LLAKES
CENTRE FOR LEARNING AND LIFE CHANCES IN KNOWLEDGE ECONOMIES AND SOCIETIES

IOE LONDON
Leading education and social research
Institute of Education
University of London

The Anatomy of Inequalities in Educational Achievements: An International Investigation of the Effects of Stratification

Tarek Mostafa

LLAKES Research Paper 3

Centre for Learning and Life Chances in Knowledge Economies and Societies

LLAKES is an ESRC-funded Research Centre.

To request printed copies of this paper or other LLAKES Research Papers,
please contact Jeremy Tayler, LLAKES Centre Administrator: j.tayler@ioe.ac.uk.

To view this and other LLAKES Research Papers online, please visit www.llakes.org.uk.

The author is grateful to Martin Weale, Andy Green, and the participants at the PISA Research Conference in Kiel for their comments.

The Anatomy of Inequalities in Educational Achievements: An International Investigation of the Effects of Stratification

Tarek Mostafa

Abstract

This paper analyses the mechanisms of stratification and inequalities in achievements. The main objective is to determine how stratification leads to unequal educational outcomes and how inequalities are channelled through student characteristics, school characteristics and peer effects.

On the one hand, a descriptive analysis is used to shed light on the education systems of the five selected countries and to provide insight into the functioning of stratification. The countries are Japan, the UK, Italy, Germany and Finland, and the dataset used is PISA 2003. On the other hand, a multilevel econometric model is elaborated in order to quantify the effects of student, school and peer characteristics on performance scores. The results on the regressions are then interpreted according to the institutional context of each country. In the last section, policy implications, based on the regression results, are derived.

Contents

Tables

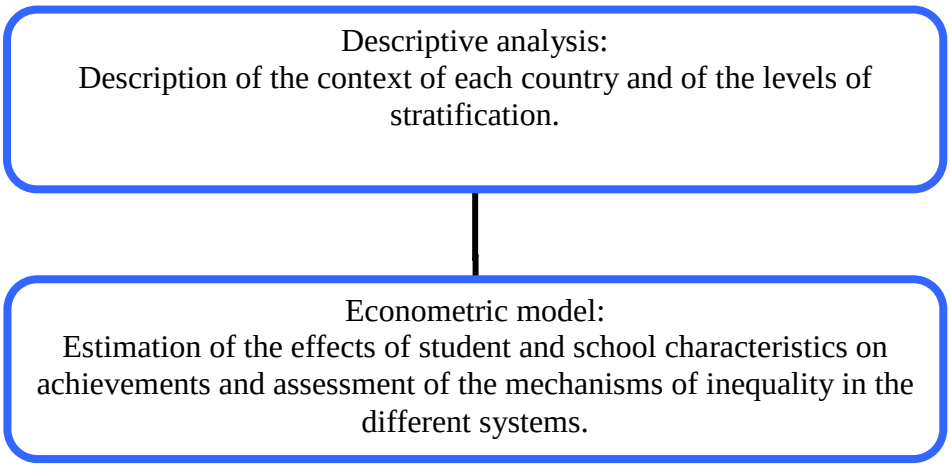
Introduction

Reducing inequalities in educational attainments has become a major preoccupation of educational reforms. Recent studies - especially the OECD's "*Education at a Glance*" - proved the existence of large disparities in outcomes and subsequently triggered a heightened interest in policy evaluation and international comparisons. In fact, education policies became a main source for political debates which made the headlines in most developed countries.

According to the traditional approach, the level of inequality is defined as the strength of the impact of social background on educational attainments. This definition is implicit in some of the empirical literature and in the international comparative reports on education, such as the "*Education at a Glance*" and the "*PISA Report*." Nonetheless, reality is more complicated. Educational achievements are not the simple direct product of social backgrounds, and in general the latter operates indirectly through intricate stratification mechanisms. For instance, less privileged households are likely to be stratified into relatively poor communities, due to the functioning of the housing market, and usually these communities are populated with other households of similar types. Hence, the social mix within local schools will reflect the social mix of the community. In such a situation, students' social backgrounds operate directly through their impact on performance and indirectly through school peer effects. Furthermore, a high level of decentralization paves the way for higher levels of disparities between schools, since the level of funding will be determined by the community's wealth. This concentration of resources and peer effects operates to the detriment of less privileged households. In other words, inequalities should no longer be considered as the mere effect of students' social and economic background on their achievements, since school characteristics are likely to be a source of inequality; and hence inequalities can be channelled through students' backgrounds and through stratification-determined school characteristics. Moreover, the strength of stratification is not the same across countries and therefore its impact on achievements may vary according to the institutional context of each education system (e.g. comprehensiveness vs. early selection). As a consequence the empirical analysis must consider several countries known for their contextual differences.

