教育研究學報 1992，第七卷， 41-49
Educational Research Journal 1992, Vol.7, pp. 41-49

A Study on School Effect: Analysis of HKAL Examination Results*

LAW Hing-chung
University of Sydney

Students from a stratified random sample of 33 schools are studied of the school effect on their HKAL results. It is found that schools with high average HKAL results are not necessarily the more effective schools. Even if adjusted for the intake of HKCE grades, in general, schools which are effective for more able students are found to be less effective for students with lower HKCE grades. Thus, it is very misleading to compare the effectiveness of school without reference to a particular range of student ability. The effects of changing schools in sixth forms and the school homogeneity are also explored. The results are not conclusive. However, it is suggested that students in the arts stream might have suffered from changing schools and the most disadvantaged students are those with low intake HKCE grades and study in schools whose students are of generally low ability.

本研究是利用從分層隨機抽樣的 33 所中學的學生，取其 1988 年香港高級程度會考成績及 1986 年香港中學會考成績進行分析。結果顯示：取得高級程度會考成績優越的學校，不一定屬效能較高的學校。學校效能會隨其學生在香港中學會考所取得的成績而不同，對成績較高及較低的學生比較顯著，而對中等程度的學生，學校的影響則不甚顯著。一般而言，對程度較佳的學生，其所屬學校的效能較高，而對程度較差的學生，其所屬學校的效能則較低。因此，即使利用中學會考成績進行調整，使用高級程度會考成績來排列學校效能的次序，必易得到使人誤導的結果。本研究未能對學生從中五到中六轉校的影響獲得結論。

There have been a growing interest in the study of school effects in the United Kingdom and the United States in recent years (Coleman, Hoffer & Kilgore 1982; Gray, Jesson & Jones 1986; Mortimore, Sammons, Stoll, Lewis & Ecob, 1988; Rutter, Maughan, Mortimore & Ouston 1979; Woodhouse & Goldstein, 1988). Some of these studies try to relate effectiveness of schools to school characteristics (e.g. average family background of students, availability of resources, class size, independent or public, single sex or co-educational etc). Some try to develop models to compare school effectiveness, especially after the Education Act 1988 in Britain, which stated that local education authorities are required to publish and rank schools based on public examination results.

In Hong Kong, although no attempts has yet been made to compare schools by public examination results officially, the results in the three major public examinations: Hong Kong Certificate of Education Examination (HKCE), Higher Level Examination (HKHL) and the Advanced Level Examination (HKAL) have, in many ways, been used as a benchmark of school achievement. It has almost been accepted by the general public that schools with good examination results are good schools. It is obvious that the outcome of education should be measured in many dimensions. Yet, even if we define measurable outcome by public examination results, schools scoring good 'outcome' scores might not be the most effective schools, if adjusted for the calibre of the intake of students and other factors. However, no empirical findings have been made in Hong Kong to support this belief. So the primary objective of this research is to find out which are the more effective schools.

The present study will be concentrated on the effectiveness of sixth-form courses of the schools. The two years of sixth form education form a coherent period in the education system, with self-contained curriculum in each of the subjects. The number of classes is typically small and school effects on the students would be comparatively uniform

* The Author would like to thank the Hong Kong Examinations Authority for granting permission and preparing the data for analysis. Parts of the study had been performed when the author was a visiting research associate at the School of Education, Macquarie University.

throughout the school. Also, for practical reasons, data for the analysis can be easily accessible. Each of the students has a standardised HKCE 'intake' score and a standardised HKAL 'outcome' score. However, restrictive as it is, the results are expected to be able to be generalised to students at other levels, or at least can be similarly analyzed, perhaps also for other attributes, if relevant data were available.

A related question of interest is the effect of changing schools. From our sample, which is believed to be representative of the population, it is found that more than 40% of the sixth form students have changed school from Form 5 to Form 6. How would this change of school generally affect the achievement of students? Of course, if the differences in school have no effect on the achievement and all results can be accounted by individual differences of students, it is expected than changing schools would have little impact on the students.

The third objective of this study is to investigate the effect of homogeneity of student ability in the school in individual students. Would students be disadvantaged or advantaged if studying in schools with students of wide range of ability? Of course, this effect of homogeneity is best studied at the class level because most of the interactions between students would have been taken place within the class rather than within the school. However, as the class level data are not available and also, in most of the schools, there would be no more than one class for a combination of subjects, the school effect would, to certain extent, reflect the class affect.

