

# Using a Multi-Level Model to Analyze the Effects of School-Level and Individual-Level Factors on Students' Academic Performance

---

Yu-Hsin Liu, National Taiwan University

## Introduction

---

In this work we analysis the Scottish School Leavers Survey (SSLS). SSLS is a representative survey conducted among young people. This survey is carried out in Scotland for students in their final year of compulsory education (ages 16-17) and includes data from several generations of students, covering 508 schools and 33,988 students.

- This analysis aims to answer:
- Are there any differences in average scores (SCORE) across schools without considering factors that affect individual grades?
- What is the proportion of variance in average school scores to the total variance in student scores? Does subsequent analysis need to consider the research level of the school (context)?
- If there is a difference in average scores among schools as mentioned in question (2), which variables among school characteristics (school type, location, and denomination) have an impact? Which variable has the greatest impact?
- After considering the impact of students' personal and family characteristics on scores, is there a difference in the adjusted average scores among schools? (In this report I assume that the impact of personal and family characteristics on scores does not vary between schools.)
- Among the impacts of students' personal and family characteristics on scores, which variables' effects will vary between school characteristics? For variables that vary, which school characteristic influences this difference?

In the analysis below, we set the significance level at 0.05, meaning significance is indicated by a  $p$ -value  $< 0.05$ .

Firstly, Without considering the factors affecting individual grades, is there a difference in average scores (SCORE) among schools?

1. Null multilevel model, *model1*

*level1*

$$score_{ij} = \beta_0 + u_{0j} + e_{ij}$$

*level2*

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

*combine*

$$score_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

- $score_{ij}$  = Score of student i in school j
- $\beta_0$  = Overall average (across schools)
- $u_{0j}$  = School's influence (school residual)
- $e_{ij}$  = Student residual

2. Null single model (remove random school effect)

$$score_{ij} = \beta_0 + e_{ij}$$

Next we conduct Likelihood ratio test, compare Null single model and Null multilevel model

Testing if  $Var(u_{0j})$  is significantly non-zero

$$LR = 2(143269.53 - -145144.42) = 3750, 1 d.f.$$

$$3750 > 3.84$$

$$3.84 = 5\% \text{ point of chi - square on } 1d.f$$

- Result: The difference between schools is significant.

Based on the previous model, what percentage of the total variance in student scores is accounted for by between-school variance? Do subsequent analyses need to consider the school (contextual) level?

Between-school variance proportion:

$$VAC = \frac{\sigma_{u0}^2}{\hat{\sigma}_e^2 + \sigma_{u0}^2} = \frac{61.02457}{61.02457 + 258.3572} = 19.1\%$$

- $\sigma_{u0}^2$  = Between school variance
- $\hat{\sigma}_e^2$  = Within school between student variance
- Result: 19.1% of the variance in student scores comes from between-school variance.

If there is a difference in average scores among schools as mentioned in question (2), then among school characteristics (type of school, location of school, and school denomination), which variables have an impact? Which variable has the largest impact?

- Added school level variables

*level1*

$$score_{ij} = \beta_{0j} + e_{ij}$$

*level2*

$$\beta_{0j} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolUrban_j + \gamma_{03} schoolDenom_j + u_{0j}$$

$$score_{ij} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolUrban_j + \gamma_{03} schoolDenom_j + u_{0j} + e_{ij}$$

- *model2*
- School type and location significantly affect student scores.
- Independent schools have better student scores compared to state-funded schools.
- Urban students perform worse than those in towns or rural areas.
- Comparing the absolute values of coefficients shows that school category has the largest impact.

After considering the impact of individual and family characteristics on grades, is there a difference in the adjusted average scores of schools?

*Individual level*

$$score_{ij} = \beta_{0j} + \beta_{1j} cohort90_{ij} + \beta_{2j} female_{ij} + \beta_{3j} sclass_{ij} + e_{ij}$$

*School level*

$$\beta_{0j} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolDemon_j + \gamma_{03} schoolUrban_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$score_{ij} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolDemon_j + \gamma_{03} schoolUrban_j + u_{0j} + \gamma$$

- *model3*

- According to the likelihood ratio test, the school level's influence remains significant after considering individual-level impacts.
- School category (schtype) and school location (schurban) significantly influence.