The objective of this paper is to study thoroughly the mechanisms of inequality in attainments by assessing the direct effects of household characteristics and the indirect effects resulting from student sorting between schools. These mechanisms are studied in the context of five

different countries: Germany representing German speaking countries (known for early selection), Italy representing the Mediterranean countries (known for its selection at the end of lower secondary schooling and for high geographical disparities), Finland representing the Nordic countries (known for their comprehensiveness), the UK for the English speaking ones (known for the liberal organization of education systems) and finally Japan for East Asia (known for its strong selection at the end of the lower secondary phase). The paper is organized as follows:



In the first section, the PISA 2003 dataset, the countries and the chosen variables are presented, and each education system is described in the light of the computed statistics. In the second section, an econometric model is constructed in order to provide quantitative answers to my investigations. First, I start with a presentation of multilevel regression techniques, their advantages and the problems related to their use; then I construct the model to be estimated. Three vectors of variables that explain performance scores are retained: students' socioeconomic characteristics, peer effects, and school characteristics. Finally, the regression results are interpreted and several conclusions are drawn. In the last section, policy implications are derived from the estimation results.

It should be noted that the theoretical literature on stratification is recent and dates back to the early 1970s with the founding articles of Barzel (1973) and Stiglitz (1974). The major developments occurred in the 1990s, when two distinct bodies of literature emerged. The first studied spatial stratification between jurisdictions differentiated by the price of housing and the level of local taxes. It includes Westhoff (1977), Rose-Ackerman (1979), De Bartolome (1990), Epple, Filimon, and Romer (1993), Nechyba (1997), Epple and Platt (1998) and Fernandez and Rogerson (1996). The second studied educational stratification between public

and private schools, where school quality is mainly determined by peer effects. It includes Arnott and Rowse (1987), Epple and Romano (1998, 2006) and Nechyba (2003). The empirical literature includes a variety of studies that assess the determinants of achievements, such as peer effects, students' ethnicity and immigrant status, students' socioeconomic backgrounds, and school and teacher characteristics. Hanushek and Welch (2006) provide a good coverage of the studies of interest.

Section One: Data, Countries and Variables

In this paper, the OECD Programme for International Student Assessment (OECD, 2003a) dataset is used. The major advantages of using it are the following. Firstly, the dataset is very convenient for international comparisons, since a large number of countries with different education systems are included. Secondly, a wide array of student and school characteristics are accounted for. Thirdly, the major subject of assessment in PISA 2003 is mathematics which is more universal than reading because it is not culturally specific or subject to cultural relativity. Fourthly, PISA uses an innovative concept of literacy which stresses the importance of certain skills for adult life instead of assessing the mastery of a particular curriculum. Fifthly, assessed students are aged between 15 years and 3 months and 16 years and two months, regardless of the grade in which they are enrolled. This coverage helps measuring the extent to which knowledge is acquired independently of the structure of national school systems (e.g. entry ages, grade repetition rules, etc).

The multilevel cross sectional nature of the PISA data only allows us to assess inequalities at the age of 15 even though the concept of inequality is dynamic and the levels of inequalities change over time. For instance, in Japan, 15 year old students have just finished lower secondary schooling and have been streamed into selective high schools. This strong induced stratification means that schools will become highly unequal even though a year before they were highly homogenous. The PISA dataset does not allow for an assessment of the comprehensive phase in Japan, or for an assessment of inequalities over time. However, it is very convenient to analyse the determinants of inequality at the end of compulsory education.
PISA 2003 covers 41 countries that represent one third of the world's population and nine tenths of its GDP. This large coverage is very suitable for comparative studies, since students are assessed and data is collected using the same techniques across countries. The two stage

sampling procedures used in PISA ensured that a representative sample of students was selected in each country. The first stage sampled a number of schools in which 15 year old students are enrolled, and the second stage drew 35 students in each school.[1]

Before undertaking any analysis, the dataset was imputed using multiple imputations with a *Marcov Chain Monte Carlo* procedure. The reason behind this decision is based on the fact that for some countries missing values represent about 30% of the data set. In addition to this, there is no explicit theory which supports that non-response is completely random. The imputation of missing data was recommended in the PISA data analysis manual on page 177 when non-response exceeded 10%. For a complete description of the MCMC method see Gill (2008), and Robert and Casella (2004), and for an examination of the EM algorithm used in MCMC see Schaffer (1997).

1. Countries

Five countries with different schooling systems were selected. They are: Finland, Germany, Italy, The United Kingdom, and Japan. This selection is motivated by two reasons. First, the purpose of the empirical part is to assess inequalities through an analysis of stratification mechanisms under different education systems. Hence, it is more reasonable to select few countries representative of major schooling systems than to select all countries with some being irrelevant to the objective of the analysis. Second, it is impractical to work with the entire sample of countries because of the lack of space and the thoroughness of the analyses to be conducted. The selection of the countries is based on the Green et al (2006) typology of education systems; the countries represent five different models: the Nordic model, the East Asian model, the English-speaking model, the German-speaking model, and the Mediterranean model.

