School effectiveness in mathematics in Sweden and Norway 2003, 2007 and 2011

Marie Wiberg[1] (marie.wiberg@stat.umu.se), Ewa Rolfsman[2] (ewa.rolfsman@edusci.umu.se), & Inga Laukaityte[1] (inga.laukaityte@stat.umu.se)
[1]Department of Statistics, Umeå School of Business and Economics, Umeå University, Sweden.
[2]Department of Applied Educational Science, Umeå University, Sweden.

Abstract

International studies like TIMSS (Trends in International Mathematics and Science Study) usually highlight students' study result changes between two assessments. The aim of this study is to identify factors that contribute to the explanation why some schools is effective but others are less effective in terms of the students' academic achievement in mathematics on TIMSS. We focus on the school efficiency in Sweden and Norway at the time points 2003, 2007 and 2011. In the core subject mathematics, Sweden's result declined in TIMSS between 2003 and 2011 while Norway's result inclined. Since Sweden and Norway have similar educational systems it is of interest to examine the incline in Norway as opposed to the decline in Sweden between 2003 and 2011. We used multilevel analysis in order to separate the effect of school level variables from the effect of student's home environment and to take care of the sampling design used in TIMSS. The results show that Norway and Sweden exhibit different trends. In Norway it was possible to identify school level factors which affect the students' achievements. In Sweden this could not be obtained. Note, in both countries at all three time points most of the student home background factors were significant suggesting that the students' background play a large role in mathematics achievement.

Keywords: Multilevel analysis, school level factors, efficient schools, country comparison

Introduction

Declining study results among Swedish students in mathematics obtained from international large scale assessments, e.g. TIMSS (Trends in International Mathematics and Science Study) have started a debate about learning outcomes and quality in education. School-effectiveness is therefore an urgent issue. Since Sweden and Norway have similar educational system among the Nordic countries and have both been part of TIMSS at several occasions they are of interest to compare at different time points. The inclining results in Norway, although lower than the Swedish result has been part of our interest. This study will contribute to the understanding of why different schools are more or less effective by contrasting school level factors influencing mathematics achievement in Sweden with Norway measured at three different time points. This is of particular interest in the light of results showing that there are significant differences between schools with reference to student performance where the school segregation has become increasingly prominent in Sweden in recent years (Swedish National Agency for Education, 2012).

A vast amount of studies have been conducted regarding school-effectiveness from different perspectives although studies based on TIMSS data occur less frequent (see e.g. Akiba 2008; Hoffman & Brahier, 2008; Dumay & Dupriez, 2007; Wiberg & Andersson, 2010). This study deal with the issue of school-effectiveness in the context of factors that contribute to the explanation why some schools in Sweden are effective but others are less effective in terms of the students' academic achievement in mathematics on TIMSS 2003, 2007 and 2011. This is conducted by investigating and comparing school-effectiveness in Norway, a country similar to Sweden in many aspects. Many studies use students' achievement in core subjects (e.g. mathematics, language and science) as criteria for school-effectiveness (Neuschmitdt, Henke, Rutkowski, & Rutkowski, 2003). We share this view, and we control for the effect of the student's home environment since it is well-known that there is a correlation between student's socio-economical background and school achievement (see e.g. Mullis, Martin, & Foy, 2008; Swedish National Agency for Education, 2008). We are also inspired by Martin, Mullis, Gregory, Hoyle, Shen, (2000) and view a school as effective if it adds value through the realization of the student body through effective instruction and efficient organization. This also means that our main focus is the school level.

The main purpose of this study is to identify factors that contribute to the explanation why some schools in Sweden and Norway are effective in terms of the students' academic achievement in mathematics. Also, what factors influence the effectiveness, and if they are different at the years 2003, 2007 and 2011.

Methodology

Data/participants: We used data from TIMSS 2003, 2007, and 2011 for grade 8th in Mathematics (IEA, 2003; 2007; 2011) from Sweden and Norway. TIMSS is given ever fourth year to one or two whole classes from randomly selected schools of 8th grade within each country. The students' mathematics achievements and their questionnaires were used together with the principals' school questionnaire.

