



IX SEMINAR

"Data from and for educational systems:
tool for research and teaching"



Andrea di Bonaiuto (Andrea Da Firenze), Allegory of Sciences, 1365-68

Rome, 17 - 19 October 2024

Under the patronage of



INDEX

INTRODUCTION	2
ABSTRACT	4
AUTHORS	123

INTRODUCTION

The Seminar "Data from and for educational systems: tools for research and teaching", now at its nine edition, has become in recent years an opportunity for meeting and discussion on the use of National Surveys INVALSI results and, in general, on the world of evaluation and school. Evaluation does not only mean standardized monitoring of learning levels, but also evaluation and comparison of its functions and potential in relation to the practices and tools through which the school system is able to carry out its educational and formative mission. This year, keynotes of national and international experts were organized on issues related to the evaluation of the education system and the use of data to support school policies.

One of the objectives of the seminar was to bring the world of scientific research and schools closer together in order to build a privileged and concrete space in which enriching the debate by sharing ideas and experiences between education and research stakeholders. The variety of topics improve a multidisciplinary approach to evaluation in the educational and school field, trying to give an account of the contribution that schools can provide to society by making possible the development of knowledge and skills.

The Seminar is organised by the research group of Area 2 - Statistical Service: Patrizia Falzetti (Manager), Paolo Barabanti, Andrea Bendinelli, Leonardo Boulay, Emiliano Campodifiori, Michele Cardone, Federica Colli, Ettore De Sossi, Alessandro Gaeta, Paola Giangiacomo, Patrizia Giannantoni, Pierangelo Grosso, Jana Kopečna, Fabrizio Lasorsa, Giuseppina Le Rose, Francesca Leggi, Lorenzo Maraviglia, Michele Marsili, Giancarlo Mastrone, Marcello Napoli, Fiammetta Noccioli, Carlo Palmiero, Monica Papini, Daniele Rowlett, Antonio Severoni, Massimo Smiraglio, Agnese Spoladore, Valeria F. Tortora.

Contacts: uff.statistico@INVALSI.it

The opinions and contents expressed within the initiative are the exclusive responsibility of the speakers and are in no way attributable to the INVALSI.

ABSTRACT

SESSION 5. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAs) METHODS AND RESULTS

ORGANIZER: UNIVERSITY OF KU LEUVEN - UNIVERSITY OF BATH

COORDINATOR: ANDRES SANDOVAL-HERNÁNDEZ

OCTOBER 17TH: 2.00 P.M. – 4.00 P.M. {ROOM 2 – RESEARCH 1}

Using INVALSI Data to Explore Beginner L2 English Learners' Performance while Reading to Grasp the General Meaning of a Text

Francesca La Russa – Clelia Cascella

Introduction

Second language (L2) reading comprehension is a multidimensional construct involving many cognitive, metacognitive, evaluative and learning processes that are difficult to conceptualize univocally (Grabe & Yamashita, 2009). This study draws on Khalifa & Weir's (2009) model of reading, which distinguishes between lower-level cognitive processes (word recognition, lexical access, syntactic parsing, establishing propositional meaning at clause and sentence level) and higher-level cognitive processes (inference, building a mental model, creating a text-level structure, creating an intertextual representation).

Lower-level processes - that are more strongly associated with linguistic knowledge (Cohen & Upton, 1998) - are usually automatically performed by proficient readers. However, beginner L2 learners, due to their limited proficiency, may struggle to automate these processes, thus expending significant cognitive resources and hampering more complex operations (Jenkins et al., 2003). This perspective aligns with the theoretical framework of the PISA 2025 Foreign Language Assessment survey (OECD, 2021:44):

From Levels Pre-A1 to A2, the focus on short, simple texts of a concrete nature suggests that these learners are primarily able to cope with the lower-level reading processes (those up to establishing propositional meaning), with the main purpose employed being careful reading at the local level due to their limited grammatical and lexical knowledge.

The INVALSI English tests assess the reading comprehension of learners ranging from PreA1 to B2 proficiency level across various reading focus: (I) reading for specific information and important details (SIID); (II) reading for main ideas and supporting details (MISD); and reading for the global understanding of the text (GIST). The latter requires higher-level cognitive processes aimed at establishing a representation of the meaning and overall structure of the text (Weir & Khalifa, 2008).

Object, aims and research hypothesis.

In this study, we observe the responses given by Grade 5 students to the INVALSI Reading test aimed at measuring Pre-A1 and A1 proficiency levels. We will examine the response behavior of students in relation to the different reading focuses, with special attention on the student's ability to read for gist. The aim is to verify if, for the previously mentioned reasons, this type of reading proves to be more difficult than others for beginner learners.

Data

To control for the research hypothesis, the responses of students in the sample classes to the INVALSI English achievement tests - administered in primary school at Grade 5 (G5) from the 2017-2018 school year to the 2022-2023 school year (included) - were analysed.

The INVALSI tests constructed for Grade 5 students are linear, paper-and-pencil tests. Each test contains 5 tasks, each consisting of a textual stimulus and an average of 6 items.

Students' reading ability is conceptualised as a unidimensional construct and operationalised through the selection of cumulative, locally independent items. Empirical verification of these characteristics is provided by deploying the Rasch model (1960/1980).

Methods

The Rasch model is one of the most widely used analytical approaches in educational surveys, especially in large-scale surveys because - if the theoretical assumptions of the model are empirically verified - Rasch estimates guarantee measurement invariance (e.g., Engelhard, 2013) and thus the comparability of results

between groups of students, between item groups, and between groups of students with groups of items. Measurement invariance is the property of the Rasch model used in the present study. Therefore, after empirically demonstrating the goodness of fit of the data to the model, Wright's maps will be presented. In addition, for further investigation of the results, the item characteristic curves (ICCs) will be presented and discussed.

Following this initial phase of analysis, the data collected by INVALSI will then be analysed using hierarchical regression to understand whether, and if so to what extent, the organisation of items into tasks (each related to a specific textual stimulus) may have an effect on their estimated difficulty.

Phase 1. The idea behind the Rasch model is simple. The probability of answering an item correctly is a function of the student's relative ability, i.e. his or her intrinsic ability relative to the difficulty of each item: the greater the student's ability relative to the difficulty of the item, the greater his or her probability of giving a correct answer.

According to this logic, both students and items can be ordered along the latent trait (ordered, that is, in relation to the property being measured) according to the amount of the trait they possess. Wright's maps show the ordering of subjects along the latent trait: the greater the amount of the trait possessed by the subjects (or items), the higher the positioning of the subjects (or items) along the latent trait represented by the map. By reading the map, it will then be possible to see at a glance which items were found most difficult by the students to whom they were administered.

The characteristic curves generated by the model are then used as an interpretation tool. For each test item, the Rasch model is able to produce a graph (the ICC) of the probability of a correct response in relation to the respondent's intrinsic ability in the trait being measured. The characteristic curve is then drawn for each of the items administered. In particular, the ICCs measuring the ability to read a text in order to grasp its overall meaning (GIST focus) are examined for diagnostic purposes.

Phase 2. The items in the INVALSI Reading test are divided into tasks, each of which relates to a textual stimulus. The organisation of items within the test is therefore hierarchical: items within tasks. Using a hierarchical regression model (Hox, 2010), we will therefore explore the (possible) relationship between the hierarchical organisation of the items within the test and the psychometric functionality of the test items.

Preliminary results

A preliminary analysis of the INVALSI data showed that in each of the years of administration examined, the items with a GIST focus are always at a medium level of difficulty (around zero, on Wright's map). In each of the years of administration, they are always scaled along the same part of the latent trait, thus forming a homogeneous block in terms of difficulty. On the other hand, items with a different focus (MISD or SIID) are scaled on different points of the latent trait and their positioning along the latent trait shows no regularity over time.

Such a preliminary result thus seems to indicate that for items with a GIST focus, the cognitive load is more or less always the same (and thus intrinsically linked to the reading strategy), whereas for items with a MISD or a SIID focus, the factors responsible for the difficulty of the items do not seem to be strictly linked to the reading strategy, but involve other aspects somehow associated, for example, to the wording of the items or the characteristics of the text stimulus.

Guidelines for future research

The results presented in this study should be seen as a starting point for further research. Indeed, it is well known from the relevant literature that the difficulty of an item depends on a number of factors. For example, an initial research development could be aimed at improving the interpretability of the results by investigating the relationship between the characteristics of a stimulus or item and its psychometric properties. For example, once these factors have been identified, specific analytical techniques (e.g. the linear logistic test model - Fisher, 1973) can be used to express item difficulty as a (linear) combination of these factors, thus producing results of interest both to educational policy-makers and practitioners (as they will improve the interpretability of educational results) and to researchers and test developers (as they will make explicit the relationship between item difficulty and text/item characteristics).

Keywords: English as a Foreign Language, L2 reading comprehension, gist, INVALSI data, item analysis

Exploring Creativity Through Standardized Tests: The PISA Framework for Assessing Creative Thinking and an Initial Analysis of the Results of Italian Students

Margherita Emiletti – Riccardo Pietracci

Exploring Creativity Through Standardized Tests: The PISA Framework for Assessing Creative Thinking and an Initial Analysis of the Results of Italian Students

This contribution aims to examine the results obtained by fifteen-year-old Italian students, mainly in their second year of upper secondary education, participating in the 2022 edition of the Programme for International Student Assessment (PISA) in the Creative Thinking tests.

For the first time, in 2022, PISA introduced the assessment of creative thinking as a content area. Sixty-four different countries, including a national representative sample of Italy, participated in this assessment.

Over the years, PISA has consistently added innovative assessment domains to the traditional content areas, as demonstrated by the inclusion of skills such as financial literacy in 2012, which subsequently became a recurring area of assessment, problem-solving in 2015, and collaborative problem-solving in 2018.

Creativity is a key factor in the evolution of human societies and in the generation of innovative knowledge, upon which international organizations and communities rely to address emerging challenges (OECD, 2010). Despite entrenched preconceptions, the OECD has emphasized that every individual has the potential to develop their own creative thinking (OECD, 2017). This capability is not simply the emanation of an unexpected idea but rather a tangible skill, grounded in knowledge and practice, enabling individuals and groups to achieve better results, especially in challenging contexts. Literature analysis supports the notion that training in creative thinking not only enhances metacognitive, inter- and intra-personal skills, and problem-solving but also promotes the development of individual identity, improving academic outcomes, and fostering professional success (Stenberg, 1999; Sawyer, R. K. (2012)). The investigation focuses on all forms of creativity expressed in everyday life through activities such as solving practical problems, producing personal works, or devising innovative solutions. It does not necessarily require specialized expertise and can be developed through practice and common experience.

Based on this approach, the process of developing the Framework for Assessing Creative Thinking began in 2017, involving international experts from participating countries in PISA. The coordination of this work was entrusted to ACT, which is a mission-driven, nonprofit organization dedicated to helping people achieve education and workplace success. Headquartered in Iowa City, Iowa, ACT is trusted as a leader in college and career readiness, providing high-quality assessments grounded in nearly 60 years of research and it was engaged by OECD for the construction of the creative thinking framework and cognitive tests. The final version of the Framework was completed in September 2022.

The Creative Thinking assessment within PISA is based on computer-based testing. All the questions in the test are open-ended, so there are multiple, and potentially infinite, ways to demonstrate possession of creative thinking skills. For the vast majority of questions (approximately 92%), scoring student responses relies on human judgment, detailed scoring rubrics, and well-defined coding procedures. These responses can take the form of graphics, presented as visual products created through a specifically designed digital interface for the test, or textual, with variable required lengths of written elaboration, ranging from a few lines to several paragraphs. The remaining 8% of questions consist of simulations solved through a trial-and-error testing interface, with automatic scoring assigned for various possible resolution strategies.

Building upon these premises, the proposed contribution is structured into three parts.

The first part aims to provide an overview of the structuring of the assessment domain of creative thinking in PISA, starting from the defined scope for the detection of the domain, articulating the description of the construct according to the three domains of competence - which include both divergent cognitive processes (the ability to generate different ideas and the ability to generate creative ideas) and convergent cognitive processes (the ability to evaluate others' ideas and identify improvements to those ideas) - and the four contexts in which creative thinking is enacted (written expression, visual expression, social problem-solving, and scientific problem-solving).

The second part presents the sample structure and describes the assessment and main data collection methods in PISA 2022, with reference to the coding process and scoring criteria for responses regarding different aspects of competence. If feasible within the scope of the contribution, it aims to complement this part with a couple of released test examples and corresponding examples of real student responses provided and coded.

The third part presents the results of Italian fifteen-year-olds in creative thinking in PISA 2022 in an international context, comparing them with the results of other relevant countries, and then analyzing them in the national context regarding strengths and weaknesses on measured aspects (which could reflect differences in the importance attributed to these domains by the school), based on some characteristics of the national sample (gender differences) and for some selected characteristics of students (sense of creative self-efficacy and expectations for career paths).

The conclusions of the paper will propose some insights into the profile of Italian students who are most inclined to creative thinking, taking into consideration some of their characteristics usually not considered in the national report, such as, for instance, the expected exercise of creative thinking in their expected profession as adults.

Keywords: OECD PISA, ILSA, Creative Thinking, Creativity, Results, Gender differences, Italy

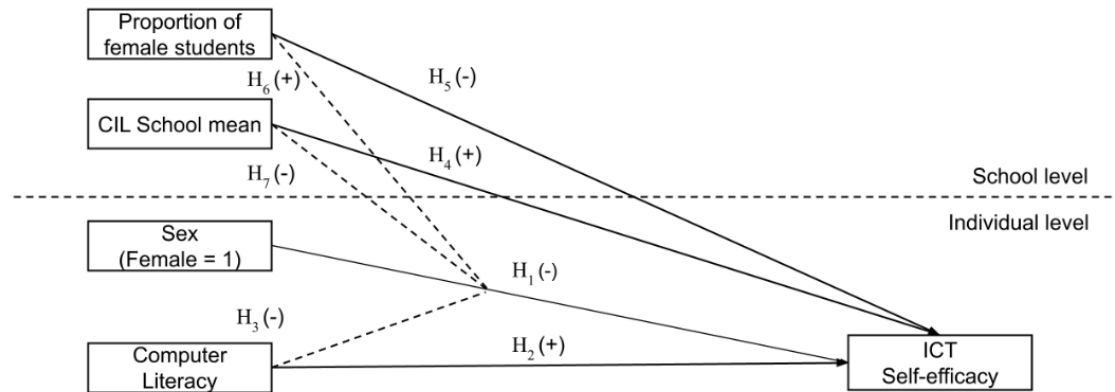
Do female students underestimate their technological abilities? Gender differences in general and specialized digital self-efficacy at school

Juan Castillo – Daniel Miranda – Nicolás Tobar

With the proliferation of digital technologies, educational systems worldwide have undergone a deep transformation, with digital literacy emerging as a crucial skill for navigating modern society. Digital skills encompass a range of competencies, including but not limited to computer literacy, coding, internet navigation, and critical thinking in digital environments. In such a context, proficiency in digital skills is increasingly recognized as fundamental component of education success and future career opportunities (Hooley & Staunton, 2020; Mahmud & Wong, 2022). However, despite the widespread integration of technology in classroom instruction, significant disparities persist in the acquisition and mastery of digital skills among school-age children, contributing unequal access and outcomes that hinder the ability of marginalized groups to fully participate and succeed in an increasingly digital society (Dodel, 2023). A primary factor that generates gaps in digital skills is socioeconomic status, where children from lower-income families often face barriers such as limited access to reliable internet connectivity, outdated technology infrastructure in schools, and insufficient resources for digital learning tools (Butcher & Curry, 2022; Mulyaningsih et al., 2021). Additionally, disparities in parental education and involvement further exacerbate these challenges, as children from families with higher levels of parental education tend to have greater exposure to technology and receive more support for developing digital skills (O'Hara, 2011). Furthermore, cultural stereotypes and gender norms can also influence the types of digital skills encouraged and valued among boys and girls, leading to differential opportunities for learning and advancement in certain domains (Clayton et al, 2009; Wong & Kemp, 2018). For instance, girls are less likely to pursue computer science and engineering due to stereotypes that these fields are male-oriented (Cheryan, 2015). Such stereotypes emerge very early and are reinforced in socialization environments as the school (Varoy et al., 2023). Research suggests that societal stereotypes and cultural norms often shape individuals' perceptions of their own capabilities and interests in digital domains, which can vary based on gender (Hargittai & Shafer, 2006). Girls, for instance, may internalize messages that associate technology and computer science with masculine traits, leading to lower confidence in their abilities and less motivation to pursue digital learning opportunities (Papastergiou, 2003). Conversely, boys may receive implicit or explicit encouragement to engage with technology, resulting in higher levels of self-efficacy and persistence in acquiring digital skills. Furthermore, socialization within peer groups and family environments can reinforce these gendered attitudes, creating a self-perpetuating cycle of differential participation and achievement in digital domains. The present paper deals with gender differences regarding digital skills at school age, focusing on self-efficacy. Relevant evidence in this regard is that girls tend to outperform their male counterparts when it comes to standardized test in technological abilities (Gebhardt et al., 2019; Tømte & Hatlevik, 2011; Tsai & Tsai, 2010; Punter et al., 2017; Fraillon et al., 2014). Therefore, our main research hypothesis is that girls tend to underestimate their technological abilities in contrast to boys. Furthermore, we argue that differences in self efficacy by gender are affected by classroom composition, whereby classrooms with a larger proportion of females would be associated with higher self-efficacy for them. Empirical studies have focused on two dimensions to explain the overestimation of men

and the underestimation of women. Firstly, on the differences in attitudes to technologies by gender, and second, on the school technological infrastructure effect on boys and girls (Colley & Comber, 2003; Tømte & Hatlevik, 2011; Cai et al., 2017; Campos & Scherer, 2023). However, few studies have taken into account the perspective of digital self-efficacy by gender at school age, let alone the classroom composition in terms of gender.

Hypothetical model



Data, variables & methods

Data and sample

The database to be used is the International Computer and Information Literacy (ICILS from now), study made by the International Association for the Evaluation of Educational Achievement (IEA) in the year 2018. The participants were students of eighth grade which were sampled using a random multistage stratified cluster design to be representative at the school and student/teacher level in all the countries measured. Schools were regarded as a "cluster," with all students and teachers nested within these clusters. Schools can therefore be referred to as the primary sampling units (level 2 in multilevel models), and teachers and students as the secondary sampling units (level 1 in multilevel models).

Variables

Student level

- Sex of the student.

The sex of the student is provided in the ICILS student database as dichotomic, where Girl is 1 and Boy 0.

- General and Specialized ICT Self-efficacy

The ICILS student data encompasses two indices of digital self-efficacy: one for general or basic applications and the other for advanced or specialized tasks. Both indices were constructed using different items from the same battery of questions. The phrasing of the battery was as follows: *'How well can you do each of these tasks when using ICT?'*

Due to the extended length of the labels for items related to ICT self-efficacy, we truncate them. Table 1 illustrates both the original and the summarized labels for the items comprising both indexes.

Item	Original label	Summarised label
<i>General Self-efficacy</i>		
is2g27a	Edit digital photographs or other graphic images	Edit graphic images
is2g27c	Write or edit text for a school assignment	Write or edit text
is2g27d	Search for and find relevant information for a school project on the Internet	Search and find information on internet
is2g27i	Create a multi-media presentation (with sound, pictures, or video)	Create multi-media presentation
is2g27j	Upload text, images, or video to an online profile	Upload multimedia to an online profile
is2g27k	Insert an image into a document or message	Insert an image into a document/message
is2g27l	Install a program or [app]	Install a program/app
is2g27m	Judge whether you can trust information you find on the Internet	Judge internet information veracity
<i>Specialized Self-efficacy</i>		
is2g27b	Create a database (e.g. using [Microsoft Access @])	Create a database
is2g27e	Build or edit a webpage	Build a webpage
is2g27g	Create a computer program, macro, or [app]	Create a computer program/app
is2g27h	Set up a local area network of computers or other ICT	Set up a local area network

- Computer and Information Literacy

The ICILS CIL reporting scale was established in ICILS 2013, with a mean of 500 (the average CIL scale score across countries in 2013) and a standard deviation of 100 for the equally weighted national samples that met IEA sample participation standards in the first cycle (2013). The ICILS database offers five possible values of the CIL score that were generated with full conditioning to derive summary student achievement statistics. Conventionally, papers based on this study sometimes occupies the first plausible value, which is coded "pv1cil". This research is no exception. As more the value of the scale, more the CIL of the student.

School level

- School Self-efficacy composition

To explore the distribution of data at the school level, both self-efficacy indexes were aggregated using means. The code for constructing the aggregated variables is accessible on the [GitHub repository of the study](#).

- Computer and Information Literacy composition

To explore the distribution of CIL at school level, the score were aggregated using means. The code for constructing the aggregated variables is accessible on the [GitHub repository of the study](#).

- Female student enrolment

2018 ICILS school database didn't include the n of girl and boy in the school, but as the student sample is representative of the student gender distribution, the ratio of girls students in the school was estimated by dividing the number of girl students in the sample by the total number of students in the sample.

- Gender student composition

Girl student ratio variable was recoded to categorical. The categories are "Masculinized school" by the range of 0 to 0.33, "Mixed school" by the range 0.34 to 0.66 and "Feminized school" by the range of 0.67 to 1.

Mains results

Although both indexes measure technological self-efficacy, it turns out that their relationship with gender is different. H1 is rejected for self-efficacy oriented towards simple or general tasks, since it is lower in boys than in girls, but H1 is approved for self-efficacy for advanced or specialised tasks, since boys obtain higher scores. This is related to the general literature in the field studied, which indicates that, in general, females have a more intensive use of digital technologies in communicational, graphic and social activities (which requires efficacy in simple tasks), while males tend to go more deeply into operational or technical uses, as they are more interested in computer development and video games. Something similar occurs with computer literacy. Student CIL has a positive effect on general self-efficacy as expected in H2 of the study. However, contrary to what was expected in H2, as students' CIL increases, their specialised self-efficacy decreases. Interaction was not significant, and moderations could not be carried out given the fixed condition of the sex variable in both efficacies. Therefore, H3, H4 and H5 were rejected. School sex composition effects were also not significant, so H6 was rejected. However, when the school is composed by students with more CIL, the individual feel less general and specialized self-efficacy. Then, H7 have to be rejected. Considering Merton's theory, probably, as the reference of the student is their pairs, when comparing him or her with them, feel less confidence of his or her own skills. Instead of boosting their confidence, students underestimate themselves when surrounded by other literate students. The implications of these results will be discussed.

Keywords: Basic and Advanced ICT Self-efficacy, Digital Gender Gaps, ILSAS, International Computer and Information Literacy Study - ICILS

Using Diagnostic Classification Models with TIMSS data to explore mathematics learning in Italy

Adam Coates

Introduction.

International Large-Scale Assessments (ILSAs) can provide valuable information about a country's education system, and this can be used to consider potential policies that might help improve learning in schools. In particular, the cognitive outcomes measured in ILSAs may be analysed at various levels: research can look at average country scores in different subjects, domains within subjects, or responses to specific items (e.g. Ajello et al., 2018; Caponera & Losito, 2016; INVALSI, 2021; M. Wu, 2009). An alternative method is to use diagnostic classification models (DCM) to identify skills which are tested in the assessments and use statistical modelling to estimate the probability that each pupil has mastered each skill. This methodology provides finer-grained information than scores within subject domains and thus allows more detailed descriptions and comparisons of learning within different countries.

DCM have been fit to ILSA data, but studies often use the data in empirical tests of new models without exploring the implications of the results (e.g. Eren et al., 2023; Gu & Xu, 2023; Oka & Okada, 2023; Yamaguchi, 2023). Studies which analyse ILSA data and aim to explore implications for learning and education tend to compare multiple countries, and can therefore only provide relatively shallow detail about the learning in each country (Jia et al., 2021; X. Wu et al., 2022; Zhu, 2023). A few studies focus on learning in a specific country, demonstrating the depth of information that can be produced by DCM (such as Chen et al., 2008 for Taiwan, and Wafa et al., 2020 for Afghanistan). However, to date, no known study has used DCM to explore learning in Italy.

Research object and hypothesis.

This study aims to use TIMSS data to produce a detailed description of mathematics learning in Italy. Patterns of mastery of mathematics skills and learning trajectories will be identified and compared with other countries to outline Italy's relative strengths and weaknesses at a fine-grained level and investigate how these have changed over time.

Data.

TIMSS mathematics response data for cycles between 2003 and 2019 will be analysed. To create the DCM, the study will use mathematics skills identified by previous research for publicly released items. Data from all countries for each TIMSS cycle will be included in each model.

Method.

To create DCM for the TIMSS data, a cognitive model for the mathematics involved in solving the test items is necessary; this involves two key steps. The first stage is identifying which skills are required to correctly solve each item. As mentioned, this can be drawn from previous research with the publicly released TIMSS items. An important feature of DCM is that at least some items are expected to draw from multiple skills, and this distinguishes the use of DCM from a more simple analysis using the TIMSS theoretical framework, since the latter classifies each item as belonging to only a single content domain.

The second step involves choosing how the skills are assumed to interact in each item: for example, all skills may be necessary or only a subset may be needed. This can be done a priori, but it is common to select a model empirically by fitting multiple different models and choosing the model which best fits the data. Previous studies indicate that additive models where each skill individually contributes to an increased likelihood of success are most suitable for TIMSS (Yamaguchi & Okada, 2018).

The main result of a DCM is estimations of which skills each pupil has mastered. These will be summarised and the findings for Italy will be compared across different TIMSS cycles and with different countries. Moreover, by inspecting the most common mastery profiles, learning trajectories can be produced which indicate the typical order in which skills are learnt and mastered.

Results and Findings.

Preliminary results from the 2011 TIMSS Grade 8 data show that, relative to 38 other countries, Italy is strongest in geometry measurement skills and weakest in algebra patterns. However, in the 2015 Grade 8 data, Italian pupils remain relatively weaker at algebra, but operations with fractions and decimals is the relatively strongest skill.

Additionally, inspection of the most common mastery profiles for 2011 shows the orders in which pupils master skills, with pupils typically mastering four base skills, followed by probability and geometry measurement, followed by algebra patterns; algebraic expressions, equations, and functions; and geometry lines, angles, and shapes. The 2015 data shows more variation in the order of mastering skills, although algebra remains a more difficult skill.

The complete findings will identify consistencies and changes in mathematics skills mastery in Italy and this may be used to point to potential areas in the Italian mathematics curriculum and teaching which may need greater support. Possible policy directions based on these findings will be discussed. Moreover, considering that differences in achievement for the different skills in the TIMSS framework have been found for gender and geographic region (INVALSI, 2021), the finer-grained breakdown of mathematics skills paves the way for a more detailed investigation of these achievement gaps.

Keywords: Mathematics education, Diagnostic Classification Modelling, TIMSS

PV or not PV. How to get random slopes in ILSA

Diego Carrasco – David Torres Irribarra

Introduction

Past research in large scale assessment have include random slope from covariates to test scores, to address questions of contextual effects moderation. In this type of research, the question of interest is if there are school factors that ameliorate socioeconomic gaps, intergenerational gaps (from parents' educations), and gender gaps among others; or if contextual factors may exacerbate certain gaps (include other gaps). A crucial element in this type of research inquiry is the ability to obtain random slopes (i.e., varying gap sizes) among units of interest (i.e., schools). In substantive terms, research of this type is interested in teaching classroom practices, school climate factors, among other school attributes that can conditioned a pre-established relationship between an outcome (e.g, socioeconomic gap of test scores), or mean difference between groups (e.g, boys and girls' differences on test scores).

International Large-scale assessment (ILSA) studies are an ideal scenario to address such research questions. These studies include many schools (about 150), around 30 students or so per school, where many of the factors of interest are available for researchers. Moreover, these kinds of studies, provide information from probabilistic samples, from different countries, and from different years, thus providing a very ideal design to endure such inquiries.

Most of the recommendations regarding the use of ILSA, emphasize the use of plausible values, when the outcome of interest is the scores of the test (Von Davier et al., 2009; Rutkowski et al., 2010; Carsten et al., 2010). And rightly so; plausible values assure secondary users can produce estimate of interest, without the need to re-estimate the measurement model that generates the latent scores of the test in the first place, saving a substantial portion of time. Moreover, plausible values carry the uncertainty of the measurement error, thus allowing secondary users to make inferences above measurement error (e.g., Pokropek, 2015). Nevertheless, there is a caveat. If the estimand of interest (Lundberg, et al., 2021) is not included in the conditional model that generates the plausible values (PV), the researcher may not obtain the estimate he or she is looking for. PV are random draws from the measurement model (Wu, 2005), while including a set of conditioning variables (Diakow, 2013). In essence, these can be thought of a special application of imputation procedures. To assure an estimand is retrieve with the imputed data, the researchers model needs to be congenial to the imputation model use to generate the plausible values (Braun et al., 2017). Otherwise, the “non congenial estimands” may be underestimates of the effect of interest. Simulations studies from Zheng (2024) assert that, the PV generating model, work well for random intercept models, but not for models including random slopes. Hence, is expected that previous literature looking for random slopes, were not able to find these.

To illustrate the present challenges, we will use data from International Civic and Citizenship Education Study from (ICCS 2009) and will estimate random slopes on civic knowledge for all participating countries on the study, using the provided plausible values, and also for students’ sex. For comparison purposes, we will use the IRT scores (weighted least estimates, WLE) included in ICCS 2009, to show how results can differed between these two approaches. IRT WLE scores, unlike PV, are not affected by the conditioning model, if these were generated with one, is considered a conditionally unbiased. Thus, even if IRT WLE could potentially lead to bias due to not accounting for measurement error, ending in attenuated estimates (Bhaktha et al, 2021), these biases are ameliorated when the test include a large number of items (e.g., more than 25 items), high reliability, and spread distribution of item locations (Diakow, 2013).

Methods. Data. We used data from the International Civic and Citizenship Education Study from 2009 (ICCS 2009). This study provides representative representative samples of 8th grade students, using a two-stage design, with intact classrooms. To illustrate the present problem, when estimating random slopes with plausible values, we fit multilevel models with a random slope for socioeconomic status (a continuous measure); and for students’ sex (a binary measure), for each country participating in the study, with the exemption of Luxemburg, Malta, and Liechtenstein due to low count of schools ($n < 60$). As such, we are including in this exercise the 35 remaining educational systems. ICCS 2009 is a convenient ILSA study, because it provides with PV and IRT WLE estimates from the same test (Schulz et al., 2011).

Dependent variable. Civic Knowledge (y_{ij}) scores represent students’ political understanding of political issues. It consists of five plausible values, generated with IRT Rasch model over a random booklet design of 79 item-test. We will use the provided plausible values (PV1-PV5), and the IRT WLE scores from each country generated with the 79 adjudicated items for the international test.

Independent variables. To illustrate the presented problems, we are using two variables. Socioeconomic Status (ses_{ij}), is a score created based on the Parents Education level, Parents Occupation, and number of books at home. Students’ sex scores (sex_{ij}) is a binary variable (girl = 1, boy = 0).

All estimates are pseudo maximum likelihood estimates, where survey design weights and plausible values are accommodated accordingly (Rutkowski et al., 2010).

Analytical strategy

To illustrate the problem of estimating random slopes we need to fit the following general model:

$$y_{ij} = \alpha + \beta_1 j (x_{ij} - \bar{x}_{\cdot(j)}) + u_{0j} + u_{1j} (x_{ij} - \bar{x}_{\cdot(j)}) + \epsilon_{ij}$$

In the present model, we include our covariate of interest (ses_{ij} , or sex_{ij}), centered at the school mean. This model specification allows to retrieve the variance of the slope of interest, without confounding, thanks to the fact that the centered covariates doesn’t carry any between variance (Enders et al., 2007). The previous model is comparing with to a similar model where the slopes are assumed to be fixed.

$$y_{ij} = \alpha + \beta_1 (x_{ij} - \bar{x}_{\cdot(j)}) + u_{0j} + \epsilon_{ij}$$

Given the variance of $\beta_1 j$, is the term of interest to make inferences regarding random slopes, we use a likelihood ration mixture, to assess if we have a random slope (Kim et al., 2013). In practical terms, one can

use a traditional likelihood ratio test, but needs to divide the obtained p value by two (Rabe-Hesketh et al., 2012).

The previous strategy is enough for making inferences regarding the IRT WLE scores. However, for making inferences regarding the plausible values, we rely on the proposal of Grund et al. (2023), to make inferences regarding variance terms from imputed data.

In consequence, we fit 2 models, for 2 covariates for 35 different educational systems, participating in ICCS 2009. Our estimand of interest is β_{1j} , and u_{1j} , and we use modified versions of likelihood ratio test to identify when the variance of u_{1j} is above zero. All estimates are pseudo maximum likelihood estimates, where survey design weights and plausible values are accommodated accordingly (Rutkowski et al., 2010). Preliminary results.

Using the methods described above, we found random slopes, using plausible values, for 2 countries when we conditioned civic knowledge on socioeconomic status, and random slopes for 6 countries when we civic knowledge on students' sex. In contrast, when we used the same procedure, using the IRT WLE estimates instead, we found 11 countries with random slopes for sex; and 12 countries with random slopes for socioeconomic status.

Conclusions

In the present study, we have shown that the study of random slopes in ILSA is compromised using plausible values. Random slopes estimate from PV and IRT WLE scores do not lead to the same conclusions.

PV imputations do not recover random slopes correctly (Zheng, 2024). IRT WLE, do find more random slopes among the different countries included in the study, yet it might not be a perfect solution either. IRT WLE scores, may underestimate the regression coefficients (the slopes or interest) (Bhaktha et al, 2021), and may struggle to recover the variance components in mixed models (Diakow, 2013). Further research is needed to identify practical solutions to the present problem.

Discussion.

The present exercise has two implications. The first pertain secondary users, and the second pertains ILSA study operations. Secondary researchers need to address the limitations of the methods they are using when making inferences. This is specially the case if the chosen methods obstacle their inferences of interest. Zheng (2024) study asserts that plausible values generated with the interim means method can recover random slopes (a method used in TIMSS 2019). For example, ICCS 2016, it seems to have used the same method (interim means) (see Schulz, et al., 2018, p133). Thus, for researchers aiming to study random slopes involved in the inquiry of protective and risk factors of educational gaps, are encouraged to identify the method of how PV were drawn, and if these compromise their methods of inference.

On the other hand, ILSA studies operations, aiming to promote the study of educational gaps with risk and protective factors at the school level are encouraged to vary their PV method generation accordingly, and provide thorough documentation of how these are generated. Thus, allowing secondary users to identify the limitations of the generated scores. ICCS studies, are good example of how thorough this documentation can be (see Schulz et al., 2011, 2018 for examples).

Keywords: random slopes, plausible values, educational gaps

SESSION 1. DIFFERENT DATA SOURCES AND THEIR INTEGRATION: POLICY INDICATORS FOR THE EDUCATION AND TRAINING SYSTEM

ORGANIZER: ISTAT

COORDINATOR: BARBARA BALDAZZI

OCTOBER 17TH: 2.00 P.M. – 4.00 P.M. {ROOM 3 – RESEARCH 2}

Social Inequality and Student Mobility from Southern Italian Regions: The Interplay between Academic Achievement, Track at Upper Secondary Education and Social Origin

Andrea Priulla – Eleonora Miaci – Nazareno Panichella – Massimo Attanasio

Abstract. The North-South disparities in terms of economic development in Italy are also reflected in the inequalities in educational opportunities. Southern regions, already plagued by significant school dropout rates, experience significant losses in the school-to-university transition. In this context, this work aims to study the interplay of sociodemographic factors and others related to previous educational attainment on the university enrolment choices of southern graduates. The results highlight how enrolment choices are differentiated based on the interaction of social class and student's previous educational attainment.

Introduction

Despite the detection of a decline of inequality in educational opportunities in several countries [1], the family background still represents an important predicting factor for individual attainments [2]. In Italy, research has revealed that university enrolment choices are strongly tied to students' social backgrounds, in addition to their abilities and personal characteristics. While the reduction in absolute inequalities in compulsory education has been noted [3], Italy's educational expansion has seen many children of non-graduates not pursuing tertiary education. Additionally, inequalities in educational opportunities are reflected in the North-South divide characterising many facets of our country. A recent report by the Bank of Italy [4] has indeed underlined the lower propensity of southern graduates to continue their studies after graduation compared to other Italian graduates. At the same time, the AlmaLaurea Report (2023) points out that around 28% of southern students choose universities in the Centre-North, with this mobility trend increasing over the last decade [5]. The choice of southern students to attend universities in the Centre-North is driven by the pursuit of enhanced socioeconomic environments, the prestige of northern universities, and the prospect of better employment outcomes [6,7]. Research has emphasized the importance of both student personal characteristics and contextual factors in the mobility of students from the South to the Centre and North [8,9,10,11].

In this work, we aim to study the school-to-university transition of southern graduates according to three potential outcomes: the choice of not enrolling at university, and the choice to enroll in the Center-North or to remain in the South. In detail, we delve into the analysis of the interplay between social origin, academic ability, and high school track on enrolment choices. While the interplay between these factors is pivotal in the literature on social stratification, there is a paucity of research focusing on how this interplay influences university mobility [12].

Theoretical framework

While the impact of social origin on educational paths has received considerable attention, its interplay with academic ability has been less thoroughly examined. Students from better social origins are more likely to excel in reading, math, and science scores, and opt for upper secondary school tracks that align with tertiary education [13,14]. Bernardi [15], in the Spanish context, investigates the compensatory effect of social background in correcting educational missteps and examines how cognitive and non-cognitive skills can influence the effect of social background on educational transitions. Family resources can increase the likelihood of educational success, whereas, only the most talented students from disadvantaged backgrounds manage to reach a higher educational level. In this framework, the choice of high school track is crucial in shaping students' academic careers, and the family is usually responsible for their children's education at this stage. Students from higher social classes tend to prefer humanistic or scientific lyceums, notably more suitable for university studies [16,17].

Literature has highlighted that mobility choices the mobility flows predominantly attract southern students with the best academic results [4]. For instance, [18] demonstrated that students from more endowed

families and those attending lyceums are more likely to leave southern regions. Moreover, track selection inevitably determines students' field of study choice in tertiary education, further impacting the choice of whether to enroll or not since some fields imply continuation and others do not, and the mobility patterns due to their uneven regional distribution [19].

Research questions and hypotheses

Our study focuses on three distinct pathways for Southern students' post-secondary education choices: non-enrollment, enrollment in a southern university, or enrollment in a central-northern university. We hypothesize two scenarios: moving to the North as a "boosting" or as a "compensatory advantage". In the first scenario, families invest in already well-prepared children, aiming to further increase their chances of success. In the second scenario, the most prepared child can establish themselves in the job market even after studying at a less prestigious university in the South. Below are the research questions and related hypotheses:

Q1: Are southern students from better-off families and with high academic ability more likely to enroll at university?

H1: We expect a positive association between social origin and enrollment, as well as between academic ability and enrollment.

Q2: Are southern students from better-off families and with high academic ability more likely to enrol in a Northern University?

H2: We expect a positive association between social origin and mobility. Given the lack of literature, we do not have expectations on the role of academic ability.

Q3: Does the effect of academic ability vary with respect to social origins?

H3: We expect higher propensity to move among most able students ("boosting effect").

Data

The analysis is based on the linkage of micro-data coming from two national administrative archives [20]: INVALSI and Anagrafe Nazionale Studenti (ANS). We consider the cohort of students on grade 13 during the academic year 2021/22, the most recent available cohort. For those students, we know about their choice to enrol at an Italian university immediately after graduation. This means that students enrolling late or abroad are considered not enrolled. Finally, students enrolling in healthcare programs and online universities are excluded.

Methodology

To investigate enrolment choices, we fit a mixed-effects multinomial logistic model

$$\text{logit}(\mu_{(ij,k)}/\mu_{(ij,1)}) = X_{(ij,k)} \beta + Z_{(ij,k)} b_{(ij,k)} \quad k=2,3$$

- **k** identifies the response level. The baseline is a not enrolled student;
- **Z_{ij}** is the vector of random covariates, that is the southern province where students attended high school;
- **b_i** ~ N(0, D) is the vector of random effects associated with the Z_{ij};
- **X_{ij}** is the vector of fixed covariates for the i-th observation in the j-th group;
- **β** is the vector of fixed covariate coefficients.

Along with the main effect of students' country of origin, we added a three-way interaction between the type of high school attended, maths scores and parents' education-

Results

Figure 1 displays the median estimated probabilities of not enrolling at university. Although we accounted for the missing data regarding the parents' educational level in our model, we opted not to include the estimated probabilities in the results.

The findings reveal a hierarchical pattern in enrollment choices: the probability of not enrolling at university is lower for scientific and humanistic students, with strong maths ability, and better-educated parents. Humanistic schools consistently show lower non-enrollment probabilities regardless of parental education and math ability, while vocational students are the least inclined to enrol at university. In detail, the gap in the estimated probabilities based on parents' education narrows as math ability increases. On the other hand, math ability seems to be more influential for technical students, especially for those with better-educated parents.

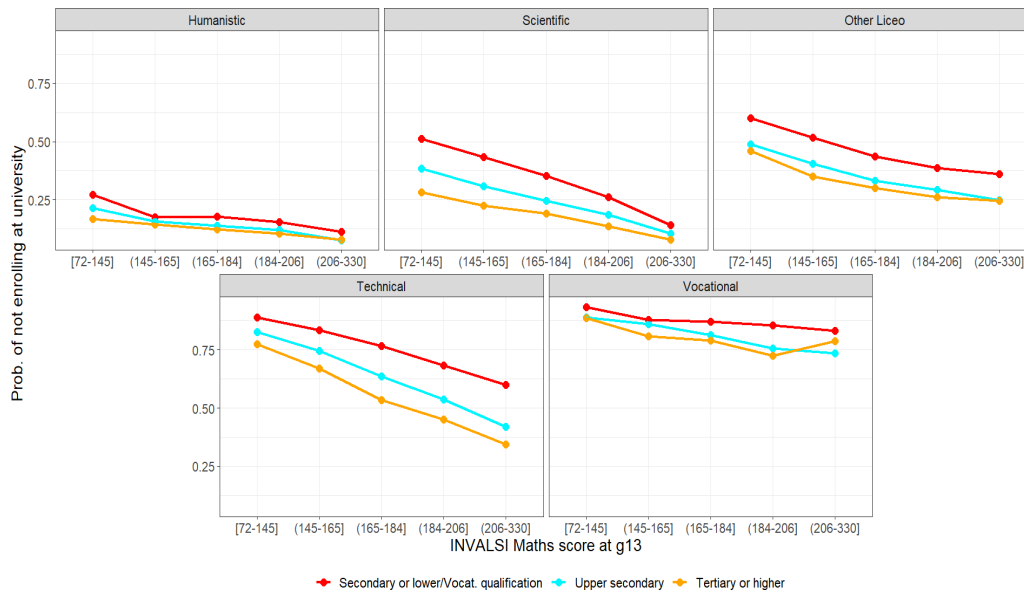


Fig. 1: Median estimated probabilities of not enrolling at university.

Figure 2 shows the median estimated probabilities of moving to Center-North. Humanistic and scientific students performing good in maths, and with better-educated parents are more likely to move to Center-North. In detail, the interplay of maths ability and parents' education seems to be more evident among top-performing students in humanistic and scientific schools, with a gap of around 0.15 between with tertiary-educated parents and those with secondary-educated parents. Moreover, the increase is stronger for those in the last quintile of math scores, highlighting that the best students are the most likely to move to central-northern regions.

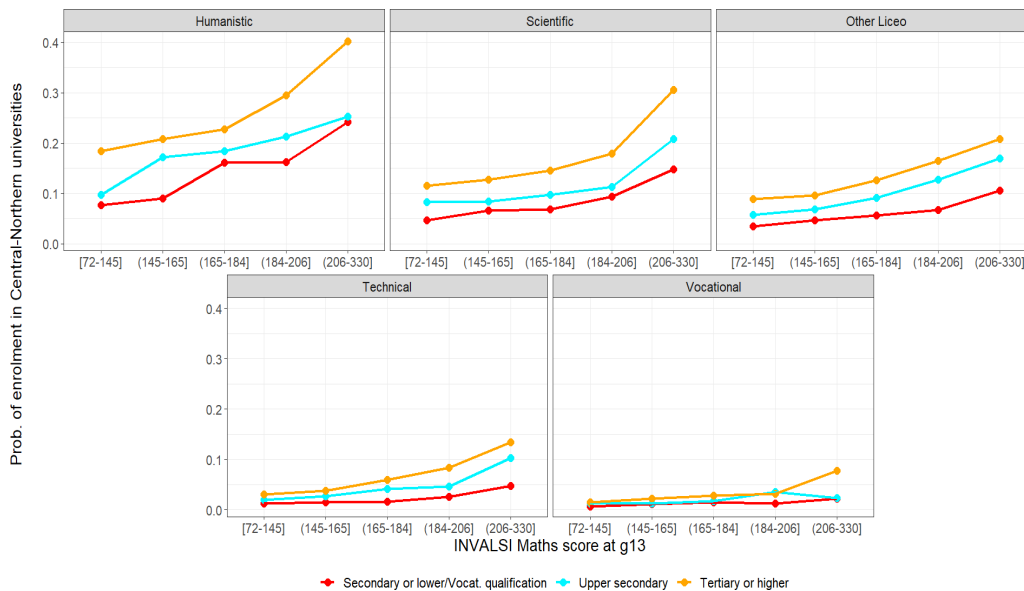


Fig. 2: Median estimated probabilities of moving to Centre-North.

Conclusions

This work provides insights into the interplay of social background, high school track, and academic ability on the mobility choices of students from southern Italy. The results of a mixed-effects logistic model align with our hypotheses. There is a strong positive association between social origin and enrolment choices, as students from more highly educated parents are more likely to enroll at university and to migrate from southern regions. Furthermore, the hypothesis of a “boosting effect” is confirmed. Migration to central-northern regions is more likely for students with higher educational achievements. Finally, the findings

underscore the necessity of adopting an intersectional approach to deal with educational data since a wide range of factors contribute to shaping the students' future trajectories from their earliest schooling years.

Keywords: social background, high school track, academic ability, university enrolment

The gender gap in school-to-university transition in Italy: Insights from the integration between INVALSI and ANS data sources

Valentina Tocchioni – Gabriele Lombardi – Samuele Milone

Introduction

The issue of the gender gap in university enrollment is widely investigated all over the world. The United States is experiencing a long-increasing trend in women's university enrollment and admissions. This involves also those disciplines traditionally sought after by men (Adebayo, 2008), and it is stimulated by the well-known evidence that females outperform males, thus being advantaged during selection procedures (Conger, 2015). Nonetheless, studies about European countries show how this long-term trend has brought only a roughly equal representation in enrollment (Francesconi and Parey, 2018; Steinmann and Rutkowski, 2023). Regarding Italy, very few studies have addressed the topic of university enrollment for women, being more concentrated on issues about their academic career. Nonetheless, females have higher enrollment and lower dropout rates, especially in the South of Italy (Contini *et al.*, 2018), but this data should be read in the light of women's lower propensity to choose a university located far away from their hometowns (D'Agostino *et al.*, 2019; Genova *et al.*, 2021). Moreover, the report on gender differences in the Italian Higher Education system highlights, on the one side, that the female rate of enrollment between the 2012 and 2021 has been reduced by one percentage point, and, on the other side, that female enrollment in Italy is consistently above the European average, mainly driven by some fields of study such as Art and Humanities and Health Studies (ANVUR, 2024). Nonetheless, since 2018, Italy has exhibited a rate of enrollment above the European average for women in the fields of Natural Sciences, Mathematics, Statistics, and Engineering, which are fields well-recognized for being crucial for the needs of the job market, but also traditionally overrepresented by men (D'Agostino *et al.*, 2020).

Given this state of affairs, we intend to proceed with a deep investigation of the main drivers of students' enrollment in tertiary education by gender. In particular, this work intends to assess the school-to-university transition rate of upper secondary schools in Italy, detecting potential gender inequalities which could be moderated by other factors. Indeed, gender inequalities in access to the higher education system could derive from the persistence and transmission of gender disparities according to different students' socioeconomic conditions, territories, types of upper secondary schools and academic fields of study. For this reason, we intend to identify schools and university degree programs that show gender-balanced enrolments and to verify if it may depend on several factors, such as a same-sex peer effect from upper secondary school or the geographical location of both the school and the university itself. For attaining our objectives, we will create several school-level and classroom-level indicators, which take into account the gender composition of students enrolled to university.

Data

In our study, we examine gender differences in the transition from high schools to universities in Italy. To do so, we draw upon two administrative data sources: the first, namely the Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione (INVALSI; in English: National Institute for the Evaluation of the Education and Training System) database, contains information about all students of grade 13 in the school year 2018-2019 who took the INVALSI test and graduated in the same year in an upper secondary school located in the Italian territory; while the second is the Anagrafe Nazionale Studenti (ANS; in English: National Student Registry) database, which includes all students enrolled in an Italian university in the academic year 2019-2020, thus in the year following graduation¹.

¹ The data used in this study have been processed in accordance with the Research Protocol for the study "From high school to the job placement: analysis of university careers and university mobility from Southern to Northern Italy" among the Ministry of University and Research, the Ministry of Education and Merit, the University of Palermo as the lead institution and the INVALSI Institution; and the data request to the INVALSI Institute for the research "Chi prosegue e chi no gli studi universitari in Italia: un'analisi delle diseguglianze dalla scuola secondaria superiore all'università", lead by the University

In order to provide a deeper understanding of the gender gap in the school-to-university transition in Italy, we will present results stratified by type of diploma (curriculum of secondary education) and by region. In the 2019 INVALSI dataset, the curricula reported by students are as follows: Artistic, Classical, Linguistic, Professional, Scientific, Scientific - Applied Sciences, Humanities, Technical Economic, Technical Technological².

By doing so, our dataset comprises a total of 438,282 students of grade 13, of which 222,525 are females (50.8%). University enrollees are 226,619 (51.7% of all students considered), whose female students are 127,475 (56.3% of all university enrollees).

Methods

As for the methodology, we will create a series of school-level and classroom-level indicators, which will take into account the gender composition of students when transitioning from upper secondary schools to university.

In the following, we present in detail two indicators which can be considered the starting point for understanding the gender dynamics in the school-to-university transition. The first indicator is the proportion of students enrolling in university, calculated as follows:

$$P_{uni} = \frac{uni}{n}$$

Where P_{uni} represents the proportion of students enrolling in university, n is the number of students who took the INVALSI test, and uni is the number of students enrolling in university. This index varies from a minimum of 0 to a maximum of +1, where higher values indicate a higher proportion of students enrolled in university. This indicator is calculated both for the female population and for the male population separately and will be referred to as the female/male entry rate.

The second index we present is the excess index of female students enrolled in university over male students enrolled in university, and will be referred as the *gender gap in the entry rate*. The formula is as follows:

$$Ecc = \frac{P(F)_{uni} - P(M)_{uni}}{P(F)_{uni} + P(M)_{uni}}$$

Where $P(F)_{uni}$ represents the ratio between female students enrolled in university to the total number of females calculated for each school, while $P(M)_{uni}$ represents the ratio between male students enrolled in university to the total number of males calculated for each school.

This index provides a relative measure of the difference in the proportion of female students and male students enrolling in university. This index can vary from a minimum of -1 to a maximum of +1: a positive value indicates an excess of female students over male students, a negative value indicates an excess of male students over female students, and a value of 0 indicates a perfect gender balance.

Results

For space constraints, we present the two indicators illustrated above among those computed for our sample of students, namely the female/male entry rate and the gender gap in the entry rate, stratified by types of upper secondary schools and region of residence (for the latter only). Table 1 shows the female/male entry rate, stratified by types of upper secondary schools. First, regardless of the school curriculum, the proportion of female students enrolling in university is consistently higher than that of male students. This trend highlights a gender disparity in the school-to-university transition, with a female prevalence in accessing higher education. Moreover, the entry rate varies considerably, ranging from a minimum in the Professional track, with a proportion of 0.178, to a maximum in the Classical track, with a proportion of 0.847. Such extreme differences can be attributed to various factors. The low value in the Professional curriculum might be influenced by its vocational nature, where the primary goal is acquiring practical professional skills, which might lead students to perceive university as less necessary for their future careers. But this stark division might also reflect the persistence of a classist model where certain schools are seen as steppingstones to leadership roles and social prestige, while others are perceived as less prestigious options, not necessarily oriented toward a university education.

of Palermo. The University of Florence is one of the institutions that signed both agreements; the reference researcher for this university is prof. Bertaccini.

² We dropped 492 students who reported that their diploma was “Other”, and 8,071 observations with a missing value in the type of upper secondary school.

Looking at the variability among schools, we observe that there are no Classical, Linguistic, or Humanities-focused schools where no female student enrolls in university (the minimum value is greater than 0 for all schools).

In conclusion, the data analysis clearly highlights a prevalence of female students enrolling in university compared to male students, as well as a significant variation in the proportion of university enrollments across different school curricula. These observations underscore the importance of considering gender dynamics in the context of higher education.

Table 1: Female/male entry rate by type of upper secondary school. Academic year 2019/2020. Descriptive statistics.

	Female entry rate				Male entry rate			
	Mean	SD	min	max	Mean	SD	min	max
Artistic	0.330	0.098	0.000	0.730	0.233	0.117	0.000	0.667
Classical	0.847	0.073	0.300	1.000	0.830	0.112	0.000	1.000
Linguistic	0.695	0.101	0.182	1.000	0.618	0.182	0.000	1.000
Professional	0.178	0.109	0.000	1.000	0.094	0.069	0.000	1.000
Scientific	0.845	0.099	0.000	1.000	0.819	0.117	0.000	1.000
Scientific – Appl. Sc.	0.798	0.143	0.000	1.000	0.785	0.124	0.111	1.000
Humanities	0.643	0.133	0.048	1.000	0.559	0.219	0.000	1.000
Tech. Economic	0.429	0.153	0.000	1.000	0.358	0.170	0.000	1.000
Tech. Technological	0.458	0.200	0.000	1.000	0.342	0.136	0.000	1.000

In Table 2, the results of the gender gap in the entry rate are reported, stratified by school curriculum. The highest values of this index, recorded for the professional and artistic curricula (0.195 and 0.192 respectively), are consistent with the lowest absolute proportions of students enrolled in university in both genders. This phenomenon suggests that, when university participation is less common within certain educational pathways, a more pronounced gender discrepancy emerges, with a noticeable amplification of the gender gap in the entry rate. This dynamic could result from multiple factors, including socio-cultural perceptions related to career opportunities, family expectations, economic dynamics, and teaching and career guidance methods.

