have a positive influence on admiration and a negative influence on envy. Since both these emotions contribute to explain social imitation, there will be two indirect effects with opposite directions (Zhao *et al.*, 2014). The importance given to admiration and envy in driving the consumption imitation determines which pathway will be most relevant. Brambilla and Leach (2014) have suggested that warmth is composed of two distinct dimensions: morality and sociality. Therefore, this article, concerning the dimensions of morality and sociality in relation to brand origin and green brand aims to investigate how the competence and the two components of the warmth, morality and sociality, of the green brand and its country of origin can positively/negatively be related to the perceived quality and purchase intentions of a green brand.

3. An Applied Analysis

In this section, in order to investigate the role of the three dimensions of the "Stereotype Content" model, competence, morality and sociality, applied to the country of origin as a function of the consumer's perception of quality and purchase intention of a green brand, we present the questionnaires and a descriptive analysis.

3.1 The Questionnaires

The quantitative analysis is performed using structured questionnaires based on a seven point Likert scale. Six different questionnaires were developed using Google Drive Modules, with a different combination of countries and products associated with the sustainable brand:

- green fashion clothing brand from the United States, Germany, Italy.
- green car brands from the United States, Germany, Italy.

The choice of these countries lies in the fact that Germany and the United States represent, for Italy, an important voice in both imports and exports[1]. In fact, in the first semester of 2019 Italian imports see Germany and the United States respectively in first and seventh place. In the same semester of 2019 as regards Italian exports the two nations are in first and third place respectively.

Moreover, this decision was also made considering some sustainability factors of the selected countries. In this respect, reference was mainly made to the EPI[2] (Environmental Performance Index). Germany occupies the thirteenth place in the global ranking, followed by Italy in sixteenth place and the United States in twenty-seventh place. These data, combined with those relating to imports and exports, have made it possible to make a choice of equipped countries considering both the economic and environmental aspects, no longer considering countries (Switzerland, France, Denmark) or less virtuous (India, Congo, Bangladesh, Burundi) from an environmental point of view but which do not have important trade with Italy (with the exception of France). In addition, it was also decided not to use countries at

[1]Source MISE website:
(https://www.mise.gov.it/images/stories/commercio_internazionale/osservatorio_commercio_internazionale/statistiche_import_export/paesi_import.pdf)
(https://www.mise.gov.it/images/stories/commercio_internazionale/osservatorio_commercio_internazionale/statistiche_import_export/paesi_export.pdf)

[2] Source Yale Center for Environmental Law and Policy:
(https://epi.envirocenter.yale.edu/2018/report/category/hlt)

opposite ends of the EPI ranking in order not to run the risk of having extreme results mainly due to this dichotomy.

As regards the brands used, we opted for the non-use of well-known brands, this to avoid any association or familiarity with some brand that might have an influence on the opinions of the participants. In all questionnaires the words "hypothetical brand green" were used leaving, therefore, a certain freedom to the imagination of individuals to try to have a certain neutrality of judgment on implicit or explicit preferences for a particular brand and trying to have a greater focus on the country of origin and sustainability. To limit excessive abstraction and give consumers more concrete information about what they had to imagine and evaluate, two green products i.e., cars and fashion clothing were used. Such products could be sufficiently known to the participants or part of the basket of goods of Italian consumers.

Once the development of all 6 types of questionnaire was completed and they have been tested on small groups of Italian customers, 6 different links were created, one for each questionnaire and, subsequently, from July 2019 to October 2019 each link was administered to 100 consumers, who were, as far as possible, predominantly workers of at least 23 years of age. Each questionnaire, in the end, had a section in which participants were required to provide feedback on their propensity for sustainable products and brands.

In the first part of the questionnaire the participants were asked to give their opinion on the degree of agreement with a sentence containing adjectives related to each of the 3 constructs of the “Stereotype Content” model so the participants had to express their degree of agreement with 12 statements containing, each, a different representative attribute of a "Stereotype Content Model" construct: competence, morality and sociality. The second section of the questionnaire consists of four parts. First of all, participants were presented with a description of what is meant by sustainable car or sustainable fashion clothing. Subsequently, the participants were asked to give an opinion on the perceived quality of the sustainable brand of one of the 2 possible products from one of the 3 possible countries, then they were asked to give their judgment on the attributes used in the first section. Finally, they were asked to give an opinion on claims to measure purchase intent. The third section was developed with the aim of measuring, the propensity to sustainability and the intention of purchasing brands and sustainable products by consumers (Environmental concern). In addition, they were asked to provide an opinion on claims aimed at assessing their perceptions about the country to be judged (CSR). The fourth and last section was implemented with the aim of collecting socio-demographic data of the participants, while maintaining anonymity. It has been requested to indicate the gender, age, last qualification obtained, profession, marital status, number of family members, family life cycle, economic situation compared

to peers, an approximate indication of annual gross family income, nationality and, finally, the country of residence. We can summarize the fourth parts of the questionnaire in the Table 1.

Table 1 – *The questionnaire.*

I part	Stereotype Content Model applied to the Country
II part	Perceived quality of the green brand
	Stereotype Content Model applied to the green brand
	Purchase intentions of the green brand
III part	Corporate Social Responsibility (CSR) of the Country
	Consumer environmental concern
	Consumer's intention to purchase sustainable brands
IV part	Socio-demographic data

3.2 The Statistical Analysis

In order to investigate the role of the three dimensions of the "Stereotype Content" model, competence, morality and sociality, applied to the country of origin as a function of the consumer's perception of quality and purchase intention of a green brand, we tried to answer the following questions:

1. What is the relationship between a green brand and the stereotype of its country of origin?
2. What is the relationship between the perceived morality of the brand's country of origin and the perceived morality of the green brand analyzed?
3. What is the relationship between a green brand's perceived morality, its perceived quality, and the brand's purchase intention?

As regards the relationship between a green brand and the stereotype of its country of origin we can considered the medians and the Kruskal-Wallis test.

As one can see from Table 2 Germany is at first place for competence while Italy is at first place for morality and sociality in the case of a green fashion clothing brand. From Table 3 we can observe that Germany is at first place for competence and morality while Italy is at first place for sociality. The differences by type of product are all significant as regards competence and sociability, not for morality.

Table 2 – *Median of the constructs Competence, Morality, Sociality of Italy, Germany, United States for the green fashion clothing brand.*

	Italy	Germany	United States
Morality	3.00	4.00	3.75
Sociality	5.12	3.25	3.75
Competence	3.50	5.75	5.25

Table 3 – *Median of the constructs Competence, Morality, Sociality of Italy, Germany, United States for the green car brand.*

	Italy	Germany	United States
Morality	3.50	4.25	3.25
Sociality	5.00	3.75	4.00
Competence	3.25	5.62	5.25

As regards the relationship between the perceived morality of the brand's country of origin and the perceived morality of the green brand analyzed and the relationship between a green brand's perceived morality, its perceived quality, and the brand's purchase intention we calculated the medians and the Kruskal-Wallis test.

As one can observe from Table 4, Table 5 and Table 6 green fashion clothing Italian brand is better as regards morality, sociality, competence, perceived quality and intention to purchase than green fashion clothing German and Unites States brand. Green car German brand is better as regards morality, competence, perceived quality, intention to purchase than Green car Italian and Unites States brand, while Green car Italian brand is the best for sociality. The differences are all significant except for the sociality of the green car United States brand.

Table 4 – *Median for green fashion clothing Italian brand and Green car Italian brand.*

	green fashion clothing Italian brand	green car Italian brand
Morality	5.00	4.25
Sociality	5.00	4.50
Competence	5.50	4.75
Perceived Quality	5.40	4.60
Intention to Purchase	4.87	4.00

Table 5– *Median for green fashion clothing German brand and green car German brand.*

	green fashion clothing German brand	green car German brand
Morality	4.25	4.75
Sociality	4.00	4.00
Competence	5.00	5.50
Perceived Quality	5.20	5.60
Intention to Purchase	4.00	4.25

Table 6 – *Median for green fashion clothing Italian brand and green car Italian brand.*

	green fashion clothing United States brand	green car United States brand
Morality	4.25	4.25
Sociality	4.50	4.00
Competence	4.75	4.75
Perceived Quality	4.60	4.60
Intention to Purchase	4.25	3.12

Specifically, with regard to the relationships between the morality of a green brand from the three countries and perceived quality and purchase intention, respectively, correlations were calculated (Table 7 and Table 8). In general, there are high correlations between the morality of a green brand and perceived quality, while this is not the case between the morality of a green brand and purchase intention (see for example the Italian green clothing brand and the German green car brand).

Table 7 – *Correlations between morality and Perceived Quality and Morality and Intention to Purchase for green fashion clothing brand.*

	Perceived Quality	Intention to Purchase
Morality brand green Italy	0.587	0.400
Morality brand green Germany	0.650	0.589
Morality brand green United States	0.625	0.554

Table 8 – *Correlations between morality and Perceived Quality and Morality and Intention to Purchase for green car brand.*

	Perceived Quality	Intention to Purchase
Morality brand green Italy	0.691	0.562
Morality brand green Germany	0.670	0.338
Morality brand green United States	0.759	0.565

4 Conclusions

In this paper we consider a revised Stereotypic Content Model in which the competence dimension is combined with the warmth dimension divided into two stereotypic sub-dimensions i.e., morality and sociality is considered.

We analyzed if and how the Country of Brand (COB) stereotypes can be related to the Country of Origin (COO). For the three constructs of the «Stereotype Content» model, matching between the country and the brand could have a positive impact on the perception of quality and consumer purchasing intentions. In particular, the results suggest that the dimension of competence related to the green brand and its country of origin seems to affect the consumer's perception of quality and purchasing

intentions; as well as morality and the combination of morality and competence, particularly for Germany and the United States. The same is not true in the case of Italy where the dimension of sociality becomes more relevant. Moreover, perceiving the moral and/or competent COO could help the green brand to be perceived positively. Finally, especially for Germany and the USA, there seems to be a correlation between green brand and COO in the perception of stereotypes of competence and morality. This research, although still in the exploratory phase, can provide useful suggestions to managers, committed to promoting green consumer styles, in terms of strategic positioning: a green brand and its origin perceived as competent and/or moral could have positive effects on the perception of quality and the consumer's willingness to buy a specific category of product. The lack of such stereotypes could be a useful indicator of possible repositioning strategies. Nowadays green brands can also be perceived as competent and not just warm, and this would be a strategic aspect on which managers should focus.

References

BLAIR I.V. 2002. The Malleability of Automatic Stereotypes and Prejudice, *Personality and Social Psychology Review*, Vol. 6, No. 3, pp. 242-261.

BRAMBILLA M., LEACH C. 2014. On the Importance of Being Moral: The Distinctive Role of Morality in Social Judgment, *Social Cognition*, Vol. 32, No. 4, pp. 397-408.

BRAMBILLA M. S. 2013. Morality and Intergroup Relations: Threats to Safety and Group Image Predict the Desire to Interact with Outgroup and Ingroup Members, *Journal of Experimental Social Psychology*, Vol. 49, No. 5, pp. 811-821.

BURGESS M., KING, HARRIS M., LEWIS E. 2013. Electric Vehicle Drivers' Reported Interactions with the Public: Driving Stereotype Change? *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 17, pp. 33-44.

CHEN C. Y., MATHUR P., MAHESWARAN D. 2014. The Effects of Country-Related Affect on Product Evaluations, *Journal of Consumer Research*, Vol. 41, No. 4, pp. 1033-1046.

CUDDY A. J.C., FISKE S.T., GLICK P. 2008. Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. In M. P. Zanna (Ed.) *Advances in Experimental Social Psychology,* Vol. 40, Elsevier, pp. 61-149.

FISKE S.T., CUDDY A., GLICK P. 2007. Universal Dimensions of Social Cognition: Warmth and Competence, *Trends in Cognitive Sciences*, Vol. 11, No. 2, pp. 77-83.

FISKE S. T., CUDDY A. J. C., GLICK P., XU J. 2002. A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow from Perceived Status and Competition, *Journal of Personality and Social Psychology*, Vol. 82, No. 6, pp. 878-902.

FENNIS B.M., PRUYN A.T.H. 2007. You Are What You Wear: Brand Personality Influences on Consumer Impression Formation, *Journal of Business Research*, Vol. 60, No. 6, pp. 634-639.

GUTMAN J. 1982. A Means-End Chain Model Based on Consumer Categorization Processes, *Journal of Marketing*, Vol. 46, No. 2, pp. 60-72.

HAN C.M. 1989. Country Image: Halo or Summary Construct? *Journal of Marketing Research*, Vol. 26, No. 2, pp. 222-229.

HOGG M.K., COX A.J., KEELING K. 2000. The Impact of Self-Monitoring on Image Congruence and Product/Brand Evaluation, *European Journal of Marketing*, Vol. 34, pp. 641-667.

HUTCHINSON J., ALBA J. 1991. Ignoring Irrelevant Information: Situational Determinants of Consumers Learning, *Journal of Consumer Research*, Vol. 18, No. 3, pp. 325-346.

LEACH C.W., ELLEMERS N., BARRETO M. 2007. Group Virtue: The Importance of Morality (vs. Competence and Sociability) in the Positive Evaluation of In-Groups, *Journal of Personality and Social Psychology*, Vol. 93, No. 2, pp. 234-249.

LUCHS M. G, NAYLOR R. W, IRWIN J. R, RAGHUNATHAN R. 2010. The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference, *Journal of Marketing*, Vol. 74, No. 5, pp. 18-31.

PEGAN G., VIANELLI D., DE LUCA P. 2020. *International marketing strategy: the country-of-origin effect on decision-making in practice*. Heldelberg: Springer.

PEGAN G., SCHOIER G., DE LUCA P. 2020. The importance of consumer perception of corporate social responsibility to meet the need for sustainable consumption: challenges in the sportswear sector. In C. Silvestri, M. Piccarozzi and B. Aquilani (Eds.) *Customer Satisfaction and Sustainability Initiatives in the Fourth Industrial Revolution*, Hershey, PA: IGI Global, pp. 212-235.

PINNA M., 2020. Do Gender Identities of Femininity and Masculinity Affect the Intention to Buy Ethical Products? *Psychology & Marketing*, Vol. 37, pp. 384-39.

ROTH M.S., ROMEO J.B. 1992. Matching Product Category and Country Image Perceptions: A Framework for Managing Country-of-Origin Effects, *Journal of International Business Studies*, Vol. 23, No. 3, pp. 477-497.

SAMIEE S., SHIMP T., SHARMA S. 2005. Brand Origin Recognition Accuracy: Its Antecedents and Consumers' Cognitive Limitations, *Journal of International Business Studies*, Vol. 36, No. 4, pp. 379-397.

SHANG J., PELOZA J. 2016. Can "Real" Men Consume Ethically? How Ethical Consumption Leads to Unintended Observer Inference, *Journal of Business Ethics*, Vol. 139, pp. 129-145.

SIRGY M.J. 1982. Self-Concept in Consumer Behavior: A Critical Review, *Journal of Consumer Research*, Vol. 9, No. 3, pp. 287-300.

SMITH E. R., DECOSTER J. 2000. Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems, *Personality and Social Psychology Review*, Vol. 4, No. 2, pp. 108-131.

STOKBURGER-SAUER N. R. 2012. Drivers of Consumer-Brand Identification, *International Journal of Research in Marketing*, Vol. 29, No. 4, pp. 406-418.

ZHAO H., GAO Q., WU Y., WANG Y., ZHU X. 2014. What Affects Green Consumer Behavior in China? A Case Study from Qingdao, *Journal of Cleaner Production*, Vol. 63, pp. 143-151.

SUMMARY

An analysis on consumer perception via "stereotype content" model between sustainable brand and country of origin

This paper aims to analyze the Country of Origin (COO) effect and the consumer's intention to purchase green brands, applying the revised Stereotypic Content Model, in which the competence dimension is combined with the warmth dimension divided into two stereotypic sub-dimensions of morality and sociality. While previous analyses have investigated the role of stereotypes of competence and warmth with regard to Country of Brand (COB) and green brands, this research aims to investigate how the competence and the two components of the warmth (morality and sociality) of the green brand and its country of origin can positively/negatively influence the perceived quality and purchase intentions of a green brand. Considering the importance of understanding if and how COB stereotypes can be related to those of the green brands and how this can influence the perception of quality and the purchase intention of the brand, an exploratory quantitative research has been conducted on a sample of Italian consumers. In particular, the empirical research has been developed considering three COO: Italy, Germany and USA and two brands: green cars and green fashion clothing. The data have been analyzed through different statistical analysis.

Gabriella SCHOIER, University of Trieste, Italy, gabriella.schoier@deams.units.it
Giovanna PEGAN, University of Trieste, Italy, giovanna.pegan@deams.units.it
Stefano FADINI, University of Trieste, Italy, stefanofadini@hotmail.com

EFFECTIVE DURATION OF THE CIVIL PROCEEDINGS IN ITALY – STATISTICAL ANALYSIS

Maria Filomeno, Irene Rocchetti

1. Introduction

Proceeding duration is one of the indicators most used in order to measure the efficiency of a judicial office, which can depend on different factors such as the nature of the cause and its complexity.

All data presented in this contribute are our analysis on Justice Ministry data coming from the data warehouse of the Justice Ministry and relying to the period 1/7/2018 - 30/06/2019.

Measured duration is the effective proceedings duration computed as difference between the exact date when a procedure is concluded and the exact date when it started; this measure differs from the *disposition time* (pending cases/resolved cases) and the *average stock formula duration* ((total initial +final pending proceedings)/2 out of (new enrolled proceedings + resolved ones)/2) which do not consider in the computation the exact date when proceedings started. Knowing that each civil proceeding is characterized by an object code and each object code belongs to a macroarea, one of the goals of this paper is of analysing the existence of a relationship between mean duration, office dimension and their geographical distribution through regression models per type of office (1° - 2° Courts). The same analysis has then been replied by considering only proceedings with the most frequent object codes in terms of total definitions per each macroarea for Courts of first and second instance. A cluster analysis has then been applied to create, where possible, an aggregation of object codes within homogeneous groups per duration classes. Finally, effective duration and disposition time have been compared with respect to the macroarea.

2. Civil proceeding duration, dimension of offices and geographical division

Our elaborations on Justice Ministry data highlight that the mean duration of proceedings resolved in the period July 2018-june 2019 is about 490 days for Courts

of first instance (greater for insolvency procedures ad executions) and 899 days for Courts of second instance (Table 1).

Table 1 – *Exploratory analysis.*

Office	Resolved cases	Mean duration Sicid (days)	Mean duration Siecic (days)	Mean Total duration (days)
Court 1°	2,160,954	448	634	490
Court 2°	141,727	899	-	899

We applied regression models (Ricci, 2006) showing that there is not a statistically significant relation between the mean duration of civil proceedings and the dimension of offices (small, medium, big and metropolitans); however the estimates of the relation between duration and geographical division are statistically significant at 95% for the Courts of first instance and at the 90% for the Courts of second instance. As far as the Courts of first instance concerns, the mean duration in northern Italy is averagely lower than the one in central and southern Italy: in the South, duration is 124 days higher compared to the one of the Centre, in the North is 165 days lower than the Centre. The same direction of estimates can be found in the Courts of second instance even if with different intensity: in the North durations are averagely 271 days lower while in the South durations are 2 months higher than in the Centre of Italy (Table 2).

Table 2 – *Statistical parameters Regression duration- geographical division.*

Office	Parameters	Estimate	Standard Error (SE)	p value
	Intercept	499.48	29.16	<0.0001
Courts I °	North	-165.66	36.06	<0.0001
	South	124.31	34.93	0.0005
	Intercept	885.75	128.07	<0.0001
Courts II°	North	-270.86	153.92	0.09
	South	58.19	143.18	0.69

3. Mean duration variability

If we look at the range (maximum-minimum) of the proceeding duration per macroarea[1], we can see very high values: for example, as far as proceedings of the labour macroarea in Courts of first instance are concerned, the mean duration is of 701 days, the minimum duration is of 83 days while the maximum duration is equal to 2284 days; in Courts of second instance the mean duration is of 788 days and goes from 215 to 1076 days.

An important heterogeneity is visible also in the Civil litigations macroarea where the mean duration goes from a minimum of 1 day to a maximum of 4414 days.

Given to the heterogeneity of mean duration between proceedings per macroarea and object codes inside of them, we decided to focus our attention on object codes related to the most frequent proceedings in terms of resolved cases. This choice is due to the fact that less frequent litigations could behave in a different way even between different offices in terms of proceeding duration, by creating a *bias* in the total distribution.

If we look at the territorial distributions of the resolved proceedings duration for each macroarea and for degree of judgemental offices, both as far as the most frequent resolved proceedings and the total resolved cases are concerned, we can see that they are not equal for all the macroarea.

For example, let us look at the distributions of the Civil litigations in Courts of first instance: the South of Italy is characterized by higher values of the proceeding duration if compared to the North where cases are resolved in about one year and half. Mean duration of proceedings obtained by considering all resolved proceedings is generally lower than the one computed by considering only the 42% of the resolved proceedings; this could be due to the fact that less frequent proceedings with a lower duration lows the total mean down. However this difference is not so high and actually can attest an acceptable level of similarity.

Thus, a different behaviour between the most frequent and total proceedings is true for the Courts of second instance but not for the Courts of first instance: in the latter case, for the majority of the macroarea, the mean duration of sampled proceedings approximates with an acceptable level of precision the mean duration of all resolved proceedings; exceptions are related to the Insolvency procedures and the Voluntary Jurisdiction (VJ) not in matter of family and persons macroarea.

We considered all the resolved cases for the following analysis.

---

[1] Macroarea is a group of homogeneous object codes per "materia". In detail, Civ.lit.=Civil litigations, Inj. ord.= injunctive orders, Sec-Exe.=Securities executions, R.Este.Exe.=Real estate executions, Ins.Proc.=Insolvency Procedures, Soc.Sec=Social security, Sp.Proc.=Special proceedings, Sep. Div.=Litigation separations and divorces, VJ in=Voluntary jurisdiction (family & people), VJ not In=Voluntary jurisdiction (not in matter of family & people).

4. A different aggregation of object codes: the cluster analysis

Given to the heterogeneity of the mean duration values of resolved proceedings related to different object codes belonging to each macroarea, we aimed at aggregating differently object codes by identifying *clusters* as homogeneous groups inside and heterogeneous groups between them in terms of mean duration of proceedings; this could have an important role in proceedings complexity classification given the assumption that higher durations are linked to more complex proceedings.

A cluster analysis (Duran and Odell, 2013) has then been applied to all object codes corresponding to resolved cases greater than 10 (to exclude unused object codes), by not diversifying according to the type of office, basing on two classification variables computed for each object codes: *mean proceeding duration and its between offices variability measured though the variation coefficient (VC).* Results have then been analysed separately for Courts of first and second instance.

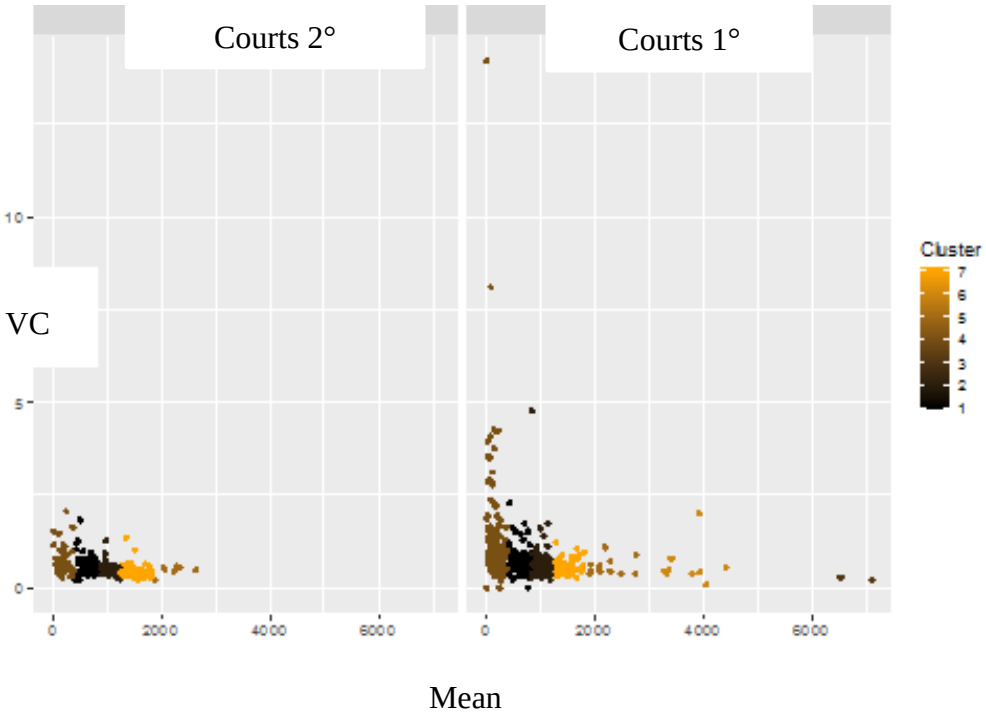
7 clusters have been identified thanking to the *Silhouette* of *Gap-statistics* (Rousseeuw, 1987) technique whose results are represented in the Figure below (Figure 1) where the x-axis shows the mean duration for each object code and the y-axis shows the variation coefficient computed on the mean duration; each point colour identifies the belonging group of every object code.

For example as far as the Courts of 1° instance is concerned, the first group of object codes coming from the Cluster analysis is characterized by the greatest variation coefficient values and the lowest mean and it is represented by the brown colour. Object codes belonging to that first group are codes belonging to different macroareas, susch as Separations and divorces, Labour, Special Proceedings, etc.

As far as the Courts of second instance are concerned, the variability of the considered measures is lower than the one of the Court of first instance offices.

The variable more affecting the clusterization procedure is the mean duration of proceedings, being the variation coefficients values more stable. Higher variability in terms of VCs is registered in the first cluster of Courts of first instance and in the first and second clusters of Courts of second instance (Figure 1).

Figure 1 – *Cluster representation for Courts of 1° and 2° instance.*



4.1 Cluster description

Table 3 summarises cluster characteristics in terms of macroarea object codes belonging to, mean duration, number of object codes and mean variation coefficient, as far as Courts of first instance are concerned. We looked also at the object code with minimum and maximum duration per each macroarea within each cluster so to better understand the cluster representation (we do not report details in this contribution for reasons of brevity).

The first cluster is the one with lower mean proceeding duration (153 days) while the 7th has greater duration (6818 days) and contains codes belonging to Insolvency procedures and Real estate executions macroarea.

Greater VCs and thus greater variability of mean duration between offices is registered for some VJ not in matter of family and persons codes for the 1th and the 2th cluster, for some Real estate executions codes for the third cluster, etc.

Table 3 – *Table Mean of the mean duration and CV per macro area per each cluster –Courts of first instance.*

	Cluster & Macroarea	Mean duration of proceedings	Number of object codes	Mean of the variation coefficient (VC)
Cluster 1	**Cluster 1**	**153**	**135**	**1.31**
	Labour	273	5	1.02
	Insolvency procedures	186	4	0.65
	Securities executions	200	3	1.09
	VJ in matter of family and persons	171	18	0.78
	VJ not in matter of family and persons	113	52	1.77
	Litigation separations and divorce	243	2	0.39
	Civil litigations	206	16	1.00
	Special proceedings	170	30	1.32
	Injunctive orders	31	5	0.77
Cluster 2	**Cluster 2**	**627**	**107**	**0.74**
	Labour	613	45	0.71
	Social securities	640	9	0.45
	Insolvency procedures	807	1	0.64
	VJ in matter of family and persons	611	1	0.76
	VJ not in matter of family and persons	544	6	1.69
	Litigation separations and divorce	653	3	0.30
	Civil litigations	654	37	0.71
	Special proceedings	593	5	0.75
Cluster 3	**Cluster 3**	**1025**	**94**	**0.65**
	Labour	989	13	0.68
	Social securities	899	2	0.60
	Real estate executions	1063	1	0.80
	Civil litigations	1033	78	0.64
Cluster 4	**Cluster 4**	**1459**	**43**	**0.58**
	Labour	1490	1	0.48
	Real estate executions	1663	1	0.29
	Civil litigations	1453	41	0.59
Cluster 5	**Cluster 5**	**2260**	**12**	**0.55**
	Labour	2284	1	0.43
	Insolvency procedures	2416	2	0.49
	Civil litigations	2223	9	0.58
Cluster 6	**Cluster 6**	**3720**	**9**	**0.61**
	Securities executions	3942	1	0.40
	Civil litigations	3692	8	0.63
Cluster 7	**Cluster 7**	**6818**	**2**	**0.23**
	Insolvency procedures	6533	1	0.26
	Real estate executions	7103	1	0.20

For the Courts of second instance proceedings are shorter than Courts of first instance, in fact there are no codes in 6th and 7th cluster, also because of absence of competence in Insolvency procedures and executions.

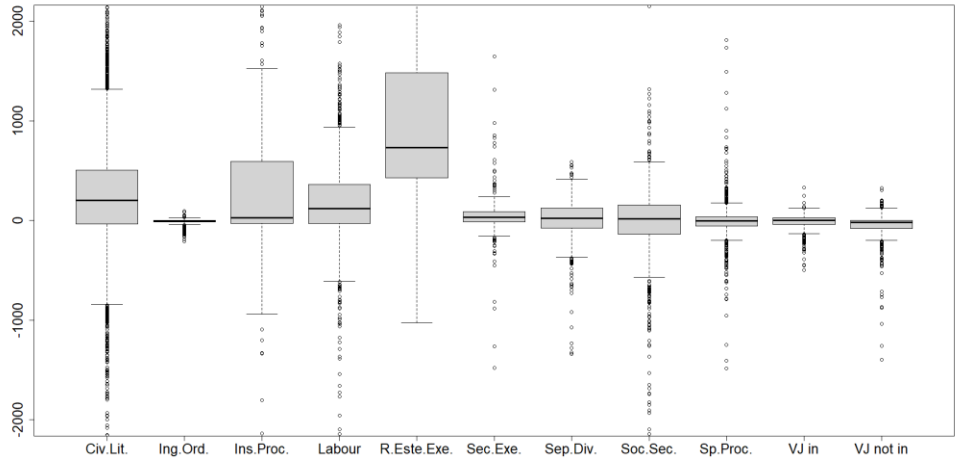
5. Effective duration and disposition time

Even if we decided to focus on the effective duration rather than on the before mentioned other kind of durations, given an assumed higher probability of reliability, we wanted to analyse the differences between the effective duration and the disposition time of proceedings, the latter used at the international level from the Cepej (European Commission for the Efficiency of Justice) as a duration estimate - which does not count for the starting and ending date of proceedings - computed as pending cases out of resolved ones times 365.25 to have duration expressed in days (Calvez *et al*., 2018).

We considered object codes corresponding to pending proceedings and resolved cases greater than 10 without considering "protections and curatorships" whose duration depends on the nature of proceedings and not on the resolving capacity of the judge.

As far as Courts of first instance is concerned, correlation between effective mean duration and disposition time is equal to 0.61 and the Anova or analysis of variance (Kaufmann and Schering, 2014) applied both over the mean effective duration and the disposition time shows that the duration means are statistically significant between the macroarea in both cases even if the model on the effective duration fits better the data. If we look at the difference between the two duration measures per macroarea, we can notice a higher variability in correspondence of the macroarea of civil litigations, executions and Insolvency procedures and labour and social security; median durations of civil litigations and real estate executions are farer away from each other (see Figure 1).

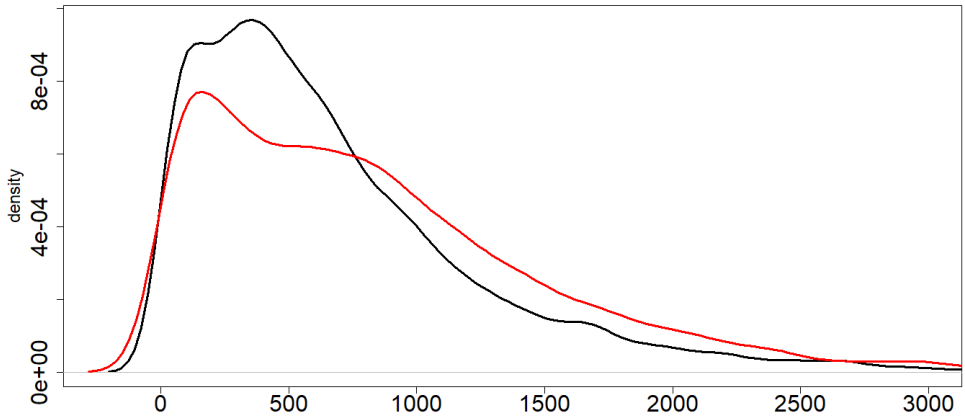
Figure 1 – *Boxplot of the differences between durations – Courts 1°.*



The percentage of difference of duration is more positive in all the macroarea but the injunctive orders, special proceedings and voluntary jurisdiction not in matter of family and person where the disposition time is greater than the effective duration.

The figure below compares the distributions related to the effective duration and the disposition time for Courts of first instance; by looking at the two densities, we can notice that in correspondence of lower durations the effective duration is lower while in correspondence of higher durations the density of the disposition time is lower.

Figure 2– *Effective duration density (red line) vs disposition time (black line).*

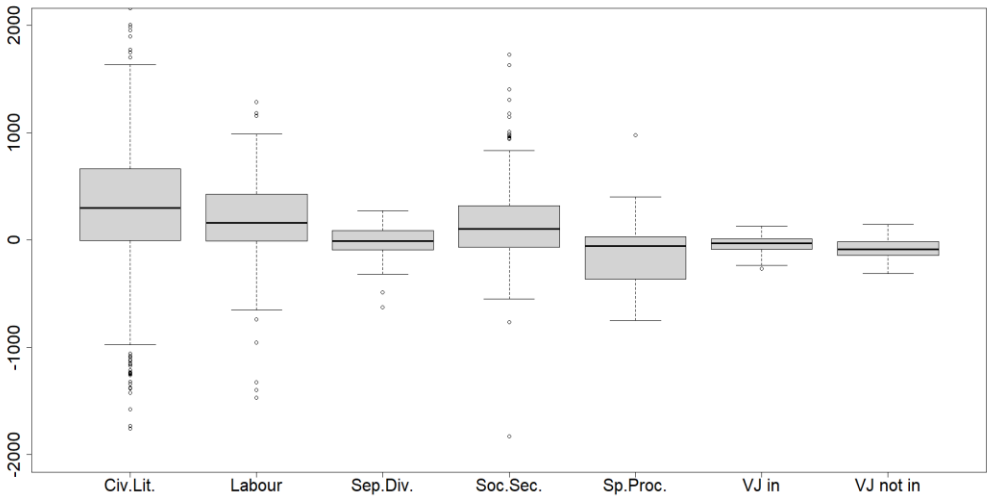


If we look to each macroarea, the two densities can behave differently, for example for the Real estate executions the two curves overlap each other in a different way and for the injunctive orders, the effective duration is greater than the disposition time within 100 days of duration; graphs are divided differently according to the different scale or better to the period in terms of days the two curves overlap each other in or are close to each other (i.e. for injunctive orders 100 days, for civil litigations 3000 days, etc.).

Effective durations and disposition times are similar for the Insolvency procedures, Social Security, Civil litigations Procedures; as far as the securities executions are concerned, the disposition time is higher than the effective time within 200 days of duration, for Voluntary jurisdiction the effective duration is higher than the disposition time within 200-400 days.

For Courts of second instance, the ANOVA (analysis of variance), (Kaufmann and Schering, 2014) applied both on the effective duration and on the disposition time per macroarea, shows that the means of durations are statistically different in both cases but the model on the effective duration fits better the data. Correlation between the two measures is equal to 0.56; the variability is greater than the civil litigations, labour and social security and special proceedings one, however for all the considered macroarea differences between duration in the Courts are in terms of medians closer to zero than the Courts of first instance (Figure 3).

Figure 3 – *Boxplot of the differences between durations-Courts 2°.*



By looking at the densities, considerations are similar to the ones for Courts of first instance even if the shapes are different especially as far as the effective duration is concerned being it the result of a mixture of two densities (we are not showing the related graph for reasons of brevity).

6. Conclusions

The aim of this contribute is of analysing the duration of civil proceedings and its possible connection with offices organization in terms of working staff dimension and their capacity of resolving civil affairs but also with the different complexity of proceedings. Duration is in fact one of the most important efficiency indicators together with ultra triennial/biennal proceedings, number of resolved case, etc., useful to evaluate the performance of judicial offices.

From the analysis on the effective duration of civil proceedings, computed as the difference between the resolving cases date and the enrolling case date, no significant relation between duration and dimension of offices emerges, while a significant difference of duration in the various geographical areas of Italy exists especially for Courts of first instance: northern and central courts are characterized by lower durations if compared to the southern ones. This may be due to a different organization in terms of resolving proceedings and in the different complexity of

proceedings but not in terms of working staff given that there not seems to be a significant relation between duration and dimension of offices according to the regression models applied.

The assumption to conduct a specific analysis by considering only most frequent object codes per each macroarea has been verified: for the majority of macroarea, the distribution of the duration of proceedings with most frequent object codes approximates well the duration of all proceedings in the macroarea. We conducted thus the study over all the object codes.

Because of the variability of object codes duration per macroarea a statistical analysis to aggregate object codes in a different way with respect to the macroarea one has been applied. Through a cluster analysis run without distinguish by type of office, 7 homogeneous groups of civil affairs have been identified; these groups are characterized by similar proceeding durations inside them and very different from each other and give a qualitative vision of the minor or major complexity of proceedings.

This clustering based classification could be considered for the eventual identification of the proceeding weights per object code.

As far as the comparison between effective duration and disposition time (the latter used by the Cepej and thus considered as valid duration estimate) is concerned, both for Courts of first and second instance, results show that the effective mean duration is greater than the disposition time in correspondence of briefs durations, viceversa the disposition time is higher than effective duration for proceedings resolved in more than 3 years. This could be because for brief durations the disposition time is not able by construction to seize the speed of the resolving process. In fact, the disposition time is a duration estimate given by unresolved out of resolved cases and thus it can be very small for brief durations because of the few unresolved proceedings or the many resolved proceedings or both of them.