Japan

The Japanese education system is traditionally characterized by high performances and low disparities between students (see: White, 1987; and Dore and Sako, 1989). Educational equality goes back to historical and social factors. For instance, during the *Tokugawa Shogunate* era, schools assumed multiple roles: they were the place for the acquisition of

[1] For details on the sampling procedures, the construction of the dataset, and the generation of test scores see the PISA Data Analysis Manuel, the technical reports and the OECD's publications (2000, 2003, and 2006).

knowledge and for the socialization of children. The *Meiji* reforms of 1868 accentuated their role as instruments of ideological unification, as well as a driving force for economic growth during the industrialization era (see Green, 1997 and1999). This industrialization led to the emergence of a large middle class; something that favoured further equality in the access to education. After the Second World War, a system led by the American occupying power emerged and was characterized by six years of elementary schooling, three years of lower secondary schooling, and three years of upper secondary education. However, the upper secondary system soon became selective and specialized, in contrast to the elementary and lower secondary systems. Other major characteristics of the Japanese compulsory education system are the centralization of control and funding, mixed-ability classes, and automatic promotion between grades, in addition to hierarchical social stratification in the upper secondary phase and the existence of a large private sector (Ichikawa, 1989). Note that the sampled Japanese students in PISA have already finished the lower secondary phase and have been stratified into higher secondary schools according to their results on the achievement tests administered by the prefectural boards of education and according to their previous records. In other words, the PISA sample of Japanese students is no longer in comprehensive compulsory education (this explains the wide between-school dispersions).

Finland

Finland possesses the same institutional structure as other Nordic countries. It retains nine years of all-through comprehensive schooling in the primary and lower secondary phases, with limited school choice and delayed first-selection. This egalitarian structure can be attributed to a number of demographic, socio-political, and historical factors.

Firstly, the common Lutheran religious traditions were favourable to the establishment of universal literacy and state regulation of education. Secondly, the Finnish education system evolved under a political democracy due to historical traditions (Boli, 1989). Thirdly, the social class structure in Finland was favourable to populist politics: the existence of weak landowner and bourgeois classes combined with strong peasant farmers and working classes paved the way towards the comprehensivization of the school system (Anderson, 1979). Finally, low population density across Nordic countries meant that school choice was limited due to the unavailability of different sources of educational supply.

Two major reforms shaped the Finnish education system. The first occurred by the beginning of the twentieth century when the lower secondary classes from the elite Latin schools were transferred to new middle schools that provided a common ladder from the primary to the upper secondary phase. The second occurred after the Second World War, under social democratic governments. It consisted of integrating middle and primary schools in order to create a nine year, all-through comprehensive schooling system.

Germany

Germany is one of the few remaining countries in Western Europe to have selective schooling in the lower secondary phase, which starts around the age of 10. Even if parents have the nominal right to choose their children's school, the latter may be relocated to other schools if they do not meet the required performance standards.

The persistence of such a selective system can be understood only in the light of a number of characteristics. Firstly, Germany has a federal political system where educational decisions are made at the Länder level, something that does not favour national reforms. Secondly, the country retains a strong apprenticeship system, as well as a tradition of cultural particularism associated with the German model of citizenship. Thirdly, the German labour market has a high level of coordination leading to the provision of training at a low cost and to mobility between firms (Marsden and Ryan, 1995). In fact, the existence of an effective apprenticeship system which is closely coordinated with the labour market provides opportunities for those graduating from lower status schools to acquire skilled qualifications and enter generally well-paid skilled jobs. These factors have each contributed to the maintenance of the selective secondary system by either reducing the pressure for reform or by making reform politically difficult.

Italy

The major characteristics of the Italian schooling system can be summarized as follows. Firstly, there is a relatively old comprehensive lower secondary school system, dating from 1962, and a differentiated upper secondary one. In fact, in 1962, a single programme for all students until the age of 14 was introduced and the *"avviamento al lavoro"* schools were abolished. Secondly, there is the use of grade repetition as an instrument for managing student heterogeneity. Grade repetition can be seen as another way of grouping students by ability,

since low performers are obliged to repeat a grade until they meet the required standards to pass. Thirdly, there is a relative absence of school choice. Italy adopted school choice in the upper secondary phase in 2000. However, in the lower secondary system schools are still required to give priority to students living in close geographical proximity. Furthermore, the new legislation of 2000 authorized the subsidy of private schools, something that was not possible according to the 1948 constitution. Finally, the Italian school system is now substantially decentralized to the level of the autonomous provinces (where some provinces have greater powers than others). In other words, there is little school autonomy but substantial regional autonomy. As with Japan, the PISA sample of Italian students is no longer in comprehensive schooling and most students are in upper secondary schools.

The United Kingdom

The United Kingdom has four distinct education systems in England, Wales, Scotland and Northern Ireland which vary in significant respects. Whilst the system in Scotland is fully comprehensive at the lower secondary stage, the other three systems retain selective grammar schools in varying degrees. Comprehensivization in England was supposed to have been achieved by phasing out grammar schools and converting them into comprehensive ones. However, this was abandoned. Parents of grammar school children were given the right to vote on whether grammar schools should be converted into comprehensive ones. In many cases the vote favoured retention of grammar schools which are still very prevalent in many counties like Kent and Buckinghamshire. In addition to this, the voluntaristic mode of educational legislation, and the introduction of competition into the education market under Thatcher's governments, impeded the move towards fully comprehensive education, particularly in England. We can say, therefore, that secondary education in the UK generally, is characterized by large territorial disparities, and an uneven spread of comprehensive schooling. England, in particular, which provides most of the schools in the PISA sample, also has an elitist private education system, and a large number of schooling alternatives differentiated by their curricula, structures, and ability grouping practices.