Method

The basic model used in the present study is to regress the 'outcome' variable y_{ij}, the HKAL result of the j-th student from the i-th school, on some student-level variables z_{tij}, one of which would be HKCE results, the 'intake' variable, together with some school-level variables z_{pj}, as follows:

$$y_{ij}=\beta_0+\sum_{p=1}^{m}\alpha_p x_{pi}+\sum_{t=1}^{n}\beta_{ti} z_{tij}+v_i+\epsilon_{ij} \quad , \qquad (1)$$

Here, separate random terms have been modelled for the school-level residual v_i and the student-level residual ε_{ij}, so that estimates can be made on for each the variances. The coefficients of the student level-variables β_{ti} can be fixed or random about the schools. If we take:

$$\beta_{ti}=\beta_t+\mu_i \quad , \qquad (2)$$

where μ_i is random across schools, the regressed 'slopes' of y_{ij} on x_{ij} would be modelled to be different for different schools.

It is well-known that multilevel models such as (1) can not be handled by ordinary least squares (discussions see, for example, Aitkin & Longford, 1986). However, in recent years, efficient estimates on such models have been developed by Goldstein (1986) using iterative generalised least squares, Longford (1988) using Fisher Scoring algorithm and Raudenbush and Bryk (1986) by EM algorithm, among others, assuming independence of random terms between different levels.

In this study, the software *ML3* (Rasbash, Prosser & Goldstein, 1989) have been used, giving efficient estimates of the parameters using iterated generalised least squares. if the random terms are multivariate normal, this is equivalent to maximum-likelihood (Goldstein, 1986).

The HKAL result (hereafter referred as AL Score) is taken be an aggregated score for the Advanced Level Examination. The AL Score is calculated from the sum of scores from fine grades of the best 3 subjects and the grade of Use of English in the 1988 Advanced Level Examination. In effect, the grade of th Use of English has been weighted as equivalent to one-third of any of the Advanced Level subjects. The HKCE result (hereafter referred as CE Score) is taken to be an aggregated score for the 1986 Hong Kong Certificate Examination, calculated in the similar way from the grades of English, Chinese and Mathematics together with the best 3 other subjecs in the 1986 Hong Kong Certificate of Education Examination. The maximum possible AL Score and CE Score would be 64 and 42 respectively.

The Sample

Schools taking part in the 1988 Hong Kong Avanced Level Examinatoin are classified into 3 strata according to the percentage of students getting an overall grade C or above in the Examination. A random sample of 11 schools from each stratum are taken, giving a total of 33 schools. The sample size is restricted by the constraint of the software at that time. However, a recent version of the software can handle a much larger data set. All candidates in the chosen schools are included in the study. The students in the sample are then matched, by identity card number, with candidates taking the 1986 HKCE Examination. Those who do not have a match or taking less than 6 subjects in the 1986 examination are deleted from the sample, since the focus of the present study would be on those who have not been repeating in two-year study after the Certificate of Education Examination. The size of the sample eventually included in the study is 1980.

The school code in 1988 HKAL Examination for each of the students is matched with his/her school code in the 1986 HKCE Examination, thus giving an indication of whether the student has changed school from Form 5 to Form 6. Also, candidates are classified into the arts or science stream. Science students are operationally defined as those who have taken Physics and/or Chemistry in the 1988 HKAL Examination. It is found that 50.6% of the candidates are in the science stream.

It is found that most of the chosen schools have about 50 candidates while 2 schools have more than 100 candidates and one school have less than 10 candidates. In the sample, 49.2% are female. 18 schools are found to be single-sex schools and there might be a slight over-representation comparing with the population of schools in Hong Kong. 46.8% of the students have changed schools in Form 6. There are 2 schools having all their students come from othe schools and 2 schools having no students come from other schools in Form 5.

The sample mean of AL Score and CE Score are 26.2 and 29.5 respectively. The correlation between the AL Scores and CE Scores is 0.73 showing that quite a substantial percentage of the variance of AL Score can be explained by the CE Score. Scatter plots of the AL Scores against the CE Scores shows no particiular 'ceiling' or 'floor effect', except that there are no points falling below CE Scores of 15, which is the minimum qualification for taking HKAL Examination at that time. The scatter is very 'normal' suggesting a linear relation between the two variables.