Furthermore, there are greater or lower differences between the two durations according to the considered macroarea: mean effective duration is lower than the disposition time for example for special proceedings, divorces and separations and voluntary jurisdiction.

Acknowledgements

We thank Stefano Cecini and Paolo Fantini to have read this contribute and given us their opinion about it, both in terms of statistical analysis and interpretations.

References

CALVEZ, F., REGIS, N. 2018. *Regis Length of court proceedings in the member states of the Council of Europe based on the case law of the European Court of Human Rights.* European commission for the efficiency of justice (Cepej), No. 27.

DURAN, B. S., ODELL, P. L. 2013. *Cluster Analysis: A Survey*. Berlin: Springer Science & Business Media.

KAUFMANN, J., SCHERING, A.G. 2014. *Analysis of Variance ANOVA*. Wiley Online Library.

RICCI, V. 2006. *Principali tecniche di regressione con R* (https://cran.r-project.org/doc/contrib/Ricci-regression-it.pdf).

ROUSSEEUW, P. J. 1987. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Computational and Applied Mathematics,* Vol. 20, pp. 53-65.

SUMMARY

Effective duration of the civil proceedings in italy - statistical analysis

Statistical analysis has been conducted on the effective duration of civil proceedings: no significant relation between duration and dimension of offices emerges, while it exists with respect to geographical divisions especially for Courts of first instance.

A cluster analysis has been applied to identify homogeneous groups of object codes and thus give a qualitative vision of the minor or major complexity of proceedings.

We finally compared effective duration and disposition time both for Courts of first and second instance, results show that the effective mean duration is greater than the disposition time in correspondence of briefs durations and viceversa the disposition time is higher than effective duration for proceedings resolved in more than 3 years.

Maria FILOMENO, Consiglio Superiore della Magistratura, m.filomeno@cosmag.it
Irene ROCCHETTI, Consiglio Superiore della Magistratura, i.rocchetti@cosmag.it

IN THE NAME OF WEAK LEGALITY? DANGEROUS RELATIONS BETWEEN CITIZENSHIP INCOME, NON-OBSERVED ECONOMY AND TAX CHEATING[1]

Umberto Di Maggio, Giuseppe Notarstefano, Giuseppe Terzo

1. Premise

The "Reddito di Cittadinanza" (RdC[2]) is an anti-poverty and active inclusion measure launched by the Italian government in 2019 at the instigation of the "Movimento 5 Stelle" party. Inspired by basic income, it is a form of support for the poor[3] through a periodic monetary payment and active labour policies to contrast marginality.

It is, therefore, a measure that addresses poverty from a one-dimensional perspective closely interrelated to lack of employment (Baldini and Gori, 2019; Van Parijs, 2018). The subsidy paid for average individual amounts up to €780 and increases concerning the household size. Considering that the price level in the South is 20% lower than in the North, the RdC discourages the acceptance of jobs, especially precarious and low paid ones (Baldini and Gori, 2019). Two years after the measure introduced, the beneficiaries who signed the employment pact and found a job are small. Of the approximately 1,400,000 recipients of the Rdc, only 350,000 have found a job[4].

There is no doubt that the RdC has enabled many households to have more disposable income. This is clear from the positive change in disposable income of consumer households (See Table 2). The Mezzogiorno is the only area that, compared to the 2018-2019 period, has increased, albeit a small one, while the remaining areas show a significant decrease over the same period.

However, the increase in disposable income has not had, at least in the first year of operation of the RdC, direct effects of great magnitude on overall consumption (See Table 3). It should be remembered that the RdC was launched in 2019 and that the data published by Istat, referring to the two years 2018-2019, refer to a time range

[1] For research articles with several authors, a short paragraph specifying their individual contributions must be provided here.

[2] https://www.redditodicittadinanza.gov.it/

[3] https://www.lavoro.gov.it/temi-e-priorita/poverta-ed-esclusione-sociale/focus-on/Reddito-di-Inclusione-ReI/Pagine/default.aspx

[4] https://www.ilsole24ore.com/art/l-anno-e-mezzo-navigator-media-colloquio-testa-giorno-ADuHI5IB

that is too narrow to make a long-range analysis on the effectiveness of the measure against poverty.

The incidence of this measure is very high in the Southern territories, where there are both the highest rates of unemployment and poverty and where there is an atavistic structural vulnerability of the productive system. With precise reference to poverty, as Istat points out, in 2020, also as a result of the pandemic, a little more than two million households (7.7% of the total from 6.4% in 2019) and more than 5.6 million individuals (9.4% from 7.7%) are in a condition of absolute poverty. After the improvement in 2019, absolute poverty thus increases to its highest level since 2005. Specifically, in 2020, the incidence of households in absolute poverty is confirmed to be higher in the South (9.4%, from 8.6%), but the largest growth is recorded in the North, where household poverty rose to 7.6% from 5.8% in 2019. This dynamic means that, while in 2019 poor households in our country were distributed almost equally in the North (43.4%) and in the South (42.2%), in 2020, they reach 47% in the North against 38.6% in the South, with a difference in the absolute value of 167 thousand households. Regarding relative poverty, households below the threshold are just over 2.6 million (10.1%, from 11.4% in 2019) (Istat, 2021).

Many of the abuses detected by judicial authorities are also concentrated in the South[5]. A lot of the recipients whose benefits were withdrawn were employed in the underground and informal economy and sometimes even in the illegal economy.

Therefore, the RdC is undoubtedly a measure to be improved, both in terms of a drastic strengthening of the social and labour activation components and the fight against irregularity. It is necessary that the RdC fuels a welfarist vision of welfare that is not consistent with the logic of social investment that, instead, is implemented in other European countries with some success (Marzulli and Pesenti, 2021).

Moreover, it is crucial to correlate the RdC with the Non-Observed Economy (NOE), which in the South, according to the Istat Economic and Territorial Accounts for 2018 (see Table 1), represents 18.8% of the added value with higher rates than in the Centre (13.8%), the North-East (10.9%) and the North-West (10.3%). In the South and the Islands, specifically, the share of irregular work stands out (7.5%) with significantly higher rates than in the Centre (5.0%), the North-East (3.9%) and the North-West (3.8%). Notably, Calabria (21.3%), Campania (19.8%), Sicily (19.3%) and Apulia (19.1%) are the regions where the overall incidence of the components of the unobserved economy on the added value is the highest[6].

[5] https://www.libertasicilia.it/palermo-illeciti-contro-il-reddito-di-cittadinanza-condannati-per-mafia-percepivano-il-sussidio/

[6]https://www.finanze.gov.it/export/sites/finanze/.galleries/Documenti/Varie/Relazione_evasione_fiscale_e_contributiva_-Allegato-_NADEF_2020.pdf

Table 1 – *Incidence of the unobserved economy on total value added - % values -Year 2018. Source Istat.*

Regions	Irregular employment	TOTAL unobserved economy (including other components)
Calabria	9,8	21,3
Campania	8,5	19,8
Sicily	7,8	19,3
Apulia	7,1	19,1
Southern Italy	7,5	18,8
Centre	5	13,8
North-east	3,9	10,9
North-west	3,8	10,3
Italy	**4,9**	**13,1**

These initial data confirm that the collection of information on the NOE is a highly complex matter. Within the NOE, there are, in fact, informal and illegal encapsulations whose precise measurement is difficult. At the most, it is possible to make estimates with information coming from tax audits that in any case, have a high degree of fragmentariness and incompleteness (Adair, 2017; Feige, 2015).

2. Objectives, assumptions, methods

The overall narrative of the abuse phenomenon has insisted exclusively on the "illegal worker" in the Sicily, Calabria, Campania, Puglia regions stigmatizing its fraudulent practices. At least one illegal worker corresponds to at least one "illegal entrepreneur" who benefits from unlawful work by lowering social security and insurance costs and reducing tax revenues for the State.

So, the RdC should therefore be analyzed with the Non-Observed Economy (NOE) that, especially in South Italy represents 18.8% of added value with higher rates than in the Center (13.8%), the North-East (10.9%) and the North-West (10.3%) (Istat, 2018).

In order to verify the link between the fraudulent perception of the RdC and the diffusion of irregular and illegal practices in the entrepreneurial and labour sphere, we hypothesise that there is a sort of propensity to fiscal infidelity.

This propensity is greater where socio-economic contexts are characterised by the presence of weak forms of legality. Specifically, “a weak form of legality occurs when an expectation or forecast of ineffectiveness, distorted, particularistic, delayed and unreliable application, avoidance, hasty reviewability is generated concerning

the legal norms in force (generally for the public policies in place) so that the prescriptions contained in them and the objectives announced in them cannot form the premises of calculations aimed at the adoption of desires by economic actors. Such norms and policies are often born with little credibility"[7] (La Spina, 2005: 162).

Our assumptions are two and are closely linked to the definition of weak legality. The first that RdC is a government measure that, paradoxically, feeds and coexists with a submerged irregular economy. This is determined by the welfarist and populist character of the measure; by the generosity of the monetary contribution transferred, which—especially in some territorial contexts such as those of the South—makes it desirable concerning obtaining a job; by the unstable labour market which allows most of the time precarious and poorly distributed jobs; by the inability of the placement system to know how to cross between labour demand and supply. The false declarations, the irregularities in the applications to obtain the subsidy and the maintenance of an odd job while receiving the subsidy at the same time confirm the widespread collective perception of the RdC, especially in some population groups, as one of the many weak rules that can be abused without incurring consistent penalties and sanctions that would act as a deterrent.

The second assumption is that propensity to demand illegal work—and therefore to give illegal job frauding the State—depends on the opportunistic behaviour of the whole economic operators (worker and entrepreneur) that use a specific value-rationality supplied by weak legality. This is linked to the institutional inability to track abuses extensively and punctually and to the provisionality and incompleteness of the measure, which makes it so obviously fragile and therefore stimulates opportunistic behaviour on the recipients and those who offer them irregularly/illegal work.

To confirm these hypotheses, we use data from:

1. 1.The Ministry of Economy and Finance referring specifically to the unobserved economy and tax and social security evasion (2020).
2. The INPS observatory on the RdC (2020).

From these two information bases at the regional level (NUTS-2), we carry out an exploratory analysis in order to correlate the number of revocations of the RdC respectively with the number of undeclared workers and the contributory irregularities detected.

[7] Our translation.

3. Results

It is interesting, in our opinion, to interpolate the data referred to the RdC to those referred to the widespread tax evasion and specifically to the tax gap, i.e. the difference between expected and actual tax revenues on all components.

Having available only the data provided by the Ministry of Economy and Finance (MEF) referring to 2018 and 2019[8], and therefore to a period prior to the entry into force of the RdC, we will proceed to interpret the propensity to tax infidelity appropriately referring to business activities developed on the whole national territory.

The data mentioned above on labour-related vigilance activities show that the South is where the most critical issues emerged (87%) followed by the Centre (83.17%), the North-East (76.33%) and the North-West (71.41%).

The figure refers to the financial volume of future INPS revenues resulting from the verification of employment relationships that would have given rise to undue social security benefits. In this case, the North-West is in first place (11.2%), followed by the Centre (9.87%), the North-East (7.88%) and the South (5.59%). However, the South of Italy is at the bottom of the ranking when comparing the lower expenses detected with the overall irregularities detected by the inspection activity.

The South is no longer in the lead when analysing the number of undeclared workers detected by the total number of workers in inspected companies. In this case, the North-East is in the first place (3.72%), ahead of the South itself (2.47%), the Centre (2.24%) and the North-West (1.46%).

The same applies when referring to the gap between VAT paid and unpaid for the total number of taxpayers. The North-West is the area with the highest value (1.32%). It is followed by the Centre (1.08%), the South (1.05%) and the North-East (1.02%).

[8] Relazione sull'economia non osservata e sull'evasione fiscale e contributiva (2018-2019-2020).

Table 2 - *Territorial indicators of entrepreneurial propensity to tax and contribution infidelity. Our elaborations on MEF 2019 data.*

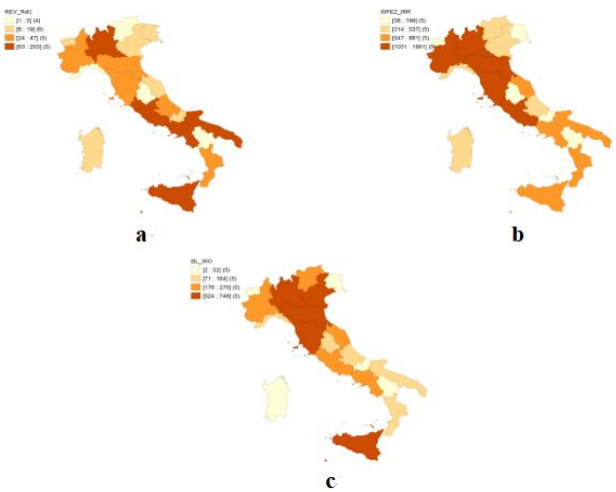
Regioni	% Irregularities on finalised GdF inspection activities	Less expenditure € * / N° Irregularities	% Illegal workers / Number of workers	% VAT Gap**/ Contributors
North-West	71,41	11,2	1,46	1,32
North-East	76,33	7,88	3,72	1,02
Centre	83,17	9,87	2,24	1,08
South	87,05	5,59	2,47	1,05

** Lower outgoings = higher future revenues for INPS resulting from the detection of fictitious employment relationships that would have given rise to undue social security benefits and therefore, for example, to undue social security contributions relief for companies and undue benefits received*
*** VAT Gap = Unpaid Value Added Tax*

Table 3 – *Correlations RdC Revocations & Illegal workers/Irregularities by inspections. Our elaboration on MEF and INPS RdC Observatory 2019.*

	RdC Revocations & Illegal workers	RdC Revocation & Irregularities by inspections
South	0,51	0,6
Center	0,47	0,83
North	0,74	0,87
Italy	0,39	0,37

Figure 1 - *Spatial distribution of RdC revocations (a), irregularities by inspections (b) and illegal workers (c). Our elaborations on MEF data (2019).*



In order to further corroborate the previous evidence, we also carry out a Principal Component Analysis on the following indicators (related to 2019):

1. The number of complaints per case.
2. The number of inspections completed per 10,000 inhabitants.
3. The number of irregularities (in percentage values).
4. The average duration of trials.

The first-factor plan explains 86% of the variability of the data. Furthermore, the first component is strongly correlated with the number of protests, the average duration of trials and the number of irregulars, while it is negatively correlated with the number of inspections completed. Thus, there is a strong positive correlation between protests, the average duration of trials and irregularities. These three indicators are, in turn, negatively correlated with the number of inspections completed.

Table 4 – *Output of Principal Component Analysis (PCA).*

Component	Autovalue	Proportion	Cumulated
1	2,6398	0,6600	0,6600
2	0,8035	0,2009	0,8608
3	0,3227	0,0807	0,9415
4	0,2339	0,0585	1,0000

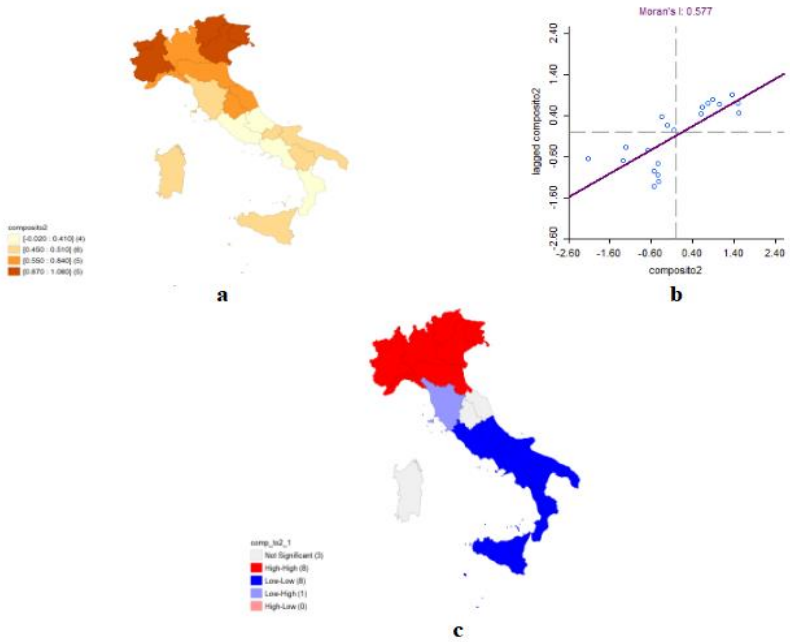
	PC1	PC2	PC3	PC4
Protests	0,505	0,515	-0,149	0,676
Inspections	-0,457	0,645	-0,549	-0,271
Irregulars	0,517	-0,365	-0,727	-0,268
Trials	0,519	0,430	0,385	-0,630

We can decipher these results by referring once again to the theoretical reference of weak legality (La Spina, 2005), according to which there are some contexts where the environmental conditions, with particular reference to specific rules and the respective activities of control and supervision of compliance with them, are fragile and favour the spread of behaviour that does not comply with the rules themselves. The inability of the institutions to control abuses and the congenital provisionality and incompleteness of the implementing rules would favour the spreading of illegal behaviours. As the analysis of the main components shows, it is possible to highlight how - also in the case of the abuses on the RdC or more generally on the lack of transparency in the working environment - the diffusion of opportunistic behaviours typical of free riders.

Using the first component we structure a composite index of weak legality, which assumes a positive polarity. In Figure 2 we can observe its spatial distribution, which shows a significant divide between Centre-North and South. This spatial distribution suggests the presence of spatial autocorrelation, which we verified performing both

Global and Local Moran Index using an inverse distance matrix[9], the results of which are also visible in Figure 2. The global spatial analysis shows, as widely expected, a strong positive autocorrelation (I=0.67), which indicates that the phenomenon of weak legality has spillover effects. while the local analysis shows the presence of two large clusters: a cluster of "high-high" values in the North and an equally large cluster of low-low values in the Centre-South.

Figure 2 - *Spatial distribution of weak illegality index (a), scatterplot of Global Moran Index (b) and cluster map of Local Moran Index (c). Our elaborations on multiple sources.*



[9] For robustness purposes, we also used alternative spatial weighting matrices, such as contiguity and KNN with different critical cut-off, without observing significant changes in the results.

4. Conclusions

We highlighted with the support of data that the RdC is a controversial measure against poverty that currently has several critical application issues. We also underlined how RdC is related to opportunistic behaviours of workers and entrepreneurs who prefer to defraud the State rather than implement virtuous behaviours in support of the common good (Bucciol and Montinari, 2019; Scott and Grasmick, 1981). In the final lines of this paper, we want to conclude with a reflection on the concept of poverty itself and the related concepts of work and consumption. We, therefore, want to move for a moment to the theoretical level of reflection but always remaining strongly linked to the empirical evidence of our study. First of all, we ask ourselves whether the transfer of money in the form of a subsidy to family spending to encourage consumption by subjects without spending capacity, as currently provided by the RdC, can be considered a decent solution to the condition of poverty. In the absence of mechanisms able to fully favour active inclusion, the RdC, per se, can become a practice that discourages employment, is inappropriate in a general sense and, even worse, feeds opportunistic, selfish and claiming behaviour towards the State. To dispel any possible misunderstanding, we would like to stress that we are certainly convinced of the need to fight poverty through a minimum income. However, we are also convinced that poverty must come through active policies that fully involve citizens and give work the central role it deserves. At the moment, however, it would seem that the centre of gravity has shifted to encouraging consumption, as if the monetary transfer, in itself, freed people from poverty. Aid in the form of money transfers does not emancipate and, on the contrary, can trap, if the recipient is in a position to choose between different purchase options. We are among those who highlight the need to combine learning and knowledge with conscious consumption in order to reduce, more generally, all the conditions that make possible social inequalities between those who possess resources (even in excess) and those who are not at all. We are also convinced of the need to free the field from guilt-based readings of poverty. The season of pauperism that looked at poverty as a fault and a stigma is undoubtedly over.

The condition of indigence, today more than ever in this pandemic scenario, is much more than the absence of money in one's pocket. It is correct to speak of poverty from the point of view of the absence of job opportunities and meaningful social relations. It is correct to consider inequality in terms of access to information, care, education and, more generally, the absence of opportunities to realise one's life plans and thus achieve well-being, not just the ability to buy as many goods as possible compulsively. Today it is certainly no longer the time to propose, as was done in the 19th century with the Poor Law, new workhouses where the poor and excluded are imprisoned and forced to work in order to obtain subsidies. Poverty

cannot be contrasted with a punitive, paternalistic and perhaps even hasty attitude. At the same time, however, the results of our research highlight the need for differently calibrated policies that move in the direction of an economic and social paradigm shift. The widespread presence of free-riders and, therefore, of dangerous relations between tax evasion and the receipt of citizenship income also by those who are not entitled to it and who are engaged in the submerged and illegal economy, impose the need to reflect, more globally, on the need to establish a new solidarity pact between citizens and the State.
Moreover, there is a need for more effective policies (which can do without the 'weak legality' mentioned in this article) and therefore less focused on extemporary solutions to structural problems such as poverty.

The Smithian invisible hand, the self-interest and the liberal selfish ethics that would guarantee, according to the dominant economic mainstream, the equilibrium of the market seem to fail in the face of the evidence of the constant presence of opportunistic behaviour among workers and entrepreneurs even in the presence of policies to support the most disadvantaged sectors of the population. What emerges, then, is the need for a more civil and less speculative approach and for virtuous relations between citizenship, the State and the market that are capable of restoring dignity to those in need and not for charitable instruments, such as the RdC implemented in this way, which instead of emancipating from the condition of need, as the data confirm, traps citizens in need by relegating them to the condition of mere consumers.

References

ADAIR P. 2017. Non-observed economy vs. the shadow economy in the EU: the accuracy of measurements methods and estimates revisited. In *4th OBEGEF Interdisciplinary Insights on Fraud and Corruption*, HAL: Post-Print.

BALDINI M., GORI C. 2019. Il Reddito di Cittadinanza, *il Mulino*, Vol. 68, No. 2, pp. 269-277.

BUCCIOL A., MONTINARI N. 2019. *Dishonesty in behavioral economics*. London: Academic Press.

FEIGE, E.L. 2016. Reflections on the Meaning and Measurement of Unobserved Economies: What Do We Really Know About the "Shadow Economy", *Journal of Tax Administration*, Vol. 2.

FREY B.S., SCHNEIDER F., KEPLER J. 2000. Informal and Underground Economy. *Working Papers 0004*, Johannes Kepler University of Linz, Department of Economics.

ISTAT. 2021. *Le statistiche Istat sulla povertà*. *Statistiche Report*, Roma, Istat. https://www.istat.it/it/files//2021/06/REPORT_POVERTA_2020.pdf

LA SPINA A. 2005. Mafia, legalità debole e sviluppo del Mezzogiorno. Bologna: Il Mulino.

MARZULLI M., PESENTI L. 2021. Il Reddito di Cittadinanza: una riforma mancata? In Boschetti B. and Vendramini E., (Eds.) FuturAP - Rapporto sul Futuro e l'innovazione dell'Amministrazione Pubblica, Milano: Educatt Università Cattolica del Sacro Cuore, pp. 183-196.

SCOTT W.J., GRASMICK H.G. 1981. Deterrence and Income Tax Cheating: Testing Interaction Hypotheses in Utilitarian Theories, *The Journal of Applied Behavioral Science*, Vol.17, No. 3, pp. 395-408.

VITTORIA A. 2020. La Scomparsa dei Poveri. Una Prima Valutazione di Policy sul Reddito di Cittadinanza, *Social Policies*, Vol. 7, No. 3, pp. 525-544.

SUMMARY

In the name of weak legality? Dangerous relations between citizenship income, non-observed economy and tax cheating

The incidence of requests for financial subsidies linked to the Citizenship Income (RdC) -especially in the Southern territories- is high. Here there is also the highest rate of unemployment and poverty and a structural vulnerability of the productive system. However, after two years of applying the anti-poverty measure, which also includes job placement, few people have found work. In addition, many fraudulent uses of the RdC benefit have been detected. Many recipients of RdC worked in the informal and underground economy (Frey *et al*., 2000), and sometimes even in the illegal economy. Using data from the Ministry of Economy and Finance referring to 2018 and 2019, and relating them to the Istat national economic accounts and those of INPS, we demonstrate the presence of tax fraud behaviour appropriately also referring to entrepreneurial activities developed throughout the national territory and not only in the South and explicitly referring also to the employment of undeclared workers. The guiding hypothesis is that the presence of a "weak legality" (specifically, rules that are unable to structurally solve the problem and the absence of capillary and adequate controls against fraud) encourages workers and entrepreneurs' opportunistic behaviour deterring it.

Umberto DI MAGGIO, LUMSA University of Rome (Palermo Campus), u.dimaggio@lumsa.it;
Giuseppe NOTARSTEFANO, LUMSA University of Rome (Palermo Campus), g.notarstefano@lumsa.it;
Giuseppe TERZO, LUMSA University of Rome (Palermo Campus), g.terzo2@lumsa.it

EXPLORING EARNINGS DIFFERENCES HIDDEN BEHIND THE GENDER PAY GAP[1]

Francesca Ceccato, Marilena Angela Ciarallo, Paola Conigliaro

1. Foreword

Gender equality is one of the fundamental values of the European Union (Article 3(2) of the EC Treaty). Anyway, it is not jet a reality in the labour market. Regarding wages and salary in particular, the gap between women and men continues to exist and there is little evidence of improvement in some countries and in some areas. Reducing possible discrimination in earnings between men and women remains one of the major objectives within the overarching goal of achieving gender equality according to the Sustainable Development Goals of the UN 2030 Agenda, and of the European Pillar of Social Rights.

A large literature (economic, sociological and psychological theories) is available aiming to explain the persistence and the determinants of the salary differences between men and women. Several theories claim that differential gender roles unconsciously adopted early in life influence much of what happens in personal relationships, family life and employment (Ochsenfeld, 2014). Following different paths in education and employment is expected lead to overall differences in pay. Some studies consider the growth in women's education a major driver of the decline in the gender pay gap (Goldin, 2002), while Olsen *et al*. (2010) show that education explains little of the gap, and its impact is also declining.

In this framework, the necessity to investigate the multiple content of that measure of gender inequalities is a growing need. One of the key indicators used to monitor imbalances in earnings between women and men is the structural indicator Gender Pay Gap (GPG), but the GPG is not exactly a measure of the incidence of discrimination in the labour market. Furthermore, it is defined 'unadjusted' since it is sensitive to composition effects (sector of economic activity, size of company, occupation, level of education, age of the employee, seniority in the company, etc.). For this reason, the main institutions (e.g., Eurostat, International Labour Office)

[1] The article is the result of the joint work of the authors. §1, 2, 6 are attributed to Marilena Ciarallo, §3 and 4 are attributed to Paola Conigliaro and §5 to Francesca Ceccato.

conducted experiments in order to develop a methodology to calculate an indicator more suitable for measuring pay differentials by controlling the composition effects.

This report presents the main characteristics of gender pay differentials in Italy in 2018, measured using the unadjusted GPG and an application of the decomposition approach (i.e., Oaxaca decomposition), used in Eurostat experiments.

2. The GPG and its main data source: the structure of earnings survey

The unadjusted GPG used in EU area to monitor possible discrimination in earnings between men and women, expressed as a percentage, is defined as follows:

$$\frac{\textit{Mean (gross) hourly earnings of men - Mean (gross) hourly earnings of women}}{\textit{Mean (gross) hourly earnings of men}} \quad (1)$$

According to Eurostat Methodology, all European Countries use the Structure of Earnings Survey (SES)[2] micro data, at employee level, to calculate the indicator. The SES allows to produce comparable information on relationships between the level of remuneration, individual characteristics of employees (sex, age, occupation, length of service, highest educational level attained, etc.) and of their employer (economic activity and geographic location of the local unit; size of the enterprise).

As the SES results are available only every four years (the last is 2018 edition), the EU countries update the indicator every year, using annual information from other sources. In the case of Italy, up to 2016 data from EU-SILC survey and since 2017 data from the Italian Labour Register based on administrative data were used. According to GPG methodology, the annual GPG estimates have to be aligned with the (*4*-yearly) SES. For this reason, estimates related to years backwards the year *t+4* have been updated distributing equally over the period the differences between the initial GPG for the year *t+4* calculated using national annual data and the SES data benchmark for the year *t+4*.

$$GPG_t \textit{ updated} = GPG_{t\,annual} + 1/4 * (GPG_{t+4\,SES} - GPG_{t+4annual}) \quad (2)$$

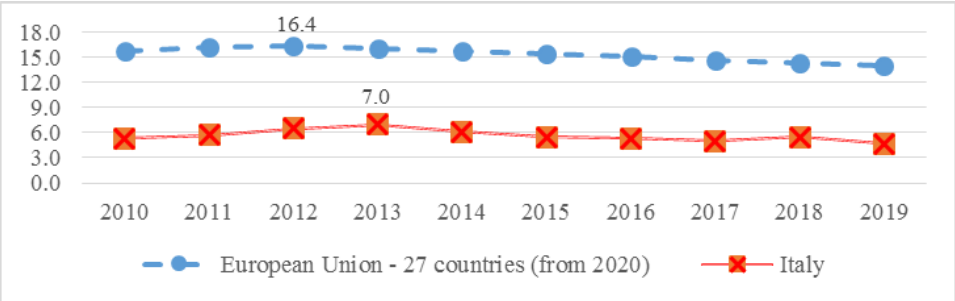
Figure 1 displays the GPG values in Italy and at EU 27 level since 2010. The Italian GPG is one of the lowest in EU: in 2018, it is 5.5%[3] and provisional data for

[2] The Structure of earnings survey is conducted every four years in the Member States and it is based on the Council Regulation (EC) No 530/1999 concerning structural statistics on earnings and on labour costs. All the enterprises and institutions belonging to Private and Public sectors with at least 10 employees compose the target population. The employees covered are those who received remuneration for the reference month (October 2018).

[3] In EU comparison, values do not consider Nace Section O (Public administration, Defence, Compulsory social security) as it is not mandatory. In the following analyses, Section O is included. In that case, the Italian GPG is 6.2% in 2018.

2019 show a further decrease. The indicator has always been many points below the EU average value (15.3% in 2018) and, as for almost all EU countries, it smoothly decreases every year. The next paragraphs deal with a closer analysis of the complexity of the indicator and of its possible interpretations.

Figure 1 – *Gender pay gap in unadjusted form EU 27 – Italy (2010-2019).*



Data source: Eurostat- Dataset Gender pay gap in unadjusted form. At 24/02/2021 update, data for 2016, 2017 and 2019 are provisional

3. What the GPG indicates and what it does not. The need of adjusting or integrating the indicator.

As defined, the GPG does not consider the earnings of employed in micro-enterprises, self-employed, professionals and agricultural workers and all workers in undeclared and informal employment relationships. Furthermore, even within the observation domain, the indicator presents some information gaps, recognized and described in much of the literature.

The ILO (2018) aiming to better understand what lies behind the GPG figure, examines a set of 80 countries very different in term of labour legislation, level of collective bargain, and methods to collect data. Most of the countries register a positive GPG (20 points on average). Considering the median, many countries manifest a wider gap. This should indicate that men record higher frequencies in the high tail of earnings distribution, while women in the low tail. As concerning the distribution of average yearly earnings, the gap widens in most of countries. The report identifies many factors of inequality hidden in the GPG measure and proposes two complementary measures to better approximate the assessment of unequal pay for equal work. The first is the factor-weighted GPG, the second is the decomposition of GPG into an “explained” and an “unexplained” part. To calculate the factor-weighted GPG the authors have: selected a set of factors that they believe affect the earning structure (e.g. education, age, working hours, private or public

sector employment); clustered women and men into subgroups defined by these main factors; estimated the GPG for each subgroup; multiplied the GPG of each subgroup by the weight of the subgroup with respect to the employee population. As concerning the decomposition procedure carried out in ILO's report, it is very near to that applied by Eurostat to calculate the "*adjusted*" GPG, described in detail in Section 5.

4. Descriptive analysis on composition effects on Italian GPG [4]

There is a broad consensus among scholars that low GPG can result even under conditions that do not support equity in earnings. For example, in Italy the higher hourly remuneration of teachers in public schools lowers the national GPG, but this information hides the fact that monthly wages of teachers are quite low, because of the low number of contractual hours. Teaching in schools is a low-paying profession, although the workers' educational level is higher than in other sectors. It is then no so attractive. The sector is highly feminized (76.8%), and despite this, the number of male employees in management positions is 1.4 times the number of females in that role. It follows that even in this sector the GPG is positive (7.9%), and assume a negative value (-8.4%) just for employees with a secondary qualification.

One of the characteristics that influences the GPG value is the form of economic and financial control of the enterprises. Data show that this factor has effect on GPG in many EU countries (Eurostat 2021). In Italy, in 2018, the GPG in the private control sector was 17.7%, while it dropped to 2% in the public sector. The GPG varies also according to the NACE section (sector of economic activity). Scholars attribute this to the "horizontal segregation", as women "choose" (i.e. have more job opportunities in) those sectors where workers receive lower wages. Another factor that affects earnings is the educational level of employees. Employees with lower educational attainment have a greater probability to work for a lower remuneration (12.1 euros on average), but women earn less than man at any educational level. However, GPG is lower among employees with a secondary educational level (10.4%) and higher for employees with basic (15%) and tertiary (18%) educational attainment. A further factor influencing both wages and GPG is age. Young employees earn 11.1 euros per hour on average, those between 30 and 49 years old 15 euros and the older ones 18.2 euros. The GPG raises with the increasing age, but the average wages of men and women do not present the same growth curves in

[4] This session presents some descriptive analyses on Italian data from the SES 2018. Figures and percentages are the results of calculations applied directly to the SES 2018 micro data. For further data and methodological information see ISTAT (2021).

relation to the age class. Hourly earnings for women increase particularly in the transition from 20-24 to 25-29 years of length of service in the enterprise (+14.2%), while the greatest improvement observed for males occurs between the first two classes (0-4 and 5-9 years) of seniority (+22.4%). The number of men in management or supervisory positions significantly exceeds the number of women, both in absolute terms (2.4 times) and in relation to the gender composition of employees (1.9 times). Furthermore, the GPG is 11.9% between employees covering those positions.[5] These differences in career paths and opportunities determine the so called "vertical occupational segregation".

Table 1 presents the results of a focus on the 45-54 age group of employees and the main NACE sections[6], splitting data according to educational attainment, sex, and section.

Table 1 – *Employees in age class 45-54, by section, educational attainment and sex; supervisory positions, average earnings and gender differences in earnings.*

Section		Educational attainment	Employees (% on total 45-54) *(a)*		Supervisory positions (% on the correspon-ding group) *(b)*		Average hourly earnings	GPG	Mean annual earnings ratio *(c)*
			F	M	F	M			
C	Manufacturing	Basic	2.59	7.67	2.29	6.26	12.89	17.74	1.38
		Secondary	2.75	7.57	8.87	16.49	16.10	16.72	1.30
		Tertiary	0.70	1.72	35.83	48.84	26.89	25.93	1.47
G	Wholesale and retail trade; Repair of motor vehicles and motorcycle	Basic	1.34	2.35	3.01	7.19	12.13	11.93	1.41
		Secondary	2.39	2.46	7.68	18.48	14.90	24.95	1.60
		Tertiary	0.46	0.58	25.55	41.78	24.71	35.20	1.63
N	Administrative and support service activities	Basic	1.90	1.78	1.97	1.82	9.49	10.41	1.71
		Secondary	1.31	1.30	4.77	6.85	11.56	15.16	1.44
		Tertiary	0.31	0.28	11.70	20.55	17.06	20.55	1.30
O	Public administration and Defence; Compul-sory social security	Basic	0.42	2.15	0.00	0.62	16.44	26.73	1.62
		Secondary	2.33	4.50	0.10	2.27	16.97	19.84	1.42
		Tertiary	1.87	1.93	6.95	17.35	21.95	19.85	1.34
P	Education	Basic	0.66	0.25	0.36	0.78	12.33	-0.41	1.05
		Secondary	3.68	0.69	0.26	0.92	17.46	-10.46	0.98
		Tertiary	5.14	1.56	4.78	20.09	24.38	6.01	1.17
Q	Human health and social work activities	Basic	1.57	0.53	2.08	2.64	11.73	8.22	1.18
		Secondary	3.70	1.00	2.57	4.01	14.65	4.90	1.12
		Tertiary	2.71	1.20	19.89	35.36	20.35	18.88	1.30

Notes: Columns (a) percentage of employees of the cell on total employees of 45-54 age class;
Columns (b) percentage of employees in supervisory position on total employees of the correspondent cell;
Column (c) ratio between the average annual earnings for men and women

[5] Employees indicated to hold a management/supervisory positions in the enterprise. It does not necessarily correspond to employees with a managerial occupation.

[6] Chosen sections represent 75.4% of employees in 45-54 age group.

Figures show that for the same educational level and sector of activity, women generally earn a lower hourly wage than men. In addition, they are less likely to hold positions of responsibility, with evident disparity effects on both hourly and annual wages. Earnings gap is in favour of women only in the education sector. However, even there, the situation reverses if we consider employees with a tertiary educational level.

Furthermore, in this subgroup, the presence of management/supervisory personnel among men is considerably higher than among women. Although women account for 3/4 of the employees, male managers are even more numerous than female ones (1.3 times). These are some examples of the information enhancement resulting from a multidimensional data description. They confirm the need for tools capable of highlighting the share of inequality "hidden" beyond the GPG.

5. Hourly wages for personal and job characteristics and Gender Gaps

As before declared, the unadjusted GPG is rather a complex indicator. Eurostat proposed an experimental method[7] to adjust the GPG indicator based on the Oaxaca decomposition[8] and to estimate the unexplained (adjusted) GPG with the Structure of Earnings Survey (SES) as the data source, aimed to better compare the indicators among the European labour markets (Leythienne *et al.*, 2018). This paragraph presents the Oaxaca decomposition that consists of two stages - a regression analysis and a decomposition analysis - based on the SES data on 2018.

5.1. The variables that affect the mean hourly wages

The ordinary linear regression analysis is carried out separately for the earnings structure of men and the earnings structure of women.

