The digitization of the private sector. A non-aggregative method to monitor the NRRP agenda at macro-area level

Susanna Traversa, Enrico Ivaldi

1. Introduction

2020 has represented a break from the economic and social policies adopted in Europe until the SARS-CoV-2 (Covid-19) emergency spread (Grasso et al., 2021). The containment measures introduced during the health crisis have produced a series of effects, including an acceleration of the digitisation both in social and in economic sphere (OECD, 2020). Within the latter in particular, Covid-19 has had a leverage function regarding private sector innovation, leading to the adoption of measures to implement the use of digital and technologies in order to ensure continuity in the production sector of goods and services (Casquilho-Martins and Belchior-Rocha, 2022). However, bringing attention to the Italian context, it is necessary to consider the critical issues related to the still-present digital divide between the northern and south-central areas of the peninsula. The digital divide, as well as digital illiteracy and infrastructural barriers, are the main obstacles that have slowed Italy's digital transition over the past few years with respect to the European scenario (Traversa et al., 2022; European Commission, 2021). It is precisely in the promotion of digitization policies that the European Union identifies one of the main drivers for a sustainable and resilient economic recovery, through which business continuity can be ensured despite lockdown policies (European Commission, 2020a,b). Based on European investment indications, great importance has been recognized by the Italian government to the theme of digitization by reserving for it targeted interventions within the first mission of the National Recovery and Resilience Plan (NRRP) for which EUR 49.2 billion has been allocated. Specifically, the NRRP commits Component 2 of Mission 1 to the strengthening of competitiveness within the private sector to be pursued by means of greater diffusion of digitization processes, technological innovation, and the strengthening of Industry 4.0 policies. The medium - to long-term recovery goals on which the NRRP is based need tools that can assess actual effectiveness on the Italian territory. Tools that can not only express the spread of digital business integration from a geographic point of view, identifying areas that show lower performance than the national context, but that can also monitor the progress of these interventions over time. In this study, a different synthesis methodology that makes use of a non-aggregative strategy was employed. The study is divided into sections where the main opportunities for developing an index to measure the effectiveness of the policies presented in the NRRP, as well as the traits and ramifications of using a non-aggregative strategy for the temporal study of socio-economic phenomena, will be outlined. Afterward, the outcomes of the index's application to the Italian context will be discussed.

2. Methodology

Given the complex and multi-dimensional nature of digitization processes, it is possible to approach their study through the construction of synthetic indices. In contrast to recent literature (Traversa et al., 2022; Benecchi et al., 2021; European Commission, 2021), the choice in this case fell on non-aggregative synthesis by means of the Partially Ordered Set (POSET). Among the main advantages of using a non-aggregative method over composite index construction is the possibility, by not carrying out the aggregation and weighting of

Susanna Traversa, University of Genoa, Italy, traversa.su@gmail.com, 0000-0001-5030-2021
Enrico Ivaldi, University of Genoa, Italy, enrico.ivaldi@unige.it, 0000-0001-6687-9378

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Susanna Traversa, Enrico Ivaldi, *The digitization of the private sector. A non-aggregative method to monitor the NRRP agenda at macro-area level*, © Author(s), CC BY 4.0, DOI 10.36253/979-12-215-0106-3.37, in Enrico di Bella, Luigi Fabbris, Corrado Lagazio (edited by), *ASA 2022 Data-Driven Decision Making. Book of short papers*, pp. 209-214, 2023, published by Firenze University Press and Genova University Press, ISBN 979-12-215-0106-3, DOI 10.36253/979-12-215-0106-3

variables, to limit the loss of information due to the flattering effect, which depends on the existence of incompatibilities between variables within a multi-dimensional system. Conversely, one of the main critical issues attributable to the POSET technique is related to computational aspects, as it requires the use of advanced statistical analysis software to enable its calculation. From a theoretical point of view, a Partially Ordered Set can be defined as a set in which there is no binary relationship between all pairs of profiles of which it is composed (Davey and Priestley, 2002; Fattore, 2017). The graphical representation tool used in the literature to represent the POSET system of relations is the Hasse diagram, according to two proprieties: (1) if $x \leq y$, the node - i.e. profile - y is placed above node x; (2) if $x < y$ then a "path" links node y to node x. By considering a pair of nodes, the one placed in the upper level is defined as being connected to the lower one according to a dominance relationship. The node (or nodes) with only descending relationships is said to be "maximum" (or maxima) of the POSET (Fattore, 2008). In order to address the synthesis of the digitization index the average height approach (avh) represents the most common method for the synthesis of multi-indicator systems through the study of POSETs (Alaimo et al., 2020; Fattore, 2017; Mazziotta and Pareto, 2020).

