Teacher Value-Added in the Absence of Annual Test Scores: Utilising Teacher Networks

Desislava Tartova
Paris School of Economics

Motivation

- **Teacher value added** (TVA) better predictor of student long-term outcomes than any observable teacher characteristics (Chetty et al., 2014b)
- Teacher recruitment, progression and compensation linked solely to observables
- **Important reason**: TVA estimation reliant on a panel of student scores in standardised exams
  - Absence of std. exams in consecutive grades in most countries
- **Panel structure** necessary to control for student unobservables correlated with teacher sorting
  - E.g. high-ability students sorted to a high VA teacher $\rightarrow$ upward bias in estimates
  - Standard method: regress exam score on teacher FEs (captures TVA) *and* lagged score (captures sorting on unobs.)

This paper in a nutshell

- **Idea:** Develop an unbiased within-school TVA estimation method not reliant on lagged std. exam grades

- **Strategy:** Use cross-sectional data of student scores in std. exams in two subjects, controlling for sorting bias within-school by exploiting “networks” of teachers

  → Two teachers in the same subject teaching in classrooms that share the same “link" teacher in another subject

- **Validation:** Simulations show method performs well compared to standard method

- **Application:** Empirical test for 9th grade students in French middle schools

Roadmap

- **Identification**
- Empirics
- Validation
- Application
- Takeaways

Theoretical determinants of student grade

- For simplicity, for now assume observable characteristics do not predict student grades

- Avg. grade of classroom c in subject f with teacher j_f as a function of TVA and avg. student ability:

$$Grade_{c,f,j_f} = TVA_{j_f} + CommonAbility_c + SubjectSpecAbility_{c,f}$$

Intuition

- **Case with only two classrooms** (c_A and c_B), each with a Math and a French teacher
- To compare VA of **two Math teachers:** ΔM as the diff. between Math grades across classrooms

$$\Delta M = \underbrace{[TVA_{j_{M_A}} + CommonAbility_{c_A} + MathAbility_{c_A,M}]}_{MathGrade_{c_A,M,j_{M_A}}} - \underbrace{[TVA_{j_{M_B}} + CommonAbility_{c_B} + MathAbility_{c_B,M}]}_{MathGrade_{c_B,M,j_{M_B}}}$$

- ΔM also captures diff. in ability across classrooms
- To **get rid of common ability**: $\Delta M - \Delta F$

$$\left([TVA_{j_{M_A}} - TVA_{j_{M_B}}] - [TVA_{j_{F_A}} - TVA_{j_{F_B}}]\right) +$$
$$[MathAbility_{c_A,M} - FrenchAbility_{c_A,F}] - [MathAbility_{c_B,M} - FrenchAbility_{c_B,F}]$$

- But $\Delta M - \Delta F$ also captures the diff. in VA of the French teachers $\rightarrow$ focus on corner case

Corner case

- Compare only classrooms with the **same French teacher** (**link teacher**)

  $\longrightarrow$ The two Math teachers are **linked** $\rightarrow$ said to belong to a **network**

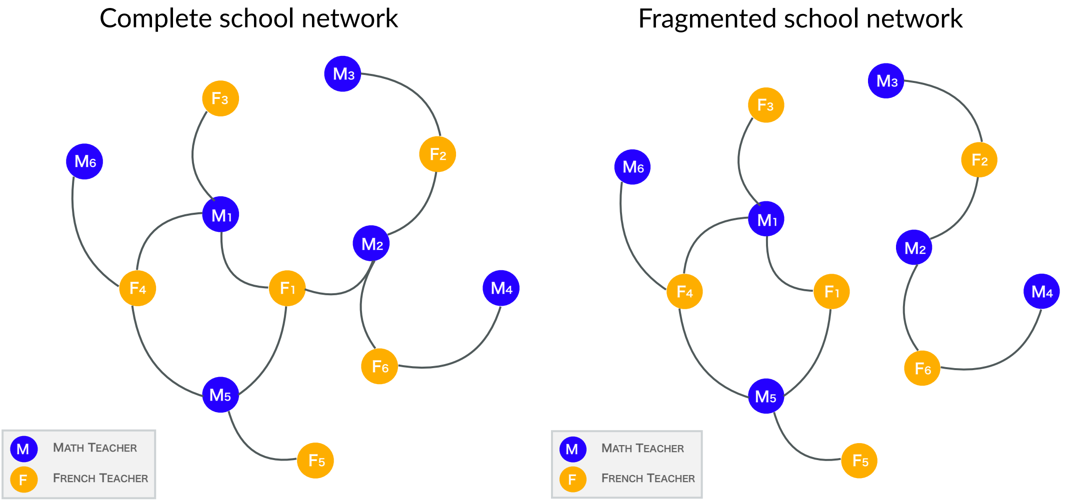
- **Remaining issue:** $\Delta M - \Delta F$ also captures the diff. in relative Math ability across classrooms