Results and Discussion

Effect of CE Score

Table 1 shows estimates of 3 models, A, B and C with AL Score regressed on the CE Scors. In Model A, no explanatory variables are included in the analysis except the constant term:

$$y_{ij}=\beta_0+v_i+\varepsilon_{ij} \qquad (3)$$

Here, the variance of the AL Score has been decomposed into between-school variance $var(v_i)=\sigma_v^2$ and within-school (between-student) variance $var(\varepsilon_{ij})=\sigma_e^2$, without making any adjustment to intake variables. The intraschool correlation from Model A can be estimated to be $\hat{\rho}=\hat{\sigma}_v^2/(\hat{\sigma}_v^2+\hat{\sigma}_e^2)=47.13/(47.13+110.30)=0.30$. This means that about 30% of the AL

TABLE 1
Effect of CE Score[a]

Parameter	A	B	C
Fixed			
Constant	28.78 (1.23)	-25.94 (1.94)	-23.82 (3.29)
CE Score	------	1.79 (0.06)	1.72 (0.10)
Random			
Level 2			
σ_μ^2	------	------	0.21 (0.09)
$\sigma_{\mu v}$	------	------	-6.42 (2.63)
σ_v^2	47.13 (12.26)	5.49 (1.79)	207.20 (2.63)
Level 1			
σ_e^2	110.30 (3.62)	81.91 (2.69)	79.64 (2.63)

[a]Standard errors of estimates in brackets

Score variance could be explained by the school differences. However, if CE Score is included as an explanatory variable, as in Model B:

$$y_{ij}=\beta_0+\beta_1 x_{ij}+v_i+\varepsilon_{ij} \ , \qquad (4)$$

the corresponding intraschool correlation is reduced to 5.49 / (5.49 + 81.91) = 0.06. That is to say, if adjusted for CE Scores, the school differences would only account for 6% of the AL Score variance. It is noted that in Model B, as compared with Model A, there are reductions of residual variances both at the school level and the student level, showing that both the between-school differences and within-school (between-student) differences can be explained, in part, by the intake CE Scores. However, the percentage reduction for the between-school variance is much greater than for the within-schools variance, thus giving rise to the notable decrease in intraschol correlation. Thus, any discussion on school effect and comparison of school effectiveness would be very misleading without taking into account of intake scores.

Model B assumes that the regression coefficient of CE Score would be the same for all schools, but as it is well-known that because of different teaching strategies, atmosphere, available of resources, experience of teachers, between-student 'contextual effect' etc, between schools, the regression of AL Scores on CE Scores might vary with schools, and it woule be more appropriate to model the regressed coefficients to be random about schools, as in Model C, as follows:

$$y_{ij}=\beta_0+(\beta_1+\mu_i)x_{ij}+v_i+\varepsilon_{ij} \ . \qquad (5)$$

Here, β_0 can be interpreted as the average intercept of all the regressed lines of AL Scores on CE Scores, with v_i being the deviate of the i-th school from the average. Similarly, β_1 can be interpreted as the average slope, μ_i being the deviate of the i-th school from the average slope. The estimates for this model is as shown in column C of Table 3. The variance of the slope is estimated to be 0.21 with standard error of 0.09. From this, we can calculate the standrd deviation $\sqrt{0.21} = 0.46$, which is quite substantial, considering the estimated mean slope being 1.72. The variance of the intercepts is estimated to be 207.20, with standard error 2.63. The standard deviation of the intercepts would then be $\sqrt{207.2} = 14.39$, comparable in size to the mean of the intercepts estimated to be -23.82 Since both σ^2_μ and $\sigma_{\mu v}$ are significant, C would be a more suitable model as compared to B. The mean intercept is found to be negative because the minimum CE Score in the sample is 15 and the predicted AL scores for CE Scores less than this value would reasonably be negative.