The analysis here treats the UK as a single unit where the PISA data do not allow us to disaggregate data for England, Scotland, Wales and Northern Ireland. However, because English students dominate the sample, where not otherwise stated the generalisations should be taken to apply principally to England.

Table 1: The number of schools and students in the PISA dataset

	Germany	Finland	The UK	Italy	Japan
Number of students	4114	5728	9045	11342	4652
Number of schools	216	197	383	406	144

2. Variables and Descriptive Statistics

The variables used in the regression analyses are grouped in three categories. They account for multiple dimensions such as: students' socio-economic backgrounds, student motivation and interest, school funding, school environment and peer effects. They are:

Student characteristics:

ESCS: Economic, social and cultural status of family.

COMPHOME: An indicator on computer facilities at home.

INTMAT: An indicator on interest in mathematics.

ANXMAT: An indicator on anxiety in mathematics.

DISCLIM: An indicator on the perception of discipline in a school.

ETR: A dummy variable taking the value of one if a student is a first generation student or a non-native. Henceforth, this category is simply called "non-natives". Note that ETR is not a measure of ethnic belonging.

Grade: a variable that controls for the grade in which a student is enrolled. Since PISA is age based and since all students were evaluated using the same test items, students' grades have to be considered in order to control for their effect on achievements.

Peer effects, school aggregates of individual characteristics:

DESCS: School average ESCS, depicting economic, social and cultural peer effects.

VARESCS: The within-school dispersion of ESCS, reflecting nonlinearities in peer effects (the impact of the dispersion of ESCS).

DCOMPH: School average COMPHOME, depicting the possession of computer facilities peer effects.

DINTMAT: School average INTMAT, depicting peer effects resulting from a generalized interest and enjoyment of mathematics within a school.

DANXMAT: School average ANXMAT, depicting peer effects resulting from a generalized feeling of anxiety and helplessness in mathematics.

DDISCL: School average DISCLIM, depicting the impact of a generalized perception of discipline in a school.

DETR: The percentage of non-natives or first generation students in a school.

Pure school characteristics:

Compweb: The proportion of computers connected to the web in a school.

Mactiv: The number of activities used to promote engagement with mathematics in a school.

Mstrel: An index measuring poor student teacher relations.

Tcshort: An index measuring principals' perception of potential factors hindering the recruitment of new teachers, and hence instruction.

Tcmorale: An index depicting principals' perception of teacher morale and commitment.

Teacbeha: An index depicting principals' perception of teacher-related factors hindering instruction or negatively affecting school climate.

Private: A dummy variable taking the value of one if a school is private (private dependent and independent schools are combined into this variable). Note that each of the selected countries, in fact, has only one of the two types of private schools. Thus, the two types have to be combined since estimation is not possible if the frequency of one of the types is close to zero. However, the interpretation of the results is made according to the predominant type.

Scmatedu: The quality of educational infrastructure in a school as perceived by the principal.

Academic: A dummy variable taking the value of one if a school selects its students according to their academic records.

The following statistics describe the education systems of the selected countries along a number of axes including: funding, private school enrolment, grade repetition, ability grouping, school autonomy, and social and performance stratification. Note that these statistics were computed using the balanced repeated replicates and the final student weights. Moreover, performance scores were computed using the five plausible values provided by PISA.[2] The variables included in the descriptive analysis are used in the interpretation of the regression results. However, not all of them are considered as regressors. Note that all these statistics are fully comparable and measured relative to OECD mean. Negative values on

[2] Plausible values provide better estimates of standard errors than the Warm likelihood estimates (see p.76 in the PISA data analysis manual). For more details on computation techniques refer to chapter 2 to chapter 7 in the same manual.

some of the statistics indicate that the country has a score lower than the OECD average which is equal to zero.

Table 2: Funding

	Germany	Finland	UK	Italy	Japan
Average quality of educational infrastructures.	0.20	-0.02	-0.07	0.13	0.01
Between school dispersion of the quality of educational infrastructures.	0.85	0.49	0.91	0.97	1.16
The proportion of computers connected to the web.	0.71	0.92	0.90	0.71	0.74
The between school dispersion of the proportion of computers connected to the web.	0.09	0.02	0.04	0.11	0.08
Teacher shortages.	0.14	-0.56	0.26	0.08	-0.04
The between school dispersion of teacher shortages.	0.78	0.42	1.00	0.86	1.34