$$[TVA_{j_{M_A}} - TVA_{j_{M_B}}] + \underbrace{[MathAbility_{c_A,M} - FrenchAbility_{c_A,F}]}_{\text{Relative Math ability in classroom } c_A} - \underbrace{[MathAbility_{c_B,M} - FrenchAbility_{c_B,F}]}_{\text{Relative Math ability in classroom } c_B}$$

- **Main assumption:** No sorting of teachers to classrooms based on relative Math ability

  $\longrightarrow$ French setting: satisfied according to tests on observables (instead likely sorting on common ability)

Full TVA distribution uncovered by transitivity



Roadmap

- Identification
- **Empirics**
- Validation
- Application
- Takeaways

Two-step procedure

- **First step:** estimate for student i in subjects $f \in \{M, F\}$ with teacher j_f in school s:

$$Grade_{i,f,j_f,t} = \mathbf{X}_i \beta_f + ExperienceFE + SchoolFE + \varepsilon_{i,f,j_f,t}$$

- where $Grade_{i,f,j_f}$ is the grade of student i in the std. exam, std. by year and subject
- $\mathbf{X}_i$ includes observable student characteristics, such as gender, age, scholarship status, socio-economic status, advanced classes taken, nationality,...
- *ExperienceFE* are teacher years-of-experience FEs (to control for the time component of TVA under assumption it is a function of years of experience)
- *SchoolFE* are school FEs (to ensure comparisons of student grades within school)

Two-step procedure

- **Second step:** Average the residuals at the classroom level and compute for each network of Math teachers j_{M_A} and j_{M_B} in each classroom c_A and c_B

$$(ResGrade_{c_A,j_{M_A}} - ResGrade_{c_B,j_{M_B}}) - (ResGrade_{c_A,j_F} - ResGrade_{c_B,j_F}) \equiv \underbrace{\widehat{TVA_{j_{M_A}} - TVA_{j_{M_B}}}}_{\text{Estimator of relative TVA of } j_{M_A} \text{ and } j_{M_B}}$$

- Moving from pairwise comparisons to a **distribution of TVA** within school: solve the system of all such equations per school

    - Note: overdetermined system, but can be written out in matrix format and solved by OLS

      Details

Roadmap

Identification

Empirics

Validation

Application

Takeaways

Simulation exercise

- Compare **my estimates to true TVA** in Monte Carlo simulations, similar to Guarino et al. (2015)
- Compare precision of my estimates **relative to the standard method**
- **Three scenarios** of interest
  - Random sorting
  - Extreme sorting on common ability: 1-to-1 corr. between true TVA and common ability
  - Extreme sorting on Math ability: 1-to-1 corr. between true TVA and Math ability

Calibration of parameters

- **Calibration of parameters** following results of Chetty et al. (2014a), Rothstein (2009) and French data descriptives

- Many **recalibration exercises** to confirm the robustness of the simulations, varying:

  → Number of teachers, classrooms, classroom size

  → Correlation between residualised Math and French grade

  → Correlation between current and lagged grade (for standard method tests)

  → Shares of total variance of student grade (TVA, common ability, subject-specific ability)

Simulation exercise

	Network Est.	Baseline Est.	Δ (NE-BE)
No sorting			
Sqr. Root of MSE	0.005	0.003	0.002
Spearman correlation	1.00	1.00	0.00
Sorting on common ability			
Sqr. Root of MSE	0.005	0.244	-0.239
Spearman correlation	1.00	1.00	0.00
Sorting on subject-specific ability			
Sqr. Root of MSE	0.464	0.246	0.218
Spearman correlation	1.00	1.00	0.00