In formula, the implementation:

$$\overline{\ln y}^M = \hat{\beta}_0^M + \sum_{k=1}^K \bar{x}_k^M \hat{\beta}_k^M \quad (3)$$

$$\overline{\ln y}^W = \hat{\beta}_0^W + \sum_{k=1}^K \bar{x}_k^W \hat{\beta}_k^W \quad (4)$$

[7] The method and results were discussed with Working Group on Labour Market Statistics (LAMAS) in October 2017 that includes delegates from national statistical institutes in EU Member States and EFTA countries.

[8] Also Called Blinder-Oaxaca decomposition (Oaxaca, 1973).

where: $\overline{\ln y}$ represents the natural log of mean of hourly earnings of men (M) and women (W); $\bar{x}_k$, from k=1 to k=K, are the means of the variables covering the observed personal, job and enterprise characteristics for men (M) and women (W); $\hat{\beta}_0$ is a costant and $\hat{\beta}_k$, from k=1 to k=K, are the parameters for the corresponding variables covering the observed characteristics for men (M) and women (W). The regression equations provide insights into the male and female earnings structures by showing the relationship between log hourly earnings and observed characteristics (personal, job and enterprise characteristics) for men and women, separately[9]. For the categorical characteristics, all categories without one (omitted category) are expressed as dummy variables. The results of the analysis should be interpreted as differentials with respect to the missing dummy. In the case analysed, the set of missing dummies that constitutes the basic profile is made up of man with primary education, clerical support worker, hired with permanent and full-time job-arrangement, working in an enterprise under private control, belonging to the class 10-49 employees and operating in manufacturing activities. Table 2 shows the results in terms of regression coefficients and percentage differences in earnings due to a unit variation of each variable, calculated as $(exp(\beta) - 1) * 100$. The performance of work experience, represented by age and job experience in the current enterprise, on remuneration is positive for men and women but while it is similar for job experience (1% for men and 0.9% for women for each year), it is very different for age (1.1% for men and 0.2% for women). As expected, looking at personal and job characteristics, higher education level than primary school leads to higher hourly wages. Among occupations, the return on earnings compared to clerical support workers is almost always worse for women (except Professionals), although it is very positive for Managers, Professionals, Armed forces and Technicians and Associate Professionals. Percentage differences in earnings are of opposite sign for Services and sales workers (-4.1% for women and +3.5% for men). The part-time working arrangements has a relevant impact on wages, reducing it by 11% among men and by 6.1% among women. The dummies associated to employment contract not permanent (fixed-term and seasonal, apprenticeship) are significantly different from zero and exhibit expected effects but in this case they are worse for men. Looking at the enterprise characteristics, the percentage effect on earnings, compared to manufacturing activities, is discriminatory for women in the enterprises of each economic sectors with few exceptions (Wholesale and retail trade-repair of motor vehicles and motorcycles, Accommodation and food service activities, and Information and communication). In particular, higher wages by 26.9% are paid to men in Financial and insurance activities, while earnings' reduction of 20.5% is observed among female employees working in Public Administration, Defence,

[9] It is a result of adjusting and expanding the standard Mincer (1974) earnings equation.

Compulsory social security. The women disadvantage in Public Administration, Defence, Compulsory social security is balanced by the determinant of Public enterprise control that leads to wages greater by 18.4% among women, comparing to 3.8 among men. Higher salaries are paid to employees who work in enterprises belonging to size classes greater than 49 employees. Even in this case, however, female workers earn less money than male colleagues.

Table 2 – *Estimate of the variables affecting the hourly earnings in Industry and Services for gender – 2018 (% effect of a unit variation of the independent variable)*[10].

INDEPENDENT VARIABLE	Men	Women	Total
PERSONAL AND JOB CHARACTERISTICS (%)			
Female			-9.9
Age (year)	1.1	0.2	0.9
Age squared	0.0	0.0	0.0
Job experience in the current enterprise (year)	1.0	0.9	1.0
Job experience squared	0.0	0.0	0.0
EDUCATION *(b)*			
Secondary education (G2)	6.4	6.2	6.5
Bachelor's degree (G3)	10.6	11.1	11.5
Beyond Bachelor's degree (G4)	29.5	25.3	27.4
OCCUPATION (ISCO-08) *(b)*			
Armed Forces (0)	34.9	24.2	33.5
Managers (1)	145.9	101.9	132.7
Professionals (2)	48.9	49.2	49.3
Technicians and Associate Professionals (3)	20.8	13.0	17.7
Services and Sales Worker (5)	3.5	-4.1	0.3 *(a)*
Skilled Agricultural, Forestry and Fishery Worker (6)	-8.4	0.2 *(a)*	-6.3
Craft and Related Trades Worker (7)	-1.6	-10.7	-4.5
Plant and Machine Operators and Assemblers (8)	0.5 *(a)*	-5.9	-1.5
Elementary Occupations (9)	-10.1	-14.1	-11.9
Part-time (c)	-11.0	-6.1	-8.3
EMPLOYMENT CONTRACT *(b)*			
Fixed-term and seasonal (B)	-4.0	-3.8	-3.8
Apprentice (C)	-16.5	-15.3	-16.2
ENTERPRISE CHARACTERISTICS (%)			
ECONOMIC ACTIVITY (NACE REV2 SECTION) *(b)*			
Mining and quarrying (B)	14.6	9.0	12.1
Electricity, gas, steam and air conditioning supply (D)	9.3	-1.1 *(a)*	5.2
Water supply; sewerage, waste management and remediation activities (E)	-4.9	-7.0	-6.3
Construction (F)	1.2	-1.3 *(a)*	1.5
Wholesale and retail trade; repair of motor vehicles and motorcycles (G)	-6.6	-2.6	-5.0
Transportation and storage (H)	-5.7	-14.2	-8.2
Accommodation and food service activities (I)	-10.3	-8.6	-9.9
Information and communication (J)	-8.1	-6.1	-7.4
Financial and insurance activities (K)	26.9	26.4	26.5
Real estate activities (L)	-8.8	-11.1	-9.8
Professional, scientific and technical activities (M)	-2.0	-5.0	-3.3

[10] The regression R-squared is 0.536 for men, 0.626 for women and 0.569 for total. The coefficients of the regression are available on request.

***Table 2** - continued*

Administrative and support service activities (N)	-13.7	-11.5	-12.9
Public administration, Defence, Compulsory social security (O)	-9.0	-20.5	-14.2
Education (P)	-3.4	-11.6	-6.4
Human health and social work activities (Q)	-9.6	-13.7	-12.0
Arts, entertainment and recreation (R)	3.7	-15.7	-5.2
Other service activities (S)	-15.2	-15.1	-15.3
Public enterprise control *(b)*	3.8	18.4	10.1
ENTERPRISE SIZE *(b)*			
1000-high employee	13.4	9.4	11.7
250-499 employee	9.5	6.2	8.2
500-999 employee	11.7	7.0	9.9
50-249 employee	6.4	4.6	5.9
ABSOLUTE OBSERVATION NUMBER	117,493	131,290	248,783

Notes: (a) Calculated on coefficient not statistically significant. (b) For the categorical characteristics, all categories without one (omitted category) are expressed as dummy variables.

5.2. GPG and its decomposition

After fitting separate regression models for men and women, a decomposition analysis of the difference between the means of log hourly earnings of men and women is carried out: $\Delta = \overline{\ln y}^{M} - \overline{\ln y}^{W}$. Within the decomposition approach, it must be decided which earnings structure constitutes the nondiscriminatory benchmark against which to decompose the difference (Bazen, 2011). For this analysis, it has been chosen the male one in accordance with the definition of unadjusted GPG[11]. As the estimated constant and coefficients in the men's equation are treated as the non-discriminatory benchmarks for the financial returns to characteristics of employees, a counterfactual equation is constructed where the constant and coefficients in the women's equation are replaced by those of the men's equation ($\overline{\ln y}^{W*}$). This equation can be interpreted as what the average female worker would have earned if she had been paid on the same basis as an equivalent male worker. The difference between the means of log hourly earnings of men and women can then be decomposed into two shares as follows:

$$\Delta = \left(\overline{\ln y}^{M} - \overline{\ln y}^{W*}\right) + \left(\overline{\ln y}^{W*} - \overline{\ln y}^{W}\right) \quad (5)$$

The first parenthesis represents the composition effects (explained GPG) caused by different average characteristics by gender, and the second (unexplained GPG) measures the difference between what a female worker with average characteristics

[11] The following other options are possible: - the female earnings structure constituting a non-discriminatory benchmark (Oaxaca 1973); - both the female and male earnings structures constituting non-discriminatory benchmarks with some weighted average applied (Cotton, 1988); - the whole population earnings structure constituting a non-discriminatory benchmark (Neumark, 1988).

would have earned if she had been treated in the same way as a typical male worker and what she actually earns (Bazen, 2011).

After a few steps, the final decomposition equation for the difference between the means of log hourly earnings of men (M) and women (W) can be written as follows:

$$\Delta = \left(\hat{\beta}_0^M - \hat{\beta}_0^W\right) + \sum_{K=1}^{k} \bar{x}_k^W \left(\hat{\beta}_k^M - \hat{\beta}_k^W\right) + \sum_{K=1}^{k} \hat{\beta}_k^M \left(\bar{x}_k^M - \bar{x}_k^W\right) \quad (6)$$

The first two component on the right side of the equation represent the Unexplained part of the difference (U), and the latter represents the Explained part (E). An interpretation of the part U as discrimination is not recommended as some other explanatory factors that were not observed in the data could change the results.

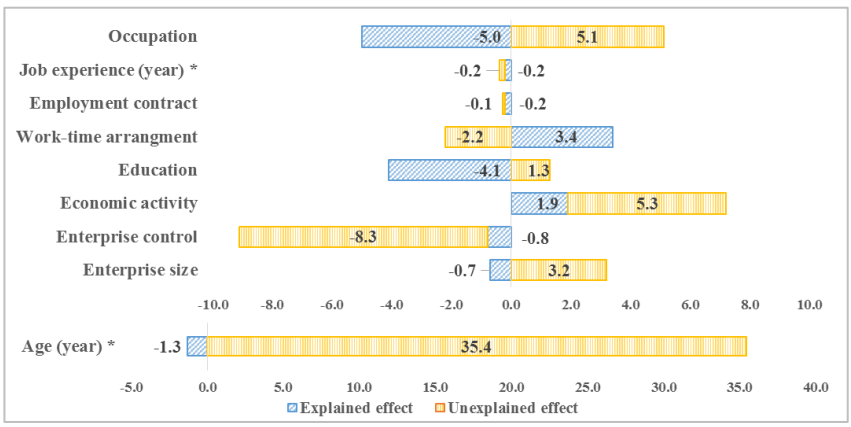
Finally, the explained and unexplained GPGs can be calculated by applying the decomposition results to the unadjusted GPG as follows:

$$GPG_{expl} = GPG_{unadj} * \frac{E}{\Delta}, \quad (7)$$

$$GPG_{unexpl} = GPG_{unadj} * \frac{U}{\Delta} \quad (8)$$

The percentages of the overall explained part and unexplained part of the difference (ln scale) are multiplied to the unadjusted GPG, using a proportional scaling, to derive the values of the two components. The values attributable to the job's characteristics are calculated as part of the total. A negative gap means that women are expected to earn more than men as possessing on average higher levels of a given characteristic than men. On the contrary, the positive signs indicate the factors that contribute to increasing inequality in favour of men. The results of the decomposition return a negative value for the explained GPG (-7%) and a positive value for unexplained (adjusted) GPG (+13.2%) given the unadjusted GPG equal to 6.2%. Figure 2 presents the results of the explained GPG for each characteristic.

Figure 2 – *Decomposition of the unadjusted GPG (in % of male hourly earnings), 2018.*



** included the variable and the variable squared.*

The overall explained GPG is mostly driven by three characteristics: occupation, education and age. The remuneration effect of the age on total unexplained GPG is prevalent (+35.4 percentage points) as consequence of the differences of observed age return shown in Table 2. The economic activities and the occupation follow (+5.3 and +5.1 percentage points).

6. Conclusions

The aim of this report was to investigate the unadjusted gender pay gap, an indicator with a simple definition but a very complex interpretation. As shown in the first part of the paper, the unadjusted GPG is extremely sensitive to composition effects arising from the different characteristics of the male and female workers, of the jobs and of the enterprise. In Italy the indicator is influenced by the value in the public economic and financial control sector while is more in line with the EU mean in the private one. On the wave of the international discussion about the opportunity to find an indicator suitable to monitor imbalances in earnings between women and men, the paper applies an experimental method, proposed by Eurostat, estimating the part of the earnings gap explained by differences in the average characteristics of male and female and the 'unexplained GPG' that come from different financial returns for the same characteristics. In Italy, results show that the explained GPG is negative, this turns into the unexplained (adjusted) GPG being higher than the unadjusted GPG. However, it is necessary to be cautions also in the interpretation of this results: it is not possible to conclude that the unexplained part measures earnings discrimination, nor that the part explained is exempt from it. Furthermore, these results are strictly related to the variables considered in the regression analysis and every change in that choice may substantially change the results. More in general characteristics included in the model are those available from SES. The discussion around the indicator is still ongoing at EU level and the necessity to find indicators able to grasp the core of the information demand to support public policies, persists.

References

BAZEN S. 2011. *Econometric methods for labour economics*. Oxford: Oxford University Press.

COTTON J. 1988. On the Decomposition of Wage Differentials, *Review of Economics and Statistics*, Vol.70, No.2, pp. 236-43.

EUROSTAT. 2021. Gender Pay Gap Statistics. *Statistics Explained*. https://ec.europa.eu/eurostat/statistics-explained/

International Labour Office. 2018. *Global Wage Report 2018/19: What lies behind gender pay gaps*. Geneva: ILO.

ISTAT. 2021. Structure of Earnings in Italy -Year 2018. *Statistiche Report*, Roma: Istat. https://www.istat.it/it/files//2021/04/SR_SES2018_EN_1_april.pdf

LEYTHIENNE D., RONKOWSKI P. 2018. A Decomposition of the Unadjusted Gender Pay Gap Using Structure of Earnings Survey Data. *Eurostat Statistical working papers*, Luxembourg: Publications Office of the European Union.

MINCER J. 1974. *Schooling, experience and earnings*. New York: National Bureau of Economic Research.

NEUMARK D. 1988. Employers' Discriminatory Behaviour and the Estimation of Wage Discrimination, *Journal of Human Resources*, Vol.23, No.3, pp. 279-95.

OAXACA R. 1973. Male-Female Wage Differentials in Urban Labour Markets, *International Economic Review*, Vol.14, No.3, pp. 693-709.

OCHSENFELD F. 2014. Why Do Women's Fields of Study Pay Less? A Test of Devaluation, Human Capital, and Gender Role Theory, *European Sociological Review*, Vol. 30, No. 4, pp. 536-48.

GOLDIN C. 2002. The Rising (and then Declining) Significance of Gender. *Working Papers 8915*, National Bureau of Economic Research.

OLSEN W., GASH V., VANDECASTEELE L., WALTHERY P., HEUVELMAN H. 2010. The Gender Pay Gap in the UK: 1995-2007: Part 1, *Research Report*, London: Government Equalities Office.

SUMMARY

Exploring earnings differences hidden behind the Gender Pay Gap

This paper deals with the issues underlying the measurement of hourly wage differentials between men and women. It introduces the GPG indicator, currently disseminated by Eurostat, on whose remodelling an international debate is underway aimed at improving its measurement. Furthermore, to understand what the GPG hides, it proposes an application on the Italian SES 2018 data of the Eurostat methodology based on the Oaxaca decomposition to adjust the indicator and evaluate how much of the differential is explained by the different discrimination factors.

Francesca CECCATO, Istat, ceccato@istat.it
Marilena Angela CIARALLO, Istat, ciarallo@istat.it
Paola CONIGLIARO, Istat, paola.conigliaro@istat.it

THE ACCURACY OF LONGITUDINAL LABOUR FORCE SURVEY ESTIMATES

Leonardo S. Alaimo, Alessio Guandalini, Antonella Iorio, Cristiano Marini, Alessandra Masi

1. Introduction

Besides cross-sectional data, the Italian Labour Force Survey (LFS) also provides longitudinal labor market data. The latter are obtained matching the members of the households who were interviewed in different time periods, due to the rotational scheme of the survey. In particular, individual records can be matched to produce 12-months and 3-months longitudinal data by involving almost 50% of the total sample. Since December 2015, the Italian National Institute of Statistics (Istat) has been providing 12-months estimates on the labour market flows, permanencies and transitions by occupational status (employment, unemployment, inactivity).

In this work, we present the methodology for computing the confidence intervals for these main indicators. Moreover, it is shown how the measures of accuracy, such as absolute or relative error, allow more precision analysis of the labour market.

The paper is organized as follows. In Section 2 the longitudinal LFS is briefly presented, the main problems are listed and the adopted solution are described. In Section 3 the methodological aspects related to the weighting procedure and moreover, the proposed methodology for computing the sampling variance are discussed. The method is applied on real data and some results are shown in Section 4. Finally, concluding remarks and further perspective are in Section 5.

2. The longitudinal data in the LFS

The Labour Force Survey (LFS) is the main source of information about the Italian labour market. It provides official estimates for a relevant number of indicators by using a sample based on a two-stage design with stratification of the first-stage units (municipalities) and a rotation scheme for the second-stage units (households). LFS aims to produce cross-sectional data. However, since each sampled household is interviewed for 4 quarters (rotation scheme 2-(2)-2: two consecutive occasions and, after a pause of two quarters, is then re-included in the sample for two other occasions), the records can be linked together to produce a rich

source of longitudinal data at 3-, 9-, 12- and 15-months (Ceccarelli *et al*., 2002). Two-quarter longitudinal data are produced linking the LFS data of each quarter. This can raise a number of methodological issues. In fact, according to the rotation scheme, the 50% of the cross-sectional household sample for each quarter should be re-interviewed 12 months after. A record linkage is carried out. However, some key variables are affected by errors of several types (response, coding, editing, etc.). In this case, a deterministic record linkage is not advisable; on the contrary, a probabilistic procedure is more advisable (Discenza *et al*., 2012).

In order to produce longitudinal datasets and the transition matrix, the following aspects must be considered:

- the longitudinal sample refers only to a specific longitudinal reference population, not to the entire one;
- LFS is not a panel survey, thus persons that move out of the selected households or household which move out of the municipality are not re-interviewed;
- household non-response may occur at subsequent waves due to refusal, non-contact, etc. (attrition);
- longitudinal sampling weights have to account for the longitudinal population, for the total non-response and to ensure coherence with the official LFS quarterly estimates.

For defining the longitudinal sampling weights, the most relevant methodological problems to be addressed are:

- definition of a suitable reference population for the longitudinal sample;
- longitudinal non-responses and eligibility;
- coherence between cross-sectional and longitudinal estimates.

2.1 The Longitudinal population for the Italian LFS

Longitudinal data for the Italian LFS concerns only people who are residents in the same municipality, both at the beginning and at the end of the period. Accordingly, the reference population[1] is defined as the resident population in the same municipality for 12 months or 3 months, thus net of deaths and of internal or international migration. The definition of the reference population directly influences the way in which the transition matrix is computed. The strategy followed by Istat is to provide flows estimates from LFS summing up two transition matrices obtained using a combination of two methods:

[1] There are several possibilities for defining the reference population for the longitudinal LFS. The choice depends on two aspects: the sample design and the availability of population totals for weighting.

- The first transition matrix contains stocks and flows estimates obtained from weighted longitudinal micro-data for the population which is resident in the same municipality both at the beginning and the end of the period. This component represents more than 96% of the total population still resident in the country; it provides very accurate estimates with many possible breakdowns (gender, age groups, NUTS region, level of education, etc.).
- The second transition matrix contains stocks and flows estimates obtained at a macro level for few domains by using retrospective questions from the cross-sectional sample at the end of the period for the population which moves across the country (internal migrants). This component, representing about 2% of the total population, has lower precision and could also contain some bias because it is based on a very small sample and it uses the main status one year before (the only which is known from retrospective questions) to model the transitions between occupational status.

2.2 The Longitudinal non-responses and eligibility

The longitudinal component of the LFS is also affected by units non-response such as:

- *Municipality non-response*: some municipalities are substituted in July at the beginning of a new annual survey cycle and some others may, for different reasons, fail to provide the interviews in subsequent waves;
- *Household and individual non-response*: some people do not fill in the questionnaire because they refuse to respond or the interviewers are unable to contact one or more individuals in a household. This kind of non-response can be assimilated to the "false negative" which occurs when the record linkage fails to match two records because of errors in the key variables.

Unit non-response may reduce the longitudinal component, thus increasing the variance of the estimates. Moreover, it can produce bias if non-respondents have significantly different labour features with respect to respondents.

Given the longitudinal population as defined in the previous pages, it is necessary to classify all the individuals interviewed at the first quarter into two groups:

- *Eligible*: they represent part of the longitudinal population that should be re-interviewed at the second wave.
- *Not-eligible*: they left from the initial population during the observed period thus they do not represent part of the longitudinal population. This is a very important theoretical concept but, since the LFS is not a panel survey, the individuals cannot be distinguished in eligible and not-eligible. Coherence between cross-sectional and longitudinal estimates.

A crucial problem is that the longitudinal component produces both cross-sectional and longitudinal estimates referred to the longitudinal population. The first ones, obtained from the longitudinal data, have to be consistent with the "official" estimates provided by the cross-sectional samples (the full sample) at the beginning and at the end of the observed period. The differences between these two kinds of cross-sectional estimates must be non-negative because they refer to the occupational status at the beginning and at the end of the period for people who left the initial cross-sectional population and for people who entered in the final cross-sectional population.

Since the longitudinal estimates have higher variability than quarterly official estimates, it is not possible to completely control their consistency. However, it is possible to reduce the risk of obtaining inconsistent results by using specific weighting strategy (Discenza, 2004).

2.3 The longitudinal main indicators

Transition rate: it is obtained as the ratio between the number of individuals who are in a different occupational status at the end of the period compared to the status at the beginning of the period and the stock relating to the condition at the beginning of the period. The rate can be seen as the probability of transition to a different occupational condition between the beginning and the end of the period.

Permanence rate: it is obtained as the ratio between the number of individuals who remains in the same occupational status during the period and the stock relating to the condition at the beginning of the period.

Workers' Separation Rate (WSR): it is equal, over a period of time, to the ratio between the people entering the occupation (UE, IE) and the sum of those who remain employed (EE), enter (UE, IE) and leave the occupation (EU, EI) in the same period considered.

$$WSR = \frac{UE+IE}{EE+(UE+IE)+(EU+EI)}$$

Workers' Hiring Rate (WHR): it is equal, over a period of time, to the ratio between people leaving employment (EU, EI) and the sum of those who remain employed (EE), enter (UE, IE) and leave employment (EU, EI) in the same period considered.

$$WHR = \frac{EU+EI}{EE+(UE+IE)+(EU+EI)}$$

with:

- UE: transition to employment from unemployment

- IE: transition to employment from inactivity
- EE: permanence in employment
- EU: transition to unemployment from employment
- EI: transition to inactivity from employment

Reallocation rate: it is given by the sum of the separation and hiring rates. It measures the well-being of the labour market and its elasticity and mobility.

3. Methodological aspects

The longitudinal data are built matching those of two quarters $\boldsymbol{Q}$ and $\boldsymbol{Q}'$. The weights used for providing the estimates are obtained solving two calibration problems. Calibration (Deville and Särndal, 1992; Särndal, 2007; Devaud and Tillé, 2021) is a widespread practice in National Statistical Institutes for several reasons. The leading reason is that it provides a system of weights that makes the sample consistent with known distributions of selected auxiliary variables. Furthermore, the calibrated weights can be used for providing all the estimates of the survey. This obscured the main reason, highlighted by Deville and Särndal (1992), which is the increase of the accuracy of the estimates when auxiliary variables strongly related with the interest variables and their totals are available.

Basically, calibration changes the design weights of the survey as little as possible for matching the totals of a set of auxiliary variables appropriately chosen. Then, the calibrated weights can be used for producing all the estimates. For longitudinal LFS, the weights of the cross-sectional sample, s, at the beginning of the period, d_k^Q are the starting point. However, the focus is just on the the *matchable* individuals ($s_{matchable}$) in the sample Q, that is, the individuals that are expected to be interviewed also in the quarter Q' because of the rotation scheme adopted in LFS. Their weights are calibrated mainly to reach consistency with quarterly estimates because the transition estimates have to be coherent with the estimates of the beginning quarter already published. Therefore, in the calibration system:

$$\begin{cases} min_{w_k^{(1)}} \left\{ \sum_{k \in s_{matchable}} G\left(w_k^{(1)}, d_k^Q\right) \right\} \\ \sum_{k \in s_{matchable}} w_k^{(1)} \boldsymbol{x_k^{(1)}} = \boldsymbol{t_x^{(1)}} \end{cases} \quad (1)$$

the auxiliary totals, $\boldsymbol{t_x^{(1)}}$, are some estimates from the quarter Q. Then, the weights $w_k^{(1)}$, are determined and used as the starting point for the second calibration in which the consistency with the longitudinal population is aimed. Only those related to the *matched* individuals ($s_{matched}$), that is individuals interviewed both in quarter Q and Q', are considered. The calibration system, in this case, is:

$$\begin{cases} min_{w_k^{(2)}} \left\{ \sum_{k \in s_{matched}} G\left(w_k^{(1)}, w_k^{(2)}\right) \right\} \\ \sum_{k \in s_{matched}} w_k^{(2)} \boldsymbol{x_k^{(2)}} = \boldsymbol{t_x^{(2)}} \end{cases} . \tag{2}$$

It is important to point out that in both the calibration system (1) and (2), $G(\cdot)$ is a pseudo-distance that measures the difference between the original and the final weights. In LFS, the truncated logarithmic distance to prevent negative or large weights is used (see, e.g., Deville and Särndal, 1992; Singh and Mohl, 1996).

The weights $w_k^{(2)}$ can be finally used for providing the estimate of a population total t_y,

$$\hat{t}_y = \sum_{k \in s_{matched}} w_k^{(2)} y_k \tag{3}$$

where y_k is the value of y variable observed on a unit k in the sample. Expression (3) refers to the calibration estimator and holds for estimating totals. However, its variance estimator cannot be directly applied to this context.

The first reason is that expression (3) takes into account just the last calibration step while, for properly addressing the variance estimation, it is necessary to consider both of them. A better approximation can be obtained writing the calibration system in (1) and (2) as a unique calibration system

$$\begin{cases} min_{w_k^{(2)}} \left\{ \sum_{k \in s} G\left(d_k^Q, w_k^{(2)}\right) \right\} \\ \sum_{k \in s_{matchable}} w_k^{(1)} \boldsymbol{x_k^{(1)}} = \boldsymbol{t_x^{(1)}} \\ \sum_{k \in s_{matched}} w_k^{(2)} \boldsymbol{x_k^{(2)}} = \boldsymbol{t_x^{(2)}} \end{cases} . \tag{4}$$

where the final weights $w_k^{(2)}$ are obtained changing the d_k^Q for matching at the same time $\boldsymbol{t_x^{(1)}}$ on the *matchable* individuals and $\boldsymbol{t_x^{(2)}}$ on the *matched* individuals. Then, the sampling variance can be approximated by

$$\widehat{AV}\left(\hat{t}_y\right) = \sum_{k \in s} \sum_{\ell \in s} \frac{\pi_{k\ell} - \pi_k \pi_\ell}{\pi_{k\ell}} (\hat{e}_k w_k)(\hat{e}_\ell w_\ell) \tag{5}$$

where π_k and π_ℓ are the first order inclusion probabilities, $\pi_{k\ell}$ are the second order inclusion probabilities and $\hat{e}_\cdot$ are the estimated residuals on y_k of the superpopulation model implicitly assumed by the calibration estimator defined by

the calibration system in (4).

Moreover, expression (5) holds for the variance of the total, while the main longitudinal indicators described in Section 2.3 are ratios, such as $\hat{R} = \hat{t}_y/\hat{t}_v$. Ratios are non-linear statistics, then, the standard formulas for the sampling variance cannot be directly used and a Taylor linearization is needed before.

The assumption on the basis of the Taylor linearization is that a non-linear statistic, such as ratios, can be approximated by its first-order Taylor. There are several ways for computing linearized estimators. All the methods are of common practice and usually lead to similar results (for further details, see: Wolter, 2007). The expression of the linearized variables for a ratio estimator computed on the sample is

$$\hat{z}_k = \frac{w_k^{(2)}}{\hat{t}_x}\left(y_k - v_k\hat{R}\right). \tag{6}$$

Replacing (6) in (5) the estimated residuals, $\hat{u}_k$, computed this time on $\hat{z}_k$ gives this expression

$$\widehat{AV}(\hat{t}_R) = \sum_{k\in s}\sum_{\ell\in s}\frac{\pi_{k\ell} - \pi_k\pi_\ell}{\pi_{k\ell}}(\hat{u}_k w_k)(\hat{u}_\ell w_\ell) \tag{7}$$

that can be used for approximating the sampling variance of the main longitudinal indicators. From expression (7) the relative error, $\sqrt{\widehat{AV}(\hat{t}_R)}/\hat{R}$, can be easily derived. Furthermore, under the assumption of normality, the 95% confidence intervals can be defined as $\left[\hat{R} - 1.96\sqrt{\widehat{AV}(\hat{t}_R)};\ \hat{R} - 1.96\sqrt{\widehat{AV}(\hat{t}_R)}\right]$.

4. First results

The methodology described in the previous section enables to measure the accuracy of the longitudinal estimates[2]. Table 1 and Table 2 show the relative error and confidence interval of longitudinal indicators, periodically disseminated by Istat, with reference to the last available data (from the 4th quarter of 2019 to the 4th quarter of 2020). Analyzing the professional condition of individuals aged 15-64 in the 4th quarter of 2020 and comparing it with that of the same period of the previous year, we can observe that 92.6% of the employed is still in employment (with an

[2] The estimates and the measures of accuracy have been computed using the package ReGenesees (Zardetto, 2015) of the R statistical software.

confidence interval ranging from 92.1% to 93%) and 21.5% of the unemployed (from 19.6% to 23.5%) and 6.6% of the inactive (from 6.1% to 7.2%) find a job; a third of the reference population (from 30.7% to 35.3%) remain trapped in unemployment. The total reallocation rate, which provides a measure of labor market mobility, is equal to 12.9% (from 12.3% to 13.4%); we observe a decrease of employment, mainly due to the separation rate (from 6.6% to 7.4%), which is significantly higher than the hiring rate (from 5.5% to 6.2%). Only 1 over 5 transits from fixed-term employment to permanent employment (ranging from 18.3% to 22%); on the other hand, more than 1 over 2 (with the upper limit of the confidence interval which is close to 60%) remains trapped in a precarious employment.

Table 1 – *Permanence and transition rate by occupational status over a 12-month period (estimate, relative error and 95% confidence interval). 2019 4th Quarter - 2020 4th Quarter.*

Permanence and transition rate in the professional condition	Estimate (%)	Relative error (%)	95% confidence interval	
			Lower bound	Upper bound
Permanence in employment	92.6	0.24	92.1	93.0
Transition from employment to unemployment	2.1	5.43	1.9	2.4
Transition from employment to inactivity	5.3	3.64	4.9	5.7
Transition from unemployment to employment	21.5	4.70	19.6	23.5
Permanence in unemployment	33.0	3.62	30.7	35.3
Transition from unemployment to inactivity	45.5	2.84	42.9	48.0
Transition from inactivity to employment	6.6	4.16	6.1	7.2
Transition from inactivity to unemployment	6.0	4.32	5.5	6.6
Permanence in inactivity	87.3	0.42	86.6	88.0
Reallocation rate	12.9	2.08	12.3	13.4
Hiring rate	5.9	2.96	5.5	6.2
Separation rate	7.0	2.92	6.6	7.4

Source: Istat, Labour force survey, longitudinal data.

Figures 1 and 2 present the lower and the upper limit of longitudinal indicators from 2013 to 2020. The quarters are considered separately in order to avoid seasonal effects. Figure 1 reports data on hiring and separation rates, from which we can see the slow overcome from the 2013 crisis; starting from the 1st quarter of 2015-2016 and (although not for all quarters) up to 2019, we observe a significant difference between the trends of the two rates (while the hiring rate tends to rise, the separation rate shows a decreasing trend). Since the 2nd quarter of 2020, the significant effect of the economic-health crisis has been very evident (the separation rate returns to be

significantly higher than the hiring rate with an important growth with respect to the previous period). Data show trends and gaps according to gender. Permanence rate in employment is always significantly higher for men than for women. For both categories, starting from 2019 and with reference to the 2nd and 3rd quarters, we observe significant decreases in the permanence rate. On the contrary, the transition from employment to inactivity rate is higher for women than for men, with a significant growth for both sex in the 2nd and the 3rd quarters of 2020.

Table 2 – *Transition from fixed-term employees over a 12-month period (estimate, relative error and 95% confidence interval). 2019 4th Quarter - 2020 4th Quarter.*

Transition rate from fixed-term employment	Estimate (%)	Relative error (%)	95% confidence interval	
			Lower bound	Upper bound
Transition to permanent employment	20.2	4.60	18.3	22.0
Transition to self-employed	2.0	17.45	1.3	2.6
Permanence in fixed-term employment	56.1	2.04	53.9	58.4
Transition to unemployment	8.5	7.20	7.3	9.7
Transition to inactivity	13.3	6.17	11.7	14.9

Source: Istat, Labour force survey, longitudinal data.

Figure 1 – *Hiring and separation rates (population aged 15-64, 95% confidence interval).*

Source: Istat, Labour force survey, longitudinal data.

Finally, the transition from inactivity to employment rate is always significantly higher for men than for women, with a significant decrease (rate from 5.8% falls to 4.6%.) only for women in the 2nd quarter of 2020.

Figure 2 - *Permanence rate in employment, by sex (employed 15-64 aged at t_0 = 100; 95% confidence interval).*

Source: Istat, Labour force survey, longitudinal data.

Figure 3 – *Transition rate from employment to inactivity, by sex (employed 15-64 aged at t_0 = 100; 95% confidence interval).*

Source: Istat, Labour force survey, longitudinal data.

Figure 4 – *Transition rate from inactivity to employment, by sex (employed 15-64 aged at t_0 = 100; 95% confidence interval).*

Source: Istat, Labour force survey, longitudinal data.

5. Concluding remarks

Accuracy is one of the main characteristics of a standardized measure, i.e., of a measure based on uniform procedures to collect, score and report numeric results. Those procedures must be subject to a verification of its proper functioning allowing to minimize the measurement errors, the random and the systematic one (Alaimo, 2020). Accuracy is a component (together with precision) of the reliability: the higher the random error the lower the level of reliability of the measuring instrument. Variables always contain a random error at different levels; this means that the same measurement process introduces this type or error and its effect of on reliability can only be estimated. The effects of random errors are totally a-systematic; an instrument affected by such an error may overestimate or underestimate the size measured in a certain object. From these considerations, we understand the importance of this study, which allows to measure the random error of LFS longitudinal data.

References

ALAIMO, L.S. 2020. Complexity of Social Phenomena: Measurements, Analysis, Representations and Synthesis. *Unpublished doctoral dissertation,* University of Rome "La Sapienza", Rome, Italy.

CECCARELLI C., DISCENZA A. R., ROSATI S., PAGGIARO A., TORELLI N. 2002. Le Matrici di Transizione della Rilevazione Trimestrale sulle Forze di Lavoro. *Nota metodologica*, Istat, Roma.

DEVAUD D., TILLÉ Y. 2019. Deville and Särndal's Calibration: Revisiting a 25-Years-Old Successful Optimization Problem, *TEST*, Vol. 28, No. 4, pp. 1033-1065.

DEVILLE J.C., SÄRNDAL C.E. 1992. Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, Vol. 87, No. 418, pp. 376-382.

DISCENZA A. R., FIORI F., LUCARELLI C. 2012 Weighting Issues in LFS Longitudinal Data, *Rivista Italiana di Economia, Demografia e Statistica*, Vol. 66, pp. 77-84.

DISCENZA A. R. 2004. Weighting Procedure for Longitudinal Data of the Italian Labour Force Survey. *Statistiche Report*, Roma, Istat.

SÄRNDAL C.E. 2007. The calibration approach in survey theory and practice, *Survey Methodology*, Vol. 33, No. 2, pp. 99-119.

SINGH A.C., MOHL C.A. 1996. Understanding Calibration Estimator in Survey Sampling, *Survey Methodology*, Vol. 22, No. 2, pp. 107-115.

ZARDETTO D. 2015. Regenesees: An Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys, *Journal of Official Statistics*, Vol. 31, No. 2, pp. 177-203.

WOLTER K. 2007. *Introduction to variance estimation*. Springer Science & Business Media.

SUMMARY

The accuracy of longitudinal labour force survey estimates

Besides cross-sectional data, the Italian Labour Force Survey products longitudinal data. Starting from December 2015, the Italian National Institute of Statistics provides 12months estimates on labour market flows, permanencies and transitions by occupational status (employment, unemployment, inactivity). In the present paper, the methodology for computing the confidence intervals for the main indicators disseminated is presented.

Leonardo S. ALAIMO, Istat, leonardo.alaimo@istat.it
Alessio GUANDALINI, Istat, alessio.guandalini@istat.it
Antonella IORIO, Istat, iorio@istat.it
Cristiano MARINI, Istat, cristiano.marini@istat.it
Alessandra MASI, Istat, masi@istat.it

EVALUATION OF CHANGING IN RESPONDENTS' PARTICIPATION IN THE SURVEYS OF INFORMATION AND COMMUNICATION TECHNOLOGIES USAGE IN ENTERPRISES (ICT) AND RESEARCH AND DEVELOPMENT FOR BUSINESS ENTERPRISES (R&D)

Claudio Ceccarelli, Gabriella Fazzi, Samanta Pietropaoli[1]

1. Introduction

This paper describes a longitudinal analysis of response rate, response burden, and paradata for the survey of Information and Communication Technologies usage in Enterprises (ICT survey, 2014 - 2020 editions) and the Research and Development for Business Enterprises (R&D survey, 2017 - 2021 editions).