Furthermore, the predicated regressed line for each of the schools could be constructed from the estimated residual slope μ_i and residual intercept v_i. The predicted regressed line for the i-the school can be expressed as:

$$y_{ij} = (\hat\beta_0 + \hat v_i) + (\hat\beta_1 + \hat\mu_i)x_{ij} \quad . \tag{6}$$

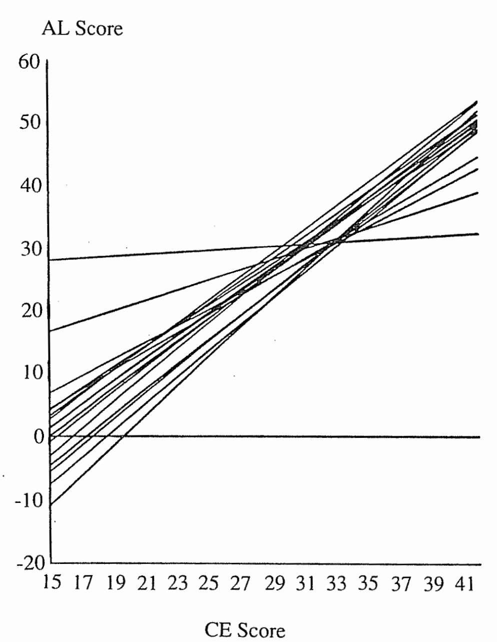

FIGURE 1. *Predicted Regressed Lines for Schools.*

Figure 1 shows the plots of the predicted regressed lines of the schools. For clarity of the graph, only 16 schools are included without any specific preference in the choice of schools. However, the fiture gives sufficient indications that schools with large intercepts tend to have smaller slopes. Indeed, if we calculate the estimated correlation between the slopes and the intercepts, $\hat\sigma_{\mu v}/(\hat\sigma_\mu \hat\sigma_v) = -6.42/\sqrt{0.21\times 207.2} = -0.97$. The value of the estimated correlation should be interpreted with caution because it depends on the choice of the origin of the CE Score. Since candidates are required to have a minimum score at CE before they can be qualified to take the Advanced Level, the CE Score could never be 0.0. If the origin of the CE Score is taken to be 15, say, the minimum CE Score in the sample, the estimated variance of the intercept is expected to be much smaller, leading to a correlation of smaller absolute value. It is not surprising to see that predicted regression lines of some schools fall below 0.0 when the CE Score is 15 because of interpolation. However, we find that all schools have positive predicted values of AL Score when the CE Score reaches 20.

Consider j-th student with CE Score x_{ij} in the i-th school, his/her predicted AL Score would be:

$$\hat y_{ij} = \hat\beta_0 + \hat v_i + (\hat\beta_1 + \hat\mu_i)x_{ij} \tag{7}$$

It is noted that for different schools, students of the same intake CE Score would have different predicted AL Scores. It would be possible to compare the school effectiveness by their predicted AL Scores. Consider s_{ij}, the difference between the predicted AL Score for the i-th school and the predicted AL Score using the overall regressed line, viz:

$$\hat s_{ij} = [\hat\beta_0 + \hat v_i + (\hat\beta_1 + \hat\mu_i)x_{ij}] - [\hat\beta_0 + \hat\beta_1 x_{ij}]$$

$$= \hat v_i + \hat\mu_i x_{ij} \quad . \tag{8}$$

We can define this value $\hat s_{ij}$ as the estimated school effect of the i-th school on the j-th student. We notice that this school effect varied with intake CE Score x_{ij}. Moreover, its variance over schools can be expressed as

$$var_i(\hat s_{ij}) = var_i(\hat\mu_i x_{ij} + \hat v_i)$$

$$= var_i(\hat\mu_i)x^2_{ij} + 2cov(\hat\mu_i, \hat v_i)x_{ij} + var_i(\hat v_i)$$

$$= \hat\sigma^2_\mu x^2_{ij} + 2\hat\sigma_{\mu v}x_{ij} + \hat\sigma^2_v \quad ,$$

which is a quadratic function of x_{ij}, with a minimum at $x_{ij} = -\sigma_{\mu v}/\sigma^2_\mu$. From Table 1, we have $\hat\sigma_{\mu v} = -6.42$ and $\hat\sigma^2_\mu = 0.21$, it is estimated that minimym variane of y_{ij}

TABLE 2

Predicted AL Scores in Different Schools for Different CE Scores[a]