Conversely, in scientific and classical secondary schools, the lowest values are recorded (from 0.005 for the Scientific - Applied Sciences curriculum to 0.013 for the Classical curriculum), indicating a similar propensity, albeit slightly more inclined towards females, to enrol in a university.

In summary, the analysis of the gender gap in the entry rate provides an in-depth look at gender differences in access to university, highlighting how these disparities are influenced not only by students' propensity to attend university but also by their respective educational pathways.

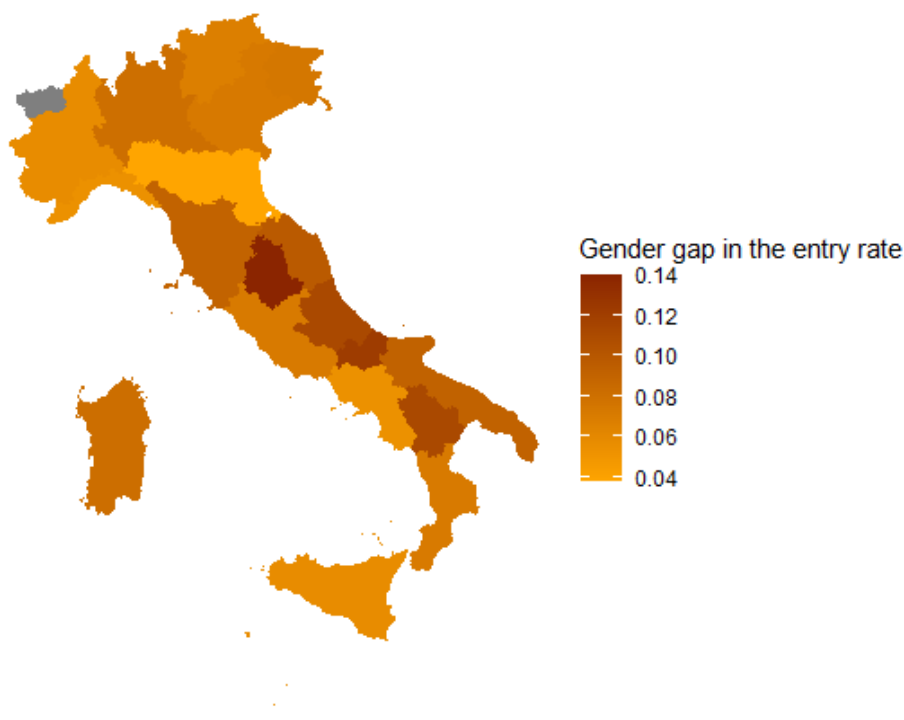
Table 2: Gender gap in the entry rate by type of upper secondary school. Academic year 2019/2020. Descriptive statistics.

	Mean	SD	min	max
Artistic	0.192	0.282	-1.000	1.000
Classical	0.013	0.089	-0.538	1.000
Linguistic	0.064	0.172	-0.600	1.000
Professional	0.195	0.411	-1.000	1.000
Scientific	0.010	0.083	-1.000	1.000
Scientific – Appl. Sc.	0.005	0.110	-1.000	0.515
Humanities	0.080	0.237	-0.511	1.000
Tech. Economic	0.063	0.260	-1.000	1.000
Tech. Technological	0.046	0.366	-1.000	1.000

Looking at Figure 1, a more homogeneous situation emerges with respect to that illustrated by type of school. Indeed, the range spans from a maximum of 0.14 for Umbria to a minimum of 0.04 for Emilia-Romagna, and this suggests that the persistence of a higher propensity of females to enrol in university compared to males does not seem to be highly influenced by geographical factors with respect to the school curriculum undertaken. However, a general pattern can be observed, with higher values on average in the centre and lower values in the north and the south of Italy.

In sum, the gap in students' propensity to enrol in university is the result of a complex interaction of factors operating at the micro-, meso- and macro- level. With our work, we intend to shed light on these dynamics looking at which factors contributed to gender-balanced enrolments at university in Italy.

Figure 1: Gender gap in the entry rate by region. Academic year 2019/2020. Mean values.



Keywords: school-to-university transition, indicators, administrative sources, Italy

Machine learning Models to identify students at risk to Academic dropout

Fiammetta Noccioli – Michele Marsili

Introduction. Machine learning techniques, a branch of artificial intelligence, are increasingly being used in various fields of social research. Indeed, the increasing availability of data, including from administrative sources, makes it possible to use these techniques to identify recurrences and relationships between variables in different application areas. In the context of educational research, enormous progress has been made in recent years. One of the main trends is the use of predictive models aimed at detecting and preventing school and university dropout phenomena. In particular, several studies (such as Berens et al., 2019) have used machine learning algorithms to try to predict, based on individual characteristics, students at risk of dropping out of university. However, data collection generally remains limited to the information level of individual universities. The availability of data therefore remains limited to the university career and socio-demographic information, including only partial data on the student's secondary school performance. Integration between the different data sources available could make a greater contribution in this direction by allowing, among other things, to include the skills measured in the INVALSI National Surveys (National Institute for Evaluation of the Educational and Training System) among the predictive factors of university failure.

Research object and hypothesis. This study aims to use supervised learning machine learning models, i.e., models trained on a set of features related to a known outcome, to explore/evaluate their performance

compared to more traditional models generally applied in education, analyzing the potential of this class of models that has seen increasing applications over the past few years. Specifically, the paper focuses on the prediction of university failure and its determinants by aiming to identify students who are at higher risk of dropping out of studies and to identify the most important predictors in this context. For the purpose of this study, it seems significant to predict this condition from the number of credits obtained at the end of the first year of the course. Indeed, relevant literature and the results of previous studies place emphasis on the transition from the first to the second year of the course as a turning point in the students' life and study paths. This critical issue is highlighted in the ANVUR's Biennial Report on the State of the University and Research System, published in 2023, which points out that the university dropout rate in Italy is among the highest in Europe.

Data. This study refers to the cohort of students graduated in the 2018/19 school year and enrolled at university in the 2019/20 academic year, coming from a dataset built on the combination of different data sources, Ministry of Education and Merit (MIM), University Student Register and INVALSI data. The analysis is based on the information relating to each student enrolled in one of the Italian universities, which are present in the National Student Registry, with the exclusion of those who do not present information for the variable number of credits at the end of the first year.

Before the analysis, the presence of outliers was checked by removing the last percentile of the distribution for the variable number of credits at the end of the first year, and the missing records identified on the context variables were treated. In particular, missing values (between 0.5% and 5%, except for the parental education around 15%) were replaced with a multiple imputation method (MICE).

Methods. Prediction of university success/failure was carried out using a set of machine learning techniques, known as supervised learning algorithms for classification problems. Several features, selected on the basis of the literature and data availability and completeness, were used to carry out the prediction, including individual-level characteristics, school type, and teacher grades. In fact, factors such as gender, age, educational career, influence the outcome of academic achievement, with teacher grades at grade 13 being particularly relevant here.

Using the number of credits earned in the first year of university studies, considered a good proxy for dropout, the binary outcome variable success/failure was constructed.

Classification models used are logistic regression, a more traditional method, compared with machine learning techniques such as support vector machine (SVR), Naive Bayes algorithm, k-nearest neighbors (KNN), Gradient Boost and Random Forest (RF). These methods were chosen as they are among the most relevant for solving classification problems according to the literature.

The performance of each model was measured through accuracy, precision, recall, F1 score and area under the ROC curve metrics, where a high value corresponds to a better performance of the model. Furthermore, a 10-fold cross-validation was carried out, therefore splitting the dataset into 10 folds.

The study is carried out following the stages of data pre-processing (imputation, standardization and transformation), model training, evaluation and prediction. The dataset was then divided into two parts: training and test sets, according to a widely spread randomized distribution, respectively of 80% of the data set for the training and 20% for the test set.

Results. Machine learning can be an effective tool in predicting university success/failure and useful for orientation interventions and policy planning/programming in the educational field, allowing students at risk to be identified early using the data available at the time of enrollment (and not career) and contextual factors.

In this context, INVALSI data introduce an extra wealth of information compared to other work predicting the phenomenon of university dropout.

Keywords: machine learning, prediction models, university dropouts

Linked administrative data and a multivariate approach to monitor attainment gaps for students with protected characteristics

Nadir Zanini

Introduction

Attainment gaps are a key indicator of educational and societal inequalities. Yet, statistical reporting of educational attainment often takes the form of simple breakdowns of students' achievement by selected characteristics. These descriptive statistics, although relevant, are potentially misleading when considered in isolation. They fail to account for the interplay between students' characteristics and other factors known to affect students' achievement but not readily available.

Aim and overall approach

This presentation will focus on the experimental work conducted by Ofqual, the Regulator of Qualifications and Exams in England, to improve the statistical monitoring of attainment gaps in secondary schools. As in other countries, the UK Department for Education (DfE) provides evidence on educational inequalities in England through their annual Official Statistics publications on the attainment of 16-year-olds and 18-year-olds. These are cross-sectional descriptive statistics in the form of bivariate distributions: simple breakdowns of achievement by some students' socio-demographic characteristics (for 2020, see respectively DfE 2020a and DfE 2020b). To tackle the limitations of descriptive statistics regularly published and retrieve more insightful evidence, we designed an analytical approach based on three main features:

- i. Broadening the set of students' characteristics through the use of linked data to include not only protected characteristics but also observable factors known to be linked to achievement (most notably, prior attainment);
- ii. Moving from a bivariate analysis to a multivariate approach to measure the impact of each student characteristic once other factors potentially affecting attainment have been held constant;
- iii. Focusing on how attainment gaps have changed over time, rather than on the gaps at a specific time, by introducing a set of criteria to identify changes we considered to be 'notable'.

Data

To broaden the set of students' characteristics included in the analysis, we pulled together data from two main administrative sources:

- i. Data collected by Ofqual from the Awarding Organisations, containing detailed information on the grades achieved by each student in each qualification awarded and their (current and prior) schooling;
- ii. Records from the National Pupil Database and the Individualised Learner Record held by the UK Department for Education containing, for each student, socio-demographic characteristics and measures of socio-economic deprivation.

We linked the two administrative sources of data so that for each student, a broad set of characteristics became available:

- i. Educational attainment – grades achieved by each student in each qualification taken at 16 and 18, since 2018.
- ii. Schooling – prior attainment measured by qualifications and grades achieved throughout their education, type of education establishment attended, and its geographical location.
- iii. Protected characteristics – ethnicity, first language spoken at home, gender, special educational needs, and disabilities (SEND);
- iv. Indicators of socio-economic deprivation – Free School Meal eligibility (FSM) and IDACI (Income Deprivation Affecting Children Index), measuring the proportion of all children aged 0 to 15 living in income-deprived families in a certain geographical area.

Methodology

We used a multivariate analytical approach. This allows us to explore the impact on overall results of each feature separately while controlling for other features. This is important because we know that there are relationships between different features (for example, ethnicity and first language) which are necessary to account for. We used regression modeling to estimate differences in results for groups of students after controlling for other variables.

We carried out mixed-effects modeling to analyze students' attainment, given the information on the characteristics of students clustered within schools. This is to be able to account for factors affecting results that only students from the same school share, such as admissions policy and teaching quality. The student-level characteristics described above were treated as fixed effects; the student's school was treated as a random effect.

Presentation of findings

Given the large sample size (totality of students), it is possible that even very small differences between groups/years are flagged as significant. For this reason, we developed a multi-step method for evaluating changes in relative outcome differences between key years of comparison. The aim is to identify practically

significant changes while taking into consideration normal between-year fluctuations. Although all the tables and charts are presented, only the notable changes flagged by the above approach are drawn out in the report.

Ofqual's equalities analyses were initially published as research reports (Lee, Stringer & Zanini, 2020; Lee, Rama & Johnson, 2022). Following user engagement that highlighted the need to make the findings of the analysis easier to navigate, from 2023 (Lee, Rama & Johnson, 2023), the analysis was made available for the first time in an interactive statistical report format available at this link (<https://analytics.ofqual.gov.uk/apps/AllQualifications/Equalities/>), similar to a dashboard. In this way, users can more easily navigate to the results they are interested in (for example, ethnicity), while still being able to find the detailed technical information made available in the previously published reports. Publishing the analysis as an interactive report allowed us to increase its accessibility.

The interactive report can be fully navigated with a keyboard alone, including navigating tabs, drop-down menus, and even interactive visualizations. All elements of these web applications work with assistive technology. For example, interactive visualizations have descriptive text explaining them in detail, going above and beyond what is available from alt-text alone. Descriptive tags explain what the interactive visualization shows, the type of visualization on display, the label and values of the axes, how many data points there are in a series, and even each data point's value. The interactive report is created using R Shiny. Graphs are created in Highcharts, which helps ensure they are as accessible as possible, using the WCAG 2.1 standard as a guideline. This allows users, even those with visual impairments, to interact with the visualization and understand the relationships of all the components of the visualization to all other components rather than just getting a description of the visualization.

Conclusions

The presentation will conclude with an analysis of the lessons learned with this work to improve the reporting of attainment gaps. The analytical approach used by Ofqual to produce and present equalities analysis has several advantages, both from a methodological and substantive perspective. Further work is needed to secure the completeness of the data and ease the interpretation of the results. Nonetheless, the use of more advanced statistical methodology to estimate indicators to be routinely reported, in education and beyond, should be encouraged as they have the potential to provide more robust evidence to inform policies.

Keywords : monitoring, attainment gaps, linked data, multivariate, policy

SESSION 7. ENHANCING FINANCIAL LITERACY THROUGH EDUCATION

ORGANIZER: POLYTECHNIC UNIVERSITY OF MILAN - UNIVERSITY OF ZURIGO

COORDINATOR: MADDALENA DAVOLI

OCTOBER 17TH: 2.00 P.M. – 4.00 P.M. {ROOM 4 – RESEARCH 3}

Evaluating a Financial Education National Campaign: Experimental Evidence from Italy

Alessia Sconti – Tim Kaiser – Annamaria Lusardi – Bilal Zia

Basic knowledge of economics and finance is essential for making informed economic choices (Lusardi and Mitchell 2014, 2023). However, in Italy, only 44% Italians shows a basic level of knowledge of concepts such as inflation, compound interest and risk diversification . Moreover, according to PISA data (OECD, 2019) , Italy shows a statistically significant gender gap in financial literacy among 15-year-old students in 2018. As a step towards improving the financial decision-making capabilities of the general population in Italy, the Committee for the Planning and Coordination of Financial Education Activities (hereinafter referred to as the Committee) promotes a nationwide financial education information campaign. We study the causal effects of a national information campaign designed to foster the financial literacy of the Italian population. The campaign involves major media such as national television, print, and social media. To this end, we conduct a randomized encouragement experiment with a representative sample of 3,798 individuals in Italy

offering monetary incentives to follow elements of the national campaign. The experiment contrasts two different edutainment treatments delivered via national television and an information treatment delivered via print, radio, and social media relative to a pure control group.

A growing branch of the literature in the field of financial literacy studies the causal effect of educational interventions of varying nature and duration, aimed at specific population targets (see the recent meta-analysis by Kaiser et al., 2022 for a comprehensive review of the most rigorous studies on the topic). Yet, there is an evidence gap regarding causal effects of low-cost information campaign interventions delivered at scale via mainstream media.

A growing literature in economics studies the causal effects of media exposure on individual behavior (see DellaVigna and La Ferrara 2015 for a narrative review). The outcomes domains studied are diverse and cover crime (Dahl and DellaVigna 2009), health (Vaughan et al. 2000, La Ferrara et al. 2012, Kearney and Levine 2015, Trudeau 2016, Banerjee et al. 2019, Breza et al. 2021), education, empowerment, and labor-market outcomes (Zavodny 2006, Chong and La Ferrara 2009, Jensen and Oster 2009, Ravallion et al. 2016, Bjorvatn et al. 2019, Kearney and Levine 2019), and financial decision making (Baker and George 2010, Berg and Zia, 2017).

In the context of financial education, evidence from South Africa suggests that edutainment is an effective way to foster individual financial decision making (see Berg and Zia 2017) with realized treatment effects being larger, on average, than what is being found in recent meta-analyses of financial education field experiments (Miller et al. 2015, Kaiser et al. 2022). Edutainment may be seen as an especially promising avenue for financial education, as marginal costs of these interventions are generally low and the interventions appear to be well suited to operate at scale. While the available empirical evidence suggests that media (especially television and social media) can impact individual field behaviors in a meaningful way, the mechanisms leading to behavior change are less well understood. Against this backdrop, we test leading hypotheses of media impact regarding three distinct channels (a) information provision, (b) role modeling and preference change via emotional connections, and (c) increased salience and basic awareness (see La Ferrara 2016).

To this end, we randomly allocate the individuals to one of four experimental conditions: (a) C: pure control (i.e., no encouragement), (b) T1: monetary encouragement to follow a popular soap-opera ("Un Posto al Sole") with financial messaging embedded into the story line (i.e., the role-modelling channel), T2: monetary incentive to follow a popular quiz show ("L'Eredità") with questions related to financial decision-making posed within the show (i.e., the information provision channel), and (c) T3: a monetary incentive to follow a low-intensity social and traditional media campaign relying on a female fictional character ("Sofia") communicating simple but salient and intuitive messages about financial decision making (i.e., the salience and basic awareness channel).

The main goal of the campaign is to raise awareness among the population regarding the importance of managing their personal finances by promoting multiple initiatives. The campaign, which started in October 2021 and ended in the first half of December 2021 (as shown in Figure 1), was structured in two blocks of activities: (1) a product placement campaign with the inclusion of educational references in the area of planning and financial choices as well as related to the Committee's knowledge within the soap opera "Un Posto al Sole" aired on Rai 3 and the TV quiz show "L'Eredità" aired on Rai 1; (2) a cross media communication campaign on television, radio, print, digital and social media spaces centered on the figure of Sofia, a female character who promotes knowledge and information on economic-financial issues to make more informed choices. The themes that the campaign addresses are related to the importance of financial planning (basic knowledge of concepts related to the following topics: savings, investments, financing, retirement and insurance) to make informed financial choices, deal with unexpected events, plan for the future and take care of one's finances.

The main outcomes of interest are changes in knowledge of the Committee and its activities, financial knowledge, financial attitudes, and financial behaviors. Preliminary results suggest small or zero effects, on average, in all outcome domains except for Committee's knowledge and activities. Yet, a (pre-registered) analysis of heterogeneous treatment effects suggests that women appear to experience a change in institutional knowledge and men reduce their self-assessed financial knowledge.

Keywords: field experiment, financial education, encouragement design

Youth characteristics, overconfidence and take-up to voluntary financial education programs

Alessio D'Ignazio – Stefania Merone

Introduction

Financial literacy worldwide is poor, with negative effects on both individuals and the society as a whole (see, for instance, Jappelli and Padula, 2013; Lusardi, 2019). Although acquiring financial education allows to improve financial well-being, take-up to voluntary financial education programs is usually low, undermining their effectiveness. The issue of low take-up is particularly pressing in the case of young people. Firstly, since young people show lower level of financial literacy (OECD, 2020), they would benefit the most from financial education. Secondly, reaching a large share of young people with financial education initiatives is going to provide a longer stream of benefits in the long run, given that young people are usually really good at spreading information among their peers and relatives (EBA Financial Education Report 2019/2020). Moreover, the training has the function of preparing young people to the increasingly complex financial decisions they will have to make (Frisancho, 2020).

The main causes of low take-up to voluntary financial education programs are of both economic and behavioral kind. For instance, as pointed out by Willis (2008), low-income individuals are less likely to enroll in voluntary financial education programs because their opportunity costs are higher (i.e. they could use those energies and time to work more and earn extra money). While behavioral barriers to take-up might take several forms, according to Yoong (2011) two behavioral biases are particularly relevant: present-bias and overconfidence. In the case of a financial education course, present-bias could lead to procrastination and, hence, low take-up; a similar outcome (i.e., low take-up) would be produced in case of overconfidence, leading to an overestimation of own ability. Such biases, however, do not just affect financial education programs' take-up. They might undermine financial education programs' effectiveness too. A stream of research studies whether financial literacy can be improved by increasing the awareness of cognitive biases, such as present bias. At this regard, Pitthan and De Witte (2022) employ a financial education program to mitigate the myopic behavioral bias and show, by means of a RCT, that it leads to an improvement of financial literacy.

Research objectives

In this paper we investigate the behavioral biases that might affect young people's willingness to participate to a financial education course and focus, in particular, on overconfidence. Arguably, individuals who are overconfident with regards to their own financial knowledge will likely refuse to participate to financial education training activities. At this regard, we investigate whether we can tackle this cognitive bias and boost youths' participation to financial education courses.

The aim of our work is twofold. Firstly, we want to estimate the characteristics that correlate with youths' willingness to participate to a personal finance course, such as gender, curiosity, expected income and financial literacy. Secondly, we want to test the effectiveness of two methods to increase youths' willingness to participate to financial education courses.

To perform this second goal we devise two alternative treatments aiming at tackling the negative consequences of overconfidence, based on the analysis of the Dunning-Kruger effect. In particular, we devise two overconfidence mitigation methods:

1. "Constructive" method: we attempt to mitigate individuals' overconfidence by increasing their knowledge on the subject (by means of a course). This treatment would work well for people on the blue path, while it might harm people on the red one (depending on their initial position on the curve and on how large the increase in knowledge is).
2. "Destructive" method: we attempt to mitigate individuals' overconfidence by asking them difficult finance questions, eventually leading them to lower their confidence. If effective, this form of treatment should produce a downward shift of the Dunning-Kruger curve (all else equal, treated individuals should experience a reduction of their confidence level).

Data and experiment setup

We exploit a rich dataset, obtained from a sample of Italian university students. The university students that participated to the experiment were divided in four groups: three groups were allocated to different forms of treatment, while the latter acted as control group. All groups were administered a questionnaire, following the treatment.

- The first group of students participated to a short course on financial education topics, organised and run by the Bank of Italy, called “Money and Payments”, and then they were administered the questionnaire; we refer to this group as T1 (constructive debiasing method).
- The second group of students did not participate to the course but were administered, instead, an additional set of tough questions on financial matters in the questionnaire. We refer to this group as T2 (destructive debiasing method).
- The third group of students (T3 group) received both the above treatments (i.e., the financial education course and the additional set of questions).
- Finally, the fourth set received the baseline questionnaire only, acting as a control group (C). The questionnaire administered to all groups included at the end a question on whether they would participate to a personal financial education course.

Results

Results show that willingness to participate to a personal financial education course is significantly larger in the case of students enrolled in an economics degree course: this result is somehow expected, since *ceteris paribus* economics students arguably share a greater interest in the subject at hand. As expected, students that display higher curiosity are also more willing to learn about personal finance. More worryingly, students that expect to have a low income in future have a lower willingness to participate to financial education course. Surprisingly, finally, financial literacy does not significantly boost the take-up rate.

Controlling for individual characteristics, students that participated to the Bank of Italy course (T1 treatment) significantly increased their willingness to participate to financial education courses. On the other hand, the second form of treatment (T2), involving the administration of a tough set of financial questions, does not significantly improve the take-up. A possible explanation for this outcome is that students taking such questions received no feedback (it was not possible from a technical point of view to implement a system of feedback) and then, as a consequence, they only partly understood any lack in their knowledge.

Keywords: financial education; behavioral biases; experiments

Understanding of pension system s Intergenerational differences and educational interventions for its promotion

Carmela Aprea

Introduction. Issues relating to reforms of the pension system have a high potential for social conflict, as the massive protests regarding the increase in the retirement age in France or the upheavals related to the pension reform of the Italian government of experts under Mario Monti impressively demonstrate. In pay-as-you-go systems, which are based on intergenerational solidarity and face particular challenges in the face of demographic change, it is reasonable to assume that these lines of conflict also run along generational lines. As research from various social science disciplines (e.g. Jaime-Castillo, 2013; Elinder et al., 2023; Fornero & LoPrete, 2023) suggests, it can moreover be assumed that the acceptance of reform measures also depends on how well the affected citizens understand how the old-age pension system works.

Research object and hypothesis. Against this background and with reference to sociological generational models, psychological concepts of mental representations (e.g., Marr, 2010) as well as economic and political approaches to preference formation (e.g., Druckman & Lupia, 2020), in this study we investigate whether different generations (baby boomers, Gen X, Y, Z) differ in their understanding of the German pension system and to what extent these differences are reflected in the perceived necessity of or preferences for reform measures. In addition, we were also interested in designing a neutral and evidence-based learning intervention for young adults and people starting out in their careers, which should arouse their interest in the topic and promote their understanding of the German pension system. Therefore, a

further research question is how educational interventions on the topic of old-age provision should be designed to support the target group in acquiring a profound understanding of the pension system.

Data/Method. To answer the above-mentioned questions, we use data from the interdisciplinary project "Understanding and attitudes towards old-age provision in Germany" funded by the German Ministry of Labour and Social Affairs, which combines a population-representative quantitative CATI survey (N = 1,000) with qualitative individual interviews (N = 24). Among other things, the survey and the interviews included questions about how well participants understand different parts and principles of the German pension system. Moreover, we asked them about their attitudes towards the pension system as well as about their preferences regarding pension reforms. These data are analysed by means of descriptive statistics as well as ANOVAs (including post-hoc tests), and Chi2 tests, respectively.

As the project had the additional goal of developing ways to translate our findings into practical actions, we carried out three design-based research cycles (Armstrong et al., 2020) to develop and test a digital educational intervention aiming at promoting the understanding of the German pension system. More specifically, this intervention was developed to raise awareness of the necessity to provide for retirement and to overcome the typical misunderstandings identified in the survey and interview studies. As part of a participatory approach, both members of the target group and pension experts were involved in the three research cycles. In the first cycle, the focus was primarily on the content of the educational intervention. Together with a group of student teachers, an initial version of the intervention was developed as part of a university seminar, which was presented to four experts and assessed by them in terms of its suitability for the target group and its accuracy. Based on this feedback, a second version of the educational intervention was created in the next research cycle and implemented with a group of 20 students, who were asked to use questionnaires to assess the extent to which the intervention and its design supported their interest in the subject matter and their motivation to engage with it. They were also asked to indicate which elements of the educational intervention they found particularly appealing or helpful for their learning and where they saw room for improvement. Closed and open-ended test questions were also used to obtain further information on the comprehensibility of the learning intervention. The suggestions obtained in this way were incorporated into the third research cycle, which focused on evaluating the single elements (texts, illustrations, learning tasks) of the intervention. For this purpose, the intervention was presented to five new entrants from different occupational groups, whose learning with the intervention were recorded by using the thinking aloud method. In addition, they were also asked about the perceived ease of learning and asked to answer the test questions already used in the second cycle.

Results/Findings. Our analyses show that the generations differ significantly in terms of their understanding of the pension system, with Gen X respondents showing the highest levels of understanding. In contrast, several specific misunderstandings are evident among younger generations. However, the differences in understanding are not consistently reflected in the perceived need for reform or reform preferences, which points to the importance of other factors. Moreover, we found differences between the generations in terms of their attitudes to the way the pension system works, with younger generations Y and Z significantly less inclined to accept social inequality.

As a result of the development and usability testing study, the so-called pension rally was developed, which is a gamified intervention based on the principle of station-based learning. At seven stations so far, young adults can acquire knowledge about the pension system in a self-organised way. The individual stations support the acquisition of knowledge through a combination of cartoons, explanatory videos, information texts and quiz elements and reflection activities. Randomised field experiments are currently being carried out to evaluate the learning effectiveness of the pension rally. The results of this evaluation study will be available by the time of the conference.

Although the studies presented here relate to the German context as an example, we assume that the findings can be transferred to other countries that are also facing the challenges of demographic change, particularly for pay-as-you-go pension systems. We would therefore also like to empirically analyse this assumption in the future.

Keywords: Financial/Pension literacy, test development and validation, individual difference, financial education intervention

No Train No Gain: Experimental Economics meets Financial Education

Introduction and Motivation

It is widely thought that affecting financial behaviors is the ultimate goal of any intervention of financial education (henceforth, FE). As we read in West & Cull (2020), the fact that bad financial habits can last over time requires educators and economic actors to support young adults in making better financial decisions. Most of the research puts financial knowledge and financial competence on the same level, but academics (Ambuehl et al. 2014) state that it may happen that consumers do not implement these principles or may deploy them incorrectly. This is particularly true in the context of trading apps that have gained considerable popularity over the last few years, exacerbated by the Covid-19 pandemic. These apps, such as Robinhood or Trade Republic, encourage users to trade frequently and allow for low stakes. Nudges, i.e., awards badges or lotteries, further blur the line between trading and gambling. Since trading apps and their gamified interface may have negative effects on decision-making competence (Chapkovski et al. 2023), there is immense potential in using FE to mitigate their dangers for vulnerable users. To achieve this, however, FE must adjust to the specific characteristics of trading apps both in teaching style and content. Testing the effects of such adjustments is the main goal of this study.

Quantitative skills are an important ingredient for financial capability but what emerges from recent results (Horn et al. 2023, Carpena & Zia 2020) is that they play little to no part in financial decision making. Therefore, deepening the nexus between financial behaviors and learning represents a key turning point in this research field to understand barriers and mechanisms that prevent individuals to translate knowledge into actions (Carpena et al. 2019). Yoong (2011) highlights the need of linking FE to concrete actions as far as possible to increase the likelihood of behavioral changes due to it. This principle is also suggested by Amagir et al. (2017), who embed behavioral economics in the design of their FE programs.

A set of interventions aiming at stimulating more conscious financial behaviors is based on active learning principles (e.g. experimental learning). According to Kaiser & Menkhoff (2020), when education is more entertaining or personalized, the impact on financial behaviors appears to be significant. Our paper wants to be a contribution to this strand of literature by conducting an online experimental game that specifically mimics trading app features. The rationale behind the adoption of this approach rely on two main motivations:

- As emerges in Lobao (2020), experimental learning activities such as virtual market participation and trading simulation practices are an effective way to acquire new information. They let participants not only to directly experience a specific situation (both in terms of contents and context) but also to feel specific emotions related to that situation (Linciano et al. 2019);
- Experimental games allow to overcome the classical survey-related issue, where hypothetical questions only lead to self-reported answers about financial behaviors and attitudes. Instead, by providing proper incentives, experiments make participants exert an appropriate level of effort to solve the assigned tasks (Muñoz-Murillo et al. 2020).

The focus of this research is to evaluate one of the many aspects of the financial sphere: asset allocation. From the EFAMA report of September 2020, indeed, it turns out that “the vast majority of EU citizens keep a disproportionate amount of their savings in bank deposits”. Reported reasons are mainly of two types. Firstly, people lack of sufficient financial literacy, and this prevents households from understanding advantages and disadvantages of the different capital market instruments available. Secondly, individuals are often risk-averse, so they prefer saving via bank deposits and insurance products that offer some form of guarantee. Moreover, it has been shown that there exists a direct link between a low level of financial literacy and a low number of assets in the portfolio (Mouna & Jarboui 2015, Abreu & Mendes 2010 and Guiso & Jappelli 2009). At the same time, however, trading apps have been shown to increase investors’ risk-taking behavior through their playful and interactive features (Chapkovski et al. 2021). This, too, can lead to suboptimal investment behaviors, such as over-trading and the disposition effect. Since all these diverse directions of suboptimal behaviors (low number of assets, underdiversification, and over-trading) are relevant in the context of modern investment tools, such as trading apps, it is of first-order importance to understand the role of FE in promoting financial literacy and decision-making competence in these relatively new and hard-to-regulate decision-making environments.

Methodology and Research Questions

In order to investigate the investment strategies in the context of trading apps, we explore the three areas of financial decision-making competence: financial knowledge, financial attitudes and financial behaviors. The latter, in particular, both self-reported and real. More specifically,

- for “financial knowledge”, we consider knowledge regarding basic financial notions such as interest rates, risk, return, portfolio diversification and financial instruments as well as knowledge about risks and dangers of trading;
- for “financial attitudes”, we consider risk attitudes;
- for “financial behaviors”, we consider portfolio investments (diversification, risk-return trade-off, and intertemporal portfolio adjustments).

The interconnectedness among these areas is explored through an RCT. Together with a control group (C), there are two treatment groups, where participants are involved in an experimental game as a simulation-based teaching tool. In particular, those assigned to treatment group 1 (T1) will be given only basic information on financial concepts, whereas those in treatment group 2 (T2) will also get pop-up messages on dangers and biases in investment decision-making during the game (see Figure 1). Additionally, every participant, regardless the group, is delivered a questionnaire before and after the course of the game. These questionnaires enable us to assess financial knowledge, attitudes, and self-reported behaviors, while the experimental game serves as a means to detect real behaviors.

The objective is to investigate how the two different approaches of delivering financial literacy affect financial knowledge, attitudes and behaviors. Moreover, by comparing these different dimensions of analysis, we can derive some insights about how they relate to each other both intra- and inter-groups.

To the best of our knowledge, this represents the first study with the objective to deepen the behavioral implications resulting from a trading app experiment. In particular, we aim at answering the following research questions:

- RQ1. Does an experimental game as a modern FE tool improve participants’ financial knowledge, attitudes, and self-reported behaviors?
- RQ2. Is there any difference in how the two types of teaching techniques affect participants’ financial knowledge, attitudes, and both self-reported and real behaviors?

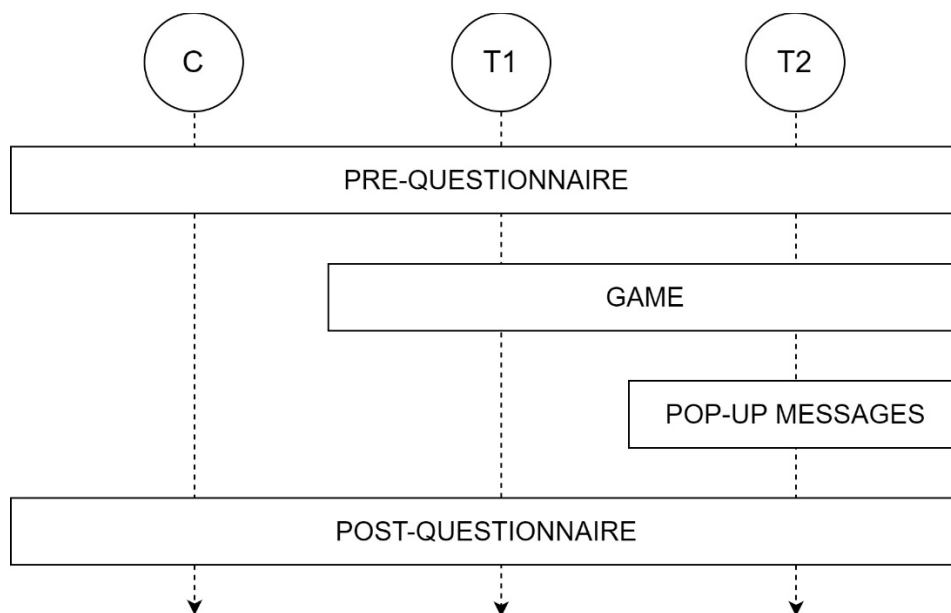


Figure 1. Experimental framework.

Expected Results

While the results of the experiment are expected for April 2024, preliminary tests are underway. If prudence is needed in terms of statistical representativeness, the initial findings appear promising. The experimental game exhibits a capacity to improve financial knowledge, attitudes, and self-reported behaviors. Furthermore, while no statistically significant disparities have emerged between T1 and C in relation to these aspects, a notable contrast is apparent in the design of T2. The presence of pop-up messages seems to lead to the cultivation of more diversified portfolios and fiscally prudent behaviors.

Consequently, should these observations be validated, the implications are twofold. First, the experimental game itself demonstrates the potential to impact comprehension of financial subjects, serving as an effective educational tool. Second, the proposed trading simulation can be expanded into a comprehensive platform, offering users a structured environment for practice and learning, before potentially transitioning to real investment scenarios.

Keywords: financial education, financial literacy, RCT, economic experiment, trading app

Financial knowledge and career aspirations among the young: a route to entrepreneurship

Massimiliano Stacchini – Sara Lamboglia – Noemi Oggero – Mariacristina Rossi

The importance of financial literacy in shaping economic behavior has been studied across various domains of household economics and finance (Lusardi & Mitchell, 2014; Kaiser et al., 2022). A solid understanding of financial literacy leads to positive economic outcomes, such as effective retirement planning (Lusardi & Mitchell, 2011; Choinière-Crévecoeur & Michaud, 2023), wealth accumulation (Lusardi et al., 2017), and efficient portfolio allocation (Abreu & Mendes, 2010; Bianchi, 2018).

Financial literacy also correlates with increased participation in the stock market (Van Rooij et al., 2011; Jappelli & Padula, 2015; Bucher-Koenen et al., 2021) and improved debt management (Lusardi & Tufano, 2015; Lusardi et al., 2020; Thorp et al., 2023). Additionally, studies have shown links between financial literacy, consumers' digital skills (Lo Prete, 2022; Marconi et al., 2022), and small businesses' engagement in digitalization and sustainable activities (D'Ignazio et al., 2023).

The paper analyses whether financial literacy contributes to steering young people's career aspirations, specifically analysing the interest in becoming an entrepreneur. Young people's aspirations are important as they can influence educational and occupational choices, affecting individual well-being and social mobility across generations.

The link between financial literacy and career aspirations and entrepreneurial spirit is scarcely explored in the literature due to limited data on these dimensions. Cumurović & Hyll (2019), Oggero et al. (2020), and Struckell et al. (2022) have investigated the correlation between entrepreneurship and financial literacy but haven't examined the intention to become an entrepreneur or established causality. To our knowledge, we are the first to assess whether greater financial knowledge influences decisiveness regarding career aspirations, including entrepreneurship.

To investigate the link between financial literacy and professional aspirations we use a dataset representative of young adults aged between 18 and 34 years old in Italy collected by the Bank of Italy in 2023 (Bank of Italy, 2023). The dataset includes 5,188 young individuals whose professional aspirations have not yet been realized or who have recently entered the workforce. It surveys professional ambitions such as the desire to become an entrepreneur as well the reasons why such desire may be absent. The data include information about the financial knowledge of respondents and a broad set of demographic and socio-economic indicators and individual preferences, particularly concerning risk tolerance.

Regarding career aspirations, 38% of young adults would like to start an entrepreneurial activity in the future, while 11% report they 'do not know'. Among the remaining respondents, reasons for not pursuing entrepreneurship are varied: some perceive it as too risky (15%), lack the necessary skills (14%) or financial resources (10%), while 11% are satisfied with their current work situation.

As far as financial knowledge is concerned, the survey asked five questions. Besides the basic concepts of simple interest rate, inflation, and risk diversification that are the topics of the well-known "Big Three" questions (Lusardi & Mitchell (2014)), the survey tested the knowledge of mortgages and interest compounding. The average number of correct answers was 3 out of 5 and only 17% of the sample correctly answered all the five questions on financial knowledge.

Through instrumental variable methods, we can establish causal relationships showing that financial literacy can shape career aspirations towards entrepreneurship. We use parental education level as instruments for individual financial literacy, considering whether the respondent's mother and father hold at least an upper secondary school diploma. Our assumption is supported by literature documenting the strong association between parental education and children's financial literacy (Lusardi et al. (2010)). Our

hypothesis is that parental education contributes to fostering a cultural environment within the family that supports the development of various forms of knowledge, including financial literacy, without directly influencing the choice to pursue an entrepreneurial profession. A placebo test is conducted to support our assumption.

The instrumental variables are used within a linear probability model, with entrepreneurial intention (a dummy equal to one if the respondent states s/he would like to become an entrepreneur) as the dependent variable and individual financial literacy (a score based on the five questions) as the main covariate. The model includes a vast set of individual-level controls and province-level dummies to account for differences in the entrepreneurial environment and geographical disparities.

Our estimates reveal that a one standard deviation increase in financial knowledge raises the willingness to become an entrepreneur by 25 percentage points. The paper also tests whether preferences play a role in ambitions and considers risk tolerance among the possible drivers. The literature has shown that risk aversion can be a deterrent against becoming an entrepreneur, as documented by De Blasio et al. (2021), and it is likely to act in the same direction also in terms of professional ambitions. Indeed, we show that risk tolerance is positively correlated with entrepreneurial intentions.

We extend the analysis and investigate how financial literacy associates with all information regarding professional intentions, including the response 'I don't know' as well as the reasons provided by those who are not interested in becoming entrepreneurs. The main result of this analysis is that financially literate individuals exhibit lower indecisiveness regarding future professional choices.

These results have significant policy implications, as they support the idea that financial literacy, along with related educational initiatives, can play an important role in guiding young people in planning their future careers by helping them achieve clarity about their professional aspirations.

Keywords: Financial literacy, Entrepreneurial Intention

SESSION 4. INVALSI DATA FOR IMPACT ASSESSMENT: LIGHTS, SHADOWS, AND TOTENTIAL DEVELOPMENTS

**ORGANIZER: UNIVERSITY OF MILAN BICOCCA - CATHOLIC UNIVERSITY OF MILAN -
UNIVERSITY OF BOLOGNA & FBK-IRVAPP**

COORDINATOR: GIANLUCA ARGENTIN

18TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 2 – RESEARCH 4}

Teachers' Characteristics and the Reproduction of Educational Inequalities: What Influences the Formulation of Guidance Advice

Elisa Manzella

The Italian educational system, like other Western countries, is characterized by a high degree of educational inequalities (Schizzerotto & Barone, 2006) and tracking in secondary school is an important factor of their reproduction (Gasperoni, 1997; Brunello & Checchi, 2007; Triventi et al., 2016). Indeed, upper secondary education track choice has a significant impact on students' further academic and occupational careers. Access to tracks differs by socioeconomic and cultural background, conditioning the choice (Gasperoni, 1997). Information gap plays a crucial causal role in this process, as shown by a previous RCT (Barone et al., 2017). In order to get a conscious educational choice, it is required an effective school guidance, a key factor for the success of students' educational paths, especially the ones from disadvantaged social background. Hence, an effective school guidance may also increase equity of the entire school system. In Italy a national school guidance system is still lacking, with the exception of a teacher-formulated guidance advice (which is not mandatory for families). Usually, the advice focuses on the choice among different tracks of upper secondary school. The guidance advice, while operating nationally, is extremely heterogeneous in its implementation.

Previous quantitative studies have highlighted the existence of robust associations between students' backgrounds and teachers' advice (Checchi, 2010; Bonizzoni et al., 2014; Argentin et al., 2017; Bonvini et al., 2023), suggesting it to be biased. Causal support for this also comes from more recent analyses based on vignette studies (Bonvini et al., 2023). Previous studies, also because they are based on small sample sizes

and limited databases, provide little information on the underlying factors of the advice and do not consider the characteristics of the actors formulating the advice themselves, namely the teachers. International studies also do not sufficiently explore the phenomenon, relying on small convenience samples, with pre-service teachers and/or indirectly measuring the existence of stereotypes and prejudices (Batruch et al., 2023). Even previous qualitative studies, which analyzed the advice's formulation processes, highlighted teachers' considerations on extracurricular and ascriptive elements (Romito, 2016), still leaving out the characteristics of the teachers themselves from the reasoning. Overall, therefore, previous studies have the important limitation of not investigating which characteristics of teachers influence their formulation of advice.

The purpose of this study is to test whether teachers' characteristics influence the formulation of guidance advice and related biases. For this purpose, data collected within the "Orientare alla Scelta" project (Manzella & Argentin, 2024) are used. In this work, a controlled randomized trial was implemented to assess whether teachers' biases can be reduced by increasing their awareness through in-service training, involving 196 Italian lower secondary schools. To pursue the knowledge objective of this work, we constructed a specific database composed of data collected through an online survey to teachers from the 196 randomized schools, linked to administrative data from INVALSI and the Ministry of Education - National Student Registry, specifically about their grade-8 students in the academic year 2021/2022 (19,136 cases). More precisely, the data include students' academic performance, final exam grades, INVALSI test scores (Italian, mathematics, and English), guidance advice formulated by teachers, actual enrollment choices for the academic year 2022/2023, and socioeconomic characteristics of students and some of their teachers.

The collected data are used in this work to test: i. how teachers' advice toward upper secondary school varies according to their characteristics; ii. if these characteristics are also correlated with teachers' biases in the provision of guidance advice.

We then analyze the influence of teachers' characteristics (age, gender, geographical area of the school where they teach, previous educational and school track, perception of guidance self-efficacy, etc.) on the guidance advice formulated to students.

The analyses conducted so far show that the most relevant characteristics of teachers are related to their previous experience in guidance, as well as their knowledge of different school paths and their direct experience of a non-lyceum upper secondary school.

Implications are drawn on the type of targeted training policies that can be implemented for different groups of teachers tasked with guiding students through a crucial transition in their educational journey.

Keywords: upper secondary school choice; educational inequalities; teachers' characteristics

Educational disparities among native and immigrant students in Italy: analysis from pre-pandemic to post-pandemic periods

Francesca Carimando – Loris Vergolini – Eleonora Vlach

Introduction

The COVID-19 pandemic has influenced educational systems globally, uncovering and, in some cases, intensifying existing disparities. Italy's educational landscape offers a compelling case for study, particularly because of its substantial population of immigrant students. This paper builds upon foundational research, such as the insightful analyses by Azzolini et al. (2012) on educational disparities and Borgonovi & Ferrara (2023) on the pandemic's specific influences within Italy. Moreover, Engzell et al. (2021) have documented the broader phenomenon of learning losses attributable to school closures, highlighting the necessity of investigating these effects within specific national contexts. This research aims to enrich the understanding of the pandemic's long-term influences on educational equity in Italy, leveraging data from the Italian National System for Student Performance (INVALSI) and insights from Triventi et al. (2022), who have documented the multifaceted challenges immigrant students face within the educational system. Through this analytical lens, the study endeavours to explore the nuanced educational performance gap between children of natives and children of immigrants, aiming to elucidate gaps underscored by the pandemic.

Aim of the Research

The foundation laid by the introduction steers this study towards dissecting the educational performance gap between children of natives and immigrants in Italy, spanning the period from before the pandemic, through its peak, and extending into the ongoing post-pandemic recovery phase. By employing data from INVALSI, this research delves into the pandemic's influence on these disparities, with a focused examination of control variables such as socioeconomic status, gender, and language spoken at home. With a focus on the 2022/2023 academic year, the research aims to investigate the educational system's recovery mechanisms, providing insights into how Italy is confronting and potentially mitigating the challenges and disparities unveiled by the pandemic.

Data

The investigation utilised population data furnished by INVALSI on students enrolled across all educational levels (i.e., grades 2, 5, 8, 10, and 13) from the 2018/2019 to 2022/2023 academic years. Employing population data, rather than a sample, ensures a robust and representative analysis of educational outcomes. Additionally, the study integrated test outcome data with responses from dedicated student questionnaires, enriching the dataset and enabling a more detailed examination of the factors influencing academic achievement.

The study selected cohorts from the 2018/2019 to 2022/2023 academic years to trace the trajectory of academic performance through periods of stability, disruption, and recovery. INVALSI's anchoring process ensured the comparability of test scores across these cohorts, facilitating the analysis of the educational landscape's evolution.

Analytical Strategy

The investigation's focus rested on the outcomes of the INVALSI assessments in reading and mathematics as the dependent variables. The primary independent variable was the migration background of students. This variable was categorised into four groups based on the birthplaces of students and their parents, namely, Natives (students born in Italy with both parents born in Italy), Mixed (students with one parent born in Italy and the other abroad, regardless of the student's birthplace), Second-generation (students born in Italy with both parents born abroad), and First-generation (students and both parents born abroad). The study also incorporated a variety of control variables, such as gender, socioeconomic status (assessed through the ESCS index and, for second-grade students, parental education and occupation), year of birth, early education attendance, and the language spoken at home. These variables offered additional insights into the disparities in academic performance, beyond those attributable to migration background alone.

The research began with an examination of descriptive statistics to outline performance disparities across different student groups. Subsequently, OLS linear regression models, adjusted for the specific grade levels and the availability of certain control variables, were employed to isolate the influence of migration background on educational achievements. This approach facilitated a nuanced investigation into the factors influencing educational disparities.

Results

The exploration of ethnic education disparities through the COVID-19 pandemic lens narrates a story of resilience, decline, and partial recovery. Data from the pre-pandemic 2018/2019 academic year unveiled a significant educational achievement gap, with children of natives outperforming their immigrant counterparts across all grade levels and subjects. This pre-existing gap set the stage for understanding the pandemic's subsequent influence on educational outcomes. Students from mixed backgrounds exhibited marginally lower performance, closely trailing children of natives. Among students with a migration background, Second-generation students generally achieved higher academic success compared to First-generation students.

As the pandemic unfolded, an initial display of resilience among students hinted at a possible mitigation of learning disruptions, particularly in the early grades. However, this resilience was fleeting, giving way to performance declines among children of immigrants as the pandemic persisted. This trend was especially pronounced in the lower grades. In higher grades, although signs of resilience were observed initially, especially at grade level 8, students experienced significant declines, leading to a growth of the educational achievement gap between children of natives and children of immigrants. These findings underscored the pandemic's multifaceted influence on students, revealing a complex interplay of resilience and vulnerability.

In the post-pandemic period, specifically in the 2022/2023 academic year, signs of recovery began to emerge across various student groups and grade levels, although this trend was uneven. The lower grades continued to underperform, with First-generation students especially enduring notable learning losses. This emphasised the enduring challenges faced by younger learners, underscoring the necessity for focused

support and resources to tackle the unique obstacles to educational equity. Conversely, higher grades exhibited signs of narrowing educational disparities, particularly among Second-generation students, who demonstrated improvements in both reading and mathematics. This trend suggests a gradual recovery process in educational outcomes following the pandemic.

Conclusion

This paper aimed to dissect the intricate landscape of educational disparities in Italy through the COVID-19 timeline. While extensive research has documented educational disparities, this paper addressed critical gaps in understanding the pandemic's long-term influence and the unique challenges posed by the native-immigrant dichotomy in education. By providing empirical evidence and insights into recovery efforts, this study contributes to a broader comprehension of educational equity, urging the implementation of comprehensive strategies to support all students in a post-pandemic era. The findings revealed not only a deepening of pre-existing gaps, but also the emergence of nuanced recovery patterns in the 2022/2023 academic year.

As Italy navigates the post-pandemic phase, this study underscores the critical role of continuous, data-driven evaluation in shaping inclusive policies and practices. Addressing the multifaceted educational challenges unveiled by the pandemic calls for a comprehensive approach that prioritises support for at-risk students—particularly in the critical stages of early and middle school years—through targeted interventions like enhanced language support and socio-emotional learning programs. Implementing such strategies offers Italy a pathway not just to alleviate the pandemic's immediate influence but also to forge a more resilient and equitable educational system.

Keywords: Ethnic Educational Gap; Generational Status; Italy; INVALSI; COVID-19

Do strict schools yield better academic results? Propensity to retain in high schools and students' academic achievement

Benedetta Valpreda – Dalit Contini – Guido Salza

Introduction

Grade repetition is a measure used in many countries to manage situations of poor academic performance by some students. In theory conceived as an opportunity to address educational deficits before proceeding in the academic path, often retention is perceived (and experienced) by students as a demonstration of incompetence or even personal "failure".

Scientific research has often debated the usefulness of retention as a recovery tool. Most of the existing literature focuses on the effect on students at risk of retention. In countries where this measure is allowed in primary schools, some studies show that retention has positive effects on the academic performance of retained students, although these effects often diminish after the first years (Greene & Winters, 2007; Schwerdt & West, 2012; Alet, Bonnal, & Favard, 2013). In high school, however, the prevailing evidence is that retention has negative effects on students' careers and well-being. For retained students, dropout rates increase (Baert, Cockx, & Picchio, 2015; Contini & Salza, 2024), the likelihood of enrolling in university decreases (Fine & Davis, 2003; Ou & Reynolds, 2010), and self-esteem, motivation, and self-concept worsen (Mathys, Véronneau, & Lecocq, 2019; Klapproth, et al., 2016).

While the effect of retention on retained students is often debated, evidence on its effects on the general student population is even more limited. Indeed, the policy of retention, simply by existing, can also have an effect on everyone. The presence of retention can serve as an implicit warning: those who do not achieve certain results risk repeating the academic year. If this situation is perceived by students as a challenge to be faced, the "threat" effect of retention can encourage students to engage more in their studies to avoid the risk of being retained; conversely, if the demand to achieve certain results is perceived as unattainable, the threat of retention can have the opposite effect and provoke anxiety, stress, and strong discomfort, especially among the most sensitive students. Grade repetition, moreover, can have a selection and allocation effect of students in 'schools suitable for them', thus making the classroom environment more homogeneous. Following this line of inquiry, the present study asks: what is the effect of the existence of retention on students who do not exhibit low academic performance?

Previous Literature

Empirical evidence on the indirect effects of retention is limited and, overall, the available literature does not provide clear indications. Koppensteiner (2014) studies the effect of a policy in Brazil that introduces automatic promotion in primary schools. His results show that removing the possibility of retention significantly reduces students' performance in the fourth year, with a more pronounced negative effect for students who would have been more at risk of repeating the year. This suggests interpreting the effect of automatic promotion (i.e., the absence of retention) as a disincentive to students' efforts. A different conclusion is reached by a study on the reintroduction of retention in high schools in Belgium (Belot & Vandenberghe, 2014). Indeed, the authors highlight that in some cases academic performance improves but typically, for students most at risk, the effect of the reform is not significant. Jacob and Lefgren (2004) study the effects of a policy introduced in the Chicago district that conditions student promotion on their outcome in a standardized test. Their results indicate that the policy has a positive impact on the following year's performance for third-grade students, but this effect fades when considering academic outcomes after two years, while for sixth-grade students the effect on subsequent performances is essentially null. In Italy, Battistin and Schizzerotto (2019) analyze the impact of a policy with similar elements, namely the 2007/2008 reform that (re)introduces the possibility of retaining students in September after remedial exams. The authors highlight a worsening in academic results in Italian and science in technical and vocational institutes, while no significant effects are observed in traditional high schools. Our article contributes to this literature, aiming to estimate the effects of retention on the population of students in the Italian context.