The purpose of this paper is to explain how the characteristics of surveys (i.e., number of questions, number of hard or soft prompts) and the characteristics of respondents (i.e., number of employees, enterprises involved in the sample of each survey edition) can affect the response rate. In section 2, the characteristics of the survey and the web application developed for the questionnaire are focused. Using paradata, the characteristics of the survey and the connection with response burden are described in section 3, a set of indicators describing the behavior of the respondent and non-respondent companies is presented, and a logit model is proposed, trying to explain the propension to answer. Finally, Section 4, some conclusions are drawn, suggesting how to reduce the burden on respondents and increase participation.

2. Data

The Italian Survey on information and communication technology (ICT) usage is carried out by Istat on annual basis on active enterprises in industrial and services sectors, with at least 10 employees[2]. The Survey on Research and Development for

[1] This article is the result of the collaboration between the authors. In particular: Introduction and paragraph 1 are attributed to Claudio Ceccarelli, paragraph 2 to Gabriella Fazzi and paragraph 3 to Samanta Pietropaoli.

[2] The principal aim of this survey is to supply users with indicators on information society: Internet activities (web site, social media, cloud computing) and connection used (fixed and mobile broadband), e-Business (use of software as ERP, CRM), e-Commerce, ICT skills, e-Invoice.

Business Enterprises (R&D) is also carried out by Istat and is an annual census of the population of the Italian companies that can be identified as "potential R&D performers" in the reference year[3]. ICT and R&D are both web surveys; the questionnaires are carried out on the Istat Business Portal[4] and designed with GX (Generalised Italian Data Collection System XML), a software tool for developing/designing electronic questionnaire, since 2014 and 2017, respectively.

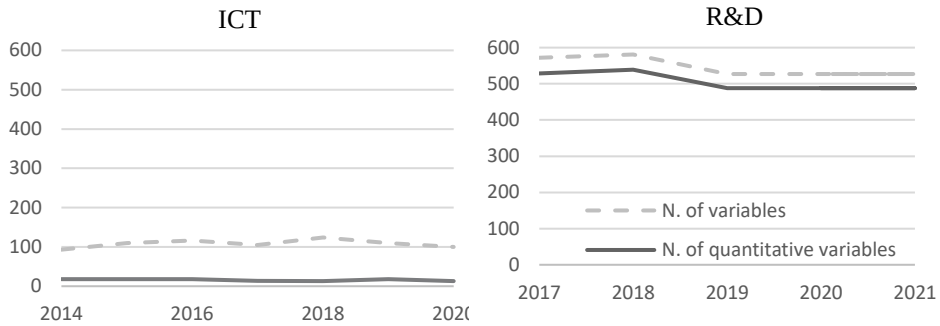
Paradata and metadata collected by the GX tool and the Business Portal starting from 2014 are analysed and we used them to try to explain the main factors for the response /nonresponse behaviour. Data on the number of requested variables, number of quantitative requested variables, number of hard and soft implemented prompts are used. The enterprises involved in the sample by size (number of employees) can classified and we are able to track whether the enterprise has done some action on the web Portal (answering the analysed surveys or other actions).

As shown in Figure 1, the ICT collects mostly qualitative information, contrary to the R&D, which collects mainly quantitative data. Also, the number of collected variables is different, in the ICT there are about one hundred for each edition, while in the R&D we work with 600 variables.

[3] It collects data about intramural and extramural R&D expenditure either within the unit (intramural) or outside it (extramural) and R&D personnel broken down by the business enterprise. Other information collected relates to the sources of funding for R&D activities and the type of research carried out (basic or applied research, experimental development).

[4] The Istat Business Portal is an integrated system for the management of data collection processes, which portal is at the same time an attempt to streamline the organization and production processes of business surveys.

Figure 1 – *Number of all type of variables vs numbe**r** of quantitative variables in ICT and R&D surveys.*



Source: our elaborations on Istat data.

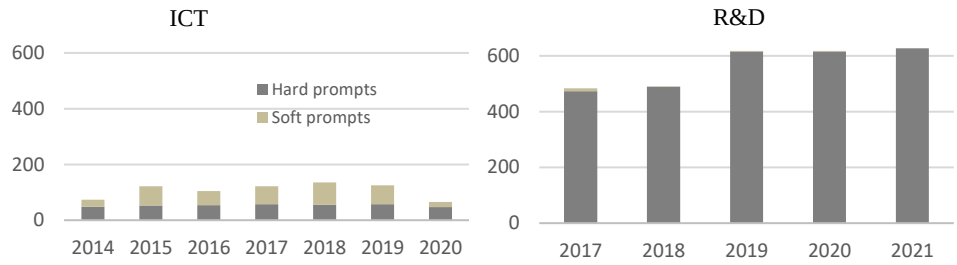
This characteristic, showed in Figure 1, highlights the different complexities of surveys: R&D is more demanding than ICT because respondents must find accounting information, referring to the current year and the two previous years, as well as expenditure forecasts for the following year.

The presence of many quantitative variables has led, over the years, to the insertion of many prompts, to allow for a consistent and higher quality data collection. In both surveys prompts are provided for unanswered questions, to check the question format and allowed character, and above all for consistency validations[5].

Nearly all prompts in the R&D survey are hard prompts, and the respondent cannot submit the questionnaire without correction. As Callegaro *et al*. (2015) note, "since validation messages interfere with the respondent's completion of the surveying task, they may be considered intrusive and annoying".

[5] The control of the consistency of responses with other data from the same survey or with the same answer in a previous survey of the same respondent.

Figure 2 – *Number of consistency validation prompts – ICT – R&D survey editions 2014-2020.*

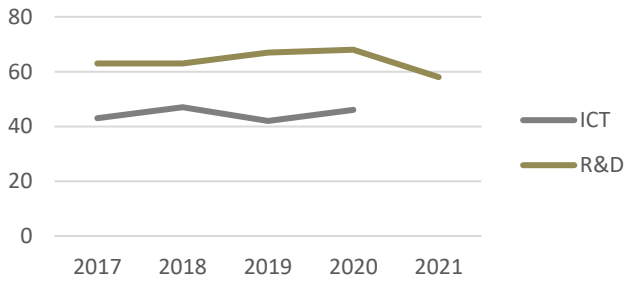


Source: our elaborations on Istat data.

Differences in the number of pieces of required information and constraints on completion lead to very different completion times for the two questionnaires as shown in Figure 3. On average it takes between 40 and 50 minutes for ICT online compilation, and usually more than 60 minutes for R&D. The commitment is heavier for larger companies, which fill in all sections of the questionnaire.

It is also worth mentioning some peculiar aspects of the response process in economic surveys, where respondents answer questions as representatives of their businesses and should be familiar with tables, matrices, and numerical information. Moreover, these surveys may need to be completed by multiple respondents, and the release of data may require approval by the company. Finally, printouts of web questionnaires are frequently used to support the preliminary process of identifying what information needs to be provided. Respondents often use paper forms as rough drafts before attempting to enter the data and answer the sequence of questions on multiple topics that appear on successive screens of a web survey (Morrison *et al*., 2008).

Figure 3 – *Compilation time online (mean in minutes) for ICT and R&D surveys*.



Source: our elaborations on Istat data.

In the ICT survey less than half of the questionnaire has been filled-in by only one referent. In one out of four cases more than two respondents had been involved in the retrieving of information (Table 1).

Table 1 – *People involved in ICT and R&D response process.*

	ICT 2020	R&D 2021
Single respondent	41.7 %	61.4 %
Two respondents	34.0 %	28.6 %
More than 2 respondents	24.3 %	10.0 %

Source: our elaborations on Istat data.

Table 2 – *ICT and R&D online and offline compilation time*.

	ICT		R&D	
	online	offline	online	offline
< 15 m	14.8	12.8	20.4	35.4
15-30 m	30.0	30.9	16.2	22.9
30-60 m	34.9	30.0	26.4	29.2
1-2 h	17.7	15.4	26.2	10.3
2-3 h	2.5	6.3	10.7	1.6
+ 3 h	0.0	4.5	3.1	0.6

Source: our elaborations on Istat data.

The time, and business cost, to consider for understanding the impact that completing the ISTAT questionnaires has on the business cannot just be the one automatically recorded by the systems. We need to add the time spent offline[6] to retrieve information or to contact other colleagues.
Comparing offline and online activities for the two surveys (Table 2), we observe that one out of four enterprises spend more than 1 hour offline working on the questionnaire for the ICT survey, compare with 12.5% for the R&D survey.

3. Response Rate Analysis

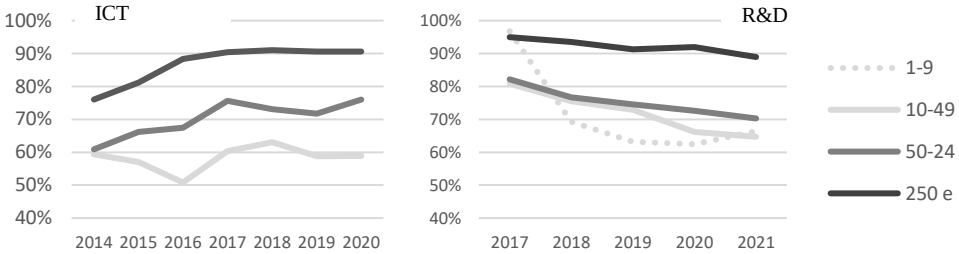
Focusing on the missing answers, we will try to understand why one enterprise out of three decides not to participate in the survey. The decision may depend on several factors: availability of the required information, lack of interest in the topics covered in the survey, authority attributed to the proposing agency, perceived response burden. Some firms (especially larger ones, with more than 250 employees) may be asked to respond to numerous surveys throughout the year, and this statistical

[6] Online time compilation, linked to paradata, is exhaustive and precise, which means it is automatically available for all enterprises who fill in the questionnaire, while data coming from the final section of the questionnaires are partial and subjective, as answering to this section was not compulsory.

burden is often significant. Understanding the incidence, but also the reasons for non-response can be important for identifying targeted strategies that can reactivate business involvement and increase participation.

The response rate trends among companies of different size ranges shows some interesting aspects: first, the response rate is always higher in larger companies; this is true not only for companies with more than 250 employees, which are subject to sanctions in the event of non-response, but also in companies that are not subject to sanctions.

Figure 4 – *Response rate for enterprise dimension (numbers of employees)- ICT and R&D.*



Source: our elaborations on Istat data.

This different response behavior may be due: a different strategy of reminders (the larger companies are also invited by telephone or by mail to fill out the questionnaire through a company that manages the outbound of respondents) or a more efficient internal organization in the larger companies, which allows time and/or resources to be dedicated to the filling out of the ISTAT questionnaires. It is interesting to observe the increases that in the ICT survey the response rate has had from 2017 onwards, and the decrease in the R&D survey (Figure 4).

We have examined some indicators that may provide some information regarding the phenomenon of missing responses:

- % of respondents who do not respond but had responded to the previous wave of the questionnaire;
- % of respondents who do not respond to a survey but respond to other surveys;
- % of non-respondents who have never accessed the Istat Business Portal.

In both ICT and R&D surveys, a portion of the sample must answer the questionnaire every year, while another portion changes from year to year. Those who had already responded to the survey the previous year are familiar with the questionnaire have already activated the information retrieval process within the company, and there are therefore somewhat facilitated to respond. On the other hand, returning to the sample each year increases the perception of statistical burden.

Using longitudinal data on sample inclusion and survey participation, we calculated for the two surveys the percentage of sample overlap (how much of the sample in *year t* was also present in *year t-1*) and the percentage of "lost" firms, i.e., missing firms that had responded the previous year.

Table 3 – *Samples overlap and "lost" enterprises in ICT and R&D survey*.

	ICT		R&D	
	Samples overlap	Lost	Samples overlap	Lost
2015 on 2014	32.0 %	10.9%	-	-
2016 on 2015	46.0 %	12.8%	-	-
2017 on 2016	33.0 %	6.1%	-	-
2018 on 2017	48.0 %	7.4%	68.0%	11.2%
2019 on 2018	62.0 %	9.3%	57.0%	11.7%
2020 on 2019	34.0%	9.4%	70.0%	6.8%
2021 on 2020	-	-	77.0%	5.3%

Source: our elaborations on Istat data.

In R&D surveys the sample overlap has been always bigger than that in the ICT survey, and the percentage of "lost" respondents is always lower. It is as if the R&D survey can count on a higher rate of loyalty from its respondents.

4. Focus on missing answers

As mentioned earlier, in both surveys approximately 66% responded to the questionnaire and completed the submission. What prevented the remaining 33% from fulfilling this task? A small percentage, 1.2% in ICT and 4.3% in R&D attempted completion but then left the questionnaire in draft form (Table 4).

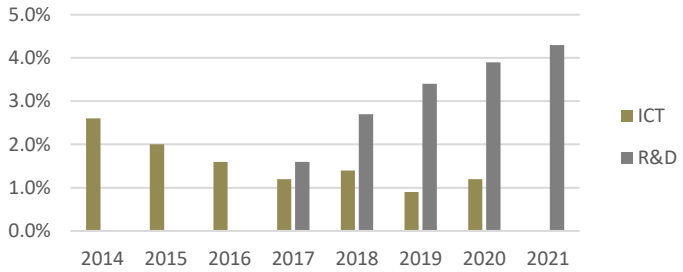
Table 4 – *ICT and R&D respondents behavior in 2021 and 2020 respectively*.

	ICT	R&D
Answered	66.1 %	66.8 %
Questionnaire in draft	1.2 %	4.3 %
Questionnaire opened and blank	3.4 %	4.1 %
Questionnaire never opened	11.9 %	15.8 %
Business Portal never opened	17.3 %	9.0 %

Source: our elaborations on Istat data.

Comparing the two surveys, we observe that the number of questionnaires left in draft has decreased over the years for ICT, while it has progressively increased for R&D (Figure 5). Every year, approximately 50% of the draft questionnaires (839 in 2021), moreover, had at least one active blocking error, that prevented their definitive sending. These are presumably questionnaires on which the high number of hard prompts had a discouraging effect.

Figure 5 – *Percentage of draft questionnaires in ICT and R&D surveys.*



Source: our elaborations on Istat data.

In fact, each company attempted the questionnaire twice before succeeding for ICT and 3.2 times for R&D. The multiplicity of attempts is due to the reporting of blocking errors for consistency checks that are triggered only at the end of the compilation, and that requires going back and correcting inconsistent data. In this regard, it is worth noting a hat is a red flag for R&D: 1172 companies (4% of respondents) changed their answer to the first question on R&D activity from "I did, I do, or I am planning to do R&D" to "I did not, I do not, and I am not planning to do R&D". This first response acts as a filter for the entire questionnaire, and a negative response allows to skip all the questions and the prompts, for a direct submission without penalty. To increase the participation of these companies, and enable them to complete the compilation, it is important to improve questionnaire design and usability principles (Couper, 2008).

Table 4 shows another interesting data point: 3.4% of the ICT sample and 4.1% for R&D merely opened the questionnaire, without beginning the compilation. To encourage their participation, therefore, it could be necessary to focus on more precise and clearer reminders. A considerable part of the sample, however, has never even opened the questionnaire webpage: this is 29.2% for ICT and 24.8% for R&D. Of these, a part ignored the invitation to access the questionnaire, yet accessed the Business Portal. They may have had difficulty in locating the web page where to begin the compilation, or they have other questionnaires to fill in and forgot or decide not to answer this one. An analysis of the user experience of the Portal could help to better understand the reasons for this dispersion, which led to the loss of 12% of the ICT sample and 16% of the R&D sample in the last edition of the survey.

The last interesting and numerically relevant category of non-respondents are those that we will call "wanted", that is, companies that have never even accessed the ISTAT Portal. These enterprises (17.3% for ICT and 9% for R&D) simply ignore the invitations and reminders they have received or have had difficulty accessing the Portal, which has discouraged them from participating in any economic survey.

5. Propension to answer: a logistic model

The data analysis concerns almost thirty-nine thousand enterprises included in the R&D sample for the year 2021 and twenty-one thousand enterprises from the ICT sample for 2020.

For each sample we studied the propension to answer through a logit model, with the dependent variable being the response registered and explanatory variables are some characteristics of the enterprises. We studied the influence of monetary sanction affecting enterprises with large dimension (over 250 employees)[7], the presence of the enterprise in previous wave of the questionnaire, and in all editions - as an indicator of low and medium statistical burden in the year (see Table 5). The indicator of statistical burden has been calculated with reference to the number of Istat statistical surveys in which the enterprise is required to participate in the year: low - if the enterprise is involved in less than five surveys; medium - if it is involved on at least five and less than twelve surveys; high – if it is involved on at least twelve surveys. In the model, our reference variable is the high burden.

Table 5 – *Variables used in the logit model.*

Variable	Description
RESPONSE	1-Respondent; 0- otherwise
SANCTION	1- Enterprise with 250 employees and over, 0- otherwise
PREC_EDITION	1- Enterprise included also in the sample of the previous years 0- otherwise
ALL_EDITION	1- Enterprise included also in the sample of all the previous years 0- otherwise
LOW_BURDEN	1- low statistical burden, 0- high statistical burden
MEDIUM_BURDEN	1- medium statistical burden, 0- high statistical burden

6. Results

Table 6 reports the change in model fit to the data obtained by omitting one explanatory variable at a time and instead retaining the others (Chi-square statistic), with the p-value associated. The results for both the survey show that each of the variables has a significant effect on the propensity to answer of the enterprise (column 3 and 5); in particular, the level of the Chi-Square statistic for both survey (column 2 and 4) shows that having participated in the previous edition is the most relevant factor among those considered. This is followed by monetary sanction,

[7] All enterprises in the sample are informed about the mandatory nature of the response, however, the sanction only affects large companies.

while, low level of burden, participation in all precedent editions seem to have a more limited impact on the phenomenon under consideration.

Table 6 – *Explanatory variables for Wald Chi-square statistic and associated p-value.*

	ICT		R&D	
	W.Chi-Sq.Test	Pr > Chi-Sq.	W. Chi-Sq.Test	Pr > Chi-Sq.
Sanction	69.7865	<.0001	65.6011	<.0001
Pre_edition	95.3255	<.0001	79.5082	<.0001
All_edition	36.7653	<.0001	0.6554	<.0001
Low_burden	41.2861	<.0001	4.6974	<.0001
Medium_burden	24.1219	<.0001	2.9872	<.0001

Source: our elaborations on Istat data.

When all variables involved in the model are categorical, as it is the case here, interpretation of the results may be more straightforward by directly examining the Odds Ratio[8], rather than the parameters.

Table 7 – *Odds Ratio for model-related regressors.*

	ICT	R&D
	Estimate	
Sanction (1 vs 0)	1.826	2.089
Pre_edition (1 vs 0)	1.361	39.445
All_edition (1 vs 0)	1.454	1.411
Low_burden (low vs high)	0.039	2.982
Medium_burden (medium vs high)	0.084	2.386

Source: our elaborations on Istat data.

All coefficients of the equation have significant values at 1% significance level. These coefficients indicate the change of the log odds of being respondent at a unit increase of the predictor variable.

Among the principal results in the ICT survey, the monetary sanction appears to be the main factor that drives companies to participate, as it corresponds to the highest probability of being a respondent. The burden of the company in the same year and the involvement in the previous edition of the survey seem to have less weight. In R&D, however, it is precisely this element that strongly encourages participation: this can be explained by the fact that the survey also requires retrospective data each year, and so for an enterprise that has already responded to edition *t-1*, the task of responding to edition *t* is, in effect, much simplified.

[8] An odds ratio (OR) is a measure of the association between an exposure and an outcome. The OR represents the probability that an outcome will occur given a particular exposure, compared with the probability that the outcome will occur in the absence of that exposure.

7. Discussion and future developments

The results of the analysis of the paradata and metadata of the two surveys are obviously not exhaustive and open the field to new ideas for analysis and research. A specific focus on the "wanted", that is on the "hard-core" of companies that have never accessed the Portal, is necessary to study strategies aimed at obtaining their collaboration, capturing their attention. Again, multivariate analysis can be conducted on a broader and longitudinal database, to try to understand the factors that lead the two surveys to have opposite trends in the response rate.

However, the main findings presented in the paper are already useful to indicate some actions that the Data Collection Directorate can take to try to reduce the burden on the companies themselves, without losing the wealth of information collected and hopefully increasing response rates and data quality.

It would be possible to try to engage the wanted with targeted outbound call campaigns, guiding them through their first access to the Portal. The internal paths within the Portal itself can be redesigned, improving the user journey, and making the call to action more immediate: highlighting the deadline of each survey and making access to the questionnaire immediate can be two first steps in this direction, improving the user experience and usability of the online data collection instruments and web Portal.

The questionnaires need to be as clear as possible, and the data consistency check required some simplification.

The burden would be reduced if different questionnaires, from different surveys, were linked together and with the administrative data that the enterprise already communicates to the PA. A firm involved in several samples would not be forced to provide the same information multiple times. This requires a work of standardization of variables and coordination of survey times between different surveys that should be discussed at Eurostat tables (EUROSTAT, 2014).

Finally, it would be important that the primary motivation for responding to the proposed questionnaires was not fear of monetary sanction. The strategies to be put in place to engage these respondents cannot be just "punitive", for example, by focusing on economic sanctions. It would be opportune to focus on more effective communication strategies and a return of information on the respondent capable of making them perceive the importance of the data collected by the Institute even for the individual business. An example is the experience of Portugal, which returns sectoral benchmark reports to companies that participate in the surveys.

For companies to understand the importance of the data collected by official statistics for the country, it is necessary that those same data return part of their value to the companies themselves.

References

CALLEGARO M., LOZAR MANFREDA K., VEHOVAR V. 2015. *Web survey methodology*. London: Sage.

COUPER M. P. 2008. *Designing effective web surveys*. New York: Cambridge University Press.

EUROSTAT 2014. *Memobust handbook - Handbook on Methodology of Modern Business Statistics.*

MORRISON R. L., STOKES S. L., BURTON J., CARUSO A., EDWARDS K. K., HARLEY D., HOUGH C., HOUGH R., LAZIRKO B. A., PROUDFOOT., S. 2008. *Economic Directorate guidelines on questionnaire design.* U.S. Census Bureau.

SUMMARY

Evaluation of Changing in Respondents' Participation in the Surveys of Information and Communication Technologies Usage in Enterprises (ICT) and Research and Development for Business Enterprises (R&D)

In the paper we analyzed paradata and metadata of two business surveys: ICT or R&D survey, with the aim to understand the response behavior of the respondents. Starting with a descriptive analysis, we focus on characteristics of the surveys (i.e., number of questions, number of hard or soft prompts) and on the characteristics of the respondents (i.e., number of employees, enterprises involved in the sample in each survey edition) for both the surveys. Then we analyzed response rate by enterprise size, and we propose a set of indicators describing the behavior of the respondent and non-respondent enterprises. We also study paradata of the users as the time taken to compile the questionnaire, online and offline, with the goal to find some actions to improve the participation of the respondents and their perception of the official statistics. For this reason, we also focused our attention on a specific subset of respondents - the so-called "wanted" - the ones who have never answered the ICT or R&D survey or any other Istat survey- and the “lost” ones, which had answered to the previous year but are missing in the current year.

Finally, we apply a logit model intending to explain which of respondents’ characteristics affect response rate and define what kind of actions can be taken to improve participation.

Claudio CECCARELLI, Istat, clceccar@istat.it
Gabriella FAZZI, Istat, fazzi@istat.it
Samanta PIETROPAOLI, Istat, pietropaoli@istat.it

THE ROLE OF HIGHER EDUCATION INSTITUTIONS IN SUSTAINABLE DEVELOPMENT: A DSGE ANALYSIS

Marina Albanese, Francesco Busato, Gianluigi Cisco

1. Introduction

This paper explores the role of higher education institutions in achieving sustainable development, employing a Dynamic Stochastic General Equilibrium (DSGE) model extended to consider the European Union (EU) educational system and households' environmental awareness. In particular, this paper addresses the following fundamental questions: (i) how HEIs productivity affect households' environmental awareness? (ii) how human capital and environmental awareness interact with the business cycle?

Economic theory on sustainable development mainly focuses on the incentives to reduce Greenhouse Gases (GHGs) emissions, but it is very important to analyze the processes of the definition of preferences of individuals and households to define a real change. In opposition to a large managerial and organizational literature on this topic, economic analysis has few contributions.

In order to fill this gap in the literature, this paper extends previous studies applying environmental DSGE models, embedding human capital accumulation, environmental variables, and households' environmental awareness.

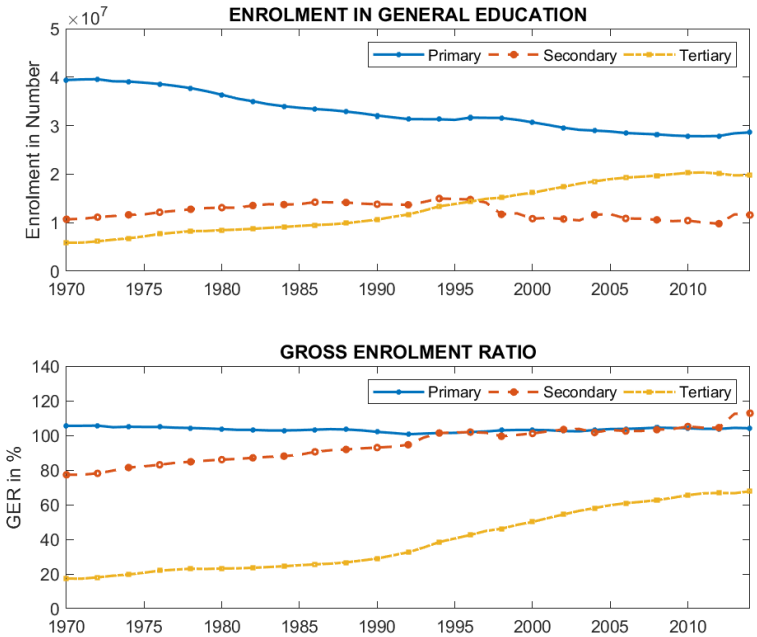
Our results can be summarized as follows. First, enrollment in tertiary education and environmental awareness are procyclical. Second, households become more sensitive to environmental issues during a positive technology shock. Finally, higher academic institutions play a key role in reinforcing the above mechanism, allowing sustainable development.

This paper is structured as follows: Section 2 discusses the main stylized facts in global and EU-27 education. Section 3 presents a DSGE model embedding time in education and household environmental awareness. Section 4 presents the model calibration. Finally, Section 5 presents the impulse response analysis. Finally, section 6 concludes.

2. Education: Empirical Evidences

This section provides an overview of long-run changes in education outcomes and outputs worldwide, focusing on the EU-27. From a historical perspective, the world went through a great expansion in education over the past two centuries. Global literacy rates have been climbing over the last two centuries, mainly through increasing enrollment rates in primary education. Secondary and tertiary education have also seen drastic growth, with the global average years of schooling being much higher than a hundred years ago. Fig. 1 displays the evolution of the European Union 27 (UE-27) education from 1970 to 2014. We consider two indicators of human capital: total enrollment and the enrollment ratio (GER) in primary, secondary, and tertiary education, all courses, and both sexes. These data are obtained from the World Bank indicator database. From Fig. 1, it is possible to identify a clear positive trend for secondary and tertiary education in the sample period. In particular, starting from 1990, the enrolment in tertiary education has increased drastically (about 40 percentage points in 1970-2014).

Figure 1 – *Enrolment and the GER in all courses, both sex, EU-27- 1970-2014*



3. The Model

This section presents a DSGE model to examine the role of higher education institutions in achieving sustainable development. This model presents three agents: a representative household, a representative higher academic institution, and a representative firm. We configure the model in the following way. Households maximize expected utility defined over to consumption, environmental quality, labor effort, and education. They can invest in abatement activity and physical capital. The representative firm produces goods employing effective labor and capital. Academic institutions transform education into new human capital. In detail, we expand a standard *Real Business Cycle* (RBC) model with human capital considering two significant features that have non-trivial implications for sustainable development analysis. First, our model embeds the environmental sector and carbon emissions dynamics. Second, households are environmentally aware and use part of their resources to protect the environment. The unique source of uncertainty is a technology shock on output.

3.1. Higher Academic Institutions

The higher academic institutions transform time devoted to education from households into new human capital. The new human capital assumes the following functional form:

$$i_t^h = A_{h,t}(h_t e_t)^{\varphi_e}[(1 - v_t k_t)]^{1-\varphi_e} \quad (1)$$

where $(1 - v_t)$ is the remaining fraction of physical capital allocated to the human capital investment sector; and $A_{h,t}$ is the human capital technology productivity. According to literature in this filed, we assume diminishing return to education; it is added to the model by the assumption that φ_e takes a value in the interval (0,1). In line with Ben-Porath (1967), an agent forms new human capital stock by combining time (e_t), talent, and current human capital stock (h_t). In addition, we assume that new human capital is linked to physical capital (k_t). Agents in the economy are investing in themselves to maximize their expected lifetime returns. In detail, households make a human capital investment to the point that the marginal return of a unit investment is equal to the alternative marginal return they can earn with the time they use for acquiring human capital.

3.2. Firms

In the model, input markets are perfectly competitive. Firms, which take factor prices as given, rent physical capital and employ labor force in order to maximize profits. The representative firm maximize their profits by choosing the optimal quantity of capital and labor:

$$max_{k_t,l_t}\Pi_t = y_t - v_t r_t k_t - w_t n_t h_t \quad (2)$$

s.t

$$y_t = A_t(v_t k_t)^{\alpha}\left(h_t\ n_t\right)^{1-\alpha} \quad (3)$$

where A_t is the total factor productivity shock and follows an AR (1) stochastic processes:

$$log(A_t) = \rho_a log(A_{t-1}) + (1-\rho_a)A_{ss} + \varepsilon_t^a \quad (20)$$

where A_{ss}, is a constant and represents the steady-state values; $0 < \rho_a < 1$ is the autoregressive parameters; ε_t^a, is Gaussian i.i.d. shocks with zero means and known variances, σ_a^2. The first order conditions for the firms are:

$$r_t = \alpha \frac{y_t}{v_t k_t} \quad (4)$$

$$w_t = (1-\alpha)\frac{y_t}{h_t\ n_t} \quad (5)$$

The representative firm hires labor until the marginal product of effective labor is equal to the wage rate, w_t, and rent capital until the marginal product of physical capital is equal to the rental rate, r_t.

3.3. Households

Households maximize expected utility defined over to consumption, leisure, and the environment investment ratio. The period utility function is:

$$U_t(c_t, l_t, m_t) = \sum_{t=0}^{\infty}\beta^t[log(c_t) + \varkappa log(l_t) + \mu log(m_t)] \quad (6)$$

where c_t is consumption per capita, l_t denotes leisure, β is the discount factor, $\varkappa$ leisure weight, μ is the environmental quality weight. As in Zhang *et al.* (2019), we consider that the environment investment ratio $m_t$affects positively the household's utility function. We use the environmental investment as a proxy of the

environmental awareness. The representative agent is confined by a unitary time endowment constraint for every period:

$$l_t + n_t + e_t = 1 \tag{7}$$

where n_t is the labor employed in goods production, and e_t is time spent in education. Households maximize expected utility subject to the flow budget constraint:

$$c_t + i_t + m_t = h_t\ w_t n_t + v_t r_t k_t \tag{8}$$

Households consume goods, invest in the production sector (i_t) and in the abatement activity (m_t). Households finance these expenditures through effective wage ($h_t\ w_t$) income from the production firms, and the return they receive from the investments in the previous period (v_t is the fraction of capital devoted to goods production). The human capital stock (h_t) evolves according to the following law of motion:

$$h_{t+1} = i_t^h + (1 - \delta_h)h_t \tag{9}$$

where i_t^h define the new human capital and δ_h is the depreciation rate of human capital. The stock of capital is usual and evolve according to the following law of motion:

$$k_{t+1} = (1 - \delta)k_t + i_t \tag{10}$$

where δ is capital depreciation rate.

Households maximize their lifetime utility. The choice variables in the maximization problem are the consumption level, the investment level, the environmental investment level, the education time, the work time, the physical, and the human capital that they plan to invest for the next period.

The following set of equations characterize the intertemporal maximization problem:

$$max_{c_t,n_t,m_t,k_{t+1},h_{t+1},e_t,v_t} \sum_{t=0}^{\infty} \beta^t [log(c_t) + \varkappa log(l_t) + \mu log(m_t)]$$

s.t:

$$l_t + n_t + e_t = 1$$

$$c_t + i_t + m_t = h_t\ w_t n_t + v_t r_t k_t$$

$$h_{t+1} = A_{h,t}(h_t e_t)^{\varphi_e}[(1 - v_t k_t)]^{1-\varphi_e} + (1 - \delta_h)h_t$$

$$k_{t+1} = i_t \ + (1 - \delta \ )k_t$$

$$given \ \ k_0 \ , h_0 \ > 0, \forall t$$

The first order conditions for consumption, environmental investments, labor, education, fraction of capital in goods production, physical capital and human capital are the follows:

$$\frac{1}{c_t} = \lambda_t \quad (11)$$

$$\frac{\mu}{m_t} = \lambda_t \quad (12)$$

$$\frac{\varkappa}{l_t} = \lambda_t w_t h_t \quad (13)$$

$$\frac{\varkappa}{l_t} = \varrho_t A_{h,t}(\varphi_e)(h_{t+1}e_t)^{\varphi_e}[(1 - v_t k_t)]^{1-\varphi_e} h_t \quad (14)$$

$$r_t \ \ \lambda_t = \varrho_t(1 - v_t)A_{h,t}(1 - \varphi_e)(h_t e_t)^{\varphi_e}[(1 - v_t k_t)]^{-\varphi_e} \quad (15)$$

$$\lambda_t = \beta\big[r_{t+1}v_{t+1}\lambda_{t+1} + \varrho_{t+1}(1 - v_{t+1})A_{h,t+1}(1 - \varphi_e)(h_{t+1}e_{t+1})^{\varphi_e}[(1 - v_{t+1}k_{t+1})]^{-\varphi_e} + 1 - \delta\big] \quad (16)$$

$$\varrho_t = \beta\big[\varrho_{t+1}e_{t+1}A_{h,t+1}(\varphi_e)(h_{t+1}e_{t+1})^{\varphi_e}[(1 - v_t k_{t+1})]^{1-\varphi_e} + \lambda_{t+1}w_{t+1}n_{t+1} + 1 - \delta_h\big] \quad (17)$$

where λ_t and ϱ_t are the Lagrangian multipliers associated to the budget constraint and the law of motion of human capital, respectively. Eq. 11, 12, 13 14, represent marginal utility for consumption, environmental protection, labor and education, respectively; Eq. 15 equates weighted factor intensities across sectors; Eq. 16 and 17 are the Euler equations for physical and human capital.

3.4. Natural Resources and Carbon Emissions

One of the novelties of this study is to consider environmental variables in a standard RBC model with human capital. As in previous studies, we assume that production activity emits carbon dioxide in the atmosphere. In particular, emissions per firm (em_t) are by-product of output:

$$em_t = (\epsilon - m_t)y_t \quad (18)$$

Emission intensity ($e\, m_t/y_t$ depends on the abatement technology ($\epsilon - m_t$), where ϵ defines the emissions for unit of output. This study considers also

environmental quality dynamics. As in Angelopoulos *et al.* (2013), environmental quality evolves according to the following law of motion:

$$q_t = \left(1 - \delta_q\right)eq + \delta_q q_{t-1} - em_t \quad (19)$$

where the parameter eq represents environmental quality without pollution, and $0 \leq \delta_q \leq 1$ is a parameter measuring the degree of environmental persistence.

4. Calibration

The model is calibrated for the EU 27 and time is measured in quarters. Table 1 lists all the parameters of the model. For conventional parameters, we use the standard estimates used in the business cycle literature (e.g., Smets and Wouters, 2003). The discount factor is set at a value consistent with a real interest rate of 4% per year, that is 0.99. The depreciation rate of capital is set at 0.025 and the capital share at 1/3. Regarding the environmental part of the model, we refer to previous environmental DSGE models for climate change to obtain plausible values for environmental parameters. As in Heutel (2012), we set the emission intensity parameter to 0.45. For the environmental quality dynamics, we refer to Angelopoulos *et al.* (2013).

Table 1 – *Baseline Parameter Calibration.*

Parameter	Description	Value	Source
χ	Leisure Weight	2.00	Endogenous Calibration
μ	Environmental Quality Weight	0.14	Endogenous Calibration
δ	Capital Depreciation Rate	0.02	Smets and Wouters (2003)
δ_h	Human Capital Depreciation Rate	0.01	Kim and Lee (2007)
α	Capital Share in Production	0.33	Smets and Wouters (2003)
φ_e	Education Technology	0.8-0.9	Kim and Lee (2007)
ϵ	Emissions Intensity	0.45	Heutel (2012)
δ_q	Natural Depreciation Rate	0.90	Angelopoulos et al. (2013)
ρ_a	Persistence of TFP shock	0.95	Smets and Wouters (2003)

More precisely, we define the persistence of environmental quality as equal to 0.95. In contrast to the other parameters, there is relatively little econometric evidence on the parameter in the human capital formation equation. Achieving effectiveness and efficiency in higher education depends on public authorities creating the right framework within which higher education institutions can operate. For this reason, we assume two alternative calibrations for education productivity in the HEI

production function. The two alternative calibrations are in line with the literature (e.g., Heckman, 1976) and allow us to explore the role of higher institutions with different educational structures. The depreciation rate of the human capital is equal to 0.01, as in previous studies in this field. Finally, for the stochastic processes of the model, we assume a high degree of autocorrelation for the exogenous shocks by setting at 0.95.

5. Results

In order to assess the role of higher education institutions on carbon emissions and the business cycle, we analyze the dynamic properties of the model under two alternative calibrations of education productivity parameters: $\varphi_e = 0.8,0.9$. We can think of this different calibration as national reform in tertiary school education improving study quality and increase returns to the education received[1].

5.1. Impulse Response Analysis

Figure 2 shows the impulse response functions for a technology shock in the goods production sector.