School	CE Score=20	CE Score=30	CE Score=40
1 (28)	7.4 (25)	28.3 (12)	40.2 (5)
2 (3)	17.9 (3)	32.1 (2)	46.4 (15)
3 (23)	11.7 (16)	26.7 (23)	41.8 (25)
4 (31)	6.3 (27)	26.9 (21)	47.5 (10)
5 (32)	5.3 (28)	25.0 (30)	44.7 (20)
6 (30)	11.8 (15)	26.3 (26)	40.7 (27)
7 (16)	9.3 (21)	27.9 (17)	46.5 (14)
8 (25)	5.1 (29)	27.5 (19)	49.9 (1)
9 (2)	10.3 (20)	28.3 (13)	46.2 (16)
10 (17)	12.9 (8)	26.4 (25)	39.9 (31)
11 (10)	12.1 (12)	31.0 (3)	49.9 (2)
12 (6)	11.3 (17)	29.6 (7)	48.0 (7)
13 (20)	20.5 (1)	28.9 (9)	37.3 (33)
14 (7)	10.5 (18)	30.0 (6)	49.4 (3)
15 (8)	4.7 (31)	25.1 (29)	45.5 (18)
16 (1)	15.8 (5)	32.6 (1)	49.4 (4)
17 (11)	12.4 (10)	29.1 (15)	43.7 (21)
18 (23)	2.9 (32)	23.8 (32)	44.8 (19)
19 (15)	6.4 (26)	26.6 (24)	46.8 (13)
20 (9)	28.9 (2)	30.5 (4)	42.0 (24)
21 (13)	8.3 (23)	29.8 (10)	49.2 (6)
22 (5)	12.1 (13)	30.0 (5)	48.0 (8)
23 (13)	15.8 (4)	28.0 (16)	40.2 (29)
24 (21)	13.3 (7)	26.7 (22)	40.2 (30)
25 (33)	4.9 (30)	23.2 (33)	41.5 (26)
26 (4)	8.8 (22)	28.2 (14)	47.7 (9)
27 (29)	12.2 (11)	26.3 (27)	40.4 (28)
28 (18)	0.8 (33)	24.1 (31)	47.4 (11)
29 (19)	15.1 (6)	29.0 (8)	42.8 (23)
30 (22)	10.4 (19)	28.7 (11)	46.9 (12)
31 (27)	12.4 (9)	25.6 (28)	38.7 (32)
32 (12)	8.3 (24)	26.9 (20)	45.6 (17)
33 (26)	11.9 (14)	27.7 (18)	43.6 (22)

[a]Number in bracket represent rank order of school mean AL Score.

would occur at $x_{ij} = -(-6.42/0.21) = 30.6$. From Figure 1, we can see that with increasing CE Scores, the spread of the predicted AL Scores decreases from CE Score at 0 to a minimum at about 30 and then increases again. This suggests a very interesting point: for those who have a CE Score of about 30, that is, an average grade C in the 6 subjects, no matter what school they are attending, the predicted AL Score would be roughly the same, but for those with higher or lower grades, the choice of school would give greater differences in the predicted scores. In other words, the school effect would be relatively prominent for students of high and low intake scores and relatively low for students with medium CE Scores at 30.

TABLE 3

Rank Correlations of Predicted AL Scores for Different CE Scores

	CE=20	CE=30	CE=40	Mean AL Score
CE=20	1.00			
CE=30	0.55	1.00		
CE=40	-0.42	0.45	1.00	
Mean AL Score	0.28	0.65	0.39	1.00

It would be informative to calçulate the predicted AL Scores for different schools for a given CE Scores. Table 2 shows the predicted AL Scores for CE Score of 20, 30 and 40. If a student has a CE Score of 20 and if he/she attend School 1, his/her AL Score is predicted to be 7.4, for example. This predicted AL Score can be used as a measure of comparing school Effectiveness. For a student with CE Score of 20, say, he/she would be most advantaged should he/she attended School 13 where the predicted AL Score is 20.5 and be least advantaged should he/she have attended school 28. Table 2 also gives the rank order of the school effectiveness for CE Scores of 20, 30 and 24. For reference purpose, the rank order in the mean AL Score of each of the schools is also listed. It can be seen that most schools would not have similar rank order of schools at the 3 levels of CE Score. A school which is more effective for students at CE Scores of 20 would usually have a lower rank order of school effectiveness for students at CE Score of 40. Table 3 shows the rank correlations of school effectiveness between the 3 score levels. It is noted that the rank correlation of school effectiveness at CE Score 20 with that at CE Score 40 is -0.42, while the other rank orders are only moderately correlated. This show that there is a strong indication that there does not exist a single index is school effectiveness even if ajusted for the intake CE Score. Some schools would be more effective to students of high CE Scores and some to those with lower CE Scores. Only a small number of schools, for example School 16, show consistently high or low school effectiveness for students in the whole range of ability.

