A quantitative study to measure the family impact of e-learning

Cristina Davino, Marco Gherghi, Domenico Vistocco

1. Introduction

The Covid-19 emergency has forced universities around the world to transfer teaching activities online. Even if online teaching allowed to carry out the planned teaching activities, it is necessary, in retrospect, to evaluate the impact of this teaching method on the different types of students, in terms of preparation, characteristics and social background. The switch from offline to online learning caused by Covid-19 is expected to exacerbate existing educational inequalities penalising more vulnerable students. The social and economic conditions of families have a major influence on the e-learning experience because less advantaged students are less likely to have access to relevant learning digital resources (e.g. laptop/computer, broadband internet connection) and less likely to have a suitable home learning environment (e.g. a quiet place to study or their own desk) (Di Pietro et al., 2020). Furthermore, according to the 2020 European Commission's annual report on the levels of digitalisation achieved by the various member states[1], Italy ranks 25th among the 28 EU Member States[2].

The aim of this paper is to analyse whether and how the distance learning activities impacted on the students' families both in terms of the organisation of spaces and daily rhythms and from an economic point of view, having required additional expenses. The study is based on the analysis of data collected at the University of Naples Federico II in June 2020. More than 19,000 students took part in a survey, carried out to monitor distance learning activities and perceptions. The paper is organised into two sections. In the first, a factorial method is exploited to obtain a composite indicator measuring the family impact of distance learning. Then, we try to explain if the family impact takes different forms and intensity depending on the students' characteristics, the availability of computer equipment and the type of teaching used. Finally, quantile regression allow to differentiate the study of effects for different levels of family impact. Some considerations on the distance learning experience in terms of family impact and the evaluation on the preferred teaching method for the future are also enclosed.

2. Measuring family impact of E-learning

The measurement of family impact is carried out following the classic steps used in research methodology for the measurement of a multidimensional and abstract concept (Freudenberg, 2003), i.e. a latent variable not directly observable and expressed as a combination of several components. The construction of such a latent variable, often referred to as Composite Indicators (CI), is done through the use of an aggregation method appropriate to the nature of the observed variables (Lebart et al., 2000).

The study proposed in this paper is based the survey conducted by the University of Naples considering only students who attended at least one distance learning course in the 2019/2020

[1] $https://ec.europa.eu/digital-single-market/en/digital-economy-and-society-index-desi$

[2] The data in the report refer to 2019 and therefore do not take into account all the initiatives taken by governments to counter pandemic.

Cristina Davino, University of Naples Federico II, Italy, cristina.davino@unina.it, 0000-0003-1154-4209
Marco Gherghi, University of Naples Federico II, Italy, marco.gherghi@unina.it, 0000-0003-0363-145X
Domenico Vistocco, University of Naples Federico II, Italy, domenico.vistocco@unina.it, 0000-0002-8541-6755

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

Cristina Davino, Marco Gherghi, Domenico Vistocco, *A quantitative study to measure the family impact of e-learning*, pp. 103-107,

DOI 10.36253/978-88-5518-304-8.21, in Bruno Bertaccini, Luigi Fabbris, Alessandra Petrucci, *ASA 2021 Statistics and Information Systems for Policy Evaluation. Book of short papers of the opening conference*,

published by Firenze University Press (www.fupress.com), ISSN 2704-5846 (online), ISBN 978-88-5518-304-8 (PDF), DOI 10.36253/978-88-5518-304-8

academic year. The sample of responses received reflects the distribution of the student population by degree course. A special section of the questionnaire was dedicated to the detection of the family impact of the e-learning experience. The following is a list of questions relating to this section with an indication of the percentage of answers for each category (labels that will be used in the tables and graphs in italics, percentages in parentheses):

- *Place* - When attending a distance learning class, you were mainly: *anywhere* (12.9), *alone* (78.6), *with_relatives*(8.3);
- *Expenses* - Did you and/or your family incur any expenses in order to follow the distance learning lessons?: *equipment* (17.1), *equipment&network* (3.2), *network* (11.1), *no* (67.0), *other* (1.6);
- *use_equipment* - To follow lessons at a distance, the device you mainly used was: *exclusive* (67.5), *shared_other* (10.5), *shared_teaching_working* (22);
- *family_habits* - Distance learning courses affected your normal habits and those of the rest of the family.: *strongly_agree* (19.3), *agree* (26.3), *neither_agree_nor_disagree* (29), *disagree* (14.7), *strongly_disagree* (10.7).