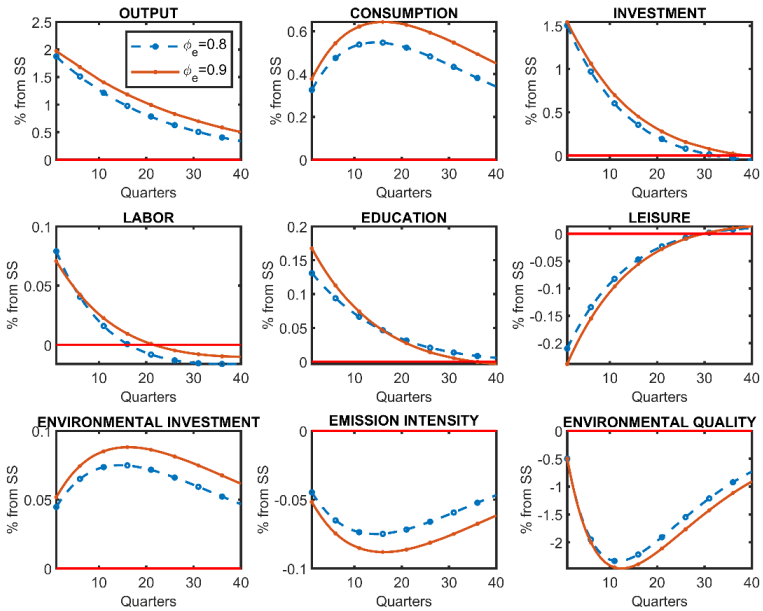
Following a positive technology innovation, both calibration, output, consumption, investment, and labor all rise persistently. Households find optimal increase investments, work harder and increase education during the early phases of the adjustment process when productivity is higher. Households allocate more time to education in order to increase their future human

capital. We find a pro-cyclical pattern for time spent in education. This finding contrasts with the results of a strand of research considering countercyclical fluctuations for the human capital (Kim and Lee, 2007, among others). However, our finding is in line with Malley and Woitek (2011) and King and Sweetman (2002), investigating the empirical relationship between college enrollment and output and evidence in favor of procyclicality. Consequently, the increase in time spent in education and improvement in firm productivity affect the household's

[1] The simulations have been obtained using numerical analysis and perturbation methods to simulate the economy and compute the equilibrium conditions outside the steady-state. We solve the model using a second-order Taylor approximation around its steady state. All results are reported as percentage deviations from the steady-state.

environmental awareness, increasing investments in environmental protection. This latter occurs for two main reasons. First, the increase in productivity induces a corresponding increase in emissions, worsening the environmental quality. Hence, households respond by increasing the abatement effort to preserve their well-being. Second, after this shock, their income increase.

Figure 2 – *Impulse response functions to a one percent technology shock*.



As a consequence, households use part of their resources to preserve the environmental quality. Our finding is in line with recent literature investigating the role of the business cycle in changing people's environmental concerns. Scruggs and Benegal (2012) find that public opinion about global warming is variable and driven by the business cycle and economic insecurity. Kahn and Kotchen (2010) find that an increase in a state's unemployment rate is associated with a decrease in the probability that residents think global warming is happening and a reduced investment in environmental protection. Turning to the national reforms in school education scenario, positively affecting returns to the education received, this study presents interesting results. In detail, when the education productivity increase households prefer to reduce their leisure and increase their investment in education. Households' education choices affect business cycle dynamics in two ways. First, it positively affects firms' productivity, amplifying beneficial effects from productivity shock on output. Second, investment in education affects workers' skills, allowing

them to earn greater wages. Consequently, households increase consumption, investments in physical capital, and environmental protection. In conclusion, a better structure of educational institutions allows achieving sustained growth economic with a lower impact on the environment. The rise in environmental protection associated with economic growth contributes to a reduction in emission intensity.

6. Conclusions

This paper provides selected insights to reason on the role of higher academic institutions in achieving sustainable development, while households choose between consumption, education, and labor and decide how to allocate their saving between environmental protection and goods production. To investigate this question (and possibly many others), this manuscript designs an equilibrium model capable of capturing the trade-off between environmental-compliant choices and those based only on crude economic drivers, focusing on how investment in education can affect it. This document is motivated by the recent rise in awareness about climate change issues and their consequences (e.g., "Fridays for Future" Climate Strike implications). However, most of the existing literature on environmental policy analysis (e.g., Fischer and Springborn, 2011; Heutel, 2012; Annicchiarico and Di Dio, 2015) neglects the role of the household in this story. Since households' behavioral changes are one of the key factors of sustainable economic development, neglecting their preferences could result in a biased calculation of environmental policies performances. Therefore, ignoring education and awareness aspects in theoretical models means disregarding an important channel for macroeconomic fluctuations and suggesting misleading policy recipes. This study suggests that time devoted to education and investment in abatement activity are pro-cyclical. First, households are willing to reduce their leisure during the economic growing phases, increasing time devoted to labor and education. Second, as documented in the literature, environmental concern is linked to the business cycle in this study. Short-term economic conditions and environmental quality anomalies affect opinions and concerns about climate change. Consequently, households become more sensitive to environmental issues during a positive technology shock. The mechanism is reinforced when higher academic institutions become more efficient in their educational activity. Tertiary education plays a strategic role in the processes of sustainable development. Some points in our future research agenda will be related to the dynamics of accumulation of a specific green human capital affecting households' choices and the impact of HEIs policies on students and households' pro-environmental attitude and their effect on sustainability, providing a welfare analysis.

References

ANNICCHIARICO, B., DI DIO, F. 2015. Environmental Policy and Macroeconomic Dynamics in a New Keynesian Model, *Journal of Environmental Economics and Management*, Vol. 69, pp. 1-21.

ANGELOPOULOS, K., ECONOMIDES, G., PHILIPPOPOULOS, A. 2013. First- and Second-Best Allocations Under Economic and Environmental Uncertainty, *International Tax and Public Finance*, Vol. 20, No. 3, pp. 360-380.

BEN-PORATH, Y. 1967. The Production of Human Capital and the Life Cycle of Earnings, *Journal of political economy*, Vol. 75, No. 4, Part 1, pp. 352-365.

FISCHER, C., SPRINGBORN, M. 2011. Emissions Targets and the Real Business Cycle: Intensity Targets Versus Caps or Taxes, *Journal of Environmental Economics and Management*, Vol. 62, No. 3, pp. 352-366.

HEUTEL, G. 2012. How Should Environmental Policy Respond to Business Cycles? Optimal Policy Under Persistent Productivity Shocks, *Review of Economic Dynamics*, Vol. 15, No. 2, pp. 244-264.

KAHN, M. E., KOTCHEN, M. J. 2010. Environmental Concern and the Business Cycle: the Chilling Effect of Recession, *NBER Working Papers* 16241, National Bureau of Economic Research.

KIM, D., LEE, C.-I. 2007. On-the-job Human Capital Accumulation in a Real Business Cycle Model: Implications for Intertemporal Substitution Elasticity and Labor Hoarding. *Review of Economic Dynamics*, Vol. 10, No. 3, pp. 494-518.

KING, I., SWEETMAN, A. 2002. Procyclical Skill Retooling and Equilibrium Search. *Review of Economic Dynamics*, Vol. 5, No. 3, pp. 704-717.

MALLEY, J., WOITEK, U. 2011. Productivity Shocks and Aggregate Fluctuations in an Estimated Endogenous Growth Model with Human Capital, *CESifo Working* Paper *Series* 3567, CESifo.

SCRUGGS, L., BENEGAL, S. 2012. Declining Public Concern about Climate Change: Can We Blame the Great Recession? *Global Environmental Change*, Vol. 22, No. 2, pp. 505-515.

SMETS, F., WOUTERS, R. 2003. An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area, *Journal of the European economic association*, Vol. 1, No. 5, pp. 1123-1175.

ZHANG, S., et al. 2019.The Effects of Energy Price, Technology, and Disaster Shocks on China's Energy-Environment-Economy System, *Journal of cleaner production*, Vol. 207, pp. 204-213.

SUMMARY

The role of higher education institutions in sustainable development: a DSGE analysis

The higher education sector is recognized as one of the major players in advancing sustainability through its research, education, and societal engagement. In order to investigate the role of higher education in achieving sustainable development, this study proposes a *Dynamic Stochastic General Equilibrium* (DSGE) model embedding human capital accumulation, environmental variables, and households' environmental awareness (i.e., investments in environmental protection). This paper studies the dynamic behavior of education and its link to environmental awareness, considering different higher education institution productivity parameters and a technology shock. This paper offers three main results. First, enrollment in tertiary education and environmental awareness are procyclical: households are willing to reduce their leisure during positive economic phases, increasing time devoted to labor and education. Second, as documented in the literature, environmental concern is linked to the business cycle in this study. Concerns about climate change are affected by short-term economic conditions and environmental quality anomalies. Consequently, households become more sensitive to environmental issues during a positive technology shock. Third, this mechanism is reinforced when the higher academic institutions become more efficient in their education activity, amplifying the beneficial effects of a technology shock on output and investments in environmental protection.

Marina ALBANESE, University of Naples Federico II, albanese@unina.it
Francesco BUSATO, University of Naples Parthenope,
francesco.busato@uniparthenope.it
Gianluigi CISCO, University of Naples Parthenope,
gianluigi.cisco@studenti.uniparthenope.it

PERIOD LIFE TABLES IN SUBURBAN AREAS: THE CASE OF THE ITALIAN MUNICIPALITY OF TARANTO

Stefano Cervellera, Carlo Cusatelli

1. Introduction

The city of Taranto has about 200 thousand inhabitants and extends for 250 km² strongly characterized by the presence of a vast industrial area that includes: an integral cycle steel mill (the largest in Europe); a large petrochemical refinery, active since 1967; a cement factory of national importance; two thermoelectric power plants; a relevant port area; two military arsenals; a NATO base; etc. Since the 1980s, together with the neighboring municipalities of Crispiano, Massafra, Montemesola and Statte, Taranto has been delimited and included among the areas at high risk of environmental crisis, and subsequently included among the top 14 sites of national interest (SIN) with high gravity environmental contamination, health risk and social alarm.

Several environmental, biomonitoring and also epidemiological studies conducted in the Taranto area have assessed industrial emissions and their impact on the territory, indicating strong air pollution originating mainly from the steel industry with the spread of particulate matter, heavy metals, polycyclic aromatic hydrocarbons and organ compounds halogenates, and showing an excess of mortality referred both to all causes and to neoplasms, in particular respiratory, pleural and bladder ones (Blangiardo and Rimoldi, 2013; Comba *et al*., 2012; Graziano *et al*., 2009; Mataloni *et al*., 2012; Michelozzi, 2012; Stafoggia *et al*., 2009; Zona *et al*., 2019).

Mortality analysis is always connected to comparative assessments and needs of a territorial and/or temporal nature, to verify how the phenomenon varies geographically and/or over time. Istat deals extensively with some important surveys and calculations, starting from the monthly and annual demographic ones, from which data on the number of the population residing in the municipalities, on the movement of the population, and also on the natural balance (through the number of births and deaths reported monthly in the D7A and D7B forms and annually in the P2 and P3 forms that all municipalities draw up). Furthermore, the D4 and D4 bis models (the latter for deaths within the first year of life, which since 2013 have replaced and integrated the old D5 and D5 bis models) show the first four causes of

death, certified by a ASL doctor, with subsequent WHO ICD-10 classification (Cervellera *et al*., 2014).

Using municipal registry sources, validated for more specific territorial disaggregation than the provincial Istat data, we calculated Taranto period life tables which represent an important tool for demographic analysis, structured by gender and neighborhood, in a decade (2010-2019) of strong socio-economic conflict, resulting in the well-known legal proceedings due to the situation of severe environmental pollution of the territory.

2. The period life tables

Period life tables represent an important tool for demographic analysis, for various reasons, to the point of being considered by many researchers to be the most complete logical-technical tool for the statistical analysis of mortality and its incidence by age and sex. They are based on a logical principle of description, in the form of tables, of the elimination by death of a generation[1] until the extinction of the last of the components, for which the main parameter of the study is the age-specific probability of death q_x: it expresses the risk that a person has of dying between the x-th birthday and the following birthday and it is an estimate of the age-specific death rate M_x.

The period life tables are obtained through different methods of detecting events: those of the first kind, *by generations* and those of the second kind, *by contemporaries*. The first type refers to the mortality of a contingent of individuals belonging to the same generation, i.e. born in the same year and followed over time until the death of all (generational analysis): given the extreme difficulty of following the contingent from the beginning and for the entire life course, one could fall back on a retrospective analysis which however would compromise the reliability of the results. To overcome these problems, reference is made to the second species period life tables, considering the set of contemporary individuals and therefore also of different ages (cross-sectional analysis). The methodology subsequently used is that of the tables for contemporaries, which is in any case the one generally used in demography for the study of the population: it would not be possible, here and for our purpose, to follow a generation of individuals up to death of the last.

Mainly, starting from the age-specific probability of death q_x, the period life tables determine some important biometric functions of the population aged x, such

[1] The initial generation of the born can be the real one, but also a theoretical one: in the latter case, a multiple of 10 (usually 10,000 or 100,000) is used as the initial value l_0, called the root.

as: the initial survivors l_x, the deaths d_x, the years lived L_x, up to the most important which is life expectancy $\dot{e}_x$ (Livi Bacci 1990, Preston et al. 2001). The period life tables are therefore a very useful tool in the comparative spatial and temporal analysis of the phenomenon in question.

3. The adopted methodology

We have elaborated abridged period life tables for suburban area such as neighborhoods, using an essentially different methodology from Istat one due to the fact that we do not re-elaborate the age-specific probabilities of death, which remain the real ones. This solution is valid as Taranto is demographically a large municipality, with about two thousand deaths every year (Tab. 1).

Table 1 – *Deaths in the districts of Taranto.*

Gender	District	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Female	Borgo-Città Vecchia	309	296	293	264	276	277	262	291	258	295
	Montegranaro-Salinella	227	217	243	216	252	236	239	237	285	267
	Paolo VI	53	45	43	48	56	66	66	52	71	73
	Talsano Lama S. Vito	125	145	124	128	171	161	143	179	133	164
	Tamburi	90	108	92	87	85	94	85	98	94	98
	Tre Carrare - Solito	218	222	215	230	210	227	226	251	251	245
Male	Borgo-Città Vecchia	236	258	246	230	238	270	221	235	229	241
	Montegranaro Salinella	195	225	211	211	225	204	221	214	233	216
	Paolo VI	54	51	72	55	67	66	74	68	73	98
	Talsano Lama S. Vito	132	169	159	164	158	167	166	172	200	165
	Tamburi	97	87	87	88	91	90	85	80	94	81
	Tre Carrare - Solito	185	208	189	200	216	206	209	188	181	228
Total		1,921	2,031	1,974	1,921	2,045	2,064	1,997	2,065	2,102	2,171

But since the number (D) of deaths is reduced in the suburban areas, we grouped data in two five-year periods: 2010-14 and 2015-19. And we also used age-classes with groups of s=5 years: $[x, x+s[$. In particular for the initial one, $q_{[0,\,5[}$, considering that in the first year of life, when mortality is generally higher, it results $q_0 \gg (q_1+q_2+q_3+q_4)$, we anyhow obtain a good estimate of life expectancy at birth: $e^0 \cong e^{[0,\,5[}$. Instead, the amplitude of the last class $[\text{int}(\omega_t/s)\cdot s, \omega_t]$ is variable, where ω_t is the highest age at death recorded in time-period t. This allows us to close the period life table at age ω_t and to determine the complete values of the last row of each table.

It differs from the methodology used by Istat which, based on a theoretical estimate built using a Kannisto model[2], closes all the tables at 125 years.

3.1. Determination of age-specific death rates

The age-specific death rate was determined with:

$$M_{[x,x+s[\,t} = \frac{D_{[x,x+s[\,t}}{\frac{1}{2}\left(P_{1.1.t}^{[x,x+s[} + P_{31.12.t}^{[x,x+s[}\right)} \tag{1}$$

for the two five-year periods t and for all age-groups with s=5 except for the class [int(ω_t/s)·s, ω_t] of amplitude ω_t–int(ω_t/s)·s.

3.2. Determination of the biometric functions

Once the age-specific death rates have been obtained, the age-specific probability of death can be determined, with a procedure similar to that of the age-specific death rates, using the Merrel and Reed method[3], getting

$$q_{[x,x+5[\,t} = 1 - e^{-M_{[x,x+5[\,t}\cdot\left(5+M_{[x,x+5[\,t}\right)} \tag{2}$$

and in particular for the last age group

$$q_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t} = 1 - e^{-M_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t}\cdot\left\{[\omega_t - \mathrm{int}(\omega_t/s)s]+M_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t}\right\}}. \tag{3}$$

This allows us to calculate also $l_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t}$, $L_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t}$ and the effective life expectancy for all classes, including the last: $\dot{e}_{[\mathrm{int}(\omega_t/s)s,\omega_t]\,t}$ (Tab. 2 and Tab. 3).

---

[2] Kannisto highlights how the rate of growth of the force of mortality decreases with advanced age and, therefore, no longer follows a linear form of Gompertz or an almost linear form of Gompertz-Makeham (Gompertz, 1825). Istat has been using it since 1996.

[3] The estimate of q_x proposed by Merrel and Reed is considered very efficient and is widely used in the international context, also used by Istat since 1992, so this use makes the results of the tables of the city of Taranto even more compatible and comparable with all the Istat period life tables.

Table 2 – *Period life table 2010-14.*

District	Age	q_x	l_x	d_x	L_x	e_x	q_x	l_x	d_x	L_x	e_x
		Female					Male				
Borgo - Città Vecchia	0-4	0.006	100000	568	498579	83.6	0.002	100000	221	499449	77.8
	5-9	0.001	99432	112	496878	79.0	0.001	99779	106	498631	73.0
	10-14	0.000	99319	0	496597	74.1	0.001	99673	103	498108	68.1
	15-19	0.000	99319	0	496597	69.1	0.000	99570	0	497851	63.1
	20-24	0.001	99319	94	496363	64.1	0.002	99570	177	497407	58.1
	25-29	0.001	99226	90	495903	59.2	0.002	99393	169	496541	53.2
	30-34	0.002	99136	164	495268	54.2	0.002	99224	238	495524	48.3
	35-39	0.003	98972	332	494029	49.3	0.007	98986	689	493206	43.4
	40-44	0.003	98640	315	492413	44.5	0.012	98297	1145	488620	38.7
	45-49	0.010	98325	996	489136	39.6	0.017	97151	1675	481568	34.1
	50-54	0.011	97329	1085	483932	35.0	0.022	95476	2081	472178	29.7
	55-59	0.012	96244	1138	478374	30.4	0.040	93395	3762	457570	25.3
	60-64	0.027	95106	2533	469196	25.7	0.078	89633	6971	430738	21.2
	65-69	0.041	92573	3788	453393	21.3	0.084	82662	6921	396009	17.8
	70-74	0.067	88784	5955	429036	17.1	0.119	75741	9045	356094	14.2
	75-79	0.117	82830	9689	389927	13.2	0.220	66696	14693	296748	10.8
	80-84	0.199	73141	14560	329304	9.6	0.324	52003	16839	217920	8.2
	85-89	0.426	58581	24962	230499	6.4	0.511	35165	17958	130927	5.9
	90-94	0.652	33619	21907	113325	4.2	0.620	17206	10668	59360	4.4
	95-ω	1.000	11711	11711	29278	2.5	1.000	6538	6538	16344	2.5
Montegranaro - Salinella	0-4	0.001	100000	131	499672	85.3	0.005	100000	489	498778	80.5
	5-9	0.000	99869	0	499344	80.4	0.000	99511	0	497556	75.9
	10-14	0.000	99869	0	499344	75.4	0.000	99511	0	497556	70.9
	15-19	0.000	99869	0	499344	70.4	0.002	99511	181	497103	65.9
	20-24	0.001	99869	81	499142	65.4	0.002	99330	240	496051	61.0
	25-29	0.001	99788	84	498730	60.4	0.003	99090	247	494833	56.2
	30-34	0.000	99704	0	498519	55.5	0.003	98843	249	493592	51.3
	35-39	0.001	99704	65	498358	50.5	0.004	98594	348	492099	46.4
	40-44	0.004	99639	351	497319	45.5	0.006	98246	587	489762	41.6
	45-49	0.008	99288	793	494460	40.7	0.008	97659	741	486443	36.8
	50-54	0.011	98496	1078	489783	36.0	0.018	96918	1725	480278	32.1
	55-59	0.007	97418	665	485425	31.3	0.028	95193	2618	469419	27.6
	60-64	0.021	96752	1987	478795	26.5	0.027	92575	2477	456683	23.3
	65-69	0.028	94766	2653	467196	22.0	0.065	90098	5864	435831	18.9
	70-74	0.055	92113	5106	447798	17.6	0.094	84234	7938	401327	15.0
	75-79	0.112	87006	9718	410738	13.5	0.171	76296	13058	348838	11.4
	80-84	0.200	77289	15466	347780	9.9	0.319	63239	20203	265685	8.2
	85-89	0.381	61823	23552	250236	6.7	0.475	43035	20449	164055	5.8
	90-94	0.640	38271	24500	130106	4.3	0.724	22587	16357	72041	3.9
	95-ω	1.000	13771	13771	34427	2.5	1.000	6230	6230	15575	2.5
Paolo VI	0-4	0.000	100000	0	500000	82.9	0.003	100000	345	499137	78.6
	5-9	0.000	100000	0	500000	77.9	0.002	99655	174	497839	73.8
	10-14	0.000	100000	0	500000	72.9	0.000	99481	0	497405	68.9
	15-19	0.000	100000	0	500000	67.9	0.000	99481	0	497405	63.9
	20-24	0.002	100000	174	499564	62.9	0.000	99481	0	497405	58.9
	25-29	0.002	99826	172	498698	58.0	0.004	99481	350	496530	53.9
	30-34	0.000	99653	0	498267	53.1	0.003	99131	309	494884	49.1
	35-39	0.000	99653	0	498267	48.1	0.001	98823	131	493784	44.3
	40-44	0.003	99653	249	497645	43.1	0.010	98691	988	490987	39.3
	45-49	0.008	99405	811	494996	38.2	0.015	97703	1452	484888	34.7
	50-54	0.011	98594	1061	490316	33.5	0.024	96252	2271	475580	30.2
	55-59	0.024	97533	2363	481755	28.8	0.041	93980	3807	460383	25.8
	60-64	0.026	95169	2521	469543	24.5	0.039	90173	3547	441999	21.8
	65-69	0.029	92648	2649	456618	20.1	0.086	86627	7444	414523	17.6
	70-74	0.080	89999	7195	432008	15.6	0.141	79183	11173	367982	14.0
	75-79	0.117	82804	9700	389770	11.8	0.199	68010	13526	306234	10.9
	80-84	0.337	73104	24631	303943	8.0	0.386	54484	21023	219860	8.0
	85-89	0.472	48473	22892	185135	5.8	0.411	33460	13762	132898	6.5
	90-94	0.761	25581	19456	79265	3.7	0.634	19699	12490	67270	4.3
	95-ω	1.000	6125	6125	15312	2.5	1.000	7209	7209	18023	2.5

Table 2 – *continued*.

District	Age	q_x	l_x	d_x	L_x	e_x	q_x	l_x	d_x	L_x	e_x
		Female					Male				
Talsano - Lama - S.Vito	0-4	0.002	100000	189	499527	84.7	0.002	100000	173	499566	80.8
	5-9	0.000	99811	0	499053	79.9	0.000	99827	0	499133	76.0
	10-14	0.000	99811	0	499053	74.9	0.000	99827	0	499133	71.0
	15-19	0.000	99811	0	499053	69.9	0.002	99827	154	498748	66.0
	20-24	0.001	99811	78	498857	64.9	0.002	99673	215	497827	61.1
	25-29	0.001	99732	78	498466	59.9	0.000	99458	0	497290	56.2
	30-34	0.001	99654	66	498106	55.0	0.002	99458	203	496784	51.2
	35-39	0.002	99588	216	497402	50.0	0.004	99255	445	495165	46.3
	40-44	0.005	99373	535	495526	45.1	0.005	98810	507	492785	41.5
	45-49	0.007	98838	739	492342	40.4	0.008	98303	785	489553	36.7
	50-54	0.006	98099	633	488913	35.6	0.014	97518	1345	484227	32.0
	55-59	0.012	97466	1197	484339	30.9	0.018	96173	1739	476516	27.4
	60-64	0.017	96269	1630	477272	26.2	0.039	94434	3697	462926	22.8
	65-69	0.037	94639	3474	464513	21.6	0.062	90737	5581	439732	18.7
	70-74	0.052	91166	4702	444073	17.3	0.106	85156	9035	403194	14.7
	75-79	0.126	86464	10887	405101	13.2	0.161	76121	12231	350031	11.2
	80-84	0.214	75577	16148	337515	9.7	0.355	63891	22686	262740	7.8
	85-89	0.398	59429	23652	238015	6.6	0.508	41205	20929	153703	5.8
	90-94	0.625	35777	22354	122999	4.4	0.679	20276	13766	66965	4.1
	95-ω	1.000	13423	13423	33557	2.5	1.000	6510	6510	16275	2.5
Tamburi	0-4	0.004	100000	402	498995	83.5	0.002	100000	208	499479	77.3
	5-9	0.000	99598	0	497991	78.8	0.000	99792	0	498958	72.5
	10-14	0.000	99598	0	497991	73.8	0.002	99792	189	498484	67.5
	15-19	0.000	99598	0	497991	68.8	0.000	99602	0	498011	62.6
	20-24	0.000	99598	0	497991	63.8	0.004	99602	369	497088	57.6
	25-29	0.003	99598	339	497143	58.8	0.004	99233	353	495283	52.8
	30-34	0.002	99259	184	495835	54.0	0.003	98880	339	493552	48.0
	35-39	0.002	99075	168	494955	49.1	0.002	98541	168	492285	43.2
	40-44	0.000	98907	0	494536	44.1	0.007	98373	659	490220	38.2
	45-49	0.006	98907	627	492969	39.1	0.020	97715	1950	483698	33.5
	50-54	0.003	98280	327	490585	34.4	0.018	95764	1713	474539	29.1
	55-59	0.017	97953	1627	485698	29.5	0.039	94051	3625	461193	24.6
	60-64	0.026	96326	2494	475395	24.9	0.070	90426	6327	436312	20.5
	65-69	0.065	93832	6139	453814	20.5	0.120	84099	10075	395306	16.8
	70-74	0.081	87694	7142	420614	16.8	0.137	74023	10135	344779	13.8
	75-79	0.108	80552	8705	380996	13.1	0.189	63888	12101	289189	10.5
	80-84	0.221	71847	15885	319521	9.4	0.395	51787	20433	207856	7.4
	85-89	0.430	55962	24054	219674	6.3	0.531	31355	16642	115170	5.6
	90-94	0.669	31908	21350	106162	4.2	0.661	14713	9731	49239	4.2
	95-ω	1.000	10557	10557	26393	2.5	1.000	4982	4982	12456	2.5
Tre Carrare - Solito	0-4	0.000	100000	0	500000	85.1	0.003	100000	276	499310	80.5
	5-9	0.000	100000	0	500000	80.1	0.000	99724	0	498620	75.7
	10-14	0.000	100000	0	500000	75.1	0.000	99724	0	498620	70.7
	15-19	0.000	100000	0	500000	70.1	0.002	99724	209	498097	65.7
	20-24	0.001	100000	92	499770	65.1	0.003	99515	271	496897	60.8
	25-29	0.000	99908	0	499539	60.2	0.003	99244	255	495583	56.0
	30-34	0.001	99908	88	499320	55.2	0.003	98989	345	494084	51.1
	35-39	0.002	99820	232	498521	50.2	0.007	98644	708	491450	46.3
	40-44	0.005	99588	496	496701	45.3	0.008	97936	792	487700	41.6
	45-49	0.006	99092	587	493993	40.5	0.013	97144	1247	482603	36.9
	50-54	0.013	98505	1241	489423	35.8	0.010	95897	930	477160	32.4
	55-59	0.013	97264	1261	483167	31.2	0.022	94967	2130	469510	27.7
	60-64	0.022	96003	2123	474709	26.6	0.037	92837	3411	455659	23.3
	65-69	0.032	93880	3022	461847	22.1	0.064	89426	5760	432731	19.0
	70-74	0.054	90858	4898	442048	17.8	0.101	83666	8430	397256	15.2
	75-79	0.098	85961	8420	408752	13.6	0.161	75236	12090	345955	11.6
	80-84	0.211	77540	16356	346811	9.9	0.274	63146	17304	272469	8.4
	85-89	0.379	61184	23190	247945	6.8	0.530	45842	24307	168441	5.6
	90-94	0.610	37994	23184	132009	4.4	0.698	21535	15027	70106	4.0
	95-ω	1.000	14810	14810	37024	2.5	1.000	6508	6508	16269	2.5

Table 3 – *Period life table 2015-19.*

District	Age	q_x	l_x	d_x	L_x	e_x	q_x	l_x	d_x	L_x	e_x
		Female					Male				
Borgo - Città Vecchia	0-4	0.003	100000	269	499329	84.0	0.011	100000	1089	497278	78.0
	5-9	0.000	99731	0	498657	79.2	0.001	98911	115	494270	73.9
	10-14	0.000	99731	0	498657	74.2	0.000	98797	0	493984	68.9
	15-19	0.000	99731	0	498657	69.2	0.004	98797	391	493005	63.9
	20-24	0.002	99731	205	498145	64.2	0.003	98405	273	491344	59.2
	25-29	0.004	99526	387	496665	59.4	0.003	98132	266	489998	54.3
	30-34	0.001	99139	97	495455	54.6	0.002	97867	178	488889	49.5
	35-39	0.003	99043	258	494569	49.6	0.004	97689	410	487419	44.6
	40-44	0.005	98785	543	492568	44.8	0.005	97279	493	485160	39.8
	45-49	0.011	98242	1107	488443	40.0	0.012	96785	1206	480912	34.9
	50-54	0.012	97135	1208	482655	35.4	0.026	95579	2514	471612	30.4
	55-59	0.012	95927	1125	476821	30.8	0.039	93066	3657	456186	26.1
	60-64	0.035	94801	3357	465616	26.2	0.047	89409	4160	436643	22.1
	65-69	0.040	91445	3623	448166	22.0	0.079	85249	6729	409421	18.0
	70-74	0.066	87822	5782	424654	17.8	0.116	78520	9082	369895	14.4
	75-79	0.111	82040	9114	387413	13.9	0.209	69438	14489	310968	10.9
	80-84	0.176	72926	12799	332630	10.3	0.326	54949	17929	229923	8.1
	85-89	0.348	60126	20913	248350	7.0	0.513	37020	18991	137623	5.8
	90-94	0.615	39214	24111	135791	4.4	0.628	18029	11326	61830	4.4
	95-ω	1.000	15103	15103	37757	2.5	1.000	6703	6703	16757	2.5
Montegranaro - Salinella	0-4	0.004	100000	431	498923	85.5	0.001	100000	145	499639	82.0
	5-9	0.000	99569	0	497846	80.8	0.000	99855	0	499277	77.1
	10-14	0.001	99569	105	497583	75.8	0.000	99855	0	499277	72.1
	15-19	0.001	99464	97	497078	70.9	0.001	99855	98	499033	67.1
	20-24	0.000	99367	0	496836	66.0	0.001	99758	92	498560	62.2
	25-29	0.000	99367	0	496836	61.0	0.001	99666	93	498098	57.2
	30-34	0.001	99367	96	496595	56.0	0.002	99573	193	497383	52.3
	35-39	0.003	99271	333	495523	51.0	0.003	99380	268	496231	47.4
	40-44	0.006	98938	584	493229	46.2	0.005	99112	509	494289	42.5
	45-49	0.004	98354	412	490737	41.5	0.011	98603	1093	490283	37.7
	50-54	0.009	97941	907	487439	36.6	0.010	97510	1012	485019	33.1
	55-59	0.008	97034	737	483330	32.0	0.024	96498	2325	476675	28.4
	60-64	0.024	96298	2315	475701	27.2	0.036	94172	3434	462275	24.0
	65-69	0.027	93983	2543	463556	22.8	0.058	90738	5283	440482	19.9
	70-74	0.053	91440	4853	445067	18.3	0.106	85455	9061	404623	15.9
	75-79	0.088	86587	7617	413891	14.2	0.137	76394	10456	355832	12.5
	80-84	0.165	78970	12999	362351	10.4	0.231	65939	15229	291620	9.1
	85-89	0.359	65971	23662	270701	6.9	0.470	50709	23831	193969	6.1
	90-94	0.622	42309	26322	145740	4.4	0.647	26878	17386	90926	4.3
	95-ω	1.000	15987	15987	39967	2.5	1.000	9492	9492	23731	2.5
Paolo VI	0-4	0.005	100000	461	498848	82.4	0.005	100000	453	498867	78.1
	5-9	0.000	99539	0	497697	77.8	0.000	99547	0	497734	73.5
	10-14	0.000	99539	0	497697	72.8	0.000	99547	0	497734	68.5
	15-19	0.002	99539	195	497209	67.8	0.005	99547	498	496490	63.5
	20-24	0.000	99344	0	496721	62.9	0.002	99049	170	494822	58.8
	25-29	0.000	99344	0	496721	57.9	0.002	98880	184	493938	53.9
	30-34	0.002	99344	182	496267	52.9	0.004	98695	383	492520	49.0
	35-39	0.002	99163	166	495399	48.0	0.007	98313	651	489936	44.1
	40-44	0.009	98997	872	492805	43.1	0.006	97662	552	486930	39.4
	45-49	0.008	98125	757	488733	38.4	0.009	97110	879	483352	34.6
	50-54	0.015	97368	1493	483107	33.7	0.031	96231	2953	473771	29.9
	55-59	0.022	95875	2078	474180	29.2	0.030	93278	2812	459357	25.8
	60-64	0.025	93797	2341	463133	24.8	0.050	90465	4493	441093	21.5
	65-69	0.037	91456	3412	448752	20.4	0.101	85972	8700	408109	17.5
	70-74	0.087	88045	7665	421060	16.1	0.110	77272	8512	365080	14.2
	75-79	0.146	80379	11761	372493	12.4	0.164	68760	11273	315618	10.6
	80-84	0.228	68618	15626	304025	9.0	0.344	57487	19747	238069	7.2
	85-89	0.421	52992	22310	209185	6.0	0.619	37740	23365	130288	4.7
	90-94	0.799	30682	24502	92153	3.5	0.838	14375	12050	41750	3.3
	95-ω	1.000	6179	6179	15449	2.5	1.000	2325	2325	5813	2.5

Table 3 – *continued*.

District	Age	q_x	l_x	d_x	L_x	e_x	q_x	l_x	d_x	L_x	e_x
		Female					Male				
Talsano - Lama - S.Vito	0-4	0.002	100000	233	499418	85.7	0.007	100000	662	498345	81.6
	5-9	0.001	99767	94	498601	80.9	0.000	99338	0	496690	77.1
	10-14	0.001	99673	90	498141	76.0	0.001	99338	84	496481	72.1
	15-19	0.000	99583	0	497916	71.1	0.000	99254	0	496271	67.2
	20-24	0.000	99583	0	497916	66.1	0.001	99254	79	496075	62.2
	25-29	0.000	99583	0	497916	61.1	0.001	99176	79	495679	57.2
	30-34	0.002	99583	162	497510	56.1	0.003	99096	342	494626	52.3
	35-39	0.001	99421	66	496938	51.2	0.001	98754	139	493423	47.4
	40-44	0.005	99355	501	495519	46.2	0.003	98615	285	492362	42.5
	45-49	0.006	98853	610	492741	41.4	0.008	98330	820	489598	37.6
	50-54	0.008	98243	753	489333	36.7	0.009	97510	923	485241	32.9
	55-59	0.006	97490	587	485983	31.9	0.019	96587	1865	478271	28.2
	60-64	0.012	96903	1149	481642	27.1	0.034	94722	3212	465578	23.7
	65-69	0.032	95754	3026	471204	22.4	0.056	91509	5151	444668	19.5
	70-74	0.046	92728	4230	453065	18.0	0.099	86358	8511	410513	15.5
	75-79	0.091	88498	8088	422270	13.8	0.151	77847	11719	359938	11.9
	80-84	0.197	80410	15857	362407	9.9	0.276	66128	18230	285066	8.6
	85-89	0.380	64553	24521	261462	6.7	0.477	47898	22871	182313	5.9
	90-94	0.634	40032	25365	136747	4.3	0.710	25027	17779	80687	3.9
	95-ω	1.000	14667	14667	36668	2.5	1.000	7248	7248	18119	2.5
Tamburi	0-4	0.003	100000	271	499323	84.0	0.000	100000	0	500000	78.0
	5-9	0.000	99729	0	498646	79.2	0.000	100000	0	500000	73.0
	10-14	0.000	99729	0	498646	74.2	0.002	100000	190	499524	68.0
	15-19	0.002	99729	211	498119	69.2	0.000	99810	0	499048	63.1
	20-24	0.000	99518	0	497591	64.3	0.006	99810	584	497589	58.1
	25-29	0.000	99518	0	497591	59.3	0.006	99226	594	494644	53.4
	30-34	0.000	99518	0	497591	54.3	0.006	98632	606	491643	48.8
	35-39	0.000	99518	0	497591	49.3	0.006	98026	555	488741	44.0
	40-44	0.004	99518	350	496716	44.3	0.012	97471	1215	484318	39.3
	45-49	0.011	99168	1067	493175	39.5	0.014	96256	1339	477933	34.7
	50-54	0.006	98101	625	488945	34.9	0.021	94917	1963	469679	30.2
	55-59	0.018	97477	1796	482894	30.1	0.035	92954	3217	456730	25.8
	60-64	0.022	95681	2136	473066	25.6	0.062	89737	5548	434818	21.6
	65-69	0.037	93545	3501	458973	21.1	0.076	84190	6382	404994	17.9
	70-74	0.058	90044	5253	437088	16.8	0.110	77808	8562	367635	14.1
	75-79	0.155	84791	13136	391115	12.7	0.173	69246	11966	316316	10.6
	80-84	0.223	71655	15975	318338	9.6	0.422	57280	24176	225962	7.3
	85-89	0.404	55680	22497	222158	6.6	0.525	33105	17366	122110	5.7
	90-94	0.611	33183	20283	115210	4.4	0.637	15739	10028	53627	4.3
	95-ω	1.000	12901	12901	32252	2.5	1.000	5711	5711	14279	2.5
Tre Carrare - Solito	0-4	0.007	100000	681	498297	84.9	0.003	100000	328	499181	81.9
	5-9	0.000	99319	0	496594	80.5	0.000	99672	0	498361	77.2
	10-14	0.001	99319	131	496265	75.5	0.000	99672	0	498361	72.2
	15-19	0.001	99187	119	495640	70.6	0.000	99672	0	498361	67.2
	20-24	0.001	99069	107	495077	65.7	0.001	99672	108	498092	62.2
	25-29	0.002	98962	194	494326	60.7	0.000	99564	0	497823	57.3
	30-34	0.000	98768	0	493842	55.9	0.004	99564	389	496849	52.3
	35-39	0.001	98768	95	493604	50.9	0.003	99175	295	495138	47.5
	40-44	0.005	98673	467	492199	45.9	0.005	98880	502	493146	42.6
	45-49	0.004	98206	419	489985	41.1	0.009	98378	894	489657	37.8
	50-54	0.010	97788	959	486541	36.3	0.012	97484	1137	484579	33.1
	55-59	0.014	96829	1394	480659	31.6	0.014	96347	1320	478436	28.5
	60-64	0.016	95435	1513	473392	27.0	0.046	95027	4370	464212	23.8
	65-69	0.034	93922	3166	461695	22.4	0.051	90658	4579	441840	19.9
	70-74	0.048	90756	4385	442817	18.1	0.095	86079	8206	409877	15.8
	75-79	0.108	86371	9307	408585	13.9	0.149	77872	11625	360300	12.2
	80-84	0.177	77063	13609	351293	10.3	0.249	66248	16519	289941	8.9
	85-89	0.355	63454	22503	261012	7.0	0.459	49729	22834	191560	6.0
	90-94	0.613	40951	25089	142033	4.4	0.700	26895	18840	87375	4.0
	95-ω	1.000	15862	15862	39656	2.5	1.000	8055	8055	20138	2.5

4 Analysis of the results

The scientific world and institutions have paid a great deal of attention to health conditions and to the analysis of the morbidity of diseases related to pollution and its related mortality in the Ionian area and, in particular, in the municipality of Taranto. The analysis with period life tables can assume great importance in ecological studies on the health of citizens, precisely when done in smaller areas than the Istat provincial aggregate data: our analysis allows to determine the period life tables for suburban area such as the 6 districts of the City of Taranto.

