INNOVAZIONE E SOCIETÀ 2009

Metodi e politiche per la valutazione dei servizi

Facoltà di Economia - Università degli Studi di Brescia
Contrada S. Chiara 24-26 Giugno 2009

A review of multilevel value-added models in education

Leonardo Grilli & Carla Rampichini

Department of Statistics, University of Florence

email: grilli@ds.unifi.it rampichini@ds.unifi.it

Methods for comparing educational institutions

- Methodology developed in different fields: educational statistics, psychometrics, sociology, econometrics ...
- In this presentation we focus on the methodological challenges connected with statistical modelling and data analysis:
  - definition of effectiveness in education
  - multilevel models and their role in assessing effectiveness
  - statistical issues arising in effectiveness evaluation
  - use of model results

Effectiveness

- The effectiveness of an organization is the degree of achievement of its institutional targets
  - **absolute** (absolute effectiveness or impact analysis): evaluation of interventions, e.g. a specific vocational training course
  - **relative** (relative or comparative effectiveness): comparison among many institutions

Effectiveness

- For ***educational institutions (schools, universities)*** the effectiveness cannot be defined in absolute terms, but only with respect to the effects on the students
- In economic terms, the customers (students) are also inputs of the production function of the educational institution
- The effects on the students are affected by the features of the students themselves: how to make a fair assessment?

Hanushek E (1986) The economics of schooling: Production and efficiency in public schools. Journal of Economic Literature 24:1141–1177
Special issue of the Journal of Econometrics (2004): The econometrics of higher education

Value added

- The analysis of the educational process is difficult → the quality of educational institutions is usually measured via an **input/output** approach:
  - the process is a *black-box*
  - the output (*outcome*) is evaluated in the light of the input → effectiveness = **value added** by the school

VALUE-ADDED = ACTUAL OUTCOME
minus
EXPECTED OUTCOME GIVEN THE INPUT

Braun H and Wainer H (2007) Value-Added Modeling. In: Rao, C.R., Sinharay, S. (eds.) Handbook of Statistics 26, Psychometrics, pp. 475–501. Elsevier.
Special issue of the J. of Educational and Behavioral Statistics (2004)

Internal/external effectiveness

The educational process leads to multiple outcomes → many measures of effectiveness

- Internal effectiveness:
  - Dropout (1=Yes, 0=No)
  - Duration of studies (time to the degree)
  - Number of credits after a given period
- External effectiveness:
  - Occupational status after degree (1=Yes, 0=No)
  - Duration of unemployment (time to first job)
  - Wage or job satisfaction

Type A and B effectiveness

- **Type A**: performance of the institution adjusted for the features of the students, irrespective of the context → to inform school choice
- **Type B**: performance of the institution adjusted ***also*** for the context (e.g. resources, local labour market, socio-economic composition of enrolled students) → for accountability

Raudenbush SW & Willms JD (1995) The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*, 307-335.

Statistical issues

- The statistical models for assessing the relative effectiveness of educational institutions must face two main issues:
  - **Adjustment**: the measures must be adjusted at least for the features of the students (necessary for a fair comparison)
  - **Quantification of uncertainty**: the measures must be accompanied by error bars (necessary to make assessments properly supported by empirical evidence)

The *raw rankings* (so called 'League Tables') ignore both issues:
Goldstein H & Spiegelhalter DJ (1996) League tables and their limitations: statistical issues in comparisons of institutional performances. *JRSS A*, 159, 385-443

Statistical issues

Adjustment & Quantification of uncertainty

Regression models

But *standard* models are not suitable!

- Standard models make unsuitable assumptions on the variance-covariance structure (independence among observations, while the results of the students of the same school usually are positively correlated) → poor quantification of uncertainty
- Standard models are unable to represent some key features, e.g. varying slopes

Multilevel models

- Multilevel (mixed, random effects) models overcome the main limitations of standard models and are well suited for assessing the relative effectiveness of schools
  - The effectiveness of a school is explicitly represented by the ***random effects***

Random intercept model: definition

Features of the student

Features of the school/context

Outcome of the student

School random effect

$$Y_{ij} = \alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + u_j + e_{ij}$$

i = student

j = school

$$= (\alpha + u_j) + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + e_{ij}$$

Intercept of j-th school

Random intercept model: value added

Actual outcome

Expected outcome given student and school/context features

$$Y_{ij} - (\alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j) = u_j + e_{ij}$$

i = student

j = school

The difference between actual and expected outcome is decomposed in two parts:

- School-level component (random effect) u_j
- Student-level component e_{ij}

The random effect u_j is the school value added, or effectiveness. It is a *residual* term → its meaning depends on which covariates are in the model

Random intercept model: Type A and B effects

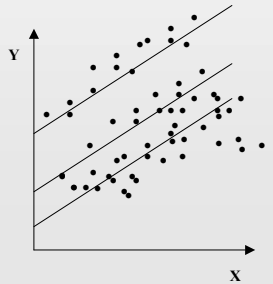
Type A effect of school *j*

$$Y_{ij} = \alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + u_j + e_{ij}$$