Change of School

As 46.5% of the students in the sample have changed schools form Form 5 to Form 6. In order to study the effect of changing school, a dummy variable w_{1ij} 'change-school' has been assigned at the student level such that it is 1.0 if he/she has not changed school and 0 otherwise. The model would then be:

$$y_{ij}=\beta_0+\beta_1 x_{ij}+\beta_2 w_{1ij}+v_i+\varepsilon_{ij} \quad . \tag{9}$$

The estimates AL scores on CE Scores and the 'change-school' are as shown in Model A of Table 4.

Comparing with Model B in Table 1, it is found that the inclusion of change-school as explanatory variable gives little difference in the estimates of the variances of the random parameters. The coefficient for change-school is 0.55 with standard error 0.56. That is to say, a student who changes school would have a predicted AL Score of 0.55 points lower that

TABLE 4
Effect of Change of Schools[a]

Parameter	A	B
Fixed		
Constant	-26.27 (1.93)	-27.77 (2.07)
CE Score	1.79 (0.06)	1.85 (0.06)
Change (Yes=0)	0.55 (0.56)	-1.03 (0.69)
Stream (Science=0)	------	-0.62 (0.64)
Change x Stream	------	3.40 (0.86)
Random		
Level 2		
σ_v^2	5.09 (1.69)	5.15 (1.70)
Level 1		
σ_e^2	81.95 (2.69)	81.00 (2.66)

[a]Standard errors of estimates in brackets

for those who have not, given the same intake CE Score. This is not significant but could worth further study with a larger sample of schools. Maybe further analysis can be performed to see whether this effect differs from schools to schools, or which kind of schools (or which school?) have better effects, by fitting the coefficient random.

Further analysis is performed to see whether the effect could be different for students of the arts and the science stream. A second dummy variable w_{2ij}, to denote the stream in which the student belongs, has been assigned so that it is 1.0 if the student is in the arts stream and 0 otherwise. This dummy variable, together with its interaction with the other dummy variable 'change-school,' is included. The model would then be:

$$y_{ij}=\beta_0+\beta_1 x_{ij}+\beta_2 w_{1ij}+\beta_3 w_{2ij}+ \beta_4 w_{1ij} w_{2ij}+v_i+\varepsilon_{ij} \tag{10}$$

The result of the estimates are as shown in Model B of Table 4. The interaction of changing schools and the stream has highly significant, showing that the effect would have differences between the arts and science students. Generally speaking, for arts students, those who have not changed schools have a predicted AL score of (3.40 - 0.62 - 1.03) = 1.75 higher than those who have changed schools. However, for those in the science stream, those who have changed schools would have 1.03 points lower than those have not.

Effect of School Homogeneity

The standard deviation of CE Scores of the schools can be used as a measure of homogeneity. A school with students of mixed ability would have a

TABLE 5
Effect of School Standard Deviations[a]

Parameter	A	B	C
Fixed			
Constant	-28.49 (2.51)	-25.13 (1.98)	------
CE Score	1.77 (0.06)	1.64 (0.10)	------
School SD	0.97 (0.65)	------	------
School SD × CE Score	------	0.042 (0.021)	------
'High'	------	------	24.62 (2.64)
'High' × CE Score	------	------	1.75 (0.12)
'High' × School SD × CE Score	------	------	0.007 (0.020)
'Low'	------	------	-23.97 (3.15)
'Low' × CE Score	------	------	1.39 (0.18)
'Low' × School SD × CE Score	------	------	0.100 (0.047)
Random			
Level 2			
σ^2_v	5.07 (1.68)	5.26 (1.73)	5.07 (2.07)
$\sigma^2_{v'}$	------	------	2.77 (1.84)
Level 1			
σ^2_e	81.90 (2.69)	81.80 (2.68)	81.69 (2.68)