Since the indicators are all qualitative and/or ordinal, a multiple correspondence analysis (MCA) was used to provide a CI measuring the family impact of e-learning. MCA can be considered one of the best known and most effective tools for the simultaneous analysis of questionnaire data. Proposed in the late 1970s by J.P. Benzècri for the case of two qualitative variables (Binary Correspondence Analysis), it has been extended to the case of many qualitative variables. MCA is a Factor Analysis that allows to identify a reduced number of variables (also called factors or latent variables or CIs) as a linear combination of the original variables. Each CI is able to explain a part of the variability of the phenomenon.

The first factor obtained from the MCA accounts for 91.34% of the total variability. It can therefore be considered an adequate measure of the Family Impact of the experience of E-Learning (from now on FIEL). The distribution of FIEL (Figure 1, left-hand side) shows a phenomenon almost equally distributed around the average value (represented by the value 35.58[3]) even if with different characteristics in the two parts of the distribution: students with a low family impact are more concentrated, while the right tail of the distribution is more dispersed. This is a signal of greater heterogeneity among those who have a family impact above the average.

The interpretation of the FIEL indicator can be deepened by considering also the contributions of the categories on the first factor, not focusing only the coordinate represented by the indicator itself. The contribution of a category to the explanation of a factor is provided by the product of the weight of the category, represented by its frequency, and the square of the coordinate of the category on the factor. Indeed, in MCA categories with vey low or high coordinates do not necessarily contribute to the explanation of the factor itself, because if they had a very low frequency, they would have a very low contribution. Similarly, categories that are more "central", but with a very high frequency, may have an important contribution to the explanation of the factor. The joint visualisation of coordinates and contributes (Figure 1, right-hand side) highlights that students who predominantly experienced a quiet e-learning experience without changing family habits (they already had all the equipment available for their exclusive use) are separate from students who were forced to share both the workstation and the device with family members engaged in smart working or other learning activities. This second group of students is forced to study in makeshift places, sometimes with other family members and distance learning has also affected their families financially.

[3]The indicator has been rescaled in the range 0-100.

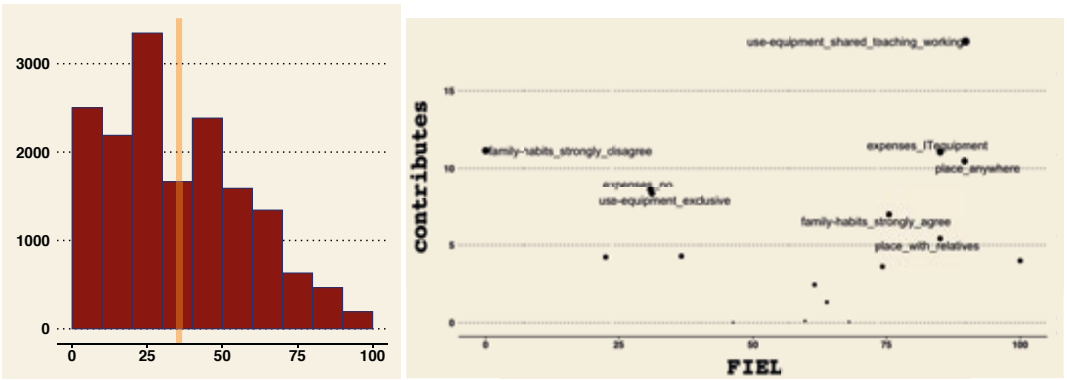

Figure 1: Distribution of the family impact of the experience of e-learning (left-hand side) and scatter plot of the categories measuring the FIEL according to the MCA coordinates and contributes (right-hand side)

3. Explaining family impact of E-learning

The interpretation of the FIEL indicator can be deepened by considering additional variables that did not contribute to its determination and that concern both personal characteristics of students, issues related more specifically to the availability of computer and network equipment (IT equipment) and also to the modality of distance learning. The former features are represented in the upper panel of Figure 2 while the latter in the bottom panel. Each point is located at the average values of FIEL in the correspondent category, the size being proportional to the frequency[4]. The vertical line represents the FIEL average. The family impact seems stronger (higher than the general average) for female students in the first years of the university experience. As might be expected, a wi–fi connection and a mobile study station (linked to the use of smartphones and tablets) can explain more complicated family situations.