Finland has the lowest level of between-school dispersions on all the proxies for funding (teacher shortages, quality of educational infrastructures, and the proportion of computers connected to the web). This reflects the comprehensiveness of the education system and the homogeneity of Finnish schools. In terms of the averages on these indices, Finland has a low level of quality of educational infrastructure while having high levels on computer and teacher availability. The UK has the lowest levels on the quality of educational infrastructures, and the highest average teacher shortages. The only domain where the UK is doing well is the availability of computers and internet. Furthermore, British schools don’t seem to be very heterogeneous on funding aspects despite the autonomy they enjoy. This can be attributed to equalization policies used in favour of schools in poorer areas. Germany is the country with the highest between school dispersions on most of the funding indicators, something that reflects the presence of both general education and vocational schools within the German sample. In terms of the averages, German schools have high quality educational infrastructures, the lowest levels of computers connected to the net and the highest number of students per teacher. Japan and Italy have a middle ranking position among the chosen countries. Japan has the highest between-school disparities on teacher shortages, and the quality of educational infrastructure, while Italy has relatively high levels of dispersions on most of the indicators. It should be noted that Italy has large territorial disparities between the

North and the South, and both the Japanese and Italian samples are already in streamed high schools.

Table 3: Private school enrolment

	Germany	Finland	UK	Italy	Japan
The percentage of students in private dependent schools.	7	7	1	1	1
The percentage of students in private independent schools i.e. less than 50% state funded.	0	0	5	4	26
Ratio of students enrolled in elitist private schools over total enrolment in the private sector.	0.47	0.04	0.71	0.3	0.43

In all our selected countries, the public system is by far the largest. However, the distribution of students in the private sector varies. Germany has a very low percentage of students attending independent private schools, since the private sector is dominated by government-dependent schools. In Finland, all private schools are government-dependent. The development of the private sector was intended to widen available choices; while public funding and control are maintained, independent private schools are not historically present in Finland, unlike in many other countries. Although there were initiatives in the 1980s to support private school expansion, the UK, in recent years, has sought to increase diversity of supply by diversifying school types and introducing school choice in the public system rather than by supporting the expansion of private schools (although these continue to benefit from privileged 'charitable' tax status). Private school enrolments have increased somewhat, despite this, but, in the absence of state funding for tuition, fees have remained high and private schools consequently remained highly elitist as well as largely independent of state control. Japan has a large private system that was conceived in order to complement public supply (Mons, 2004). Most private schools are independent (on our definition here) and largely tuition financed (26% of the Japanese PISA sampled students are in private independent schools which receive less than 50% of their core funding from government agencies). Finally, Italy has a very small private sector, and most schools are independent and Catholic. It should be noted that the constitution of 1948 did not allow the subsidy of private schools. However, this legislation was dropped in 2000 and some private dependent schools emerged, even though they still account for less than 1% of total enrolment.

The third statistic accounts for social elitism in the private sector. This indicator is constructed using the within-school average and dispersion of ESCS (Economic Social and Cultural Status). Elitist private schools, both dependent and independent, are considered to have a high average ESCS and a low dispersion. The distance between the maximum and the minimum for each country, on both the average and the variance, is divided into 10 units.[3] For a school to be socially elitist, it has to belong to the highest 3 units on average ESCS and to the lowest 3 units on its variance. The percentage of students enrolled in these schools is then computed.

In the UK, 71% of the students enrolled in the private sector attend socially elitist schools. The UK is then followed by Germany with 47%, by Japan with 43% and then by Italy with 30%. Note that most private schools in the UK, Italy, and Japan charge tuition fees and are government-independent, something that explains why they are elitist. On the other hand, in Finland, only 4% of students enrolled in the private sector attend socially elitist schools. This is due to the fact that most Finnish private schools are government-dependent, and do not charge fees or select students. It should also be noted that elitism in Japan concerns a larger fraction of the population than in the UK, since the private sector in the former is substantially bigger.

Table 4: Ability grouping and grade repetition

	Germany	Finland	UK	Italy	Japan
Percentage of schools with no ability grouping.	53	44	2	36	56
Percentage of schools with ability grouping for some classes.	21	42	28	43	30
Percentage of schools with ability grouping for all classes.	26	14	70	20	14
Percentage of grade repetition in the lower secondary phase.	4	0	0	6	5
Percentage of grade repetition in the upper secondary phase.	5	3	1	9	7

Ability grouping and grade repetition are alternative methods used in the management of student heterogeneity. These policies have always been the subject of debate, since they may negatively affect students by labelling them.

[3] $unit = \frac{Max - Min}{10}$

In general, grade repetition has been used in Mediterranean countries, whereas tracking and setting policies have been used in English-speaking ones. The Nordics, on the other hand, have made limited use of both. As my results show, the UK has the highest value on ability grouping for all classes followed by Germany, Italy, Finland and Japan; while Italy has the highest grade repetitions in the lower and upper secondary phases followed by Japan and Germany. Finland has limited complete ability grouping and grade repetitions in the lower secondary phase, and the UK has quasi-completely abandoned the practice of grade repetitions in both phases.