Statistical analysis: Multilevel analysis (e.g. Gelman & Hill, 2007) with ideas from Ma, Ma & Bradley (2008) was used to examining school level factors when controlling for the student home backgrounds. This strategy was chosen due to the sampling procedure used in TIMSS (Kyriakides & Charalambous, 2005) and since we wished to control for factors which are not connected to school effectiveness. In the descriptive statistical analysis IEA IDB analyzer was used to prepare the files and analyze them together with SPSS 20.

We constructed two home background factors (see Table 1 for descriptive statistics) based on previous studies of TIMSS (Mohammadpour and Ghafar, 2012; Wiberg & Andersson, 2010; Wiberg & Rolfsman, 2013) together with two variables (student's sex and if the father was born in Norway/Sweden or not).

[MSC] Mathematics-self-concept: Includes four items focusing on student's own opinion about mathematics. 1) I usually do well in mathematics, 2) Mathematics is more difficult for me than for many of my classmates, 3) Mathematics is not one of my strengths, and 4) I learn things quickly in mathematics.

[SES] Socioeconomic status: This is based on the items: A) Number of books at home. B) Home resources (study desk, computer).

In general, missing data was low in the home background variables, ranging from 0.02% (for SEX in Sweden 2003) to 7.29% (for MSC in Sweden 2007). Thus we used listwise deletion to exclude missing data (Tabachnick & Fidell, 2007). Although it might have been better to impute missing data this choice was made due to the small amount of cases which were removed.

Table 1. Student level factors descriptions with mean and standard deviation within parenthesis for the first two factors and proportion (%) for sex and native fathers.

Factor	Description of factor	Norway			Sweden		
		2003	2007	2011	2003	2007	2011
MSC	Mathematics-self-concept	2.78 (0.02)	2.31 (0.01)	2.38 (0.02)	2.82 (0.02)	2.34 (0.02)	2.26 (0.01)
SES	Socioeconomic status	2.07 (0.02)	1.99 (0.02)	1.98 (0.02)	2.10 (0.02)	2.01 (0.02)	1.91 (0.02)
FB	% native fathers	89.8	85.1	84.4	86.5	79.5	77.0
SEX	% male	50.1	50.5	48.6	50.4	52.2	48.1

We examined the effectiveness of schools by assuming that a school is more effective if its mean mathematics achievement is higher than predicted from multiple regressions of the identified home background factors. This means that we examined the mean difference between the five mathematics plausible values and the expected scores from multiple regressions using the student home background factors and variables at each time point within each country. We regarded schools as effective if they were in the top third, mid-effective if they were in the mid third and less effective if they were in the bottom third in mathematics achievement in their country. Next, we examined school level factors which had potential to influence school effectiveness (see Table 2 for descriptive statistics). To decide which school level factors to use, we constructed factors based on their theoretical construct, partly inspired by Mohammadpour and Ghafar (2012), who has conducted a study which identifies school level factors which are significant for some of the countries participating in one of the TIMSS-studies and partly based on variables available in TIMSS over the years. Especially we chose variables which we could find at all three time points in both countries. This step was challenging since many items are rephrased at a later time point or a country use the possibility of national adaptions. Here we describe all school level factors we initially constructed. Note that two of the initially constructed school level factors, denoted TC and PP, could not be included in the later analysis, since the questions in the school questionnaire in 2011 were not comparable to those of 2007 and 2003. Descriptive statistics is given in Table 2 for the school level factors which were significant for at least one time point in any of the countries.

[PSA] = *Poor student attendance at school*. This category is based on the principals' estimates of what degree students arriving late at school and absenteeism occur or is a problem at the school.

[SCT] = *School climate teacher*. This category is based on the principals' estimate of how the teachers at the school view their job satisfaction and the principals' estimates of the teachers' abilities to understand the schools' curricular goals, their success in implementing the school's curriculum and the teachers' expectation for student achievement.