From a practical point of view, the synthesis vector is obtained by following a stepwise procedure: (1) Extract all line extensions of P by creating Ω_p; (2) For each element $p \in P \cdot$ and for each $l \in \Omega(P \cdot)$, assign a rank $r_l(p)$ of p in l, which represents 1+ the number of angle covers joining p to the maximum of l; (3) The third step is the calculation of the average $r(p)$ of $r_l(p)$ over $\Omega(P \cdot)$ for each $p \in P \cdot$.
The avr can be represented through a graph that shows the min-max avr range for each profile $(\perp \top)$. As in the case of aggregate indices, the POSET technique also allows for the study of a phenomenon over time, thus complying with the requirements that will be presented in the course of the research design from the point of view of the selection of synthesis methods (Alaimo et al., 2020). The development of a temporal POSET consists in the merge of two (or more) POSETs related to the years under study. Each statistical unit is measured with respect to variables referring to two (or more) different t times, calculating for each year the average rank of the POSETs. From a graphical standpoint, it is possible to obtain a visual representation of the temporal POSET by merging the corresponding Hasse diagrams. Following the merging of the POSETs and rebuild of the dataset, an intertemporal POSET is obtained which must again be subjected to the calculation of average height. Continuous recourse to avrcalculations could lead to a loss of information in POSET. To solve problems of comparability between nodes a possible solution should be the use of a reference system commonto the whole POSET - embedded scale - which will represent the benchmark through which the evolution of the POSET can be interpreted.

1. Research design

As a result of the analysis of the targets enshrined by the NRRP within M1.C2. "Digitization, Innovation and Competitiveness in the Production System", the variables were selected through the "Rilevazione sulle tecnologie dell'informazione e della comunicazione nelle imprese" (source: National Institute of Statistics – I.Stat) data-warehouse by following a formative approach (Diamantopoulos et al., 2008). Four variables expressing the digitization of the private sector were selected and presented below: (1) percentage of enterprises with fixed or mobile broadband connection (BBC); (2) percentage of enterprises using robots (ROB); (3) intra-muros research and development expenditure (thousands of euros at current prices) (R&S); (4) percent- age of enterprises that organized training courses in the previous year to develop or upgrade theICT/IT skills of their employees (FDS). The composition of the dataset needs some specification. The unavailability of access to I.Stat data with respect to ultrabroadband deployment in the macroareas led to the use of broadband connection data as a proxy. As for the 2020 figure for R&S spending, it was imputed from 2019 data, as it was not

available at the macroarea level but only nationwide. Variables were also selected based on spatial and temporal availability requirements. Specifically, the variables are presented on the basis of a geographical breakdown into macroareas: Northwest (NO); Northeast (NE); Center (CE); and "Meridione" (ME), which includes southern regions and islands. Italy (IT) was also considered as a comparison of macro-areas against the national average. In this way, it was possible to put more emphasis on the issue related to the digital divide and how it has evolved over the three-year period 2018-2020 considered. This time period makes it possible to capture the pre-pandemictrend of digital implementation in the productive sector by comparing it with the scenario realized in 2020, following the shocks produced by the pandemic. Computation of the results was conducted by means of the statistical software Rstudio and the package "parsec" (Fattore and Arcagni, 2014).

2. Discussion

Following the construction of the temporal POSET, the main results obtained from the application of the non-aggregative method are presented and discussed below. First, the Hasse diagrams of the individual years examined were constructed. As can be seen from Figure 1, the graphical representation of the 2018 and 2019 POSETs exhibit the same structure. IT, NE, NO are placed as maximal nodes, which are connected by a cover relation to the two lower nodes: CE and ME.