The districts of Taranto show a strong inequality of mortality detected in the biometric functions. The northern districts (Paolo VI, Tamburi and Borgo-Città Vecchia) have a much shorter life expectancy at birth in the entire 2010-19 period than the southern districts (Tre Carrare-Solito, Montegranaro-Salinella and Talsano-Lama-S.Vito). As for women (Fig. 1), Paul VI represents the most problematic district, with the lowest expectation already in the five-year period 2010-14 (82.9 years) which even decreases in 2015-19 (82.4 years) demonstrating the presence of adverse public health situations. For men (Fig. 2), however, the Tamburi district is the most problematic in the two periods (77.3 and 78.0 years). The range of variation in life expectancy at birth for women goes from 2.35 years in 2010-14 to 3.32 years in 2015-19, while for men it goes from 3.49 to 3.98 years. All indicates how in the city of Taranto there is an increase in internal territorial health inequality for citizens residing in the north (therefore closer to the notorious industrial area) compared to those living in the south districts.

This evidence available for the suburban areas of Taranto shows a health and environmental picture with multiple critical aspects, claiming for primary prevention interventions and continuous monitoring of the health status of the population. This would require a feasible tool to rapidly perform a preliminary assessment of population health status every year, preferably based on official data provided by the same municipality and able to allow epidemiological comparison between different years, age groups, gender and neighborhoods.

Furthermore, it comes out how the analysis at suburban level shows that for medium-large cities it is necessary to analyze the mortality at the district level, given that the higher mortality in the entire municipality is essentially attributable to a significant excess of mortality among residents in some neighborhoods. This could be particularly true in cities that suffer from severe environmental pressures that are not widespread but located in specific areas of the city such as the large steel plant in the north-western part of Taranto.

Figure 1 – *Female life expectancy at birth in the districts of Taranto.*

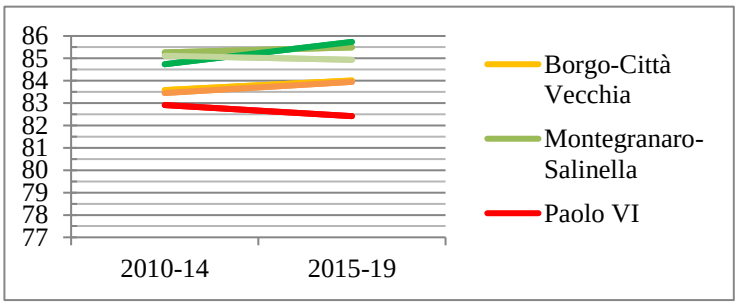
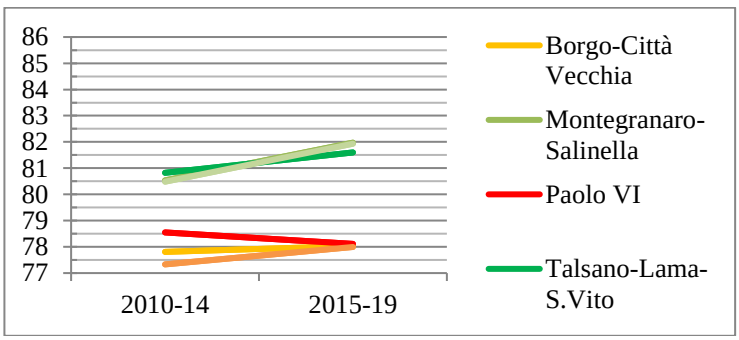


Figure 2 – *Male life expectancy at birth in the districts of Taranto.*



5 Final remarks

In cities suffering from heavy environmental pressure or pollution, it is extremely important to rapidly access municipal demographics that can be used as indicators of population health status. Among those, mortality rates represent the most reliable data as they are officially retained and available to municipality with high level of details, thus allowing epidemiological comparison between different neighborhoods of the city across several years.

Our study was aimed at validating and propose as universally applicable approach the use of municipal demographics as first-line tool to rapidly assess population health and drive health policies or urban planning in cities characterized by heavy environmental pressure. The case study of Taranto has been chosen due to the presence of the biggest European steel plant since 1960s resulting in heavy burden on environment and population health.

References

BLANGIARDO G.C., RIMOLDI S. 2013. Vivere (e morire) a Taranto, *Statistica & Società*, SIS, Vol. 1, No. 3, pp. 16-18.

CERVELLERA S., CUSATELLI C., SALINAS U. 2014. Costruzione ed analisi delle tavole di mortalità a Taranto dal 2003 al 2013 e comparazione tra fonti di dati. In URICCHIO A. (Ed.) *L'emergenza ambientale a Taranto: le risposte del mondo scientifico e le attività del Polo "Magna Grecia". Collana del Dipartimento Jonico, Università degli Studi di Bari*, Vol. 12, Bari: Cacucci, pp. 375-396.

COMBA P., PIRASTU R., CONTI S., DE SANTIS M., IAVARONE I., MARSILI G., MINCUZZI A., MINELLI G., MANNO V., MINERBA S., MUSMECI L., RASHID I., SOGGIU E., ZONA A. 2012. Environment and health in Taranto, southern Italy: epidemiological studies and public health recommendations, *Epidemiologia & Prevenzione*, Vol. 36, No. 6, pp. 305-320.

GOMPERTZ B. 1825. XXIV. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies, *Philosophical Transactions of the Royal Society of London*, Vol. 115, pp. 513-583.

GRAZIANO G., BILANCIA M., BISCEGLIA L., NICHILO G., POLLICE A., ASSENNATO G. 2009. Statistical Analysis of the Incidence of Some Cancers in the Province of Taranto, *Epidemiologia & Prevenzione*, Vol. 33, No. 1-2, pp. 37-44.

LIVI BACCI M. 1990. *Introduzione alla demografia*. Torino: Loescher.

MATALONI F., STAFOGGIA M., ALESSANDRINI E., TRIASSI M., BIGGERI A., FORASTIERE F. 2012. A Cohort Study on Mortality and Morbidity in the Area of Taranto, Southern Italy, *Epidemiologia & Prevenzione*, Vol. 36, No. 5, pp. 237-252.

MICHELOZZI P. 2012. Disastro Ambientale a Taranto: Il Ruolo dell'Epidemiologia, *Epidemiologia e Prevenzione*, Vol. 36, No. 5, pp. 231-233.

PRESTON S.H., HEUVELINE P., GUILLOT M. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishing.

STAFOGGIA M. et al. 2009. Inquinamento Atmosferico e Mortalità In Dieci Città Italiane. Risultati Del Progetto Epiair, *Epidemiologia & Prevenzione*, Vol. 33, No. 6, Suppl. 1, pp. 65-76.

ZONA A. et al. 2019. SENTIERI - Studio Epidemiologico Nazionale dei Territori e degli Insediamenti Esposti a Rischio da Inquinamento. Quinto Rapporto, *Epidemiologia & Prevenzione*, Vol. 43, No. 2-3, Suppl. 1, pp. 1-208.

SUMMARY

Period life tables in suburban areas: the case of the Italian municipality of Taranto

Using municipal registry sources, validated for more specific territorial disaggregation than the provincial Istat data, we obtained Taranto period life tables which represent an important tool for demographic analysis, structured by gender and neighborhood, in a decade (2010-2019) of strong socio-economic conflict, resulting in the well-known legal proceedings due to the situation of severe environmental pollution of the territory. The use of municipality data can be considered a methodological approach that allows a timely, reliable and costless first-line assessment system of the population status and drive urban planning policies in cities suffering from heavy environmental pressures: from the period life tables in suburban areas, a situation of strong inequality emerges, as Borgo-Città Vecchia, Paolo VI and Tamburi districts are united by the levels of life expectancy very detached from the others, as if they were different cities. The epidemiological and environmental evidence available for the Taranto area shows various critical aspects, claiming for primary prevention interventions and continuous monitoring of the health status of the population which can provide important indications for urban planning policies.

Stefano CERVELLERA, Municipality of Taranto, s.cervellera@comune.taranto.it
Carlo CUSATELLI, University of Bari "Aldo Moro", carlo.cusatelli@uniba.it

WHY DO BOYS PERFORM WORSE THAN GIRLS IN READING LITERACY? EVIDENCES FROM PISA SURVEY 2018

Giovanna Di Castro, Valentina Ferri[1]

1 Introduction

The reading gender gap is little studied and explored even though it is much higher than math gender gap, on which instead there is a wider literature and greater political attention, as such a gap affect the quality of women's education and on the choice of their career paths (Mostafa, 2019). The boys' underachievement in reading and lower reading engagement relative is typical of all countries that participate in the OECD's Programme for International Student Assessment (PISA) tests, since the first assessment in 2000 (OECD, 2019). The importance of increasing the competence and involvement of boys in reading literacy has a dual purpose, not only achieving greater equity between girls and boys, but also improve overall the basic level of reading skills regardless of the field of study or professional that students will choose. Indeed, reading literacy is an increasingly important element in a society in a rapidly changing, in which both the quantity and variety of written information are increasing and in which the ability to use, communicate and written information is critical to solving complex problems (Binkley *et al.*, 2012).

Reading literacy constitutes not only a specific subject area, but a prerequisite for understanding and communication for all educational and as well as social areas.

In the end, in the opportunity that better prepared male students could have in the choice of tertiary studies even in humanistic subjects where the male presence is small or almost non-existent.

The present study analyses the main factors behind the differences in test scores on reading literacy in Italy between male and female 15 years old students performing the Oaxaca-Blinder decomposition method (Blinder, 1973) on the data from the OECD PISA 2018 survey, in which reading was the main subject assessed.

[1] Le opinioni espresse in questo lavoro impegnano la responsabilità delle autrici e non necessariamente riflettono la posizione dell'Istituto Nazionale per l'Analisi delle Politiche Pubbliche (INAPP). Pur essendo frutto della collaborazione tra le autrici, i paragrafi 2, 3 e 4 sono da attribuire a Giovanna Di Castro e i paragrafi 1, 5 e 6 a Valentina Ferri.

The study wants to examine the weight of variables that measure the quality of the school environment, the socio-cultural context, the quality of teaching, and other non-cognitive elements that typically influence the academic success of adolescent students. The results of this study will be useful to offer indications for educational policies and practices, highlighting both those factors that feed gender inequality in boys, and those factors that in general are more related to the acquisition of reading literacy, a fundamental skill for full participation in school and social life, for which Italy is still below the international OECD average.

2 Descriptive analysis

PISA is a three-year survey that measures 15-year-old students' skills in reading, math and science. In each cycle of PISA, a subject is tested in detail, with a larger and more thorough number of items in the test.

The main subject in 2018 was reading literacy, with math and science as minor assessment areas, and saw the participation of 79 countries and education systems (Fig. 1). The PISA 2018 framework conceptualises reading as an activity where the reader interacts with both the text that he or she reads and with the tasks that he or she wants to accomplish during or after reading the text.

To be as complete as possible, the assessment covers different types of texts and tasks over a range of difficulty levels, and different types of cognitive processes

"*Reading literacy is understanding, using, evaluating, reflecting on and engaging with texts in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society*" (PISA, 2018)[2].

In PISA, the purpose is to respond to questions about texts in order to provide evidence of their level of reading literacy.

"*The reading process is assumed to be influenced by different factors related to the reader (e.g. motivation, disposition, and experience), the text (e.g., different text formats or sources), and the tasks or items (e.g., item difficulty)*" (Khorramdel *et al.*, 2020)

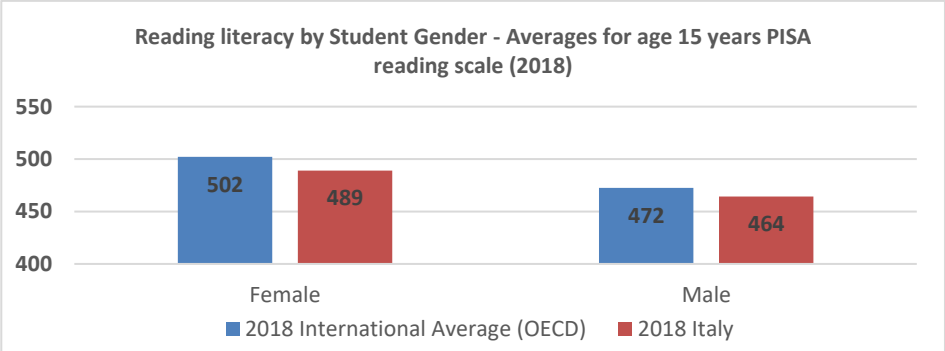
[2] Definition of reading literacy OECD (2019), PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris.

Figure 1 – *The countries participating in PISA survey*.

Source: OECD, 2019. PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris.

Analysing reading competence by student gender, we observe that the average scores of Italian women are 13 points lower than the OECD average scores. The male average is 8 points lower than the OECD average (Fig. 2).

Figure 2 – *Reading literacy by student gender*.

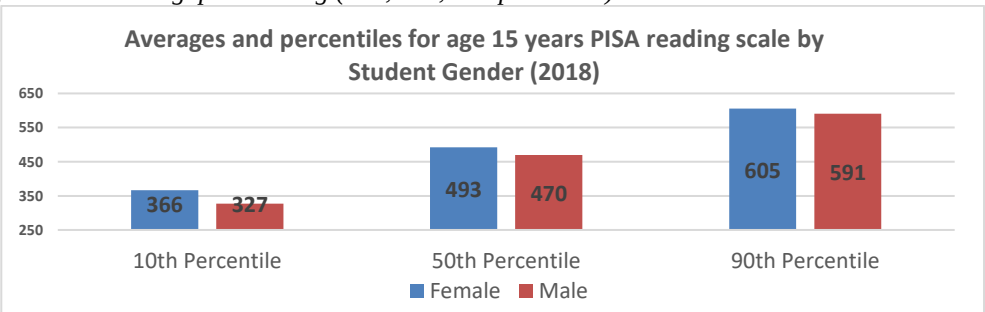


Source: Authors' elaborations on OECD-PISA 2018 data.

As for the gender gap, we observe that in Italy female have 23 points more than the male average, a difference far from small since it has been estimated that 25/30 points of difference on the PISA scale correspond to about 1 year of learning (Woessmann, 2016). That gap widens among with lower scores students, and at the

10th percentile we observe 39 points of difference while at the 90th percentile is 14 points (Fig. 3).

Figure 3 – *Gender gap in reading (10th, 50th, 90th percentile).*



Source: Authors' elaborations on OECD-PISA 2018 data.

3 Literature

Although reading literacy can be considered fundamental for learning, the difference performance between men and women has not been analyzed much (Linnakylä and Malin, 2000). A large literature, on the other hand, focuses on the analysis of the implications of the gender gap in mathematics in terms of under representation for women in math-related fields or gender inequalities in the labor market (Card and Payne, 2021). Instead, the determining factors and educational and professional implications of boys' poor reading performance are less clear.

The following studies regard the state of art of research concerning reading literacy in OECD countries.

Alvarado (2017) using the Blinder-Oaxaca decomposition on gender score gap of Colombian students in the math and reading components of the PISA 2012 test, finds that observable characteristics account for 34% in reading. This effect is due mainly to individual factors and school characteristics.

Munir *et al.* (2018) look at reading scores for all countries included in the OECD's PISA 2012 test and through decomposition methods these score differences at different percentiles of the distribution. The analysis shows that girls have a larger advantage in reading over boys. This advantage is particularly large for low-achieving individuals. Over the distribution of talent, boys' scores increase more than girls at the highest percentiles – they find a smaller reading advantage for girls.

Brozo *et al.* (2014) provide a summary of major gender differences in PISA 2009 along with relevant trends since 2000, analysing five countries: Finland, Korea, Germany, Ireland, United States. The authors' findings add further evidence to

support poor overall performance of boys and less involvement in reading than girls. They conclude with some policy implications.

Some authors have also investigated the relationship between some personality traits and performance, as well as the role played by motivational and attitudinal factors and the scores in financial literacy (another subject evaluated in the PISA tests), to explain part of the gender gap in favor of male boys in Italy, the only country where the score is significantly higher than that for girls (Longobardi, *et al*., 2018).

Some literature approaches argue that how students spend their time can affect performance, and like many other areas in adolescence, the use of free time differs widely between boys and girls (McHale *et al.*, 2001; Downey and Yuan, 2005). Boys spend more time playing video games than girls and less time reading for fun, especially complex texts, such as narrative texts (Weis and Cerankosky, 2010). Reading proficiency is very important because is the basis for other skills; when students do not read well, their performance in other school subjects is also affected (OECD, 2015).

Borgonovi (2016) suggests that the lower performance of boys in reading at the age of 15 is widespread in all the countries. She suggests an explanation linked to the habit of the boys that play video games every day more likely than girls.

The poor performance of boys in reading is lower when the assessment is provided on a computer. Boys' computer-based advantage is associated with gender differences in video games.

At the best of our knowledge, despite the importance of the reading gender gap for Italy, empirical evidence is very limited. Moreover, we observe that women perform better in reading in almost all the countries included in the PISA survey. This study allow to investigate the determinants of these differences between man and women in reading literacy that are not yet investigated in literature and in particular with regard to Italy.

4 Methodology and variables

In order to estimate the amount of the differential between male and female average wages, we have applied the decomposition of Oaxaca and Blinder (Oaxaca, 1973, Blinder, 1973). Through this method we distinguish which part is due to differences in characteristics included in model estimations and which part is due to gender "discrimination".

We estimate the twofold decomposition that divides the reading scores differences in two components: explained and unexplained.

The first component represents the explained effect (E) due to differences in the predictors; the second component (U) is the contribution of discrimination.

$$R = Q + U \quad (1)$$
$$Q = \{E(X_A) - E(X_B)\}'\beta^* \quad (2)$$
$$U = E(X_A)'(\beta_A - \beta^*) + E(X_B)'(\beta^* - \beta_B) \quad (3)$$

The equation (1) represents the twofold decomposition. The equation 2 indicates the gap in the outcome means that can be attributed to the characteristics of the groups A (male) and B (female). The equation 3 is the unexplained part. We perform the decomposition from the viewpoint of female group. Group differences in predictors are weighted by the coefficients of female's group in order to calculate the endowments effect. The variables included in our model are described in Table 1.

Table 1 – *Included variables in the Oaxaca Blinder decomposition*.

Reading score – the mean score achieved by students in reading literacy
Emosups – The index of parents emotional support
exp_par – Importance for decision about future occupation: my parents o guardians expectations about my occupation Very important (1) important (1) not important (0) somewhts important(0)
Disclima - Index of disciplinary climate
Istitut1 – general
Istitut2 – pre-vocational
Istitut3 – vocational
Age – age from 15 years and 3 completed months to 16 years and 4 completed months (14 possible options: ex. 15 years and 3 completed months; 15 years and 4 completed months; 15 years and 5 completed months)
Try_Job - Trying hard at school will help me get a good job (dummy 0/1)
Try_Coll - Trying hard at school will help me get into a good college (dummy 0/1)
Clear Ideas - what do you think you will be doing 5 years from now I will be studying because the occupation i want requires a study degree (1) 0 otherwise
Escs – Pisa Index of economic, social and cultural status
mean_escs – Mean for school of Pisa Index of economic, social and cultural status
Misced - Mother's education
Fisced - Father's education
Cultposs – Index Cultural Possession of the family
Joyread - The index of enjoyment of reading activities
Joyreadp - The index of enjoyment of reading activities of parents
Eudmonia -index of meaning in life
Dirins - The index of teacher-directed instruction
Perfeed - The index of teacher feedback
Lmins - Learning time reading minute
Teachint -index of teacher enthusiasm
Teachsup - The index of teacher support

Table 1 – *continued.*

AbleDiffTest- Agree: I am able to understand difficult test. Agree = 1; Disagree =0
NoDiffTest - I have always had difficulty with reading Agree = 0; Disagree =1
ParentHomework- how often: Help my child with his/her reading and writing homework (Never or hardly ever= 0; Once or twice a year= 0; Once or twice a month= 0; Once or twice a week= 1; Every day or almost every day= 1)
Videogames - Agree: I like to meet friends and play computer and video games with them; Disagree = 0 Agree = 1
UsePCHome - using computers at home (dummy 0/1)
Internet - Available for you to use at home: Internet connection Yes and I use it(1) – yes but I don't use it (0)|No (0)
100Pages- how many pages was the longest piece of text you had to read this year for reading lessons? more than 100 = 1; less than 100 = 0

5 Results

Estimates indicate that the gender gap in reading test in Italy averages around 22 points (Tab. 2). Gender difference associated with observable characteristics contributes to explain female positive performance (around 28 points). The observed characteristics contribute to increasing the gender gap in favor of girl, while the gap's unexplainable portion is insignificant as a whole.

Table 2 – *Oaxaca Blinder decomposition.*

	Reading score
Group_1 (F)	511.4690***
	[3.3141]
group_2 (M)	489.5794***
	[3.9109]
Difference	21.8895***
	[3.6880]
Explained	27.6850***
	[2.8013]
unexplained	-57.955
	[3.5988]

Source: Authors' elaborations on OECD-PISA 2018.

The variables that contribute to greater reading performance of girls in the explained part are described below.

Parental expectations, emotional support and help from parents with reading homework have a significant role in the increase in the differential between girls and boys (exp_par, emosups, parent_homework).

A better disciplinary climate (variable disclima) in reading lessons perceived by the student (e.g., a silent lesson, less chaotic atmosphere) increases the differential in favor of girls.

Table 3 – *Explained Oaxaca Blinder decomposition*.

Explained	
Individuals characteristics	
try_Job	-0.3571*
	[0.2031]
try_Coll	0.8338**
	[0.3846]
Eudmonia	1.5083***
	[0.3646]
ableDiffTest	1.3894***
	[0.4493]
noDiffTest	1.6788***
	[0.5247]
Joyread	5.9714***
	[1.4941]
clear Ideas	2.2418***
	[0.5979]
Habits	
Videogames	2.5929***
	[0.9759]
Internet	1.2798**
	[0.5500]
usePcHome	-0.6161*
	[0.3161]
Parents	
emosups	1.1511***
	[0.3696]
exp_par	0.9735***
	[0.3657]
parentHomework	2.1665***
	[0.6488]
School context	
mean_escs	3.2669***
	[1.1690]
dirins	0.6980**
	[0.2785]
disclima	1.1363**
	[0.5266]

Source: Authors' elaborations on OECD-PISA 2018.

Explained variables (Tab. 3) show that girls have significant advantages over boys in their enjoyment of reading and in their perception of reading skills (joyread,

abledifftest, nodifftest), while it is interesting to note that the enjoyment of reading index found in parents has no significant effect.

Having a clear meaning in life (index greater meaning in life) also favors girls (eudmonia).

Among the variables relating to teachers, only the perception of receiving very clear instructions and objectives from them during lessons (dirins) is significant, increasing the differential, while the other variables relating to teaching practices previously illustrated do not seem to affect reading performance.

Motivation and engagement aimed at a specific goal in terms of preparation for university (already clear in mind at the age of 15) increase the difference in reading performance in favor of girls, where we assume that girls are more likely to choose paths and jobs not related to mathematical-scientific skills, and therefore are a priori focused more on reading literacy subjects.

Finally, also the quality or socio-economic level of the school environment contributes on average to significantly increase the reading gender gap. On the contrary, the socio-economic level of the family has no effect on the variables

Among the factors that instead play a role in narrowing the gender gap between female and male students, we find scholastic commitment aimed at obtaining a good job a good job (instead of a "good college"), and also to have read during the 'year, in school practice, a text at least 100 pages long (or more) and no shorter.

Although the overall significance is not observed in the inexplicable part, among the variables analyzed, often playing video games with friends (video games) seems to strongly shorten the girl-boy gap (-13 points) as an inexplicable component (Tab.4). The perception of having no difficulty in reading (NodiffTest) also contributes to significantly reducing the coefficient effect (-17 points).

Receiving feedback from the teacher (perfeed) on one's academic performance in the reading lessons increases the unexplained part of the differential.

Table 4 – *Unexplainable factors - Oaxaca Blinder decomposition results.*

Unexplained	
Dirins	-0.9273*
	[0.4696]
Perfeed	0.4623*
	[0.2643]
NoDiffTest	-17.0816**
	[7.7644]
Videogames	-13.3411***
	[3.3170]

Source: Authors' elaborations on OECD-PISA 2018.

6 Conclusions

Despite his social and pedagogical importance, the gender gap in reading has been poorly studied in the literature, and educational outcomes have usually focused on the underachievement of girls.

Our results on the analysis of the reading gender gap in the OCED PISA 2018 assessment, relative to Italy, suggest that observable variables account for most of the gap. The gender gap in reading is mostly attributable to explained component, while the variables related to the discrimination effect are those related to different habits between women and men. In particular, consistent with the literature, it seems that the habit of spending time playing videogames has a lot of relevance in the discrimination effect.

Therefore, the analysis highlights the importance of "joy of reading" and emphasizes the need to encourage boys to read in order to improve in their reading abilities. Joy of reading is important for reading skill (OECD; 2021) and the school is the context in which stimulate this interest. In fact, it emerges that in the classes where the hours spent in reading are more, the gap seems to decrease. We believe that the well-known propensity of boys towards more technical-scientific paths (science, technology, engineering, and mathematics, STEM) leads to underestimate the importance of this competence, which instead constitutes a very important basic skill also in these fields of education. The results highlighted in this study acquire greater relevance in the Italian context where the average scores in reading literacy obtained by Italian students in international reading assessments are significantly below the OECD average. This means that male students in Italy on average have insufficient reading skills to understand texts and information. Reading, decipher information, and communicate is increasingly important in a digital world and it is necessary to deepen the study of the factors that influence the lower involvement and performance in reading, in a society increasingly based on the correct understanding of information.

References

ALVARADO L. K. A. 2017. Gender Score Gaps of Colombian Students in the Pisa Test, *Vniversitas Economica*, Vol. 17, No. 8.

BINKLEY M., ERSTAD O., HERMAN J., RAIZEN S., RIPLEY M., MILLER-RICCI M., RUMBLE M. 2012. Defining twenty-first century skills. In *Assessment and teaching of 21st century skills* (pp. 17-66), Springer, Dordrecht.

BLINDER, A.S. 1973. Wage Discrimination Reduced Form and Structural Estimates, *Journal of Human Resources*, Vol. 8, No. 4, pp. 436-455.

BORGONOVI F. 2016. Video Gaming and Gender Differences in Digital and Printed Reading Performance Among 15-Year-Olds Students In 26 Countries, *Journal of Adolescence*, Vol. 48, pp. 45-61.

BROZO W. G., SULKUNEN, S., SHIEL, G., GARBE, C., PANDIAN, A., & VALTIN, R. 2014. Reading, Gender, and Engagement: Lessons from Five PISA Countries, *Journal of Adolescent & Adult Literacy*, Vol. 57, No. 7, pp. 584-593.

CARD D., PAYNE A. A. 2021. High School Choices and the Gender Gap in STEM, *Economic Inquiry*, Vol. 59, No. 1, pp. 9-28.

DOWNEY D. B., VOGT YUAN A. S. 2005. Sex Differences in School Performance During High School: Puzzling Patterns and Possible Explanations, *Sociological quarterly*, Vol. 46, No. 2, pp. 299-321

KHORRAMDEL L., POKROPEK A., JOO S. H., KIRSCH I., & HALDERMAN L. 2020. Examining Gender DIF and Gender Differences in The PISA 2018 Reading Literacy Scale: A Partial Invariance Approach, *Psychological Test and Assessment Modeling*, Vol. 62, No. 2, pp. 179-231.

LINNAKYLÄ P., MALIN, A. 2000. *3 How to reduce the gender gap in reading*. Northern Lights on PISA, 39.

LONGOBARDI S., PAGLIUCA M. M., REGOLI, A. 2018. Can Problem-Solving Attitudes Explain the Gender Gap in Financial Literacy? Evidence From Italian Students' Data, *Quality & Quantity*, Vol. 52, No. 4, pp. 1677-1705.

MCHALE S. M., CROUTER A. C., TUCKER, C. J. 2001. Free-Time Activities In Middle Childhood: Links With Adjustment in Early Adolescence, *Child development*, Vol. *72*, No. 6, pp. 1764-1778.

MOSTAFA T. 2019. Why don't more girls choose to pursue a science career? *PISA in Focus n.93*, Paris, OECD Publishing.

MUNIR F., WINTER-EBMER R. 2018. Decomposing International Gender Test Score Differences, *Journal for Labour Market Research,* Vol. 52, No. 12, pp. 1-17.

OAXACA R., 1973. Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review,* Vol. 14, No.3, pp. 693-709.

OECD, 2015. *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence.* PISA, OECD Publishing, Paris.

OECD, 2019. *PISA 2018 Results (Volume I): What Students Know and Can Do*. PISA, OECD Publishing, Paris.

OECD 2021, *21st-Century Readers: Developing Literacy Skills in a Digital World*. PISA, OECD Publishing, Paris.

WEIS R., CERANKOSKY B. C. 2010. Effects Of Video-Game Ownership on Young Boys' Academic and Behavioral Functioning: A Randomized, Controlled Study, *Psychological science*, Vol. 21, No. 4, pp. 463-470.

WOESSMANN L. 2016. The Importance of School Systems: Evidence from International Differences in Student Achievement, *Journal of Economic Perspectives*, Vol. 30, No. 3, pp. 3-32.

SUMMARY

Why do boys perform worse than girls on reading literacy? Evidence from PISA survey 2018

The present study analyses the main factors behind the differences in test scores on reading literacy in Italy between male and female 15 years old students performing the Oaxaca-Blinder decomposition method (Blinder, 1973) on the data from the OECD PISA 2018 survey, in which reading was the main subject assessed.

In order to estimate the amount of the differential between male and female average wages, we have applied the decomposition of Oaxaca and Blinder (Oaxaca 1973, Blinder 1973). Through this method we distinguish which part is due to differences in characteristics included in model estimations and which part is due to gender "discrimination".

Our results suggest that the gender gap in reading is mostly attributable to explained component, while the variables related to the discrimination effect are those related to different habits between girls and boys.

Valentina FERRI, Istituto Nazionale per l'Analisi delle Politiche Pubbliche (INAPP), v.ferri@inapp.org
Giovanna DI CASTRO, Istituto Nazionale per l'Analisi delle Politiche Pubbliche (INAPP), g.dicastro@inapp.org

AN EMPIRICAL EVALUATION OF COMMON CROSS-NATIONAL MEASURES OF STATE CAPACITY

Andrea Vaccaro

1. Introduction

State capacity is of broad and current interest in many social science subfields. Most researchers and policymakers agree that a capable state apparatus with effective institutions is crucial for the creation of long-term socioeconomic prosperity. Conversely, "weak or failed states are close to the root of many of the world's most serious problems" (Fukuyama, 2004). Therefore, today, building well-functioning state institutions is a high-priority global objective: it is one of the targets of the UN 2030 Agenda Sustainable Development Goals (SDGs), and more generally, considered as a necessary condition for the achievement of the SDGs.

Despite the widespread consensus on the importance of state capacity, there is much less agreement on how the concept – usually understood as the state's ability to reach its objectives (e.g., Acemoglu and Robinson, 2019) – should be quantified. Plenty of different measures have been used to capture state capacity. Yet, these measures are hardly ever evaluated comparatively. In particular, the empirical differences between these measures and the quality of these measures are seldom discussed in the literature.

If measures of broadly related concepts such as democracy (e.g., Knutsen, 2010; Boese, 2019; Vaccaro, 2021) and rule of law (e.g., Skaaning, 2010; Møller and Skaaning, 2011; Versteeg and Ginsburg, 2017) have been analysed extensively, only a handful of studies have compared measures of state capacity. Furthermore, most of these comparative studies have taken primarily a conceptual (e.g., Cingolani, 2018) rather than empirical (e.g., Hendrix, 2010) approach to the issue.

While conceptual questions are important, empirical aspects ought not to be treated superficially. As well put by Jerven (2013), "scholars pay great attention to defining the concepts ad devote great effort to theorizing the existence of the phenomenon and spend comparatively little time critically probing the numbers that are supposed to represent them". My study aims then to tackle this shortcoming in the literature by providing one of the first in-depth comparative empirical assessments of common measures of state capacity. The central contribution of my study is to help researchers and policymakers to make better choices among

competing measures of state capacity by reducing current "informational uncertainties" (Mudde and Schedler, 2010).

My study proceeds as follows. First, I conduct a survey of existing cross-national measures of state capacity and select the most "currently relevant" measures for further analysis. Second, I examine the basic statistical properties of the selected measures and assess the measures according to two criteria: normality of distribution and ability to discriminate between observations. Additionally, the most problematic measures are further discussed through empirical examples. Third and last, in the conclusive section, I sum up the main findings and provide ideas for future research on the topic.

2. Selection of data

Social scientists have used a variety of different cross-national measures to quantify state capacity. Since it is impossible to analyse comprehensively all these measures in a single study, I pick out some of the most relevant ones for further analysis. The selection process is carried out according to four criteria: (1) academic relevance, (2) type of data, (3) coverage, (4) accessibility.

The first criterion is met if a measure has been used to quantify state capacity in several comparative social science studies. The second criterion is met if a measure is based – at least partially – on subjective (i.e., perception-based) data. The third criterion is met if a measure provides yearly country-level scores over time and across countries in the world. The fourth criterion is met if a measure is publicly and freely available for use. An extensive review of recently used datasets suggests that seven measures meet the selection criteria. Table 1 presents these measures.

Table 1 – *Selected measures of state capacity.*

Measure	Producer
Quality of government index (QOG)	Quality of Government Institute
Capacity index (HSI)	J. Hanson and R. Sigman
Government effectiveness (WGI)	The World Bank
State fragility index (SFI)	Center for Systemic Peace
Failed states index (FSI)	Fund for Peace
Corruption perceptions index (CPI)	Transparency International
Impartial public administration (VDM)	Varieties of Democracy

Before proceeding with the analysis, it is important to stress that the selected measures differ – at least up to a certain extent – in terms of content and intended purpose. That said, assessing the content validity and/or the intended purpose of the measures is out of the scope of my study. What matters for the study at hand instead is the effective use of these measures to quantify state capacity. Following the

aforementioned selection criteria, regardless of their content and their intended purpose, the chosen measures have been all frequently used to quantify the concept of state capacity in comparative cross-national social science research.

Quality of government index (QOG) is published by the Quality of Government Institute (Teorell *et al.*, 2020). It aggregates three indicators – *Bureaucracy quality*, *Corruption*, and *Law and order* – from PRS Group's International Country Risk Guide into a single multidimensional index. The data is thus entirely coded by PRS Group's country experts. QOG ranges from 0 (low) to 1 (high) and provides data for most countries (140 in 2015) in the world from 1984 onwards.

Capacity index (HSI) is developed by Hanson and Sigman. It synthesises pre-existing data on the extractive, administrative, and coercive capacities of the state (Hanson and Sigman, 2021). The underlying data contains both subjective and objective indicators. HSI runs from low to high and provides annual data for up to 177 countries from 1960 to 2015. In our sample, the final index ranges from –2.31 to 2.96, with a mean of 0.68 and a standard deviation of 0.93.

Government effectiveness (WGI) is one of the six Worldwide Governance Indicators (Kaufmann *et al.*, 2011). It synthesises perception-based data related to the quality of public administration and the quality of public services from nearly 20 sources into a single composite index. WGI runs from low to high on a standardised (z-score) scale with a mean of 0 and a standard deviation of 1. It provides yearly data from 2003 onwards (biannual data from 1996 to 2002) for virtually all countries and some territories in the world (209 in 2015).

State fragility index (SFI) is published by the Center for Systemic Peace (Marshall and Elzinga-Marshall, 2017). It is based on 14 sub-indicators related to four aspects of state effectiveness and state legitimacy: political, social, economic, and security. The final index combines both subjective and objective data. SFI is scaled from 0 to 25, where 0 represents the highest and 25 the lowest level of state capacity. It provides yearly data for all countries in the world with a population of at least 500,000 from 1995 onwards (167 in 2015).

