THE ROLE OF SAMPLE CHARACTERISTICS IN THE STABILITY OF VALUE-ADDED ESTIMATES OF SCHOOL EFFECTS: SAMPLE SIZE, STUDENT MOBILITY AND SAMPLE HETEROGENEITY

THE ROLE OF SAMPLE CHARACTERISTICS IN THE STABILITY OF VALUE-ADDED ESTIMATES OF SCHOOL EFFECTS: SAMPLE SIZE, STUDENT MOBILITY AND SAMPLE HETEROGENEITY

ADMIN DETAILS

Project Name: The role of sample characteristics in the stability of value-added estimates of school effects: sample size, student mobility and sample heterogeneity.
Grant Title: G023823N
Principal Investigator / Researcher: Koen Aesaert
Description: Many countries use value added (VA)-estimates as indicators of the effectiveness of their schools. A major issue is the instability of schools' VA-estimates over successive years. Instable VA-estimates may indicate a lack of reliability and are therefore less usable for intended purposes, e.g., identification of (in)effective schools or school choice. A challenge is to provide evidence of the characteristics that drive the artificial (in)stability of VA-estimates, such as the statistical model that is being used or the covariates that are integrated in the model. We will study the relationship between the (in)stability of VA-estimates and one specific type of characteristics: sample features. This relationship between sample characteristics and the stability of VA-estimates will be explored in a variety of scenarios. As such, the project aim is threefold:

(1) We will answer the question of minimally required sample sizes to estimate stable VA-estimates, under the conditions of a joint value added model and the traditional approach of using separate VA-models.
(2) Combining the advantages of joint value added modeling and multiple membership modeling, we will investigate if VA-estimates of schools are more stable over time when the degree in which students have attended multiple schools is explicitly modeled.
(3) Finally, we will study the relationship between stability of VA-estimates and school heterogeneity.

Each of these research objectives will be explored, using simulation studies and real country data from Flanders and the United Kingdom.
Regarding the simulated data, we will use Monte Carlo simulation techniques to stochastically generate 2000 datasets per condition, assuming a population model and specific parameter values. After data

analysis, we will compare the resulting 2000 estimates for each parameter with the population value used to generate data. A big advantage of simulation studies is that, because we know the 'population' values used to generate the data, we have strong criteria to compare the performance of multiple models and to evaluate the quality of the inferences. A drawback is that conclusions are in theory conditional on the way data were generated. Therefore, we will adjust the assumptions that we will use to generate data, i.e., we will simulate data for a wide variety of conditions to reveal the underlying sample characteristics that drive differences in the stability of VA-estimates across multiple models.

For each research objective and study, the performance of the different models and the impact of the sample characteristics on this performance will be illustrated using two real datasets from Flanders and England. In both applications we will focus on primary education, as these schools deal with smaller sample sizes, more mobility and more volatile intakes (due to small size). Regarding the Flemish data, we will analyze two cohorts of approximately 4,500 students from 150 schools, who will take a digital version of the standardized SIBO-tests for reading (Vanwynsberghe et al., 2017) and the standardized reading test of Bogaerts et al. (2022). In total, each student will complete a vertical equated version of the test four times, i.e., at the start of grade 4 (age 8-9), at the end of grade 4, at the end of grade 5 and at the end of grade 6 (age 11-12). The data collection is part of a larger data collection of the Centre of Educational Effectiveness and Evaluation (KU Leuven), started from June 2021 onwards. The stratification procedure of the Flemish national assessments will be used, i.e., schools will be stratified for school size and educational network. Using this procedure, schools will be representative for the background and school composition characteristics (based on home language and SES) of this project (Denis et al., 2019). To maximize the degree to which we can capture student mobility, all schools will be located in the province of the university. In Flanders, differences in student achievement according provinces are relatively small (Denis et al., 2019). Information on background characteristics and student mobility will be gathered using short student, parent and teacher questionnaires. The GDPR assessment and ethical review of this larger data collection confirmed that the project meets the standards for academic research and that the proposed processing of personal data meets the requirements of the General Data Protection Regulation (reference number: *G-2021-4176-R2(MAR)).*

Regarding the data for England, we will analyse student data from the National Pupil Database (NPD), a longitudinal administrative census of all students in state-funded education. We will model primary schools' value-added between the end-of-primary school 'Key Stage 2' national standardized achievement tests in reading and math (age 10/11, year 6) and reading and math 'Key Stage 1' tests taken four years earlier (age 6/7, year 2). The data include basic student demographic and socioeconomic characteristics. We observe student mobility annually. We will investigate the stability of the schools' VA-estimates across four cohorts of students who completed primary schooling in 2016, 2017, 2018, and 2019. The full data relate to approximately 15,000 primary schools and 550,000 pupils per cohort. The real data sets will be used to explore the practical impact of different models and approaches, and illustrate how data can be analyzed to minimize artificial instability. They will also be used to inform the simulation studies, e.g., specification of sensible true parameter values.
Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)
Koen Aesaert (PI)
Wim Van Den Noortgate (coPI)
George Leckie (coPI, Bristol University)

FWO Project number & title
G023823N
The role of sample characteristics in the stability of value-added estimates of school effects: sample size, student mobility and sample heterogeneity.

2. DATA DESCRIPTION

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

- Generate and use simulated data
- Collect new 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
Reuse student and school background data; student test scores from the national standardized achievement tests in England	.xlsx	<1GB	Retrieved from the National Pupil Database/English Department of Education
Create and use student and school background data; student test scores from 4th, 5th and 6th graders from approximately 150 schools in Flanders	.xlsx	<1GB	Information about student background characteristics will be gathered by administering student (digital using Qualtrics) and parental questionnaires (paper and pencil). Teachers also receive a questionnaire to gather information about the schools (paper and pencil). Digitalized standardized tests will be administered to the

			students in grades 4, 5 and 6.
Simulated data	R-table	<1GB	Data will be simulated in R using runmlwin and lme4.

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.

Yes

We will use data from the following categories:

- Identification data: names
- Identification data: e-mail
- Personal data: age
- Personal data: gender
- Personal data: date of birth
- Personal data: nationality
- Personal data: occupation
- Sensitive personal data: physical and mental health
- Sensitive personal data: learning difficulties and disorders

Regarding the Flemish data, all data are pseudonymized and the pseudonomisation key is held by the PI and main researcher of this study at the Centre of Educational Effectiveness and Evaluation (PI's research center).

Regarding the NPD-data, we will ask the English Ministry of Education to assure that students are not identified and are not identifiable when receiving data.

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 NPD-data provided by the English 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 simulating data, using the Flemish data and conducting secondary analyses using the English NPD-data. Regarding the NPD-data, we refer to the original codebook and coding rules that were developed during the original project. Regarding the Flemish, codebooks will be created for each dataset (xlsx-format).
We will provide the syntaxes and algorithms that will be created to 1) simulate the data, 2) link the real datasets, and 3) estimate the different value added models.

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 3 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?
Regarding the Flemish data, all data from 2021 and 2022 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 and only the PI and PhD student have access to the pseudonymisation key of the Flemish datasets (stored on a separate network drive). The same procedure will be followed for the subsequent parts of data collection. The NPD-data should already be non-identifiable when it is provided by the English 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 NPD-data of the English 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 NPD-data cannot be shared.

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

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

- Upon request by mail

The Flemish data will be available on request after signing a data sharing agreement. Access to the data should be asked directly to the PI 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 Flemish 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 (Tom Van Ransbeeck) and the PI (Koen Aesaert) .

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

The main researcher (Tom Van Ransbeeck) 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).