Among the effects of individual and family characteristics on grades, which variables' effects vary between school characteristics? For the variables that vary, which school characteristic influences this difference?

### Individual level

$$score_{ij} = \beta_{0j} + \beta_{1j} cohort90_{ij} + \beta_{2j} female_{ij} + \beta_{3j} sclass_{ij} + e_{ij}$$

### School level

$$\beta_{0j} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolDemon_j + \gamma_{03} schoolurban_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} schoolType_j + \gamma_{12} schoolDemon_j + \gamma_{13} schoolurban_j + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21} schoolType_j + \gamma_{22} schoolDemon_j + \gamma_{23} schoolurban_j + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31} schoolType_j + \gamma_{32} schoolDemon_j + \gamma_{33} schoolurban_j + u_{3j}$$

$$score_{ij} = \gamma_{00} + \gamma_{01} schoolType_j + \gamma_{02} schoolDemon_j + \gamma_{03} schoolurban_j + u_{0j} + (\gamma$$

Family background (whether a student comes from a managerial or professional class family, and whether a student belongs to an unclassified class), generation, and school type have significant interactions.

- *model4*

- Interaction between `sclass` and `schType` shows that overall, students from managerial and professional class backgrounds perform better than those from middle classes, who in turn perform better than unclassified students. However, the benefit of being in a private school over a state-funded school is greatest for unclassified students, less so for middle-class students, and least for managerial and professional class students.
- Interaction between `cohort90` and `schType` shows that overall, private schools perform better than state-funded schools across all generations, but the advantage of private schools diminishes with newer generations.

## Summary

---

See table 1 for a comprehensive model summary.

Among school-level factors, school type, particularly private schools, have a significant positive impact on student scores, while urban locations have a significant negative impact compared to rural or town settings.

On an individual level, females and newer generations perform better. Moreover, the advantage of private schools over state-funded schools diminishes with each new generation.

Furthermore, students from managerial and professional class families perform better than those from middle classes. Students from labor class and unclassified backgrounds perform worse than those from middle classes, especially those from unclassified backgrounds, for whom the negative impact of socio-economic background is greater. However, it is these unclassified students who benefit the most from school type, followed by middle-class students, with managerial and professional class students benefiting the least.

Tabel 1. Model estimated coefficient and standard error

	(1)	(2)	(3)	(4)
	score	score	score	score
score				
state				
independent		10.25*** (8.62)	4.398*** (5.42)	7.341*** (7.13)
town or rural				
urban		-3.128*** (-4.22)	-1.462** (-3.02)	-1.347** (-2.78)
non- denominational				
Roman Catholic		-0.970 (-1.06)	0.170 (0.28)	
cohort90			1.182*** (48.81)	1.212*** (49.57)
male				
female			1.967*** (12.75)	1.977*** (12.84)
intermediate				
managerial and professional			5.186*** (26.30)	5.442*** (26.87)
working			-5.847*** (-28.64)	-5.786*** (-28.15)
unclassified			-9.588*** (-33.63)	-9.703*** (-33.72)
intermediate # state				
managerial and professional # independent				-3.915*** (-4.47)

working # independent				3.883 (1.63)
unclassified # independent				9.463*** (4.66)
independent # cohort90				-0.614*** (-5.89)
Constant	30.60*** (82.83)	32.01*** (52.20)	31.10*** (73.32)	30.90*** (72.16)
lns1_1_1 Constant	2.056*** (56.06)	1.952*** (51.86)	-0.955	-0.988
lnsig_e Constant	2.777*** (718.32)	2.777*** (718.28)	2.631	2.630
lns1_1_2 Constant			1.497	1.508
atr1_1_1_2 Constant			-0.274	-0.221
Observations	33988	33988	33988	33988

*t* statistics in parentheses

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