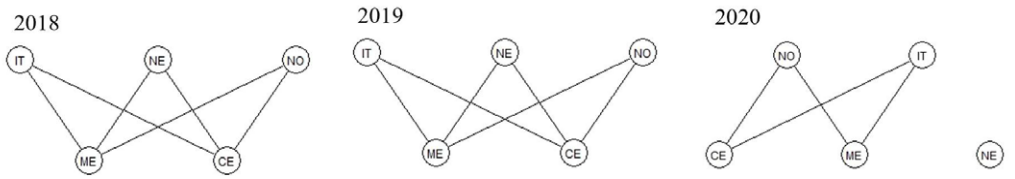

Figure 1: Comparisons between single years Hasse diagram. Period: 2018-2020.

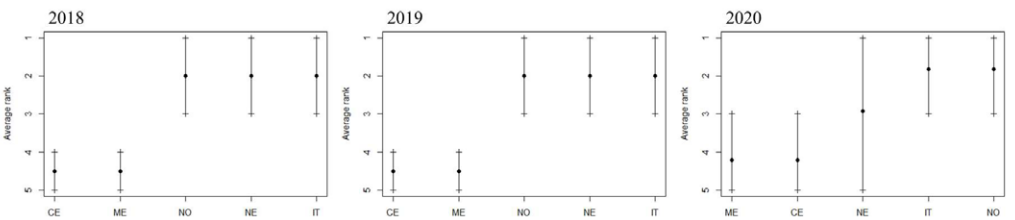

Figure 2: Comparisons between single years avr plot. Period: 2018-2020.

The dominance relationships between the profiles expressing northern and south-central regions are also confirmed following the calculation of the average ranks of the two-year pre-pandemic period (Figure 2). In the average rank plot, y-axis scale expresses the total number of observations sorted in descending order, attributing the best condition to the profile corresponding to rank 1. The points on the graph indicate the average value of the simulations obtained during the calculation of the average ranks, while the vertical bands express the variability of the profiles. A high range between the minimum and maximum value expresses the variabilitywith respect to the identification of a unique average rank for the profile. The CE macro area confirms the worst performance in the first part of the period, presenting together with ME an average rank fluctuation range between the fourth and fifth rank. On the other hand, the situation differs for NO, NE, IT where the distance between the two whiskers is greater and ranges from 1-3. IT has better values, positioning itself on the far extreme right of the graph, followed by NE and NO. A different scenario for 2020 is reported. During the pandemic year, only IT and NO are in the maximum positions, maintaining a dominance relationship with the lower nodes of CE and ME. More peculiar, however, is the case of the

NE profile, which stands outside the Hasse as an anti-chain, reporting no comparability in its digitization to the other macro areas (Figure 1). Moving the focus to the avr plot, there is a reversal in ranking between ME and CE, with an increase in variability in the average ranks. In contrast to the previous two years, NO reports an improvement ranking ahead of IT. NE also confirms in the avr plot an "anomaly" with respect to digitization in the four variables derived from the M1.C2 presented before, with a range of variation in the avr maximum. An analysis of the original data obtained from I.Stat shows that compared to the pre-pandemic two- year period, the elementary indicators considered do not show significant changes (positive or negative), with the exception of some macroareas that have suffered more in certain dimensions from the impact of Covid-19. As can be deduced from the preliminary study of the POSETs of individual years, the Northeast macro-area experienced the greatest fluctuations during 2020. If in the case of ROB no significant - albeit positive - changes are observed, BBC percentage for NE ranks below the national average as well as being the only area that experiences a re- duction in the percentage (not significant). However, a greatest change is experienced for the FDS dimension, which loses 6.6 percentage points in 2020 compared to 2019. As for NO - along with NE - it tends to perform better over the period, trending above the national average in all variables, although there is a slight deterioration in the FDS dimension between 2019 and 2020, offset by a 3.5 percentage point improvement in BBC. Finally, the CE and ME macro areas show values below the national average for each dimension with deterioration in FDS in both macro areas and in ROB for the "Meridione" area. After reconstructing the background of digitization in the private sector in Italy, the results that emerged from the construction of the temporal POSET are, below, addressed, for which three benchmarks expressing MIN, MEDIAN and MAX were calculated. The best performances are identified for NO_19, NO_20, NE_19, NE_20 and IT_19.