Data and Analytical Strategy

The ideal setting for evaluating the indirect effects of retention is the introduction of a policy that introduces or limits its use. This allows comparing similar students in different contexts: one where retention is a possibility and another where it does not exist. Estimating a causal effect is more complex when there are no significant variations in school policies and retention is assigned in the absence of a transparent threshold (e.g., based on the result of a standardized test) and with a considerable degree of discretion by teachers.

Given the peculiar characteristics of the Italian institutional setting, in line with Contini and Salza (2024) and Valpreda (2023), this work focuses on the indirect causal effect of retention on learning, exploiting the variation in schools' "strictness," understood as the propensity of schools to retain students, net of the observable characteristics of the schools themselves and of the students.

The research uses a rich longitudinal dataset that includes the population of students who started grade 9 (first year of high school) in the three most populous regions of Northern Italy (Lombardy, Piedmont, Veneto) during the academic year 2014/15. The data cover a period from the academic year 2013/14 to 2016/17. The dataset contains a wide range of information on students and schools, coming from two main sources: the National Student Registry (ANS) and the National Institute for the Evaluation of the Education and Training System (INVALSI) registry. This information includes details about schools and students, as well as end-of-year grades in Italian, Mathematics, and conduct, and the results of the standardized INVALSI tests for eighth and tenth grades. The dataset also collects information on students' family backgrounds.

Considering that, as already observed, in Italy the decision to retain is discretionary, schools may differ in "strictness." It is therefore possible to exploit the variability in strictness among schools, comparing the level of learning of promoted students in more severe schools with others comparable in less severe schools. The population under analysis consists of students attending grade 9 in 2014/15 who are successfully promoted to grade 10 and thus participate in the INVALSI tests.

However, strictness is not observed and must be estimated. The starting point is a multilevel logistic model for the probability of being retained, with random intercepts at the school level. Within each school, the decision to promote or retain students depends mainly on the vector of the student's current abilities, which, given the richness of the measurements observed in our dataset, captures a wide range of cognitive and non-cognitive abilities. However, some schools retain more students than others, even with the same student characteristics. The strictness of the school is defined by the random component of the model, conceptualized as the propensity of the school to retain net of the observable characteristics of the school itself and its students.

Once the strictness of each school is estimated, it is possible to estimate the main model, which has as the dependent variable the score obtained in the INVALSI test in grade 10, i.e., the academic year following the potential retention, and as the main explanatory variable the strictness itself. The logic is as follows: if among the promoted students learning is better in more severe schools, retention can be considered a measure that allows preserving the level of learning of all students.

It is appropriate to note that the identification of the causal effect of retention on learning is based on the assumption that strictness can be considered an exogenous treatment. In the work, the conditions for which such assumption can be considered valid are discussed, solutions are hypothesized to overcome the problem if such assumptions were not to be fully convincing, and the direction of the distortion in this latter case is analyzed.

Preliminary Results

In general, the strictness of the school does not significantly influence INVALSI performances in grade 10. The effect of strictness on the Italian standardized test is statistically indistinguishable from zero. However, considering different types of schools separately, in the case of non-traditional high schools (artistic and human sciences), a positive effect is found: with the same individual characteristics and school context, as the strictness level of these schools increases, performances in Italian increase. The picture is similar considering the dependent variable as the tests in Mathematics. As a robustness check, the analyses were repeated excluding very small schools, with very similar results. However, limiting the analysis to students who would have been promoted even in the most severe schools, the effect found within non-traditional high schools remains positive and significantly different from zero but decreases in its intensity.

In conclusion, the (still preliminary) analyses suggest that retention, in addition to having negative effects on retained students (Contini & Salza 2024), in the majority of cases does not even contribute to preserving the level of learning of others.

Keywords: grade repetition; academic performance; policy evaluation

Reflections on teachers' practices for an effective and conscious use of standardized assessments

Maria Chiara Cibien – Federica Ferretti – Stefania Pozio – Matteo Viale

The systemic impact of standardized assessments has been well studied for a long time. More recent is the interest on their impact on classroom practices, except for the discussion on teaching to test. There is evidence that there is an increasing use of materials from national standardized assessments in pre-service and in-service teacher professional development programs, but no organic research has been performed in the Italian context (and very few in the international one) on this topic. This research aims to fill this gap.

The main outcome of this project is the development of the model for pre-service and in-service mathematics teachers, based on an effective and conscious use of standardized assessments. We plan to develop a model for teachers' professional development, and ad hoc materials to promote the professionalism of teachers in the use of standardized assessment within teaching practices.

Mathematics teachers' professional development has become a research topic that entails theoretical, methodological and educational issues. Several researches (Shulman, 1986; Wenger, 1998; Ball, Thames, & Phelps, 2008; Jaworski, 2006; Da Ponte & Chapman, 2006; D'Amore & Fandiño Pinilla, 2009; Bartolini Bussi et al. 2017; Carrillo-Yañez et al. 2018) have developed different theoretical perspectives and methodological design that allow teachers to develop their specialized knowledge in order to implement in their school practice effective and forefront teaching methodologies and to create new cultures of mathematical activity. Our study develops within this line of research focusing both on a theoretical perspective for the design of effective methodologies and on the introduction of tools and resources used and designed for teacher collaboration.

The framework for the design of the models and the development of tools interfaces three main components and hence results from the networking of three well-established theories.

The first two are Jaworski's notion of Community of inquiry (Jaworsky, 2006) and the Mathematics Teacher's Specialized Knowledge (MTSK) model (Carrillo-Yañez et al., 2018). We believe that these perspectives take into account two defining characteristics of a teacher PDP model. On one hand we need to conceptualize appropriate social interaction within a community of practice between teachers and between teachers and researchers. We consider teachers' professional development a developmental process that entails teacher change (Guskey, 2002) in terms of a transformation of beliefs and attitudes regarding mathematics, teaching-learning processes, the students and the political and social role of the education system. Within this frame, assessment is a key factor. Such a change cannot be a solitary,

individual and autonomous process, instead it is constitutively a sociocultural activity whose outcome is the transformation of the individual's identity as a teacher. On the other hand, we need to conceptualize and outline the specific knowledge and professional skills that we would like teachers to achieve as a result of the change they undergo in their training process. We are referring to a wide range of knowledge that includes mathematics, epistemology, pedagogy, didactics, psychology. Sociocultural development within a community of practice requires tools and, in a broader sense, resources that mediate the activity, contribute to the interpersonal exchange within its actors and bring to the fore cultural and conceptual objects. Adler (2000) claims that mathematics teacher education should focus on resources intended as instruments that allow teachers to broaden their impact in the classroom mathematics activity.

As a third component, we use the Meta-didactic Transposition model of Arzarello and colleagues (2014). It will allow to describe and interpret variables in teacher education processes and to be able to analyze their reciprocal relationships and their evolutions. This model considers the practices and processes that are being developed during teacher education programs. Mathematical knowledge lives in the institutional dimension where mathematical objects emerge from socio-cultural activities shared by individuals belonging to one or more institutions. The relationship with mathematical knowledge is both personal and institutional (Chevallard, 1992). The development of mathematics, the teaching and learning of mathematics and the teacher education are characterized by the dialectic between the personal and institutional relationship with knowledge.

In the process of meta-didactic transposition, educational practices and theoretical lenses are shared between teachers and teacher educators. The term 'meta-didactics' emphasizes that these practices consist of reflections on teaching activities. These reflective actions can be facilitators of particular practices involving different types of tasks (in our case, the a priori analysis of mathematics INVALSI tasks), together with the techniques available to solve them (e.g. the development of methodologies and patterns for analyzing INVALSI tasks). Within the meta-didactic transposition, practices become meta-didactic practices as they refer to the practices and reflections that characterize teacher education processes. Meta-didactic practices refer to practices and theoretical reflections developed within a community of inquiry in the sense of Jawroski (2006).

Our model can be seen as an instance of the Meta-didactic transposition model. As far as the institutional dimension is concerned, our model will involve the Italian Ministry of Education through the INVALSI institute, Italian schools and universities; the role of mediator is typically played by researchers involved in our project, who are responsible for courses in Mathematics Education for pre-service teachers at university level and for teachers' professional development programs for in-service teachers. Researchers and teachers share praxeologies and reflect on them. A key element is the intermediation process carried out by INVALSI tests and related tools, which acts as a boundary object and facilitates border crossings between communities.

According to our framework, the models and the materials that the project will develop and validate will have these key features:

- a focus on Mathematics contents, whose learning macrophenomena are enlightened by the INVALSI test;
- a focus on transversal skills, in particular argumentative skills of students;
- a focus on transdisciplinary aspects, such as the role of language in mathematics teaching and learning, for students and teachers;
- a focus on professional skills of teachers, in particular assessment;
- a focus on the role of media, in particular how far a switching from paper-and-pencil assessment to a computer-based one, at national level, entails changes all along the teaching-learning flow;
- a focus on the use of tools such as websites and repositories (invalsiopen.it; gestinv.it);

We expect a better understanding of teachers' beliefs, attitudes and needs about assessment and standardized assessment in particular. This is of particular importance, since assessment is a key component of teachers' identity, that catalyzes beliefs, practices, and implicit assumptions. Assessment makes learning visible (Hattie, 2009), but we can also say that it makes teaching visible.

This is a part of a larger project "Mathematics standardized assessment as tool for teachers' professional development", that won the PRIN 2022, funded by the European Union and the Italian Ministry of Education and University.

Keywords: Large-scale international assessments, teachers' professional development, classroom practices, mathematics education.

SESSION 3. WEAKNESS AND EXCELLENCE IN EDUCATION: FOR A POLYSEMOUS AND MULTIDIMENSIONAL IDEA OF EQUITY OF EDUCATIONAL OPPORTUNITIES

ORGANIZER: INVALSI

COORDINATOR: PAOLO BARABANTI

18TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 3 – RESEARCH 5}

Academic and socio-emotional skills of high achieving students. An analysis of both cognitive tests and student questionnaire in OECD-Pisa 2022

Paolo Barabanti – Laura Palmerio – Carlo Di Chiacchio

Introduction

High achieving students have been for long a marginal issue within the wider debate on learning processes in the European context and in the Italian school system as well. In recent years, however, in addition to the constitutional duty to help weak and disadvantaged students, there has been an increasing to the institutional duty to promote talented students as a “right to excellence”. This change of perspective is driven not only by a multidimensional concept of equity (Benadusi & Niceforo, 2010) but also by the need of exploiting, and not wasting, students’ potential, useful for the whole community (OECD, 2009) in terms of meaningful contribution to society and as a key for economic competitiveness (European Commission, Directorate-General for Education, Youth, Sport and Culture, 2024).

Although there isn’t a common definition, reflecting overall the characteristics of “good students” already highlighted about half a century ago by Parsons (1972), nowadays teachers’ and school principals’ opinions converge to a “plural” idea of excellence in which particular cognitive and performance skills are combined with specific behavioural characteristics, attitudes and orientations (Barabanti, 2018). The most shared idea is the one that considers as a high achiever a student who has not only academic skills but also socio-emotional ones, qualities which have both important effects on education, health, success in adult private life and on the future job market (Borghans et al., 2008; Chetty et al., 2011; Moffitt et al., 2011; Heckman & Kautz, 2012; Giancola & Lovecchio, 2018; Maccarini, 2019).

Over recent decades, large-scale evaluation programmes have globally received a great deal of attention because they can provide policy makers and researchers with useful and comparative information on educational outcomes. In this type of survey, traditionally, large amounts of background information are collected, by means of various questionnaires, on both the individual and the family and school levels. They often serve as a pioneering ground for experimentation within innovative research areas (see e.g. creative

thinking in OECD-PISA 2022), both from a cognitive and affective-behavioural perspective. Therefore, they may also be of great importance in order to investigate both dimensions of high achieving students.

Research Object

This paper has two aims. The aim of the first theoretical part of the study, building on the literature and previous surveys and research, is to offer a few comments in order to try to question whether and how it is possible to use large-scale assessments to properly investigate both dimensions of excellence. The second part of the paper wants to provide some information about this group of students by analysing international large-scale assessment.

The main research questions are: Are international surveys a useful tool to shed light on both dimensions of excellence? Which social and emotional skills can be investigated and by means of which tests and questionnaires? Do students who scored higher in academic skills also reported top levels in the social and emotional ones? Are there significant differences among Italian students and students from other European Union countries?

Data and Method

OECD PISA data (edition 2022) from both cognitive tests and student questionnaire will be used. The academic dimension of excellence can be operationalized by referring to “top performers”, that is students with high ability who reach level 5 or level 6 (the highest ones) in one of the three literacies assessed (reading, mathematics or science). The social and emotional dimension may be highlighted referring to students who are in the fourth quartile in some indices about emotional-relational-social domains constructed with sets of items of the student questionnaire.

Results and Preliminary Findings

The dimension of excellence related to academic skills is more easily quantifiable and measurable, being at the core of several large-scale standardized surveys at national and international level (Viteritti, 2018). PISA has been extensively measuring student achievement for over 20 years. But academic performance is only one aspect of success at school (and in life); another is the one related to social and emotional domain (Gardner, 1993; Goleman, 1996). Unlike academic skills, the socio-emotional ones are more difficult, although not impossible, to be operationalized (Di Francesco et al., 2015). The student questionnaire is a useful tool to investigate the socio-emotional dimension because it includes some sets of items which are related to the Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism): the five domains investigated in PISA are Open-mindedness, Task performance, Engaging with others, Collaboration and Emotional regulation.

The items that make up the social and emotional skill scales are statements about the student’s emotions, attitudes and behaviours on which they are asked to report their agreement, using five response options ranging from ‘strongly disagree’ to ‘strongly agree’. The responses to the items belonging to the same skill are standardized and summarised with a score on a psychometric scale.

Preliminary findings show that Italy does not have a statistically different share of top performers in mathematics literacy compared to other EU countries. With regard to the social and affective dimensions, Italian 15-year-old students belonging to the group with the highest values are present to a lesser extent in all indices, while equally present in the Open-mindedness (Curiosity) index.

Keywords: high achieving students, academic skills, socio and emotional skills, OECD-PISA tests

Exploring the Influence of Maternal and Teacher Role Models on Gender Differences in Mathematics

Alice Bertoletti

Significant gender disparities favouring boys in mathematical performance persist in Italy. The recent PISA 2022 results reveal that, among all 81 school systems analysed, Italy reports the largest gender difference between boys and girls in mathematics. Fifteen-year-old male students outperform their female counterparts by 21 points (the OECD average is only 9), equivalent to one year of schooling (OECD, 2023). The underachievement of girls in mathematics raises concerns as it influences their career choices, with female students less likely to opt for STEM (Science, Technology, Engineering, and Mathematics) field careers (Cheryan et al., 2017; Stoet & Geary, 2018). Consequently, gender inequality in mathematics

remains an unresolved issue affecting female representation in certain labour market fields, salary disparities, and the persistence of gender stereotypes in career choices (Weiner et al., 2010).

Evidence from INVALSI data indicates that these gender differences in mathematics achievement become more pronounced as Italian students' progress through the educational system. Notably, a gender gap in math becomes evident as early as grade 2, where boys achieve an average score of 193 compared to girls who score an average of 190 (see INVALSI, 2022). This gender gap widens to over 13 points by grade 5 and persists consistently throughout students' academic careers (see INVALSI, 2022).

Among possible channels that can explain gender differences in education, a recent stream of the literature highlights the relevance of role models in inspiring students to overcome gender stereotypes and pursue careers in STEM fields, particularly parental and teacher role models (De Gendre et al., 2023).

Based on this, the present paper focuses on the influence of role models in explaining and mitigating the gender gap in mathematical achievement. Specifically, the aim is to investigate the influence of both mother and teacher role models on student performance, with a particular focus on the standardised math test scores of female students in Italy.

In detail, the study addresses the following research questions:

- What is the extent of the gender gap along the Italian students' careers?
- How effective are female role models in influencing the standardised math test scores of girls compared to boys?

The paper employs a longitudinal analysis of INVALSI data, covering the period from the school year 2012/13 until 2018/19, which represents the last pre-pandemic wave. Both maternal and female-teacher role models have been investigated in this study. When examining the role of mothers, data from a cohort of over 400 thousand students from second to fifth grade and third grade middle school were analysed, resulting in a total of over 1.2 million observations. This sample corresponds to the overall population enrolled in second grade during the 2012/13 academic year and can be followed up to third grade (approximately 13 years old). When considering the role of teachers, teacher questionnaire data were matched with student data. Since information about teachers is only available for the representative sample of the population (INVALSI sample), three different student cohorts were considered to increase the sample size. Students were followed from second grade (starting from the 2013/14 cohorts, the first cohort available for the teacher questionnaire) to fifth grade (up to the 2018/2019 cohort). This allows for studying the effect of the role model played by teachers in a homogeneous context, within the same primary school, while eighth grade would involve the introduction of not-observable factors, given the transition from elementary to middle schools.

The panel structure of the data allows for a more specific understanding of the evolution of the gender gap and the influences played by role models. The longitudinal dimension of the data represents an important novelty of this study. Exploiting the panel structure of the data, analyses adopt a dynamic panel approach as a baseline model, which considers past students' achievement as a regressor. This approach enhances the accuracy of the gender gap estimates (Contini and Grand, 2015; Contini et al., 2017).

Results reveal that the gender gap in mathematics for Italian students exhibits significant variation across regions and grades, with opposite patterns between the North and South of Italy. Maternal role models seem to play an important role, with maternal characteristics exerting a larger influence on female performance in mathematics compared to that on males, with notable differences between working mothers and non-working mothers. Regarding teachers' role, students in classes where there has been a gender change of the mathematics teacher from second to fifth grade are compared with other students where this change has not occurred, controlling for both teacher and student characteristics. These estimates suggest a positive effect of the teacher's gender change, with the effect size varying between female and male students.

The findings offer valuable insights for addressing the gender gap in mathematics and promoting equality in educational and occupational opportunities. The results shed light on potential mechanisms that can lead to gender disparities in mathematical performance throughout the students' school careers. Indeed, the potential of role models as a potent policy tool for reducing educational inequality has been highlighted, underscoring the need for targeted interventions to encourage greater engagement of female students in STEM fields.

Keywords: gender gap; mathematics achievement; role models

Teachers' Learning to Learn conceptions in four countries: foundational support for students' L2L acquisition to close the achievement gap

Cristina Stringher – Hugo Armando Brito Rivera – Ana Castro Zubizarreta – Salvatore Patera

“Capabilities beget capabilities” (Heckman, 2008: 34). Nobel prize James Heckman underlines the generative nature of skills acquisition through three main mechanisms: self-productivity, dynamic complementarity and cross-productivity. Self-productivity refers to the idea that a high stock of one skill in a given period generates higher levels of that skill at a later stage. Dynamic complementarity refers to the synergistic effect in capabilities investments at different ages that reinforce each other, whereby early childhood education experiences should be followed by later investment for the former to be productive. In addition, Heckman maintains there is also a cross-productivity effect, and cognitive skills affect the formation of other skills and vice versa. In his view, these mechanisms in skills formation also account for the differences among individuals and the poorer the investment is on skills since early childhood, the greater the achievement gap.

Socio-economic status, with its different levels of investment on skill formation, may thus be a determinant of educational outcomes and later adult earnings (Van der Berg, 2008), because instruction impacts capabilities. It is therefore crucial to provide adequate education to contrast inequalities and, in this framework, Learning to Learn (L2L) is key.

Recent PISA 2022 data show that on average Mathematics, Reading and Science scores have declined in most countries compared to 2018, yet some are resilient, and some have even improved their results during the COVID-19 pandemic (OECD, 2023a). Among the many lessons that can be learnt from PISA 2022, three are worth attention: a) the difficulty of students in motivating themselves to do schoolwork during school closures; b) the varying ability of students to learn autonomously and remotely; c) additional support to students from their families which generated greater confidence in their capacity for self-directed learning (OECD, 2023b). Motivation, autonomous learning, self-confidence and self-directed learning are crucial aspects of a common notion: the hyper-competence of Learning to Learn (L2L), whose functions are learners' empowerment in order to autonomously face new challenges and confer meaning to learning experiences with self-confidence (Sala et al, 2020; Deakin Crick, Stringher, Ren, 2014). Interestingly, during PISA 2022, the change compared to PISA 2018 in the percentage of students reporting that their teachers give extra help in most or every lesson is the highest in Italy among OECD participating countries and it is increasing (OECD, 2023b). This attention of Italian teachers towards their students is a very positive legacy after the pandemic and it could be enhanced, especially if they could give students the rod to fish instead of the fish, i.e., if they could equip students with the learning competences implied in L2L. However, in order to do so, teachers themselves should apply a wide L2L conception to support this competence in their students during the daily school routine.

Our study, carried out just prior to the pandemic, sheds light on 108 teachers' L2L conceptions in four countries: Brazil, Italy, Mexico, Spain. Our main research questions were: how do teachers in these countries define L2L? Do they hold a wide or narrow L2L conception (Hounsell, 1979)? Are there differences in these countries' conceptual maps on L2L? This analysis is part of a broader study framed within a sociocultural perspective and employs a qualitative methodology through thematic coding of teachers' responses to a specific question of a semi-structured interview guide (Brito et al, 2020, Torti et al, 2021) on their L2L conception. An abductive analysis (Timmermans & Tavory, 2012) has been performed on transcript segments pertaining to this question: initially, in vivo coding (bottom-up approach) aimed at tracing L2L components, and these segments have been subsequently reconciled with theoretical categories that have been previously derived from the international literature on L2L. The analyzed segments were thus coded through a theory-driven strategy (top-down approach). For this purpose, the components of learning to learn were identified and coded according to the theoretical framework of Stringher (2021b). We then determined if each definition was wide or narrow. Following Hounsell (1979), a narrow definition refers only to study strategies, techniques, and methods, while in a broad conception the learner is deeply involved with the material to be learnt mobilizing cognitive and socio-affective-relational resources to start and sustain learning, in the face of learning difficulties. The resulting comparative conceptual map helps identifying commonalities and differences in the L2L conceptions of teachers within the examined countries segmented by education level (preschool, primary and secondary education).

Overall results show that the vast majority of teachers in all countries hold a narrow L2L conception, with frequencies ranging from 72% approximately in Italy and Spain to almost 79% in Mexico, through 75% in Brazil. This result, albeit with the limitation of a qualitative methodology, is quite similar to a study carried out in Belgium (Waeytens et al., 2002)). Examples of a narrow definition include a reduced number of L2L components cited by teachers (less than three theoretical components), while broad conceptions not only include a higher number of components, but also articulate them in a coherent whole.

In all countries, L2L components cited by teachers are Learning modes, mechanisms and strategies; Learning for life and lifelong learning; Self-regulation; Competence, capacity, ability; Autonomy, agency and learning identity; Curiosity; Learning relationships; Problem-solving in learning.

Several other theoretical categories do not appear in all countries and show a situational pattern in teachers' responses: Self-confidence and resilience; Learning styles; Innate desire to learn; Creativity; Developmental dimension; Meaning making. We will also present and discuss our analyses by education level to comment on similarities and differences among countries and within a single system among teachers of different levels (preschool, primary and secondary education). National contextual factors will be considered in the interpretation of our results, such as national curricula and local values, according to Inglehart e Welzel's World Values Survey 7 (2020).

Although L2L has been present in the international educational debate for the last 50 years (Deakin Crick, Stringher & Ren, 2014) it is evident that for the interviewed teachers certain components are present across countries (e.g., learning for life and lifelong learning; self-regulation), while other categories receive less attention in some countries (e.g., self-confidence and resilience or innate desire to learn). On the one hand, such results are indicative of the cultural nature of the L2L construct (Stringher, 2021), representing a concept with potential variations, according to teachers' native culture, with respect to the different ways in which participants conceptualize education and value certain theoretical components. On the other hand, even considering the framework of cultural diversity to comment on the components indicated by participating teachers, data point to the need to provide them with a broader L2L notion, conceptualized as a hyper-competence composed of an ample set of components. Specifically, the prevalence of narrow conceptions points to the need to train teachers in the management of this competence in their classrooms, in order to acquire a broader perspective and more didactic expertise on L2L, both of which are conducive to supporting L2L in students.

In this sense, the data show the need to design and implement local teachers training processes aimed at raising teachers' awareness of all the basic components and functions of learning to learn, with special attention to the orchestration of learners' own resources for learning sustained by their autonomous agency, which sometimes may be at odds with a traditional way of teaching. One aspect of our work seems key: if teachers do not hold such broad conception, it will be harder for them to support their students in the acquisition not only of the technicalities of learning, such as memorization strategies or concept mapping, but also in their motivation to learn with intrinsic curiosity and critical thinking. Technicalities are important, but a good learning climate is central. If education is to aspire to close achievement gaps, it should equip all students with the rod to fish.

In this respect, there is ample opportunity for improvement, because schools, through their self-evaluations, are just starting to understand the potential of L2L in closing the achievement gap in basic competences, at least in Italy. Therefore, our study points to the need to conceive teacher professional development on L2L for their empowerment as a basis for their capacity to help students learning how to learn in the face of adversity and inequalities in educational opportunities. A second implication concerns the need for concrete whole school L2L interventions to reach a critical mass of enabling teachers for the benefit of all children, pupils and students, no matter their initial social circumstances. This is actually what we are already planning for the future. Last but not least, a teacher survey on their L2L conception to understand if the proportion of teachers with a wide notion is similar to what we found in our qualitative study, could help measure the need of L2L training interventions.

Keywords: Teachers' professional identity; Learning to Learn; inequalities; international study; qualitative methodology.

An efficiency analysis of Italian schools: differences between private and public schools

Giovanni Antonini – Tommaso Agasisti

Education systems play a key role in the development processes of countries, both locally and globally, fostering long-term economic growth (Hanushek et al, 2015). The literature has extensively highlighted the benefits of a quality education and training system that enhances students' capabilities and promotes the development of human capital. In this context, the analysis of the performance of educational institutions and the education system assumes fundamental importance.

In the literature on the economics of education, the analysis of school systems is often conducted within a theoretical framework based on the concept of the production function of education. This approach involves the modelling of inputs, such as school resources and teacher quality, and outputs, commonly identified as student knowledge outcomes (Hanushek, 2020). Within this framework, it is possible to identify and measure the performance of educational institutions in terms of efficiency defined as "the ability of an organisation to produce the highest possible level of output given the level of available inputs" (Johnes, 2004). For a state, having an efficient school system means responding to both educational and financial needs at the same time, which is quite important in the current context of limited public spending. Therefore, among the aims of a school is mainly the development of students' capacities, creating knowledge and making the best use of the available financial and human resources, trying to achieve more than proportional results (Johnes et al., 2017).

The countries present a very heterogeneous structure and organisation of education systems. Among the differences is the presence of mixed systems, such as Italy, where state and private parochial schools coexist. Several international studies have focused on the differences in performance between the various institutions. The literature points out that public schools, on average, perform better than state schools due to some specific factors (Agasisti and Zoido, 2019; Agasisti and Zoido, 2018; Cordero et al., 2022). Indeed, public schools operate in an environment of competition that is not present for public organisations. The competition, therefore, responds to the need on the part of public schools to attract its students and respond to the needs of the community, understood as parental interests and values. A feature of the private system is its typicality to attract a high percentage of students with a higher socio-economic background (Dronkers and Robert, 2008). Intrinsic to this competition dynamic is the relationship between efficiency and competition in the free market (Epple et al., 2016). A second characteristic of the public school is a greater incentive on an efficient use of resources based on greater autonomy and a greater focus on quality and not only quantity of investments. At the level of autonomy that schools can manage, there is one related to the knowledge that is transmitted and one related to the staff employed. If the former refers to a greater autonomy that public schools have to manage in the choice of the educational pathway, the latter refers to the possibility of hiring more freely and without adherence to national rankings for teachers. Faced with these characteristics, public schools find themselves with greater incentives that justify their greater efficiency compared to public schools (Friedman & Friedman, 1981).

Starting from this framework, this paper asks whether there are differences in efficiency between public and public schools in Italy. It analyses territorial differences and gaps in efficiency results between geographical areas, especially for public schools.

The data used in the study comes from the National Institute for the Evaluation of the Educational System of Education and Training (INVALSI), which, through standardised tests, annually assesses the student population in reading and mathematics for certain grades. The present research focuses on the surveys carried out for the school year 2018/19 for grade 8, i.e. the last year of lower secondary school. Also included in the dataset are student achievement data from three years earlier, when they were attending the last year of primary school, i.e. grade 5. In addition to the test score, a proxy for the student's socio-economic status, called SES, is used, measured through the ESCS index, which encompasses parental education and occupation, along with household assets.

As the original INVALSI dataset is constructed at the student level in the next step we aggregated all data at the school level, as empirical analyses are conducted using the school as the unit of analysis. The dataset includes the entire student population for the 2018/19 school year with a total of 346,834 students enrolled in public schools and 13,705 students enrolled in public schools (3.8% of the entire sample). In the next stage of analysis, the data were aggregated to have schools as a unit.

Inputs and outputs were chosen for the analysis based on two criteria: the existing literature and the available data. In fact, while data on human and financial resources are necessary for a complete analysis of the efficiency of the educational organisation, due to a lack of these data, we only use the average student socio-economic status index (ESCS) as input. Both reading and mathematics test scores are considered as outputs.

Context variables relating to student characteristics were then included to capture a potential association with efficiency scores. The variables taken into account are the percentage of gender present in the school, having attended pre-school, being native Italian or being of immigrant origin (a distinction was made between first-generation and second-generation students), proportion of repeaters and finally the reading and mathematics grade taken three years earlier.

The methodology applied for the analysis is Data Envelopment Analysis (DEA). This represents the most widely used technique in the literature to calculate the Decision-Making Units (DMUs, in this case schools) efficiency frontier, based on the combination of input and output.

In line with the literature, the results show a high heterogeneity in efficiency scores among schools, both by type and geographical composition, with many schools having room to improve their performance.

Keywords: schools' efficiency; private schools; Data envelopment analysis (DEA)

SESSION 8. RELYING ON ADMINISTRATIVE DATA TO ANALYZE STUDENT CAREERS

ORGANIZER: UNIVERSITY OF PIEMONTE ORIENTALE

COORDINATOR: CARMEN AINA

18TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 4 – RESEARCH 6}

Net gains: an investigation of telematic universities' academic performance

Carmen Aina – Luca Bonacini – Chiara Mussida – Giuseppe Pignataro

In recent years, the rise of telematic universities has emerged as a noteworthy phenomenon in Italy. These institutions offer a wide range of educational programs, leading to a shift in the landscape of tertiary education in Italy and altering the dynamics of competition with traditional public universities. To date, there are 11 telematic universities offering online degree courses. By now, more than 10% of university students are enrolled in a distance learning institution. According to USTAT data, in 2008/09 there were 20,874 enrolled in degree courses in these universities out of approximately 1.8 million; in 2011/12 they were 40,164 out of 1.75 million, in 2015/16 they were 62,276 out of 1.65 million. In the last academic year 2022/23, there were 251,017, accounting for 13.15% of those enrolled in a degree course, while the previous year there were 223,937 and in 2020/21 (the year of the pandemic) 184,901 (ANVUR 2023). A large part is mature students. Among the enrollees in distance universities, there is a significant presence of workers over 31 years old. According to the ANVUR report 2023, indeed, in the academic year 2021/22, 80% of students in traditional universities were under 26 years old, but in distance learning institutions they were only 34%, while about 57% (almost 2/3) were at least 28 years old and over 45% (almost half) were over 31. In traditional universities, 80% of graduates were 23 years old (compared to 66.5% in the academic year 2011/12), while in distance learning institutions they were only 20.6% and almost 60% of graduates were at least 28 years old. An important part, therefore, is already integrated into the job market and enrolls in a distance learning institution to complete their university career: over 70% of them come from state universities (ANVUR 2023). These universities are regulated by the decree of April 17, 2003, issued by the Ministry of Education, University and Research, which outlines all the requirements that Italian universities must meet to be recognized. This ensures that online degree qualifications are equivalent to those obtained from traditional universities (i.e. with the same legal value). Telematic degree courses are diverse and include popular disciplines such as Engineering, Computer Science, Economics,

Law, and Psychology. Many questions remain unanswered regarding the significance and consequences of these alternative educational routes. In this paper, we aim to leverage a 2016 reform that introduced the recruitment of permanent faculty members within telematic universities. Prior to 2016, these universities primarily relied on short-term teaching contracts, frequently hiring individuals from other universities or the workforce. However, since 2016, telematic universities have been required to employ staff at various levels such as researcher, associate professor, and full professor, albeit incurring higher costs. Distance learning institutions are indeed characterized by a reduced permanent staff, and beyond the minimum requirements, distance universities show an abnormal ratio between students and permanent faculty. The problem lies not only in the minimum number of tenured professors required to activate a study program, according to the criteria and accreditation platforms set by ANVUR and the Ministry. The issue is the ratio between the total number of permanent faculty and the number of students they must serve. This ratio is considered central in evaluating any university and, in general, any university system. This ratio not only pertains to teaching (i.e., the direct relationship between teacher and student in education) but also to the ability to participate and interact during educational activities that extend beyond the classroom in academic paths (such as laboratory work, seminars, or support in constructing and preparing the final project). In Italy, the overall university system has a student-to-faculty ratio of about 1/30. Italian distance learning universities, however, have a student-to-permanent-faculty ratio of a different order of magnitude: according to the ANVUR Report (2023), it increased from 152.2 in 2012 to 384.8 in 2022 (almost thirteen times higher than traditional universities). However, the 2016 reform, by increasing the proportion of structured faculty, not only raised personnel costs but also led to greater continuity in teaching. Utilizing the administrative data from the National Archive of Students and Graduates (i.e. Archivio Nazionale degli Studenti e dei Laureati, ANS) provided by the Ministry of University and Research (MUR). We investigate the administrative data regarding the academic careers of students between 2011 and 2019, so that we can test this change in the faculty on academic outcomes without being influenced by the effects of Covid-19. The sample analysed includes about 2,500,000 students at their first enrolment in the Italian HE system. This natural experiment enables us to investigate how student performance evolved in both the short and long run. We hypothesize that educational continuity might have contributed to influencing several academic outcomes, namely dropout, number of credits achieved, average grades, graduation. As private institutions, telematic universities could provide a variety of services to their students. A preliminary analysis indicates that the continuity of teaching has resulted in a decline in student grades in recent years (excluding factors related to the pandemic). The reasons for this decline can be varied and are currently under investigation. Among the pre-university enrollment control variables, we are also verifying whether we can include, at least for some cohorts, the INVALSI data recorded during the school path.

Keywords: academic outcomes, students' performance, telematic universities, administrative data

General Schools and University Performance

Cristina Specchi – Luca Bonacini – Giuseppe Pignataro

The Italian education system, like many other countries in Europe and the United States, consists of two paths that greatly affect the post-secondary options of students. On the one hand, general high schools (licei) are designed to facilitate entry into universities and encourage further studies. On the other hand, vocational high schools prepare students for direct entry into the workforce (Hanushek and Woessmann, 2006). We aim to investigate the effect of school choice, specifically on the academic performance of students who enroll in universities, and the role of attending a general school program in this regard. There is a consensus that the dual-track system has led to unequal opportunities and affected school readiness, university choices, and employment prospects (Brunello and Checchi, 2007). Vocational education and training (VET) can equip individuals with skills that may facilitate access to the job market, but these skills may quickly become outdated. The choice of enrolling in higher education is a determinant in the future choice of students to enroll in university, and the high school diploma in a general school favors this kind of choice (Camanho et al., 2021). Measuring the impact of upper secondary education on further education is crucial for policy evaluation, considering the significant role of education in promoting economic growth (Hanushek and Woessmann, 2020). Our research aims to investigate how the choice of a general school impacts university performance after the decision has been made (intensive margin) rather than the

probability of enrolling in university (extensive margin). As a source of information, we use data from a group of students who are enrolled at the University of Bologna. Through these data, we can capture a qualitative dimension connected to the grade obtained in an exam or the rejection of it, and a quantitative dimension related to the number of credits taken. We have access to a comprehensive dataset that covers administrative records for over 80,000 students over five years, from the academic year 2017/2018 until 2021/2022. By examining the variation in the share of general schools at the municipal level, we can account for any unobserved selection. To address the challenge of identifying the impact of school choice on students, we utilized instrumental variable estimation. This helped us to account for other factors related to the communities in which students grew up. We were able to find a useful variable by gathering precise information about the name of the high school, its location in Italy, and whether it was a general or vocational school. Specifically, the instrumental variable we used was the share of general schools in the municipality where the student attended high school. We derived this variable from the Ministry of Education, University, and Research's list of secondary schools in 2012. Our rationale was that the more prevalent general schools were in comparison to the total number of upper secondary schools in an area, the more likely it was that students would enroll in them. To account for the varying factors that can lead to biased results, such as class size, subjects studied, and cultural background of students, we utilize a two-step procedure, based on the work of Bosquet and Overman (2019). In this procedure, we first remove the average effect of each program from the student outcomes, which helps us focus on the pure effect of the exogenous IV. This enables us to measure the difference between general and vocational schools in university performance, net of the effect of the program. Our study has revealed that general education has a positive impact on various outcomes in universities. General education has a positive effect on all the main outcomes of interest, starting with the grades obtained and the number of exams (credits). We have observed, through unconditional quantile regression, that the effect of general education is consistent across all students, without any sorting effect related to their previous background, for both credits and grades. Additionally, the analysis of heterogeneity has identified policy issues related to school distribution, living in deprived areas, scholarship reception, and first-year enrollment. In summary, our research sheds light on the discrepancy in university performance between students who pursued general education and those who pursued vocational education throughout their academic journey. This disparity is measured by both the number of exams taken as well as the quality of grades obtained. One possible solution to address this issue is to consider introducing compulsory introductory courses during the summer of the first year of enrollment. This could be implemented regardless of the program type, to bridge the preparation gap for students who have a background in vocational education.

Keywords: General School, Municipality, Tertiary education, Degree Programs

The effect of anticipating performance requirements in need-based grants: a natural experiment

Giuseppe Pignataro – Luca Bonacini – Cristina Specchi

Given the broad social and private benefits of tertiary education, ensuring universal access for all interested students is central to achieving social equity and economic efficiency. The effective completion of tertiary education is becoming increasingly important due to the growing demand for highly skilled individuals in the labor market in OECD countries, La vecchia et al. (2020). Indeed, Europe has decided that the proportion of people aged 25-34 with a tertiary education should be at least 45% by 2030. A goal that many countries will find extremely difficult to achieve.

Within this framework, the need to ensure that tertiary education is truly more accessible to disadvantaged students has led most countries around the world to intervene in recent years with massive financial programs aimed at facilitating their access. Many countries, particularly in Europe, have chosen to provide scholarships as the main solution to facilitate access to university. Almost all of these scholarships are offered in different countries according to a combination that responds to the economic needs of the most disadvantaged combined with meeting certain academic requirements related to the number of credits in order not to lose the right to renew the scholarship (Scott and Clayton 2011). As principal-agent theory would suggest, the introduction of performance requirements for scholarship confirmation is necessary to reduce moral hazard behavior on the part of students, with the goal of holding them accountable and

monitoring their commitment to further study. While this has a positive effect on incentives by ensuring that students perform better at university, it may more easily induce a proportion of students to drop out because the requirements are too high for them to meet. Combining this mechanism - related to the traditional trade-off between efficiency in the allocation of scholarships and equity in their distribution - is not easy, although very interesting studies have tried to estimate the causal impact of scholarships and their effects on educational inequality, exploiting some reforms in different countries (Dynarski 2003).

We use comprehensive administrative data on the universe of students who applied to the scholarship program at a large university between 2017/2018 and 2021/2022 to observe a variety of academic outcomes. First, we use a Difference in Differences approach to assess whether scholarship students in cohorts enrolled after the reform experienced changes in academic performance due to the proposed alternative timing. In particular, the reform was approved in June 2019 to take effect in the following academic year 2019/2020. In our case, we cannot ignore the impact of COVID-19, which could have affected students with scholarships differently from students without scholarships in some way as of March 2020. In this context, we propose a set of methods for both general and local analysis of a Triple Difference, including the Regression Discontinuity Design in Time -- a sharp regression discontinuity with DID analysis -- which allows us to control for the effect of the pandemic, confirming that both the treated group (students with scholarships) and the control group (students without scholarships) were affected by the event in the same way. This allows us to control for the effect of the reform with certainty, without observing any bias related to the pandemic shock caused by the transition to the online world.

Furthermore, our additional contribution is to apply the changes-in-changes model proposed by Athey and Imbens (2006) -- a nonlinear generalization of the DID model that focuses on the full distribution of outcomes. The idea would be to observe the impact of the reform on each quantile of the distribution of the student outcome by looking at the differences between students with scholarships and students without scholarships. This is not possible in the standard context because the adjustments are made to the mean, not to the entire distribution, and are implemented linearly.

The results of the main analysis suggest a radical change in student behavior between semesters. In particular, the DID analysis shows an increase in credits in the first semester, which is associated with a significant increase in the probability of scholarship confirmation for both semesters.

In contrast, there is no effect on the average grade, while failed exams increase in the second semester. In our analyses, we consider both an immediate effect in the difference between 2019/2020 (reform year) and the previous year but validate each result with a series of robustness checks.

Moreover, we add that no COVID-related effects are found in both local and global analysis, so we can confirm that the effect of the pandemic did not lead to different changes between the group of scholarship and non-scholarship students. Finally, we found that scholarship students with the greatest need to earn credits increased the number of exams they passed in the first semester, while others with less need to earn credits decreased the number of exams they passed in the second semester. This effect is robust to different extensions and grows over time.

Trying to dig deeper into this result, we even observe a shift in the type of exams taken, with a clear incentive to take smaller exams in the first semester and leave larger exams for the second semester. In fact, in Italy, exams are measured by the number of credits assigned to them. There are exams from 2 to 12 credits. Each credit corresponds to 6-10 hours of class time, and this necessarily leads to a variation in the number of topics that can be covered on average, which is greater for exams with higher credits. The overall result suggests that, on the one hand, the reform has created a strong incentive for students to earn part of their scholarship in advance in the first semester. As a result, students choose shorter exams in the first semester - perhaps guided by the perception that they are more manageable to study or easier to prepare for, given the smaller number of topics covered on average. In the second semester, however, we observe that there are more failures. In fact, students are left with larger exams - those with a higher number of credits required - and this may lead to lower average grades for some of them.

In conclusion, the reform seems to have significantly improved students' careers by increasing the number of credits and the likelihood of scholarship renewal by shifting to relatively smaller exams, especially in the first semesters. However, with the primary goal of maintaining the scholarship over time, the reform induced students to register for more exams in the second semester -- even though they were not sufficiently prepared to take the exam -- with the side effect of reducing the number of exams passed and the average grades for a portion of the student distribution.

Keywords: Math Exposure, Twins, High School Reform, University, Major Choice, STEM

Math Exposure and University Performance: Causal Evidence from Twins

Luca Bonacini – Graziella Bertocchi – Majlinda Joxhe – Giuseppe Pignataro

Taking mathematics courses in high school has far-reaching effects on various educational outcomes, including high school graduation rates, college enrollment (Aughinbaugh, 2012; Bertocchi et al., 2023) and performance (Long et al., 2009), and future earnings (Altonji, 1995), while even influencing the likelihood of pursuing a career in STEM (Science, Technology, Engineering, and Mathematics) (De Philippis, 2023). Our work contributes to this research line by examining the effects of exposure to mathematics hours during high school on college major choice and performance. In order to identify the causal effect of high school math on our outcomes of interest, we exploit two complementary empirical strategies: The first is based on a twin fixed effects estimator and the second on a difference-in-differences estimator.

First, we leverage a kind of natural experiment represented by the occurrence of twins within the population of students enrolled at Italy's second largest university, the Alma Mater Studiorum University of Bologna, in the period 2011-2021. Thanks to a unique identifier, we extract a sample of 1,396 twins, about whom we have information on the high school they attended. One advantage of our sample of twins, which we obtain from administrative data rather than from twin registries, derives from the fact that most of the latter depend on the voluntary participation of twins, with an implied recruitment bias which we instead avoid. To capture the math content of high school curricula, we divide high school curricula in two types, respectively characterized by high and low math content. By including in the regression within-twin fixed effects, i.e., fixed effects at level of each pair of twins, we can then exploit differences in high school choices within each pair, in such a way to effectively neutralize the influence of latent factors, such as genetics or family background, that potentially result into a bias of the standard OLS estimator (Ashenfelter and Krueger, 1994; Li et al., 2012; Ye et al., 2023). Our results indicate that attending a low math school reduces the probability of enrolling in a STEM major and increases grade performances during college. Moreover, the fixed effects estimates reveal a downward bias in the standard OLS estimates, suggesting an even more powerful role for high school math than what has been pointed to in the literature.

Next, we exploit yet another kind of a natural experiment that occurred in Italy through a national reform that increased the number of hours in mathematics only for high schools with lower mathematics content. In order to evaluate the impact of this reform, we employ a difference-in-differences framework using once again data on students from Bologna. Our results indicate that, for treated students, an increase in math hours reduces the probability of enrolling in a STEM major. Moreover, treated students tend to achieve higher grades during their college careers. Interestingly, our heterogeneity analysis shows that these results are primarily driven by high-ability students who exit high school with higher final grades. High-achieving students tend to shy away from STEM majors in response to increased mathematics exposure, potentially due to perceived difficulties, lack of interest, and/or to an aversion to math (math anxiety) that they likely developed earlier on and that being exposed to more math at a later stage actually intensifies (Pellicioni et al. 2016). In addition, we point out that the improvement in the academic performance is driven by high-achieving students.

In our analysis of heterogeneity, we also observe that the reduced likelihood of enrolling in a STEM course due to the treatment is homogeneous by gender, but at the same time, the positive effect of the treatment on academic performance is higher for female students. This latter result is also interesting and paves the way for further future investigations in light of the fact that, as is well known, women tend to have poorer performance in mathematics and avoid STEM courses.

These results underscore the importance of considering differential effects based on students' academic aptitudes and aspirations when implementing educational reforms. Moreover, they highlight the complex interplay between curricula adjustments, university choice, and academic outcomes. Understanding these dynamics is crucial for designing effective educational policies aimed at fostering both equitable access to higher education and academic excellence.

Keywords: CBT, digital skills, learning Italian, learning mathematics

SESSION 5. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAs)

METHODS AND RESULTS

ORGANIZER: UNIVERSITY OF KU LEUVEN - UNIVERSITY OF BATH

COORDINATOR: ELLEN CLAES

18TH OCTOBER: 2.00 P.M. – 4.00 P.M. {ROOM 2 – RESEARCH 7}

Exploring international learning loss patterns: a ML investigation using PISA data before and after COVID -19

Melisa Lucia Diaz Lema – Melvin Vooren – Ilja Cornelisz – Chris Van klaveren

Background

The global education landscape has been significantly disrupted by the Covid-19 pandemic, prompting schools worldwide to shutter their doors and adopt remote learning measures. In the participating nations of the Programme for International Student Assessment (PISA), a substantial number of students experienced school closures due to Covid-19, some lasting three months or more (OECD 2023). During and after this period of educational turmoil, there is a growing sense of urgency among policymakers, educators, and researchers to fully comprehend the extent of learning setbacks experienced by students.

While the disruptions caused by Covid-19 are widely acknowledged, efforts to quantify their impact have predominantly focused on individual countries (Bertoletti et al. 2023a; Maldonado and De Witte 2022). This narrow focus underscores the necessity for broader international assessments. This study seeks to bridge this gap by utilizing PISA data to assess the scale of learning loss at an international level.

Data

To comprehensively model students' performances while considering both longitudinal trends and cross-national variations, we leveraged data from three consecutive editions of PISA, conducted in 2015, 2018, and 2022.

PISA assesses the skills and knowledge of 15-year-old students in reading, mathematics, and science, providing a framework for international comparisons. To ensure a consistent analysis, we focused on the subset of 61 countries that participated in all three PISA editions, comprising 85% of participants in 2015, 77% in 2018, and 75% in 2022. This deliberate selection yielded a substantial sample size, with over 400,000 students included in each edition.

Beyond academic performance, PISA also captures a wide array of contextual information, encompassing student, parent, and school characteristics. In our study, we meticulously filtered these variables to include

only those consistently present across all three editions. These covariates spanned students' demographics and backgrounds as well as indicators of school and home support.

Methods

The study utilizes various machine learning techniques to develop predictive models for students' performance in reading, mathematics, and sciences across specific cohorts. Subsequently, it utilizes these models to generate out-of-bag predictions for other cohorts. This approach enables the assessment of the impact of the Covid-19 pandemic, treated as an exogenous shock, on the Educational Production Function (EPF) (Hanushek 2020a). ML model predictions are expected to reflect the underlying changes to the EPF across different cohorts over time, namely, pre- and post-pandemic periods. To evaluate this, the study compares the out-of-bag predictive performance of models before and after the pandemic. For example, it compares the predictive power of model A, trained on PISA-2015 data and tested on PISA-2018 data, with that of model B, trained on PISA-2018 data and tested on PISA-2022 data. Additionally, as a robustness check, it examines the out-of-bag performance of model A, trained on PISA-2015 data, using PISA-2022 data as a training set.

Utilizing a variety of statistical and ML techniques enables the study to address different model assumptions. Generalized Linear Models (GLMs) rely on parametric forms to establish associations between covariates and the dependent variable. Alternatively, Support Vector Machines (SVM), Decision Trees, and Random Forests (RF) can capture complex relationships and potential interactions among covariates without imposing strict parametric assumptions.

Expected results

The out-of-bag predictive performance, when using the PISA-2022 data to test the model, is anticipated to be lower when compared to the out-of-bag predictive performance when using the PISA-2018 data set to test the model. This stems from the understanding that the pandemic likely altered the EPF. Consequently, a model trained on pre-Covid data is predicted to exhibit poorer predictive performance when applied to post-Covid-19 data than when applied to pre-Covid-19 data.

Furthermore, it is projected that this disparity in predictive performance may vary when the data is categorized according to different background characteristics. For instance, variations among different countries are anticipated, attributable to disparities in the severity of the Covid-19 pandemic experienced within each nation. This can also be extended to different geographical areas, such as rural versus urban. It is likely that the impact of Covid-19 was more severe in urban areas compared to rural areas.

There may also be differences across different groups of students. Socio-economic status is expected to affect the influence of Covid-19 on learning loss through different channels. One mechanism is parental education. When spending more time at home during a lockdown, it is expected that this has an influence on parent-child interactions. With higher levels of parental education, there are expected to be more positive spillovers from parent to child compared with lower levels of parental education. The same holds for family income. Lower levels of family income are expected to be associated with more difficulties with working and studying in a lockdown situation.

The heterogeneous effects of Covid-19 on learning outcomes and the diverse channels through which it influences learning have significant and broad policy implications. For instance, if Covid-19 disproportionately affects students from families with low incomes or lower levels of parental education, tailored policy instruments can be developed and implemented to mitigate these adverse effects. This study, by examining international learning loss patterns before and after Covid-19, contributes to a deeper understanding of the pandemic's global impact on education. The insights garnered will inform evidence-based policy responses aimed at addressing learning setbacks effectively.

Keywords: Learning loss, Covid-19, Machine Learning, PISA data

How does ICT affect school efficiency? Evidence from OECD PISA 2022

Mara Soncin – Tommaso Agasisti

Introduction

In recent years, there has been a pronounced interest surrounding the evaluation of educational institutions' efficiency. This interest stems from the increasing awareness by policymakers and educators about how crucial is to ensure effective resource allocation to enhance overall educational outcomes (De

Witte & López-Torres, 2017). Furthermore, with the pervasive integration of digital technologies into the educational setting especially after the COVID-19 period, investigating the role of these tools in shaping school efficiency has become a matter of crucial interest, yet it is largely underexplored (Mergoni et al., 2023). Taking an international perspective can be particularly informative, as recent research in Latin America highlighted that the degree of access to digital tools influences school efficiency more than its use (Agasisti et al., 2023), while evidence from European countries indicate that ICT use (in terms of instructional time using technologies) matters more than the availability of digital resources in shaping students' results (Mergoni et al., 2023).

The current research further explores the relationship between technology utilization and school efficiency for 23 European countries by addressing the following research questions:

- What is the level of school efficiency in European countries, and how much variation exists across and within countries?
- How can the level of school efficiency be explained by access to and the use of ICT resources?

By examining these questions, the study aims to contribute to a better understanding of the relationship between school efficiency and technology utilization, extending the academic literature and providing insights that may inform educational policies and practices across countries.

Data and Methodology

To achieve the objectives outlined above, this study leverages on data from the OECD Programme for International Student Assessment (PISA) 2022. PISA provides a comprehensive dataset that allows for the assessment of educational outcomes in both cognitive and non-cognitive terms, as well as data on school and student background. In addition, an ad hoc student questionnaire tracks the use of digital devices at school and at home, with available data for 23 EU countries. The study adopts a two-stage procedure. To address the first research question, a robust Data Envelopment Analysis (DEA) model is employed, with the use of country-specific frontiers to account for structural differences in the educational production function of each country (Charnes et al. 1981). The robust specification allows to account for outliers, while meta-frontiers approach enables the proper assessment of the within-country schools' efficiency score. In this setting, inputs used refer to the level of human and infrastructural resources available to schools as well as to the average school socioeconomic background, while outputs refer to student outcomes as measured by cognitive results (in reading, mathematics and science) as well as social and emotional outcomes (as measured by indexes computed by the OECD). In a second stage and to address the second research question, machine learning algorithms are employed to explore the association between the school efficiency and the availability and use of digital devices, also controlling for a set of student, school and context level characteristics.