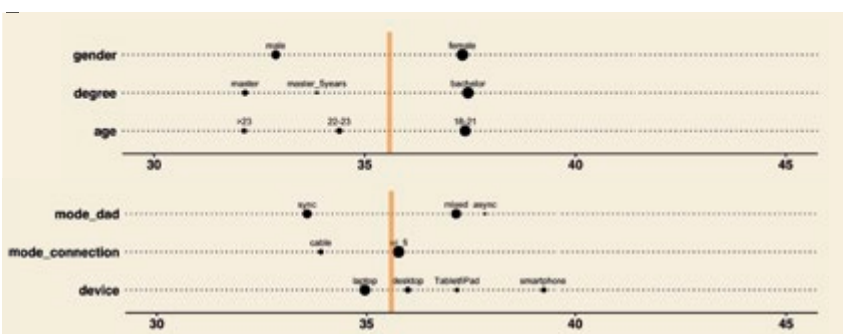

Figure 2: FIEL averages according to socio-demographic features and type of IT equipment

A comparison among the average values of FIEL does not allow to capture possible differences in the impact of the considered variables for different levels of family difficulties. Quantile regression (Koenker and Basset, 1978) allows us to complement the results of a classical OLS regression by exploring the effects of the regressors on the entire distribution of FIEL. In fact, although the number of quantiles that can be explored is theoretically infinite, it is shown that a sufficiently dense grid can be enough to reconstruct the entire dependent variable (Davino et

[4]For each variable considered, the averages are significantly different.

al., 2013). Nevertheless, in many cases we explore a small number of quantiles that represent parts of the distribution important for the particular analysis. In Figure 3, QR coefficients equal or greater than the conditional median are graphically represented for the different considered regressors. The horizontal axis displays the different quantiles, while the effect of each feature holding the others constant is represented on the vertical axis. The horizontal solid lines show the OLS results while the piecewise lines refer to the coefficients at different quantiles. The aim is to graphically catch the coefficient trends moving from lower to upper quantiles.