The Hasse diagram of the temporal POSET confirms an improvement despite Covid-19 for the macro-areas located in the northern part of the peninsula, although it achieves the result in ways that are not comparable with each other. Regarding national results, the nodes expressing national digitization show a deterioration for IT_20 with implementation of digital integration not comparable to what was achieved in previous years. In addition to the effect of Covid-19 that could contextualize the worsening of the national average, the incomparability with IT_18 and IT_19 could be attributable to the performance of NE_20. Lastly, as for ME and CE, the response to the shocks produced by Covid-19 has opposite effects. While ME after a positive trend in the pre covid two-year period observes a deterioration in digitization performance, CE shows an improvement in 2020 in line with what is realized for IT_20. The impact produced by Covid-19 is evident from the average height ranking (Table 1).

The best ranking is attributed to NE_19 followed by NO_19 and NO_20, underscoring on the one hand the better digital implementation within the private sector in the northern regions, and on the other hand a greater sensitivity with respect to the effects produced by Covid-19 on the NE macro area, which for the year 2020 reports a ranking lower than the IT 20 average. Coherent, on the other hand, are the performances of the southern regions, for which performance is observed to be positioned below the benchmark expressing the MEDIAN, with an improvement over the three-year period more rewarding for CE, which ranks at a higher Hasse diagram level than profiles in central and southern Italy (with the exception of NE_19) and digitization in linewith the national average IT_20.

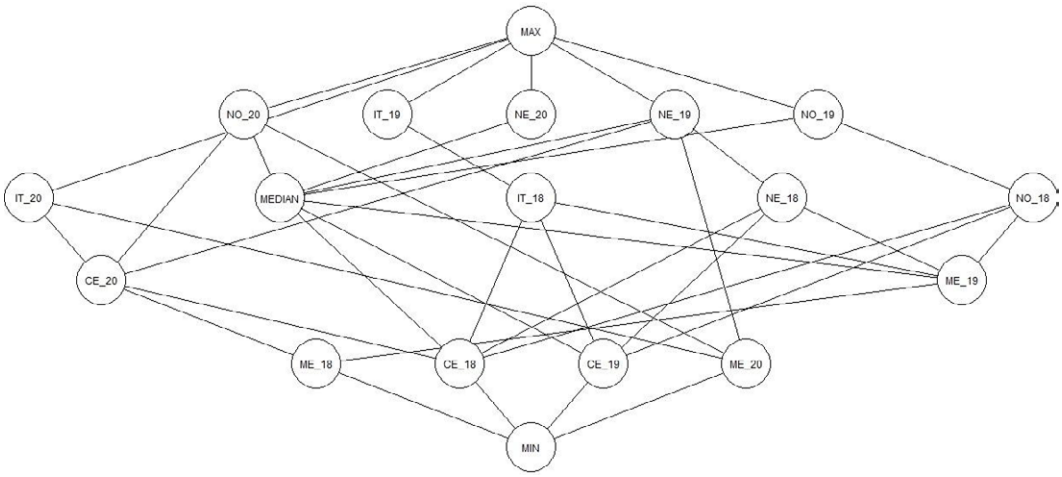

Figure 3: Temporal Hasse diagram.

COD.	2018	2019	2020
IT	9,792	13,902	13,463
NE	10,204	14,758	13,216
NO	10,035	14,386	14
CE	3,611	4,006	7,986
ME	2,847	4,926	6,407
MIN	1	1	1
MEDIAN	8,459	8,459	8,459
MAX	18	18	18

Table 1: Average height distribution of the Temporal POSET with benchmarks.

5. Conclusion

The development of quantitative tools to monitor thedigitization goals of the private sector within the scope of the NRRP goals, is a timely issue worthy of further investigation. The still pronounced existence of the digital divide and digitalilliteracy, represent a major obstacle to proper digital integration within enterprises, highlighting the negative impact produced by the rapid spread of digital in some areas of the country due to Covid-19. The POSET technique allow to contextualize rank positioning based on order relationships, which leads one to lean toward further exploration in the use of non-aggregativeapproach.