[a]Standard errors of estimates in brackets

larger standard deviation in the CE Scores. The results of some of the models are as shown in Table 5. In model A, the CE Score and the standard deviation of the school are included as explanatory variables as follows:

$$y_{ij}=\beta_0+\beta_1 x_{ij}+\beta_2 z_i+v_i+\varepsilon_{ij} \ , \qquad (11)$$

where z_i is the standard deviation in HKCE of its students in the i-th school. The effect of school standard deviation is 0.97 with standard error 0.65, which is not signigicant at the 0.05 level. However, the effect is too substantial to be neglected. First of all, the effect is positive. That is to say, a heterogenous school would in general be beneficial to students. The standard deviation in CE Scores of schools ranges from 1.12 to 4.52 with an average of 3.09. A difference of standard deviation of 3 points would have an average effect of nearly 3 points for the student. This could almost be unexpected as many hold the view that a homogenous class or school would produce better teaching effects and that is the main argument to support streaming in school/class. Anyway, this deserves furrther exploration, so more refined models are used. It might also be possible the effect of homogeneity are not the same for all students, perhaps more favourable to the more able people and unfavourable to the weaker students. We may replace the standed deviation of schools by the interaction of CE Score and the standard deviation as follows:

$$y_{ij}=\beta_0+(\beta_1+\beta_2 z_i)x_{ij}+v_i+\varepsilon_{ij} \ , \qquad (12)$$

where z_i is the standard deviation of the i-th school and x_{ij} is the CE Score of j-th student in the i-th school. The result of this analysis is as shown in Model B of Table 5. It is found that the estimation of the interaction is significant and positive, equal to 0.042. That is to say, a school of a wide range of ability, would favour the more able students. Or could be interpreted in the other way, a school of smaller standard deviation would be unfavourable to students of low ability. For the same 2 schools of difference in standard deviation of 3 points, say, the predicted increase in AL Score due for a student of CE Score of 40 is 0.042x3x40=5.04 whereas that for another student of CE Score of 20 would be 0.042x3x20=2.52, a difference of about 2.5 points. It might be argued that the effect of homogeneity would be different for those schools with students of average higher intake score from those with lower intake scores. It might be worthwhile to give separate estimates for these two types of schools. The schools were divided into two groups, those mean CE Score greater than 29.5, the overall mean of CE Scores, and those schools with mean CE Score less than 29.5. A dummy variable 'high' w_{1ij} is so assigned that if the student belonged to the former group, w_{1ij} is 1.0 and 0 otherwise. Another dummy variable 'low' w_{2ij} for the latter group in the similar way. It is found that the number of schools in 'high' and 'low' groups are 22 and 11 respectively. The number of schools in the 'low' group has been relatively small because these schools generally have more students. The model would then be

$$y_{ij}=(\beta_{01}+\beta_{11}x_{ij}+\beta_{21}z_i x_{ij}+v_{1i})w_{1ij}+ (\beta_{02}+\beta_{12}x_{ij}+\beta_{22}z_i x_{ij}+v_{2i})w_{2ij}+\epsilon_{ij} \quad . \tag{13}$$

For cases where a student belongs to school in the 'high' group, that is by taking $w_{1ij}=1.0$, $w_{2ij}=0$, equation (11) would become (10), similarly for the 'low' group. Thus equation (11) gives separate estimates for the 'high' and 'low' group, in a more efficient way than performing two separate regressions. Note that while we have the same student level residuals, we can have separate school level residuals. The covariance of the school level residuals must be modelled to be 0 since there would not be a single term with both $w_{1ij}v_i$ and $w_{2ij}v_i$ non-zero. The estimates are as shown column C of Table 5.

It is seen that for the 'high' group, the estimate for the interaction of standard deviation and CE Score is very small and insignificant. It is found that for schools with intake of better students, the homogeneity does not seem to have any significant effect on the students. There would be no difference whether the student is placed in a homogenous or heterogenous school. But, the estimate for the 'low' group, 0.100, is significant and substantial, despite the small number of schools in this group. That is to say, for a student of CE Score 20, the difference of predicted AL Score for school of standard deviation 2.0 to a school of 4.0 would be 0.100x(2.0-4.0)x20=4. It seems that the most disadvantaged group is those students of low CE Scores and being admitted to a school of heterogenous ability (probably generally low). If he/she is being admitted to a more homogenous school, some of the better students could give positive 'contextual' effect on the student.