Table 5: School autonomy

	Germany	Finland	UK	Italy	Japan
Degree of school autonomy.	-0.70	0.02	1.14	-0.46	0.11
Degree of resource autonomy.	1.37	3.01	5.42	1.41	2.54
Degree of curricular autonomy.	3.00	3.89	3.93	3.70	3.93

The use of decentralization in the management of education systems can be attributed to historical and socio-political factors. For instance, in the German-speaking countries (Germany, Switzerland, and Austria), a historical decentralized management of education was put in place, due to the federal nature of the states. On the other hand, the UK has always had somewhat different education systems in its constituent countries, but decentralization to the regional (county or metropolitan area) level dates from the 1902 Education Act which made the local authorities the main administrative unit of the public school system. In the Mediterranean and Asian countries (Japan and South Korea), a centralized management was adopted, due to republican traditions in the first instance and to the central role played by governments in the conception of public education in the second (Ichikawa, 1989). However, in Italy, management was decentralized to a regional level.

PISA 2003 accounted for the degree of school autonomy through question 26 in the school questionnaire. This question is used to describe how much autonomy schools have over a number of aspects. However, it does not describe how much decentralization is there at a higher level. In other words, if educational management is centralized at a regional level, schools would still have low autonomy but the country may still be considered as decentralized.

From table 5, we can see that the UK has the highest score on the degree of school autonomy. This is due to strong decentralization policies that transferred responsibilities from central governments to schools. On the other hand, Japan and Finland have a low level on this indicator. Japan is well known for its centralized organization of education, except in teacher hiring and budget allocation, while Finland is known for decentralization at a municipal level; thus schools still have a low level of autonomy. Germany, a federal state, seems to have an unexpectedly low level of school autonomy. In fact, only 6% of the federal states in the OECD have applied decentralization policies at a school level and most of them transferred responsibilities to local or regional authorities, such as the Länder in Germany (see Mons, 2004). Finally, Italy resembles to Germany with decentralization at a regional level.

When the degree of school autonomy is broken into resource and curricular autonomy, the UK still has the highest level on both. In terms of curricular autonomy, all countries seem to have similar levels; this indicates that all of them have transferred some responsibilities in the definition of curricula and the selection of text books towards schools. In terms of resource autonomy, Germany has the lowest score, since funding is centralized at a regional level. Italy has a low level as well, whereas Finland and Japan have higher values.

Table 6: Social stratification

	Germany	Finland	UK	Italy	Japan
Average ESCS	0.14	0.24	0.13	-0.11	-0.09
Total ESCS variance	0.98	0.68	0.82	0.97	0.57
ESCS between-school variance	0.31	0.11	0.19	0.33	0.18
ESCS between school variance over total variance	0.32	0.16	0.23	0.34	0.32
ESCS Average within-school variance	0.67	0.57	0.63	0.64	0.39
ESCS dispersion of within-school Variances	0.09	0.03	0.04	0.05	0.02
Percentage of students in comprehensive schooling	3.1	100	97	1.5	0

Note that stratification levels cannot be understood unless we know where are the 15 year old PISA students within the education system of each country. In Finland, all students are in comprehensive schooling; in Germany they all have undergone the first selection (at around the age of 10); in the UK 97% of students are still in comprehensive schooling; while in Japan and Italy all of them have finished the comprehensive phase. It is important to bear in mind

that the PISA database does not allow for comparisons between the comprehensive phases across countries, since these phases start and end at different ages and grades. What PISA is designed for is to assess performance levels and the state of inequalities at the age of 15 independently of the grade students are attending. Thus, any interpretation of the regression results shall be made within this context.

ESCS is perhaps the most important variable that describes stratification in the PISA dataset. This variable was constructed using three indicators: parental education, parental occupation, and home possessions. In table 2, a number of statistics that describe the distribution and levels of ESCS across countries, between and within schools have been computed.

On the one hand, Finland has the highest average ESCS and one of the lowest total dispersions. This is due to a long tradition of social democracy and a tendency for egalitarianism. Japan has a low average as well as a low dispersion on ESCS, this can be considered to be the result of the modernization process, which eliminated old elites and led to a less class-divided society, (see Green et al 2006 for a description of East Asian countries). Germany, the UK and Italy have high dispersions, indicating wide social disparities within the population. The UK, Germany and Finland have higher values than the OECD's average ESCS (which is equal to 0). Italy and Japan scored less than the OECD average and have negative values.

On the other hand, Finland has the highest ratio of average within-school ESCS variance over total variance, indicating high levels of social diversity within schools and the lowest ratio of between-school variance over total variance. Finland is followed by the UK, then by Germany, Italy and Japan. The last three countries have similar levels on both ratios, with high "between-school variance/total variance" and relatively low "average within-school variances/total variance".

