

TOWARDS ENRICHED STANDARDIZED TESTING: CONSTRUCTING RELEVANT SCHOOL-LEVEL INDICATORS OF LEARNING PROGRESS IN SECONDARY EDUCATION

ADMIN DETAILS

Project Name: Towards enriched standardized testing: Constructing relevant school-level indicators of learning progress in secondary education

Grant Title: C24M/21/013

Principal Investigator / Researcher: Koen Aesaert

Description: In the context of national tests, value-added estimates (VAE) are often used to indicate the extent to which schools foster their students' learning progress. Although VAE are statistically valid, they 1) do not show how much progress schools make between measurement occasions; 2) reduce the quality of a school to a univariate performance measure; and 3) require complex modeling, making the results less transparent and usable for its purposes, i.e., school accountability, improvement and choice. Moreover, it is hard to identify a school's true learning progress due to changes in school populations between measurement occasions, irregular student trajectories and the problem that short-lived administrative schools do not always coincide with real stable schools. The project aims are 1) to identify real, stable, secondary schools, 2) to develop a multivariate indicator of learning progress in these schools, taking into account student trajectory interferences such as inter-school mobility and tracking, and 3) to identify local school clusters providing complete secondary education trajectories and to investigate whether the role of a school within its local cluster is a suitable criterion to select schools for comparison in terms of learning progress.

To explore these research aims, student/school administrative data from the Flemish government (2020) and available data from the LISO-project (2013-2019) will be used. Regarding the first research aim, linked pupil / school campus administrative data for all the secondary education trajectories in the school years 2009–2010 to 2018–2019 will be requested from the administrative datawarehouse of the Ministry of Education and Training. This range of school years is chosen so as to encompass the measurement interval (2013–2014 to 2018–2019) of the LiSO study (from which the data are used for the second research aim) and the largest part of the transition of its main birth cohort (2001) through secondary education. In this dataset, a given pupil in a given school year is represented, conceptually, by one record, containing pupil and school campus identifiers (anonimised), the grade, an indicator of the pupil's track (possibly the administrative group), the locations (community) of the pupil's domicile and of the school campus, and an indicator of the socio-economic status of the pupil. An operational definition of a school will be combined with a pupil-flow linkage method to identify real stable schools from these data.

Regarding the second research aim, the data from the recently completed LiSO study (Dockx et al., 2019) and its predecessor, LoSO (Van Damme et al., 2002) will be used. LiSO (LoSO) has followed 6148 (6411) pupils in 48 (57) schools who started secondary education in 2013–2014 (1990–1991). The studies have used regional sampling with complete enumeration: nearly all the pupils of the cohort in all the schools in the regions in question were included. Achievement in mathematics and Dutch has been measured at several occasions, with longitudinal measurement invariance. In addition, pupil and family background variables and teacher, class, and school characteristics have been collected. Both datasets will be used to construct and validate the school-level indicator of learning progress.

Regarding the third research aim, the complete set of real stable schools identified in the administrative data (first research aim) will be partitioned in local clusters. afterwards, the internal structure of the local clusters will be analysed by means of the pupil flows between schools in the same cluster, partitioned according to grades and tracks, in order to derive a classification of the different roles of schools within the cluster. The proposed approach in quality assurance of comparing a school to schools with similar roles in their local clusters will be evaluated by means of the LiSO and LoSO datasets. Within the framework of a traditional league table approach to the analysis of learning progress, the ranking of the school within the league table will be compared to its ranking within the subset of schools with similar roles within their local clusters. Finally, the relevance of the local clusters as a level of grouping in educational effectiveness analysis will be explored. Early school leaving will be used as the outcome measure in this analysis, using the LiSO and LoSO datasets.

Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)

Koen Aesaert (PI)

Wim Van Den Noortgate (coPI)

Internal Funds Project number & title

C24M/21/013

Towards enriched standardized testing: Constructing relevant school-level indicators of learning progress in secondary education.