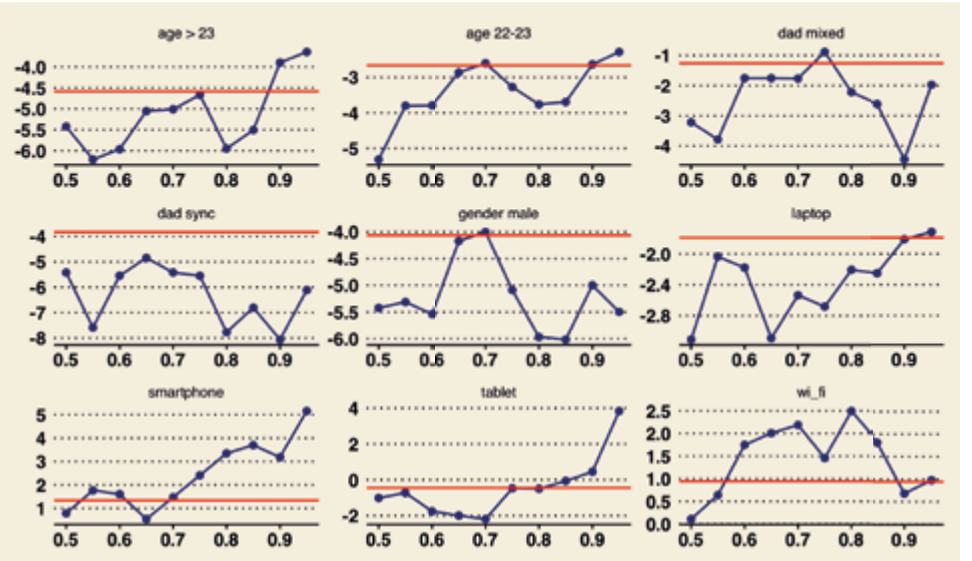

Figure 3: OLS (horizontal solid lines) and QR (piecewise lines) coefficients

Coefficients have been estimated for a sequence of quantiles from 0.5 to 0.9 with a step of 0.5. It was decided to explore the results of only the top 50% of the distribution as the aim is to investigate the situations of discomfort in order to understand what levers can be used to intervene. By the way, in the remaining part of the distribution (students with FIEL below the median) the effects of the regressors considered are practically null. A positive trend of the quantile curves emerges from the plot. This correspond to an increased variability of the FIEL variable (i.e. an increased difference between for instance the 25% and 75% conditional quantiles) with increasing values of the regressor or when the category changes. The interpretation of the results must take into account, apart from possible fluctuations in the values of the coefficients at the different quantiles, the sign of these coefficients and the possible presence of patterns from the lowest to the highest quantiles. For example, the negative effect of age on FIEL is less amplified in cases where the family impact is very high. The increasing trend suggests that this effect is gradually disappearing. More interesting is the interpretation of the results concerning the device used for distance learning (the reference category for the regression is desktop). In particular, the use of a smartphone compared to a fixed location has a consistently positive and increasingly strong effect moving towards the top of the distribution. As regards the use of Tablet/Ipad, the sign is even reversed starting from quantile 0.85. In addition to the above information, it should be noted that all the coefficients are always significant, with the exception of tablets and mixes, which are never significant, and wi–fi and smartphones, which contribute significant coefficients at the top of the distribution, at quantile 0.65 and 1uantile 0.85 respectively.

The results shown in this paper, although in many cases expected, allow to quantify and

visualise relationships among different elements that can contribute to highlight heterogeneity in the conditions and characteristics of students, an element that, in non-emergency conditions, is ignored when the same teaching strategies are adopted for all the students. Moreover, a complete understanding of a phenomenon cannot be achieved without measuring it. In this sense the results here illustrated can provide a quantitative measure of a multidimensional and abstract concept, the family impact of e-learning. The use of quantile regression allows to explore if student characteristics or IT equipment have different effect among those who have suffered a stronger family impact.

Looking to the future, students' preference for the different teaching modes changes according to the family impact of the experience. The boxplots in Figure 4 show the distribution of FIEL in the group of those who would prefer lessons exclusively at a distance (online), who believe that they can still benefit from an appropriate combination of the two modes (mixed) or who would prefer a total return to normality (onsite). There is an increase in the FIEL quartiles from the online category to the mixed and then onsite category, a sign that lived experience influences, hopefully only in part, the vision of the future.

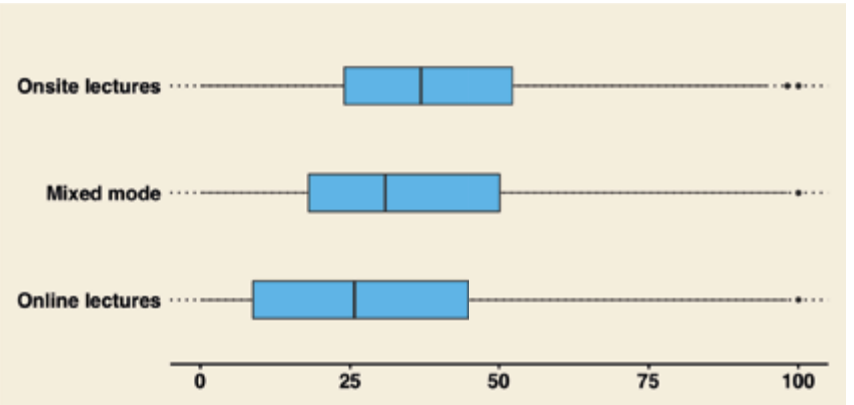

Figure 4: FIEL distribution according to the future vision of the students

References

Davino, C., Furno, M., Vistocco, D. (2013). *Quantile Regression: Theory and Applications*. Wiley, Chichester, UK.

Di Pietro, G., Biagi, F., Costa, P., Karpiski Z., Mazza, J. (2020). *The likely impact of COVID-19 on education: Reflections based on the existing literature and international datasets.* EUR 30275 EN, Publications Office of the European Union, Luxembourg.

Freudenberg, M., (2003). *Composite Indicators of Country Performance: A Critical Assessment*. OECD Science, Technology and Industry Working Papers 2003/16, OECD Publishing.

Koenker R., Basset G.W. (1978). Regression Quantiles. *Econometrica*, **46**, pp. 33–50.

Lebart, L., Morineau, A., Piron, M. (2000). *Statistique Exploratoire Multidimensionnelle*. Dunod, Malakoff Cedex, FR, pp. 181–184.