Indeed, by cross-referencing the information that can be obtained from the trend of basic indicators in individual regions, with the POSETs of individual years and the temporal one, it is possible to gain a greater understanding of the ways in which NRRP digitization goals are carried out in relation to territory and time. This provides an enhancing of the complexity ofthe phenomenon and not limiting it to a simplification as is occurring with aggregative synthesis techniques. Although the study is not free from limitations, due in part to the scarsity of available data inherent in digitization, it represents a possible starting point for the development of statistics assessing NRRP performance on a national scale. A further possibility to consider is the replicability of the study at the NUTS-2 territorial level, in order to highlight the performance of those regions that tend to show performance that is not in line with the performance of the macro-area to which they belong (Traversa et al., 2022).

References

Alaimo, L. S., Ciacci, A., Ivaldi, E. (2020). Measuring sustainable development by non-aggregative approach. *Social Indicators Research*, 1(22).

Alaimo, L. S., Arcagni, A., Fattore, M., Maggino, F. (2020). Synthesis of multiindicator system over time: A poset-based approach. *Social Indicators Research*, 1(23).

Benecchi A., Bottoni C., Ciapanna E., Frigo A., Milan A., Scarinzi E. (2021), La digitalizzazione in Italia: evidenze da un nuovo indicatore regionale, in *Questioni di economia e finanza*, eds. Banca d'Italia Occasional papers, n.662.

Casquilho-Martins, I., Belchior-Rocha, H. (2022). Responses to COVID-19 Social and Economic Impacts: A Comparative Analysis in Southern European Countries. *Social Sciences*, 11(2), pp. 36.

Davey, B. A., Priestley, H. A. (2002). Introduction to lattices and order. *Cambridge university press*.

Diamantopoulos, A., Riefler, P., Roth, K. P. (2008). Advancing formative measurement models. *Journal of business research*, 61(12), pp. 1203-1218.

European Commission. (2020a). 2030 digital compass: the European way for the digital decade.(*Available on Eurlex*).

European Commission. (2020b). European moment: Repair and prepare for the next generation. *Communication from the commission to the European parliament, the European council, the council, the European economic and social committee and the committee of the regions*, Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0456from=EN (Available on line).

European Commission. (2021). Digital economy and society index (DESI) 2021. *Thematic chapters,* Retrieved from https://ec.europa.eu/newsroom/ dae/redirection/document/80563 (Available online)

Fattore M. (2008). Hasse diagrams, poset theory and fuzzy poverty measures. Hasse Diagrams. Poset Theory and Fuzzy Poverty Measures, *Rivista Internazionale Di Scienze Sociali*, 116(1), 63–75. http://www.jstor.org/stable/41625201

Fattore M. (2017). Synthesis of indicators: The non-aggregative approach, in *Complexity in society: From indicators construction to their synthesis*, eds. Springer, pp. 193-212.

Fattore, M., Arcagni, A. (2014). Parsec: An r package for poset-based evaluation of multidimensional poverty, in *Multi-indicator systems and modelling in partial order*, eds. Springer, pp. 317-330.

Grasso, M., Klicperova´-Baker, M., Koos, S., Kosyakova, Y.,Petrillo A., Vlase, I. (2021). The impact of the coronavirus crisis on European societies. What have we learnt and where do we go from here? Introduction to the COVID volume. *European Societies*, 23(sup.1), pp. S2-S32. Routledge.

Italian Government. (2021). Piano Nazionale di Ripresa e Resilienza. *Presidency of the Council of Ministers*. Retrieved from https://www.governo.it/sites/governo.it/files/PNRR0.pdf.

Mazziotta, M., Pareto, A. (2020). *Gli indici sintetici*. Giappichelli.

OECD. (2020). OECD digital economy outlook 2020. *OECD Publishing*.

Traversa, S., Ciacci, A., Ivaldi, E., Gonzalez-Relanõ, R. (2022), Measuring the Digital Gap in Italy: A NUTS-2 Level Index, in *Globalization, Income Distribution and Sustainable Development*, eds. Chandra Das, R. Emerald Publishing Limited, Bingley, pp. 265-281.