Expected Results

By delving into these research questions, this study generates several key findings. Firstly, it is expected that the research will uncover varying levels of school efficiency across the 23 involved countries, highlighting differences both between and within nations. The average level of (in)efficiency provides informative evidence on the schools' ability to use the current inputs to maximise students' outputs. Secondly, the study expects to reveal correlations between school efficiency and the accessibility and utilization of ICT resources. Given the under-exploration of this topic, available evidence would inform both the academic and practitioner community transnationally. Understanding these associations can provide valuable insights into the mechanisms through which technology influences efficiency and inform strategies for optimizing the use of digital tools in educational settings. Moreover, the use of machine learning algorithms represents a novelty in this kind of explorative studies. By providing insights into the mechanisms through which digital tools contribute to educational efficiency, this research seeks to inform educational policies and practices across countries, ultimately enhancing the quality of education provided to students within the European countries and beyond.

Keywords: Efficiency analysis; ICT; digital; OECD PISA

School Leadership Support and Socioeconomic Status (SES) Inequalities in Mathematics and Science Achievement: Evidence from TIMSS 2019

Nurullah Eryilmaz

Despite global efforts to promote equitable education, the socioeconomic achievement gap persists, with students from lower socioeconomic backgrounds facing significant educational disparities (Chmielewski, 2019). Research has consistently demonstrated that students from low socioeconomic status (SES) households experience lower academic achievement, higher dropout rates, and reduced likelihood of college graduation compared to their high-SES counterparts (Ballard Brief, 2022). The "opportunity gap," as it is sometimes referred to, has been correlated with a lack of resources, leading to adverse long-term impacts on individuals and the economy (Carnoy, & Rothstein, 2013). For instance, a study by Reardon and Portilla (2016) revealed a widening achievement gap between students from high and low SES backgrounds in the United States over the past five decades. Furthermore, the Programme for International Student Assessment (PISA) has highlighted the strong association between students' socioeconomic status and their performance in reading, mathematics, and science across various countries (OECD, 2018, Strello, 2023). This persistent achievement gap underscores the need for continued efforts to address the educational disparities associated with socioeconomic status, despite the existing international agreements and national laws emphasizing the right to quality, inclusive education for all children (Bower, 2013). Additional research by Duncan and Murnane (2011) provides further insights into the economic and social impacts of the achievement gap, emphasizing the importance of closing this gap for the well-being of individuals and the society at large.

Efforts to reduce the socioeconomic achievement gap have been a focus within the education sector, with significant attention given to the roles of teachers, schools, and principals (Cabral-Gouveia et al., 2023). Research has consistently demonstrated that this gap has been correlated with a lack of resources, leading to adverse long-term impacts on individuals and the economy. While various reforms and interventions, such as high-quality preschool, effective teachers, and a challenging curriculum, have shown demonstrable effects on reducing the achievement gap (Olson et al., 2017; Mayer, 2008), relatively less study has been conducted on how school leadership support moderates this gap (Smith & Gumus, 2022; Gumus et al., 2022). This highlights an opportunity for further investigation into the specific ways in which school leadership support can contribute to narrowing the socioeconomic achievement gap. Research has shown that stronger instructional leadership can diminish the relationship between school-level socioeconomic status (SES) and average student achievement, thereby reducing the achievement gap between students in high- and low-SES schools (Gumus et al., 2022). This underscores the potential for school leadership to play a significant role in mitigating the impact of socioeconomic status on student achievement. The existing knowledge gap concerning the specific moderating effects of school leadership on within-school inequality in student achievement calls for further research to better understand the mechanisms through which school leadership can contribute to reducing socioeconomic achievement gaps.

In order to address the existing knowledge gap and provide more evidence on the mechanisms influencing the socioeconomic achievement gap, this study investigated the moderating effects of school leadership using data from the Trends in International Mathematics and Science Study (TIMSS) 2019 across 46 countries for both student mathematics and science achievement scores. The TIMSS 2019 data provides a robust and comprehensive foundation for examining the relationship between school leadership and student achievement across diverse socioeconomic and cultural contexts. By focusing on the specific ways in which school leadership support can contribute to narrowing the socioeconomic achievement gap, this study aims to offer valuable insights that can inform the development of more targeted and effective interventions to reduce educational inequalities and promote equitable access to high-quality education for all students.

Design/methodology/approach: Using data from the Trends in International Mathematics and Science Study (TIMSS) 2019 across 46 countries for both student mathematics and science achievement scores, this research estimated regression models to investigate whether the effect of school leadership support on student achievement in math and science students with the different individual levels based on socioeconomic status.

Analytical Strategy

To answer the research questions, we first conduct a regression analysis using pooled data and then estimated separate regression models for each educational system. In these models, we examine the relationship between school leadership support and student discipline variables, teacher characteristics and student characteristics and interaction between socioeconomic status and leadership support. The regression model estimated for each education system is outlined below.

The dependent variable Maths/Science (Mathematics/Science Achievement) was regressed on the eleven teacher measures and three student characteristics:

$$\text{Math/Science} = \beta_0 + \beta_1 * \text{lead_support} + \beta_2 * \text{Stu_Dis_var} + \beta_3 * \text{ses} + \beta_4 * \text{stu_Gender} + \beta_5 * \text{Teach_Years} + \beta_6 * \text{Maths/Science} + \beta_7 * \text{Maths/Science_Ed} + \beta_8 * \text{PD} + \beta_9 * \text{teacher_Gender} + \beta_{10} * \text{native} + \beta_{11} * \text{leadxs} + \varepsilon$$

In this study, our primary focus lies in examining the impact of the relationship between school leadership support and student discipline variables, teacher characteristics and student characteristics and interaction between socioeconomic status and leadership support. As this research objectives revolve exclusively around teacher-level variables and link student level variables, I made a deliberate choice to employ teacher weights in our analysis. The use of teacher weights is typically relevant when considering student-level outcomes or characteristics, which fall inside the scope of this investigation. This decision aligns with the specific goals and parameters of this research, ensuring that our analysis accurately reflects the relationships between teacher-related factors and student mathematics and science achievement. All estimations were performed using IDB Analyzer (IEA, 2022).

Findings: School leadership support exhibits a significant moderating effect on the relationship between student-level socioeconomic status (SES) and achievement in both mathematics and science across the aggregated data. However, the moderation effect for individual SES and school leadership support does not attain statistical significance when examined on a country-specific basis.

In conclusion, this study has provided valuable insights into the complex relationship between student academic achievement, socioeconomic status (SES), and school leadership support. The findings revealed a significant positive impact of student socioeconomic status on mathematics and science achievement, highlighting the influence of SES on academic performance. Additionally, the study identified a negative relationship between leadership support and students' science and mathematics achievement, particularly for mathematics, suggesting the need for further investigation into the specific aspects of leadership support and their potential impact on student achievement. The results also demonstrated a significant moderating effect of school leadership support on the relationship between SES and academic achievement for pooled data, indicating that strong school leadership support is associated with a reduction in the achievement gap related to socioeconomic status. While these findings emphasize the crucial role of school leadership in promoting more equitable academic outcomes for students, irrespective of their socioeconomic backgrounds, it is noteworthy that this study did not uncover consistent evidence across all countries. Consequently, caution is warranted in generalizing the impact of school leadership support in this regard.

The study's implications for educational policy and practice emphasize the potential of proactive leadership interventions to contribute to a reduction in the achievement gap linked to socioeconomic status. Future research should further explore the specific mechanisms through which school leadership support influences student achievement, as well as the potential impact of different dimensions of leadership support on academic outcomes.

Keywords: school leadership support, instructional leadership, socioeconomic status gap, mathematics achievement, science achievement, TIMSS 2019

Inquiry vs General Cognitive Activation Strategies: Shaping Teacher Profiles and Student Science Learning in Italy

Zhijun Chen – Andres Sandoval Hernandez

Introduction. Learning science goes beyond the acquisition of scientific facts, concepts and principles. It establishes and improves students' interests in exploring the world. As students learn new science content, they start to build up their ability to generate scientific enquiries, design and conduct experiments towards those questions, and explain the results based on their learning (Teig, 2022). To stimulate science learning, teachers must encourage students to be more engaged in high-order thinking by providing learning opportunities with cognitively activating activities (Teig et al., 2019). Cognitive-Activation Strategies (CAS) is a set of approaches that contributes to activating students' cognitive development and processing using complex tasks and challenging questions (Praetorius et al., 2018) Among different domains of classroom instructions, CAS is the most dependent aspect on different subjects. For instance, focusing on science instruction, CAS refers to engage students learn through investigation and experimentation using inquiry

activities. Combined with (Pedaste et al., 2015)'s framework (shown in Figure 1), science-specific CAS could also be seen as inquiry-based CAS in which students investigate the scientific knowledge from designing to interpreting experiment. Except for the inquiry-based CAS, some other key aspects of CAS that are situated in all subjects are referred to as general CAS, such as relating the course to student's life experiences and encouraging students to think of solutions independently (Schlesinger & Jentsch, 2016). Studies have examined the joint and independent contributions of general and inquiry-based CAS on student achievement yet have encountered challenges in establishing definitive findings. One potential explanation for this variability is the heterogeneous responses of educators to these CAS practices. Not all teachers adopt all elements of general and inquiry-based CAS, with some demonstrating a preference for inquiry-based teaching while others adhere more closely to conventional teaching methods. This study seeks to address this gap in the literature by identifying distinct teaching profiles based on general and inquiry-based CAS items and examining their association with science outcomes. Research object and hypothesis. The research aims to identify distinct teaching profiles based on items pertaining to both general and inquiry-based CAS within science teaching classes. Subsequently, we will investigate whether any of these identified teaching profiles exhibit consistency with science outcomes, utilizing data from the Trends in International Mathematics and Science Study (TIMSS) 2019. Due to the multifaceted design of TIMSS and its focus on science achievement, Italy was chosen as our study's focus. This is because Italy has a distinctive feature among participating European countries: a one-teacher-per-classroom structure (see further explanations in the method section). The research questions are:

1. What are the characteristics of different teacher types identified through LCA?
2. are those identified teacher typologies more consistent with student science outcomes?

Data. The current study made use of data sourced from the Trends in International Mathematics and Science Study (TIMSS) conducted in 2019. TIMSS is a globally conducted assessment held every four years, designed to gauge the academic achievements of students in mathematics and science at the fourth and eighth-grade levels. It covers a wide array of countries and involves gathering contextual information through questionnaires distributed to students, teachers, and school administrators. Specifically, this analysis focused on Italy's participation in the eighth-grade assessment.

Measures. Science achievement. Our investigation employed eighth-grade science achievement as the dependent variable. The assessment encompassed four content domains: Earth science, chemistry, physics, and biology. To account for measurement uncertainty, five plausible values were generated for each student's science achievement. All five values were incorporated into our statistical models.

Cognitive-Activation Strategies (CAS)

In TIMSS 2019, teachers were presented with two sets of questions regarding their classroom strategies in both general classrooms and specifically in science classrooms. Utilizing a four-point Likert scale (0 = never, 3 = every/almost every lesson), teachers rated the frequency of occurrence for those teaching and learning activities employed within their science lessons. Drawing from existing literature, we identified six items that reflect general CAS, such as linking new content to students' prior knowledge. Additionally, we selected five items representing inquiry-based CAS, such as presenting data from experiments or investigations.

Method. We will utilize the *Autonation* package (version 1.1.1) within R Studio (version 2023.12.1) to conduct all analyses. To address our first research question, which aims to identify distinct latent profiles of teachers based on their use of cognitive activation strategies (CAS), we will employ a special type of latent class analysis called the Factor Mixture Variant of Continuation Ratio Model (M-CRM) (Carrasco et al., 2023). We will follow a procedure similar to Carrasco et al., 2023 to determine the optimal number of latent classes. First, we create a set of $k-1$ pseudo items. These represent all possible response combinations for the original 11 CAS items. We achieve this by converting the original responses into dummy-coded variables based on the continuation ratio framework. These newly generated variables serve as the dependent variables in our M-CRM model. Next, we will fit a series of M-CRM models with increasing numbers of classes (from 1 to 5). The Vuong-Lo-Mendell-Rubin Likelihood Ratio test (VLMR-LRT) will be used to compare models with k versus $k+1$ classes (Nylund-Gibson & Choi, 2018). We will select the best-fitting model based on a combination of the VLMR-LRT results and overall model fit indices like AIC, BIC, and aBIC. To answer the second research question, we will incorporate the distal outcomes (e.g., student science achievement) into the final class model. Notably, the unique feature of the TIMSS Italy 2019 data, where each classroom has only one teacher, allows us to equate teacher profiles with classroom profiles. Therefore, we will utilize

class-average science achievement scores for our comparisons. This will allow us to examine whether students from different latent classes (teacher/classroom profiles) exhibit statistically significant differences in their science outcomes.

Results and Findings

The result showed the M-CRM with three classes performed significantly better than the one with four classes (VLMR-LRT $p = .0323$; 3-class indices: AIC = 3868.334, BIC = 3985.148, aBIC = 3874.251). This suggests we identified three distinct profiles based on the application of general and inquiry-based CAS. Profile 1 represents teachers who utilize inquiry-based activities more frequently than general activities. Profile 2 identifies teachers with a moderate level of both activity types. Finally, Profile 3 describes teachers who primarily focus on general CAS activities with less emphasis on inquiry-based approaches. Our analysis of the second research question revealed significant differences in science achievement scores across the teaching profiles. Notably, students taught by teachers in Profile 1 demonstrated the highest mean science achievement score. Building on our findings, this study suggests a valuable implication for Italian education. Recognizing the diverse teaching profiles in science classrooms equips policymakers and educators to tailor interventions and professional development, ultimately improving teaching practices and student learning.

Keywords: Cognitive Activation Strategies (CAS), Latent Class Analysis (LCA), Science Education

SESSION 3. WEAKNESS AND EXCELLENCE IN EDUCATION: FOR A POLYSEMOUS AND MULTIDIMENSIONAL IDEA OF EQUITY OF EDUCATIONAL OPPORTUNITIES

ORGANIZER: INVALSI

COORDINATOR: PAOLO BARABANTI

18TH OCTOBER: 2.00 P.M. – 4.00 P.M. {ROOM 3 – RESEARCH 8}

Observing and enhancing talent: the first results from the SOPI tool

Sara Mori – Jessica Niewint – Alessia Rosa – Michela Zambelli

Observing and enhancing talent: the first results from the SOPI tool

The INVALSI 2023 report emphasised that "the equity of a school system cannot be considered only in terms of equal opportunities, but it is also necessary to monitor whether learning opportunities are good and effective for everyone" (INVALSI, 2023, p. 108.). It is therefore necessary to support teachers in observing and monitoring students to provide them with learning opportunities that meet their needs and potential. Observation serves as a critical pedagogical instrument designed to understand students on a personal level and to assess the impact of educational practices, covering both cognitive content and aspects related to relationships and well-being (Baumgartner, 2004). Observation is defined as a process dedicated to exploring and understanding a specific phenomenon through the detailed and accurate description of the features of a particular event, behaviour, or situation, along with the circumstances of its occurrence (Braga and Tosi, in Mantovani, 1995, p. 84). From this standpoint, which places a significant emphasis on nurturing and focusing on the individual, observation appears as a crucial element of educational ability. Its primary aim is not to simplify the inherent complexities of educational relationships but to explore and amplify everyone's capabilities while addressing any challenges.

Taking up then the famous phrase of Don Lorenzo Milani according to which "There is no greater injustice than making equal parts among unequal" the observation becomes "conditio sine qua non" of any process of equity. Only by knowing the student through the observation of his or her behaviour in multiple situations and by setting aside predetermined beliefs and biases, we can develop approaches that move beyond the "one size fits all" mentality towards a more ambitious goal: maximizing the potential and achievements for each individual, addressing their unique talents.

The INVALSI test report also states that "there is no antithesis between excellence and fragility" (INVALSI, 2023, p.117). The question posed by Renzulli (1998), whether talent is an absolute or relative concept, and whether it is static or dynamic, has yet to be answered unequivocally. Theories are divided between those who see talent as a fixed attribute, owned or not by the individual, and those who interpret it as something more flexible and dependent on context and opportunity. The model of Subotnik, Olszewski-Kubilius and Worrell (2021) proposes a complex approach to talent development that considers the importance of context, opportunity, and interaction with mentors. This theory suggests that talent development occurs through the transition from high potential to competence and experience to eminence, emphasizing the key role of domain-specific skills and commitment in the process. Dai's (2021) Evolving Complexity Theory of Talent Development (ECT) emphasizes the complex nature of talent development, seeing it as adapting to environmental challenges and creating a personal niche. Gagné's (2011) theory focuses on how various exceptional abilities are transformed into talent through learning, and practice, emphasizing the importance of environment in nurturing these abilities. Gardner (1999) with his theory proposes an expansion of the concept of intelligence itself, beyond the traditional IQ parameters, by identifying nine distinct types of intelligence. This perspective emphasizes the importance of recognizing and nurturing different intelligences, or potentials, to tailor educational approaches to individual strengths and preferences, promoting a more inclusive and effective learning environment. In this context, inclusion in education must also include a focus on individual educational needs and the personalization of learning processes to enhance the potential of each student.

Subject, objectives and research hypothesis

This paper presents the first results of the experimentation of the SOPI (Instrument for Observing Potential INDIRE) tool which aims to measure the potential of students in the nine domains proposed by Gardner's theory of multiple intelligences, through the observation of the student's behaviour and attitude during a series of educational and relational activities. The tool aims to support teachers in identifying potential by

focusing on two key aspects: the observed ability and the attitude toward specific activities within a particular competency area. It is designed to help teachers in fostering the potential of students from early childhood through late adolescence. The underlying hypothesis suggests that providing teachers with effective tools to observe students' abilities, interests, and motivations can play a significant role in enhancing the customization of teaching strategies and educational goals (Mori et al, 2021). These are crucial for reducing school dropout rates and promoting greater equity within the educational system (Benadusi and Giancola, 2021).

Data used

The report summarizes findings from initial and final questionnaires linked to the classroom testing of the SOPI instrument, involving 69 teachers across various educational levels (14 preschool, 23 primary, 14 secondary I, 16 secondary II, and CPIA) from 24 institutions throughout Italy. These teachers, including both curricular and support roles (N=16), initially completed a questionnaire to assess their interest in the study and their expectations for using the SOPI tool. Out of these, 56 teachers went on to use the SOPI tool with at least one student, resulting in 170 individual assessments. Subsequently, 49 teachers provided feedback via a final questionnaire on the tool's effectiveness, usability, and ease of use.

Method or approach

The study contributes to a larger research and training initiative (Magnoler, 2012; Asquini, 2018), which primarily aims to achieve two goals: firstly, to develop and provide teachers with a tool for recognizing students' potential; and secondly, to collaboratively design and evaluate personalized educational pathways, using various tools with a special focus on the use of technology and artificial intelligence. This paper reports the analyses of the questionnaires involving closed and open questions used before and after the experimentation of the SOPI tool, to understand how and whether the proposed interventions were perceived as useful for improving student observation and designing more personalised teaching.

Results or argumentation

Among the expectations for the SOPI tool highlighted in the initial questionnaire by teachers, the most notable included: enhancing the ability for systematic observation of students' potential; providing support for educational and teaching initiatives through the design of personalized interventions; offering a flexible and user-friendly tool; and ensuring applicability across various contexts. 88% of teachers expressed overall satisfaction with the SOPI tool, with 94% finding it user-friendly and 100% appreciating the clarity and completeness of the instructions. Furthermore, 88% confirmed that the SOPI met their expectations, and 92% recognized it as an effective tool for identifying student potential.

Key factors contributing to the SOPI's effectiveness, as identified by teachers, encompassed detailed and clear domain explanations, domain descriptions relevant across different school subjects, a comprehensive list of both structured and unstructured activities for potential observation, and the tool's broad adaptability to various contexts and situations. Lastly, 92% of teachers expressed an interest in integrating the SOPI tool into their teaching practices.

Both the teachers' initial expectations and the initial results of the experimentation of the SOPI tool seem to support the fact that this type of intervention responds to the teachers' need to have useful tools for observing students to foster their potential. The tool's transversal perspective also makes it possible to explore domains that are sometimes less valued, thus fostering discussion among the teachers themselves. These reflections will be able to contribute to the debate opened in the seminar session, with respect to the definition of useful strategies to improve the fairness of the system; it will also be interesting as a future development of the experimentation, where possible, to associate the observations deriving from the SOPI tool, also with the results of other tools, or with the learning outcomes.

Keywords: talents, observation, tools, professional support to teachers

Beyond Barriers: Gender, Migration Background and Socio- Economic Inequalities in School Performance in Italy

Enrico Deleo – Paola D'Elia – Sergio Di Sano

Introduction. Education is a fundamental basis for individual and social development, but inequalities in access and academic performance are a major challenge on a global scale. It is important to consider gender differences and the impact of students' migrant backgrounds on their academic performance, as well as the

influence of their socio-economic-cultural backgrounds (ESCS). Research has extensively documented the impact of these factors (Goni & Bello, 2016), but there is a specific gap in the understanding of the intersection between gender, origin and ESCS, especially in relation to mathematics and Italian language proficiency. This study aims to investigate these dynamics and explore how they influence the academic performance of Italian student and first-second generation migrant students. Furthermore, recent research has examined the impact of ESCS, gender and immigrant status on English proficiency, highlighting the importance of taking these factors into account when analysing academic performance (D'Elia et al., 2024). Similarly, other studies have investigated the impact of gender, school location and SES on academic performance in mathematics, highlighting that male students living in urban areas and with higher SES tend to perform better (Alordiah et al., 2015). In line with this framework what has already emerged in the literature, this research aims to enrich our understanding of the intersection between gender, migrant background and SES, and to provide new perspectives for promoting inclusive education policies.

Aim and hypothesis

The aim of this research is to investigate differences in performance in Italian and mathematics between native, first and second generation immigrant students in Italy, focusing on the effect of gender and ESCS. It is hypothesised that academic performance in Italian and mathematics differs significantly between the groups, with possible gender advantages varying according to the discipline.

Methodology

WLE scores in Italian and mathematics from the 2022/23 academic year, referring to grade 8 students, were used for data analysis, integrating information on gender, students' immigrant background and ESCS. Analysis of variance (ANOVA) and regression analysis were used to assess differences in performance and the effect of ESCS on performance in Italian and mathematics, controlling for gender and origin variables.

Results

The ANOVA analysis and regressions carried out revealed significant differences in academic performance between native, first-generation and second-generation immigrant students in Italy. Specifically, the ANOVA analysis revealed significant differences in performance in Italian ($F(2, N) = 45.78, p < 0.001$) and mathematics ($F(2, N) = 32.56, p < 0.001$), highlighting that the immigrant background of the students has a significant impact on the results. Furthermore, in mathematics, first-generation female migrants scored higher than their male counterparts, with $M = 178.10 (36.67)$ and $M = 159.58 (47.4)$ respectively.

Further investigation with regression analyses revealed that socio-economic-cultural background had a significant effect on performance in both Italian ($\beta = 0.37, p < 0.001$) and mathematics ($\beta = 0.29, p < 0.001$), after controlling for gender and students' origin.

Conclusions

This study adds to the understanding of the combined effects of gender, origin and ESCS on academic performance in Italy. The results highlight the importance of considering the complexity of the intersections between these variables in the development of inclusive educational strategies. Studies have shown that the intersection of gender, migrant background and ESCS shapes students' educational opportunities and experiences in complex ways, drawing attention to the importance of educational policies that are sensitive to these dimensions (Green, 2015). Integrating the findings with the theoretical context provided by the literature, including the study by D'Elia et al. (2024), highlights the need for policies that promote equity in education.

Finally, the results presented here paint a picture in which native students generally tend to outperform their immigrant counterparts, but the performance gap is partly due to differences in socio-economic-cultural backgrounds. Importantly, there are also significant differences between immigrant students: second-generation immigrant students perform closer to native students than their first-generation peers, suggesting a possible assimilation of educational advantages over time. These data underline the importance of targeted educational interventions that take into account not only differences in origin, but also the specific challenges posed by socio-economic-cultural background.

Keywords: Migration background - School inclusion - Academic performance

Challenges and opportunities for an inclusive school: the potential of Evidence-Based Education (EBE)

Giusi Antonia Toto – Giorgio Mori

In the context of advancing digitization and the introduction of artificial intelligence systems, the education sector is exploring ways to provide quality education that is accessible to all students, especially as questions about the concept of inclusion and best teaching practices to make this perspective effective emerge with increasing insistence. One hundred years after the theories of Russian scholar Lev S. Vygotsky laid the foundation for school and social inclusion and triggered profound transformations in modern pedagogy and special education, Italy and many other national education systems are still working to fully realize inclusivity within their school structures. It is also of paramount importance to maintain constant attention to the needs of learners, to continually evaluate the effectiveness of teaching practices, and to design curricula that truly meet the needs of learners. Among the educational methodologies that best support the improvement of teaching practice and particularly in heterogeneous classrooms, the paper highlights how Evidence Based Education (EBE) provides innovative teaching strategies and effective strategies capable of meeting the individuality of each learner. The educational approach proposed by EBE, that is, "the integration of professional wisdom with the best available empirical evidence in making decisions about how to deliver instruction" (Whitehurst, 2002), is guided by empirical evidence, in that its application occurs as teachers choose teaching methods supported by reliable data, going on to integrate them, then, into a curriculum that is adapted to the specific instructional environment in which they operate (Moran & Malott, 2004). EBE is thus shown to be particularly akin to principles of individualization and personalization that are typical of special pedagogy and inclusive education, in that attention to the specificity of the individual student allows the teacher to tailor educational interventions according to the needs and real abilities of each individual. Approaches such as EBE and Data Based Decision Making (DBDM) promote planning for transition that is based on observable and measurable post-secondary goals that are constantly monitored to identify any necessary modifications to the transition path; implementation of an internship program that integrates theory and practice; active participation of families in the transition process; teaching social skills essential for workplace relationships; and collaboration with community agencies, can ensure effective work and social transitions for students with intellectual disabilities into adulthood. By resorting to EBE-related practices, teachers are able to implement methodologies that are successful in the context of inclusive classrooms. Therefore, it would be desirable for such a research and teaching model to be implemented in teacher training pathways: allowing, for example, TFA support participants to learn the methods and strategies experimented with EBE during their training would ensure an improvement in the teaching environments in which they work, an increase in their own sense of self-efficacy in classroom management, and a greater awareness of their role (Coggi & Bellacicco, 2023). Teachers, in fact, are facilitators of learning and, therefore, must possess skills to promote inclusion. Classroom teaching, in this sense, would undergo a significant change, resulting in the well-being of teachers and students, as well as an inclusive ecosystem that is attentive to valuing everyone's specificities (Batini, 2023). From this perspective, EBE is a clear driving force when it comes to the advancement of inclusive processes, as it is able to offer teachers such tools and resources to readily address the challenges that heterogeneous classrooms present every day. Based on the present reference scenario and with a view to an in-depth study on educational inclusion in the field of specialized teacher training, an exploratory survey was conducted at the University of Foggia's TFA Support involving a representative sample of 596 participants, corresponding to 81.9 percent of the target population. The study aimed to explore the attitudes and beliefs of future specialized teachers toward inclusion, making use of the SACI questionnaire (Castellana et al., 2023). The survey results indicate that the majority of the sample (52 percent) has never received specific training regarding processes and tools to foster inclusion. In addition, a significant proportion of respondents (38.4 percent) perceive promoting the participation of all students in a context of educational diversity as a daunting task, while another substantial proportion (43.6 percent) do not believe that heterogeneity in learning levels necessarily results in obstacles in teaching practice. Evidence emerges from the research on the importance of Evidence Based Education and practical experimentation as a driving force for the advancement of inclusive processes. Such approaches provide teachers with fundamental tools and resources to deal effectively with the complexities of teaching situations characterized by heterogeneity of students and educational needs. The analysis highlights challenges in the practical implementation of inclusion, emphasizing the need for ongoing support and

training for teachers. Despite frustration in organization and limited recognition of the role of support teachers, attention toward BES students remains high. Although there is positive support toward BES students, there is less effectiveness in involving colleagues in instructional planning. This highlights the importance of a collaborative school culture to maximize learning opportunities for all students, regardless of their needs.

Keywords: Inclusion; Evidence Based Education (EBE); Special Educational Needs; Heterogeneous Classes; Teacher Perception

Inequality of educational opportunity in Italy: A machine learning approach

Moris Triventi – Paolo Brunori – Emanuele Fedeli – Giovanna Scarchilli

A strong relation between students' ascriptive characteristics and their educational attainment challenges educational systems, both in terms of their equity and efficiency. First, a society in which the socioeconomic background is a strong predictor of educational outcomes may be considered to be unfair. Second, a society characterized by a low level of social mobility, discriminating against competent and hard-working students from economically disadvantaged families misallocates human capital investment. While for a long time social scientists have analysed educational attainment as the main outcome, the increasing importance of human capital in contemporary economies led researchers to focus also on more direct measures of individuals' competencies. Today, the availability of data sources with standardized measures of pupils' scholastic performance allows to investigate how students' competencies are shaped by a set of individual and contextual characteristics. A first research stream investigates how students' achievement is affected by socio-economic resources of the family of origin. Empirical evidence, both at the international level (OECD, 2018) and in Italy (Invalsi, 2019) reports that social origin is strongly associated to children's competencies in key subjects, both at the mean and along the distribution (Giambona & Porcu 2015; Costanzo & Desimoni 2007), that these differentials start early in life and continue to be substantially strong also in upper secondary school and during young adulthood (Dammrich & Triventi, 2018). Additionally, gender and migration background are key predictors of students' competencies level. While boys tend to perform relatively better than girls in mathematics, the gender gap is larger and in favour of girls in Italian (Contini et al. 2017). There is agreement across studies that students with a migration background perform less well in standardized tests than natives, a gap that is especially pronounced for first-generation pupils (Triventi, Vlach & Pini 2022). A second line of inquiry examined how student performance varies across geographical areas, reporting large heterogeneity in test scores across Italian regions (Bratti, Checchi, & Filippin, 2007; Agasisti & Cordero-Ferrera, 2013; Argentin & Triventi, 2015) and educational standards across schools (Argentin et al., 2017), usually favouring the North.

With this work we aim to develop a bridge between these two research streams, which, so far, have been rather detached from each other. Namely, we analyse whether the role of ascriptive characteristics is similar or systematically varies across geographical areas. This allows us to establish whether not only the effectiveness of education differs across Italian regions but also whether inequality of educational opportunity (IEOpp) is heterogeneous across territories. A second contribution of our work consists in adopting a machine learning approach to shed light on the role of students' circumstances on student achievement in Italy, which has interesting potentials in the study of educational inequality but it has not been widely applied (Masci, Johns, Agasisti, 2018). Third, we provide a comparison with more traditional regression-based methods to assess similarities and differences across analytical approaches.

We use data from Invalsi-SNV, which monitors students' competencies in Italian and mathematics across key school grades. In this work, we put together data on the 5th grade (end of elementary school) from 2012 to 2022. By exploiting the large size of the Invalsi dataset, we are in the position to disaggregate the analysis by provinces, thereby gaining a more fine-grained perspective on territorial heterogeneity in educational inequality. Moreover, contrary to what is done in the larger part of the analysis of the Invalsi data we do not focus on the summary index of economic, social and cultural status (ESCS) (Campodifiori, Figura, Papini, Ricci, 2010). Indeed, by applying machine learning algorithms we are able to consider all items contained in students' questionnaires about parental education, parental occupation, and home educational resources without risk of overfitting the data.

Empirically, the identification of social types is achieved through a data-driven approach using transformation trees, a learning algorithm introduced by Hothorn and Zeileis (2021). These trees partition the population based on observable features and describe the outcome distribution for each type. To address the variability inherent in tree-based methods, we employ a bootstrap aggregation technique known as bagging, which provides insight into the uncertainty surrounding the estimated inequality of opportunity structure. It is important to note that embracing a data-driven approach implies becoming agnostic about the data generating process which mean that the algorithms may select different models across contexts, introducing important flexibility in the modelling strategy. The Shapley decomposition will be used to compute the relative contribution of individual circumstances in the measurement of inequality of opportunity. By considering all possible combinations of circumstances and their impact on IEOp, the Shapley decomposition provides a fair allocation of importance to each circumstance based on its unique contribution. In our setting, this approach offers an effective tool for understanding how different circumstances influence the extent of IEOp and whether their importance changes across territories and school grades.

Our analysis aims at producing maps showing the heterogeneous magnitude of the effect of social background on students' outcomes across the Italian provinces, which will provide useful insights on which socioeconomic variables are more conducive to academic achievement and in which contexts (Gregorutti et al., 2017).

Keywords: inequality of opportunity; educational achievement; social inequalities; geographical differences; machine learning

A Teacher Professional Development Research to Contrast Educational Poverty

Giorgio Asquini – Gabriella Agrusti – Ira Vannini

The severe difficulties faced by Italian schools in recent years as a result of the Covid-19 pandemic have contributed to increasing educational inequalities, especially where they were already substantial. The results of the first available surveys show the effect of a further limitation in access to schooling and an increase in so-called "implicit dropout" for children and youth in educational poverty. The secondary school system shows an increasing differentiation between the learning prospects of students from more affluent socio-cultural backgrounds and those of students from disadvantaged socio-family backgrounds (see Invalsi Open website and OCSE Pisa, 2018, with obvious negative impacts on the education system desired equity. Acting promptly is essential to succeeding in creating, in the school, the conditions for the recovery of this socio-cultural disadvantage, creating inclusive and welcoming school contexts, capable of promoting a climate of well-being, participation and collaboration, leveraging the individualization in pedagogy and in teaching practice as the foundation of all policies on national curricula. This entails, first and foremost, orienting the actions of the school and teachers toward the diversification of teaching-learning proposals with a view to achieving unquestionable goals for all students, in the logic of guaranteeing the right to active citizenship, through an equitable and significant development of key competencies. To address these educational emergencies, one of the CRESPI Center's research groups (www.site.unibo.it/crespi/en) is carrying out a project (PRIN 2022 Next Generation funding, Project Code 20224WRLF0) that aims to define and develop strategic and empirically grounded intervention models to support the professional development of Italian secondary school teachers carrying out experimental interventions in schools located in urban and suburban areas of five cities in Northern, Central and Southern Italy (Milan, Bologna, Perugia, Rome and Palermo), in order to contrast the phenomena of educational poverty and school failure.

Theoretical Framework

Educational poverty and socio-economic, linguistic, and cultural disadvantage are key multidimensional concepts that refer to numerous areas of deprivation concerning the individual and the socio-economic context in which he or she is placed (Sen, 2000). It is possible to trace back to situations of educational poverty all those economic, social and individual factors that are obstacles to individuals' access to educational experiences that enable the acquisition of multiple literacy processes (New London Group, 1996) needed for the achievement of a healthy and well-integrated development in the social environment of reference (Save the Children, 2014). Educational poverty is linked to several dimensions within a complex framework, where economic poverty, understood as the lack or total absence of resources

necessary for the satisfaction of needs related to a decent life (Sen, 1994, 2000; Nussbaum, 2012), acts as a concomitant variable to the formation of contexts that limit educational opportunities and exacerbate conditions of disadvantage and social inequality (Benvenuto, 2011). Educational poverty may also occur due to life contexts marked by adult inattention, family conflict, and excessive use of social networks. Personal contexts can also determine limits to children's right to education. To address these educational emergencies, it is important to refer to the conceptual framework of WHO's ICF (2011), which requires identifying, within contexts, the material, cultural, and social barriers that must be broken down in order to promote the implementation of facilitating interventions. Taking action sooner rather than later is critical for two reasons:

1. Socioeconomic, linguistic, cultural disadvantage is frequently observable in Italian classrooms: 9% of pupils in Italy have a special educational need and 35% of these pupils are affected by socioeconomic, linguistic, cultural disadvantage (ISTAT, 2020)
2. Scientific literature shows that these pupils are at high risk of educational failure, which can lead to reduced employment opportunities and an increased risk of social exclusion (OECD, 2012). In Italy, in the 2016/17 and 2017/18 school years, 1.17 percent of lower secondary pupils and 3.8 percent of upper secondary pupils dropped out at the beginning of the school year (MIUR, 2019). The percentage of NEETs also reached 20.7 in the second quarter of 2020 (European Union, 2020). These data suggest the importance of properly considering this target group of students.

Objectives of the research project

1. To thoroughly analyze and understand, through the implementation of observational and correlational research designs within each of the identified school cases, what has happened in some urban and suburban school settings since the Covid-19 crisis, focusing on the main social and educational factors that may foster or counteract inequalities in student learning outcomes
2. To design and validate an integrated intervention model based on teachers' participation aimed at promoting their planning, teaching and assessment skills. This intervention model is in line with the methodological approach of Teacher Professional Development Research (Asquini 2018) and it attempts to support teachers in overcoming current obstacles in implementing active teaching-learning strategies tailored according to students' needs. This should allow to achieve good key competency levels for all students, ensuring quality and equity of outcomes.

State of the art and proposed presentation at the Invalsi Conference

The Research Group is currently carrying out the observational and correlational survey phase (with case studies and a convergent qualitative-quantitative plan, Creswell et al., 2011) related to Objective 1. In particular, a wide set of macro-variables useful for investigating risk and protective factors with respect to socio-cultural disadvantage in secondary school have been considered, both through the voice of teachers and grade 6 and 7 students (with ad hoc cognitive tests and specific questionnaires for the two populations). Scientific collaboration agreements have been established with the 10 participating schools to enhance their contribution to the research. The observational survey phase will also allow to define clearly the educational needs of each school, so that during academic year 2024-25 a series of innovative educational interventions aimed at neutralizing the factors that determine educational poverty and strengthening the factors that can determine equity and quality of learning can be applied.

The interventions in the different contexts will be accompanied by a continuous exchange between researchers and teachers, in the spirit of Teacher Professional Development Research. Activities will start with the shared design of the interventions, will continue with the monitoring of the teaching activities, and will be completed with the validation of the intervention project and the evaluation of the results obtained. In addition to aspects related to learning, those related to teachers', students', and families' well-being will also be considered. The overview of the interventions implemented in the 10 schools will allow for the production of a set of guidelines on combating educational poverty, in line with one of the impact objectives set out in the PRIN 2022 project.

The availability of Invalsi (8th grade) data on the achievement levels in Italian Language and Mathematics from 2019 to 2023 of students in the selected secondary schools constitutes an important opportunity to outline - in a longitudinal sense - the strengths and weaknesses in the promotion of basic skills in those educational contexts. This increases the chances of a valid, coherent and timely diagnosis of teacher training needs, which is useful to define targeted projects for each school context involved, also cross-referencing the data with local internal variables (such as teacher turnover and the incidence of non-tenured staff). Later, in the validation phase, it will be useful to compare the data from the Invalsi tests in the exiting classes involved in the project, particularly with regard to the variance of results between and within classes. At

the end of the project, it is also expected the opportunity to promote, among teachers in the selected schools, an improved ability to read, interpret, and use the results of the national tests.

Keywords: Educational poverty; socio-cultural disadvantage; Teacher Professional Development Research; Case Studies

SESSION 9. INEQUALITY AND EDUCATIONAL OUTCOMES
SESSION 8. RELYING ON ADMINISTRATIVE DATA TO ANALYZE STUDENT CAREERS

ORGANIZER: HARVARD AND BOCCONI UNIVERSITY - UNIVERSITY OF PIEMONTE ORIENTALE

COORDINATOR: LINDE STALS

18TH OCTOBER: 2.00 P.M. – 4.00 P.M. {ROOM 4 – SESSION 9 – SESSION 8}

Longitudinal effect of early digital experiences on standardized learning outcomes during school career

Marco Gui – Giovanni Maria Abbiati – Chiara Respi – Sofia Ercolanoni

There is a heated debate both among scholars and in the media about early access to digital media by children and adolescents: parents and educators seem disoriented and ask if anticipating access to the Internet is a good or a bad thing (Livingstone&Blum-Ross 2020). In the meantime, early access to the Internet is more and more common: for example, the moment of arrival of smartphones is constantly anticipated among early adolescents, and the covid-19 pandemic has further accelerated this trend (Mascheroni & Siibak 2021).

Simultaneously, evidence is emerging that digital media practices during childhood and early adolescence constitute a new dimension of social inequality. Smartphones, video gaming consoles and other forms of autonomous access to the Internet seem to be anticipated - somehow paradoxically - especially among deprived families in terms of socio-economic background (Gui et al. 2020). Research is also highlighting that problematic digital media use is more likely to manifest in families where socio-economic conditions are more deprived (Gerosa et al. 2021). In turn, early access has been found to be associated with problematic media use once children enter adolescence (Jalouk & Boumosleh 2018). In particular, problematic digital media use is more frequent among adolescents from immigrant families, females and young people from less advantaged socio-economic backgrounds (Vitullo et al. 2021; Gui 2015). The risks of an early, unguided and excessive use of digital media, therefore, are not equally distributed in society, making this issue a potential new source of social inequality.

However, the empirical results obtained on the relationship between early digital media use wellbeing in the long-run are scarce (see Gerosa et al. 2024). In particular, the negative impact of early digital use on learning outcomes that have been highlighted are often reproached for not being solid from a methodological point of view, as they are based on correlational or short-term longitudinal data (Amez & Baert 2020). Also, research gaps emerge as to the impact of age at first smartphone on learning outcomes, measured with standardized and solid methodological approaches (Gui & Büchi, 2021). More importantly, there is a very urgent need to gather robust empirical evidence about why and how online problems can be largely predicted by young people's vulnerabilities offline (Odgers & Jensen 2022), that is how early access to permanent connection interacts with social inequality in the lives of youths.

To fill these gaps, we aim to provide the most rigorous estimation so far of the impact of an early autonomous access to smartphones, videogames and social media on the learning outcomes of high school students, and its interaction with gender, migratory background and parental education. This becomes possible as - for the first time since INVALSI (the Italian National Institute for the Assessment of the School and Training system) started to collect data on students' outcomes - it is possible to follow the entire school careers of current 10th and 11th graders. Exploiting this possibility, we merge an ad-hoc retrospective survey data about screen use during childhood and early adolescence with students' longitudinal learning outcomes at grades 2nd, 5th, 8th and 10th. A dataset with 6738 high school students in northern Italy, representative by school type, is analyzed to answer the questions above by means of longitudinal analysis. Preliminary results confirm that early digital use is more likely in socio-economically deprived families and that age of smartphone ownership has a negative impact on learning outcomes in low and high school for those with pre-existing habits of intensive media use during childhood. The paper discusses the implications of these findings for communication and education theory, in particular for what concerns digital inequality in technologically mature contexts.

Keywords: Educational inequality; Student achievement; Children, adolescence and media; Digital parenting

Indicators for assessing secondary school effectiveness and peers' impacts on facilitating tertiary education access and completion

Isabella Sulis – Mariano Porcu – Silvia Culumbu – Cristian Usala

It is largely argued in the educational framework the role that coming from a disadvantaged socioeconomic background plays in the intergenerational transmission of inequalities. Achieving high levels of education is a strategic asset to reduce disparities. Higher-educated people have a greater awareness of their abilities and opportunities to build a better future, and this ensures better economic rewards in the long-term period: higher educational levels are generally associated with higher employment rates and salaries, and in general, education is associated with less inequality in income distribution and human development. Data from Italy show that the achievement of a tertiary education degree in Italy is still strongly associated with the family of origin and disparities and persists across advantaged and disadvantaged people. The probability of getting a university degree for individuals in the cohorts 1960-69, 1970-79, and 1980-89 coming from highly educated families is about .5 higher than for those coming from low-educated families (Busetta et al., 2023). Furthermore, the transition rate from high school to university is strongly associated with the kind of schools attended, which is conditioned by parents' level of education and family socioeconomic status. Moving from this framework, the main aim of this research is to inspect factors affecting the school's capability to foster fairness and inclusion in education by maximizing the opportunities for students to reach the highest level of education regardless of their socioeconomic and migratory background (Field et al., 2007; OECD, 2012a; Agasisti & Longobardi, 2014; OECD, 2016a; Sulis et al., 2020) and to detect disparities in schools' behavior at geographical level. Taking inspiration from the system of indicators built up by OECD using the results of PISA's surveys to monitor the state of quality of equity of the educational systems and to make cross-national comparisons, in line with the PRIN 2022 project "From high school to university: Assessing peers' influence in educational inequalities and performances", this research aims to provide a set of indices to measure the efficiency of the secondary school system and the peers' influence on access and completion of students' tertiary education cycle. For this sake, multiple data sources have been merged: the MOBYSU.IT data which contains the Italian National Student Archive (ANS) microdata related to the population of students enrolled in Italian universities and the INVALSI grade 13th surveys that provide information about student families' socioeconomic conditions. Using both sources several indices of the school effect have been built up at high school level to assess the socioeconomic gradient with respect students' performances, which measures the effect on an increase in students' family socioeconomic and cultural status on the expected results of the INVALSI tests in the main skill areas investigated by the survey and the indicators of students' performance during the 1st year at the university. Moreover, the PISA and OECD definition of resilient students, which identifies as resilient those students who achieve a high level of proficiency in the PISA test (score in the third quartile) despite facing adverse socioeconomic circumstances (ESCS index in the first quartile), was translated to be adopted to identify resilient students at the university, by focusing on the regularity of students' career at the end of the 1st year. We defined resilient students as those coming from families with low ESCS index who achieved a number of formative credits higher than the median observed in the disciplinary field where they enrolled. The joint system of indices has been used to profile schools based on their capability to foster quality (measured in terms of ensuring high performance in tertiary education), equity (measured in terms of capability to promote inclusion and fairness apart the socioeconomic status), and resilience, measured in terms of capability to allow disadvantaged students to reach a high level of performance. The paper also provides an explanatory analysis to assess the influence of peers' characteristics in terms of advantaged and disadvantaged backgrounds of students attending the same schools or the same degree programs at the university in fostering quality, equity, and resilience. For this sake, the following indices have been built: the average ESCS of classmates/schoolmates who choose the same university; the average ESCS of classmates/schoolmates who choose the same degree program; the average ESCS of classmates/schoolmates who choose the same degree program at the same university. The combined use of multilevel modelling approaches to deal with complex data structure (i.e. nested observations in schools, field of study and universities) and latent profile regression analysis allowed us to detect the role played by

school's characteristics in determining the capability of the institutions to boost resilience and to create profiles (i.e. latent classes) of schools which are able to combine quality and equity in terms of expected results of their students at the university.

Keywords: Index of socioeconomic gradient, university performance, school effect, MOBYSU.it

The resilience in test scores and the mitigating effect of summer schools one year after the COVID-19 school closures

Letizia Gambi – Kristof De Witte

Introduction & research hypothesis

To mitigate the adverse impacts of COVID-19 related school closures (see reviews by Betthäuser et al., 2023; De Witte & François, 2023; Donnelly & Patrinos, 2021; Moscoviz & Evans, 2022; Patrinos et al., 2022) and to enhance the resilience of their education systems, several countries implemented targeted remedial programs throughout the school year and summer holidays (see review of policy responses in Europe by De Witte & Smet (2021) and the European Commission (2022)). While most countries have allocated part of their budget to such initiatives (UNESCO et al., 2021), there is limited evidence regarding their effectiveness. In this paper, we examine the evolution of the between-cohorts learning deficits and the dispersion in student test scores within schools one year after the COVID-19-related school closures. A key contribution and innovation of this paper is our examination of the effects of summer schools, a policy initiative by the Flemish government in 2020 primarily targeting low socio-economic status (SES) students. This study extends the current literature by exploring the role of these educational interventions as a tactical response to the disruptions caused by the COVID-19 pandemic, thereby providing critical insights into the effectiveness of these programs in mitigating the adverse impacts on education. We used a rich dataset with administrative data and standardised test scores from private Catholic schools representing the largest primary education provider in the Flemish region of Belgium. We identify between-cohorts learning deficits compared to 2019 levels in all subjects but social science.

Our analysis reveals that, although summer schools targeted at low SES students have effectively mitigated further learning deficits in these schools, disparities in educational attainment persist. A quantile analysis suggests that mathematics test scores of the best-performing students within a school have significantly declined. This paper contributes to the understanding of education impacts stemming from educational disruptions, highlighting the role of remedial interventions in mitigating these effects.

Data

The analysis relies on a panel data-set covering a seven years period from 2015 to 2021. The data-set combines: (1) data on standardised tests that are administered every year in June by the network of Catholic schools in Flanders (Katholiek Onderwijs Vlaanderen), in the last year of primary school (grade 6); (2) administrative data comprising several school characteristics, characteristics of students in grade 6, and teachers' characteristics; (3) and administrative data on summer schools, for the year 2020. These administrative data do not allow to identify individual student characteristics and, as such, these data are aggregated at the school level.

Methods

We examine the effect on test scores and measures of dispersion one year after the first wave of COVID-19 school closures. To distinguish between the overall impact of COVID-19 and its change in 2021, we construct two variables.

First, we construct a dummy variable 'COVID – 19' to indicate if the test scores were measured after the COVID-19 crisis. Hence, the dummy equals 1 in the years 2020 and 2021, and 0 otherwise. This variable measures the overall effect of the COVID-19 school closures.. The estimated coefficient can be interpreted in terms of change in standard deviations due to the COVID-19 crisis. Second, we construct a dummy variable 'Change in 2021' to indicate if the test score was measured in 2021. This dummy, equal to 0 in all other years, measures the change in test scores one year after the first wave of COVID-19 school closures. Combining the two dummy variables in one regression model provides the net effect of the extent to which the resilience in test scores after the closures (i.e., 'Change in 2021' dummy), ceteris paribus the impact of COVID-19 on test scores (i.e., COVID-19 dummy).

The immediate effect of the COVID-19 school closures, and its average change of standardised test scores in 2021, is investigated by means of the following regression model:

$$y_{i,t} = \alpha + \beta(\text{COVID} - 19)_{i,t} + \gamma(\text{Change in 2021})_{i,t} + \delta X_{i,t} + \lambda \text{Test}_{i,t} + \theta \text{Trend} + v_i + \epsilon_{i,t} \quad (1)$$

where $y_{i,t}$ denotes the analysed outcome variable (i.e. test score in each subject, measures of dispersion in test scores) of school i at time t . The outcome is regressed on a constant α , the dummy (COVID-19) and the (Change in 2021) dummy. β and γ represent the two main coefficients of interest as β identifies the immediate effect of the school closures, while γ isolates the sole effect on test scores 2021 relative to 2020 results (a measure of resilience in test scores). The school fixed effects (v_i) remove any omitted variable bias arising from unobserved time-constant heterogeneity across schools (e.g. school and school board management quality, location of the schools, etc.) that may be correlated with regressors. Student fixed effects and, similarly, classroom or family fixed effects, cannot be included in model (1) as we do not observe the same students over time (i.e. we cannot account for student-specific characteristics). Next, Test are dummies capturing differences between the test versions. Trend is a continuous linear time trend that controls for the overall change in test scores over the years with θ measuring the unit change in test score for a year change in the trend variable. $X_{i,t}$ denotes a vector of time-varying control variables at school, grade 6 and teacher level. Finally, the standard errors $\epsilon_{i,t}$ are clustered at the school level to capture the intra-school correlation of the outcome variable.

Results

The results indicate a significant between-cohorts decrease in all test outcomes, except for social science, since the start of the pandemic. The point estimates vary from -0.15 SD for the native Dutch language, to -0.10 SD in sciences and -0.07 SD in math. Furthermore, the estimates provide a mixed picture on the resilience of test scores one year after the COVID-19 related school closures. The significant learning deficits halted for mathematics and science, and significantly improved for social sciences (+ 0.06 SD). For the Dutch language, the learning deficit that occurred in 2020 further deepened with -0.09 SD in 2021, indicating an additional decline of about 70% as compared to 2020 decline of -0.15 SD. Given the larger variation in test scores within schools than between schools at the student level, we observe smaller point estimates, although overall the estimated coefficients have the same direction and significance as the school level estimates. With respect to the dispersion in test scores, the results suggest that test score dispersion within schools rose somewhat for the Dutch language in 2021, and decreased for mathematics.

However, not all students and schools are affected in the same way. Notwithstanding the halted learning deficits in math in 2021, a quantile analysis suggests that the math test scores of the best-performing students in a school (i.e., quantile 70 to 95) have significantly declined, while those of low-performing students do not significantly differ from those of 2020. For the Dutch language, the quantile analysis suggests that all students showed learning deficits, irrespective of the quantile. When considering marginal effects of the 'Change in 2021' dummy based on socioeconomic indicators, we find that learning deficits increase in 2021 for those schools that count more students with lowly educated mothers. Lower average decreases in 2021 attainment (though insignificant) are found for schools with students from a more disadvantaged neighbourhood. At the same time, schools with a more advantaged socioeconomic status composition in terms of neighbourhood of residence, and home language, seem to experience higher (significant) learning deficits one year after the pandemic. Importantly, we observe heterogeneity in test scores with respect to the availability of summer schools, which are publicly financed targeted remedial actions aiming at reducing COVID-19 related learning deficits. In 2021, the learning deficits observed in the postcode areas with a summer school have halted for both the Dutch language and mathematics. At the same time, the learning deficits in Dutch increased in those postcode areas where no summer school took place.

The findings of this paper have important policy implications considering the unprecedented challenge posed by the school closures on the educational system and the policy actions undertaken so far (De Witte & Smet, 2021; UNESCO et al., 2021). First, it seems that disadvantaged students have partially recouped the learning deficits upon returning to school, or at least at a faster rate than the best-performing students. Our finding suggests that targeted policy actions, especially summer schools aimed at the most vulnerable students, were relatively successful. Nevertheless, our findings suggest that significant policy attention should also be directed toward the best performing students, who appear to lag behind one year after the pandemic. Thus, the current set of policy actions should be supplemented with the implementation of remedial measures designed for all students, including the best-performing students. Second, the COVID-19 pandemic seems to accentuate existing trends in school outcomes. Specifically in the Flemish education

system, the COVID-19 pandemic reinforced the already downward trend of Flemish students' performances in international tests (e.g. PIRLS, PISA, TIMSS).

Keywords: School closures (COVID-19); Summer schools; Standardised tests (Flanders, Belgium); Inequality; primary school

Do Financial Incentives Enhance University Participation and Achievement? Evidence from the Great Recession

Loris Vergolini – Nadir Zanini

Introduction and motivation

In this paper, we illustrate the main empirical results of a research project aimed at evaluating the effectiveness of a grant provision, named Grant 5B, implemented in the Province of Trento – an area in the North-East of Italy – from 2009 to 2012. It is a merit-based financial aid for students from low-income families and it should cover the indirect costs connected with the participation to HE, while the direct costs are covered by national level grants. The target population comprises students resident in the Province of Trento for at least three years who have successfully completed the last year of secondary school, obtaining a final grade above 93/100, and whose family equivalent income is below € 30,000.