Roadmap

- Identification
- Empirics
- Validation
- **Application**
- Takeaways

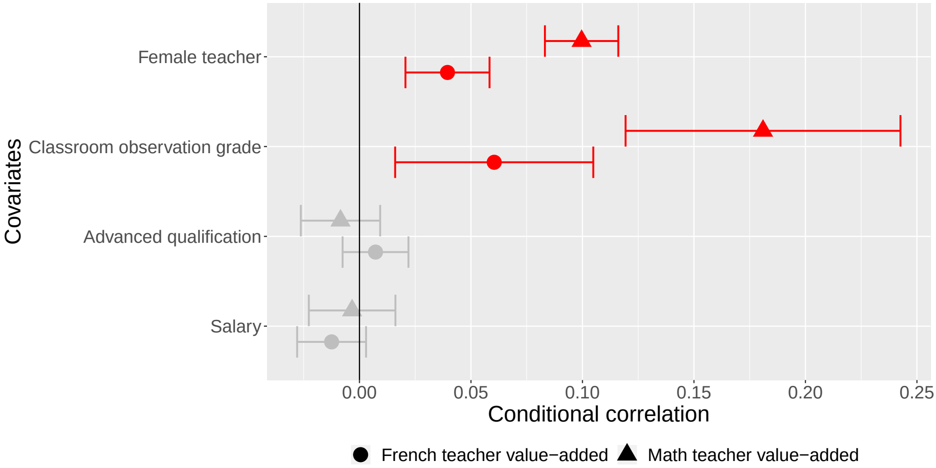
Application for French middle schools

- **Test method** using French administrative matched student and teacher data
- **Standardised exams:** DNB national exams in Math and French held at the end of 9th grade
- **Available networks:** 86% of schools with complete school network over 2009-2010 to 2018-2019
  - 40,000 unique networks in Math + 60,000 in French (incl. across time)
  - On average 135 (117) students per network in Math (French)
  - On average 6 (5) networks per Math (French) teacher within school
- **Unbiasedness:** Identifying assumption holds based on test on observables

Value-added estimates

- +1 s.d. in Math (French) TVA within a school $\Longrightarrow$ +0.175 (+0.164) s.d. student score
- Moving 5th→95th teacher quality within school $\Longrightarrow$ +0.58 s.d. (+0.54 s.d.) student score
- Estimated s.d. on the high side compared to the US (0.10-0.15 s.d. in Math, 0.05-0.15 s.d. in Literature), e.g. Jackson (2014), Bacher-Hicks and Koedel (2022)
- Closer to estimates in developing countries, e.g. Bau and Das (2020), Buhl-Wiggers et al. (2017)

Correlation of value-added and teacher characteristics



Roadmap

Identification

Empirics

Validation

Application

Takeaways

Takeaway

- New method for uncovering fixed within-school estimates of TVA not relying on panel data of student grades

- Method performs well compared to the standard method in simulations

- Plausible unbiasedness in French middle school setting

Thank You!

Desislava Tartova

desislava.tartova@psemail.eu

Bibliography

Andrew Bacher-Hicks and Cory Koedel. Chapter 2 - Estimation and interpretation of teacher value added in research applications. *Handbook of the Economics of Education*, 6:93–134, 2022.

Natalie Bau and Jishnu Das. The impact of individual teachers on student achievement: Evidence from panel data. *Americal Economic Journal: Economic Policy*, 12(1):62–96, 2020.

Julie Buhl-Wiggers, Jason T. Kerwin, Jeffrey A. Smith, and Rebecca Thornton. The impact of teacher effectiveness on student learning in Africa . 2017.

Raj Chetty, John Friedman, and Jonah E. Rockoff. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9):2593–2632, 2014a.

Raj Chetty, John Friedman, and Jonah E. Rockoff. Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679, 2014b.

Cassandra M. Guarino, Mark D. Reckase, and Jeffrey M. Wooldridge. Can Value-Added Measures of Teacher Performance Be Trusted? *Education Finance and Policy*, 10(1):117–156, 2015.

C. Kirabo Jackson. Teacher Quality at the High School Level: The Importance of Accounting for Tracks. *Journal of Labor Economics*, 32(4), 2014.

Jesse Rothstein. Student Sorting and Bias in Value-Added Estimation: Selection on Observables and Unobservables. *Education Finance and Policy*, 4(4):537–571, 2009.

Roadmap

Appendix

Transforming pairwise differences into coefficients

- Estimates direct comparisons between pairs of teachers
- To obtain VA distribution we need VA estimate per teacher
- Many ways to find an estimate for $VA_j \implies$ overdetermined system of linear equations
  - $VA_{M_1} - VA_{j_{M_2}} = a$, $VA_{j_{M_3}} - VA_{j_{M_2}} = b$, and $VA_{j_{M_1}} - VA_{j_{M_3}} = c$, but $c \neq a - b$
- **Solution:** Write out in matrix form and solve by OLS $\implies$ instead of trying to equate each equation to zero, $Ax - v = 0$, it minimises the sum of squared distances from zero

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Back