It should be noted that even if the ratios are equal for two countries, this does not mean that they have the same levels of dispersion. For instance, Finland has one of the lowest average within-school dispersions as well as a very low total dispersion. The result is the highest ratio of average within-school variance over total variance. When the levels of dispersions are not

standardized over total variance, we have the following hierarchy of countries. Japan has the lowest average within-school variance, followed by Finland, the UK, Italy and Germany.[4] Moreover, Finland has the lowest between-school variance, followed by Japan, the UK, Germany and Italy.[5]

The interpretation of the dispersion of within-school variances is slightly more involved. The within-school variance indicates how diverse is a social mix within each school; its dispersion indicates if schools are similar or not in terms of their social diversity. Germany has the highest level on this indicator which means that some German schools have a wide social intake, while others have a more homogenous one (thus the within-school variance is dispersed between schools). Italy and the UK are middle ranking, while Finland and Japan have low levels on this indicator. In other words, schools in Japan tend to be equally homogenous in social terms, whilst differentiated by their average ESCS. In conclusion, it is possible to describe social stratification in each of the five countries:

Finland is characterized by a high level of average ESCS and a low total dispersion of ESCS among students. It retains the highest average within-school variance over total variance and the lowest between-school disparities (between school variance/total variance). Schools tend to be homogenous with similar levels of average ESCS and similar social diversity. Finland can be described as a country with very low social stratification. **Japan** is characterized by a low level of average ESCS and a low total dispersion of ESCS among students. Schools tend to diverge on their average ESCS levels while having homogenous social intakes. In other words, schools have similar levels of within-school social diversity. Japan can be described as a socially less stratified country with a socially hierarchical upper secondary education system. In other words, schools have different ESCS averages and each school enrols students from a particular social class. Note that the Japanese education system is very comprehensive and egalitarian up to the age of 15 when students are stratified (the Japanese PISA sample is entirely in the upper secondary phase).

The UK is a middle ranking country within this set of countries; it has a relatively high level of average ESCS and lies in the middle of the rank order of comparator countries in terms of

[4] Note that since Japan and Finland have low total dispersions on ESCS, it is normal for them to have low average within-school dispersions because the width of the total distribution of ESCS is smaller.
[5] The non-standardized variances do not allow for international comparisons because they do not control for the width of the distribution of ESCS for each country.

total dispersion (i.e. lower than Italy and Germany but higher than Japan and Finland, which have the lowest total dispersions). It retains a high level of within school dispersions and low between school disparities. However, even if schools in the UK differ only moderately in terms of average ESCS, compared with the four comparator countries here, they differ on the level of within-school ESCS dispersions. Hence, some schools have homogenous social intakes (low dispersion), while others are more diverse (recall that the UK has the highest percentage of elitist private school enrolment). The UK can be described as having moderate stratification relative to our other countries here. This result should, however, be seen in context. Even though the UK ranks directly after Finland in terms of its level of stratification, this happens because the three other countries are selective at the point where the students are sampled in PISA. More precisely, Germany has early selection and the sampled Japanese and Italian students have already finished comprehensive schooling. Therefore, when compared with comprehensive systems, the UK will rank rather high in terms of between-school disparities.

Italy has a low level of average ESCS and a high total dispersion among students. It has low within-school social diversity and high between-school disparities. Schools are heterogeneous on both average ESCS and its dispersion. **Germany** has a relatively high level of average ESCS and the highest level of total dispersion among students (note that ESCS is an economic social and cultural status measure and not an income dispersion measure). It has low within-school diversity and high between-school disparities. Schools are heterogeneous on both average ESCS and its dispersion. Italy and Germany have the highest levels of social stratification, since the social status of schools is highly dispersed. Moreover, some schools are elitist, while others are diverse. This phenomenon can be explained through differences between vocational and general education tracks in Germany and through geographical disparities in Italy in addition to stratification in the upper secondary phase.

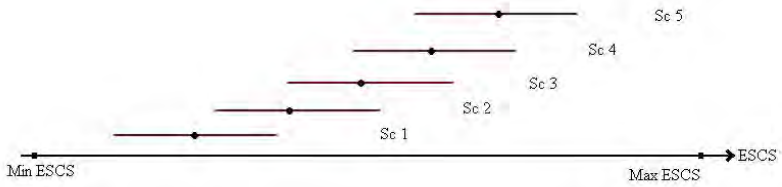
The patterns of stratification are illustrated through the following figures:

- The dots represent average ESCS in five schools. If the projections of the dots on the ESCS axis are close, this indicates low between-school dispersions. The reverse is true when the projections are far apart.
- The length of the lines represents the width of the distribution of ESCS within schools. A short line corresponds to a socially homogenous school, while a long one corresponds to high social diversity. Finland and Japan have lines of the same length,

even if averages are different for Japan. The UK, Germany and Italy have lines of different lengths. More precisely, Finnish and Japanese schools have similar within-school dispersions, while British, German, and Italian schools differ in the degree of social heterogeneity. In some countries the projection of the lines covers the whole range of ESCS, indicating wide social intakes, while in others it does not.