From a theoretical point of view, financial aid policies rely on the idea that the role of students' family income and social origins is crucial for their transition to HE (Shavit et al. 2007; Triventi 2013). Students and their families have to face several direct and indirect costs if they want to enrol at the university and these costs are less affordable for people from lower social strata. This means that they have a liquidity constraint problem and, therefore, a generous grant could foster HE participation by reducing direct and opportunity costs. Another approach suggests that social origins shape both ability and expectations of children. Hence, students from disadvantaged background do not enrol for a lack in abilities and motivations, while the role of liquidity constraints is marginal (Carneiro & Heckman, 2002). This research strand suggests that it will be more valuable to intervene on children's motivations and competencies during their school career, when they are very young. Indeed, cognitive abilities are formed very early in life and it is more difficult to intervene as children grow up. Our goal is to study the effects of a specific grant introduced in a period of economic recession at local level that aims to increase university enrolment, academic performances and to reduce social inequalities by manipulating financial costs.

Data, variable and methods

The dataset used in this paper is the result from a linkage procedure of administrative and survey data covering four consecutive cohorts of students (2009-2012). The combination of the two sources allows the creation of a unique database on more than 10,000 students that account for about the 80% of the target population. In order to identify the causal effects of the Grant 5B we exploit a regression discontinuity design strategy (Cattaneo et al. 2019) comparing the outcomes around the threshold values that is determined by the final mark (93/100).

Empirical results

The main findings are summarised in Tables 1 and 2. The first row of Table 1 reports the results pooling the four waves. It is clear that the Grant 5B does not have any impact on enrolment probability, but it is able to change the choice of the university location. Indeed, eligible students have a higher probability to enrol outside Trento in faculties that are not present at the University of Trento (i.e., Medicine). Table 2 also show the trend over time of the effect of the Grant on the main outcomes. We have two main results: the increase on the effects on enrolment probability and the disappearance of the effects for the other two outcomes. Our explanation of these changes refers to the persistence/worsening of the economic crisis that strikes Italy starting from 2009. In the province of Trento, the effects of the economic crises on youth condition start to be dramatic from 2011. Indeed, in the province of Trento youth unemployment rate raised dramatically from 11.5% in 2009 to 20.5% in 2012, while in 2010 and 2011 it is respectively 15.1% and 14.5%. In 2008, before the economic crisis and before the implementation of the measure, youth unemployment rate was equal to 8.5%. It is evident that the largest jump in unemployment happened in 2012 (6 percentage points), while from 2008 to 2009 the variation was smaller (3 percentage points). It seems reasonable to suppose that the change in the effects of the Grant 5B effect could be due to the persistence of the unfavourable economic conditions.

Table 1. Sharp (ITT) Rdd estimates of the treatment effect on enrolment choices.

	(1) Enrolment			(2) Enrolment outside Trento			(3) Enrolment in faculties absent from UniTN		
	ITT	SE	N	ITT	SE	N	ITT	SE	N
Pooled	0.0566	0.0422	1,233	0.2329***	0.0597	1,185	0.1766***	0.0519	1,292
2009	-0.0351	0.1028	259	0.3199***	0.1113	228	0.3592***	0.1117	172
2010	-0.0044	0.1059	241	0.2867***	0.1370	283	0.2779***	0.1310	300
2011	0.1938*	0.1080	261	0.2095*	0.1209	232	0.1061	0.1130	216
2012	0.0954	0.0821	311	0.1827	0.1325	263	-0.0057	0.1008	243

Note: Levels of significance are reported as follows: *** p<0.01; ** p<0.05; *p<0.1.

Table 2 reports the effect of Grant 5B on students' behaviours and on academic performances for the first two cohorts that are the cohorts for which we are sure that there are non-effects on enrolment probability. In fact, in case of effects on enrolment rate we won't be able to disentangle the effects of the Grant from the effects due to change in the composition of enrolled students induced by the Grant. The first set of outcomes pertain to the behaviour of students during their time while at university. For both the time spent studying and working, the estimates show that there is no significant impact of the Grant 5B. Moving to the outcomes pertaining measures of academic performance, the first dimension of interest is students' drop-out rate. Whereas, for the Bachelor's students, there does not seem to be an impact of the Grant 5B on the probability of drop-out at the end of the second and third year, Grant 5B recipients seem to have a 20-percentage-point decrease in the probability of dropping-out by the end of the first year. If this seems to suggest that the financial aid provided by the Grant 5B reduces the risk of drop-out, the fact that ITT is not statistically significant hinders the possibility to make a definitive conclusion. On the other side, for students enrolled in single-cycle degree the estimates are smaller than in the previous case.

In the first year for Bachelor's degree, Grant 5B recipients seem to perform only slightly worse, in terms of GPA, than other students, but in the third year the estimate of the impact is positive and much more higher than in the first year, leading to around an extra 2/30. However, similarly to the impact of drop-out, also for performance the estimates are not statistically significant.

Table 2. Sharp (ITT) Rdd estimates of the treatment effect on academic performances, time use and timely graduation. Bachelor's and single-cycle degrees.

	Bachelor's		Single-cycle	
	ITT	S.E.	ITT	S.E.
Average hours studied	-5.926	(5.591)	8.943	(8.102)
Average hours worked	1.923	(3.842)	-5.634	(4.987)
Drop-out at the end of the 1st year	-0.234	(0.155)	-0.003	(0.042)
Drop-out at the end of the 2nd year	0.076	(0.132)	-0.038	(0.040)
Drop-out at the end of the 3rd year	0.091	(0.173)	0.129	(0.310)
GPA at the end of the 1st year	-0.171	(1.040)	1.407	(1.695)
GPA at the end of the 2nd year	0.866	(1.156)	0.379	(1.735)
GPA at the end of the 3rd year	2.283	(2.564)	-0.615	(1.183)
Credits at the end of the 1st year	9.144	(8.885)	-12.44	(9.054)
Credits at the end of the 2 nd year	-5.034	(16.777)	-55.60***	(17.324)
Credits at the end of the 3rd year	20.75	(23.417)	-37.22*	(22.058)
Timely graduation	0.426*	(0.233)		

Note: Levels of significance are reported as follows: *** p<0.01; ** p<0.05; *p<0.1.

The results concerning the number of credits obtained show a clear difference between Bachelor's and single-cycle degree. In the first case, we do not find significant results and the estimates are in any case positive in the first and third year, with a slight decrease in the second year. On the other side, the effect of the Grant 5B for single-cycle students is negative and statistically significant during the second and third year.

The final outcome considered is timely graduation that is calculable only for the Bachelor's degrees. In this case it is apparent that the impact of the Grant 5B is positive and statistically significant: Grant 5B recipients are, on average, 4.6% more likely to complete their university course than their counterparts who are not entitled to the financial aid. This result is consistent with the, albeit not statistically significant, findings about the number of credits (+20.75 credits at the end of the third year).

Conclusions

The findings show that the measure does not enhance enrolment probability and academic performances, but it has remarkable effects on the choice of the university location. The overall picture is tangled by the diverging results emerged for what concern the last two cohorts. The different findings suggest that the persistence of the economic crisis affects students' decision regarding university participation.

Keywords: Financial aid, inequality, impact evaluation

Uni.Co: a diachronic model for analyzing the transition paths to employment of graduated students

Pietro Lucisano – Andrea Marco De Luca – Giulio Lucentini

The present work aims to present the potentials and limitations of the UNI.CO model for analyzing the transition paths to employment of graduates. The UNI.CO model, activated experimentally in 2013 following an agreement between Sapienza University of Rome and the Ministry of Labor and Social Policies, develops a research model based on the integration of administrative data on the careers of graduates and the Mandatory Communications of the Ministry of Labor and Social Policies.

The availability of administrative data, i.e., data recorded during administrative procedures, and the possibility of integrating large databases allow for a different approach to the analysis of transition paths to employment. This is because it allows observing the phenomena under examination across entire populations, not just samples, relying on administratively verified data rather than declarations, and providing a diachronic view that allows verifying phenomena longitudinally rather than punctually. In particular, Mandatory Communications represent a significant resource for this type of analysis. Since 2008, the Ministry of Labor and Social Policies has mandated employers (both private and public) to electronically communicate the activation, variation, and termination of all subordinate and para-subordinate employment contracts.

The development of the model, which has extended to universities in Lazio and Tuscany, allows for a diachronic reading of data related to subordinate and para-subordinate employment contracts obtained by graduates, enabling the examination of their effective duration and contractual characteristics (duration, contract type, professional qualification, economic sector, coherence with the academic degree, territorial location).

In this contribution, the potentials of the system will be illustrated by referring to the two observatories established for the Tuscany and Lazio regions. The observatories allow each university to have a comprehensive view of the paths of their own graduates, while at the regional level, only an overview for the systems is possible. The reference populations consist of graduates between 2008 and 2018 for the Lazio region (262,327) and graduates between 2008 and 2020 for the Tuscany region (248,686), who during the observation period obtained a total of 2,110,142 employment contracts.

The purpose of the Observatories is, on one hand, to provide the Regions with a tool to analyze the transition paths to employment of graduates, thus understanding how to support orientation processes both entering and exiting university studies through appropriate active policies, in relation, for example, to the type of university path undertaken and the demand for subordinate and para-subordinate employment in the region and the national territory. On the other hand, the Observatories provide universities with a service to increase their awareness of the outcomes of their educational offerings, useful for supporting educational

planning, orientation, teaching, and third mission activities, also in relation to the demand for employment in the regional territory and the country.

The possibility of following transition paths over time allows, for some cohorts, not only information on post-graduate employment contracts but also any contracts before and during the study period, providing an indicative overview of the job demand directed at graduates in the two territories and examining the outcomes of graduates from each degree program.

The investigation has allowed the development of some synthetic indicators such as the employment rate, defined as the ratio between days worked and days observed from the date of graduation until the end of the observation, or the coherence index, which allows examining the vertical mismatch between the academic degree and the days of work in jobs consistent according to ISTAT indicators. These indicators can then be compared with the careers of graduates.

Another aspect concerns the types of contracts and job stability. Typical contracts, indefinite contracts, represent a minimal percentage of the activated contracts and often extinguish during our observation. However, the model allows working towards an optimal definition of transition and verifying to what extent different university paths favor an optimal transition to employment for graduates.

The model allows, respecting privacy regulations, the reconstruction of the transition paths of each graduate, thus integrating the reading of statistical indicators with useful insights to understand some aspects of the transition to employment.

In this contribution, alongside the illustration of the model, the main constructs and elaborated variables, some focuses will be presented on problematic aspects of the transition paths of STEM graduates in relation to the employment rate and coherence index and in relation to the origin and locations of the contracts activated by graduates from southern regions. The aim of these focuses is to highlight some of the contradictions in the job demand, such as forcing graduates to choose between job security and quality or to move en masse from their regions of residence to regions with higher job demand, resulting in actual educational poverty in southern regions.

The goal is to draw attention to the potentials and informative content of this approach and its ongoing developments, both for better educational planning and for the implementation of more effective active labor policies in territories, but also to highlight the limitations of our country's job demand, which seems unable to valorize the potential of graduates, even from highly demanded degree programs, and subjects a high number of graduates to a transition marked by long periods of uncertainty and precariousness.

Keywords: University to Work Transition, Employment Rate, Vertical educational mismatch, Types of Employment Contracts, Valorization of University Studies

SESSION 5. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAs)

METHODS AND RESULTS

ORGANIZER: UNIVERSITY OF KU LEUVEN - UNIVERSITY OF BATH

COORDINATOR: JAN-PHILIPP WAGNER

19TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 2 – RESEARCH 10}

Italian students' opinions on political and social issues in the IEA ICCS 2022 survey

Ornella Papa – Michele Cardone – Rita Marzoli

Introduction

The International Civic and Citizenship Education Study (ICCS), sponsored by the International Association for the Evaluation of Educational Achievement (IEA), collects data on citizenship education and civic knowledge of students around thirteen and a half years old, as well as their opinions and attitudes towards political and social issues (Schulz et al., 2023). IEA ICCS 2022 is the third edition of the survey (the previous ones took place in 2009 and 2016) and involved 22 education systems worldwide, with an increased percentage of European countries compared to previous editions (INVALSI, 2023). Indeed, the European Union, considering active citizenship a key competence, recommends educational pathways to promote it and monitoring the levels of competence achieved (Council of the European Union, 2015; 2018). For countries like Italy that have participated in all ICCS editions, the national data are also comparable with those of previous editions allowing us to detect the evolution of the results achieved over time. Although the adolescents' political socialization are deemed crucial for their future political participation, there are not many in-depth studies on this topic at national and international level (Myoung & Liou, 2022).

Subject, aim and research hypothesis

This study analyses Italian student data from the IEA ICCS 2022 survey, comparing them with international data and data from previous ICCS editions. The data analysed are mainly those collected by the student questionnaire, which investigates students' characteristics, opinions and attitudes; the results of the cognitive test, more specifically related to civic and citizenship knowledges, are also considered. The sample consists of 4347 students attending 8th grade from 226 schools (of which 154 administered the test in digital format and 72 in paper format). The focus is on students' opinions and attitudes in social and political areas, specifically: satisfaction with the current political system, trust in institutions and media, interest in political issues and dialogue with friends and parents about them, whether they intend to vote when they come of age. The aim of this study is to deepen the point of view of the adolescent on relevant political issues, it is hypothesized that their opinions reflect the political disaffection behind the growing abstention at the polls, especially of young people which reaches up to 50% among non-graduates (Schäfer, 2021)

Methodology

The data are analysed from a comparative perspective, both with other participating countries and with previous editions, especially the most recent one in 2016. Comparative analyses are also carried out between students grouped by socio-economic background and gender. Students' political opinions and attitudes are deepened and the strength of their association with civic knowledge is checked. Data analysis is carried out using SPSS.

Findings

The principal results show that Italian students are more dissatisfied and distrustful of the political system and institutions, especially Parliament, than students in other countries. In fact, only 44% of Italian students, compared to 55% of the international average, believe that their country's political system works well; only 36% of Italian students, compared to 44% of the international average, believe that parliamentarians represent the interests of young people well. Nevertheless, interest in political and social issues is more prevalent among Italian students (39%) than in students from other countries (30%), and discussions with parents about these issues occur more frequently than internationally. Compared to the 2016 edition, trust in the institutions and media has decreased internationally but especially in Italy, where trust both in Parliament and media has fallen by 10 percentage points. The propensity to vote also decreases in Italy, compared to the 2016 edition of ICCS, but remains frequent among students with high levels of civic and citizenship knowledges, especially among girls.; in fact, in Italy as internationally, girls achieve better results than boys in the cognitive test. Altogether Italian student have an average cognitive test score (523)

significantly higher than the international one (508) and not significantly different from the average score of the two previous editions of ICCS.

Conclusions

Although Italian students have an average score for civic knowledge higher than the international one and not significantly different from the ones of the ICCS 2016 and 2009 edition, they are less satisfied with the current political system, institutions, and media. However, they are interested in political and social issues and discuss them with their parents more than students from other countries. Similarly, in the ICCS 2016 edition, Italian students reported talking to their parents about sociopolitical issues more frequently than in the ICCS 2009 edition, although their trust in institutions had meanwhile declined (Malak-Minkiewicz & Torney-Purta, 2021). These findings reveal an attitude that is increasingly critical of politics but far from apathetic or disinterested. Deepening what students think about political issues is important first because it gives a voice to the new generations, who generally do not have one (Tuorto, 2018), but also because it can offer insights into understanding the reasons why young people have shifted, over the decades, from representing the group with the highest voter turnout to the one with the lowest turnout (Schäfer, 2021). Of course, this negative trend has an impact on the overall percentage of abstention, which reached 36% during the 2022 general elections, placing Italy at the bottom of Western Europe in terms of voter turnout (Emanuele, V. & Marino, B., 2022); a complex phenomenon that cannot be underestimated, but it has to be faced and countered by planning wide-ranging strategies and interventions.

Keywords: Students' opinions and attitudes, Political and social issues, Citizenship education, IEA ICCS 2022, Comparative research, International survey.

Comparative Analysis of Civic Education Perceptions: Insights from Mexican and Chilean Teachers

Citlalli Sanchez-Alvarez

Introduction

The interest in civic and citizenship matters has surged among researchers, government agencies, and nations over the past few decades. This rise is primarily attributed to growing concerns about the noticeable decline in levels of civic engagement among young people worldwide (Brooks, 2009; Hustinx, Lucas, Meijs, Handy, Femida & Cnaan, 2011). These forms of participation encompass what are traditionally considered formal or conventional civic activities, such as voting or actively engaging in political affairs (e.g., working for a political party or advocating for specific candidates). Numerous studies have provided evidence indicating a growing distrust among young people towards the decision-making processes of governments, institutions, political parties, and politicians, all of which significantly impact them as citizens (Ostrander, 2004; Torney-Purta & Amadeo, 2004).

While many studies confirm this trend, research also demonstrates that young people are increasingly turning to new forms of civic engagement within their schools and communities. They are opting for arenas of participation that hold personal significance, such as volunteering or engaging in activities focused on pressing issues like climate change and human rights. This shift reveals new avenues for civic education.

In this context, civic education emerges as a crucial gateway for educators to equip young people with the knowledge, skills, and attitudes necessary to understand, engage with, and participate in the democratic construction of society. Education for global citizenship and democracy presents a formidable yet essential challenge. It requires a focus on promoting new forms of education that foster critical thinking, democratic discourse, active engagement through proposals and actions for societal improvement, and values like solidarity, collaboration, and empathy.

Preparing students for their future roles as citizens requires that institutions and schools recognize the multitude of factors influencing the development of the knowledge and skills required to cultivate positive civic attitudes and engagement. Education delivered in schools plays a pivotal role in this process. When schools provide conducive environments for students to discuss various topics of interest in a safe and respectful atmosphere, students are more likely to develop heightened political interest and participation levels. Additionally, studies have shown that operating schools in a democratic manner, involving teachers, students, and parents in decision-making processes, positively impacts student engagement and fosters the development of positive civic attitudes (Amadeo, Torney-Purta, Lehmann, Husfeldt & Nokolova, 2002; Buk-

Berge 2006; Cox, Bascopé, Castillo, Miranda & Bonhomme, 2014; Schulz, Ainley, Fraillon, Kerr, & Losito, 2010; Schulz, Ainley, Fraillon, Losito & Agrusti, 2016; Treviño, Béjares, Villalobos & Naranjo, 2017). Teachers play a central role in citizenship education (Conde, 1997; Reyes, Campos, Osandón & Muñoz, 2013; Huertas, 2023), often tasked with instilling valued attitudes in students that society deems essential for future generations. They should actively encourage student participation in school activities to nurture their skills and civic competencies.

This study aimed to explore and compare teachers' perceptions of civic and citizenship education in Mexico and Chile. These two countries were selected due to a significant difference in their civic education curriculum. For many decades, Mexico has mandated a specific and compulsory subject matter in civic education, while Chile, did not until 2020 when an explicit Citizenship Education subject matter was incorporated as a product of a curriculum reform. The results reveal disparities in (a) the type of community activities teachers report doing with their students during the school year, (b) their professional development regarding civic and citizenship education topics, and (c) their perception of the things that are needed to improve civic and citizenship education in their schools.

Method

Research design. This empirical comparative study adopts a descriptive approach to compare the perceptions of Mexican and Chilean teachers regarding citizenship and citizenship education within their respective schools. It delves into their views on classroom climate and interactions, the types of activities conducted related to civic and citizenship education, their opinions on professional preparation and training, and the instructional practices employed in these two educational systems.

Instruments. The study utilizes data from the 2016 cycle of the International Civic and Citizenship Education Study (ICCS), which examines how young people across different countries and regions are being prepared for their roles as citizens. The ICCS collects information on students' knowledge of civic and citizenship topics, along with their attitudes, perceptions, and related activities. The teacher questionnaire, chosen for this study, comprises 22 Likert-type and multiple-choice items. These items cover teachers' demographic information, attitudes, and perceptions regarding civic and citizenship education, their observations of student behavior and attitudes, pedagogical methods employed, classroom climate, teaching practices related to citizenship education, and professional development experiences.

Participants. Data from the teacher questionnaires of two Latin American countries, Mexico and Chile, were analyzed for this research study. Teachers from both private and public secondary institutions were surveyed. The sample size for the study comprised 3370 teachers, with 1918 from Mexico and 1452 from Chile. The ages of the participants ranged from 23 to 63, with mean ages of 41.76 for Mexico and 42.05 for Chile.

Data analyses. Variables derived from the teacher questionnaire were analyzed using descriptive statistics with SPSS and categorized into two groups. The first category includes six of the ten existing scales that capture information derived from teachers' responses to questions concerning students' activities in the community, classroom climate, classroom activities, and professional development. These scales were developed by the statistical team of the International Association for the Evaluation of Educational Achievement (IEA) and are detailed in the corresponding technical report. The second category encompasses a set of teacher variables selected based on their contextual relevance and nature.

1) Scales

- a. Teachers' perceptions of student activities in the community (α 0.76 Chile, 0.69 Mexico).
- b. Teachers' perceptions of classroom climate (α 0.90 Chile, 0.84 Mexico).
- c. Teachers' reports on civic-related activities in class (α 0.79 Chile, 0.67 Mexico).
- d. Teachers' confidence to teach citizenship and civic education (CCE) topics (α 0.92 Chile, 0.88 Mexico).
- e. Teachers' professional development activities addressing topics related to CCE (α 0.93 Chile, 0.89 Mexico).
- f. Teachers' professional development activities related to teaching methods for CCE (α 0.89 Chile, 0.85 Mexico).

2) Additional variables

Two additional variables (Q14, Q22) were analyzed to explore teachers' perceptions of the main purpose of CCE at school and their opinions on aspects needing improvement in CCE, such as the quantity or quality of materials, in-service training, cooperation among teachers, and other relevant factors.

Results and Key Findings. This study aimed to empirically examine the perceptions of Mexican and Chilean teachers regarding citizenship education within their schools, including their views on classroom climate, instructional practices, and professional preparation. Analysis of data from the 2016 ICCS cycle yielded several key findings:

a) Teachers in both countries share similar perspectives on the primary aims of civic and citizenship education, emphasizing the promotion of knowledge about citizens' rights and responsibilities, the development of conflict resolution skills, fostering respect for the environment, and nurturing critical and independent thinking among students.

b) Similarities were observed between both countries regarding the frequency and nature of activities related to civic and citizenship education conducted during lessons. These activities include discussing current issues, utilizing textbooks, engaging in small group work on various topics, and taking notes during lectures.

c) Another commonality among teachers in both countries is their perceived level of preparedness in teaching various civic and citizenship-related topics, particularly those concerning citizens' rights and responsibilities, gender equality, conflict resolution, human rights, critical thinking, and electoral processes.

d) However, differences emerged in teachers' perceptions and practices related to community activities conducted with students, professional development opportunities in civic and citizenship education topics, and their opinions on improvements needed within their schools' civic education programs.

These findings offer valuable insights into how schools can contribute to civic education and the development of civic competencies among students, underscoring the importance of understanding teachers' perspectives and addressing areas for improvement within civic education curricula and practices.

Recommendations. Future research should continue to compare civic education perceptions across nations, examining both similarities and differences. Understanding these variations can inform education planning and policy development at both national and international levels. Comparing teachers' views on citizenship across nations in research could include analyses of data from various national or regional studies, as well as in longitudinal studies with data derived from the different cycles of the ICCS project. This could offer insights to researchers, educational authorities, and policymakers regarding the varied understandings of citizenship among teachers worldwide.

Keywords: Civic and Citizenship Education, Teachers' Perceptions, International Civic and Citizenship Education Study (ICCS), Comparative International Research

Students' future education pathways and their occupational aspirations

Beti Lameva - Zaneta Chonteva

North Macedonia has one of the highest proportions of students failing to demonstrate basic proficiency (Level 2) in all three domains of science, mathematics and reading among PISA (Programme for International Student Assessment)-participating countries (52.2% in PISA 2018 testing; 57.4% in PISA 2022). Young citizens of North Macedonia continue to leave education with among the lowest learning outcomes in Europe. On the other hand, between 2018 and 2023, North Macedonia's public spending on education as a percentage of GDP fell from 2.80% to 2.72%. The share of total government expenditure allocated to education also declined. While poverty rates have fallen in recent decades, low educational performance is limiting the employment and life opportunities of many individuals and impeding national development.

The activity of youth in the labor market of North Macedonia is relatively low (46.7% in the first three quarters of 2022), either compared to the adults or their peers from the EU countries. One in four people over 15 are unemployed, compared to less than one in ten across OECD countries. Low activity of youth illustrates generally low employment probabilities in the country, and the difficulty of school-to-work transition, that can be attributed to (i) unwillingness of employers to bear the costs of on-the-job training of inexperienced youth (ii) skills mismatch between employer's needs and skills produced by the education system, as well as (iii) the increasing tendency of youth to stay longer in formal education.

Students' academic performance on the PISA 2022 testing is not aligned with their expectations for further education and career. They hold ambitious expectations of future education, 72% of students expect to complete tertiary degree (34% expect to finish doctoral studies, ISCED 8). Students (83%) reported that they have a clear idea of their future job, and they expect to work in high-skill occupations, such as software developers, medical doctors, managing directors and chief executives.

Object and research hypothesis

This research aims to determine if there is a significant relationship between Students' Education and Career Expectations and their Academic achievement in Math, Science and Reading, Index of Economic, Social, and Cultural Status, Gender, and Language of instruction. Data obtained from the research is further discussed with students in focus group discussions. Recommendations from the research will be shared with the state representatives responsible for the reforms in secondary education.

Method(s) employed

The research used questions that refer to the students' future educational pathways and preparation, and their occupational aspirations. They are integral part of the questionnaire intended for students, as part of the research procedure prescribed by the protocol of the international research study PISA 2022 (OECD, 2023). To assess the relationship between Students' Education and Career Expectations with their Academic achievement in Math, Science and Reading, Index of Economic, Social, and Cultural Status (ESCS), Gender (Female and Male), and Language of instruction (Macedonian and Albanian) Chi-Square Test of Independence was performed in SPSS Statistics 21. In the analysis we applied weights on student-level data.

Variables in the research:

Students' Education Expectations, the variable is coded in the following categories: category 1- ISCED level 2; category 2- ISCED level 3, ISCED level 4 and ISCED level 5, category 3- ISCED level 6, ISCED level 7 and ISCED level 8.

Students' Career Expectations, the variable is coded in 14 categories based on the Classification of Occupations in VET schools in North Macedonia.

Academic Achievement in Math, scores are coded in 6 levels of proficiency (OECD, 2023).

Academic Achievement in Science, scores are coded in 6 levels of proficiency (OECD, 2023).

Academic Achievement in Reading, scores are coded in 6 levels of proficiency (OECD, 2023).

Index of Economic, Social, and Cultural Status (ESCS) is a composite score derived from three indicators related to family background: parents' highest education, in years, parents' highest occupational status and home possessions. Based on the score 3 groups are created: group 1-students with lowest ESCS; group 2-students with middle ESCS; group 3- students with the highest ESCS (OECD, 2023).

Gender, two categories are defined: Female and Male.

Language of instruction, two categories are defined: Macedonian and Albanian.

Results achieved

Students' Education Expectations: Results of Chi-Square Test of Independence

There is a significant relationship between the Students' Education Expectations and Achievement in Math, $\chi^2(10, N=11141) = 408.32, p = .001$.

There is a significant relationship between the Students' Education Expectations and Achievement in Science, $\chi^2(8, N=11140) = 441.57, p = .001$.

There is a significant relationship between the Students' Education Expectations and Achievement in Reading, $\chi^2(6, N=11141) = 421.42, p = .001$.

There is a significant relationship between the Students' Education Expectations and Achievement in Index of Economic, Social, and Cultural Status (ESCS), $\chi^2(4, N=11114) = 615.75, p = .001$.

There is a significant relationship between the Students' Education Expectations and Achievement in Gender (Female and Male), $\chi^2(2, N=11142) = 141.04, p = .001$.

There is a significant relationship between the Students' Education Expectations and Achievement in Language of instruction (Macedonian and Albanian), $\chi^2(2, N=11140) = 40.28, p = .001$.

Students' Career Expectations: Results of Chi-Square Test of Independence

There is a significant relationship between the Students' Career Expectations and Achievement in Math, $\chi^2(65, N=7875) = 798.34, p = .001$.

There is a significant relationship between the Students' Career Expectations and Achievement in Science, $\chi^2(52, N=7875) = 575.37, p = .001$.

There is a significant relationship between the Students' Career Expectations and Achievement in Reading, $\chi^2(39, N=7879) = 453.54, p = .001$.

There is a significant relationship between the Students' Career Expectations and Achievement in Index of Economic, Social, and Cultural Status (ESCS), $\chi^2(26, N=7858) = 449.62, p = .001$.

There is a significant relationship between the Students' Career Expectations and Achievement in Gender (Female and Male), $\chi^2(13, N=7875) = 1588.10, p = .001$.

There is a significant relationship between the Students' Career Expectations and Achievement in Language of instruction (Macedonian and Albanian), $\chi^2(13, N=7875) = 348.96, p = .001$.

Keywords: PISA, occupational aspirations, education

Results of INVALSI standardized assessment as a tool for professional development of mathematics teachers

Camilla Spagnolo – Maria Chiara Cibien – Marta Saccoletto – Carlotta Soldano

The European Parliament and the Council launched a recommendation on lifelong learning in 2006 in which Member States are invited to promote key competences for all citizens in their lifelong learning strategies.

There is serious worrying data coming out from international research focusing on the Mathematical competences. OECD data show that students who are fifteen years old and have reached Level 1 acquire limited mathematical knowledge on their own, which they are only able to apply in some familiar situations. Referring to INVALSI data of 2022, a similar situation occurs in Italy. Already at the end of the last century, the Curriculum and Evaluation Standards for School Mathematics (National Council of Teachers of Mathematics, 1989) and the Professional Standards for Teaching Mathematics (NCTM, 1991) have promoted a vision for teaching and learning that has growing support from the mathematics education community. There is increasing recognition that without carefully planned professional development programs, the chance of widespread implementation of this vision of mathematics is small. Support is essential for practicing teachers who express an interest in teaching in a way described in the Standards, with all that such teaching implies in terms of content, pedagogy, and assessment.

The topic of teacher professional development is increasingly gaining ground in the field of educational research, also in mathematics education.

The project stems from a firm and shared conviction that standardised assessments, used in a formative perspective, can become a potentially revolutionary element for the educational and didactic function of school systems. In a democratic school perspective (equity and quality together), we embraced the bloomian belief (Bloom, 1972) that all students can learn basic skills in each topic and we trust in the link between the development of students' skills and the teachers' professional development paths. Literature highlight that schools show an evident fragility with respect to developing an effective ability to promote good mathematical competences. It is well known that teachers have difficulties in adopting valid practices for individualized teaching techniques in order to be able to pursue quality and equity for the results of every and each one of the student population, but also in finding strategies to improve the learning for the fundamentals. The possibility for a teacher to use in his classroom the insights on mathematical learnings provided by the standardized tests - through their theoretical frameworks, the released items with the interpretation of macro-phenomena observed, the valuation of their results - in a formative way is an important opportunity to renew the praxis in teaching and assessment.

The use of standardized surveys to substantiate the formative assessment practices in the classroom is of top-notch priority and of great importance opportunity. The impact of standardized assessment is traditionally a top-down impact: the results of the surveys affect the public opinion; this moves the policy makers acting on the structure of the system at different levels (curriculum, teacher training, recruitment). Only in a final phase, the surveys and their results influence teachers' local action. The innovative idea of our project is to provide instruments for a bottom-up approach in which the theoretical frameworks/test/results are used immediately as a tool by teachers, to accomplish formative assessments and individualized teaching (from the perspective of quality and equity of students' achievement).

To make this approach virtuous and fruitful some conditions should be fulfilled: first, there must be a link between assessments and curriculum (Mons, 2009; Meckes & Carrasco, 2006, Martignone, Ferretti, & Lemmo, ICMI 24). Second, this link should be perceived by the teachers, since assessment is a key component of teachers' identity (Hannula et al., 2016; Ferretti et al., 2021).

In Italy, the first of these conditions is met: as a matter of fact, the national system of standardized assessment (INVALSI) framework and tests are aligned with National Guidelines goals. There is a need for tools to develop educational activities to achieve the second condition as well.

To reach this goal we use articulate our methodology as a multiphase research design. In particular, we start identifying the training needs of pre-service and in-service mathematics teachers. This empirical investigation phase will consist of the design and the administration of six questionnaires, three aimed at primary and secondary in-service teachers, and three aimed at primary and secondary pre-service teachers, each consisting of three sections based on three macro-categories of variables. The focuses of investigation of all questionnaires will be on the field of mathematical knowledge, the affective sphere and the beliefs and knowledge of formative and standardised assessment. In particular, the three macro-variables on which the questionnaires will be constructed are:

- mathematical knowledge and mathematics teaching (including how teachers interpret the Invalsi items and their results)
- affective factors, Di Martino and Zan TMA Model (perception of one's own competence in mathematics and mathematics teaching, view of mathematics and mathematics teaching, emotions towards mathematics and mathematics teaching).
- beliefs and knowledge about the purposes and possible uses of standardised assessment in mathematics, about formative assessment; we will use a battery concerning beliefs about the formative function of assessment, developed within the FAMT&L research Project (Ferretti, Michael-Chrysanthou, & Vannini, 2018) and error awareness (interpretation of the causes of errors highlighted by macro-phenomena in standardised assessment).

Each questionnaire will be declined according to the group of teachers involved (in-service and pre-service) and the school level in which they are teaching/will be teaching. The mathematics INVALSI tasks in the questionnaires will be items that are symptoms of national educational macro-phenomena and will focus on problem solving skills.

The questionnaire will be administered in telematic mode to the in-service teachers collaborating with the researchers of the different project units and to all the in-service teachers registered in the GESTINV database. Furthermore, it will be administered to future teachers of mathematics at secondary school (to the students of the courses of mathematics education given by the Department of Mathematics and Informatics of the University of Ferrara, by the Department of Mathematics of the Sapienza University of Rome and by the Department of Mathematics of the University of Eastern Piedmont) and to future teachers of primary school (to the students of the mathematics education courses of the Faculty of Primary Education of the Free University of Bolzano, of the Department of Primary Education of the University of Turin).

In this paper, we will present first analysis of the questionnaire data, with a focus on three macro-variables compared across school grades. This is a part of a larger project "Mathematics standardized assessment as tool for teachers' professional development", that won the PRIN 2022, funded by the European Union and the Italian Ministry of Education and University.

Keywords: international large-scale assessment; teachers' professional development; mathematics education

SESSION 4. INVALSI DATA FOR IMPACT ASSESSMENT: LIGHTS, SHADOWS, AND POTENTIAL DEVELOPMENTS

ORGANIZER: UNIVERSITY OF MILAN BICOCCA - CATHOLIC UNIVERSITY OF MILAN - UNIVERSITY OF BOLOGNA & FBK-IRVAPP

COORDINATOR: LORIS VERGOLINI

19TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 3 – RESEARCH 11}

Informing secondary school students about financial aid: does it affect undergraduate enrollment rates?

Federica Laudisa – Samuele Poy

In Italy, student grants are the primary means for removing economic barriers to enable low-income students to access higher education. However, inadequate information among the stakeholders causes a relevant flaw in implementation (Laudisa, 2022). Previous research by IRES Piemonte revealed that a large proportion of first-time tertiary education entrants in Piedmont universities fail to apply for student grants, despite meeting eligibility requirements. We hypothesize that this is due to a lack of awareness of the financial aid opportunities available. The problem of not applying for grants despite eligibility appears to be a concern not only in our country but also in other parts of the world. (King, 2004; Kofoed, 2017; Herber & Kalinowski, 2016).

This study employed a randomized controlled trial to examine the impact of an informational session about financial aid on university enrollment and grant application rates. In other words, we investigated whether raising awareness of financial aid opportunities influences students' choices to pursue higher education and, if they do enroll, whether it increases the number of grant applications submitted.

We started by assessing students' current understanding of the grant scheme by interviewing 6,500 randomly selected high school seniors in Piedmont. Our data confirmed a disparity between the students' professed familiarity with the grant and their actual knowledge. Despite 95% claiming to know about the grant, only 6% could accurately identify the eligibility requirements, and a mere 9% had heard of the grant organizers (EDISU Piemonte). A significant correlation between grant knowledge and student characteristics, such as gender, grade point average, school type, family background, and participation in guidance activities, was also observed.

Female students, students with parents who have tertiary education, those with higher grade point averages, and students following academic tracks—considered the most advantaged—are more informed and significantly more likely to aspire to further education. This finding is not particularly surprising (INVALSI, 2023; AlmaDiploma, 2023). However, this study also shows that being more actively engaged in guidance activities, all other factors mentioned being equal, significantly increases awareness of the benefits provided by EDISU, leading to a greater desire to enroll in university. This suggests that guidance activities may play a crucial role in mitigating the influence of family background in socioeconomically disadvantaged contexts. Such activities likely represent the only means of fulfilling this function.

Participating students were randomized into treatment and control groups. The treatment group attended a 30-minute talk providing them with detailed information about the grant, including eligibility requirements, application procedures, and award amounts. Conversely, the control group received no such information. Follow-up interviews were conducted after the students' graduation from high school to determine how many had enrolled in university and/or applied for a grant. By comparing the rates of the treatment and control groups, we were able to evaluate the impact of having attended the informational session.

This randomized controlled trial contributes to a relatively sparse strand of literature. At the national level, Abbiati et al. (2017) conducted a preliminary study where upper secondary students attended three separate informational sessions covering various aspects such as costs, employment returns, drop-out rates by degree course, and scholarships. The aim was to ascertain whether these guidance activities could influence enrollment rates and alleviate the impact of social background. Although no impact was observed, at least regarding enrollment rates, the international literature shows different but not incontrovertible outcomes. Overall, the literature tends to suggest that the impact is most likely when informational sessions

are coupled with personalized counseling activities (Herbaut & Geven, 2020). The only exception is a study by Loyalka et al. (2013), which shows that a single informational meeting increased university enrolment. Similarly, our study revealed that a one-time intervention on financial aid had a positive and significant effect on enrollment rates, estimated at +2.8 percentage points. This intervention was particularly successful in increasing higher education participation among students whose survey responses had indicated indecision about continuing their studies due to economic considerations (+8.3 p.p). This means that grants can make the difference in choosing whether to enroll in university, particularly when economic concerns contribute to indecision.

However, the effect was null for vocational track students, suggesting that a more intensive information program may have been necessary since their training is specifically geared towards entering the labor market. This hypothesis finds support in research by Martini et al. (2021), who observed the impact of the “Percorsi” program on the probability of university enrollment among vocational school students.

Informing students about grant opportunities not only influences college enrollment but also significantly boosts the number of grant applications submitted (+7.4 p.p). This impact is even more pronounced (+15.9 p.p.) among students hailing from lower socioeconomic family backgrounds, who are the main beneficiaries of financial aid programs.

In conclusion, our findings underscore the importance of providing information about financial aid to students as a strategic approach to increasing access to higher education in Italy.

Keywords: Information, Financial aid, Randomized controlled trial, School-University transition

Case Study of Policy Linking National Assessments to Measure SDG 4

Daniel Shephard – David Rutkowski

Less than six years remain for achieving Sustainable Development Goal 4 (SDG 4) and the current assessment is that improvement on learning indicators for SDG 4 are disappointing (UNDESA, 2023). There are further concerns that gaps in data systems are leaving behind the learners that are furthest from achieving the promises of SDG 4 (Delprato & Shephard, 2024). One proposed pathway forward is to expand the availability and comparability of national assessment data in the hopes that doing so will contribute to faster progress on learning outcomes in reading and mathematics (UIS, 2023; UNDESA, 2023; World Bank, 2021).

To support this effort, the UNESCO Institute for Statistics (UIS) and the Global Alliance to Monitor Learning (GAML)—along with other bilateral and multilateral actors—have design and piloted a methodology known as “Policy Linking” (UIS, 2023). The methodology involves an adapted Angoff method to “link” existing national assessments with the Global Proficiency Benchmark cut-off points for reading and mathematics. While the methodology has been piloted in 2019 and 2021/2022, it is not yet used for reporting on SDG 4.1.1 and there is no published, peer-reviewed literature that documents how the guidance for policy linking is carried out in practice.

In this paper, we present the results of an in-depth case study of one country in sub-Saharan Africa that has piloted Policy Linking. We document how the panel of experts is identified, how they carry out the policy linking guidelines in practice, their perceptions of the policy making methodology, and the ways in which national education actors, especially those at the Ministry of Education, make use of policy linking.

We used snowball sampling until we had reached saturation among individuals who had been involved in the policy linking. With each respondent, we carried out a series of semi-structured interviews to document their experience of the methodology and their perspectives on its validity and use. We then created verbatim transcriptions of each interview and analyzed them in four steps. First we completed descriptive memos of one to two pages after each interview documenting the most important points raised from our perspective and according to extant literature on the use of the Angoff method and standard setting (Fechter & Yoon; Hurtz & Auerbach, 2003; Tannenbaum & Wylie, 2008). Second, once all interviews were completed and transcribed, we conducted first cycle eclectic coding to identify emerging themes. Third, we conducted second cycle categorical coding in which we merged the first cycle codes into broader categories and recoded the transcripts according to those broader categories. Finally, we conducted iterative code-weaving to create a short summary of the main findings, to discuss those among the research team, and then to return to the data and code-weaving until we converged on the primary findings. Each of these steps was done

independently by the two authors with results compared after each step and adjustments made—any substantive deviations in coding or interpretation are noted in the paper.

The findings illustrate several important dynamics with implications for Comparative Education scholars and for measuring SDG 4's learning targets using policy linking. First, transnational education consultants play a role in disseminating the policy linking methodology. Second, national education actors adapt the methodology in ways that suite their context and their goals but might influence the validity of any cross-national comparisons. Finally, there are important divergences in terms of the perceive utility and validity of policy linking at the national and international level. We discuss the theoretical implications of these findings for both the policy borrowing and lending literature and the critical cultural political economy literature (Robertson & Dale, 2015; Steiner-Khamsi, 2012). Practically, it raises questions and make proposal for a path forward for policy linking if it is to be used for reporting on SDG 4 and, more importantly, if it is going to be used to inform policy improvements that assist us in achieving equitable education for all. In future work, we will explore how variations in national assessments and policy linking can influence error and inference.

Keywords: policy linking, sdg 4, assessment, comparative education

Learning losses during pandemic times: what role for school closures choices?

Silvia Duranti – Giada Garbini – Leonardo Grilli

Introduction.

Italy was the first Western country hit by Covid-19: the virus first broke out in the North and then spread throughout the peninsula. The Government shut down schools of all levels in the North in late February 2020 and in the rest of the country in early March 2020. Virtual classroom lessons then replaced in-person classroom lessons until the end of school year 2019-20. Following the sharp drop in Covid-19 cases and deaths in the summer, schools reopened their gates in September 2020. However, the pandemic soon regained strength in some areas of the country, and the Government then introduced some restrictions in a bid to prevent another national lockdown. Contrary to the first wave, in school year 2020-21 the Government allowed school closures to be differentiated according to grade and region. This new management model was intended to safeguard the needs of the youngest pupils, and to make the severity of the measures proportionate to the Covid-19 aggressiveness within a Region. Indeed, Regions had different reactions to the pandemic emergency, first of all because they were differently hit, but also because they relied on more or less robust health services and school systems. This resulted in differences both in the policies concerning the length of school closures and in the reactions from the school system. In some areas long closures are to be ascribed to strong COVID waves, but in other areas they are just the political response to the fear that the health service and the school system could not be able to cope with the emergency.

The duration of school closures has raised many concerns about the effects in terms of educational attainments and a growing international literature has investigated their impact (Blaskò et al., 2021; Donnelly and Patrinos, 2021; Hammerstein et al., 2021; Storey and Zhang, 2021; Thorn and Vincent-Lancrin, 2021; Moscoviz and Evans, 2022; Patrinos et al., 2022; Zierer, 2021; Betthausen et al., 2023; Di Pietro, 2023). Also the literature on the Italian case is constantly growing (Contini et al.; 2022, 2023; Borgonovi and Ferrara, 2023; Bazoli et al., 2022; Carlana et al.; 2023; Aparicio Fenoll; 2022; Bertolotti et al., 2023), but, as far as we know, only Battisti and Di Maggio (2023) have analyzed the learning losses in relation to the heterogeneity of the closing days between territorial administrative units, although for a single Region, Sicily.

Objective. The aim of the present paper is to contribute to the empirical analysis of the educational consequences of Covid-19 on school-age students in Italy. Italy is certainly one of the most interesting cases among European countries. Firstly, Italy had the longest school closures in Europe, being the country hardest hit by the pandemic during the first and subsequent waves. Secondly, Italy was largely unprepared for remote schooling because of its technological backwardness (ISTAT, 2020).

In particular, we want to exploit the regional variability in the number of school closures to assess the impact of distance learning on students outcomes at the end of each school cycle.

Data. The analysis relies on a unique dataset which was made up by collecting the number of school closure days imposed at regional and national level for primary and lower secondary schools. This information have been merged to the INVALSI databases containing average test scores in Italian and Mathematics at the municipal level from s.y 2015/2016 to s.y. 2022/2023, for a total of seven years. The limits of the use of these data consists in the fact that Invalsi only releases information on municipalities having at least three schools per each grade. This means that to perform our analysis we had to work on a subsample of municipalities for which it was possible to observe data for a sufficient number of year.

Methodology. Two different methodological approach have been tested. The first estimation rests on a fixed effects panel model with a second-degree polynomial of time, while the second one relies on a difference-in-differences counterfactual evaluation in which the control group is the closest generation before the Covid-19 pandemic.

Discussion. Because of the aforementioned Invalsi rule on the release of data by municipality, we needed to restrict the analysis to a subsample of municipalities, which represent a different percentage of the universe of municipalities according to the methodology used and of the grade analyzed. In the case of difference-in-difference analysis, the percentage of municipalities covered by the analysis is 31% for primary education and 14% for lower secondary education. Although these percentages might seem very low, the municipalities included in the analysis account for 77% of total primary education students and 59% of total lower secondary education students. In the panel analysis we tested two different types of restrictions on the number of municipalities. In the first case, we included in the final sample those municipalities being present at least five year and covered 41% of municipalities and 78% of students in primary education and 15% of municipalities and 59% of students in lower secondary education. In the second case we included in the final sample those municipalities being always present in the seven-year period of observation and covered 34% of municipalities and 75% of students in primary education and 14% of municipalities and 58% of students in lower secondary education. However, the coverage of municipalities is very heterogeneous between regions, ranging from 11% of Molise to 56% of Toscana.

Results obtained with the two methods indicate a reduction in learning outcomes following the Covid-related closures in lower secondary education and in primary for the mathematics test only. The extent of the reduction is greater in lower secondary schools than in primary schools and in the math test than in the Italian test. The estimate carried out using as a variable the number of days of school closures imposed at regional level allows us to highlight a differentiated impact between regions which ranges, in the case of math score at the end of lower secondary education, between -2% and -5% (did estimation). The impact of regional choices regarding school closures on students' performance in the Invalsi tests is therefore confirmed.

Keywords: Covid-19, school closures, distance learning, learning loss, INVALSI

The calculation of added value as a measure of social performance

Vincenzo Nunzio Scalcione

Introduction

Quality evaluation research involves the analysis of systemic models aimed at planning programs and objectives, research skills that have the intention of optimizing procedures according to the expected outcome. Quality control thus shifts from ascertaining the conformity of the results with standard objectives to the exploratory verification of the conditions that ensure coherence and alignment of the training interventions with the evaluation process.

The variability is produced by factors extrinsic to the institution and to the educational action, i.e. "exogenous" variables, which influence the results achieved.

Object

The relationship between a student's performance and his previous performance is measured through a simple linear regression, which offers the possibility of "establishing to what extent the independent variable influences the dependent variable or, more technically, detecting the intensity of the effect exerted by the independent variable on the dependent variable" (Silva, 2016, p. 39).

This relationship is then expressed through the linear equation of the line: " $\hat{Y}_i = \alpha + \beta X_i + \varepsilon_i$ " where: \hat{Y}_i =expected exit return from student i;

α =intercept or constant (value assumed by Y when X is equal to zero); β =regression coefficient (average change in Y associated with a unit change in X);

X_i =incoming performance of student i;

ϵ_i =residual of the regression (difference between Y_i and \hat{Y}_i , i.e. between observed return and expected return).

The effect of schools and classes through added value will therefore be expressed as the part of unexplained variance, i.e. the sum of the regression residuals aggregated at school and class level. $\hat{Y}_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$

Attention is also paid to the aggregation of the residuals of the regression ϵ_i . These can be aggregated at class or school level and thus average indicators of added or subtracted value can be obtained which depict their effectiveness or, on the contrary, their ineffectiveness.

When the measurement takes place at class level, the effectiveness of the teachers will therefore be determined. When this occurs to pay attention to the progress of students, the measurement will take place at class level, on the contrary, when reporting of results needs to be implemented, it will proceed at institute level (Corsini, 2010).

Data

The INVALSI tests are administered based on goals for the development of competence and on learning objectives established and foreseen by the regulations for students.

Considering the documents mentioned above, the tests include a subdivision relating to the Italian, Mathematics and English subjects, as well as aspects relating to citizenship.

It is possible to divide the outcome of a standardized test into different parts, namely:

- A part dependent on external conditions on which the school cannot intervene directly (general social context, social origin of the students, previous preparation of the students, etc.), i.e. dependent on the so-called exogenous factors;

- A part determined by the school effect, i.e. by the set of actions implemented by the school to promote learning (didactic-methodological choices, organization of the school, etc.).

INVALSI has taken steps to return the schools' results, in terms of added value, since 2016. The added value model that INVALSI uses to measure the scholastic effectiveness of Italian schools is based on two levels, namely students and school.

The equation becomes the following: $Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + u_{0j} + \epsilon_{ij}$

From the 2018 surveys, different models were created based on the subject areas (Italian and Mathematics), as well as on the last two classes of grades 5 and 8. For upper secondary schools, separate models are used for subject areas and for the different types of educational institutions.

The dependent variable for each grade level is the Rasch scale score. The variables referring to the two levels (school and students) explain the variability of the dependent variable. In Table 1, next to the categorical variables, the class considered in the analyzes is indicated (student gender). The level 2 variables are obtained thanks to the principle of aggregation from the level 1 variables. In upper secondary schools, the percentage of female students in the institution is inserted since there is often an imbalance in the student population. The level 1 variables are estimated on the average of a group and therefore of the schools, unlike those of level 2. The lack of the variable that refers to the results of the pupils in the previous grade is noted and this often leads to not being able to carry out the estimate of added value.

In order to measure the school effect, the results of the tests administered in primary schools throughout the country were taken into consideration, making a comparison of the data for the years between 2019 and 2023.

The results of the second primary class in Italian show the comparison of 2019 with 2023 (see figure 1). From their reading it is possible to deduce a lowering of the average level at national level since there are 5.9 points of difference between 2023 and 2019. Even if the administration after the pandemic, in 2021, had not brought to light negative results, among There is a decline in 2021 and 2022 which remains stable between 2022 and 2023 and which could be caused by the long-term effect of the pandemic on learning. In figure 2 it is also possible to observe a movement of students towards lower and medium-low bands and it can be deduced that in the second primary class there are approximately 69% of students who reach the lowest band. Compared to 2022 there are 4 percentage points less.

Figure 1 - Results of class II primary school students - Italian. Average score

Italian grade 2

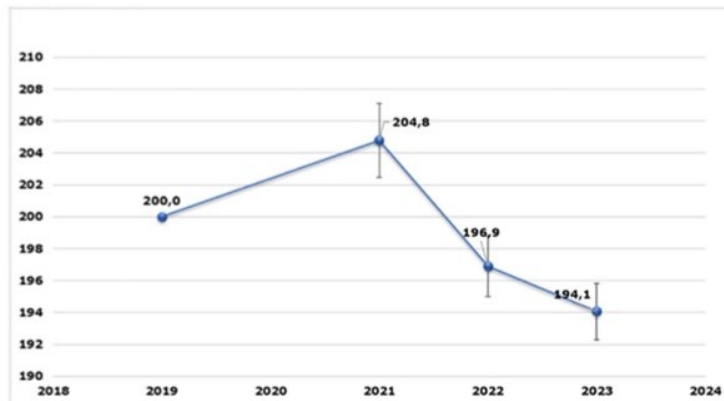
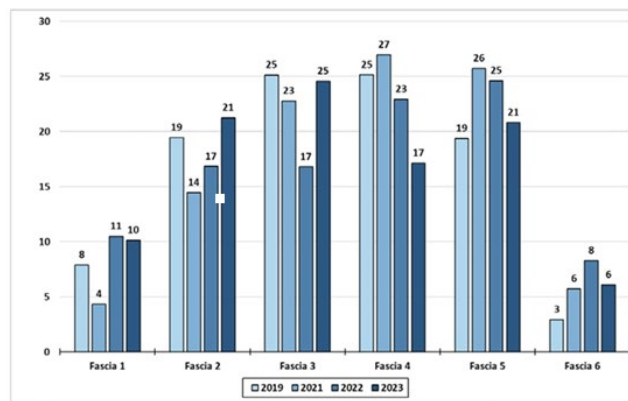


Figure 2 - Students by result range in Italian primary school class II. Percentage distribution



As regards the results of the second primary class in Mathematics, there is a national average drop of 9.9 points between 2019 and 2023, greater than that of the Italian tests and it is hypothesized that the cause of this drop is due to the pandemic (see figure 3). A shift of students towards lower bands is observed and approximately 64% reach the lowest band. There are 7 points less than in 2022 (see figure 4).