[LMR] = *Lack of school resources for mathematics.* This category is based on the principals' estimates of the extent the schools capacity to provide instruction is effected by a shortage or inadequacy of resources, such as computers, software, calculators, library materials and audio-visual materials for mathematics instruction

[SLOC] = *School location.* This is based on the principals' answer to the question; how many people live in the city, town, or area where your school is located?

[NSB] = *Negative student behavior.* This category is based on the principals' estimates of what degree the kind of student negative behavior described as classroom disturbance, intimation or verbal abuse among students or of teachers or staff, occurs or is a problem at the school. Not significant at any time points.

[PI] = *Parental Involvement* This category is based in the principals' estimates of the extent parents are asked to be involved in school activities, such as volunteer for school projects, programs, trips and serve on school committees. Not significant at any time points.

[TC] = *Teachers competence.* This category is based on the principals' view of the teacher professional developing opportunities for mathematics targeting issues such as improving teaching skills, improving content knowledge and using information and communication technology for educational purposes. Not used in the later analysis.

[PP] = *Principals pedagogical activities*. This category is based on the principals' view of his/hers own pedagogical activities such as time spent on instructional leadership and supervising and evaluating teachers and other staff. Not used in the later analysis.

Table 2. School level factors' with mean and standard deviation within parenthesis for the first three factors and proportion of urban schools for school location.

Factor	Norway			Sweden		
	2003	2007	2011	2003	2007	2011
PSA	1.74	1.87	1.72	2.09	2.24	2.16
	(0.07)	(0.06)	(0.06)	(0.07)	(0.05)	(0.07)
SCT	2.29	2.30	2.51	2.46	2.40	2.38
	(0.05)	(0.07)	(0.08)	(0.03)	(0.05)	(0.07)
LMR	2.24	2.12	1.58	1.74	1.73	1.85
	(0.10)	(0.12)	(0.07)	(0.06)	(0.07)	(0.07)
SLOC	45.40	40.16	45.31	54.60	59.84	54.69

In the multilevel analysis we used the software MPLUS 7 (Muthén & Muthén, 2011). For each year we examined the null, context and full model. We used the five mathematics achievement plausible values as dependent variables and grand mean centering of the variables. Before evaluating and comparing the different multilevel models we ran the null model and calculated the intraclass correlation (ICC) to test how homogeneous the data are within group level clusters. In the null model mathematics achievement for each student was estimated as a function of the school average with a random error.

Level 1 (within schools): $Y_{ij} = \beta_{0j} + r_{ij}, \; i = 1, \ldots, N,$

Level 2 (between schools): $\beta_{0j} = \gamma_{00} + u_{0j}, \; j = 1, \ldots, J,$

where Y_{ij} denotes mathematics achievement for student *i* within school *j*, β_{0j} is the average mathematics achievement for school *j*, and r_{ij} is the error term representing a unique effect associated with student *i* in school *j*. Level 2 term γ_{00} denotes the grand mean of mathematics achievement and u_{0j} is the error term representing a unique effect associated with school *j*.

The ICC was used to calculate the design effect, which shows how much the standard errors are underestimated. The design effect can be obtained from

$$\text{Design effect} = 1 + (\text{average cluster size - 1}) \cdot \text{ICC}.$$

A design effect greater than two indicates that the clustering of the data needs to be taken into account during estimation (Kish, 1965). In all the models the design effect was greater than two: from 3.31 (in 2003) to 4.36 (in 2011) in Norway, and from 4.50 (in 2007) to 6.72 (in 2003) in Sweden. This indicates that the grouping of the data needs to be taken into account during estimation.

In the home context model the mathematics achievement for each student was estimated as a function of the school mean achievement and the student effect. At the first level of the model, students' home background factors ($H_1, \ldots, H_f$) and at the second level aggregated student level factors ($aH_1, \ldots, aH_f$) were included. Student level factors were weighted by total student weights (TOTWGT).

Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(H_1) + \beta_{2j}(H_2) + ... + \beta_{fj}(H_f) + r_{ij}, \; i = 1, \ldots, N,$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(aH_1) + \gamma_{02}(aH_2) + ... + \gamma_{0f}(aH_f) + u_{0j}, \; j = 1, \ldots, J,$$

$$\beta_{1j} = \gamma_{10}, \beta_{2j} = \gamma_{20}, \ldots, \beta_{fj} = \gamma_{f0},$$

In the full model the association of school level factors with student mathematics achievement was quantified while controlling for students’ home background. The previously defined home context model was supplemented by school level factors ($S_1, ..., S_l$) weighted with school weights (SCHWGT). The student level weights were recalculated by multiplying weighting factors and nonresponse adjustments for class and student.

Level 1 (within schools):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(H_1) + \beta_{2j}(H_2) + ... + \beta_{fj}(H_f) + r_{ij}, \; i = 1, \ldots, N,$$

Level 2 (between schools):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(aH_1) + \gamma_{02}(aH_2) + ... + \gamma_{0f}(aH_f) + \gamma_{0(f+1)}(S_1) + ... + \gamma_{0l}(S_l) + u_{0j}$$

$$\beta_{0j} = \gamma_{00} + \mu_{0j}, \; \beta_{1j} = \gamma_{10}, \; \beta_{2j} = \gamma_{20} \ldots \; \beta_{lj} = \gamma_{lj}.$$

For the best model selection Bayesian information criterion (BIC) was chosen. The advantage of BIC (Schwartz, 1978) is a possibility to compare competing models (whether or not they are nested) as long as the sample remains constant (McCoach & Black, 2008).

Results

The results clearly show that Sweden and Norway are different from each other with reference to exhibited trends in student achievement, and with reference to school level factors associated with student performance. First, by examining the effectiveness in school (see Table 3), it is evident that the less effective schools in Sweden are almost at the same level of student achievement in 2003, 2007 and 2011 even though the average achievement is declining. This is also shown both in the mid effective and the more effective schools. In Norway, the upward trends in student achievement are shown at all levels, i.e. regardless of whether the school is effective or not.

Table 3. Average mathematics achievement with standard errors in parenthesis; and in less-, mid-, and more effective schools in Norway and Sweden in 2003, 2007 and 2011. The ICC attributable to schools, between school variances, with proportion of between school variances explained by the chosen multilevel model within parenthesis.

Country	Norway			Sweden		
	2003	2007	2011	2003	2007	2011
Average	461 (2.5)	469 (2.0)	475 (2.4)	499 (2.6)	491 (2.3)	484 (1.9)
Less effective	434 (2.9)	448 (2.0)	449 (2.7)	458 (3.7)	467 (2.5)	459 (2.9)
Mid effective	462 (2.5)	467 (1.8)	479 (2.3)	504 (2.7)	494 (2.0)	485 (1.7)
More effective	487 (2.7)	493 (2.4)	499 (2.7)	535 (2.6)	515 (2.4)	508 (2.5)
ICC schools	0.088	0.085	0.126	0.238	0.121	0.123
Between school variance	445.61 (0.58)	354.43 (0.60)	548.80 (0.44)	1267.85 (0.66)	567.42 (0.58)	596.43 (0.67)