2. DATA DESCRIPTION

2.1. Will you generate/collect new data and/or make use of existing data?

- Reuse existing data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

Type of data	Format	Volume	How created
--------------	--------	--------	-------------

Linked pupil and school campus administrative data	.csv	< 1 GB	Retrieved from the Ministry of Education
Student, teacher and school background data; student test scores	.xls	< 100MB	Retrieved from the LiSO and LoSO database available at the Centre of the PI: data were originally retrieved from standardized tests and student, parent, teacher and school leader questionnaires

3. ETHICAL AND LEGAL ISSUES

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

No

Regarding the administrative data, the data protection officers of the Ministry of Education assured that students were not identified and are not identifiable.

Yes

The LiSO-data and LoSO-data can be considered as personal data. Regarding the LiSO-/ LoSO-data, all data are pseudonymized and the pseudonomisation key is held by the LiSO coordinator at the Centre of Educational Effectiveness and Evaluation (PI's research centre). A PRET-application was submitted and approved: nr. G-2022-5683-R2(MIN)

3.2. Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s).

No

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

No

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

Yes

The administrative data provided by the Ministry of Education can only be used for this project.

4. DOCUMENTATION AND METADATA

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

All research aims will be tackled by conducting secondary analyses using the LiSO-/LoSO- and administrative data. We refer to the original codebook and coding rules that were developed during the original projects in which these datasets were created.

We will provide the syntaxes and algorithms that will be created to 1) identify the real stable schools, 2) construct and validate the school level indicator of learning progress, and 3) identify the local school clusters.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

No

5. DATA STORAGE AND BACKUP DURING THE PROJECT

5.1. Where will the data be stored?

All data is stored in the secured Onedrive for Business of KU Leuven, with multifactor authentication for confidential data (up to 2TB). Only the PhD student and PI have access to the data.

5.2. How will the data be backed up?

All data is stored in the secured Onedrive for Business of KU Leuven with automatic daily back-up procedures.

5.3. Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

Yes

Data size will be no more than 40 GB. KU Leuven's Onedrive has a capacity up to 2TB.

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

None

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All data is already pseudonymised and encrypted. It will be stored in the secured Onedrive for Business of KU Leuven. Only the PhD student and PI have access to the data. Only the PI and the LiSO-dataset coordinator have access to the pseudonymisation key of the LiSO- and LoSO-dataset (stored on a

separate network drive). The linked student-school administrative data is already non-identifiable when it is provided by the Ministry of Education.

6. DATA PRESERVATION AFTER THE END OF THE PROJECT

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

Due to contractual restrictions the administrative data of the Ministry of Education will not be preserved once the project ends.

All other data will be preserved for 10 years.

6.2. Where will these data be archived (= stored for the long term)?

After the study, we will transfer the data to the PI's Onedrive for at least 10 years.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

None

7. DATA SHARING AND RE-USE

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

Due to contractual restrictions, the administrative data cannot be shared.

7.2. Which data will be made available after the end of the project?

(Parts of) the LiSO-data.

7.3. Where/how will the data be made available for reuse?

- Upon request by mail

Data from LiSO will be available on request after signing a data sharing agreement. Access to the data should be asked directly to the LiSO-team, not to the PI and researcher of this project.

7.4. When will the data be made available?

- Immediately after the end of the project

7.5. Who will be able to access the data and under what conditions?

Access of the LiSO data will be considered after a request is submitted explaining the planned reuse.

Only uses for specific research purposes will be allowed and commercial reuse will be excluded.

7.6. What are the expected costs for data sharing? How will these costs be covered?

None

8. RESPONSIBILITIES

8.1. Who will be responsible for the data documentation & metadata?

The main researcher (Georges Van Landeghem) and the PI (Koen Aesaert) .

8.2. Who will be responsible for data storage & back up during the project?

The main researcher (Georges Van Landeghem) and the PI (Koen Aesaert).

8.3. Who will be responsible for ensuring data preservation and sharing?

The PI (Koen Aesaert).

8.4. Who bears the end responsibility for updating & implementing this DMP?

The end responsibility for updating and implementing the DMP is with the supervisor/PI (Koen Aesaert).