Figure 3 - Results of students in primary school class II in Mathematics. Average score

mathematics grade 2

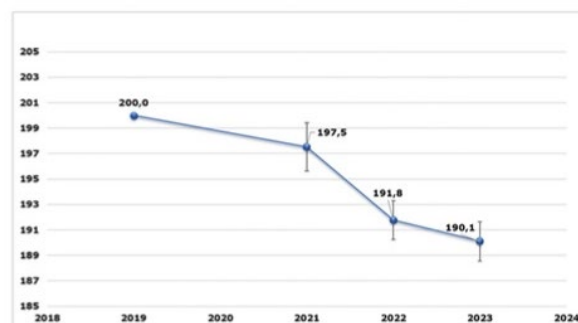
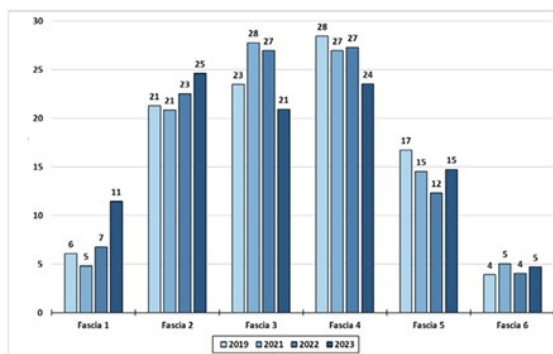


Figure 4 - Students by result range in Italian in primary school class II. Percentage distribution



The results of the Italian test of the students of the fifth primary class represent an important element for having awareness of the exiting skills possessed by the students. Comparing the results of 2023 with those of 2019, there is a decrease of 4.2 points. The decline continues until 2022 and records a further decrease in 2023 (see figure 5). The reason refers to a post-pandemic situation. Furthermore, a shift of students towards lower brackets can be observed, around 74% with 6.4 percentage points less compared to 2022 (see figure 6).

Figure 5 - Results of students in primary school V in Italian. Average score

Italian grade 2

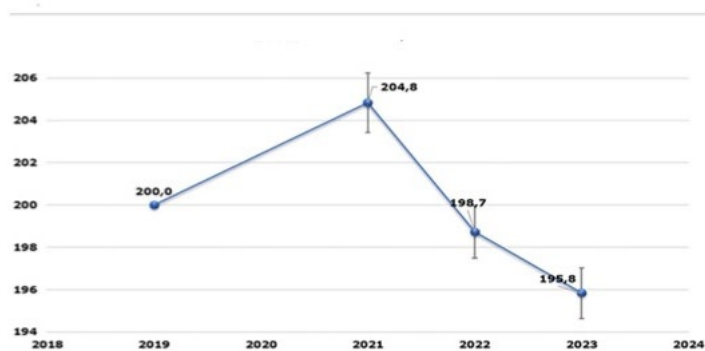
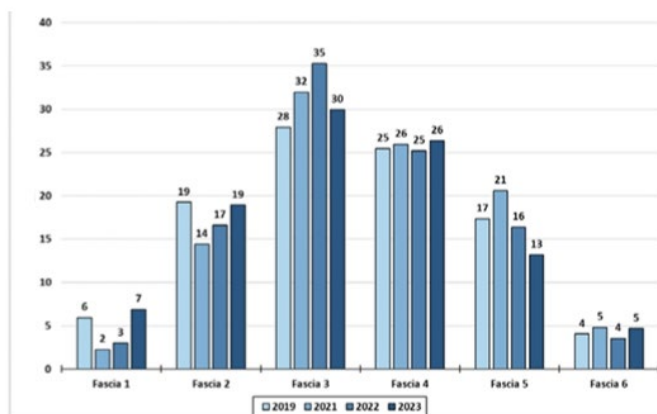


Figure 6 - Students by result range in Italian in primary school V. Percentage distribution



The comparison from 2019 to 2023 of the results in Mathematics of the fifth primary class records a significant drop of 9.8 points (see figure 7). It follows that approximately 63% of students reach the low and medium-low bands with 3 points less than in 2022 (see figure 8).

Figure 7- Results of students in primary school 5th grade in Mathematics. Average score³
 mathematics grade 5

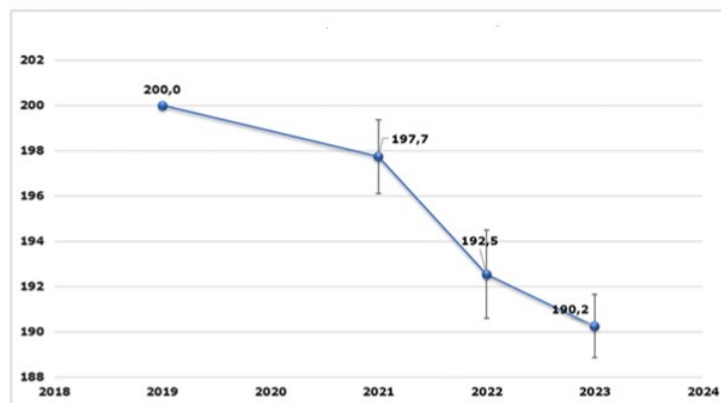
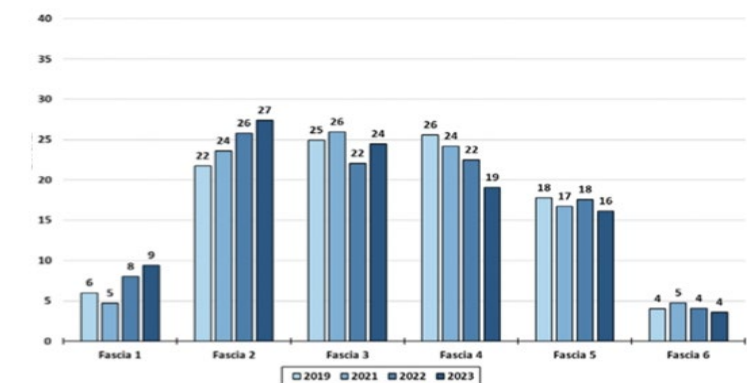


Figure 8 - Students by achievement band in Mathematics in primary school V. Percentage distribution



Methodology

To evaluate the effectiveness of a school it therefore becomes essential to define whether and to what extent it has succeeded in ensuring that its pupils have been able to learn more than what was achieved on average by 'comparable' pupils, in the same period of time, in other schools (Martini, 2020).

The measurement of added value occurs through complex technical-statistical procedures that take into consideration aggregations of students (class or school). A fundamental role is attributed to performance expectations as schools add value when a pupil's performance goes beyond expected expectations.

We must consider how the relationship between a student's performance and his previous performance is measured through a simple linear regression, which offers the possibility of establishing the influence of the effect of the independent variable on the dependent variable, measuring the intensity that this exerts .

Results

The results offered by the added value calculation vary depending on the model adopted. Even though it uses a numerical language, it can in fact present itself in different ways based on the database used and the type of analysis. However, it is necessary to consider how "engaging in the measurement of educational effectiveness creates added value" (Minello, 2012, p.216)

In the scientific community the debate is therefore of importance, focused on the models to be adopted, but above all on their educational use, and the related fields of application (Capperucci, 2017), having to consider variables linked to measurement methods and tools.

Keywords: quality, control, evaluation, added value, school effectiveness

SESSION 10. MATHEMATICAL EXPERIENCES TO CONTRAST LEARNING LOSS AND SCHOOL DROP-OUT IN DISADVANTAGED CONTEXTS

ORGANIZER: UNIVERSITY OF SALERNO - POLYTECHNIC UNIVERSITY OF MILANO

COORDINATOR: GIOVANNINA ALBANO

19TH OCTOBER: 9.00 A.M. - 10.30 A.M. {ROOM 4 – RESEARCH 12}

Inclusive education, learning loss and implicit dropout: how to interpret a relationship

Elisabetta Robotti – Alessandra Boscolo

The pandemic has taught us how the phenomenon of learning loss can be linked to implicit school dropout. Although the phenomenon of implicit dropout in high schools is decreasing, as evidenced by data collected through the INVALSI tests, in the INVALSI National Report 2022 it can be observed that implicit dropout is rapidly increasing among students from less advantaged families. Furthermore, these data are accompanied by INVALSI results in Mathematics at the primary school level much lower than those recorded in 2019 and 2021, in line with those of 2022 (for second-grade classes), while in fifth-grade primary classes, the results of 2023 are lower than those of previous years, including 2022.

These results convince us of the importance of teaching Mathematics in primary school to address Mathematics teaching in higher grade levels and demonstrate how the context of students continues to play a very important role in learning levels.

Preventing and countering dropout has become a priority, and an effective way to do so lies in the configuration of inclusive teaching approaches, already at the primary school level, capable of accommodating, supporting, and guiding the learning process of a plurality of students.

International context also suggests such an approach: the UN Agenda 2030 outlines 17 goals to which UN member countries have committed to promote sustainable development. Among these goals, Goal 4 focuses on the right to education, calling on nations to ensure quality, inclusive, and equitable education at all levels and for all. It is evident how this goal, aiming to provide all students with equal learning opportunities, also addresses school dropout (implicit or explicit), considered an indicator of educational inequality and lack of equity (Benvenuto, 2016). Our intervention aims to illustrate a possible implementation of mathematics teaching activities that reflect the characteristics of equity, quality, and inclusivity promoted by the UN Agenda.

What characterizes inclusive education, at least from the perspective of mathematics education? To answer this question, we will present the experience of a research-action group, the "Mathematics Teaching Laboratory" at the University of Genoa (<https://sites.google.com/view/labddm/home?authuser=0>), illustrating the potential of collaborative design (Morselli & Robotti, 2023), based on the principles of Universal Design of Learning (CAST, 2011), and on some results of research in mathematics education as well as in neuroscience and cognitive sciences. In particular, Embodied cognition suggests how the body and mind are closely interconnected in the development of mathematical thinking, highlighting how many mathematical concepts are developed and contained in motor, visual, and sensory memories (Núñez, R., & Lakoff, G., 2005). Furthermore, research in Mathematics Education has shown that the processing of mathematical meanings occurs through multimodal activities: doing, touching, moving, seeing are essential components of mathematical thinking from the early stages of conceptual development to more advanced learning processes (Arzarello et al., 2007; Nemirovsky, 2003; Radford, 2006). Based on these theoretical premises, we will illustrate a path on fractions, designed for the primary school level, which reflects our idea of inclusive education by adhering to the principles of Universal Design of Learning (UDL).

The hypothesis on which we develop our research work is to counteract the phenomenon of implicit dropout and learning loss, by designing teaching activities that allow all students to solidly build the mathematical meanings required by the preparation goals of their study path in the "Indicazioni Nazionali".

Keywords: Universal Design of Learning, Multimodality, School Inclusion, Embodied Cognition, Implicit Dropout

Contrasting the development of learning loss in mathematics: a longitudinal challenge

Federica Ferretti

By learning loss we can refer to the gap between the level of knowledge, skills or competencies assessed before and after the interruption of school activities. Until 2020, learning loss basically coincided with “summer learning loss” since, prior to the COVID-19 pandemic, the only interruption large enough to allow an estimate of the effects was summer break. Beginning in 2020, the focus shifted to trying to measure the impact of COVID-19 lockdown on students' learning. Studies conducted suggest that the “loss” does not affect the entire student population equally, but is related to the learning opportunities provided by the extracurricular setting. In detail, research shows that the socioeconomic component strongly influences achievement and that the “loss” is more pronounced in secondary school than in primary school. Another finding that emerged is that this learning loss had more pronounced effects for Mathematics than for Reading skills.

In order to estimate the learning loss caused by the pandemic, INVALSI compared tests administered in 2021 with those administered in 2019. Although INVALSI surveys are not longitudinal, the results collected are consistent with the relevant literature. For instance, the “losses” measured by INVALSI tend to be concentrated in secondary schools. In line with the literature and what the data revealed, it can be particularly relevant to investigate the characteristics of the INVALSI mathematics items that highlighted difficulties that occur longitudinally, and seem to persist as the years of schooling progress. As the literature has shown, analyzing and interpreting with solid theoretical lenses macro-phenomena that have emerged at the national level can provide valuable information on the learning processes of Italian students. In particular, analyzing the characteristics of INVALSI items that highlight difficulties that persist at the longitudinal level can provide valuable information on directions to pursue as early as primary school to try to avoid, at least with regard to aspects purely concerning the mathematics, the development of learning loss.

Keywords: extracurricular setting, learning loss, statistical analysis, INVALSI surveys

Learning loss in mathematics: the case of evolution in the period 2020-2024 in two comprehensive schools

Monica Alberti – Lucia Cirina – Maria Polo

Our contribution explores the phenomenon of learning loss in mathematics within socio-economic disadvantaged communities, particularly in the context of school dropouts (Lai-Polo, 1999; Ricci, 2019; Battini, 2023). We have investigated the evolutionary path (maybe “developmental trajectory”?) of mathematical skills among primary school students from 2020 to 2024. The challenges posed by the lockdown significantly altered the cognitive, affective-relational, and psychological dynamics of school life for both teachers and students. Drawing from firsthand experiences in classrooms, we have analyzed the aftermath of these disruptions and the subsequent efforts towards recovery: the “obligatory” “mandatory” selection of teaching activities and the repercussions on the construction of mathematical knowledge and attitudes towards the subject (Di Martino & Zan, 2011).

To investigate the cognitive aspects and types of mathematical activities, we analyze the results of the INVALSI tests administered to second and fifth-grade primary school students who experienced the pandemic phase. These data are placed in relation with various intervention measures implemented by the

Institutes upon the return to in person classes, such as: Improvement Plan, Inclusion Plan, PTOF, Risk Assessment Document, Social Reporting.

Specifically, we analyze data from classes within two comprehensive institutes in Southern Sardinia: the "Generale Luigi Mezzacapo" comprehensive institute of Senorbì and the "Monsignor Saba" comprehensive institute of Elmas. These Institutes, situated in distinct geographical areas, encompass diverse socio-economic contexts and student populations. The former is located in a town in the hinterland of Cagliari while the latter currently includes the territorial communities of four municipalities: Senorbì (including the hamlets of Arixì and Sisini), Barrali, San Basilio and Suelli. Overall, these two institutes represent and reflect the two most characteristic conditions of the scholastic realities in the Sardinia region. This dichotomy creates distinct educational challenges and opportunities, and affect educational resources, student opportunities, and community support systems, thereby shaping students' educational experiences in each context. During the pandemic and its afterwards, the equilibrium of the territorial system was disrupted impacting the integrated education system. Signs of youth distress have escalated, evidenced by rising rate of addictions, dropouts and school disengagement, all contributing to social exclusion. An inadequately targeted use of digital tools has emerged, hindering many students, particularly the most vulnerable ones, from fully participating in school activities and social interactions.

For all these reasons, the inclusive school model has encountered a crisis. The improvement measures implemented by the two Institutes, such as teacher training in evaluation, self-evaluation and inclusion, align with the critical issues identified in research in Mathematics Education (Bakker, A., Cai, J., & Zenger, L., 2021). System initiatives, including projects aimed at expanding educational offer through intensified extracurricular activities, have progressively reduced the percentage of students in the lower performance range, while increasing those in the middle range and solidifying the high achievement group. It is noted that this improvement coincides with the resumption of in-person teaching activities, despite significant heterogeneity observed within the same class and across different classes.

Keywords: primary school, mathematics education, learning loss, INVALSI test

SESSION 5. INTERNATIONAL LARGE-SCALE ASSESSMENTS (ILSAs) METHODS AND RESULTS

ORGANIZER: UNIVERSITY OF KU LEUVEN - UNIVERSITY OF BATH

COORDINATOR: MARIA-MAGDALENA ISAC

19TH OCTOBER: 2.00 P.M. - 4.00 P.M. {ROOM 2 – RESEARCH 13}

Civics Education During a Lockdown: COVID-19 Containment Policies and Grade 8 Student Civic Outcomes

Alec Kennedy – Ana María Mejía-Rodríguez – Rolf Strietholt

Introduction

In early 2020, nations around the globe responded to the rapidly developing COVID-19 pandemic by implementing emergency containment measures to slow the spread of the virus. Examples of such policies included school or business closures, public event cancellations, restrictions on gatherings or travel, and

stay-at-home orders. Decisions to implement such measures highlighted the benefits of such policies in battling against a global pandemic, yet it was clear that they were weighed against the potential costs to daily life and the well-being of citizens. Most sectors felt the impact of such measures, especially the education sector where decisions to close schools for in-person learning affected over 1.5 billion students worldwide (OECD, 2021).

A body of research has emerged to study, and mainly quantify, the impact of the pandemic on student learning. Two recent meta-analyses have synthesized the findings of this research, revealing a significant decline in academic achievement following the outbreak of the pandemic, with effect sizes equivalent to approximately one-third to half a year's worth of learning. (Betthäuser et al., 2023; Di Pietro, 2023). The extent of learning declines has been found to be associated with the duration of school closures with larger declines being found in areas where schools were closed for longer periods of time (Patrinos, 2023; Jakubowski et al., 2023; Kennedy & Strietholt, 2023; Jakubowski et al., 2024). It should be noted that research has focused on the academic impact of the pandemic on core subjects (i.e., mathematics, science, and reading) while impacts in other domains, such as civics education, remain understudied. Considering how one core task of educational systems is to develop and prepare students to become engaged and active members of society it is important to examine how the disruption in education impacted students' civic outcomes. Moreover, in recent years, the importance of civic education has been emphasized for multiple reasons including increasing socioeconomic inequalities, growing polarization, violent extremism, rising levels of misinformation, declining levels of civic engagement, and various threats to democracy (Joris & Agirdag, 2019; Savage et al., 2023; Winthrop, 2020). The COVID-19 pandemic might have exacerbated many of those issues and some worry about its lasting impact on democratic attitudes (e.g., Amat et al., 2020; Bol et al., 2021). It should also be noted that civic education has many outcomes that extend beyond cognitive knowledge and skills, as providing students with knowledge must be linked with developing civic attitudes and behaviors that allow them then to participate actively and successfully as citizens in society (Blaskó et al., 2019; Schulz et al., 2023; Winthrop, 2020).

Research objectives and hypotheses

In the present study, we examine the relationship between the stringency and/or duration of COVID-19-related containment measures and Grade 8 student civic outcomes. We examine not only students' civic knowledge but also several affective-behavioral outcomes. Our goal is to better understand the impacts of the COVID-19 pandemic and restriction measures on students. Specifically, we seek to measure the relationship between the duration or stringency of COVID-19 lockdown measures and: (1) trends in student civic knowledge and (2) changes in student perspectives on civic engagement.

Data

We combine data from multiple sources. We use data from over 90,000 students from 15 countries or educational systems that participated in the International Civic and Citizenship Education Study (ICCS). ICCS provides internationally comparable data of civic knowledge and other civic and citizenship measures across several cycles (2009, 2016, 2022). ICCS collects data from eighth-grade students across several countries, including Italy, making it a valuable resource to examine the relationship between national variation in containment policies and student civic outcomes. The outcomes used in this study are: civic knowledge, trust in civic institutions, expected political and electoral participation, and expected participation in legal or illegal protest activities.

We also use data from the Oxford COVID-19 Government Response Tracker (OxCGRT) which collects daily information on policy measures that were implemented in response to the COVID-19 crisis including measures of school closures and other containment policies. From the data, we calculate the duration of school closures during the time between January 2020 and the end of data collection in ICCS. In addition, we use an OxCGRT derived measure of policy stringency which captures the extent to which several COVID-19 containment measures are in place. The stringency index ranges from 0 (no measures in place) to 100 (all measures in place aimed at the full population) and we average it over the time period between January 2020 and the end of data collection in ICCS.

Method

To estimate the relationship between COVID-19 containment policies and student civic outcomes we estimate the following model on a dataset containing both ICCS 2016 and ICCS 2022 data (before and after the onset of the pandemic):

$$Y_{ict} = \alpha + \beta \text{COVIDPolicy}_c * D_{2022} + \mu_{c,t} + \tau D_{2022} + \gamma X_{ict} + \varepsilon_{ict}$$

where Y_{ict} is civic outcome for student i in country c at time t . It is important to note that the data are not longitudinal and come from a repeated cross-section. COVIDPolicy_c is our measure of COVID-19

containment policies (i.e., duration of school closures or average stringency index). D_{2022} is a dummy variable that is equal to one when the data come from the year 2022. μ_c and $\tau_{D_{2022}}$ represent country and time fixed effects. X_{ict} is a vector of demographic variables at the student-level (student background, COVID-19 restriction measures, and participation rates). In estimating the model, we utilize sampling weights, plausible value methodology, and jackknife repeated replication techniques as advised by ICCS (Schulz et al, 2018).

Results and Findings

Table 1 summarizes the estimates of β (standardized for comparison) from the model presented above. Each row indicates the measure of COVIDPolicy (i.e., average stringency index, school closure duration) on several civic outcomes shown in the columns.

The first column shows that declines in average civic knowledge (CIVKNO) were larger in countries with higher average stringency and longer school closures. A one standard deviation increase in the stringency or duration of COVID-19 measures are associated with 0.054SD and 0.066SD declines in civic knowledge, respectively.

Examining the impacts of COVID-19 policies and the other student civic outcomes, we find several interesting patterns. Longer school closures and more stringent COVID-19 measures were associated with lower trust in civic institutions, increased intentions to protest, and greater expectations to participate in elections or politics.

While research has documented the negative effects of the pandemic on core subjects such as mathematics, science, and reading, it appears that learning in civic education was also affected during this time. In Italy, which had the highest average lockdown stringency of the countries examined, the findings from this study have profound implications. Evidence presented provides valuable insights into the associations between COVID-19 lockdown measures and early secondary student civic outcomes, highly relevant for understanding impacts of the pandemic on the future citizens of the world.

Table 1: Estimated standardized coefficients on school closure duration and average stringency index on civic outcomes

	CIVKNO	INTRUST	LEGACT	ILLACT	ELECPART	POLPART
Average Stringency	-0.054*** (0.012)	-0.044*** (0.011)	0.111*** (0.011)	0.045*** (0.013)	0.054*** (0.011)	0.051*** (0.013)
School Closures	-0.066*** (0.013)	-0.061*** (0.014)	0.071*** (0.009)	0.038** (0.014)	0.038*** (0.009)	0.028* (0.014)
N student	99,587	97,527	96,407	95,898	95,937	95,899
N country	15	15	15	15	15	15

CIVKNO = civic knowledge; INTRUST = trust in civic institutions; LEGACT = expected participation in legal protest activities; ILLACT = expected participation in illegal protest activities; POLPART = expected political participation; ELECPART = expected electoral participation.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Note: Coefficients are from separate models and have been standardized for comparison. All models control for student background (student gender, home socioeconomic status, immigrant status), COVID-19 severity measures (case and death rates), study participation and exclusion rates, and country and time fixed effects.

Keywords: COVID-19, Civics Education, Secondary Education, School closures, Lockdown stringency, International Comparisons

Can Civic Education compensate social inequalities in civic outcomes? A literature review focused on Large-Scale Assessment Studies

Daniel Miranda

Introduction. The proper functioning of democracy rests on its citizens equally participating in the decision-making processes, being interested in issues of social and political relevance, adhering to democratic principles, and developing attitudes consistent with life in a democratic society (Hoskins & Janmaat, 2019;

Rouse et al., 2022; Schlozman et al., 2012). However, the literature shows that multiple gaps threaten the egalitarian ideal, such as socioeconomic, gender or ethnic gaps, affecting the access to political voice, interest in politics, knowledge, adherence to democratic values or principles, etc. The family and the school are the most relevant institutions in the political formation process during childhood and adolescence. On the one hand, the family transmits beliefs, attitudes, knowledge and skills pertinent to exercising citizenship, directly socializing the descendants. Additionally, the family transmits the advantages of its social position to the next generation, which translates into a transmission of political inequality. Inequalities in the social structure mark the experience of family political socialization. The family experience in different social positions also translates into other ways of experiencing politics, since within the family, the conditions of increasing stratification typical of modern societies are manifested (Peña, 2015). On the other hand, education is considered the critical space for promoting democracy and training citizens to function, compensating for inequalities of origin. The school system generates a series of activities, defined both formally and informally, that have the potential to form those aspects relevant to the exercise of citizenship. Contrary to family political socialization, school has the potential to generate a common experience for all those who attend the school system. In this context, the compensation hypothesis is raised. This suggests that the school experience, similar for all who attend, can level, moderate, mitigate or compensate for the gaps caused by social origin (Campbell, 2019). In that sense, political inequalities are expected to decrease or disappear due to access to the citizenship training experience at school. Thus, for example, two students from very different social positions will develop similar levels of knowledge about the school system or propensity to participate politically because they have experienced similar political socialization processes at school. The idea of compensation for inequalities through the school experience is not only proposed for inequalities in the political sphere but, for example, for inequality in academic achievement. Various authors have argued that this ideal is not met. Still, on the contrary, the school tends to operate as an agent that reproduces inequalities of origin instead of correcting them (Bernstein, Coleman, Bourdieu). In this sense, it is essential to ask whether the school experience of citizenship training compensates, accelerates or maintains inequalities of origin. According to Hoskins & Janmaat (2019), the school can fail in its equalized role in two ways: by generating unequal teaching and learning for students through different opportunities according to their socioeconomic conditions (unequal provision) and by accelerating differences. That is to say, socioeconomic conditions of origin benefit learning opportunities for those students who are advantaged by their conditions of origin (for example, socioeconomic, gender or immigration status). Considering the previous context, the aim of this paper is to evaluate what the literature that use Large-Scale Assessment data says about citizenship education's role over origin gaps, such as compensation, acceleration or maintenance. Data and methods. To address the objective, several challenges arise. First, it is necessary to distinguish what is being compensated. Given that when talking about citizenship, multiple aspects must be considered to exercise the role of citizen, for example, different types of participation, democratic attitudes, adherence to values or principles or relevant knowledge. Second, the characterization of social origin can take multiple definitions, such as social class, occupational status, cultural capital, immigration status or gender. Given that each of these forms has the potential to relate differently depending on the citizenship outcome in question, it is relevant to distinguish them. A third challenge refers to the characterization of the school experience. Characterizing citizenship education in school can take multiple forms. For example, citizen training activities are defined as curricular subjects, distributed in various activities or depicted as a type of climate. Considering these challenges, this review distinguishes between multiple aspects when describing the evidence for the compensation hypothesis. The review considers 26 papers that specifically address the compensation hypothesis between 1997 and 2023. The main inclusion criterion is that the paper analyzes the compensation hypothesis of inequalities of social origin concerning civic outcomes based on some school activity aimed at civic education, such as courses, volunteering or council elections. Main results and conclusions. The main result is that several papers indicate that the evidence supports the compensation hypothesis. They show that school civic education can take multiple forms and compensate for gaps of social origin for civic outcomes. It is interesting to note that, although there is a good amount of evidence that supports the idea that the school compensates for some inequalities of origin, the evidence also shows that this effect does not occur for all types of school civic education activities nor for all civic education activities or outcomes considered. This in some sense indicates that this hypothesis requires further research to unravel under what conditions it occurs and for which variables it is most generative. In that sense, it seems to me that there are several challenges that must be considered. Firstly, the problem of measuring the variables involved in the analysis. As could be observed for both the social origin variables, the civic outcome variables and the

characterization of activities aimed at citizen education present a multiplicity of approaches. Different concepts are observed, different ways of measuring them, and also some relevant absences. This challenge combines 3 big discussions. The analyzed papers show a wide range of dependent variables on which focus has been placed. Each subdimension, knowledge, attitudes or participation; They have their own conceptual challenges when evaluating gaps of social origin and the compensatory role of the school. On the one hand, inequalities of origin are not homogeneous depending on the variable in question. The relationship between socioeconomic origin and civic knowledge has ample evidence that shows its association (Collado et al., 2015; Schulz, 2019), while for attitudinal variables this association is less robust (Miranda et al., 2018) or even reverse, as in the case of illegal activism (Hoskins & Janmaat, 2019; Sánchez-Barría & Miranda, 2022). This fact challenges the selection of which school variables can or cannot compensate for origin gaps. In that sense, the justification of the school involved is more dependent on the type of gap that is evaluated.

Keywords: Citizenship education, youth citizenship inequalities, mitigation/acceleration effect of civic education

An Intersectional Approach to Students' Endorsement of Gender Equality using Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy - MAIHDA

Natalia Lopez-Hornickel – Diego Carrasco – Andres Sandoval-Hernández

Students' endorsement of gender equality is a crucial citizenship outcome. To protect democracy, future citizens must know about civic structures and participate, but also be tolerant to respond thoughtfully in front of discrimination, such as racism or sexism (Kennedy, 2019). Furthermore, the adherence to traditional gender attitudes can affect boys and girls (Kågesten et al., 2016) because these can be associated with stereotypes about gender, implying evaluative representations such as “boys are violent”, creating evaluative responses (Zosuls, Martin, Ruble, Miller, Gaertner, England & Hill, 2011), risky behaviours associated with hegemonic masculinities (Varela et al., 2022), and affect the wellbeing and academic achievement of girls at schools (Nuamah, 2019).

Like most educational problems, attitudes toward gender equality are associated with inequality. Access to information and formation about how to endorse more equitable statements are subject to family and origin factors, as well as the access they have in school to these topics. Socioeconomic backgrounds, such as having more educated mothers, can imply a higher adherence to these attitudes (Dotti Sani & Quaranta, 2017), being a girl or a woman (Dotti Sani & Quaranta, 2017; Yu & Lee, 2013), and have an immigrant status (Isac et al., 2021).

Nevertheless, previous research on citizenship is not often concerned with this affective behavioural aspect (Isac, Sandoval-Hernández & Miranda, 2018). Most of the studies in this area have focused instead on more “cognitive” dimensions, such as civic knowledge or more related to the “engagement”, such as political or civic participation (Amnå, 2012; Ekman & Amnå, 2012; Miranda et al., 2020). Furthermore, earlier pieces of research that have studied attitudes toward gender roles have tried to identify how factors such as gender, socioeconomic status and education are related together to influence this outcome but using traditional analysis techniques such as interactions (Keller et al., 2022).

An intersectional perspective would help us address the inequality problem since the simultaneous use of social categories can reveal disadvantages in the adhesion to attitudes toward gender equality. Notably, a quantitative approach can help us understand the problem's dimensionality. Not simply using interaction, which can carry parsimony problems (Keller et al., 2022), but applying the Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA) approach to solve the statistical challenges that this kind of analysis represents (Evans, 2019; Keller et al., n.d.; Merlo, 2018). This approach uses intersectional social strata as second-level units. Therefore, their membership implies exposure to certain advantages and disadvantages (Evans, 2019).

The present work has two objectives. First, frame the adhesion to attitudes toward gender equality as an inequality problem that could be understood under the approach of intersectionality. Therefore, understanding how the intersection of different social categories can make the lack of adhesion to these attitudes more pronounced. Second, to replicate the MAIHDA framework from the achievement educational field (Keller et al., n.d.) to the area of citizenship attitudes. Thus, to understand how social categories, such

as gender, socioeconomic status, and immigrant status intersect and are related to attitudes toward gender equality.

The data used is the International Civic and Citizenship Education Study (ICCS) 2016 (Schulz, Carstens, Losito & Fraillon, 2018), conducted by the IEA. For this research, it was selected Chile from Latin America data, and Italy from Europe data.

Regarding the preliminary findings, for the case of Chile, the results indicate that social groups constituted by immigrant and native women with a higher socioeconomic background (parents with university education and more literacy resources at home) tend to present the strongest support for attitudes toward equality. About previous literature or research, this is concordant with evidence about women and their support for more egalitarian behaviour. Being a woman continues to be a critical axis for the support of gender equality. There is still an organisation of social life and status beliefs, as indicated by Ridgeway (2011), which affects women more.

The hypothesis is that there is an interest-based approach according to which women or girls would benefit from gender equity, and their support comes from that fact (Bolzendahl & Myers, 2004). Also, an exposure-based hypothesis according to which women would be more exposed to situations that make explicit the presence or absence of gender equality (Bolzendahl & Myers, 2004).

In the case of socioeconomic background, it is confirmed once again that parents' education influences the adhesion to democratic values among children (Miranda et al., 2018). Previous antecedents differentiate between mother and father education (Dotti Sani & Quaranta, 2017; Farré & Vella, 2013). Nevertheless, in this study, it was used a measure that represents the highest level of education of either parent. Also, it was included the variable of literacy resources as an indicator of socioeconomic background at home, whose higher level was also associated with a higher adhesion to equality between men and women.

Finally, regarding immigrant status, groups conformed by immigrant women were the most supportive. Particularly immigrant women with a high socioeconomic background tended to fully support democratic values. This is consistent with previous evidence regarding the support of non-native students for democratic and more integrative participation (Isac et al., 2021).

It is interesting to recognise the relevance of intersectionality in these results. It is different – and it could imply opposite results – if you are an immigrant woman or immigrant man. If you are a female immigrant, you will most likely adhere to greater equality, in contrast to being a male immigrant.

Thus, in this research, among the social categories of concern are those with lower support for this outcome: immigrant men whose parents show university education, with no literacy resources at home. Understanding which groups can be in a weaker position regarding supporting human rights should not be a premise to separate them or discriminate against them. But instead, to take action to implement educative programs and give them access to the necessary information that, otherwise, they would not access on their own.

The presentation will include results for Italy and discuss with those obtained for Chile.

Keywords: Gender Equality, Intersectionality, MAIHDA, Chile, Italy

Scoring the Education for Sustainable Development (ESD) Subscale on ICCS 2022: Methodological and Conceptual Challenges of Subscores

David Torres Iribarra – Diego Carrasco

Introduction

Understanding students' proficiency in specific domains like Education for Sustainable Development (ESD) is highly desirable for informed policy-making and educational interventions. ESD is a complex collection of attributes encompassing different topics such as environmental awareness, willingness to behave pro-environmentally, and sustainability consciousness (e.g., Berglund et al., 2014).

The International Civic and Citizenship Education Study (ICCS) serves as a valuable platform for exploring students' knowledge and attitudes towards global citizenship and sustainability, as it includes ESD as a specific content subdomain (Damiani, 2021).

However, despite its inclusion in ICCS as a subdomain, generating ESD specific scores poses methodological challenges that warrant careful consideration. Particularly when these are generated after the fact, instead in conjunction with the public data release. For example, can a sub score be interpreted as a separate dimension (e.g. as in geometry ability within mathematics), or is it a “mini test” of the same content knowledge? What is the more suitable measurement model to generate such scores, a unidimensional

response model, a multidimensional model, or a higher ordered multidimensional model? (e.g., Mughogho, 2021). What are the limitations of these different choices, on how scores can be generated, and what are their sensible uses?

ICCS 2022 includes a total list of 141 items to generate a Civic Knowledge score, using ICCS 2009 anchored delta parameters from a item response model (Schulz et al., 2023). Out of the complete set of items, a subset of the items were designed to assess Education for Sustainable Development (ESD), an emphasis in the last ICCS cycle (Schulz et al., 2023). Assuming, the collection of ESD items be well spread between the different rotated booklets, and mode of applications (i.e., paper and pencil, and computer based applications).

Despite this conditions inputs it is in principle possible to generate individual estimates based that would represent the level of knowledge/proficiency of students on ESD issues if we take into consideration the specific list of items classified as tapping into ESD issues, their mode of application (papers based, and computer-based) and we apply an appropriate latent variable model.

In this presentation we outline a methodological framework for scoring the ESD subscale on ICCS 2022, emphasizing conceptual consistency, reliability, and validity. We address these methodological concerns that can arise from the generation of subscores, discussing considerations regarding the use of IRT based scores in this context; and its conceptual challenges (Torres Iribarra & Arneson, 2023) and technical issues associated with the dimensionality of subscores, as well their technical challenges (Sinharay, 2010; Haberman & Sinharay, 2010). We aim to provide researchers and policymakers with an analysis of the evidence that could support the use of these ESD subscores, an analysis of their limitations, and a sound assessment of its reasonable uses.

Methodological Framework

Data: Leveraging the dataset from ICCS 2022, which includes items specifically designed to assess students' knowledge and attitudes towards sustainable development, we will identify and extract the relevant ESD items for scoring purposes. We will analyze its internal structure as a separate attribute, its relation with the original construct and associated attributes in ICCS, and we will examine its reliability.

Procedures: Item Selection: We will conduct a thorough review of the ICCS 2022 items to identify those that specifically address ESD concepts and principles. Items selected for inclusion in the ESD subscale will undergo a review to ensure conceptual alignment with established frameworks for sustainable development education.

Scoring Models: We will employ latent variable modeling techniques, including multigroup Item Response Theory (IRT) models, to generate individual estimates of students' proficiency in ESD. This approach allows for the estimation of Expected A Posteriori (EAP) and Weighted Likelihood Estimates (WLE) individual scores, providing a comparison of the properties of different methods to generate individual ESD proficiencies.

Validity: We will examine validity evidence to support the interpretation of ESD scores by comparing the content coverage of selected items with established frameworks for sustainable development education and its relation to the original overarching scale and other related attributes.

Reliability: We will assess the reliability of the ESD subscale scores using person separation reliability estimates using person separation reliability estimates (Adams, 2005; Verhavert et al., 2018), examining the degree to which scores consistently allows inferences about respondents.

Implications

The proposed methodological framework holds several implications for research and practice in the generation of subscores from broader attribute in the context of large scale assessments.

By examining different scoring methods and collecting different forms of validity evidence, the proposed study will inform the extent to which we can make more specific and reliable inferences about students' proficiency in thematic subsets of larger assessment instruments.

On a more substantive note, the availability of robust ESD subscale scores will enable policymakers and educational stakeholders to make informed decisions regarding the design and implementation of interventions and educational programs aimed at promoting sustainable development education.

The proposed methodological framework lays the groundwork for future research endeavors aimed at further refining and validating measures of ESD proficiency and other relevant subscales in ICCS, thereby advancing our understanding of additional and more specific inferences based on international large scale assessments.

Conclusion

In conclusion, this proposal presents a methodological framework for scoring the ESD subscale on ICCS 2022, emphasizing the analysis of evidence of reliability, and validity, and evaluating how it supports (or

not) different uses. We aim to provide researchers and policymakers with robust estimates of students' proficiency in ESD, thereby facilitating informed decision-making and targeted interventions to promote sustainable development education, and to present a documented approach that could be used to generate new subscores in other parts of the ICCS instruments.

Keywords: Subscores; Multidimensionality, Education for Sustainable Development, Item Response Theory

Pro-Environmental Dispositions Among Secondary School Students

Ernesto Trevino – Diego Carrasco

The environmental degradation requires sustained actions to engage students in understanding that the human species is an integral part of the environment (Núñez Tobar et al., 2023). The Brundtland report (World Commission on Environment and Development, 1987) states the many challenges that countries' economic development and underdevelopment have carried with them: resource depletion and pollution, coupled with poverty and hunger in less resourceful areas happening jointly in a common world (Keeble, 1988). "Sustainable development" is a conceptual tool to face these challenges, pushing forward for a different way of development that cares for its consequences and allows for future generations to also meet their needs. It also includes a central role for education "to empower people to take responsibility for shaping their future" (Walid & Luetz, 2018). As such, school systems are viewed as an important vehicle to promote mental models regarding the relationship between humans and their environment, and they may offer opportunities to lead a change of mentality that considers humans not as users of the environment but as part of an interconnected ecological systems (Carr, 2016).

Education for sustainable development entails an action-oriented pedagogy, which instills among students the understanding that knowledge and skills require to be accompanied by actions to transform reality (González-Salamanca et al., 2020). However, the research agenda has focused on how effective schools promote sustainability consciousness (e.g., Berglund et al., 2014; Boeve-de Pauw et al., 2015; Olsson et al., 2016), instead of focusing on engagement into either present or future actions. Although bibliometrics analysis of education for sustainable development (ESD) attest a noticeable growth in publications, since the Brundtland report from 1987 till the present (Grosbeck et al., 2019), research inquiring into students' willingness to act pro-environmentally and its relation to school practices is scarce (e.g., Verhelst et al., 2021). Moreover, contemporary educational effectiveness research regarding schools' implementation of education for sustainable development (ESD) emphasizes the need to cover two aspects: a) the inclusion of students' perceptions of relevant ESD teaching practices, and b) an action-oriented focus (Isac et al., 2022; Sass et al., 2023; Sinakou et al., 2019). The present paper aims to bridge this latter gap. It looks into different school factors that may influence students' willingness to protect the environment. In essence, it focuses on learning opportunities relevant to education for sustainable development. The study focuses on the opportunities to learn in the school according to different informants. First, principals reported on the pro-environmental school actions, the activities related to environmental sustainability in the school, and the promotion of respect and safeguard of the environment in the school. Second, the study analyzes students' perspectives on their opportunities to learn on how to protect the environment in the school, as well as the openness of classroom discussions. Finally, the study analyzes the sense of preparedness of teachers to teach about the environment as well as their training opportunities. There are two foreseeable limitations of the present study. On the one hand, the outcome variable is a single item in the ICCS 2016, a feature that may restrict the variability when measuring students' future intended personal efforts to protect the environment. Therefore, the present study only covers a single focus on students' intention to take action to protect the environment in the future, that contrasts with the full array of components related to action competence as a desired learning outcome of education for sustainable development. Moreover, it is necessary to keep in mind that the results come from an observational cross-sectional design, which limits the ability to make direct causal claims based on its results. Thus, the interpretation of the presented results should be done with caution. In the present study, we use a series of ordinal logit models, to assess (Rabe-Hesketh & Skrondal, 2012) the relationship of all selected factors. We inquire if different opportunities to learn at school and teacher preparation contribute to students' willingness to act pro-environmentally. In the following section, we first describe how students' current behavior, norms endorsement, and threat awareness is related to their willingness to protect the environment. Then, we assess the relationship to the

learning opportunities schools provide. The main findings show that only 10% of students who did not endorse this social norm, could be expected to certainly be willing to engage in personal efforts to act pro-environmentally. In contrast, we can expect that 40% of students show a higher willingness to act pro-environmentally, when these also considered it important to protect the environment for good citizenry. This positive result was observed across all participating countries in the study. Hence, social norms endorsement related to environmental sustainability is positively related to students' intention to behave pro-environmentally. The results for the pooled sample show that a high proportion of students presented a high willingness to act pro-environmentally (75%) in schools where they had learned to a large extent how to protect the environment. We expect that only 38% of students present a high willingness to act pro-environmentally in schools where they have learned how to protect the environment to a lesser extent. This large difference in proportions of students with high pro-environmental intentions varied between countries by 15% to 55% across all countries with the exception of Slovenia, where we did not see this pattern of results. Conditional to the open classroom discussion level of the schools, we obtained increments of students with a high willingness to act pro-environmentally that varied between 7% to 35% between the students with average classroom discussion levels, and students in schools with one standard deviation of more open classroom discussion. In the pooled sample, for example we expect that 38% of students present a high willingness to act pro-environmentally when exposed to average levels of classroom discussion. In contrast, we expect that more than half of the students (52%) present a high willingness to act pro-environmentally if these are exposed to one standard deviation more of open classroom discussion. The results of the models suggest that teaching practices and action-oriented educational approaches are associated to students' intentions to make efforts to protect the environment in the future. The adjusted models include the entire sample of participating countries, however, there is both country and school variability in the results. Therefore, it is worth including some discussion on country and school differences. Students' willingness to make personal efforts to help the environment varies between schools, and the intra class correlation, close to 10%, is similar to that found by Sass et al. (2022). Research results therefore suggest that school and teacher practices are important ways of leveraging the engagement of students in future actions to protect the environment. However, the potential of such variables may be shaped by the contextual conditions of countries and schools within each country. This finding poses the question of how education for sustainable development may blend perspectives on teaching about the environment and teaching for the environment.

Keywords: Sustainable development, education, environment, school effects

SESSION 9. INEQUALITY AND EDUCATIONAL OUTCOMES

ORGANIZER: HARVARD AND BOCCONI UNIVERSITY

COORDINATOR: MICHELA CARLANA

19TH OCTOBER: 2.00 P.M. - 4.00 P.M. {ROOM 3 – RESEARCH 14}

Teacher effectiveness: insights from Italy

Giulia Assirelli – Giovanni Maria Abbiati

Introduction

Teacher effectiveness is a concept elaborated in Economics of Education literature that describes the contribution of individual teachers to the cognitive growth of their students, as measured by standardized competence tests (Goldhaber et al. 2015). Sociological studies of educational inequalities have often overlooked the importance of this factor and its impact on school-related processes. This concept, in fact, poses challenges for sociologists for several reasons, including its alignment with an efficiency-driven educational model and the risk of reinforcing negative teaching practices (such as teaching to test), as outlined by many commentators (e.g. Grimaldi and Barzanò 2014). However, the established consensus on the critical influence of teachers in shaping student success and job trajectories (and associated disparities)

calls for Sociologists to address this factor and ultimately integrate it into their empirical as well into their theoretical works (Thrupp 2001; Argentin 2018; Abbiati 2021).

In particular, the recent literature on the tertiary effects of social origins on educational inequality (i.e., the influence of those micro-level mechanisms operating within the school system that amplify inequalities of opportunity already in place), highlights the centrality of teachers in these processes, as actors who shape - often in a non-neutral way - the educational trajectory of students and amplify (almost always unconsciously or unintentionally) already existing gender, migratory and social class differentials (Argentin and Pavolini 2020). In the Italian case, numerous studies found that how teachers suffer from various types of biases that can undermine or conversely promote their students' careers (e.g. Alesina et al. 2018), elaborate differentiated guidance advice based on students' ascribed characteristics (Argentin et al, 2017; Aktaş et al., 2022; Manzella, 2023), and devote efforts in teaching related to their prospects of staying in school, the latter decision being taken while considering the socioeconomic composition of students (Barbieri et al. 2013). Related to the latter element is the peculiar functioning of the Italian teacher labor market, in which lacks monetary incentives that are not tied to seniority is compensated by the possibility of selecting the school to work in as seniority progresses. It has been observed how the combination arising from the composition of the student body (determined partly by phenomena related to residential segregation) and teachers preferences for better working conditions tends to reinforce perverse allocative mechanisms between teachers and their students, systematically at the disadvantage of students coming from lower socio-economic backgrounds (Barbieri et al. 2011; Abbiati et al. 2017).

Objectives

In this paper we provide teacher effectiveness estimates for Italy, a country characterized by marked educational inequalities despite its high degree of centralization, particularly in the lower and upper secondary school cycles (Falzetti 2019). Such estimates are studied in order to: a) test the robustness of this method when applied to the Italian case and to the Italian data structure, b) study the relevance and the distribution of teacher effects across subjects, regions, and schools, c) study how teachers effects reinforce or alleviate educational inequalities, d) test the predictive power of these estimates on later educational stages, and ultimately d) show to the community of social researchers the potential that this method has to address important issues in both educational research and public policy.

Data

To this purpose we analyze data on a nationally representative samples of Italian 8th and 10th grade students for the school years 2017-18 and 2018-19, which include competence scores in language and mathematics and socio-demographic information. We merged this data with a survey administered to their language and mathematics teachers.

The choice of this school years is motivated by the inclusion of a questionnaire item that allows the identification of teachers who have followed sample students from the beginning of their educational cycle (the target of the analyses). Therefore, we are able to distinguish teacher effects from the negative externalities associated to turnover.

Methods

In line with the literature, effectiveness estimates have been retrieved by regressing students' scores on a teacher dummy variable, students' competences at the end of the previous school cycle (i.e. 5th grade for models on eight graders and 8th grade for model on 10th graders) and a wide set of individual-level regressors. This method has been subject to some criticisms, mostly because a) it might be subject to confounding factors, such as teacher-student matching and b) it is estimated on a very limited number of subjects. However empirical research showed that such estimation method is sufficiently robust to fulfill its function (Chetty et al. 2014).

Results

These estimates have been analyzed to assess their distribution among schools, families, and territories, and to study their predictive power in shaping students' careers in later educational stages. This study will contribute to a deeper understanding of teacher effectiveness and its implications for addressing educational disparities.

Keywords: educational inequalities; teacher effectiveness; student achievement

Inclusive Teaching and Classroom Relational Climate: Spotting Social Isolation in the Classroom

Social cohesion is a crucial factor in achieving sustainable economic development and prosperity (Alesina and La Ferrara, 2005; Easterly et al., 2006). A cohesive society possesses a good collective character supported by the well-developed social skills of its members (e.g., trust, reciprocity, cooperation). These skills allow the formation of social capital to effectively communicate with others, engage in ethical economic interactions, and work together for common goals. In settings where social skills do not develop properly, social exclusion and isolation may occur, leading, over time, to violence, intolerance, and identity-based segregation. Moreover, when feelings of exclusion persist over extended periods of time, chronic expectations of exclusion may result (Mendoza-Denton et al., 2002). People who develop such expectations are more likely to feel depressed, have low self-esteem and experience more negative affect than those who expect to be accepted (Ayduk et al., 2001; Downey and Feldman, 1996; Mendoza-Denton et al., 2002). Therefore, intervening early to shape social skills correctly is crucial.

Recent evidence indicates that schools are important in shaping these skills, especially for socioeconomically disadvantaged children. There is now ample evidence suggesting that early educational interventions targeting children's socio-emotional skills may have long-lasting impacts on individual outcomes (Alan and Ertac, 2018; Alan et al., 2019; Sorrentiet al., 2020) and collective outcomes related to social cohesion and intergroup relations (Alan et al., 2021). Given that the opportunities to develop such skills are limited for children of poor families, the curricular and pedagogical school inputs are especially valuable for them. For these children, schools represent the only opportunity to develop social and emotional skills and learn how to form healthy social relationships (Gradstein and Justman, 2002). In this study, we design and evaluate a "feedback treatment" (FT) that helps teachers recognize the social isolation that exists within their classrooms, especially among migrant and low socio-economic status children. The intervention also highlights the negative impact of social exclusion on children's socioemotional and cognitive skills. It emphasizes the vital role of teachers in creating inclusive classroom environments and offers practical suggestions for implementing inclusive teaching methods. We implemented this study in public primary schools in three regions of Italy. The intervention was implemented as a randomized controlled trials. We recruited schools through an open call sent to all primary public schools in the target areas. Schools teachers were offered to participate to a training program for the development of social skills (DeSS).¹ To this purpose we collected baseline data in October 2021. Teachers participated to the DeSS training program during the 2021-2022 school year. We used data collected at baseline to build a "feedback treatment" directed at teachers. We randomly allocated 23 schools to receive this treatment and the remaining 23 schools to a control group that did not receive any feedback. In November 2022 we invited teachers in treatment schools to an online meeting. The purpose of the meeting was to highlight the important role teachers have in the formation of children's social skills and in the development of social relationships. We revealed to teachers the average prevalence of social exclusion within the sample that was well above the teachers's perceived prevalence. We then highlighted that isolation was prevalent among children from low socio-economic status or migrant status (i.e. first or second generation migrants) and showed that social exclusion was positively correlated with low mental health status and a worse perception of the school environment. After the meeting, we sent class-level reports to all teachers in treatment schools and invited them to join WhatsApp groups to receive weekly practical tips on adopting inclusive practices during their teaching days. We collected endline data in May 2023. The objective of this study is to examine the effects of the intervention on (anti) social behaviors, social isolation and ethnic segregation and on children's socio-emotional status and perception of school climate. Our hypothesis is that by providing teachers with information that draws attention to social exclusion patterns in classrooms, emphasizing their role in developing inclusive classroom networks, and equipping them with practical tools to adopt inclusive teaching practices, we can reduce antisocial behaviors, social exclusion, and ethnic segregation. This, in turn, can lead to an improvement in children's socio-emotional status.

Our goal with this study is to enhance our comprehension of how the school environment affects our collective character as a community, shapes our interactions with out-groups, and develops crucial social skills that contribute to our social capital. We are focusing on the role of teachers in a context where ethnic and socio-economic diversity can be significant, such as the classroom. Specifically, we are interested in how teaching practices influence children's social skills and peer relationships, and in how these skills

develop differently among immigrant and native students. To evaluate the intervention we collect a wide range of outcomes that characterize these dimensions. We implemented incentivized games to elicit antisocial behaviors using a performance sabotage game. We elicited classroom social networks by asking children to nominate at least three classmates as their best friend. We use these data to construct measures of social isolation and ethnic segregation. We implemented item-response questionnaires to measure children's mental health status, perception of the school environment, sense of belonging to school and perception of teachers behaviors and attitudes. These are all self-reported measures.

Along with the outcomes of interest we collected demographic data, parents characteristics and a battery of teachers characteristics, attitudes towards social inclusion and teaching practices. We document the following main results. First we show that the treatment reduces antisocial behaviors elicited with the performance sabotage game. By simulating payoffs of the games in two scenarios with and without sabotage using the empirical distribution, we show that the treatment generates higher class-level wealth and lower inequality relative to the control group. These results suggest that on average the treatment improved the relational climate of the classroom which translated into lower sabotage, higher outputs and lower inequality. Second, we show that the treatment reduces social isolation and class level ethnic segregation. In particular the treatment increases the number of nominations students receive or give to other students. Finally, we find that the intervention improves the socio-emotional well-being of children measured with some dimensions mental health and sense of belonging to school.

This study makes two important contributions. First, to the best of our knowledge this is the first study that shows how revealing to teachers the extent of social isolation within a group, providing information about its consequence and giving practical tools on how to reduce isolation can reduce antisocial behaviors and social exclusion. Second, we investigate our research question thanks to the collection of rich and high quality data. We measure antisocial behaviors through an incentivized sabotage game that, to the best of our knowledge, was not previously implemented with young children. We obtain objective children's antisocial behavioral outcomes that are usually measured with subjective scales developed in the social psychology literature. Moreover, thanks to the collection of rich network data we objectively measure social exclusion and ethnic segregation

Keywords: Social Isolation, Inclusion, Ethnic Segregation, Sabotage, Teachers effects, Feedback

School Closures and Education Inequality: Is Online Teaching a Great Equalizer?