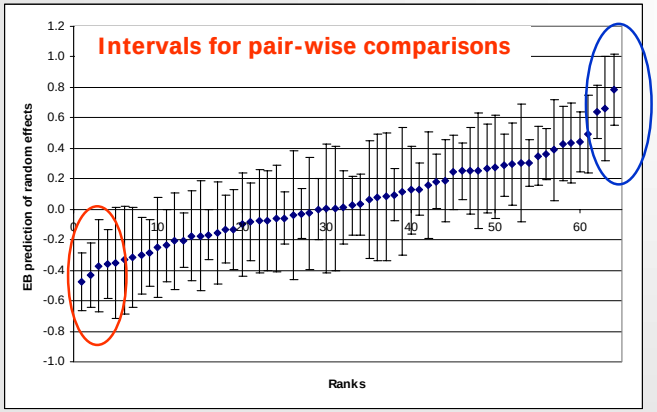
Type B effect of school *j*

Both effects are uniform (same effect for all the students)

- Constant slopes → parallel regression lines
- Unique ranking of the schools
  - ranking on Type A effects to inform potential students
  - ranking on Type B effects for accountability



Uncertainty about the school rankings



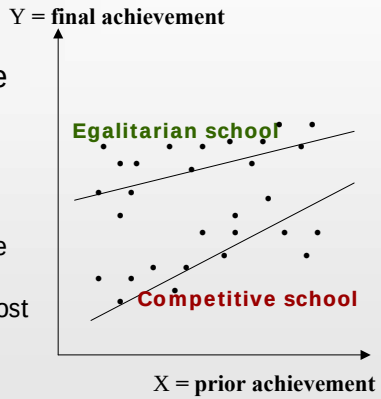
Only top and bottom schools are significantly different!

To inform school choice we need future rather than past effectiveness → larger error bars → comparisons are even more inconclusive

Leckie G, Goldstein H (2009) The limitations of using school league tables to inform school choice. JRSS A (forthcoming)

From uniform to varying school effects

- Uniform effects are often a restrictive assumption
- Typically a given school practice has more or less impact on student learning depending on the kind of student under consideration:
  - **Egalitarian schools** try to reduce the gap in the prior achievement
  - **Competitive schools** tend to boost the initial differences
- In statistical terms: competitive schools have an higher slope on prior achievement

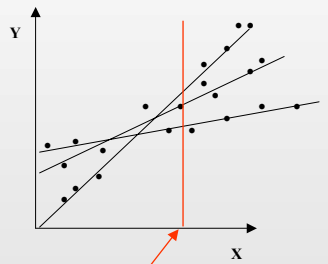


Random slope model

$$Y_{ij} = \left(\alpha + u_{0j}\right) + \left(\beta + u_{1j}\right) x_{ij} + e_{ij}$$

Random intercept of school j — Random slope of school j

- Random slopes → crossing regression lines
- Varying effects → different school effects, depending on student characteristics
- No unique ranking of the schools → different rankings conditionally on student characteristics



1. Define student profiles
2. Build rankings by profile

Models for non-hierarchical structures

- **cross-classified**, e.g. pupils are classified by primary and secondary school

	Secondary1	Secondary2	...
Primary1			
Primary2			
...			

- **multiple membership**, e.g. pupils change their school

$$\text{e.g. student } i \in \begin{cases} \text{school A} & \text{for 4/5} \\ \text{school B} & \text{for 1/5} \end{cases}$$

Goldstein H, Burgess S, McConnell B (2007) Modelling the effect of pupil mobility on school differences in educational achievement, *JRSS A*, 170, 941-954.

Achievement progress and measurement error

- Value-added models are based on measures of student achievement usually obtained through standardized tests
- The score of a test is a fallible measure of the true achievement (measurement error depends on reliability)
- The prior score is often used as a covariate in value-added models, causing measurement error bias (attenuation)
  - the school ranking may change: the effect of the prior achievement in not fully controlled for → schools with disadvantaged students are penalized

Ladd H.F. and Walsh R.P. (2002) Implementing value-added measures of school effectiveness: getting the incentives right. *Econ. Educ. Rev.*, 21, 1–17.

Ferrao ME, Goldstein H (2009) Adjusting for measurement error in the value added model: evidence from Portugal. Quality and Quantity (forthcoming)

Limitations of the value added approach

- Need more information to understand why some schools are more or less effective
- Studies of school effects are quasi-experiments → causal conclusions are questionable
- An effective adjustment for the input requires several good-quality covariates
- Measurement error in the covariates (especially prior achievement) may bias the slope estimates
- Difficult to fully account for all the uncertainty
- Difficult to communicate the results to a non specialized audience

Volumes from italian research projects on the evaluation of universities

- Chiandotto B, Grilli L, Rampichini C (Eds) (2005) Valutazione dei processi formativi di terzo livello: contributi metodologici, Collana Valmon n. 12, Università di Firenze. **http://valmon.ds.unifi.it**
- Boero G. and Staffolani S. (Eds) (2006) Performance accademica e tassi di abbandono. Un'analisi dei primi effetti della riforma universitaria. CUEC, Cagliari
- Fabbris L (Ed) (2007) Effectiveness of University Education in Italy: Employability, Competences, Human Capital, Heidelberg: Springer-Verlag.
- Capursi V, Ghellini G (Eds) (2008) Dottor Divago. Discernere, valutare e governare la nuova università. Franco Angeli.
- Bini M, Monari P, Piccolo D, Salmaso L (Eds) (2009, to appear), Satistical methods for the evaluation of educational services and quality of products. Physica-Verlag.

… where the present review is going to appear: ask me a copy at **grilli@ds.unifi.it**