Conclusion

It must be reiterated that the study is only based on HKAL results and school effect is only based the on measurement of outcome adjusted for income model. It is well known that there are other attributes, possibly in other domains and maybe more important, should be included. Or perhaps school effect should be evaluated through the analysis of the process inside schools/classrooms. Yet, however crude our models are, there are some interesting results, showing a general trend and it is expected that measurements in other domain may show similar picture.

First of all, we have found that there does not exist a unique index of school effectiveness. If we use the predicted AL Score for a particular school as a measure of its effectiveness, we found that the rank orders of effectiveness in the schools would be very different at different levels of student intake score. In particular, the rank correlation of those with CE Score 20 (low achievers) and CE Score (high achievers) is found to be -0.42. Moreover, we observed that the rank order correlation of school AL Score mean correlates only very mildly with the school effectiveness at any level of intake CE Score. This could have been expected. As schools in Hong Kong are relatively homogenous in intake and teaching is based on a narrow range of students' ability, a student would be very disadvantaged if he/she is being admitted to a school not gear to his/her ability. It is very misleading to compare schools by the performance of AL results, without taking into account of the intake scores, and also it would be meaningless to compare school effectiveness without reference to the range in intake scores of the students under consideration. It is hoped that this would clear off, to a certain extent, the perpetual belief that schools/classes can be compared simply by their average examination achievements. The same would apply to school effects in terms student behaviour, for example. People tend to rate schools by observing student behaviour as it appears. Without considering the intake and other factors, say family background, it would be very difficult to evaluate what the schools have achieved. If the data for intake score and outcome score are available for a number of schools or classes or any cluster under study, instead of comparing their effectiveness, it could be more worthwhile to establish a chart for student guidance purposes, so that a student, for a given intake score, would know which school(s) would be most beneficial to him/her.

The effect on change of school is not conclusive. Yet, at a lower confidence level, arts stream students tend to be benefited from the change of schools while those in the science stream would disadvantaged from the change of school, for reasons deserve further explanation. Further exploration on this area could be possible to provide more information for the debate on pooling resources in sixth forms schools, in which case more students would have to change school. Preliminary results might show that this would have more impact on the arts students than the science students.

As for the effect of homogeneity of schools, for schools with above average intakes, the homogeneity of school does not have significant effect on the AL Score. However, the general pattern may suggest that, in general, heterogenous schools would be more beneficial to students. In particular, for schools of average low intake, there is a significant effect on the interaction of intake score and the school standard deviation. That is to say, the most disadvantaged students would be those with low intake score and attending a school of heterogeneously low ability.

This could only serves as a preliminary study of the school effect. Further study can be performed in a number of possible directions. School effectiveness can also be measured through other educational outcomes in social and behavioural arenas. Addition information such as type of school, resources available to school, teaching methods, school atmosphere, teacher qualification etc, might be obtained and included as school level variable. Furthermore data for a number of years can be collected so that three level models can be constrcted: Student at level 1, year at level 2 and school at level 3, so that the stability of school effects over the years can be estimated.

References

Aitken, M., & Longford, N. (1986). Statistical modelling in school effectiveness studies. *Journal of Royal Statistical Society, A, 149* (1), 1-43.

Coleman, J. S., Hoffer, T., & Kilgore, S. B. (1982). *High school achievement: Public, Catholic and other Schools compared.* New York: Basic Books.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika, 73* (1), 43-56.

Gray, J, Jesson, D., & Jones, B. (1986). The search for a fairer way of comparing school's examination results. *Research Papers in education, 1*(2), 91-122.

Longford, N. (1987). *Fisher scoring algorithm for variance component analysis with hierarchically nested random effect.* Princeton: Education Testing Services (Research Report No. 87-32).

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters.* Londan: Open Books.

Rasbash, J., Prosser, R., & Goldstein, H. (1989). *ML2: Software for two-level analysis (Three-level supplement).* London: Institute of Education.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*(1), 1-17.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen Thousand hours.* London: Open Books.

Woodhouse, G., & Goldstein, H. (1988). Educational performance indicators and LEA league tables. *Oxford Review of Education, 14* (3), 301-320.

Author

LAW Hing-chung, Associate Lecturer, Department of Behavioural Sciences, Faculty of Health Sciences, University of Sydney, Sydney, Australia.