Massimo Anelli – Vincenzo Galasso – Silvia Griselda – Herman G. Van De Werfhorst

Important questions in the social stratification literature are how much education contributes to learning, and whether it equalizes opportunities across children of different socioeconomic backgrounds. One way to study these questions is through the comparison of skill acquisition between periods when schools are in session and periods when schools are not in session. A sizeable body of research has compared progress in academic achievement scores during the summer and during periods when children go to school, the so-called seasonal variation design (Downey 2023). A common finding is that the learning progression curve goes steeper in school time compared to the summer period, indicating that children's cognitive development goes faster when schools are in session than when schools are not. Moreover, most studies also find a compensatory functioning of schools because the gains of schooling (relative to home learning) are higher for children of less resourceful families. Exposure to schooling, in other words, equalizes opportunities of children of different backgrounds. Recent reviews conclude, however, that the seasonal variation models do not always replicate well (Workman et al. 2023). One problem with this design is that the measurement of exposure to schooling is imperfect, for instance because the period between June and October is used to define the summer, while in fact there is quite some exposure to schooling in those months.

Other designs have been proposed to study exposure effects of education on learning by looking at the precise length of schooling that students have been subjected to. One example is to measure exposure by examining the exact date of achievement tests at the start of primary school (Passaretta & Skopek 2021). Comparing school entry dates and test dates, it appeared that schooling contributes to cognitive development, but equally so for students of different socioeconomic backgrounds. Hence, there was no compensatory effect of schooling in that study. Other exposure designs have examined the reduction of the length of secondary school by one year in Germany (Marcus and Zambre 2019), the reduction of the length of the school year (Grätz 2023), or the variability in length of school careers up to an intelligence test as part of the army conscription examination (Carlsson et al. 2015).

While the exposure designs contrast school-based development with development in the home, contemporary learning practices are more varied. Certainly with the rise of educational technology, today's spectrum of modes of cognitive development include not only school-based and home-based learning, but also learning through online environments. Schools were rapidly drawn into the development of online education as an alternative to in-person classes, when the COVID-19 pandemic happened and schools had to be closed. Estimates are that learning progression was much reduced during the school closures relative to regular school years (Engzell et al. 2021; Betthäuser et al. 2023).

In this paper we use the experiences during the pandemic to engage with the question how cognitive development can depend on three types of input: the school while in session, the family outside schooling hours, and the online school environment while at home. It can be assumed that school-based classes are most compensatory: in-school education reduces the impact of family background on learning. But how effective is exposure to online education in comparison to in-school education? Given that summer learning goes faster among high-SES students presumably because parents contribute to learning at home, and knowing that parents of more advantaged socioeconomic positions have more resources to employ in the home to assist with home-based online learning (Bol 2020; Van de Werfhorst et al. 2022), it is unlikely that online education is as effective for everybody as in-school education.

It is important to know more about the effectiveness of online learning not only in the context of the pandemic. If online education is effective, it can be a useful tool for after-school tutoring, for the recuperation of the learning delays associated with the pandemic, or for summer programmes targeting low performing students (Kim & Quinn 2013). Online education may also offer a solution to cope with teacher shortages, or to make efficient use of available technology. If the effectivity of online education is homogeneous across socioeconomic groups, such skill-enhancing interventions will moreover raise the standards without increasing inequalities. If, however, online education is more effective for higher-SES students than for lower-SES students, inequalities go up in online environments, which could raise concerns by educational policy makers and school leaders about the feasibility of digitalization of learning.

We collected new survey data among parents in .. high-income countries, at the end of 2020, the first year of the pandemic. We assessed detailed information on the number of hours of schooling in a regular pre-pandemic school day, and the number of hours of online education on days when schools were closed during the pandemic. This information was combined with estimates of the effectiveness of school-based and online learning taken from other publications, heterogeneous across socioeconomic groups. Together, this approach enables us to assess the exposure effects of in-class and online learning for different subgroups. Moreover, we can calculate the learning loss (by group) when in-school education was (partly) replaced by online learning.

Keywords: School closures, Inequality, In-person and remote learning, COVID-19

Apart but Connected: Online Tutoring, Cognitive Outcomes, and Soft Skills

Michela Carlana – Eliana La Ferrara

Tutoring has been proved to be one of the most effective tools for enhancing educational outcomes and mitigating disparities in learning (Nickow et al., 2020; Fryer Jr, 2017). This potential is particularly relevant in the aftermath of the Covid-19 pandemic, which generated significant learning losses across countries, estimated at 0.17 standard deviations in achievement test scores (Patrinos et al., 2022). One of the main challenges of tutoring programs is their scalability, both in terms of costs and in terms of supply of qualified tutors –especially when offered in disadvantaged settings. This paper presents the results from an

innovative online tutoring program that we designed and evaluated at the outset of the Covid-19 pandemic and subsequently continued implementing during regular school periods.

The "Tutoring Online Program" (TOP) is a pioneering policy experiment launched in Italy in the spring of 2020. To the best of our knowledge, TOP is the first online tutoring program created in response to the pandemic and it has become a systematic policy to mitigate learning gaps in Italy. It is offered free of charge and targets underprivileged students in grades 6 to 8 (middle school). TOP has two distinguishing features. First, all tutoring sessions are conducted exclusively online. As mentioned above, the circumstances of the lockdown and the limited availability of qualified tutors in underserved areas make remote delivery a particularly interesting feature to test. Second, the tutors in TOP are not professional educators but volunteer university students, who undergo training and receive support from pedagogical experts. While teachers and professionals undoubtedly possess the necessary qualifications, volunteer tutors offer advantages from a budgetary perspective and may also enhance the quality of interpersonal interaction. Indeed, TOP harnesses the intrinsic motivation of university students to volunteer and helps them develop valuable soft skills in the process.

Since its creation, TOP has reached over 4,000 disadvantaged students. In this paper, we focus on the editions implemented in 2020 and 2022, which involved a clean evaluation design and are broadly comparable. TOP 2020 was implemented during a period in which all schools were closed, while TOP 2022 during normal school times. In both these editions, the demand of tutors by middle school students could not be fully met due to budgetary and administrative constraints. As a result, tutors were allocated randomly among eligible students, taking into account the subjects requested by the student (specifically, Math, Italian, and English), as well as each tutor's availability to teach a particular subject. This allows us to causally estimate the impact of being assigned a tutor on the students' performance, as well as on a number of non-academic outcomes. To this end, we collected administrative data as well as detailed individual surveys from students, parents, tutors and teachers before and after each intervention.

We find sizeable and significant improvements in math performance for students who were assigned an online tutor, compared to those who were not. Teacher-assigned math grades increased by 0.11 SD and 0.09 SD in TOP 2020 and TOP 2022 and the probability of failing the subject fell by 24 and 21 percent, respectively. TOP also improved treated students' performance in a standardized math test that we administered at endline, by 0.24 SD in 2020 and 0.16 SD in 2022. When these three outcomes are aggregated into an overall math performance index, the effect of the program is a 0.3 SD increase in the index in 2020, and +0.16 SD in 2022. These are remarkable effects, given that the median length of the online tutoring was about 6 weeks in both rounds.

Our experimental design allows us to investigate the channels of impact, following the conceptual framework by Nickow et al. (2020). Starting with the dosage of the intervention, we can exploit random variation in the number of hours of online tutoring per week. While the majority of students in TOP 2020 received 3 hours per week, a random subset of those who needed help in more than one subject was assigned a tutor for 6 hours per week. We find that math performance gains double and grade retention halves with the more 'intense' tutoring. The positive impact on test scores persists in the longer term, with an improvement of +0.24 SD in the national standardized math test score for students receiving high-dosage tutoring, more than one year after the end of the intervention. Second, we investigate the impact of tailored instruction by randomizing students to individual vs. small group tutoring in TOP 2022. While small group tutoring may be more cost effective, one-on-one tutoring enables more personalized instruction. We find larger effects on the math performance index for individual tutoring (+0.2 SD) vs. group tutoring (+0.11 SD). The tutor/tutee relationship may extend beyond academic content, involving a mentorship aspect that may influence students' aspirations, socio-emotional skills, and psychological well-being. This was particularly valuable during the pandemic, when opportunities for social interactions were reduced and students experienced higher levels of depression and slower development of socio-emotional skills (Orgilés et al., 2020; Golberstein et al., 2020). We find that TOP 2020 had sizeable and significant effects on an index of educational aspirations (+0.19 SD), a socio-emotional skills index capturing perseverance, grit and locus of control (+0.16 SD) and an index of psychological well-being measuring happiness and depression (+0.16 SD). No significant effect on these outcomes was found in TOP 2022. This aligns with the notion that, during periods when students' aspirations and soft skills are shaped by many other in-person interactions, the role of the tutor becomes less critical.

We test for heterogeneity of treatment effects by student and tutor characteristics. We do not detect significant differences in impact between boys and girls, nor immigrants and natives – except for the effect on psychological well-being which is entirely driven by immigrant students in TOP 2020. Improvements in

academic performance tend to be somewhat more concentrated among students from disadvantaged backgrounds, though the differences are not always statistically significant. Interestingly, tutor characteristics such as gender, GPA, pro-social attitudes and motivation do not systematically affect the effectiveness of the tutoring. For pro-social attitudes and motivation, though, it should be observed that our tutors are already positively selected along these dimensions, implying relatively low variation along these dimensions.

Finally, we investigate how the experience of being a TOP tutor during the pandemic affected the tutors themselves. We can do so because in 2020 we randomly selected the university students to whom we offered the job from the pool of those who applied to be volunteers. Four months after the end of the program, we find that volunteers who were included in the TOP program have significantly higher empathy than those who were not. The effect corresponds to a 0.27 SD increase. We instead do not find significant effects on tutors' beliefs regarding the relative role of luck versus hard work in determining success in life. Our paper contributes to several strands of literature. A considerable body of work shows that in-person tutoring is highly effective for improving academic outcomes. Recent meta-analyses find that the impacts are sizeable (a pooled effect size of 0.37 SD in Nickow et al. (2020)), and robust across a wide array of contextual factors (Fryer Jr, 2017). The importance of small group or individual tutoring has been underlined for students who struggle (Ander et al., 2016; Cabezas et al., 2011) and in order to teach at the right level (Banerjee et al., 2015). Also, the tutor-student relationships are often close to a mentorship connection that may affect the development of cognitive as well as social skills, such as prosociality (Falk et al., 2024; Kosse et al., 2020; Resnjanskij et al., 2023; Gallego et al., 2023). On the other hand, tutoring is much costlier than classroom instruction and it may not be easy to arrange individual, in-person tutoring in the presence of geographical constraints. Also, tutoring may sometimes be attached with the stigma of being identified as a student in-need and pulled out from regular classes (Coie and Krehbiel, 1984; Richmond, 2015). We contribute to this literature by providing evidence on large-scale online tutoring done by volunteer tutors. Our model allows to substantially reduce costs –one of the most significant barriers to large-scale implementation– but also to efficiently reach students located in disadvantaged areas through virtual learning. Finally, online tutoring is less observable from peers than in-person tutoring, which may reduce the stigma possibly attached with this intervention.

Our results are of course directly relevant to the debate on effective strategies to mitigate the effects of Covid-19 on education. The existing evidence suggests that students who lag behind the most during the pandemic are from low-income families with limited access to technology, and that they receive less support from parents and lower quality of remote learning from schools (Bacher-Hicks et al., 2020; Chetty et al., 2020; Engzell et al., 2020; Agostinelli et al., 2022; Stantcheva, 2022; Carlana et al., 2023; Jack et al., 2023; Werner and Woessmann, 2023; Lichand et al., 2022). Different forms of remote instruction have been adopted across countries and, although the evidence on interventions before the pandemic is mixed (Escueta et al., 2017; Malamud and Pop-Eleches, 2011; Fabregas and Sola, 2023), the impact of digital technology may differ during school closures compared to normal school years.

To the best of our knowledge, very few policy experiments have attempted to use remote tools to improve learning during the pandemic and replicated them after the school closures (Angrist et al., 2023). Angrist et al. (2022) evaluate two low-tech interventions in Botswana that use SMS text messages and direct phone calls to support parents in the education of their children, finding sizeable effects on student outcomes and parental beliefs. Following our early work in 2020, other remote tutoring interventions have been implemented finding positive effects on academic outcomes. Gortazar et al. (2023) find positive impacts of an online group tutoring program implemented in Spain; Hardt et al. (2022) evaluate a remote peer mentoring intervention at a German university during the pandemic, where peers met online to discuss self-organization; Hassan et al. (2022) evaluate the impact of randomizing an over-the-phone learning support intervention on primary school students in Bangladesh; Kraft et al. (2022) conduct a pilot in which they assign a (volunteer) college student tutor to a middle school student. Our work contributes to this body of work in various ways. First, we evaluate the effects of the first online tutoring intervention implemented in response to Covid-19. TOP is an innovative and low cost online tutoring program targeting teenage students who were adversely affected by school closures. Second, we provide evidence on the effectiveness of the same intervention implemented during school closures (2020) and in-person schooling (2022). Third, we show impacts on learning outcomes as well as soft skills and psychological wellbeing. Fourth, we study within the same intervention the effects on the students and on the tutors themselves.

Finally, recent work on organizations highlights the power of intrinsic motivation and social recognition for improving public service delivery (e.g., Ashraf et al., 2014; Gauri et al., 2019). Levitt et al. (2016) underline

that such behavioral aspects can be leveraged to improve educational performance. While we cannot directly speak to this question –as we did not vary the recruitment method, or the incentives provided to tutors– the fact that our tutors self-selected into volunteering for TOP and their intrinsic motivation may have contributed to the effectiveness of our intervention. We do provide evidence that volunteering as a tutor increased empathy compared to university students who applied but were not assigned a student.

Keywords: tutoring, COVID-19, socio- emotional skills, well-being

HERO Teachers: Harboring Empathy to Raise Opportunities

Noemi Facchetti – Selene Ghisolfi – Giovanna Marcolongo – Marco Le Moglie

Introduction and Literature

In the second quarter of 2020, about 13.5% of Italian people aged between 18-24 years old left the school pathway before its completion (Eurostat, 2021). The number, only three percentage points higher than the European average, hides a divergence between the North and the South of the country. In the South, the dropout rate reaches almost 20%. The increase in dropout today poses challenges that transcend educational attainment and extend to a higher risk of inequality in the future through decreased participation in the job market and worsened physical and mental wellbeing (García-Santillán et al, 2021). The risk of dropout is higher among students from fragile socio-economic backgrounds. According to ISTAT (2021), the rate of students leaving school before obtaining a diploma is nearly three times higher among foreign students, and almost twice as high among those whose parents are unqualified workers or unemployed. Higher dropout rates among disadvantaged contexts jeopardize the role of the school as “social equalizer”.

In the educational system, teachers play an essential role in fostering students' motivation and shaping their aspirations (Alesina et al. 2018). This is even more true in fragile contexts, where children are less likely to receive parental support, but more likely to come from single-parent families. The economic and psychological literatures have widely explored how motivated teachers can discourage the risk of students' dropout and improve their academic performance (Zein et al., 2020, Wondimu et al, 2010). A healthy and positive teaching environment is key for students to thrive academically and reduce the risk of leaving school too early (Collie et al. 2012; Shefi and Chis, 2016; Deming, 2011; Wondimu et al., 2010, Roorda, 2007). To the contrary, relationships that are characterized by conflict and lack of trust jeopardize students' achievement (Spilt et al. 2011, Hamre and Pianta 2001).

While the role teachers play is particularly pivotal in more disadvantaged schools, it is also true that their efficacy and motivation face higher challenges in these contexts. Teachers are more likely to encounter high-stress situations (Kyriacou, 2001) and to face students' undisciplined behavior. These factors may wear away their job motivation, increase their frustration and, with it, their risk of burnout. Lower teachers' efficacy frays teacher-student relationships: teachers are more likely to see their students negatively and to face conflict (Grayson and Alvarez, 2008, Hamre and Pianta, 2001, Roorda et al., 2011, Shefi and Chis, 2016). Students facing exhausted teachers perceive lower support and face more frequent interpersonal difficulties (Shen et al., 2015, Grayson and Alvarez, 2008). They no longer seek support from their professors, but rather antagonize them. This, in turn, affects students' motivation causing worse academic achievement (Arens and Morin, 2016; Herman et al., 2018; Madigan and Kim, 2021; Zee and Koomen, 2016) and increasing the risk of dropout (Zein et al. 2019).

Contribution, Research Subject, Objective and Hypotheses

Different contributions have suggested and studied the effectiveness of various policies to decrease the risk of students' dropout. Typically, these policies target students and their families, for example by providing incentives for school attendance or helping students not lagging behind via tutoring (Baird et al. 2011, Dearden et al., 2009, Van der Steeg et al. 2015). Despite the variety of descriptive literature analyzing the relevance of positive relationships between students and teachers to discourage the risk of dropout (e.g. Hamre and Pianta, 2001, Roorda et al., 2011, Shefi and Chis, 2016), less attention has so far been provided on how supporting teachers may reduce early school leaving by students. In our work, we study how a policy aimed at improving teachers' wellbeing can reduce students' risk of dropout.

In this project, we measure the impact of an intervention designed to improve teachers' wellbeing and student-teacher interactions on class climate and students' dropout. The intervention provides teachers

with a group of mutual support to teachers, similar in spirit to the “Balint” (Balint, 1955) group, from the psychological literature. In Balint groups - typically implemented in medical contexts - physicians share cases and challenges they face with patients and look for support and suggestions from other clinicians, so as to improve their relationships with patients and take care of their own psychological wellbeing. In our context, we asked teachers to join eight group-sessions led a trained psychologist. The sessions are design to create a safe space to discuss the challenges and struggles teachers face with students. The psychologist complement the structure of the Balint group, providing guidance on how to best address the challenges teachers face with students.

We test whether the intervention improves teachers’ wellbeing, students/teachers interactions and, ultimately, students’ motivation to attend schools and therefore their absences and drop out rates. Teacher’s mental health is an important outcome per se, but also it holds relevant consequences on students. Teachers in middle schools suffer from high risk of burnout (Chang, 2009) which hurts their wellbeing and motivation, consequently affecting children’s learning (Madigan and Kim, 2021). The literature offers extensive evidence of the positive effects of tutoring programs for students and teachers on students’ academic performance and aspirations (e.g. Alan et al., 2019). Carlana et al. 2022 show that teachers’ behavior, attitudes and biases impact students’ achievement. However, little is known about the possibility of improving students’ outcomes combining psychological support to teachers with training, especially in challenging environments.

Data and Methodology

We use a randomized control trial (RCT) to measure the impact of offering psychological support to teachers on dropout rates of students and collect information on how such effects may be mediated by students’ aspirations and achievements, students’ and teachers’ well-being in school, class climate and teachers’ motivation. Figure 1 explains our theory of change.

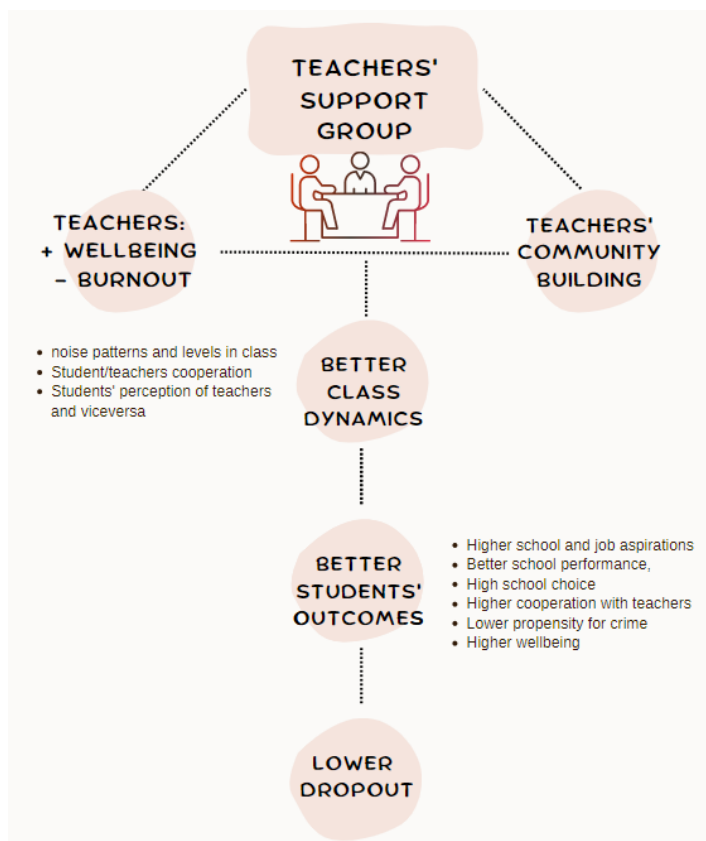


Figure 1: Theory of change

Between July and December 2023 we contacted and recruited 60 middle schools in the municipality and surrounding areas of Naples interested in participating in the project Su.Per.Prof (Support To Teachers). Each school’s principal identified the stream (i.e., Sezione) with the highest concentration of children at risk of drop-out. We randomly assigned half of the schools to the treatment group (scheduling of eight group meetings with a psychologist) and the remaining half to the control group (no intervention).

Teachers of treated schools met eight times in group session led by a psychologist discussing eight topics varying from students' motivation, to how to better work with colleagues and to how to develop effective communication with family, teachers and students.

We will evaluate the effectiveness of the intervention on students' dropout, measured through administrative data, and assess how the effect is mediated by several measures collected through baseline and follow-up surveys: i) students' aspirations and achievements; ii) students' and teachers' wellbeing in school; iii) class climate and teachers' motivation. Before the start of the intervention, students and teachers in all enrolled schools (i.e., both treatment and control) have filled a baseline questionnaire including validated psychometric measures of wellbeing, self-efficacy, class climate borrowed by the psychological literature. The same survey is being administered between April and May 2024 to all the students and teachers that participated. We plan to complement the analysis with administrative data from INVALSI and from the Ministry of Education. In particular, data on students' and teachers' absences will be key to our analysis.

We will run regressions as "diff-in-diffs" to measure the impact of our intervention, that is we will compare the change in the outcomes listed among treated schools versus control schools.

Results

At the current moment, no result is available as the intervention is still ongoing. However, we plan to have preliminary results available by the Fall and for the IX Seminar "I dati del e per il sistema educativo: strumenti per la ricerca e la didattica".

Keywords: inequality, drop out, teachers, class climate, psychological support

SESSION 6. THE ACCUMULATION OF HUMAN CAPITAL: A COMMON GOAL OF SCHOOL, VOCATIONAL TRAINING, AND THE LABOUR MARKET

ORGANIZER: INVALSI – IRPET - FBK-IRVAPP

COORDINATOR: DAVIDE AZZOLINI

19TH OCTOBER: 2.00 P.M. - 4.00 P.M. {ROOM 4 – RESEARCH 15}

School to Work Transitions. A longitudinal study on three cohorts of students in the province of Trento

Davide Azzolini – Sonia Marzadro – Mattia Oliviero

Introduction

As has already emerged from other research, in Italy individual characteristics strongly influence their educational and employment outcomes, starting with school choice in secondary education. The present work attempts to investigate the role played by some of these characteristics, such as gender, migration background, and area of residence. The research is based on a single archive obtained by integrating various administrative data archives relating to the province of Trento.

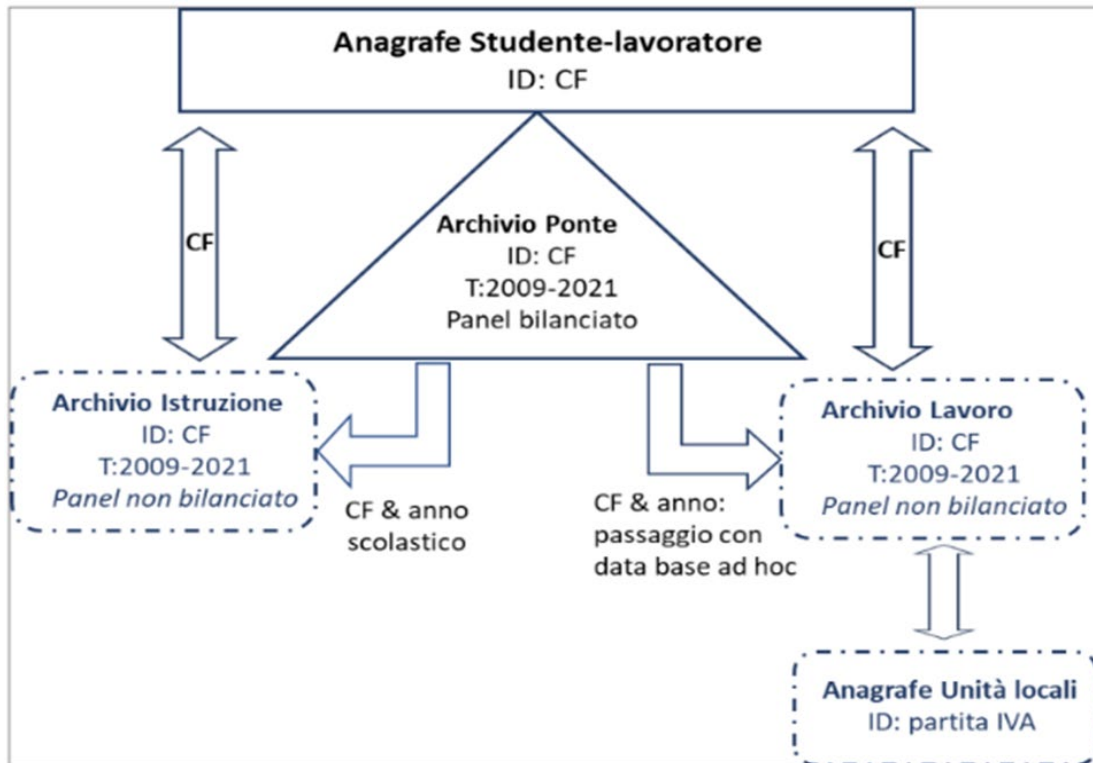
Purpose

Our research focuses on three cohorts of lower secondary school leavers in the province of Trento who, following their graduation, continued their studies in upper secondary school. Among the outcome variables considered, upper secondary school choice and the type of school completed are those we mainly focus on, although we also consider participation in the labor market and post-secondary education. Three individual characteristics are taken into account: (a) gender; (b) country of birth (whether Italy or abroad); (c) size of municipality of residence. In addition, the role played by school performance is investigated, such as: (a) the grade obtained in the first-cycle state exam; (b) and an indicator of delay in the school career.

Data

The data come from three sources: the Education DataWarehouse (DW) of the Province of Trento, the records of the University of Trento, the Register of Compulsory Communication (COB). The analysis is conducted on three cohorts of secondary school leavers in the school years 2008/2009, 2009/2010 and 2010/2011. The total number of observations is 15,845 subjects

Figure 1- The archive combining different administrative data



Methods

To establish the extent to which the variables listed influence the probability of enrolling in (and completing) the various secondary school courses, multinomial logistic regression models are run, which allow us to estimate the predicted probabilities of enrolling in (and completing) the different type of schools.

Findings

Gender appears to be strongly associated with school choice: girls are more likely than boys to choose licei. In contrast, boys are more likely to choose technical schools or vocational training.

Individuals' place of birth makes the difference in the likelihood of choosing a "traditional" liceo (i.e., scientifico or classico) or vocational training, while it plays a smaller role in the other licei and almost no role at all in technical institutes.

Moreover, different choices can be discerned according to the size of the municipality: the greatest differences concern the choice of traditional licei (greater among those living in large municipalities) and of technical schools or vocational training (in both cases greater among those living in small municipalities). In addition to being influenced by personal characteristics, the choice of the type of secondary school is also correlated with school performance (the grade obtained in the state exam and the delay in studies). Using a multinomial regression model in which we also control for the grade obtained in the final state examination at the end of the first cycle, we learn that this (together with 'school delay') are positively and statistically significantly correlated with the choice of secondary school track. Considering school performance cancels out the gender differences between those who choose traditional licei (the difference between females and males in the predicted probability goes from 0.10 points to 0.01) and drastically reduces those observed in vocational training (which go from -0.10 to -0.02 points). The gender gaps remain "qualitatively" unchanged for the chosen technical and other licei. The effect of place of birth is, moreover, very limited. A result in line with the scientific literature on the role of migratory background, which has highlighted that, net of previous scholastic performance, the children of immigrants may also show more ambitious educational choices than natives. Finally, the effect of the municipality of residence on educational choices does not seem to be affected by school performance, as these do not vary appreciably with the size of the municipality of residence, contrary to gender and migration background. Instead, the differences between small and large municipalities seem to be more due to other factors linked to the different local school offerings.

Looking at the status of students five years after the completion of lower secondary school, we observe that girls are more likely to graduate from upper secondary education within the legal duration of the course of study than boys. They are less likely to be delayed in their school career and to drop out within five years of first enrolling in secondary school. The influence of place of birth appears to be much stronger than that

of gender. The probability of obtaining a diploma is 18.8 percentage points higher among natives than among those born abroad. On the contrary, it is the latter who are most likely to stop attending school prematurely without obtaining a degree. If we look at the role played by the municipality of residence, we find that boys and girls living in smaller municipalities show a slightly higher probability of having obtained a diploma five years after the first cycle state examination (76.5% vs. 72.6%).

When we recalculate the same probabilities as above, controlling for school performance and school of enrolment, the inequalities reported above are drastically reduced and, if significant, remain smaller. For example, for degree attainment, the gap between males and females is halved (from +0.089 to +0.049) and that related to migration background is reduced threefold. This confirms the importance of the transition from lower to upper secondary school even for outcomes observed five years after enrolment in upper secondary school.

The analysis of entry into the labor market during and after the completion of upper secondary school (or vocational training) is still ongoing. While, as far as post-secondary education (university and higher technical education) is concerned, the archive currently suffers from the limitation of only including students who have enrolled in the University Trento, therefore an integration of the archive with data from other Italian universities is necessary.

Keywords: school-to-work transition, administrative archives, inequalities

Participation and Dropout in Piedmont: How Integrating Data Sources Enhances Understanding of the Education and Training System

Luisa Donato – Carla Nanni

Introduction

A broad and extensive sociological tradition has focused its analysis on inequalities in education and equal opportunities, demonstrating the extent to which school careers, access to higher levels of education, and work are determined by inequalities stemming from social class and family background (Bolivar, 2005). The monopoly established in the literature on this issue (Boudon, 1983; Bourdieu & Passeron, 1977) has over time led to broadening this view with questions that include a greater complexity of factors. Those proposed in this contribution relate to the role that integrating data sources on education, training, educational guidance and school-to-work transition paths can play in improving system knowledge, encouraging participation, and reducing school dropout.

The analysis of inequalities in the school context has historically gone through various phases (Grisay, 1984). The current phase is characterized by a central interest in social achievement equality. Individual, family, and contextual characteristics are identified, but there is no hierarchy among these dimensions. It is assumed that certain inequalities are inevitable and must be considered; in fact, as Sen (1995) states, "treating everyone as equals can turn into unequal treatment towards those who are in a disadvantageous situation", therefore, to ensure equality of opportunity, attention must be directed to the most vulnerable groups as a priority (Rojas, 2004).

It is in this phase that school dropout enters the debate on educational opportunity inequalities. The focus is on the very concept of school dropout, over which there is debate among those who consider it a term specifically referring to leaving the education and training system (Morgagni, 1998; Sempio Liverta, 1999) and those who understand it as an umbrella concept under which are grouped the difficulties and incidents that students may encounter in their career (Zurla 2004, Besozzi 2017; INVALSI, 2020, Nanni, Donato 2020). The former, understood as explicit dropout, means interruption of attendance, the aspect most fraught with negative consequences because the student leaves the system before completing the course, without a diploma. The indicators for this involve interruption during the school career (Drop-out) and downstream dropout, i.e., at an age later than that of attending secondary paths (ELET - Early Leaving from Education and Training). The latter, called implicit dropout, monitors the qualitative aspect of the phenomenon: the portion of students who manage to complete the paths but with learning levels not adequate to the titles obtained (INVALSI, 2023).

At the regional level of Piedmont, the possibility of integrating standard indicators with information on participation in professional Education and Training paths under regional management, on activities

offered by the regional Educational Guidance System, and information provided by the Labor Force Survey, allows for providing additional elements in the debate on how to encourage participation in the training system and combat school dropout in its various expressions.

Objectives

The objective of the contribution is an integrated analysis to systematize the results of Piedmontese studies on the education system, on education and professional training paths, on orientation, and on the school-to-work transition. The integrated perspective helps to understand the trends of participation in the system in its complexity and of school dropout, offering keys to interpretation enriched by the complementarity of the available information.

The available sources in the Piedmont region are of different types: administrative, sample-based, and include:

- School survey of the Piedmont Region: preschool, primary and lower, and upper secondary school.
- Administrative data of the professional training courses of the Piedmont Region: IeFP (Initial Vocational Education and Training) paths in training agencies.
- Data provided by the Statistical Office of the Ministry of Education
- Eurostat-ISTAT, ELET rate.
- SNV-INVALSI, learning levels (low performers) and implicit dropout.
- Forze Lavoro Survey for education levels, age groups of the Piedmont population with provincial details.
- Regione Piemonte, SERSE data, enrollments in paths in the regional Educational Guidance System OOP (Obiettivo Orientamento Piemonte).

Methodological Approach

The analyses use different techniques and quantitative methodologies. Descriptive analysis includes the collection, cleaning, and management of secondary data, analysis and representation, and commentary on the results.

Results

The analysis starts from the definition of school dropout and the family of indicators available to monitor the phenomenon in Piedmont. Among the indicators included in explicit dropout are those of school failure (failures, repeaters, delays) and school dropout (frequency interruption, ELET), while among those related to implicit dropout are included low performers (students with insufficient learning levels) and the INVALSI indicator. Data on activities and participants in the Piedmontese regional Educational Guidance System, monitored through a recently developed platform, as well as those made available by the Labor Force Survey, useful for identifying Neets, employment and unemployment rates of the youngest segment of the Piedmontese population, provide a context in which school careers develop and the school-to-work transition is faced in Piedmont.

The contribution examines the trend of participation in the education system and shows how, in the second cycle, the contribution of Professional Education and Training (IeFP) paths positively affects the schooling rate in Piedmont. Particularly benefiting from this increased schooling are males and students with foreign origins.

School failure indicators increase as the level of school increases, and it is from lower secondary school that greater difficulties and performance differences by gender and nationality begin to emerge. By the end of lower secondary school, one in ten students is behind. This information is important for identifying at-risk students. But it is in the first two years of upper secondary school that failures and second thoughts concentrate, with new disparities related to the type of school. School dropout, as interruption during school, records a more contained rate in lower secondary school and a higher rate in upper secondary school. Downstream dropout, the ELETs, are in a downward trend over the last five years. In 2022 in Piedmont, 11% prematurely left their studies. Young males, overall with less brilliant school performances than their female peers, have a higher dropout rate than girls: 13.3% compared to 8.5% (Nanni et al 2023). Over the long term, this rate has decreased for both genders: the launch of Professional Education and Training (IeFP) paths has provided a significant contribution to containing youth dropout, particularly for male teenagers.

The qualitative aspect of dropout is identified through the distribution of results in the 'Learning Levels' of the SNV-INVALSI Survey. Low performers are those who do not reach basic levels on the learning scale detected by the survey. At the end of lower secondary school, SNV-INVALSI results show in Piedmont a portion of students who do not reach basic levels in Italian and mathematics, respectively 36% and 39%. The learning gaps are particularly weighed down by the socioeconomic background of the students'

families. What do these data tell us about implicit dropout? First, that differences in results increase along the school path, which are articulated by gender, origin, socioeconomic status. Second, that delay negatively weighs on learning levels and, third, that differences increase depending on the course of study and the territory in which the school is located.

The action to balance disparities can pass through the containment of implicit dropout, which often precedes explicit dropout. If so, how? Starting from prevention actions, among these, an example is guidance. In Piedmont, there is a regional educational guidance system that coordinates and implements permanent orientation interventions. The actions promote initiatives to counteract school dropout through the dissemination and diversification of interventions capable of acting actively and preventively towards young people, families, and operators. The planned interventions include individual interviews (at desks in schools and in the territory) and group activities in schools, with the aim of helping young people to continue or resume their educational path, to develop orientation skills to make autonomous and informed choices, to navigate the transition phases between systems (school-professional training-university-work). During the academic year 2022/23, the regional Educational Guidance System involved about 94,200 people among adolescents, young people, and adults, in more than 11,000 activities provided. The coverage rate in the 11-15 age group stands at 35% at the regional level. In particular, the participation of 13-year-olds reaches 77%, with almost 30,300 Piedmontese adolescents reached.

On the labor market front, an in-depth analysis of the school-to-work transition of young Piedmontese highlights a competition and a progressive displacement of lower qualifications by higher ones in the youth labor market (Donato, 2023). There is a structural trend of the system that, on the one hand, pushes boys and girls to a greater investment in training, focusing on participation and countering dropout, and on the other, struggles to absorb the youth labor force with higher qualifications that come onto the market. This different evolution of the productive system compared to that related to the training of young generations has triggered a mismatch between demand and supply in accessing the labor market in Piedmont (IRES, 2019).

Keywords: Participation – School Dropout – Professional Education and Training – Educational Guidance – School-to-Work Transition

The origins of the gender wage gap: the role of school-to-work transition

Giulia Bovini – Marta De Philippis – Lucia Rizzica

Some recent literature indicates how, in modern and increasingly specialized economies, a full understanding of returns from education should go beyond considerations related to the level of education and focus on the type of human capital investments, in particular on the choice of fields of study. For instance, Altonji et al. (2012) and Altonji et al. (2016) show that university graduates face sizable earnings differences depending on their selected majors and that these disparities can be as large as the overall university–secondary school premium. At the same time, a growing body of evidence points out the increasingly important role of firms in explaining wage heterogeneity; Card et al. (2013) and Song et al. (2018) document that differences in firm pay premiums account for a large share of the observed earnings inequality in Germany and in the US.

This paper connects these two strands of the literature in the context of studying the determinants of earnings gaps between men and women. While a growing body of research indicates that both channels are crucial determinants of the existing gender pay gaps (see, among many, Zafar, 2013 and Kahn and Ginther, 2018 for field of study choices; Card et al., 2015, Sorkin, 2017, Casarico and Lattanzio, 2023b and Morchio and Moser, 2021 for firm pay premia), to the best of our knowledge, they are usually studied separately. Thanks to a unique administrative dataset assembled for Italy, we are able to consider them jointly.

We first study gender heterogeneity in field of study choices and analyze to what extent it can explain the aggregate gender earnings gap. Next, we look at within-field differences in the type of employers men and women work for and in the type of jobs they hold, exploring to what extent they account for within-field gender earnings gaps. Evaluating the relative importance of these two channels helps target policy interventions, i.e., whether to focus on the matching of female and male workers to their jobs or on the moment when boys and girls make educational choices.

We focus on gender pay gaps at the beginning of the career. While the recent literature has pointed out the importance for aggregate gender gaps of the gendered effect of parenthood (i.e., the so-called child penalty; see Kleven et al., 2019 and Kleven et al., 2020, Casarico and Lattanzio, 2023a, De Philippis and Lo Bello, 2022), some recent research also highlights the prominence of gaps earlier in the career (Arellano-Bover et al., 2023). Assessing the magnitude and the determinants of gender gaps already in the first years on the labor market is crucial, as they may widen over the career and they call for policies that are different from those that target the child penalty.

We address this question using a unique dataset newly assembled for Italy from multiple administrative registers. The data contain detailed information on both educational careers and early labor market outcomes, together with a vast array of family characteristics. We document several important facts. First, already 1 year after graduation, there is a significant gender gap in average annual labor earnings, amounting to 32% among high-school graduates and 26% among university graduates. This does not depend exclusively on differences in the intensive margin of labor supply (days or hours worked): when we focus on the daily wages of full-time workers - the closest metric to hourly pay rates that is available in our data - the gap is still 16% among high-school-educated and 13% among university-educated. 5 years after graduation, disparities remain large. Second, there are pronounced gender differences in the choice of high-school tracks and university majors: all in all, girls are more likely to select fields with lower expected average financial returns. This reflects the fact that the firms recruiting graduates from the fields of study most selected by girls tend to be smaller, have lower value added per worker and lower wage, are closer to home. The choice of the field of study is in turn a major determinant of the early career gender gap. Making use of an Oaxaca-Blinder decomposition, we show that it explains 30% of the gap in daily wages for high-school graduates and 60% for university graduates. The highest explanatory power for university graduates could in part reflect that university degrees tend to be more specialized than high-school tracks.

Third, even within majors and high-school tracks, there remain sizable gender gaps in earnings and wages upon entering the labour market. A rich array of firms' and jobs' attributes can explain 45% of this residual within-field gender wage gap at the average among high-school graduates, but only 20% of it among college graduates. Furthermore, the share of the within-field gender gap that is not explained by observable workplace characteristics is larger at the top of the wage distribution. It also emerges that, among university graduates, girls tend to graduate from majors where the unexplained component of the within-major gap is lower.

Overall our results indicate that actions should be taken at the time when boys and girls choose their field of study. Indeed, pre-market heterogeneity in field of study choices explains most of the observed gender gaps at the beginning of the career, which are a crucial driver of future disparities (Arellano-Bover et al., 2023). Also because of the different degree of school specialization embedded in the choice of high school tracks and university majors, the latter is found to be a particularly key moment for policymakers to focus on.

Keywords: gender, school-to-work transition, fields of study, firms

The (many) pillars of human capital accumulation

Lorenzo Maraviglia

The percentage of Early Leavers from Education and Training (Early Leavers from Education and Training) is often used as an indicator of the efficiency of the school system, i.e. its ability to lead all students to the achievement of a diploma. Consequently, the goal (set at EU level) of reducing this percentage below the 9% threshold by 2030 – a target which appears within the reach of our Country - is almost exclusively placed on the shoulders of schools.

Therefore, in addition to experiencing a structural reduction in public funding, schools also find itself overburdened by the problem of counteracting the trend towards a decline in student skills and learning, a phenomenon reported by the OECD-PISA surveys and, more recently, also by INVALSI.

In reality, the model of human capital accumulation underlying the notion of Early Leavers from Education and Training is more complex and presupposes the integration - or at least the simultaneity - of services provided by schools, vocational training agencies, enterprises and public administration.

ELETs are in fact those who, at the age of 18, do not have a diploma (issued by a school) or a qualification (issued by a vocational training agency) and who are not attending a course - at a school or at a training agency, possibly as part of a qualification path linked to an apprenticeship or work-training contract - aimed at obtaining one of these qualifications.

As far as our country is concerned, if in some regions (for example, in the non-insular South) the training and apprenticeship channels are effectively residual - and the entire burden of human capital training falls on schools - in others, on the other hand, non-scholastic paths are a reality and represent for many students a plausible alternative to obtaining a school diploma.

Where the training of human capital is distributed across multiple pillars, there is room for a redefinition of the tasks of the various systems, for the development of forms of specialization, of competitive but also cooperative relationships. In other words, the conditions are created for the emergence of formal or informal regulatory mechanisms that need to be investigated on a theoretical and empirical ground.

This contribution tries to reconstruct the factual framework of this scenario with the aid of statistical data (for example the elementary data from the labour force survey, ISTAT/ISFOL data on the diffusion of vocational training, INVALSI data) and administrative data (the MIM data on students attending different grades and school levels).

Particular attention will be dedicated to the reconstruction and analysis of territorial differences, in the belief that the different level of articulation of the composite system of human capital accumulation is part of the set of structural differences that contrast different areas of our Country.

Keywords: ELET, drop-out, human capital accumulation, territorial heterogeneity

Regional vocational education and training pathways: A bridge to the completion of secondary education or a shortcut to the labour market?

Silvia Duranti – Valentina Patacchini – Francesco Bogazzi

Introduction. Vocational Education and Training (IeFP) courses came into force in the 2010/2011 school year, becoming an alternative channel to upper secondary school. These pathways are a regional competence and allow the achievement of professional qualifications (with two- or three-year courses) and professional diplomas (with four-year courses) through laboratory activities, internships and, not infrequently, dual apprenticeships. VET courses are carried out by training agencies accredited by the Regions, or by Professional Schools, under a subsidiarity regime; the weight of the two components in different regions has a high variability.

Once the professional qualification has been obtained, it is possible to enter the labor market directly, to continue the training course by attending the fourth year of the IeFP courses for the achievement of the professional diploma or to continue to study in the education system in order to obtain an upper secondary school diploma.

The objective of VET pathways is therefore twofold. On the one hand, they aim to reduce early school leaving, through a shorter duration than upper secondary education courses and teaching methods more focused on laboratories and learning by doing. On the other hand, they aim to provide the local production system with skills that can be directly spent in the labour market, according to local needs.

Objective. The aim of this paper is to analyze the pathways of young people after the achievement of a regional IeFP qualification. In particular, we want to identify two different components among outgoing young people: those who decide to continue their studies in the education system and those who decide to continue their studies in the education system.

Data. The data used refer to IeFP courses held in Tuscany; in particular, the focus is on those who obtained a qualification or diploma in the period 2016–2021. The analysis is based on three different databases:

- The database of training interventions funded by the Tuscany Region, which contains the personal data (gender, age, citizenship, residence) of those who have obtained a qualification or diploma and information relating to the professional qualification obtained
- the Student Registry, which contains information on students enrolled in Tuscan schools, useful for verifying their return to the education system;
- the Labour Information System, which contains the data relating to the Mandatory Communications that employers, both public and private, are required to transmit when they make a new hire, modify or terminate an existing employment relationship.

Methodology. The analysis starts from an examination of the characteristics of the students, to be compared with those of upper secondary school students, in order to highlight differences and peculiarities. Subsequently, through the union of the database containing data on those who have obtained a qualification or diploma with the databases Student Registry and Labour Information System, the qualification or post-diploma paths are analyzed. In particular, we will estimate:

- the percentage of re-entry into the school system for the achievement of the upper secondary education diploma,
- the percentage of people entering the labour market,
- the percentage of those who work on permanent contracts,
- the number of days worked in the three years following the achievement of the qualification or diploma.

The analysis will be carried out considering the heterogeneity by: the field of qualification or diploma, provider (training agency or school), duration of the course.

Discussion. The results show that VET is in most cases a bridge to an upper secondary education degree. In fact, more than 80% of those who have obtained a VET qualification or diploma choose to continue their studies; Young people who leave the education system and find a job within 12 months of obtaining their degree account for 5% of the total, while the remaining 7% is not found in the Employment Information System of the Tuscany Region. The heterogeneity is high between different qualifications, with some offering more opportunities than others in the labour market.

The analysis shows that the value of VET courses lies mostly in their ability to offer the possibility of obtaining a degree to those young people who do not intend to continue in the education system after obtaining the middle school diploma, intercepting potential drop-outs and bringing them back into an institutional path. In addition to this function, it is important to plan the regional training offer well, so that they train professionals who can easily be integrated into the labour market.

Keywords: leFP pathways, employment outcomes, continuation of studies, regional courses

**SESSION 10. LEARNING TO LEARN IN SCHOOLS, IN ADULT EDUCATION
AND IN VOCATIONAL EDUCATION AND TRAINING**
ORGANIZER: UNIVERSITY OF SALERNO - POLYTECHNIC UNIVERSITY OF MILANO
COORDINATOR: DOMENICO BRUNETTO
19TH OCTOBER: 4.30 P.M. - 6.30 P.M. {ROOM 1 – RESEARCH 16}

Problem solving laboratory: thinking in action in disadvantaged contexts

Pier Luigi Ferrari – Annarita Monaco

The problem of problems in primary school appears to be linked to the fact that pupils give up thinking when they are given a problem. Often this attitude is linked to teaching methods adopted by teachers who believe it is useful to "facilitate students' solution paths" focusing on the product rather than on the processes. The path presented here, on the contrary, is aimed at bringing out and enhancing the strategic, representative and communicative intuitions of the pupils engaged in an educational setting characterized by the resolution of problems specifically designed to encourage discussion and the development of argumentations in small groups, but also means to help pupils to feel "good solvers", which is crucial in order to gratify them socially and create the motivation which in turn activates the will to engage themselves in the activities.

Keywords: collaboration, discussion, strategies and spontaneous representations, semiotic systems, metacognition

From classroom teaching practice to significant linguistic experiences. How to support mathematics learning in heterogeneous primary school classes

Emanuela Atz – Pier Luigi Ferrari – Giovanna Mora

The development of paths with objectives common to multiple disciplinary fields (GISCEL 1975) is the focus of the activities aimed at supporting learning in mathematics in the primary school of a Comprehensive School in the province of Bolzano. Starting from the students' linguistic repertoires, the longitudinal path of knowledge organization is developed both through the argumentation in mathematics problems and

through the exploration of the language system and the related disciplinary languages (Ferrari 2021) in a context of endogenous multilingualism - teaching in Italian and German - and exogenous - languages of the students (Iannaccaro 2019).

Argumentation is an end and a means for the development of linguistic and mathematical skills, leveraging the different linguistic repertoires present in the classroom and proposing texts in the three languages of instruction, in English and in some of the students' L1s.

The strategies used (translation, peer tutoring, in-depth study in L1, ...) for the comprehension and production of texts can support learning with significant linguistic experiences. The latter strengthen self-esteem and awareness of one's linguistic and mathematical competence.

Keywords: plurilingualism, mathematical problem solving, interdisciplinarity, argumentation

Co-disciplinarity to reduce early leaving school: math education in a innovative and integrative dialogue with other disciplines

Antonella Montone – Michele Giuliano Fiorentino

This research addresses the problem of early school leaving in Italy and Europe, especially in lower and upper secondary schools, analyzing the different causes of the phenomenon which is recorded through the increasingly high number of school dropouts, since primary school.

Particularly relevant is the students' loss of meaning and perception of uselessness of mathematics in everyday reality.

The phenomenon of early school leaving is more accentuated in vocational schools, where disciplinary specificity is linked to increasing professionalism for future work. Anyway mathematics curriculum is the same as in other upper secondary schools, even if it's required mathematical tools suitable for vocational subjects.

Furthermore, while on the one hand mathematics teachers often do not know which mathematical contents are useful for solving applied problems in other disciplinary fields, on the other hand teachers of other disciplines need more detailed information regarding mathematics to address some of their characteristic topics.

The general aim of this research is therefore to give meaning (Wake G., 2014) to Mathematics with respect to other disciplines and vice versa to give meaning to the problematic situations of other disciplinary fields, in which Mathematics is necessary, designing co-disciplinary educational paths, with the greater involvement of students at risk of early leaving school. It is therefore considered appropriate to propose a continuous training project for teachers of mathematics and other disciplines in which teachers are involved in the co-designing of co-disciplinary teaching paths.

A framework that integrates well and allows for more detailed analyzes in this context is the Theory of Semiotic Mediation (Bartolini Bussi & Mariotti, 2008). The students' activity with artefacts, in this context, coming from other disciplines allows them to develop a multiple semiotic potential: the one linked to mathematical meanings and the one linked to the characteristics of other disciplines.

The interactions between different disciplines, in this scenario, require the need to overcome the different approaches proposed in the literature (multi-disciplinarity, trans-disciplinarity, interdisciplinarity), considering the co-disciplinary approach as the most adequate one (Blanchard-Laville, 2000).

In accordance with the aim of the research study, the analyzes of the experimental activities conducted in class seem to show how the co-disciplinary approach has favored the development of mathematical concepts linked to other disciplines. In the discussions, the evolution of the artefact signs towards the mathematical ones and towards the disciplinary ones is completed through the request for cross-interpretation, and this allows the construction of the relationship of meaning between the two disciplines.

Furthermore, active participation in the proposed activities was observed, especially among students reported at risk of dropping out of school.

Keywords: early leaving school, co-disciplinarity, Math Education, Boundary objects artefacts

SESSION 9. INEQUALITY AND EDUCATIONAL OUTCOMES (GENDER GAP)

ORGANIZER: HARVARD AND BOCCONI UNIVERSITY

COORDINATOR: AGOSTINO PORTERA

19TH OCTOBER: 4.30 P.M. - 6.30 P.M. {ROOM 2 – RESEARCH 17}

The gender gap among Italian students: PISA 2022 mathematics results

Angela Martini – Maria Teresa Siniscalco

Introduction

In the latest edition of the PISA survey, Italian boys scored 21 points higher in the mathematics test than their female peers, recording the widest gap among all 37 OECD countries, where the average difference is 9 points. The figure itself is not new: in all previous rounds of PISA, in Italy males have always achieved significantly higher results in mathematics than females (with the sole exception of 2008, when the difference, even then in favour of males, was not significant). Although the gender gap in mathematics is greater in Italy, it is also shared by the other OECD countries: among the 37 participants in PISA 2022, 34 saw males outperform females in mathematics mostly in a statistically significant way, while there were only three countries in which females scored higher than their peers: Norway, Slovenia and Finland; in the first two, however, the difference is small and not significant and only in the third did females score significantly higher than males by 5 points.

The explanation generally given for the lower ability of girls in mathematics, attested by the PISA survey and, to varying degrees, by other international surveys on achievement levels, but also by the national INVALSI surveys, traces it back to gender stereotypes that condition girls' interests and choices and are also said to be at the root of their lower performance in this discipline. Against this hypothesis and, more generally, against the 'false theory' of women's inferiority in mathematics, the sociologist Luca Ricolfi has recently spoken out. On the basis of an analysis of the academic results of females and males in lower secondary school and in its final examination, he argues instead that females perform better than males in all subjects, including mathematics.

Research object and objectives

The objectives of our research are the following:

- 1) to analyse in detail the gender gap in PISA 2022 mathematics test, comparing the scores of males and females in the content and process subscales as well as on the overall scale;
- 2) to quantify the percentage of males and females at the various levels of the PISA proficiency scale and to investigate whether and how the size of the gender gap varies according to the macro-area where the students reside and the type of school they attend;

- 3) to compare, on the one hand, the school grades obtained in mathematics by boys and girls at the same level on the proficiency scale and, on the other hand, the average grade and the average PISA score of boys and girls in each type of school;
- 4) to examine whether and how much males and females differ as to a set of individual variables having a relation with the mathematics performance;
- 5) to investigate whether and to what extent - given the unequal distribution by gender in the different types of upper secondary education, characterised by a different mathematics curriculum - the gender gap decreases when controlling for the type of school and other variables affecting mathematics results in which males and females differ.

Data used

The data used were the PISA scores in mathematics of the students in the Italian sample attending the second grade of secondary school (modal class in Italy) who took both the PISA test and the INVALSI mathematics test in 2022, corresponding to 72% of the unweighted and to 82% of the weighted data of the total sample. This allowed to integrate the PISA database with other information, such as grades in mathematics, not present in it.