Fragile states index (FSI) is developed by the Fund for Peace. Its scores are based on qualitative expert assessment, content analysis of articles and reports, and quantitative secondary data relevant to 12 dimensions of the state such as security and rule of law (Fund for Peace, 2017). FSI provides yearly data for most countries in the world (177 in 2015) from 2005 onwards. The final index is scaled from 0 to 120, where 0 represents the highest and 120 the lowest level of state capacity.

Corruption perceptions index (CPI) is published by Transparency International. It is based on pre-existing perception-based data (from 12 sources in 2015) on public sector corruption and closely related aspects such as transparency and bureaucratic professionalisation (Transparency International, 2015). CPI provides annual scores from 1995 onwards, and its 2015 edition covers 168 countries in the world. The

index ranges from 0 to 10 until 2011 and from 0 to 100 from 2012 onwards. A higher score indicates less corruption, and thus, more state capacity.

Impartial public administration (VDM) is produced by the Varieties of Democracy Institute (Pemstein *et al.*, 2019). The measure answers the question: "Are public officials rigorous and impartial in the performance of their duties?" (Coppedge *et al*., 2019). Its scores are based on expert coding and its annual data goes back as far as 1789 for nearly all countries in the world (177 in 2015). In our sample, VDM ranges from –3.22 (low) to 3.61 (high), with a mean of 0.43 and a standard deviation of 1.47.

3. Research strategy

Now that I have selected some of the most relevant measures of state capacity, it is time to proceed to compare and evaluate them. This section describes briefly my research strategy. The next section presents and discusses the empirical results.

I start the empirical evaluation of the measures of state capacity by exploring their basic statistical properties through violin plots. Approximately normally distributed variables are preferable to completely non-normal variables, not because we expect real world state capacity to be distributed Gaussian, but because many common statistical tests and analyses assume that variables follow more or less a bell-shaped curve. The visual assessment of the measures is complemented with a formal Shapiro-Wilk test for normality (Royston, 1992).

Acquiring information on basic statistical properties and distributional characteristics of measures of state capacity is important in itself. Nevertheless, violin plots allow us to discover also some less apparent measurement issues, which are further examined through individual country scores.

To be more specific, considering that all sciences aim to describe reality accurately (Goertz 2020) and that "a key purpose of measurement in social sciences is to identify, quantify and possibly explain the differences that exist between units of analysis" (Gnaldi *et al.*, 2017), then all other things equal a measure of state capacity that is unable to discriminate between countries must be inferior to a measure that successfully describes true cross-national differences.

If such lack of precision causes scores to clump at extreme values, a measure additionally fails to capture interesting variation in its entirety, and thus, its scale is not extended enough (Goertz, 2020). On these grounds, it seems reasonable to assume that the inability to discriminate between different cases and clumping at extreme values are characteristics that a flawless measure of state capacity should not possess.

As we have seen in the previous section, the selected measures of state capacity do not have equal scales. Hence, to ease the comparability among the measures, I normalise[1] (min-max) all indicators to range from 0 to 1, where a lower score indicates a weaker state and a higher score indicates a stronger state. This means also that throughout the empirical analysis the original scales of FSI and SFI are reversed. Missing data is deleted listwise. Therefore, the empirical analysis includes only country-years that are common to all the selected measures. This ensures that our results are not even minimally driven by differences in samples.

4. Results and discussion

4.1 Basic statistical properties

Violin plots (Figure 1) illustrate the basic statistical properties of the selected measures in all common country-years. The grey outlines of the "violins" show the frequency distribution of each measure. The black-bordered box in the middle of each violin stretches out from the first to the third quartile of each variable. The whiskers stretch out to the lowest and highest observations that are not considered unusual in the data. Single observations that do not fall inside this range of the data (i.e., outliers) are represented by dots above or below the whiskers. The small black rectangle inside the box represents the median.

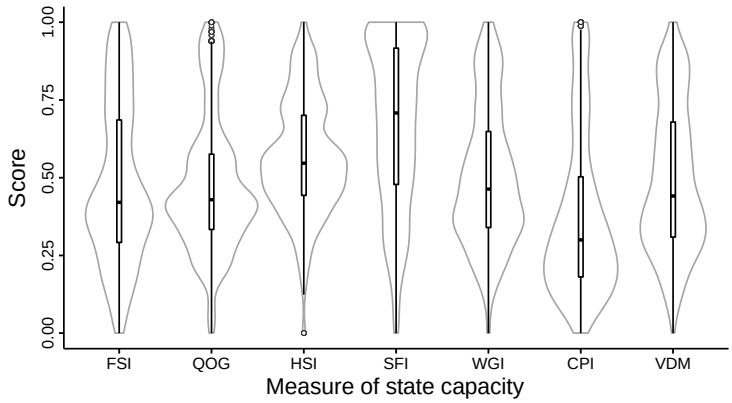
First, a visual inspection of the violin plots suggests that all the measures deviate from a perfectly bell-shaped distribution. The results of a formal Shapiro-Wilk test for normality confirm that none of the measures is normally distributed. In comparative terms, however, there are some interesting differences between the measures. FSI, QOG, HSI, WGI, and VDM seem to be more normally distributed than CPI and SFI. They have a main peak at intermediate levels of state capacity, the median approximately at halfway of the scale, and a lower frequency of observations at the two extremes of the scale. CPI and SFI instead seem to be far from having a reasonably bell-shaped distribution.

CPI is heavily skewed to the right and its mode and median are remarkably low. As many as 41.1% of CPI's observations are in the bottom quarter of the scale and 74.5% of its observations are below the mid-point of the scale. In practice, this means that many countries receive much lower scores with CPI than with the other measures.

[1] See Mazziotta and Pareto (2021) for a comprehensive analysis of the advantages and disadvantages of common normalisation methods.

SFI, instead, suffers from the opposing problem: it has a heavily left-skewed distribution and a comparatively high mode and median. As many as 43.6% of its observations are in the topmost quarter of the scale, but only around 8.1% of its observations are in the bottom quarter of the scale. This means that in general countries are more "capable" with SFI than with the other measures.

Figure 1 – *Violin plots of measures of state capacity (2005-2015).*



N = 1411.

Overall then, FSI, QOG, HSI, WGI, and VDM are more normally distributed than CPI and SFI. Their main modes, medians, and means are relatively close to the halfway of the scale. FSI's, QOG's, WGI's, and VDM's main peaks and medians are closer to the low end of the scale, whereas HSI's main peak and median is closer to the high end of the scale. QOG and CPI have some outliers at the upper extreme of the scale, whereas HSI has a single outlier observation at the low extreme of the scale. FSI, WGI, SFI, and VDM do not have any outliers.

4.2 Ability to discriminate between countries

As already said, acquiring information on the above features of the data is helpful in itself. Additionally, however, violin plots are useful in revealing certain less obvious empirical shortcomings in the measures. In particular, a careful inspection of the shapes of the distributions of our measures, suggests that SFI compresses too many observations at the upper extreme of the state capacity scale. From 2005 to 2015, SFI rates nearly 200 country-years with the maximum possible level of state capacity. In 2015, the most recent year of common observations, SFI assigns the maximum possible score to as many as 20 countries (out of 130).

As shown by Table 2, SFI is by far less sensitive than the other six measures in describing differences between high capacity countries. For instance, there is abundant evidence that the German state apparatus is more capable than the Italian one (Fukuyama, 2014) and that state institutions in Sweden function more effectively than state institutions in Spain (Dahlström and Lapuente, 2017). Yet, SFI is the only measure that fails to discriminate between these countries.

Different countries should have the same score only if such equivalence reflects reality. Since the other measures of state capacity are able to distinguish not only between Germany, Italy, Sweden, and Spain but also between all the other countries in Table 2 nearly without exceptions, we are induced to conclude that SFI has severe limitations in its ability to distinguish high state capacity countries one from another. Its scores are thus relatively imprecise and do not reflect well reality.

Table 2 – *State capacity in countries with maximum score with SFI in 2015*.

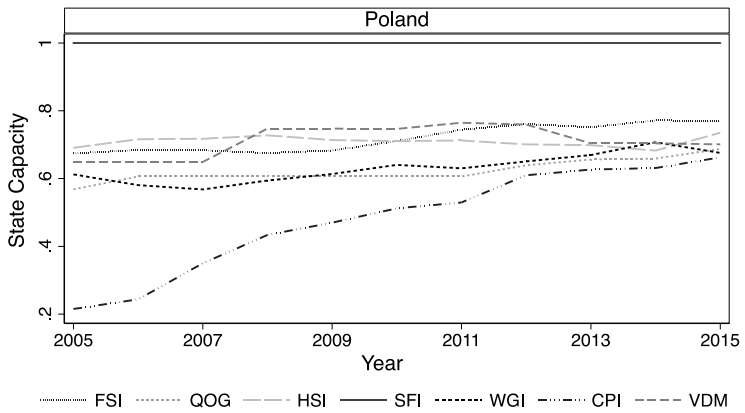
Country	SFI	FSI	QOG	HSI	WGI	CPI	VDM
Austria	1.000	0.909	0.938	0.955	0.829	0.819	0.838
Canada	1.000	0.947	0.938	0.874	0.893	0.904	0.929
Czech Rep.	1.000	0.769	0.688	0.768	0.732	0.578	0.775
Denmark	1.000	0.972	1.000	1.000	0.912	1.000	1.000
Estonia	1.000	0.742	0.641	0.792	0.736	0.747	0.928
Finland	1.000	1.000	1.000	0.925	0.904	0.988	0.903
France	1.000	0.835	0.781	0.780	0.820	0.747	0.800
Germany	1.000	0.897	0.906	0.917	0.887	0.880	0.949
Hungary	1.000	0.644	0.625	0.754	0.608	0.518	0.679
Ireland	1.000	0.961	0.938	0.856	0.841	0.807	0.828
Italy	1.000	0.745	0.547	0.755	0.597	0.434	0.664
Japan	1.000	0.829	0.875	0.806	0.898	0.807	0.762
Latvia	1.000	0.700	0.641	0.745	0.741	0.578	0.819
Netherlands	1.000	0.901	0.969	0.899	0.909	0.916	0.774
Poland	1.000	0.770	0.688	0.735	0.676	0.663	0.701
Portugal	1.000	0.891	0.750	0.807	0.771	0.675	0.726
Slovenia	1.000	0.841	0.688	0.770	0.715	0.627	0.789
Spain	1.000	0.779	0.719	0.855	0.760	0.602	0.884
Sweden	1.000	0.960	1.000	0.928	0.905	0.976	0.886
UK	1.000	0.857	0.906	0.808	0.889	0.880	0.941

Min-max normalised scores from low to high. Missing data deleted listwise.

The inability to differentiate between countries is not the only problem caused by such an agglomeration of observations at extreme values. Given that so many countries have the maximum score with SFI, as a consequence, the index is also unable to detect any possible increase in the level of state capacity over time in multiple countries. We would obviously expect an ideal cross-national measure of state capacity to be informative both about differences across countries and over time changes within single countries.

As many as 12 countries have the maximum possible score with SFI every year from 2005 to 2015. Yet, generally speaking, it is unrealistic to assume that state capacity has not changed at all in any of these countries in more than ten years. Figure 2 shows more detailed evidence of the evolution of the level of state capacity in one of these countries: Poland. As we can see from the line plot, all measures except SFI identify an increase in the level of state capacity in Poland from 2005 to 2015.

Figure 2 – *State capacity in Poland (2005-2015).*



This imprecise representation of reality is one of the practical consequences of SFI's inability to discriminate between high capacity countries. Since the index assigns the maximum possible score to so many countries, it is simply not as sensitive as the other measures in detecting changes in high capacity countries, and especially, it is not able to capture improvements over time in any of these countries. Considering also that according to case studies the level of state capacity in Poland has actually increased since the country became a member of the European Union in 2004 (Charasz and Vogler, 2021), there are no doubts that SFI describes high capacity countries less accurately that the other six measures.

Clumping at the upper extreme of the scale and the inability to distinguish between high capacity countries are thus weaknesses of SFI. Luckily, the other six measures of state capacity do not have such an agglomeration of equivalently rated observations at either of the two extremes of the scale. Nevertheless, some of these measures do have some less severe problems in distinguishing observations one from another.

CPI is not finely grained enough to discriminate between many countries at low levels of state capacity. For instance, in the most recent year of common observations

(2015), Brazil, Burkina Faso, India, Thailand, Tunisia, and Zambia have exactly the same score with CPI. QOG, instead, is not finely grained enough to discriminate between many countries at intermediate levels of state capacity. Just to give an example, in the most recent year of common observations, Albania, Algeria, Bangladesh, Egypt, Kazakhstan, Pakistan, Panama, Papua New Guinea, Peru, Uganda, and Zambia have exactly the same score with QOG.

With both measures, despite the empirical equivalence in country-scores, it is unlikely that there are no actual differences in the level of state capacity of so many countries. Otherwise the remaining measures would not be able to capture the differences between these countries virtually without exceptions.

FSI, HSI, WGI, and VDM are able to distinguish well between different countries. In fact, in 2015, WGI and VDM are able to distinguish between all common observations. With FSI and HSI instead no more than two countries have an identical score in 2015. If we hold on to the assumption that the ability to differentiate between countries is an asset of any measure of state capacity, from this specific perspective, FSI, HSI, WGI, and VDEM can be considered to be more informative than SFI, CPI, and QOG. Overall, SFI seems to have less desirable empirical features than the other surveyed measures.

5. Conclusions

This study has compared and evaluated seven frequently used measures of state capacity, in terms of basic statistical properties and the ability to discriminate observations one from another. Most existing comparative studies on the measurement of state capacity have focused mainly on conceptual issues. To address this shortcoming in the literature, the approach adopted in this study has been empirical, not conceptual.

My findings indicate that SFI is the most problematic of the evaluated measures. First, its values are far from being normally distributed. Second, it is not able to discriminate between countries with a high level of state capacity. Third, it is not able to capture any possible increase in the level of state capacity in many high capacity countries, and thus, its scale should be further extended. Given these findings, at least in respect of our evaluation criteria, researchers should not use SFI as a measure of state capacity, unless they have strong theoretically justifiable reasons to do so.

My analysis reveals that the other six measures of state capacity are less tricky in terms of our evaluation criteria. Of the six remaining measures, CPI seems to have the least desirable features. It has too many observations at low levels of state capacity and has some difficulties in discriminating between these low-capacity

countries. QOG is not very effective in discriminating between some of the countries at intermediate levels of state capacity, but apart from that, the remaining five measures are fairly bell-shaped and able to distinguish well different countries one from another.

My findings provide a first look into the empirical aspects of cross-national measures of state capacity. In general, they should be considered as an initial step on the path towards a more comprehensive understanding of existing data on the state and state capacity. My study provides valuable guidance for the users of this data, but it should not be seen as a final and conclusive analysis on the topic.

On the contrary, many important questions on measures of state capacity and their quality remain to be addressed in future comparative studies. For instance, since I found that some of the measures rate countries with high divergence, future studies should assess whether these empirical differences affect the results of inferential research. Future research should also explore the causes of these empirical differences and assess the quality of the input side of the data generation process (e.g., transparency, replicability, aggregation). The comparative analysis at hand can be used as a starting point for these future studies.

References

ACEMOGLU, D., ROBINSON, J.A. 2019. *The narrow corridor: States, societies, and the fate of liberty*. New York, NY: Penguin Press.

BOESE, V.A. 2019. How (Not) to Measure Democracy, *International Area Studies Review*, Vol. 22, No. 2, pp. 95-127.

CHARASZ, P., VOGLER, J.P. 2021. Does EU Funding Improve Local State Capacity? Evidence from Polish Municipalities, *European Union Politics*, Vol. 22, No. 3, pp. 446-471.

CINGOLANI, L. 2018. The Role of State Capacity in Development Studies, *Journal of Development Perspectives*, Vol. 2, No. 1-2, pp. 88-114.

COPPEDGE, M., GERRING, J., KNUTSEN, C.H., LINDBERG, S.I., TEORELL, J., ALTMAN, D., ZIBLATT, D. 2019. V-Dem Codebook v9. Varieties of Democracy (V-Dem) Project.

DALHSTRÖM, C., LAPUENTE, V. 2017. *Organizing Leviathan*. Cambridge: Cambridge University Press.

FUND FOR PEACE. 2017. *Fragile states index and CAST framework methodology*. Washington, DC: The Fund for Peace.

FUKUYAMA, F. 2004. The Imperative of State-Building, *Journal of Democracy*, Vol. 15, No. 2, pp. 17-31.

FUKUYAMA, F. 2014. *Political order and political decay*. New York, NY: Farrar, Straus and Giroux.

GNALDI, M., DEL SARTO, S., MAGGINO, F. 2017. The role of extended IRT models for composite indicators construction. In MAGGINO F. (Ed) *Complexity in Society: From Indicators Construction to their Synthesis*, Cham: Springer, pp. 213-227.

GOERTZ, G. 2020. *Social science concepts and measurement: new and completely revised edition*. Princeton, NJ: Princeton University Press.

HANSON, J.K., SIGMAN, R. 2021. Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research, *Journal of Politics*, Vol. 83, No. 4, pp. 1495-1510.

HENDRIX, C.S. 2010. Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict, *Journal of Peace Research*, Vol. 47, No. 3, pp. 273-285.

JERVEN, M. 2013. *Poor numbers*. Ithaca, NY: Cornell University Press.

KAUFMANN, D., KRAAY, A., MASTRUZZI, M. 2011. The Worldwide Governance Indicators: Methodology and Analytical Issues, *Hague Journal on the Rule of Law*, Vol. 3, No. 2, pp. 220-246.

KNUTSEN, C.H. 2010. Measuring Effective Democracy, *International Political Science Review*, Vol. 31, No. 2, pp. 109-128.

MARSHALL, M.G., ELZINGA-MARSHALL, G. 2017. *Global report 2017: Conflict, governance, and state fragility*. Vienna, VA: Center for Systemic Peace.

MAZZIOTTA, M., PARETO, A. 2021. Everything You Always Wanted to Know About Normalization (but Were Afraid to Ask), *Rivista Italiana di Economia Demografia e Statistica*, Vol. 75, No. 1, pp. 41-52.

MØLLER, J., SKAANING, S.E. 2011. On the Limited Interchangeability of Rule of Law Measures, *European Political Science Review*, Vol. 3, No. 3, pp. 371-394.

MUDDE, C., SCHEDLER, A. 2010. Introduction: Rational Data Choice, *Political Research Quarterly*, Vol. 63, No. 2, pp. 410-416.

PEMSTEIN, D., MARQUARDT, K.L., TZELGOV, E., WANG Y., MEDZIHORSKY, J., KRUSELL, J., MIRI, F., VON RÖMER, J. 2019. The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data. *The Varieties of Democracy Institute Working Papers 2019-21*.

ROYSTON, P. 1992. Approximating the Shapiro-Wilk W-Test for Non-Normality, *Statistics and Computing*, Vol. 2, No. 3, pp. 117-119.

SKAANING, S.E. 2010. Measuring the Rule of Law, *Political Research Quarterly*, Vol. 63, No. 2, pp. 449-460.

TEORELL, J., DAHLBERG, S., HOLMBERG, S., ROTHSTEIN, B., ALVARADO PACHON, N., AXELSSON, S. 2020. The QoG Standard Dataset, version Jan20. University of Gothenburg: The Quality of Government Institute.

TRANSPARENCY INTERNATIONAL. 2015. Corruption Perceptions Index 2015: Full Source Description. Available at: https://www.transparency.org/en/cpi.

VACCARO, A. 2021. Comparing Measures of Democracy: Statistical Properties, Convergence, and Interchangeability, *European Political Science*. doi: 10.1057/s41304-021-00328-8 (advance online publication).

VERSTEEG, M., GINSBURG, T. 2017. Measuring the Rule of Law: A Comparison of Indicators, *Law and Social Inquiry*, Vol. 42, No. 1, pp. 100-137.

SUMMARY

An Empirical Evaluation of Common Cross-National Measures of State Capacity

Today, most experts have no doubt that *state capacity* is crucial for sustainable development. Anyhow, there is no agreement on how to quantify state capacity and there are plenty of different measures of state capacity to choose from. Ideally, one should pick a measure that closely represents the chosen theory. If that is not possible, however, existing literature offers little guidance in helping scholars to select one measure over another. The study at hand contributes to fill this gap in the literature by comparing and evaluating selected empirical characteristics of seven frequently used measures of state capacity in commonly available years (2005-2015). Ultimately, thus, this study provides new valuable guidance to the users of measures of state capacity.

By drawing on previous comparative research on social science measurement, the author of this study analyses common cross-national measures of state capacity mainly along two desirable features: the normality of distribution and the ability to discriminate between cases. The results of this study show that measures of state capacity are not all the same: some measures have more desirable empirical features than others. In particular, *State fragility index* seems to have the least advantageous characteristics. Therefore, unless theoretically justifiable, the study at hand does not recommend its use in statistical analyses on the topic.

Andrea VACCARO, Department of Social Sciences and Economics, Sapienza University of Rome, andrea.vaccaro@uniroma1.it

STILL BELIEVE IT? AN ANALYSIS OF PARTNERSHIP TRAJECTORIES AFTER FIRST UNION DISSOLUTION IN ITALY [1]

Alessandra De Rose, Eleonora Meli

1. Introduction

The deep transformations in conjugal behaviour, such as the growing number of consensual unions, but also the recent introduction of the "brief" divorce and other legislative innovations, have changed the context of re-partnering in Italy. In order to provide an updated picture of the Italian situation, we examine the patterns of conjugal behaviour after the first union dissolution by analysing data from the Survey "Families, Social Subjects and Life Cycle", carried out by Istat in 2016.

The aim of this work is to examine the characteristics of re-partnership, focusing on the second union type, including LAT (Living Apart Together), among the various forms of coupling. Since re-partnering is a field not extensively investigated, and LAT is a very uncommon way to consider a partnership, we intend to contribute with original results to the research on consequences of marital dissolution in contemporary societies.

The topic of re-partnering has long been the subject of attention, mainly in countries with decades of experience of separations and divorces (Thornton, 1977; Oppenheimer, 1988; Amato, 2000, 2010; Wu and Schimmele, 2005; Beaujouan, 2012; Ivanova *et al.*, 2013; Gałęzewska *et al.*, 2017; Mortelmans, 2020). In Italy, studies on the consequences of marital dissolution are relatively recent, due to the late diffusion of marital breakdown (Salvini and Vignoli, 2011). Despite this delay, the Italian divorce rate is on the increase, although still below that of the countries of central and northern Europe: the most recent data show a level about 1.5 divorce per 1,000 inhabitants against 1.9 on average in EU (Impicciatore and Guetto, 2021). As a consequence, the incidence of second marriages is on the increase (Istat, 2021). Meanwhile, couples based on informal unions have spread and, because of their higher instability, also re-partnering of previously cohabiting individuals is on the increase (De Rose *et al.*, 2008).

[1] This article is the result of the collaboration between the authors. In particular, paragraphs 1 (Introduction) and 4 (Conclusion) are attributed to Alessandra De Rose and paragraphs 2 (Data and methods) and 3 (Results) are attributed to Eleonora Meli.

The past decades trend in union formation and dissolution is spurring research on living arrangements after the break down and on their characteristics. Findings of the analyses conducted on Italian data so far, also in a comparative perspective, suggest that the speed and frequency of re-partnering are very different between men and women and highly dependent on age, on presence of children from the previous union, on socio-occupational status, namely on education and work, as well as on cultural traits (Rettaroli, 1997; Angeli and De Rose, 2003; 2007; Meggiolaro and Ongaro, 2008; Gałęzewska *et al.*, 2017).

The behaviour after first union dissolution can be very heterogeneous. Some individuals never enter a new couple, while others follow a variety of different paths: some meet a new partner very quickly, while others take longer; some form a lasting union while others separate after a short time; some marry (or remarry) while others simply cohabit; and a few may experience several different unions in succession. Among the possible ways of living a new romantic relationship after the failure of the first one, cohabitation does not necessarily have to be foreseen. A different living arrangement, which is summarized by the term LAT, proves to be an alternative to cohabitation (Levin, 2004). As outlined by Liefbroer and colleagues (2015): "*Most people in LAT unions intend to live together but are apart for practical reasons. LAT is more common among young people, those enrolled in higher education, people with liberal attitudes, highly educated people, and those who have previously cohabited or been married. Older people and divorced or widowed persons are more likely to choose LAT to maintain independence*". The propensity to establish a shared home decreases with increasing levels of educational qualification and with income, and vice versa, it increases at the lower end of the social hierarchy (Régnier-Loilier, 2019). In fact, money constraints can push couples to move in together to benefit from certain economies of scale associated with cohabitation (one rent rather than two, etc.). The presence of cohabiting children also affects the likelihood to form another cohabiting union, and this especially true for women (Vanassche *et al.*, 2015). In these cases, couples often pursue LAT relationship rather than cohabit or marry (de Jong Gierveld, 2004; de Jong Gierveld & Merz, 2013; Duncan *et al.*, 2013). Moreover, cohabitation is not the predominant choice of re-partnering for those experiencing union dissolution at older age, who want to maintain own independence (Schimmele and Wu, 2016; Brown and Wright, 2017). In Italy, studies on LAT are quite rare and the latest evidence suggest that LAT relationships are essentially concentrated in the early phases of the life course, among young couples who are waiting to start a cohabitation or to marry (Régnier-Loilier and Vignoli, 2018).

In this paper, we will provide an overview of the re-partnering behaviour of Italian individuals after the eventual break-up of their first union and try to enlighten the role of LAT as an alternative way of living a new romantic experience. Our goal

is twofold: we first estimate the probability that people who had dissolved a first union get into a new union and which factors are mainly associated with this risk; then, we concentrate on those entering a new union and try to disentangle among cohabiting and non-cohabiting unions and factors associated with the alternative choices.

2. Data and methods

Our analysis examines the union histories collected by the Istat Survey "Families, Social Subjects and Life Cycle" in 2016. Using the sample of people aged 18 and over, the unions' trajectories are investigated to reconstruct the sequence of the unions; moreover, many other pieces of information have been considered in order to build a reference frame. Our sample includes all individuals who have had their first cohabiting union dissolved (n=4,997); the sub-sample of those individuals who have had a second union (n=1,682) has been then selected in order to estimate the propensity to choose one type of union or another. Weighted data have been shown in the analysis to account for the stratified sampling design used, as well as nonresponse.

Retrospective union histories offer a complete data collection on the partnership path. Our study starts considering the individual's first cohabiting union that could have been an informal cohabitation, a direct marriage (without a previous cohabitation) or an indirect marriage (with a previous cohabitation). For those who dissolved their first union with the end of the cohabitation, a separation, or a divorce we consider any subsequent union, with or without cohabitation. In this second step, unions studied are informal cohabitations, second marriages and LATs.

In explaining patterns of re-partnering in Italy, and according to the empirical literature in the field, we consider the role of the following characteristics[2]: type of first union, territory, gender, birth cohort, age at union break-up, duration of first union, presence and age of the youngest child from previous union, educational level, support network.

Territory is the geographical area of residence at time of the interview, that could be different from those at the union formation or dissolution.

Birth cohorts are grouped into ten-year periods (1939 and before, 1940-1949, 1950-1959, 1960-1969, 1970-1979, 1980 and after).

Duration of the first union is the discrete time, in years, from the union's beginning – start of cohabitation for informal union and indirect marriages or year of marriage – until the first date of union break-up.

[2] In this paper we only included time-constant variables.

Presence and age of the youngest child (if any) from the previous union are also considered. We built a three-category variable: No children, youngest child less than 18, youngest child 18 or older.

Educational level is the higher educational attainment obtained, categorised in three classes: "High" is at least a university degree, "Medium" is secondary school diploma, "Low" is middle school diploma or less.

Support network is the dimension of the community that could offer help in case of need. It is defined as a "Strong network" if the person can count on at least two types of networks (neighbours, friends, relatives, or someone who would lend 800 euro in case of need), "Weak" only one network and "No support network".

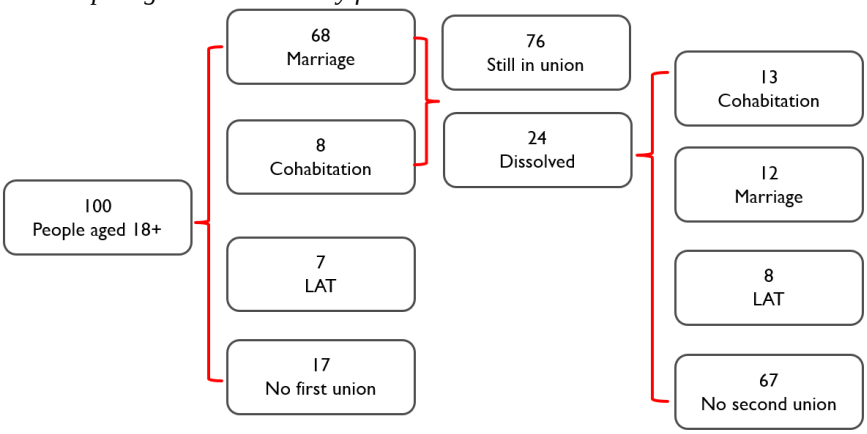
To estimate the risk of entering a second union and investigate on the association between the survival time and our predicting variables we apply the standard Cox proportional-hazards model. In this work, duration is defined as the time interval between date of first union break-up and second union (if any) or date at the interview.

Then, we selected only those who entered a second union, and we analysed the differences between types of union (marriage, cohabitation, and LAT). Performing separated logistic regression analyses, firstly we estimate the likelihood to choose a non-cohabiting relation (LAT) rather than a cohabiting one (considering both marriages and cohabitations); then, we discriminate among those who choose to live together, estimating the probability to getting married rather than starting an informal cohabitation. All models are performed in a multivariate setting, including all the variables previously illustrated.

3. Results

Paths of unions' formation and dissolution can be very different. In Figure 1 reducing the complexity, the common transitions from one status to the next are shown. Based on data collected on union histories, we estimated that 10 million people have experienced a union dissolution in the Italian population; in quite 80% of the cases the first union was a marriage, leftover from a cohabitation. Among those who dissolved the first union, more than one third of them engaged in a new partnership. Cohabitation has been chosen by 13.3% and marriage by 12.1%, while 8.1% of people chose to have a relationship without cohabiting.

Figure 1 – *People aged 18 and over by first and second union - Year 2016.*



Source: Istat, Families, social subjects and life cycle.

In Table 1 a more detailed analysis of the partnership trajectories is reported: 77.7% of those who dissolved a marriage are still single, while only 26.2% of those who dissolved a cohabitation did not enter a new union. Over one third of former cohabitants choose again as second union a cohabitation. Entering a new union without cohabitation (LAT) is chosen by a minority of former partners (8.1%), however the percentage among the former cohabiting is about three times that among former married individuals.

Table 1 – *People who dissolved their first union, by following trajectories Year 2016.*

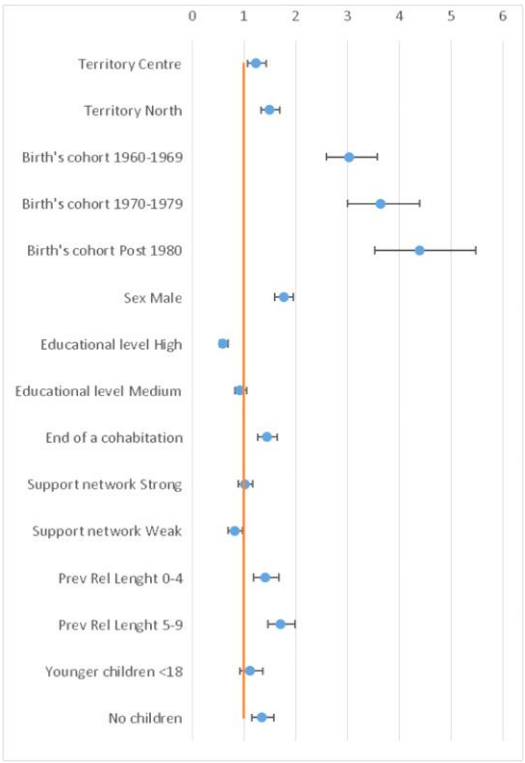
	Second union				
	No second union	Marriage	Cohabitation	LAT	All
First union					
Marriage dissolution	77.7	9.2	7.5	5.6	100
End of a cohabitation	26.2	22.6	34.0	17.3	100
All	66.5	12.1	13.3	8.1	100

Source: Istat, Families, social subjects and life cycle.

Re-partnering depends on birth cohort and gender, as well as on different factors related to personal and contextual characteristics. In Figure 2 the hazard ratios from the multivariate Cox's regression analysis are reported. Results suggest a higher probability to enter earlier in a second union for men, for those who live in the Centre and North of the country with respect to the South. Younger people show higher

probability to enter a new union[3]. There are conditions which appear to facilitate re-partnering such as having ended a cohabitation instead of a marriage, having had a shorter union (less than 10 years) and not having children from the previous union. Education, instead, plays a negative role, that is people with high educational level have less chances to enter a new union, while no clear pattern of effect is observed for the support network.

Figure 2 − *Cox's regression results for entering in a second union.*



Source: Istat, Families, social subjects and life cycle.

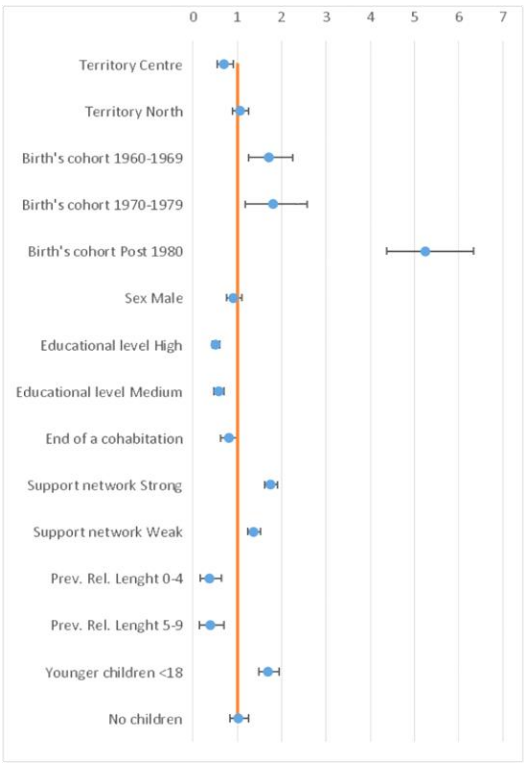
In a second step of the study, we deepen the analysis of the second unions. Thus, we concentrate on those individuals who did enter a second union and perform two distinct logistic regression models in order to: 1) single out the different

[3] In a different model - here not shown but available upon request - we also included age at separation, that, as expected, has a strong negative effect on the re-partnering risks, the other effects unchanged, including that of the birth cohort. For the seek of model parsimony - especially in the following logistic regressions based on a lower number of cases - we retain only the birth cohort as indicator of age.

characteristics associated with the choice of entering a LAT rather than a cohabiting union; 2) evaluate, among those who started a new cohabitation, what are the differences between marriage and informal unions.

In Figure 3 the odds ratios of the model that evaluates the probability to choose a LAT rather than cohabitation as a second union are shown.

Figure 3 – *Logistic regression model on the probability to choose a LAT instead of marriage or a cohabitation.*

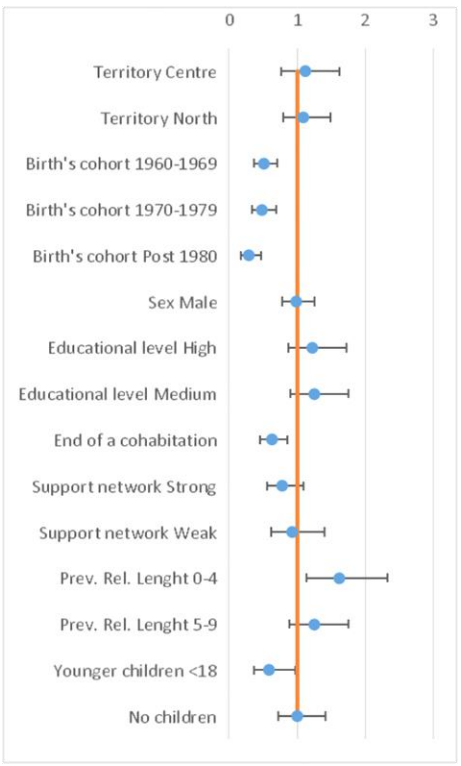


Source: Istat, Families, social subjects and life cycle.

The youngest people (those born after 1980) are more likely to choose a LAT-type union, as well as those having at least one child aged less than 18. Another characteristic that favours LAT is having a strong support network, a factor that has apparently no role in time to re-partnering, as previously noted. Even duration of previous union plays a different role: people coming from unions that lasted 10 years or more are more likely to enter a LAT as a second union instead of a cohabiting one.

Focussing on the individuals entering a new cohabiting union (Figure 4), the probability of getting married rather than starting a new informal cohabitation is higher among those who dissolved a previous marriage than among those coming from a previous informal cohabitation and among those whose experienced a very short first relationship (less than 5 years); instead, it is lower for younger generations and for those who have at least one child under 18 years of age. Interestingly, no gender differences turn to be statistically significant in both of these two last models.

Figure 4 − *Logistic regression model on the probability to choose a marriage instead of a cohabitation among those living together a partner in a second relationship*.



Source: Istat, Families, social subjects and life cycle.

4. Discussion

Transition to the second union is quite high: our data confirm that one in three people who exit a first union enter a new one. This would seem a clue that people "still believe it" and are willing to try again. Another interesting finding in this respect is the persistence of the partnership model: although the likelihood to re-partner is higher among those coming from an informal union, the probability of getting married rather than informally cohabit is higher among those who have already a marriage behind them, and those who have dissolved an informal union tend to repeat the same experience.

An alternative is emerging, that is starting a new sentimental relationship without cohabitation. LATs, although a still marginal reality, are becoming a true partnership form especially in certain circumstances, among the youngest, among those with small children and among those who can count on a solid network of help.