The ICC from the multilevel analyses, showing how much of the total variance in mathematics achievement that is attributable to schools in Norway and Sweden at the three time points is presented in Table 3. The ICC for Norway was about 9% (in 2003 and 2007) and 13% (in 2011). For Sweden the ICC were 24% (in 2003) and only 12% (in 2003 and 2007). Also note the relative large proportion of between school variance explained by the chosen multilevel model. A high proportion of between-school variance indicates the existence of school effects (Ma, Ma & Bradley, 2008). In Table 4 the results from the full or/and in some cases the home context models are summarised. To facilitate comparisons of the models we used only those home background factors that are significant in all the models, except FB which was not significant in the full model for Norway 2011 data. We decided not to use the variable SEX because it was only significant in the models for Sweden 2007, Norway 2003 and 2007. Note, the student level factor mathematics self-concept adds a larger value in both countries at all three assessments than e.g. socioeconomic status. The significance of the aggregated student level variables was different at different years and countries. Aggregated socioeconomic status (aSES) was significant in both analysed countries in all the models. However, aggregated mathematics self-concept (aMSC) was significant only in 2003 for both countries, and aggregated father born (aFB) was significant only in the model for Sweden 2003. A surprising result was that school level factors didn't show any strong relation to school effectiveness in Sweden at any of the time points. This pattern was not seen in Norway where different school level factors were significant at different time points. It is, however worth noting, that all significant school level factors in Norway seem to decrease the mathematics achievement, except the school climate from a teacher's point of view (SCT).

Table 4. Parameter estimates for the home context and full multilevel models.

	Level 1			Level 2						
	MSC	SES	FB	aMSC	aSES	aFB	PSA	SCT	SLOC	LMR
Norway										
2003	44.73***	17.92***	22.04***	35.22***	23.81**				-8.02*	
	(1.77)	(2.31)	(5.42)	(9.31)	(8.62)				(4.00)	
2007	42.66***	18.55***	12.95***		37.62***					-9.97**
	(1.74)	(2.03)	(3.53)		(8.68)					(3.63)
2011	54.59***	17.47***	6.81		33.02**		-9.72*	7.81*		
	(2.49)	(2.36)	(3.50)		(10.52)		(4.25)	(3.93)		
Sweden										
2003	41.04***	21.50***	22.41***	34.06***	28.74**	49.16*				
	(1.49)	(1.57)	(3.30)	(9.53)	(8.57)	(19.32)				
2007	44.01***	20.67***	15.60***		31.76***					
	(1.75)	(1.56)	(2.71)		(6.76)					
2011	55.17***	17.99***	12.30***		48.33***					
	(1.40)	(1.42)	(2.32)		(5.97)					

*** - p-value < 0.001, ** - p-value < 0.01, * - p-value < 0.05.
MSC = Mathematics self-concept, SES =socioeconomic status, FB = father born, aMSC, aSES, and aFB are aggregated student level factors. PSA = Poor student attendance at school, SCT = School climate teacher, SLOC = School location, LMR =Lack of school resources for mathematics.

Discussion and concluding remarks

Our study makes a contribution to the urgent issue of enhancing the effectiveness in our schools manifested by student achievements. We contrasted Norway and Sweden, since they are the only Nordic countries participating in the three preceding TIMSS-studies, i.e. in 2003, 2007 and 2011, with grade 8th students. We aimed at identifying core school level factors associated to student achievement results, while controlling for student level factors associated with student success. We used socioeconomic status, if the father was born within the examined country and mathematics self-concept to control for student home background.

It was especially interesting to note the high coefficient value of mathematics-self-concept at all three time points in both countries.

The results from the multilevel analysis were somewhat surprising for Sweden because, in contrast to Norway, we have not found any significant school level factors. Our previous studies on TIMSS data from 2003 and 2007 has shown that there are some school level factors which are influential in Sweden (Rolfsman & Wiberg, 2012). This is a result that may be related to our method, i.e. comparisons of the two countries at three time point, where we were forced to exclude some of the questions due to alternation of existing questions or response options or simply that some of the questions in the questionnaires has been replaced by new items. An example of this, are questions related to the principal's pedagogical activities (PP) and teacher competence (TC) which are not included in this study, due to alterations of the questionnaires over the years.