Method

The data were disaggregated according to the objectives of the study and the scores of males and females compared. Regression analyses were conducted to test the weight of a number of variables, starting with school type, on the PISA mathematics score and to estimate how the gender gap varies all other conditions being equal.

Results

Analyses show that males outperform females on both the overall mathematical proficiency scale and the content and process subscales. Disaggregating the data by macro-area and school type, it can be seen that the gender gap, compared to the national average, increases in geographical areas and schools where the overall result is higher and decreases in those where it is lower. Moreover, it is not the same across the whole proficiency scale but widens at the highest levels and narrows at the lowest, as shown by the proportion of males and females in the six levels of the scale, where females are in the majority at the lowest levels and in the minority at the highest.

The profile of girls, from a social and psychological point of view, differs from that of boys particularly as to some respects: firstly, in all types of school, the socio-economic-cultural status of females is lower than that of males, especially in academic schools, while, at the same time, the effect of the Escs index on the mathematics achievement is stronger for males than for females. Secondly, females show higher levels of anxiety about mathematics and lower self-confidence in their ability to succeed in this subject than males. When controlling for Escs and the type of school attended, in order to take into account the different importance attached to mathematics by the various branches of upper secondary school, the difference between the two sexes diminishes considerably, and it diminishes further when also anxiety, and sense of self-efficacy are controlled.

Comparing school grades in mathematics with scores on the PISA test shows that, at the same level on the proficiency scale, females are overestimated in comparison with males, and the same is true if one compares the average score obtained in each school by boys and girls with their average score. This also explains why, when looking at teachers' evaluations, as in Ricolfi's analysis of lower secondary school results, females appear superior to males in all subjects, including mathematics

Keywords: PISA 2022 score in mathematics; gender inequalities; school grades and tests

Unmasking gender gaps in reading and mathematics competences: the role of time and space

Emanuele Fedeli – Andrea Pietrolucci

Introduction

The literature on gender gaps in educational achievement consistently reports that girls outperform boys in reading test scores, while the opposite is true for mathematics test scores. This pattern has been broadly confirmed in different contexts by several comparative studies (Marks, 2008; Stoet and Geary, 2013; Breda and Napp, 2018). Despite some cross-country variation, the gender gap is reported in the vast majority of

countries, with an average gap of around 0.15 standard deviations in favour of boys in mathematics and a larger gap of 0.35 standard deviations in favour of girls in reading (Breda and Napp, 2018). Contributions examining the timing of gender gaps document that differences in reading and mathematics achievement follow different trends. While the boys' gap in reading is found to emerge as early as in the kindergarten and narrows between grades 3 and 8 (Robinson and Lubienski; 2011), the average gap for girls in mathematics is negligible in kindergarten and gradually widens across grades in primary school (Penner and Paret, 2008; Robinson and Lubienski, 2011). Fryer and Levyyt (2011) report that the gender gap in mathematics is likely to increase by 0.2 standard deviations between grades 1 and 5. Other contributions have focused on the role of the contextual level in explaining the gender gap. In a cross-country comparative analysis, Ma (2008) documents that the school level is likely to play a key role in the production of gender gaps. The author reports that in more than half of the countries studied for the gender gap in reading and in more than a third of the countries studied for the gender gap in mathematics, the gender gap is likely to vary considerably between schools.

The majority of contributions, however, examines the temporal and spatial variation of gender gaps as separate dimensions. Yet, to get a deeper understanding of gender gaps dynamics, it is important to account for their variation over grades while also evaluating the relative importance of different contextual levels in which such gaps arise. Moreover, while most studies focus on US or adopt a cross-sectional comparative perspective, empirical evidence on European countries and Italy, in particular, is more limited. Contributions adopting Italy as a case study show that the gender gap in mathematics is likely to be comparatively large and to widen over grades. Contini et al. (2017) report a female penalty in mathematics achievement of 0.16 standard deviations at grade 5, which increases to 0.28 standard deviations at grade 10. Giofr  et al. (2020) document that gender gap in mathematics is also likely to vary across regions, with northern regions showing larger gaps compared to southern regions. This north-south gradient, however, is not observed for the males' gap in reading achievement.

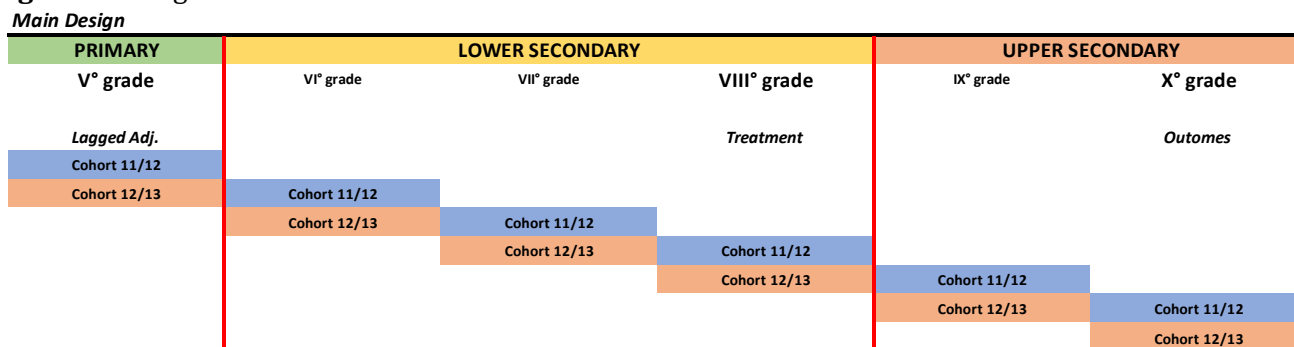
Theoretical estimand

This paper has two main goals. Firstly, it aims to contribute to the literature by identifying trends in gender gaps in reading and mathematics competencies in Italy across students' entire educational trajectories. Additionally, we address the spatial dimension of gender inequality, which has been underexplored in the literature that predominantly focuses on national-level trends. Italy serves as an intriguing case study due to its historical delay in forming a national state compared to other European countries, leading to the persistence of profound territorial inequalities along three main axes. These axes include the typical North-South divide, the disparity between internal and external areas, and the divide between the Tyrrhenian and Adriatic regions. It is theoretically reasonable to consider the role of local contexts when studying gender differences in the educational system.

Research Design and methods

Our theoretical question aims to describe a simple skills gap between boys and girls, how this gap varies over time across school grades, and how it differs depending on geographic contexts. Given the purely descriptive nature of this gap, we consider individual characteristics such as ethnic background and socio-economic origin as simple checks or side analyses in understanding what happens in the gender gap over time across different spaces. Currently, our research design (Figure 1) involves following two cohorts over time from the 5th grade of elementary school to the 2nd grade of middle school, including the 3rd grade of middle school, thus covering the three main phases of the Italian school system. In the subsequent phase, we expand our scope to also include the 2nd grade of elementary school and the final year of high school. For the methods we use multilevel models.

Figure 1: Design



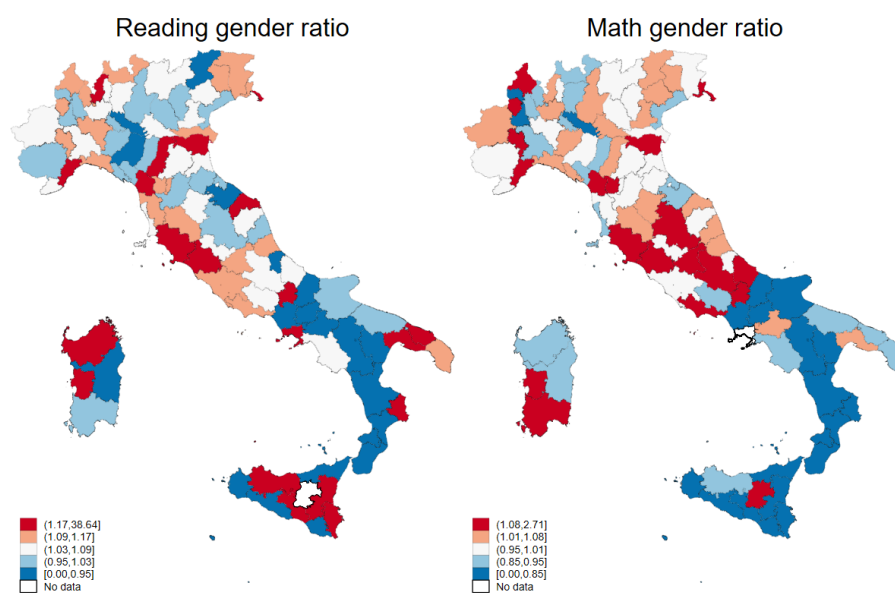
Data

To investigate gender-based competencies disparities across regions in Italy, we utilize data from the INVALSI-SNV dataset along with geographical classifications provided by ISTAT for Italian provinces, regions, and labor market areas (commuting). Specifically, we track two cohorts from the academic years 2011/2012 and 2012/2013, employing standardized tests and basic socio-demographic characteristics such as gender, ethnic origin, and socio-economic background. In a subsequent phase, we aim to expand the dataset by constructing indicators that measure the exposure of local contexts to school investment policies or, conversely, austerity measures.

Preliminary results

Preliminary findings indicate three primary trends. Firstly, there is a gender disparity favoring males in standardized tests, both in reading and mathematics. This gap tends to widen over time, notably during the transition from middle school to high school. In the subsequent phase and in preparation for the conference, we will systematically examine these disparities at a geographical level. As an initial step in our geographic analyses, we present a provincial-level cross-sectional analysis, where we compute the ratio of male competence to female competence for each subject. This ratio takes a value of 1 if competence is equal between the two groups, increases if males outperform females (resulting in the province being shaded in red), and decreases if females outperform males (resulting in the province being shaded in blue). As evident, there are geographical clusters warranting further exploration. In fact, it appears that in the southern regions, females perform better than males in both mathematics and Italian, while the situation is more varied in the central and northern regions.

Figure 2: Gender ratio by subject



Keywords: gender gap; time; space

The progress of women in science: a longitudinal population study on gender differences into university career in stem

Patrizia Giannantoni – Patrizia Falzetti

CONTEXTUAL FRAMEWORK

A major interest in the most recent years has been devoted to STEM disciplines, for the increasing job demand and higher wages associated to this field of studies. In this perspective particular attention has been directed to gender differences into STEM studies as responsible of the widening gender gap in work career and life opportunities.

The data about Italian situation show a different composition by gender at University: among STEM graduates the male component is higher, reaching 59%, of students while among non-STEM graduates women prevail (they are almost two out of three).

Preliminary research on the same joint dataset investigated the impact of performance at INVALSI test in Math for students in the last year of upper secondary school (so called grade 13) on the propensity to enrol in a STEM course at university the following year. Findings showed that there was an important influence of cultural status of the family, measured by educational attainment of the mothers, and an influence both of teachers marks and performance at INVALSI test in math, with a slightly higher impact of the former, on the probability to choose a highly scientific course of study at university. A second line of research has investigated career into STEM courses according to gender, to provide an answer to question such as “do boys perform better in terms of scores and or speed of career in highly scientific disciplines at university?”. First results in this line of research exhibit better achievement for female students compared to their male peers in STEM university courses at university in terms of “survival” into STEM, marks and also speed of career. Survival rate into STEM at 3 years from enrollment at national level is 91% for girls and 88% for boys. In terms of achievements, average number of credits (CFU) gained each year is consistently higher for girls (54 vs 50).

With this contribution we aim at provide a more comprehensive analysis of scientific career at university in a gender perspective, taking into account the effect of a wide set of socio-demographic and contextual characteristics.

OBJECTIVES

The aim of this contribution is to deepen the study of factors predicting a STEM career, by looking more to the university success in terms of exams and credits gained, rather than merely to enrolment in different university courses. Some questions of interest are: do the mathematical skills achieved in high school count in a different way for a highly scientific career of boys and girls? With the same mathematical skills, how important are context factors, such as the socio-economic background of the family of origin, and / or the evaluation by the teachers, the country of origin in proceeding into a STEM path at university? Once inserted in the STEM University courses, do the girls have an equal performance with their male colleagues or even higher?

DATA

We used as a base the original dataset already created ad hoc for the previous steps of this research project, i.e. a dataset built on the combination of different data sources: MIM (Ministry of Education and Merit), University Register of students and INVALSI (National Institute for the Evaluation of the Education and Training System) data. This joint data-source allowed us to follow different cohorts of students in a longitudinal perspective that covers different school grades (INVALSI data), with an enrichment of information from Ministry data, and is eventually linked to university enrolments and to the number of educational credits obtained for each year of university course.

It will therefore be possible to have data about students' performance, in all subjects (Italian, mathematics and English) and throughout the whole school career, as well as to have information about the student's family background (e.g. parents' educational level, socio-economic status of the family, country of origin, language spoken at home).

Furthermore, data will be available at the University level and with a very specific course classification, which can be recode into a binary variable (STEM / non-STEM).

METHODS

Analyses about the development of career of students into scientific tracks will be carried on by looking at similarities and differences between boys and girls in the speed and the average mark they obtain in scientific courses.

More in depth studies would be related to regression analyses and survival analysis.

On the one hand, regression models allow the estimation of the weight of the various factors in determining the progression into the university career, as measured in terms of credits gained.

Having a wide set of socio-demographic and contextual characteristics it would be possible to estimate the impact of the different characteristics on the speed and success in progression through the university career for boys and girls, keeping as a key variable that of "mathematics skills" measured as a continuous variable (WLE score) or as a categorical variable based on the WLE score levels (ordinal scale from 1 to 5) during the INVALSI test of level 13. On the other hand, survival analysis, could improve the understanding of STEM

promoting factors, highlighting characteristics of those students who survive compared to those dropping out. Of a major interest would be at this point, to observe whether the scientific career for boys and girls has significant differences.

EXPECTED RESULTS

After giving a first descriptive analysis about gender differences in academic performance and university choices, we aim with this contribution to analyze more in depth the characteristics of students having more success into a STEM discipline at university, keeping constant the interest on gender in the analyses. We intend to use regression models, to study the impact of different factors, both individual and contextual, on the university success into STEM discipline. Particularly, we estimate the weight of the mathematical skills, which we assume to have an expected and legitimate influence on the choice of academic orientation, compared to the weight of the "social" factors that instead lower the propensity of girls towards the more scientific-technological university paths.

Previous results showed that math skills are a strong predictor of choosing STEM degree: students with very high mathematics performance, i.e., Level 5 at INVALSI test, during the last year of high school are 5 times more likely to enroll in a STEM-university degree compared to students who have only reached the lowest level. However, gender remain a characteristic that hinder the access to STEM courses.

Despite the barrier in access, when we look only to STEM students at university and compare girls and boys, the former show better performances in terms of "survival" into STEM, marks and also speed of career. Survival rate at 3 years from enrollment into STEM is 91% for girls and 88% for boys. Average number of credits (CFU) gained each year is consistently higher for girls and final mark at university degree is on average higher for girls both in humanities and science.

With this contribution we will explore more in depth factors affecting success in an academic highly-scientific career in a gender perspective.

Keywords: STEM, University, career, gender gap, math, INVALSI, predictive factors

AUTHORS

Giovanni Abbiati is currently research fellow at the department of economics and management of the University of Brescia. His main research interests are related to methodology of the social sciences, educational inequalities education and evaluation of public policies.

Tommaso Agasisti is a Professor at the School of Management, Politecnico di Milano. He also serves as the Associate Dean for International Relations and Quality at PoliMi GSOM - Graduate School of Management. His research interests lie in the field of Public Economics and Finance, Public Management and Policy, and Public Administration, with a particular focus on schools, universities, and local governments.

Gabriella Agrusti, Full professor of Educational Research. Current research topics: evaluation models and procedures in school contexts; Civic and citizenship education; international large-scale assessment surveys; teacher education. Co-directed by Cadmo. International Journal of Educational Research (Ed. FrancoAngeli). Scientific manager of the CRESPI Interuniversity Center.

Carmen Aina is a Professor at the School of Management, Politecnico di Milano. He also serves as the Associate Dean for International Relations and Quality at PoliMi GSOM - Graduate School of Management. His research interests lie in the field of Public Economics and Finance, Public Management and Policy, and Public Administration, with a particular focus on schools, universities, and local governments.

Sule Alan is a Professor of Economics at the European University Institute, Florence and an Adjunct Professor of Economics at Bilkent University, Turkey. She is a board member of the Innovation in Government Initiative and co-chair of the Humanitarian Protection Initiative at the Abdul Latif Jameel Poverty Action Lab (J-PAL), a research fellow at CEPR, an affiliate at the Innovations for Poverty Action (IPA) and CESifo. Her research interests include education, migrant integration, gender and household behavior. She is a co-editor of the Economic Journal (2021-present).

Monica Alberti obtained her Doctorate in History, Philosophy and Science Teaching at the University of Cagliari in 2008. Currently a primary school teacher. He collaborates as a trainer in CRSEM activities and with the USR in the workshops for new hire.

Massimo Anelli is Associate Professor of Economic Policy at Bocconi University. He is also Research Fellow at the CESifo, IZA, Dondena, Baffi and Fondazione Rodolfo De Benedetti research centers. His studies focus

on the Economics of Education and the impact that technological progress has on labor markets, politics and demography.

Giovanni Antonini is a master student at the Catholic University of the Sacred Heart in Milan in Public Policy. He holds a three-year degree in Economics, Markets and Institutions from the University of Bologna. He did a curricular internship at the office of Associate Professor Gianluca Argentin where he did a thesis on teacher precariousness and inequalities in student learning. (New evidence from INVALSI data, 2017-2022).

Carmela Aprea holds the Chair of Business Education - Design and Evaluation of Educational Systems at the University of Mannheim and is the founding director of the Mannheim Institute for Financial Education (MIFE) (together with Prof. Tabea Bucher-Koenen). She is also an associate professor at the Department of Taxation at the University of South Africa and a member of numerous national and international committees, such as the scientific advisory board of the OECD's International Network on Financial Education (INFE), the expert committee of the OECD's PISA Working Group on Financial Literacy, and the OECD's Global Money Week working group. Her work focuses on the development and evaluation of innovative, also digitally supported, learning modalities in the fields of financial, economic education and sustainability education. She also works on the promotion of resilience in learning and working contexts. She is currently conducting research, for example, in interdisciplinary project networks on financial concerns and fears in times of multiple crises and on understanding and attitudes towards retirement provision in Germany.

Massimo Attanasio is a Full Professor in Social Statistics at the University of Palermo, Italy. He is currently PI of the national research project "From high school to job placement: micro-data life course analysis of university student mobility and its impact on the Italian North-South divide".

Emanuela Atz is a secondary school teacher and teacher trainer since 1992. Her fields of interest are language learning/teaching, the plural approaches and minority languages. She currently works at the Language Centers in the province of Bolzano to support the integration of pupils with migration background into school.

Davide Azzolini is a research fellow at the Research Institute for the Evaluation of Public Policies of the Bruno Kessler Foundation (Italy) and an affiliated scholar at the Urban Institute (Washington DC, US). He has carried out research in the areas of student achievement, migrant education, financial aid and education technology.

Paolo Barabanti, PhD in Evaluation of Educational Processes and Systems. He has been a teacher at primary school and lecturer in Sociology of Education; now he is a researcher at INVALSI. His main studies are about learning outcomes through standardized tests, high achieving students, multicultural schools and the use of INVALSI tests data returned to schools.

Graziella Bertocchi is a Professor of Political Economy at Unimore and President of the Einaudi Institute for Economics and Finance. She is a Research Fellow at CEPR and IZA. She coordinated the GEV13 for the VQR 2011-14. Her research activity focuses on topics such as macroeconomics, institutional economics, and political economy.

Alice Bertoletti is a researcher at the Joint Research Center of European Commission in Seville (Spain). She holds a PhD in Management Engineering from Politecnico di Milano (Italy), with a dissertation on the evaluation of the impact of higher education systems on regional economic development. Her research interests involve the field of economics of education and the use of quantitative methods to evaluate educational policies.

Luca Bonacini is Assistant Professor (RTDa) at the University of Bologna and is a research fellow at Global Labour Organization. His main research interests concern income distribution and inequality of opportunities, with a focus on education and gender issues.

Alessandra Boscolo is a researcher in Mathematics Education at the Department of Mathematics of the University of Genoa, and a CRESPI collaborator. After graduation in Mathematics, she did an international double-degree PhD at the LUMSA University in partnership with the Australian Catholic University, with a dissertation in Mathematics Education.

Giulia Bovini works at the Households and Labour Market Division of the Structural Economic Analysis Directorate of the Bank of Italy. After graduating in Economics from the University of Bologna, she obtained her MRes (2015) and PhD (2019) in Economics from the London School of Economics and Political Science. Her main research interests concern labour economics, economics of education and public economics.

Hugo Armando Brito Rivera, PhD in Psychology of Interaction, Socialization and Communication (University of Rome La Sapienza). Researcher in the project “Learning to learn in Italy, Europe and Latin America, coordinated by INVALSI. Member of the Mexican Network of Researchers in Organizational Studies (REMINEO) and of the Inter-American Society of Psychology (SIP). Independent Researcher, Research Professor at UNIPUEBLA and International Consultant at UNICEF.

Paolo Brunori is Associate professor of Economics at University of Florence and Associate Professorial Research Fellow at the London School of Economics. He was previously assistant professor at University of Bari and research assistant at the European University Institute. His research activity focuses on methodological aspects of inequality measurement. Recently he acquired extensive experience in using machine learning algorithms to complement standard econometric tools.

Michele Cardone holds a degree in “Statistics for demographic and social sciences” and a Master (I level) in ‘Statistics for the management of information systems’ (Università di Roma “La Sapienza”). Working for INVALSI since 2004, member of the Statistical Service since 2010, mainly involved in the analysis of the school tests data and in the management of the annual data return to schools.

Francesca Carimando is a student about to finish a two-year master’s in Politics and Social Policy at the University of Bologna. Her research interests include ethnic inequalities and social policies.

Michela Carlana is an Associate Professor of Public Policy at the Harvard Kennedy School and a Visiting Associate Professor of Economics at Bocconi University. She works on topics related to inequality and education, with a focus on gender and immigration. She is also a Faculty Research Fellow of the NBER, Faculty Affiliate of J-PAL, and a Research Affiliate of IZA, CESifo, and CEPR.

Diego Carrasco is a full time academic researcher at the Centro de Medición MIDE UC, at Pontificia Universidad Católica de Chile. His research is devoted to the study of contextual effects and learning environments comparisons, using international large scale assessment studies (ILSA), specially for non-cognitive outcomes in civic education. He has served as a methodological expert for instrument design, scoring and/or results production using context background questionnaires in ILSA studies for UNESCO-OREALC, UNESCO-UIS, UNICEF-MENA, and the OECD. He received the IEA Richard M. Wolf Memorial Award in 2020, for its journal article “Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America”. He often participates as instructor of workshop on statistical modelling and results production with secondary data of ILSA, and as a lecturer of quantitative research methods.

Clelia Cascella, Before joining INVALSI as a researcher in 2014, she held a Ph.D. in Research Methodology and a Ph.D. in Economics, and then qualified as an Associate Professor in Mathematics and in Sociology. She has spent several years abroad, mostly in the UK, first as a Marie Curie Fellow and then as a Lecturer in Social Statistics.

Ana Castro Zubizarreta, Professor of Theory and History of Education at the University of Cantabria (Spain). Her research interests focus on Early Childhood Education, learning to learn competence and educational transitions. She has published nationally and internationally and has extensive experience in university management.

Zhijun Chen is a PhD candidate at the Department of Education at the University of Bath, UK. She also works with Dr. James KO at the Education University of Hong Kong, focusing on teaching assessment in multiple cultures. Her research interests include educational effectiveness, education inequality, and leadership.

Zaneta Chonteva, PhD, works at the Bureau of Education Development, Ministry of Education and Science, Republic of North Macedonia, as a advisor for research and research projects. Also, Chonteva is engaged as an external associate at the American University College-Skopje for the courses: Introduction to Statistics for Psychologists and Educational Psychology.

Maria Chiara Cibien, PhD student in Mathematics Education at the University of Ferrara, whose main research interests concern the use of large-scale assessment in the context of teacher education (pre- and in-service) and the use of technology in the teaching-learning process of mathematics.

Lucia Cirina, obtained her Doctorate in History, Philosophy and Science Teaching from the University of Cagliari in 2008. Currently a support teacher at primary school. She collaborates as a trainer in CRSEM activities.

Adam Coates is a PhD students at the University of Bath. His research focuses on mathematics learning and achievement, and the use of cognitive diagnostic assessment and modelling. His research interests also include the applied use of philosophical assumptions in social science research.

Silvia Columbu is Assistant Professor in Statistics at the University of Cagliari. Her research interests include methodological and applied aspects of the statistical modeling. In particular, her work has been recently focused on extensions of mixture and latent class models for complex data structures, including multilevel and network data, and on the definition of discrepancy measures for hypothesis testing in the bayesian framework.

Dalit Contini is professor of social statistics at the University of Torino. Her research interests comprise social and gender inequalities in achievement and attainment, educational choices, the role of school systems, impact evaluation of social and educational policies. Her work has been published in many international journals including European Sociological Review, Sociological Methods and Research, Social Science Research, Economics of Education Review, Higher Education, Journal of Youth Studies.

Andrea Marco De Luca, Graduated in Mathematics in 1980 from La Sapienza University of Rome, since 1981 he has worked in ICT companies. Since 2017, he has been involved in the UNI.CO research project for analyzing the transition to employment of graduates. In the following years, he collaborated with other Italian universities that applied the analysis model to their own context.

Marta De Philippis has been working as a researcher at the Bank of Italy since 2014. After graduating in Economic and Social Disciplines from Bocconi University, she received her Ph.D. from the London School of Economics in 2016. Her main research interests are in the economics of labour and education.

Kristof De Witte is full professor at KU Leuven and professor by special appointment at Maastricht University. He is program director of the Educational Master in Economics, and director of the faculty research group 'Leuven Economics of Education Research' (LEER). The research interests of Kristof De Witte comprise education economics, performance evaluation, and political economy.

Enrico Deleo is a Postdoctoral Research Follow at the Department of Neuroscience, Imaging and Clinical Sciences (University 'G. d'Annunzio' of Chieti-Pescara). He holds a Master's degree in Developmental Psychology and a PhD in Social, Developmental and Educational Psychology at Sapienza University of Rome. His fields of research are Theory of Mind, School Climate and School Psychology.

Paola D'Elia is Research Fellow at the Department of Neuroscience, Imaging and Clinical Sciences (University "G. d'Annunzio" of Chieti-Pescara). Postgraduate in School Psychology, her research interest focuses on digital learning and school climate, and how to promote inclusive education and effective learning.

Melisa Diaz Lema is a Junior Assistant Professor at Politecnico di Milano, Department of Management Engineering. Her research falls within the intersection of public management, economics of education, and digital innovation. Her main research interest focuses on the usage of data analytics for decision making in the public sector.

Alessio D'Ignazio is a researcher at the Bank of Italy, Directorate General for Consumer protection and financial education. His main research field covers financial education and program impact evaluation.

Luisa Donato, holds a PhD in Comparative Social Research and Sociology and is a researcher at IRES Piemonte. She is the Co-Head of the Observatory on the educational system of Piedmont (SISFORM), alongside Dr. Carla Nanni. Donato is involved in studies, analysis, and evaluation of learning outcomes, regional Guidance System, professional needs, and school-to-work transitions.

Silvia Duranti is a researcher at the Istituto Regionale per la Programmazione Economica della Toscana (irpet) which supports the region in developing and evaluating public policies. Her main areas of interests are the analysis of the education system and the evaluation of the effectiveness of labour and education policies.

Margherita Emiletti, she is Researcher at INVALSI - international assessment team. She has been a member of the OECD PISA working group since 2001. Since 2016, she is the National Research Coordinator for IEA PIRLS. For PISA 2022, she collaborated on translating the Creative Thinking test items, worked on the revision of international coding guides, and coordinated the national group for test corrections.

Sofia Ercolanoni is Research fellow at the Department of Sociology at the University of Milano-Bicocca, where she works on quantitative methods to study educational poverty, territorial disparities, and specifically digital educational poverty.

Nurullah Eryilmaz is a Research Analyst at the IEA, Hamburg. He holds an MA degree from the University of East Anglia. His research interests focus on the domains of comparative education, sociological theory and psychometric analysis.

Patrizia Falzetti, Technologist Director, she is the Head of INVALSI area of Evaluation Research, of the SISTAN Statistical Office and of the INVALSI Statistical Department, which manages data acquisition, analysis and return about both national and international (OCSE and IEA) surveys on learning. She coordinates and manages the process of returning data and statistical analysis to every school and to the Ministry of Education and Merit.

Emanuele Fedeli Emanuele Fedeli is a Postdoc research fellow at University of Trento within the INEQUALITREES project funded by Fondazione della Compagnia di San Paolo. He defended the thesis at University of Trento. Previously, he studied at LUISS University and Collegio Carlo Alberto.

Pier Luigi Ferrari, He has served at various universities (Genova, Catania, Torino, Piemonte Orientale). He held a position as full professor of Mathematics Education at the University of Piemonte Orientale from 2007 to 31 October 2022, the date on which he quitted service. After graduating he carried out studies in Logic and Foundations of Mathematics. Subsequently he oriented his interests towards research in Mathematics Education, dealing with various topics, such as problem solving, the high school-university transition and the use of technology in mathematics education. Since 1998 he has been working on the role of language in learning mathematics, which has been his main research topic since then. He is the author of several publications including the volume *Educazione matematica, lingua, linguaggi. Costruire, condividere e comunicare matematica in classe* (published by UTET).

Federica Ferretti, Researcher in Mathematics Education at the University of Ferrara, Department of Mathematics and Computer Science. Her main research interests are didactical contract in all school grades, formative assessment and the formative use of standardized assessments. He has been conducting university and teacher training courses for all school grades for many years.

Michele Giuliano Fiorentino, Researcher in Mathematics Education (SSD MAT/04) at University of Bari. Among the main research topics are: the contribution of mathematics teaching to reduce early leaving school; innovative teaching technologies; teacher training and professional development for mathematics teachers.

Vincenzo Galasso is full Professor of Economics at Bocconi University. Director of the Analysis in Pension Economics unit of the Baffi-Carefin Center, Research Fellow at the Center for Economic Policy Research (CEPR), at the CES-Ifo, at IGER, and at the Center for Economic and Political Research on Aging (CEPRA) at the University of Lugano.

Letizia Gambi is a doctoral researcher at the research group 'Leuven Economics of Education Research' (LEER) at KU Leuven. Her research focuses on the organization of schools, students' performance evaluation and educational policy evaluation.

Giada Garbini graduated in Economics and Development from the University of Florence in 2023, and she is now about to complete her Master in Economics at the Collegio Carlo Alberto in Turin. She is yearning to enter a PhD program, and her research interests include Economics of Education, Economics of Gender, and Family Economics.

Selene Ghisolfi is Economics Assistant Professor at Catholic University in Milan, and faculty affiliate at LEAP, Laboratory for Effective Anti-poverty Policies, Bocconi. She obtained her PhD in Economics at the Institute for International Economic Studies, Stockholm University. She works in Development and Behavioural Economics, with a focus on education, agricultural markets, education, fairness preferences and bargaining.

Patrizia Giannantoni, holds a PhD in statistics and demography from a multinational program with University of Rome and Lund and Max Plank Institute in Rostock. She has worked on psychometric evaluation of developmental tests in collaboration with CNR and University of Parma and participated in research projects on migration as research fellow at University of Naples. Since 2017 she has joined the statistical office of INVALSI, keeping her research interests and educational inequalities, particularly based on migration background and gender.

Katharina Hartinger, PhD, postdoctoral researcher at Johannes Gutenberg University Mainz. Her research is about the tools we use to navigate decisions related to education, modern technology, and sustainability: from the role of culture in life-long learning to curiosity and future thinking in pro-environmental behavior.

Gabriele Iannotta, PhD, postdoctoral researcher at the Politecnico di Milano. He is a behavioral and experimental economist, and his current work is mainly focused on studying how attitudes and behaviors influence, and are influenced by, different educational practices.

Tim Kaiser, Full Professor of Economics at the University of Kaiserslautern-Landau (RPTU). Affiliated with CESifo, IZA Institute of Labor Economics, Bonn, and the Mannheim Institute for Financial Education (MIFE), a joint initiative of the University of Mannheim and the Leibniz Centre for European Economic Research (ZEW).

Elia La Ferrara is a Professor of Public Policy at Harvard Kennedy School. She is President of the Econometric Society and Program Director of Development Economics for the Center for Economic Policy Research (CEPR). Elia received a PhD in Economics from Harvard University in 1999.

Francesca La Russa, She got her Ph.D. in Foreign Language Teaching at Bordeaux MONTAIGNE University and Roma Tre University. She works as a Researcher at INVALSI and is Associate member of the *Laboratoire Cognition, Langues, Langage, Ergonomie* (CLLE, UMR 5263) of Toulouse - Jean Jaurès University.

Sara Lamboglia works as analyst at the Directorate general for Consumer Protection and Financial Education at the Bank of Italy. After receiving her PhD in Mathematics at the University of Warwick (UK), she was a researcher in Mathematics at Goethe University in Frankfurt and at the ICERM institute located at Brown University (USA).

Beti Lameva, PhD is head of the Sector for Exams, IT and Research, National Examination Centre, Republic of North Macedonia. She is responsible for organizing state matura and national assessments, also she is the National Project Manager for PISA, TIMSS and TALIS studies. She has more than 20 years of professional experience in research in educational.

Federica Laudisa holds a sociology degree and a master's in public policy analysis. She is a researcher at IRES Piemonte. Her main fields of study are student support policies, both at the regional and national levels. From 2017 to 2019, she was a member as an expert at the National Observatory for University Student Affairs.

Marco Le Moglie is Associate Professor of Economics at Catholic University in Milan. He is also a research fellow at CLEAN (Crime and Law Economic Analysis) at Bocconi University. His research focuses on the Economics of Crime, Education, Political Economy, and Public Economics.

Marinella Leone is Associate Professor at the University of Pavia and affiliate of the Laboratory for Effective Anti-poverty Policies (LEAP-Università Bocconi) . She received her PhD in economics from the University of Sussex and worked as a Research Fellow at the Institute of Development Studies. Her research is in development economics. Her more recent research focuses on the study of early child development parenting programs, the development of socio-emotional skills in primary school children and gender-based violence. In previous research, she investigated the short, long-term and intergenerational impact of conflicts.

Gabriele Lombardi is a Research Fellow in Social Statistics at the University of Florence. His research interests include internal migration of Italian university students, university-to-work transition, adolescent gambling addiction, and the propensity to adopt pro-sustainable behaviors among adolescents.

Natalia López Hornickel is Ph.D. Researcher in Advanced Quantitative Methods in Social Sciences and she holds a Master's degree in the same pathway, at the University of Bath, United Kingdom. She has a master's degree in Sociology, and she is a Sociologist from the Pontificia Universidad Católica de Chile. She has a Diploma in Intersectional Focus, Gender and Public Policies, Universidad Mayor, Chile. Her areas of interest are gender inequality, citizenship education, large-scale assessments, work trajectories, and social research methodologies.

Pietro Lucisano, former full professor of Experimental Pedagogy at the University of Rome La Sapienza. He has coordinated two IEA surveys and national projects for Italy. At the university, he served as vice-rector and was responsible for orientation and placement. He coordinated the doctoral program in Experimental Pedagogy. He is the past president of the Italian Society for Educational Research (SIRD). He coordinates the UNI.CO project.

Annamaria Lusardi is Senior Fellow at the Stanford Institute for Economic Policy Research (SIEPR), and the Director of the Initiative for Financial Decision-Making, a collaboration between SIEPR, the Graduate School of Business (GSB), and the Economics Department at Stanford University. She is also Professor of Finance (by courtesy) at the GSB.

Elisa Manzella is a research fellow at the University of Milano-Bicocca and holds a Ph.D. in Sociology from the Catholic University of Milan with a thesis on the role of teachers in the reproduction of educational inequalities. Her research interests include inequalities in education and school guidance, as well as the experimental evaluation of interventions in the educational field.

Lorenzo Maraviglia, PhD in Sociology (title obtained at the University of Pisa). In the last twelve years I have directed the statistical office of the province of Lucca and I have participated in various research projects in the social and statistical fields promoted by ISTAT and various research bodies. Actually I work as a statistical researcher at INVALSI.

Michele Marsili, graduated in Statistics at Sapienza University of Rome. He worked in Business Intelligence consulting, providing software development solutions for analysis and support for company's decision

making in insurance and pharmaceutical industries. Since January 2018 he has been working in the Statistical Service of INVALSI.

Angela Martini graduated in Philosophy and in Experimental Psychology at the University of Padua. She deals with research in the field of students assessment and schools evaluation, comparison between educational systems and analysis of data from international surveys on learning levels and is the author of numerous essays and articles published in Italian and foreign magazines.

Sonia Marzadro is a research fellow at the Research Institute for the Evaluation of Public Policies (IRVAPP).

Rita Marzoli is director of INVALSI Library; she graduated from the Vatican School of Librarianship. She is editorial board coordinator for 'INVALSI per la ricerca', a double-blind peer-reviewed series published by Franco Angeli. Her research interests are: systematic review, bibliometrics, open access, research evaluation, information literacy.

Ana María Mejía Rodríguez is a researcher at the IEA's Research and Analysis (RandA) Unit. Prior to joining the IEA, she was an Early-Stage Researcher of the European Training Network OCCAM. Her interests include educational effectiveness research and the dissemination of results beyond the core research audience.

Stefania Merone is a PhD Student at the University of Oxford. Her research interests include micro theory, behavioral economics and experimental economics.

Eleonora Miaci is a PhD candidate in Statistical Sciences, curriculum in Demography, at La Sapienza University of Rome and holds a Research Fellowship in the MEI (Multiple Ethnic Inequalities in Italy) project at the Department of Social and Political Sciences, University of Milan.

Daniel Miranda is a full-time research assistant professor at the Centro de Medición MIDE UC at the Pontificia Universidad Católica, research associate at the Center for Social Conflicts and Cohesion Studies - COES and adjunct researcher at the Millennium Nucleus of Opportunities and Digital Inequalities - NUDOS, Chile. His research focuses on citizenship education, political inequalities and their intergenerational transmission.

Annarita Monaco, She is a primary school teacher, holds a PhD in Social Psychology, Development and Educational Research and is a member of various research associations, author and trainer of mathematics teaching for the first cycle of education. She is also an instructor in university laboratories and tutor in specialization courses. In 2019 she won the Lucia Ciarrapico prize from UMI. She is co-author of the book *Matematica per allievi con discalculia* (published by CAROCCI, 2024).

Antonella Montone Associate Professor (SSD MAT04) in Mathematics Education, at the University of Bari. Main research topics: Reduction of the phenomenon of early leaving school; Training and professional development of Mathematics teachers: the value of collective discussion; Technological innovation.

Giovanna Mora, Since 1996 L2-Italian teacher at German primary schools in the province of Bolzano, class teacher at IC Bassa Atesina (BZ) since 2000; since 2013 speaker as expert teacher in mathematics in courses at national level for public and private bodies; from 2014 to 2022 preparation of items for standardized tests for primary school for INVALSI; school year 2014/2015 collaboration to the preparation of the Mathematics Indications for primary schools in the province of Bolzano; from 2014 to 2018 coordinator and contact person for a number of projects related to specific difficulties in primary school; since 2020 author of mathematical paths for the journal *Nuovo Gulliver NEW*; from 2018 to today participation to the project " CO15truzioni del significato nei problemi matematici della scuola primaria " in collaboration with the Bassa Atesina Linguistic Center and Pier Luigi Ferrari.

Sara Mori, PhD in Evaluation of Educational Systems and Processes, psychologist-psychotherapist. She works at INDIRE: she carries out research activities in the field of the development of students' transversal competences and the promotion of well-being at school.

Giorgio Mori Full Professor of Histology and Head of the related laboratory at the University of Foggia, he also serves as Scientific Referent of the Confocal Microscopy core facility and is the Dean's Delegate for Teaching. He is the author of 85 published papers in international journals. H-index 36, Citations 3300.

Chiara Mussida, Associate Professor of Economic Policy at the Università Cattolica del Sacro Cuore (Piacenza). She received her BA and Ph.D. in Economics from Università Cattolica. She also holds a Master of Arts in Development Economics from University of Sussex (Brighton, UK). Her research interests include empirical labour economics, gender inequality, education, and microeconometrics.

Carla Nanni is Graduate in Humanities and is a researcher at IRES Piemonte. She coordinates the annual report of the Observatory on education and vocational training in Piedmont. Nanni is in charge of the Observatory on the educational system of Piedmont (SISFORM), where statistics and research from the IRES working group are gathered.

Jessica Niewint-Gori is a Technologist at INDIRE and the Head of the Research Department on innovative methods for laboratory education and new technologies and principal investigator in international and national research projects based on STEM education and didactics supported by Artificial Intelligence tools.

Mattia Oliviero, holds a PhD in Sociology and Social Research at the University of Trento. Currently collaborating with IPRASE on various educational projects.

Laura Palmerio, senior researcher at INVALSI, is head of the International Surveys department. Italian Responsible for OECD and IEA projects. Main research interests: equity in education, relations between literacy in reading and in Mathematics.

Nazareno Panichella is a Full Professor of Economic Sociology at the Department of Social and Political Sciences of the University of Milan. He is the Principal Investigator of the projects Geography and Social Inequality in Italy (GESI) and Multiple Ethnic Inequalities in Italy (MEI).

Ornella Papa is Specialist in Psychological Evaluation and Researcher at INVALSI Library. She has been Subject Matter Expert at the Chair of "Introduction to Educational and Evaluative Research" at the University of Rome Tor Vergata. Her current research interests are: School libraries, Information literacy, Digital divide.

Valentina Patacchini is a researcher with statistical specialization at the Istituto Regionale per la Programmazione Economica della Toscana (Irpel) which supports Tuscany Region in developing and evaluating public policies. Her main areas of research concern the analysis of training policies and the characteristics of Tuscan firms.

Salvatore Patera, Associate professor in Didactics, special pedagogy and educational research at UNINT (Rome). President of the master's degree course in "Languages for innovative teaching and interculturality" at UNINT. Director of the specialization course for educational support activities for students with disabilities in the same university.

Riccardo Pietracci is a Research Collaborator at INVALSI. He has served as coordinator of computerized assessment for international surveys since 2018. He has coordinated national groups for test corrections across various areas of assessment for large-scale surveys conducted by IEA and OECD (reading comprehension, civic education, digital literacy, computational thinking, creative thinking).

Andrea Pietrolucci is a Postdoc research fellow at University of Milan within the EDUPOL project funded by European Commission. He defended the thesis at University of Trento. Previously, he studied at Sapienza University.

Giuseppe Pignataro Associate Professor at the Department of Economics of the University of Bologna and member of the Social Impact Agenda for Italy. His research interests focus on the study of inequalities of opportunity and their measurement, as well as on the causal inference evaluation of public policies, with particular reference to the issues of education and educational processes in general.

Maria Polo obtained her Doctorate in Mathematics Education at the University of Rennes I in 1996. She is an associate professor at the Department of Mathematics and Computer Science at the University of Cagliari. He deals with mathematics education and how the teacher's epistemology influences teaching practice.

Samuele Poy is an Assistant Professor (RTD-B) in Economic Policy in the Department for Sustainable Development and Ecological Transition and Adjunct Professor at the University of Piemonte Orientale (UPO), Vercelli (Italy). His research mainly focuses on counterfactual impact evaluation, public policy analysis, labour and social policies, welfare.

Andrea Priulla is a PhD from the University of Palermo, Italy. He holds a research fellowship at the University of Palermo at the Department of Economics, Business, and Statistics.

Chiara Respi is Technologist at the Department of Sociology at the University of Milano-Bicocca, where she works on the MUSA project about digital wellbeing and on the EYES UP project about the impact of digital devices on students' learning. Her research interests focus on the social implications of digitalization and on surveys.

Elisabetta Robotti is an Associate Professor in the Department of Mathematics at the University of Genoa, where she teaches Mathematics Education for Inclusion. She also teaches Mathematics and Mathematics Education at the Department of Primary Education Sciences. Her main research interests include teaching and learning early mathematics in kindergarten and primary school, as well as inclusive education.

Mariacristina Rossi is currently serving as Commissioner at COVIP and she is Full Professor on leave from the University of Turin. She is involved in research on household finance as well as gender economics and development economics. She previously worked as a consultant at the World Bank, Unicef and ILO.

David Rutkowski, Ph.D., is a Professor specializing in Qualitative and Quantitative Research Methods at Indiana University. Before joining IU, he served as a Professor of Education at the Center for Educational Measurement (CEMO) at the University of Oslo, Norway, and worked as a researcher for the International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, Germany. David earned his Ph.D. in educational policy with a focus on evaluation from the University of Illinois at Urbana-Champaign. His research centers on educational assessment and policy, particularly in the context of international large-scale assessments. He explores the methodologies of large-scale assessments, their role in policy debates, the influence of background questionnaire quality on achievement results, and achievement estimation in heterogeneous populations.

Guido Salza is a post-doctoral researcher at the Department of Social Sciences, University of Luxembourg. He is interested in the institutional and informal constraints that prevent equal educational opportunities for all, with a focus on quantitative methodology. His work has been published in international journals such as the European Sociological Review, Research in Social Stratification and Mobility, and Socio-Economic Planning Sciences.

Citlalli Sanchez-Alvarez, bachelor degree in psychology from the Autonomous University of Baja California, Mexico; Master's and PhD in Educational Sciences from the Institute for Educational Research and Development (IIDE-Mexico) with a focus on large-scale educational/psychological constructs, language proficiency/certification assessment instruments. Teacher in educational and social sciences undergraduate and postgraduate programs since 1997. Instructor in professional teacher development programs since 1999 aimed at primary, middle school, and university level teachers for their skill development in: • the design of classroom and formative assessments, • teaching practices and strategies, and • student learning skills development. Research interests: • Development and validity of large-scale psychological constructs, and student achievement assessment instruments, using AIG and traditional methodologies. • Development and validity of student context questionnaires, and surveys that focus on professional development and the teaching profession. • Analysis of personal, school, and teacher factors associated with student achievement in national and international assessments, in subject matters such as Mathematics, Language, Science, and Civic education.

Andres Sandoval Hernandez is currently a professor at the Department of Education at the University of Bath, UK. Prior to that, he worked as Head of the Research and Analysis Unit at the International Association for the Evaluation of Educational Achievement (IEA) in Germany.

Vincenzo Nunzio Scalcione teaches theories and methods of school planning and evaluation at UNIBAS. Ph.D. in Human Relations Sciences, degrees in Philosophy and Primary Education Sciences. Member of national and international scientific committees, author of publications on the topics of skills assessment and promotion.

Giovanna Scarchilli is a Postdoctoral Research Fellow at the University of Milan, with a PhD in Economics. She is specializing in the use of machine learning models to carry on research on socio-economic inequalities. Her research interests include inequalities of opportunity, multidimensional poverty, and redistributive capacity of a welfare system.

Alessia Sconti is a Researcher at the University of Kaiserslautern-Landau (RPTU) in Germany and Bocconi University in Italy. She is affiliated with the Global Financial Literacy Excellence Center at Stanford University and is a member of the Executive Committee of the G53 Financial Literacy and Personal Finance Research Network (G5 Network).

Daniel Shephard, Ph.D., is Lecturer in Comparative Education and Education Policy at the University of Edinburgh. He is also a former Postdoctoral Fellow in International Sustainable Development at Indiana University Bloomington. This research is at the intersection of comparative education, analytical sociology, development, and forced migration. His current research projects are focused on studying the sociological dimensions of refugee education inclusion in low- and middle-income countries and research on gaps in the international educational data regime related to SDG 4. He holds a PhD and MPhil in Comparative and International Education from Columbia University and a MSc in Evidence Based Social Intervention from the University of Oxford.

Maria Teresa Siniscalco is the National President of ADi, Association of Italian Teachers' and School Heads, and works as expert for INVALSI, where she was previously National Project Manager of PISA 2003. Before, she worked as Senior Researcher at the IEA Data Processing Centre in Hamburg. She holds a degree in philosophy and a doctorate in experimental pedagogy from the University of Rome, La Sapienza.

Carlotta Soldano, university researcher (RTDA) at the Department of Philosophy and Educational Sciences, University of Turin, Savigliano campus. Disciplinary scientific area MAT/04.

Mara Soncin is a Senior Assistant Professor at Politecnico di Milano, School of Management. Her research interests are in the educational field, with a focus on (i) digital learning, (ii) school management and (iii) the use of quantitative models for the evaluation of public policies.

Camilla Spagnolo is a researcher in Mathematics Education at University of Ferrara. Her main research interests include argumentation processes in mathematics, didactics by competence and teacher professional development. She is involved in teacher professional development courses for all school levels.

Cristina Specchi is a PhD Student in Public Governance, Management and Policy at the University of Bologna. Her research interests encompass the study of inequality of opportunities and their measurement, as well as effective public policies to contrast them. Her main focus regards educational and economic inequalities.

Massimiliano Stacchini, PhD in Economics, has worked for several years in the DG for Economics, Statistics and Research of the Bank of Italy and is currently heading the Analysis and Surveys unit of the DG for Consumer Protection and Financial Education of the Bank of Italy. In 2015 he spent one year as Visiting Scholar at Boston College.

Rolf Strietholt is the Co-Head of the IEA's Research and Analysis Unit (RandA). He coordinates research on educational measurement and consults external researchers and international organizations. His main research interests include school and educational effectiveness research and research methodology.

Cristina Stringher, senior researcher INVALSI responsible for the ECEC sector, has been Chair in the *Data Development Group* of the *OECD ECEC Network*, component in the Advisory Board *OECD Study on Social and*

Emotional Skills and invited expert by the EU Commission on: quality indicators for ECEC; key competences and Learning to Learn.

Isabella Sulis is Professor of Social Statistics at the University of Cagliari. Her research interests include multilevel analysis, item response theory, and measurement models to build-up adjusted indicators of performance, with main focus toward the evaluation of the education system.

Nicolás Tobar, Sociologist from the University of Chile. Interested in studying the connections between theories of social acceleration, democracy, civic education and political participation. She works as a teaching support in the sociology department of the University of Chile and is an assistant in research centers such as NUDOS (Nucleus of Inequality and Digital Opportunities), OLES (Observatory of Violence and Social Legitimacy) and Social Theory Today.

David Torres Iribarra, is an Associate Professor at the Pontificia Universidad Católica de Chile. His research areas include: the application of latent variable models to measurement contexts; the theoretical foundations of social sciences measurement (see his book "A Pragmatic Perspective of Measurement" (2021, Springer); and the use of digital technologies to enhance educational assessment.

Giusi Antonia Toto, Full Professor of Didactics and Special Education, Delegate for Teacher Training and Orientation at the University of Foggia. She has lectured in areas pertaining to educational sciences and cognitive-cultural development of the individual. She has acquired solid experience in teacher training and educational innovation.

Ernesto Treviño, Doctor in Education from Harvard University, professor at the Faculty of Education of the Pontificia Universidad Católica de Chile, director of the UC Center for Educational Transformation (CENTRE UC) and principal researcher at the Center for Educational Justice. He has collaborated with IEA, OAS, OEI, UNESCO, UNICEF and other international organizations, as well as in different countries, supporting the development of studies and making educational policy recommendations. Her lines of research are interested in connecting public policies and school and classroom practices with a focus on educational justice.

Moris Triventi is Full Professor in the Department of Social and Political Sciences at the University of Milan. Previously he was Associate Professor at the University of Trento and Research fellow at the European University Institute (Fiesole). His research interests comprise social inequality and mobility, educational inequalities, causal inference and policy evaluation. His publications appear in a number of international journals, including Review of Educational Research, Annual Review of Sociology, European Sociological Review.

Herman van de Werfhorst is professor of sociology at the European University Institute and visiting fellow at the Wissenschaftszentrum Sozialforschung Berlin (WZB). His work concentrates on inequalities in and through education, using both comparative and longitudinal data and research designs.

Chris Van Klaveren is an Associate Professor at the Faculty of Behavioural and Movement Sciences of the Vrije Universiteit Amsterdam and director of ACLA.

Ira Vannini, Full professor in Educational Research. Current research topics: assessment and educational evaluation; curriculum design; teacher training methodologies; monitoring of educational experiments; relationship between teachers and formative assessment practices. Director of the CRESPI Interuniversity Center.

Loris Vergolini is currently Assistant Professor at the University of Bologna. His research activity deals with the impact evaluation of educational policies with a focus on tertiary education. His research interests include the study of inequalities in educational opportunities and social cohesion.

Matteo Viale, Full professor of Italian Linguistics (scientific-disciplinary field L-FIL-LET/12) in the Department of Classical Philology and Italianistics at the University of Bologna. He is currently coordinator of the master's degree in Italianistics, European Literary Cultures, Linguistic Sciences.

Melvin Vooren is an Assistant Professor at the Faculty of Behavioural and Movement Sciences of the Vrije Universiteit Amsterdam and affiliated with the Amsterdam Center for Learning Analytics (ACLA). His research focuses on education, health, crime, and predictive modelling.

Michela Zambelli is a research fellow in psychometrics and adjunct professor at the Catholic University of the Sacred Heart. Her research activity involves the study of the dynamics of change of the meaning-making process through the application of intensive longitudinal designs, and the development and validation of tests and self-report measures.

Nadir Zanini is currently Associate Director for Research Data and Analytics at Ofqual, the Office of Qualifications and Examinations Regulation in England. He has previously worked within the Research Division of Cambridge Assessment (Cambridge, UK), and at FBK-IRVAPP, Research Institute for the Evaluation of Public Policies (Trento, Italy).

Bilal Zia is the Head of Data Science & Analytics at Duolingo, an EdTech company which aims to develop the best, universally available, education in the world. He led an interdisciplinary Central Science team at Amazon, after a decade and half in the Research Department of the World Bank in Washington D.C. pursuing an applied academic career.

INVALSI

Via Ippolito Nievo, 35 - 00153 - Tel. 06941851

Mail: protocollo@INVALSI.it PEC: protocollo.INVALSI@legamail.it Website: www.INVALSI.it
C.F. 92000450582