Our study confirms that men re-partner more often than women, but no gender or territorial differences appear as far as the type of second union is concerned. Also, territory differences persist: the risk to re-partner is higher in the North than in the rest of the country. The presence of children and their age act as an important constraint for the formation of a new union and condition the form of union chosen, also controlling for gender, age and duration of the union. Level of education also plays a non-trivial role, with the more educated people less likely to enter a second union: this result can be interpreted with the narrower marriage market that the most educated people have access to (de Graaf and Kalmijn, 2003; Wu and Schimmele, 2005) and given the persistent high homogamy by education level in our country (De Rose and Fraboni, 2016). An alternative explanation of this result is the independence hypothesis, that is the major human capital of more educated individuals makes them less in need of a new partner. Both interpretations deserve further investigation, namely on two directions: first, performing separated analyses by gender, since second marriage/unions market is highly different between men and women (Theunis *et al.*, 2015); second, by including other and more objective indicator of economic independence (Pasteels and Mortelmans, 2017) - such as income and employment status - which are however time-varying, requiring a more complex analytical approach, that go beyond the explorative purposes of this paper.

References

AMATO P.R. 2000. The Consequences of Divorce for Adults and Children, *Journal of Marriage and Family*, Vol. 62, No. 4, pp. 1269-1287.

AMATO P.R. 2010. Research on Divorce: Continuing Trends and New Developments, *Journal of Marriage and Family*, Vol. 72, No. 3, pp. 650–666.

ANGELI A., DE ROSE A. 2003. Donne e uomini dopo lo scioglimento della prima unione, in PINNELLI A., RACIOPPI F., RETTAROLI R., *Genere e demografia*, Bologna: Il Mulino, pp.367-392.

ANGELI A., DE ROSE A. 2007. Women and men after the first union dissolution. In PINNELLI A., RACIOPPI F., RETTAROLI R. (Eds) *Genders in the Life Course*, Dordrecht: Springer, pp. 185-204.

BEAUJOUAN É. 2012. Repartnering in France: The Role of Gender, Age, And Past Fertility, *Advances in Life Course Research*, Vol. 17, No. 2, pp. 69-80.

BROWN S.L., WRIGHT M.R. 2017. Marriage, Cohabitation, and Divorce in Later Life, *Innovation in Aging*, Vol. 1, No. 2.

DE JONG GIERVELD J. 2004. Remarriage, Unmarried Cohabitation, Living Apart Together: Partner Relationships Following Bereavement or Divorce, *Journal of Marriage and Family*, Vol. 66, No. 1, pp. 236-243.

DE JONG GIERVELD J., MERZ, E. M. 2013. Parents' Partnership Decision Making After Divorce or Widowhood: The Role of (Step) Children, *Journal of Marriage and Family*, Vol. 75, No.5, pp. 1098-1113.

DE ROSE A., RACIOPPI F., ZANATTA A.L. 2008. Italy: Delayed adaptation of social institutions to changes. In Family Behaviour, *Demographic Research*, Vol. 19, No.19, pp. 665-704.

DE ROSE A., FRABONI R. 2016. Educational Assortative Mating in Italy: What Can Gini's Homogamy Index Still Say?, *Genus*, Vol. 71, No. 2-3, pp. 53-71.

DUNCAN S., CARTER J., PHILLIPS M., ROSENEIL S., STOILOVA M. 2013. Why Do People Live Apart Together?, *Families, Relationships and Societies*, Vol.2, No.3, pp. 323-338.

GAŁĘZEWSKA P., PERELLI-HARRIS B., BERRINGTON A. 2017. Cross-National Differences in Women's Repartnering Behaviour in Europe: The Role of Individual Demographic Characteristics, *Demographic Research*, Vol. 37, pp. 189-228.

DE GRAAF P.M., KALMIJN M. 2003. Alternative Routes in The Remarriage Market: Competing-Risk Analyses of Union Formation After Divorce, *Social Forces*, Vol. 81, No. 4, pp. 1459-1498.

IMPICCIATORE R., GUETTO R. (Eds) 2021. *50 anni dalla legge sul divorzio. Cause e conseguenze dell'instabilità matrimoniale in Italia*. Firenze: Neodemos.

ISTAT 2021. Matrimoni, unioni civili, separazioni e divorzi. Anno 2019, *Statistiche Report*, 18.

IVANOVA K., KALMIJN M., UUNK W. 2013. The Effect of Children On Men's And Women's Chances Of Re-Partnering In A European Context, *European Journal of Population*, Vol. 29, No. 4, pp. 417-444.

LEVIN I. 2004. Living Apart Together: A New Family Form, *Current Sociology*, Vol. 52, No. 2, pp. 223-240.

LIEFBROER A.C., POORTMAN A.R., SELTZER J.A. 2015. Why Do Intimate Partners Live Apart? Evidence on LAT Relationships Across Europe, *Demographic Research*, Vol 32, No. 8, pp. 251-286.

MEGGIOLARO S., ONGARO F. 2008. Repartnering After Marital Dissolution: Does Context Play a Role?, *Demographic Research*, Vol. 19, pp. 1913-1934.

MORTELMANS D. (Ed.) 2020. *Divorce in Europe. New insights in trends, causes and consequences of relation break-ups*. European Studies of Population No. 21, eBook: SpringerOne.

OPPENHEIMER V.K. 1988. A Theory of Marriage Timing, *American Journal of Sociology*, Vol. 94, No. 3, pp. 563-591.

PASTEELS I., MORTELMANS D. 2017. The Socioeconomic Determinants of Repartnering After Divorce or Separation in Belgium, *Demographic Research*, Vol. 36, pp.1785-1812.

RÉGNIER-LOILIER A., 2019. New Partner, New Living Arrangements? The Process of Repartnering After Separation, *Population*, Vol. 74, No. 1, pp. 71-100.

RÉGNIER-LOILIER A., VIGNOLI D. 2018. The Diverse Nature of Living Apart Together Relationships. A France-Italy Comparison, *Journal of Population Research*, Vol. 35, No. 1, pp. 1-22.

RETTAROLI R., 1997. *Le seconde nozze*. In BARBAGLI M. SARACENO C. *Lo stato delle famiglie in Italia* Bologna: Il Mulino.

SALVINI S., VIGNOLI D. 2011. Things Change: Women's and Men's Marital Disruption Dynamics in Italy During a Time of Social Transformations, 1970-2003, *Demographic Research*, Vol. 24, pp. 145-174.

SCHIMMELE C.M., WU Z. 2016. Repartnering After Union Dissolution in Later Life, *Journal of Marriage and Family*, Vol. 78, No. 4, pp. 1013-1031.

THEUNIS L., PASTEELS I. VAN BAVEL J. 2015. Educational Assortative Mating After Divorce: Persistence or Divergence from First Marriages? *Zeitschrift für Familienforschung*, Vol. 27, pp. 183-202.

THORNTON A. 1977. Decomposing The Re-Marriage Process, *Population Studies*, Vol. 31, No. 2, pp. 383-92.

VANASSCHE S., CORIJN M., MATTHIJS K., SWICEGOOD G. 2015. Repartnering and Childbearing After Divorce: Differences According to Parental

Status and Custodial Arrangements, *Population Research and Policy Review*, Vol. 34, No. 5, pp. 761-784.
WU Z., SCHIMMELE C.M. 2005. Repartnering After the First Union Disruption, *Journal of Marriage and Family*, Vol. 67, No. 1, pp. 27-36.

SUMMARY

Still believe it?
An analysis of partnership trajectories after first union dissolution in Italy

We examine the patterns of partnership behaviour following first union dissolution analysing the individual union histories from the Survey "Families, Social Subjects and Life Cycle", carried out by Istat in 2016. The aim of the work is to analyse the frequency and characteristics of re-partnering, and to explore the role of LAT (Living Apart Together) as an alternative for those who form a new union without cohabitation. The Risk of entering a second union by time elapsed since first union dissolution has been estimated with a Cox proportional-hazards model. Then, we selected individuals who entered a second union and applied multivariate logistic models to analyse factors associated with the likelihood to enter a certain partnership form (marriage, informal cohabitation, or LAT) rather than another. More than one third of those who dissolved the first union engaged in a new partnership. Male partners, living in the North, with low-medium level of education, with no children and exiting from a first consensual union show the highest risk to form a new union. LAT proves to be an alternative to a new cohabiting union (marriage or informal). The odds of choosing this new form of union increase with the younger birth-cohort, those with children and with strong family support. Among those who enter a second cohabiting union, marriage is generally chosen by the already married and informal partnership among those who experienced the same living arrangement.

Alessandra DE ROSE, Sapienza Università di Roma, alessandra.derose@uniroma1.it
Eleonora MELI, Istat, eleonora.meli@istat.it

DYNAMICS AND STABILITY IN AN OLG MODEL WITH NON-SEPARABLE PREFERENCES

Giorgia Marini

1. Introduction

In the literature on steady state stability in OLG models, Diamond (1965) model represents the benchmark model with separable preferences. Under the assumption of separable preferences, de la Croix (1996) proves that the optimal solution in the Diamond (1965) model is always characterised by monotonic convergence to the steady state. Michel and Venditti (1997) and de la Croix and Michel (1999) both depart from Diamond (1965) model assuming non-separable preferences: Michel and Venditti (1997) provide sufficient conditions for stability of the equilibrium in an OLG model with separable preferences across generations only and prove that the optimal solution may be oscillating and optimal cycles may exist; de la Croix and Michel (1999) provide sufficient conditions for existence and uniqueness of the equilibrium in an OLG model with separable preferences within generations only and prove that the optimal solution may display damped oscillations even when the social planner does not discount the utility of future generations.[1]

The main contribution of this paper is to provide sufficient conditions for existence and uniqueness of a steady state equilibrium in a production economy of overlapping generations with non-separable preferences, i.e. habits are transmitted from one generation to the next one (intergenerational spillover) *and* from one period to the next one (intragenerational spillover), and to analyse the implications of non-separable preferences (*both* across and within generations) for the local stability of the steady state equilibrium. The present OLG model is by no means trivial as it provides completely new results in terms of dynamics and stability under the assumption of non-separable preferences, both across and within generations, and it therefore fills in a gap in the existing literature.

The rest of the paper is organised as follows. In section 2 we describe the model, in section 3 we analyse the competitive setting, while in section 4 we analyse the optimal solution and derive conditions under which the optimal solution is stable.

[1] Both Michel and Venditti (1997) and de la Croix and Michel (1999) assume a production economy of overlapping generations of agents. Two references for dynamics and stability in a pure exchange economy of overlapping generations agents with non- separable preferences are instead Lahiri and Puhakka (1998) and Orrego (2014).

We prove that, under the assumption of intergenerational and intragenerational spillovers, convergence of the optimal solution to the non-trivial steady state is not always assured and that the optimal solution may display either locally explosive dynamics or damped oscillations. We provide concluding remarks in section 6.

2. The model

The model is a simple extension of the Diamond (1965) economy without outside money. At each date, the economy is populated by three generations (young, adult and old), each living for three periods. The growth rate of population is zero. The young generation has no decision to take and only inherits habits h_t from the previous adult generation according to the following equation

$$h_t = c_{t-1}^a \tag{1}$$

where c_{t-1}^a is the consumption of the adult generation at time $t-1$. The adult generation draws utility from consumption of the quantity c_t^a, given its own stock of habits h_t. When old, each agent draws utility from consumption of the quantity c_{t+1}^o, given her own past consumption c_t^a. The intertemporal utility function of each adult agent is

$$U(c_t^a, c_{t+1}^o; h_t) = u(c_t^a - \theta h_t) + v(c_{t+1}^o - \delta c_t^a) \tag{2}$$

where $\theta \in (0,1)$ measures the intensity of the intergenerational spillover effect due to the inherited habits (passive effect) and $\delta \in (0,1)$ measures the intensity of the intragenerational spillover effect due to the persistence of own preferences over time (active effect). In other words, we assume that adult consumption at time $t-1$ determines a frame of reference against which adult individual consumption at time t is judged and that the depreciation rate of these inherited habits is so high that it no longer affects the evaluation of consumption when old. We also assume that adult consumption at time t determines a frame of reference against which old individual consumption at time $t+1$ is judged and that persistence of preferences is so high that neither young consumption at time t nor that at time $t-1$ affect in any possible way the evaluation of consumption when old.

Moreover we assume that the utility function is strictly increasing with respect to consumption and decreasing with respect to the stock h: $u_{c^a} > 0$, $v_{c^o} > 0$, $u_h < 0$, $u_{c^a c^a} < 0$, $v_{c^o c^o} < 0$, $u_{hh} < 0$ and $u_{c^a h} > 0$. The assumption $u_{c^a h} > 0$ amounts to postulating that an increase in the stock h rises the desire for consumption. We also assume that starvation is ruled out in both periods

$$\lim_{c_t^a \to 0} u_{c^a} + v_{c^a} = \lim_{c_{t+1}^o \to 0} v_{c^o} = \infty \tag{3}$$

and that the utility function is strictly concave under the following condition

$$\delta < \frac{u_{c^a}}{v_{c^o}} \tag{4}$$

Note that if preferences are separable as in Diamond (1965), $\delta = 0$ and strictly concavity is always ensured by the standard set of assumptions on marginal utility, i.e. $u_{c^a} > 0$ and $v_{c^o} > 0$. If preferences are non- separable, $\delta > 0$ and concavity is ensured only if condition (4) holds. Otherwise, the utility function is flat ($\delta = \frac{u_{c^a}}{v_{c^o}}$) or convex ($\delta > \frac{u_{c^a}}{v_{c^o}}$).

At each date a single good is produced. This good can be either consumed or accumulated as capital for future production. Production occurs through a constant return to scale technology. Per capita output y_t is a function of capital intensity k_t

$$y_t = f(k_t) \tag{5}$$

in which $f(\ )$ is a neoclassical production function with $f_k > 0$ and $f_{kk} < 0$. Assuming total depreciation of capital after one period, the resource constraint of the economy is

$$y_t = c_t^a + c_t^o + k_{t+1} \tag{6}$$

At date 0 the economy is endowed with a given quantity of capital per capita k_0 and a level of inherited habits h_0.

3. The competitive economy

The competitive behaviour of firms leads to the equalisation of the marginal productivity of each factor to its marginal cost:

$$R_t = f_k(k_t) \tag{7}$$
$$w_t = f(k_t) - k_t f_k(k_t) \tag{8}$$

where R_t is the interest factor paid on loans and w_t is the real wage paid to workers.

The adult generation works during the period t and sells one unit of labor inelastically at any real wage w_t, consumes the quantity c_t^a and saves s_t for the next period by holding capital

$$c_t^a = w_t - s_t \tag{9}$$

while the old generation spends all her savings s_t plus interest matured and consumes c_{t+1}^o

$$c_{t+1}^o = R_{t+1} s_t \tag{10}$$

The maximisation program of each individual is thus to choose c_t^a, c_{t+1}^o in order to

$$\max_{c_t^a, c_{t+1}^o} u(c_t^a - \theta h_t) + v(c_{t+1}^o - \delta c_t^a)$$

$$\text{subject to } \begin{matrix} c_t^a = w_t - s_t \\ c_{t+1}^o = R_{t+1} s_t \end{matrix}$$

where w_t, $R_{t+1}$and h_t are given to the agent. Assuming an interior solution, under rational expectations, the above decision problem leads to following first order condition

$$u_{c^a} - \delta v_{c^a} = R_{t+1} v_{c^o} \tag{11}$$

With respect to a standard Diamond (1965) model in which $\delta = 0$, marginal utility of the young is lower, as $v_{c^o} > 0$: in order to achieve the same level of satisfaction when old, adults need to correct their satisfaction by the (negative) habit effect.

Equation (11) allows to define the following saving function

$$s_t = s(w_t, R_{t+1}, h_t) \tag{12}$$

The partial derivative of the saving function (12) are

$$s_w = \frac{u_{c^a c^a}}{u_{c^a c^a} + R_{t+1}^2 v_{c^o c^o}} > 0, s_r = \frac{-(v_{c^o} + v_{c^o c^o} c_{t+1}^o)}{u_{c^a c^a} + R_{t+1}^2 v_{c^o c^o}}, s_h = \frac{-\theta u_{c^a h}}{u_{c^a c^a} + R_{t+1}^2 v_{c^o c^o}} < 0$$

Since the utility function is concave and there is no wage income in the last period of life, savings increase with wage income. The effect of the interest rate is instead ambiguous and depends on the value of the intertemporal elasticity of substitution, $\frac{v_{c^o c^o} c_{t+1}^o}{v_{c^o}}$. Finally, the effect of rising inherited habits is negative: when the passive effect is low, the agent has a sober lifestyle and savings are high; when the passive effect is high, the agent spends much on consumption to maintain a life standard similar to the one of their peers and their propensity to save is low.

The equilibrium condition in the capital market implies

$$k_{t+1} = s_t \tag{13}$$

Combining equations (1), (7), (8), (12) and (13), the competitive equilibrium is defined as a sequence $\{k_t, h_t; t > 0\}$ which satisfies

$$k_{t+1} = s(f(k_t) - k_t f_k(k_t), f_k(k_{t+1}), h_t) \tag{14}$$

$$h_{t+1} = f(k_t) - k_t f_k(k_t) - s(f(k_t) - k_t f_k(k_t), f_k(k_{t+1}), h_t) \tag{15}$$

Equation (14) is the clearing condition of the asset market, given that the labour market is in equilibrium (i.e. that (8) holds). It reflects the fact that savings are to be

equal to the capital stock of the next period. Equation (15) is the equation (1), given that the asset and the labour markets are in equilibrium.

As in de la Croix and Michel (1999), the competitive equilibrium is characterised by spillovers from one generation to the next and from adulthood to old age. The main components of the intergenerational spillovers are: savings by old people and past consumption levels of the previous generation. While the process transforming the savings by the old into income for the adult displays decreasing returns to scale due to the characteristics of the production function, the process transforming past consumption of the adults into consumption of the next generation displays constant returns to scale due to the characteristics of the utility function. The intragenerational spillovers is only given by the individual past consumption that feeds individual's habits from adulthood to old age. This process displays constant returns to scale, again due to the characteristics of the utility function. Thus, even though the intergenerational bequest in terms of higher wages will not be sufficient to cover the intergenerational bequest in terms of higher inherited habits, the intragenerational spillover leaves a bequest in terms of higher persistence. The combination of the positive bequests in terms of higher wages and higher persistence is sufficient to offset the negative bequest in terms of the higher externality. This leads to an increase in saving to maintain future standards of consumption that induces an expansion. When the enrichment is strong enough, the externality has already reverted to higher levels, allowing a fall in savings and the start of a recession. As the effect due to persistence is stronger than the effect due to the externality, the model is characterised by converging cycles. Thus, the competitive equilibrium still displays fluctuations, but the bifurcation corresponds to different critical values of θ and δ. Depending on the parameters θ and δ, the economy may converge to or diverge from the steady state.

As such results are perfectly in line with the existing literature on growth when preferences are transmitted from one generation to another (de la Croix and Michel, 1999), the rest of the paper focuses on the optimal solution and derives conditions under which the sub-optimality and instability caused by the externalities (intergenerational and intragenerational spillovers due to non-separable preferences) can be overtaken.

4. The optimal solution

As inherited habits introduce an externality in the model, the decentralised equilibrium is obviously sub- optimal compared to the equilibrium that would maximise the planner's utility. Thus, hereafter, we focus our attention to the optimal solution and consider a social planner who chooses the allocation of output in order to maximise the present discount value of current and future generations.

Assuming that the social planner's discount factor is γ, the social planner maximisation program is thus to choose $\{c_t^a, c_t^o\}$ and $\{k_t, h_t\}$ in order to

$$\max_{c_t^a, c_t^o, k_t, h_t} \sum_{t=0}^{\infty} \gamma^t \left[u(c_t^a - \theta h_t) + \frac{1}{\gamma} v(c_t^o - \delta c_t^a) \right] \quad (16)$$

$$\text{subject to } \begin{matrix} y_t = c_t^a + c_t^o + k_{t+1} \\ h_t = c_{t-1}^a \end{matrix}$$

and given k_0 and h_0. First order conditions are

$$u_{c^a}(c_t^a - \theta h_t) + \gamma u_h(c_{t+1}^a - \theta h_{t+1}) = \frac{1}{\gamma} v_{c^o}(c_t^o - \delta h_t) - v_h(c_{t+1}^o - \delta h_{t+1}) \quad (17)$$

$$\frac{1}{\gamma} v_{c^o}(c_t^o - \delta h_t) = v_{c^o}(c_{t+1}^o - \delta h_{t+1}) f_k(k_{t+1}) \quad (18)$$

Equation (17) is a condition for optimal intergenerational allocation of consumption between adult and old alive at the same time. Marginal utility of the adult, corrected by the social planner to internalise the taste externality, is equalised to marginal utility of the old is equal to the marginal utility of the old. Note that, due to the presence of the taste externality and contrary to the standard Diamond (1965) model, this social planner's first order condition does not respect the individual first order condition (11). Moreover, with respect to the standard Diamond (1965) model in which $\delta = \theta = 0$, marginal utility of the adult, $u_{c^a}(c_t^a - \theta h_t) + \gamma u_h(c_{t+1}^a - \theta h_{t+1})$, is lower, as $u_h < 0$, while marginal utility of the old, $\frac{1}{\gamma} v_{c^o}(c_t^o - \delta h_t) - v_h(c_{t+1}^o - \delta h_{t+1})$, is higher, as $v_h < 0$. Equation (18) sets the optimal intertemporal allocation.

The optimal equilibrium is defined as a sequence $\{c_t^a, c_t^o, k_t, h_t; t > 0\}$ which satisfies equations (17), (18), (6) and (1) simultaneously:

$$u_{c^a}(c_t^a - \theta h_t) + \gamma u_h(c_{t+1}^a - \theta h_{t+1}) = \frac{1}{\gamma} v_{c^o}(c_t^o - \delta h_t) - v_h(c_{t+1}^o - \delta h_{t+1}) \quad (19)$$

$$\frac{1}{\gamma} v_{c^o}(c_t^o - \delta h_t) = v_{c^o}(c_{t+1}^o - \delta h_{t+1}) f_k(k_{t+1}) \quad (20)$$

$$h_t = c_{t-1}^a \quad (21)$$

$$k_{t+1} = f(k_t) - c_t^a - c_t^o \quad (22)$$

It appears from the system above that the steady state (c^a, c^o, k, h) of this optimal economy is defined by

$$u_{c^a}(c^a - \theta h) + \gamma u_h(c^a - \theta h) = \frac{1}{\gamma} v_{c^o}(c^o - \delta h) - v_h(c^o - \delta h) \quad (23)$$

$$\frac{1}{\gamma} = f_k(k_{t+1}) \quad (24)$$

$$h = c^a \quad (25)$$

$$k = f(k) - c^a - c^o \quad (26)$$

Equation (23) shows that in an economy with passive habits, the marginal utility of the adult is lower than the corresponding marginal utility in the standard Diamond (1965) model: the inheritance represents a benchmark from which individuals want to depart. Even the marginal utility of the old is higher than the corresponding marginal utility in the standard Diamond (1965) model: the same interpretation carries on. Once the externality associated with parents' habits is internalised, persistence affects marginal utility in the same way as the externality: they both induce consumers to save. Equation (24) is the modified golden rule: the introduction of intergenerational and intragenerational spillovers does not modify the optimal steady- state stock of capital which remains fixed at the modified golden rule level. Equations (25) and (26) have been already discussed in the paper.

Proposition 1. *A positive steady state equilibrium exists and is unique if and only if* $\det(\mathbf{I} - \mathbf{J}^{SO}) \neq 0$, *where* $\mathbf{J}^{SO}$ *is the Jacobian matrix associated to the optimal equilibrium (19)-(22) and evaluated at steady state* (c^a, c^o, k, h).

Proof. See Appendix A.1.

Proposition 2. *Assume that* k *and* h *are state variables and that* c^a *and* c^o *are jump variables. Locally explosive dynamics is possible, depending on the sign of the trace* $\mathbf{T_J}^{SO}$ *and of the element* Z *of the Jacobian matrix* $\mathbf{J}^{SO}$. *If* $\Delta \geq 0$, *the eigenvalues are real and local dynamics is either explosive or monotonic. If* $\Delta < 0$, *the eigenvalues are complex and conjugate and local dynamics displays either explosive or damped oscillation.*

Proof. See Appendix A.2.

The above proposition identifies all possible dynamics of the optimal steady state equilibrium. Under the assumption that the trace $\mathbf{T_J}^{SO}$ and the element Z are both positive, locally explosive dynamics is identified by an unstable node if the eigenvalues are real and by an unstable focus if the eigenvalues are complex and conjugate. Under the assumption that $\mathbf{T_J}^{SO}$ and Z are both negative, the optimal solution is a stable saddle point, only if the constraints on the elements of the Jacobian matrix $\mathbf{J}^{SO}$ respect the condition on negativity of the trace. Under the assumption that $\mathbf{T_J}^{SO}$ and Z have opposite sign, the optimal solution may be either stable or unstable: if stable, dynamics displays damping oscillation to the steady state; if unstable, locally explosive dynamics occurs when the constraints on the elements of the matrix $\mathbf{J}^{SO}$do not respect the condition on negativity of the trace.

The stability of the optimal steady state equilibrium depends on the assumption that habits are trans- mitted both across and within generations, assumption that affects the sign of the trace $\mathbf{T_J}^{SO}$and of element Z. Monotonic convergence to the optimal steady state equilibrium is ensured only under the assumption that the stock of inherited habits does not persist into their old age, i.e. only if $\delta = 0$ as in the standard Diamond (1965) model. Contrary to the competitive equilibrium, the

optimal solution is only characterised by a positive intragenerational spillover: savings by the old, that directly finance the capital stock required for production in the next period and indirectly sustain wages of the adult. The intergenerational spillover due to habits is *a priori* internalised by the social planner in the maximisation problem (16). As in the competitive equilibrium, the process transforming the savings by the old into income for the adult displays decreasing returns to scale, due to the characteristics of the production function. However, the intergenerational bequest in terms of higher wages does not interact with any other spillover. The intergenerational bequest in terms of higher wages will lead to a constant increase in saving that induces a permanent expansion. The model might thus be characterised by diverging explosive dynamics. A numerical example is provided in Marini (2015).

5. Conclusions

This paper derives sufficient conditions for existence of a steady state equilibrium in an OLG model with non-separable preferences and analyses the implications of non-separable preferences for the local stability of the steady state equilibrium.

It studies conditions for the existence and stability of optimal equilibrium and it proves that the optimal solution may display damped oscillations or locally explosive dynamics. This result crucially depends on the assumption that habits are transmitted from one generation to the next one and from adulthood to old age.

This paper shows that combining different forms of non-separable preferences is not innocuous: when we introduce persistence of individual tastes in the contest of an OLG model in which habits are inherited, dynamics of the model and stability of the equilibrium are considerably affected. The results presented in this paper are therefore fundamental to understanding the mechanisms underneath models with habit formation and habit persistence, as habits seem to play a significant role in many aspects of economic theory.

Appendix

A.1 Proof of Proposition 1

First, linearize the non linear dynamic system (19)-(22) around the steady state (23)-(26) and get

$$\begin{bmatrix} \mathrm{d}h_{t+1} \\ \mathrm{d}c^a_{t+1} \\ \mathrm{d}k_{t+1} \\ \mathrm{d}c^o_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \dfrac{[B(1-\gamma E)-1]}{\gamma} & A-B(D-E) & \dfrac{DB}{\gamma} & C-B(1+D) \\ 0 & -1 & \dfrac{1}{\gamma} & -1 \\ E & D-E & -\dfrac{D}{\gamma} & 1+D \end{bmatrix} \begin{bmatrix} \mathrm{d}h_t \\ \mathrm{d}c^a_t \\ \mathrm{d}k_t \\ \mathrm{d}c^o_t \end{bmatrix}$$

in which $A \equiv -\frac{u_{c^a c^a}+\gamma u_{hh}+v_{hh}}{\gamma u_{c^a h}} > 0$, $B \equiv \frac{v_{c^o h}}{\gamma u_{c^a h}} > 0$, $C \equiv \frac{v_{c^o c^o}}{\gamma^2 u_{c^a h}} < 0$, $D \equiv \frac{\gamma v_{c^o} f_{kk}}{v_{c^o c^o}} > 0$, $E \equiv \frac{v_{c^o h}}{v_{c^o c^o}} < 0$ under the assumption that in steady state $f_k(k_{t+1}) = f_k(k_t) = \gamma^{-1}$. As the Jacobian matrix evaluated at steady state (c^a, c^o, k, h) is

$$\mathbf{J}^{\mathrm{SO}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \dfrac{[B(1-\gamma E)-1]}{\gamma} & A-B(D-E) & \dfrac{DB}{\gamma} & C-B(1+D) \\ 0 & -1 & \dfrac{1}{\gamma} & -1 \\ E & D-E & -\dfrac{D}{\gamma} & 1+D \end{bmatrix}$$

it is immediate to show that the matrix $\mathbf{I} - \mathbf{J}^{\mathrm{so}}$ is equal to

$$\mathbf{J}^{\mathrm{SO}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -\dfrac{[B(1-\gamma E)-1]}{\gamma} & 1-A+B(D-E) & -\dfrac{DB}{\gamma} & -C+B(1+D) \\ 0 & 1 & 1-\dfrac{1}{\gamma} & 1 \\ -E & -D+E & \dfrac{D}{\gamma} & -D \end{bmatrix}$$

and that its determinant is

$$\det(\mathbf{I}\text{-}\mathbf{J}^{\mathrm{SO}}) = (A+B-C)D - D - \tfrac{(1-B)D}{\gamma} = D\left[(A+B-C) + \tfrac{B}{\gamma} - \tfrac{1+\gamma}{\gamma}\right]$$
$$= \tfrac{\gamma v_{c^o} f_{kk}}{v_{c^o c^o}}\left[\tfrac{v_{c^o h} - (u_{c^a c^a}+\gamma u_{hh}+v_{hh})}{\gamma u_{c^a h}} + \tfrac{v_{c^o h} - v_{c^o c^o}}{\gamma^2 u_{c^a h}} - \tfrac{1+\gamma}{\gamma}\right] \neq 0$$

under the assumptions that the utility function is concave, that the production function is neoclassical, that equations (23)-(26) hold and that conditions (25) and (26) are met.

A.2 Proof of Proposition 2

The characteristic polynomial P in the eigenvalues σ associated to the Jacobian matrix $\mathbf{J}^{\mathrm{SO}}$ evaluated at steady state (c^a, c^o, k, h) is

$$P(\sigma) = \sigma^4 - \mathbf{T}^{\mathrm{SO}}_{\mathbf{J}}\sigma^3 + Z\sigma^2 - \gamma^{-1}\mathbf{T}^{\mathrm{SO}}_{\mathbf{J}}\sigma + \det(\boldsymbol{J}^{SO}) = 0$$

in which

$$\det(\boldsymbol{J}^{SO}) = \frac{1 - B + \gamma CE}{\gamma^2} = \frac{1}{\gamma^2}$$

$\mathbf{T}_{\mathbf{J}}^{SO} = 1 + \gamma^{-1} + A - B(D - E) + D \gtrless 0$

if $1 + \gamma^{-1} + A + D \gtrless B(D - E)$

$Z = 2\gamma^{-1} + (1 + \gamma^{-1} + A - B(D - E) + D) \gtrless 0$

if $1 + 3\gamma^{-1} + A + D \gtrless B(D - E)$

In order to study the polynomial P, factorize the polynomial P into

$P(\sigma) = (\sigma - \sigma_1)(\sigma - \sigma_2)(\sigma - \sigma_3)(\sigma - \sigma_4) = 0$

which is equivalent to

$$\left(\sigma^2 - \phi_1\sigma + \frac{1}{\gamma}\right)\left(\sigma^2 - \phi_2\sigma + \frac{1}{\gamma}\right)$$

in which $\phi_1 \equiv \sigma_1 + \sigma_2$ and $\phi_2 \equiv \sigma_3 + \sigma_4$. Then analyse all possible scenarios, due to the sign's ambiguity of the trace $\mathbf{T}_{\mathbf{J}}^{SO}$and of element Z.

First, assume that $\mathbf{T}_{\mathbf{J}}^{SO}$and Z are both positive and analyse the two possible cases:

1. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} \geq 0$, $i = 1,2$. The four eigenvalues are real and they can be:
   a. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates $\mathbf{T}_{\mathbf{J}}^{SO} > 0$.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is excluded as it violates $Z > 0$.
   c. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is accepted as it respects both conditions on $\mathbf{T}_{\mathbf{J}}^{SO}$ and Z.
2. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} < 0$, $i = 1,2$. Look at the real parts only. Since the real part $a = -\frac{1}{2}\phi_i \neq 0$, $i = 1,2$, the eigenvalues are complex and conjugate and they can be:
   a. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates $\mathbf{T}_{\mathbf{J}}^{SO} > 0$.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is excluded as it violates $Z > 0$.
   c. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is accepted as it respects both conditions on $\mathbf{T}_{\mathbf{J}}^{SO}$ and Z.

Under the assumption that $\mathbf{T}_{\mathbf{J}}^{SO}$ and Z are both positive, the only admissible case is (c). It identifies an unstable node if the eigenvalues are real and an unstable focus if the eigenvalues are complex and conjugate. Locally explosive dynamics is highly likely.

Then, assume that on $\mathbf{T}_{\mathbf{J}}^{SO}$ and Z are both negative and analyze the two possible cases:

1. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} \geq 0$, $i = 1,2$. The four eigenvalues are real and they can be:
   a. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates both conditions on $\mathbf{T}_{\mathbf{J}}^{SO}$ and Z.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is admissible only if $\phi_1 + \phi_2 < 0$ as $\mathbf{T}_{\mathbf{J}}^{SO} < 0$.
   c. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates condition on Z.

2. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} < 0$, $i = 1,2$. Look at the real parts only. Since the real part $a = -\frac{1}{2}\phi_i \neq 0$, $i = 1,2$, the eigenvalues are complex and conjugate and they can be:
   a. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates both conditions on $\mathbf{T}_\mathbf{J}^{SO}$ and Z.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is admissible only if $\phi_1 + \phi_2 < 0$ as $\mathbf{T}_\mathbf{J}^{SO} < 0$.
   c. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates Z.

Under the assumption that $\mathbf{T}_\mathbf{J}^{SO}$ and Z are both negative, the only admissible case is b), but only if $\phi_1 + \phi_2 < 0$. It identifies a stable saddle point that ensures monotonic local convergence.

Finally, assume that $\mathbf{T}_\mathbf{J}^{SO}$ and Z have opposite sign and distinguish two possible cases:

1. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} \geq 0$, $i = 1,2$. The four eigenvalues are real and they can be:
   a. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates both conditions on $\mathbf{T}_\mathbf{J}^{SO}$ and Z.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is admissible only if $\phi_1 + \phi_2 > 0$ as it ensures $\mathbf{T}_\mathbf{J}^{SO} > 0$.
   c. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is admissible only if $\mathbf{T}_\mathbf{J}^{SO} < 0$ and $Z > 0$.
2. $\Delta \equiv \phi_i^2 - 4\gamma^{-1} < 0$, $i = 1,2$. Look at the real parts only. Since the real part $a = -\frac{1}{2}\phi_i \neq 0$, $i = 1,2$, the eigenvalues are complex and conjugate and they can be:
   a. four positive roots. This case implies that $\phi_1 + \phi_2 > 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is excluded as it violates both conditions on $\mathbf{T}_\mathbf{J}^{SO}$ and Z.
   b. two negative and two positive roots. This case implies that $\phi_1 + \phi_2 \gtrless 0$ and $\phi_1 \cdot \phi_2 < 0$. This case is admissible only if $\phi_1 + \phi_2 > 0$ as it ensures $\mathbf{T}_\mathbf{J}^{SO} > 0$.
   c. four negative roots. This case implies that $\phi_1 + \phi_2 < 0$ and $\phi_1 \cdot \phi_2 > 0$. This case is admissible only if $\mathbf{T}_\mathbf{J}^{SO} < 0$ and $Z > 0$.

Under the assumption that $\mathbf{T}_\mathbf{J}^{SO}$ and Z have opposite sign, case (b) identifies an unstable solution as $\mathbf{T}_\mathbf{J}^{SO} > 0$. Locally-explosive dynamics is highly likely. Case (c) identifies a stable node for real eigenvalues and a stable focus for complex and conjugate eigenvalues, and therefore it ensures damped convergence to the steady state.

References

DE LA CROIX D. 1996. The dynamics of bequeathed tastes, *Economics Letters*, Vol. 53, No. 1, pp. 89-96.

DE LA CROIX D., MICHEL P. 1999. Optimal growth when tastes are inherited, *Journal of Economic Dynamicsand Control*, Vol. 23, No. 4, pp. 519-537.

DIAMOND P. 1965. National debt in neoclassical growth model, *American Economic Review*, Vol. 55, No. 5, pp. 1126-1150.

FERSON W.E., CONSTANTINIDES G.M. 1991. Habit persistence and durability in aggregate consumption: Empirical tests, *Journal of Financial Economics*, Vol. 29, No. 2, pp. 199-240.

LAHIRI A., PUHAKKA M. 1998. Habit persistence in overlapping generations economies under pure exchange, *Journal of Economic Theory*, Vol. 78, No. 1, pp. 176-186.

MARINI G. 2015. Dynamics in an OLG model with nonseparable preferences, *Public Finance Research Papers* 15, Istituto di Economia e Finanza, DIGEF, Sapienza University of Rome.

MICHEL P., VENDITTI. A. 1997. Optimal growth and cycles in overlapping generations models, *Economic Theory*, Vol. 9, No. 3, pp. 511-528.

ORREGO F. 2014. Habit formation and indeterminacy in overlapping generations models, *Economic Theory*, Vol. 55, No. 1, pp. 225-241.

SUMMARY

Dynamics and stability in an OLG model with non-separable preferences

This paper presents sufficient conditions for existence and uniqueness of a steady state equilibrium in an OLG model with non-separable preferences and analyses the implications of such assumption for the local stability of the steady state equilibrium. The conditions for a stable solution are derived under the assumption that habits are transmitted both across and within generations. Under this assumption, the paper shows that monotonic convergence to the steady state is not always assured. The paper thus proves that also the optimal solution may be affected by instability and explosive dynamics, under particular conditions on the relevant parameters.

Giorgia MARINI, Sapienza University of Rome, giorgia.marini@uniroma1.it

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: Prof.ssa 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. Libraria Editrice Università di Padova"
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

ATTIVITÀ DELLA SOCIETÀ

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).
LVII	Metodi e misure politiche per una ripartenza sostenibile (Evento on-line, 27 maggio 2021).