The fact that we failed to identify similar or same school level factors at the three different time points in Norway may also be related to our method. However, this may also be a result reflecting that there are areas that seem to lack importance for school efficiency in regard to factors that has been identified but also whether the questions in the questionnaires actually measure the intended concept or processes. Still, the comparison of the two countries is of interest because the study results among students in Norway and in Sweden seem to move in different directions. In addition, the results indicate that the schools in the two countries seem to play different roles in this development. While the Swedish schools appear to have less influence over the results in 2011 compared to in 2003 and in 2007, because the total variance in mathematics achievement that is attributable to schools has declined over the years, the Norwegian schools has moved in the opposite direction. Still, the total amount of variance attributable to the schools is at the same level in both countries in 2011. Nevertheless, the results raise the questions whether studies using TIMSS questionnaire data can (or should) include comparisons over more than two time points?

References

Akiba, M. (2008). Predictors of student fear of school violence: a comparative study of eight graders in 33 countries. *School Effectiveness and School Improvement, 19* (1), 51-72.

Dumay, X., & Dupriez, V. (2007). Accounting for class effect using the TIMSS 2003 eight-grade database: Net effect of group composition, net effect of class process, and joint effect. *School Effectiveness and School Improvement, 18* (4), 383-408.

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University press.

Hoffman, L. R., & Brahier, D. J. (2008). Improving the planning & teaching of mathematics by reflecting on research. *Mathematics Teaching in the Middle School, 13* (7), 412-417.

IEA. (2003). TIMSS 2003 international database. Retrieved September 2011 from http://timss.bc.edu/timss2003.html

IEA. (2007). TIMSS 2007 international database. Obtained from the official TIMSS 2007 International database CD.

IEA. (2011). TIMSS 2011 international database. Retrieved April 2013 from http://timss.bc.edu/timss2011/

Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.

Kyriakides, L. & Charalambous, C. (2005). Using educational effectiveness research to design international comparative studies: turning limitations into new perspectives. *Research paper in Education, 20 (4)*, 391-412.

Ma, X., Ma, L., & Bradley, K. (2008). Using multilevel modelling to investigate school effects. In A. A. O´Connell & D. B. McCoach (Eds.) *Multilevel modelling of educational data*, Chapter 3, pp. 59-110. Information Age Publishing: Charlotte, NC.

Martin, M.O., Mullis, I.V.S, Gregory, K.D., Hoyle, C.D., Shen, C. (2000). *Effective Schools in Science and mathematics. IEA´s Third International Mathematics and Science Study.* TIMSS International Study Center. Boston: Boston College.

McCoach, B.D. & Black, A.C. (2008). Evaluation of model fit and adequacy. In A.A. O'Connell & D.B. McCoach (Eds.), *Multilevel modeling of educational data* (pp.245-272). Charlotte, NC: Information age publishing inc.

Muthén, L. K. & Muthén, B. O. (2011). MPLUS (Version 7) [computer software]. Los Angeles, CA.

Mullis, I.V.S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 International Mathematical Report. Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eight Grades.* Boston College: TIMSS & PIRLS International Study Center, Lynch School of Education.

Neuschmidt, O., Henke, J., Rutkowski, L., & Rutkowski, D. (2003). *Effective schools in Arab educational systems: An analysis of TIMSS 2003.* Retrieved October 26, 2009 from IRC 2008 – TIMSS–Mathematics. International Association for the Evaluation of Educational Achievement. http://www.iea.nl/irc2008_timssmath.html?&no_cache=1&sword_list[]=countries

Rolfsman E., & Wiberg, M. (2012). School effectiveness and mathematical achievement results from TIMSS in Sweden and Norway. *Manuscript submitted for publication.*

Schwarz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics*, 6, 461-464.

Swedish National Agency for Education (2008). *TIMSS Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv* [*TIMSS Swedish students' knowledge in mathematics and science in an international perspective.*] Report 323. Stockholm: Skolverket.

Swedish National Agency for Education (2012): *Likvärdig utbildning i svensk grundskola? En kvantitativ analys av likvärdighet över tid* [Fairness in Education. A quantitative analysis of fairness over time] Report 374. Stockholm: Skolverket.

Wiberg, M. & Rolfsman, E. (2013). School effectiveness in science in Sweden and Norway viewed from a TIMSS perspective. *Accepted manuscript.*