Finland

Japan



UK

Germany and Italy

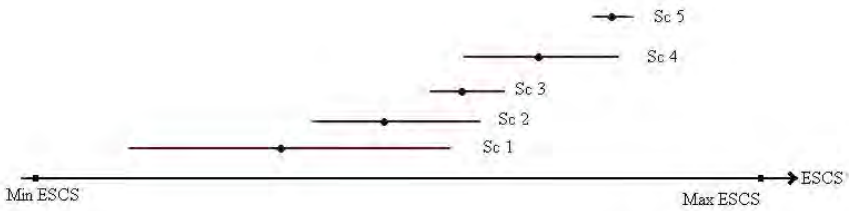


Table 7: Stratification according to performance scores and to immigrant status

	Germany	Finland	UK	Italy	Japan
Average Performance scores	503	544	509	466	534
Total performance scores' variance	10435	6927	8380	9695	10212
Performance scores' between-school variance	6220	620	2836	5398	5608
Performance scores' between-school variance over total variance	0.60	0.09	0.34	0.56	0.55
Performance scores' average within-school Variance	4216	6306	5544	4298	4604
Dispersion of performance scores' Within-school variances	2150665	2892138	3471417	1895955	2574521
Percentage of non-natives and first generation immigrants	14	1.86	7.6	2	0.1
Between school variation of the percentage of non-natives and first generation immigrants	268	9.7	109.2	70.5	0.3

The same statistics were also computed for performance scores. **Finland** has the highest average performance and the lowest variations in performance scores. Finnish schools have quasi-identical average performances while having wide performance intakes. However, some schools tend to be less heterogeneous then others. In general, Finland can be described as a country with low performance stratification, something that can be attributed to the comprehensiveness of the system. **The UK** follows Finland, but with lower achievements and higher variations in individual performance. Schools are only moderately differentiated by average performances, relative to the comparator countries, and have relatively wide ability intakes, though they are very dispersed in terms of their within-school variances. In other words, some have wide performance intakes, while others are more homogenous (they may be using some form of academic selection).

Japan has a high level of average performance as well as a high variation in individual performance scores. Japan seems to be strongly stratified at the upper secondary stage, with strong between-school dispersions and limited within-school diversity. Traditionally, Japan is considered to be a very equal country, a claim that may still be made at its lower secondary

system. However, in the upper secondary phase, schools are highly heterogeneous in terms of their peers due to the placement test undertaken at the end of lower secondary schooling. **Germany** and **Italy** have the same characteristics as Japan. However, Germany's average performance is closer to the OECD mean, while that of Italy is far below.[6] These two countries have high levels of stratification, with Germany being the one with the highest between-school dispersion and lowest within-school diversity. Germany is closely followed by Italy on these two aspects. One should bear in mind that Italy, Japan and Germany are selective at the age of 15, something that influences the ranking of countries.

When it comes to immigrant status, the percentage of non-natives is the highest in Germany (14%) and in the UK (7.6%), followed by Italy and Finland with about 2%. The percentage of non-natives in Japan is close to zero, and only 5 schools among 144 have non-native students (thus, the result for Japan are not interpreted). The between-school variation of the proportion of non-native students indicates if they are concentrated in particular schools. In other words, this indicator has a high value when some schools retain the majority of non-natives, while others only have native students. Germany has the highest concentration, followed by the UK and Italy. In these countries, non-native students are not uniformly distributed across schools. Conversely, Finnish schools have relatively the same percentage of non-natives.

In what follows, the education production function is estimated for the five selected countries and the results are interpreted.

Section Two: The Empirical Model

1. Multilevel Regressions and Econometric Analyses.

The PISA dataset does not allow for the estimation of time or value-added education production functions (EPF), since data are collected only over a single period of time. Furthermore, value-added EPFs are not suitable for the analyses of inequalities in achievements. In fact, value-added equations relate the variations in inputs over a period of time to variations in output over the same period. They can be written as follows:

$$y_{it} - y_{it*} = f\left(F_i^{t-t*}, P_i^{t-t*}, S_i^{t-t*}\right) + \left(\varepsilon_{it} - \varepsilon_{it*}\right)$$

[6] The OECD average performance score is equal to 500.

where y is performance of student i at two different time periods t and t^*. F, P and S are vectors of student, peer, and school characteristics taken at the same periods. However, what is needed in this model is to relate variations in inputs across students and schools to variations in outputs. Multilevel regressions allow for this type of estimations. The education production function can be written as follows:

$$y_{ij} = f\left(F_{ij}, P_j, S_j, A_i\right) + \varepsilon_{ij}$$

Where y_{ij} is performance of student i in school j, F_{ij} is student characteristics of student i and P_j and S_j are peer effects and school characteristics of school j. Note that this formulation eliminates the need for historical information and has the advantage of being adapted to international databases, such as PISA. The two formulations are very different in their conceptual frameworks. The first analyzes the change in outcomes over time, while the second analyzes the variations in students' performances across schools and students during one period. In other words, the first assesses the growth of student performance while the second analyzes the inequalities between students and schools.

Multilevel regression techniques were retained for the estimation of the model. They acknowledge that students nested within larger units (schools) bear some resemblance. Moreover, multilevel regressions allow for the estimation of school specific effects both on the intercept and on the regression coefficients on student variables. In other words, this estimation technique acknowledges the fact that student level variables may not necessarily function the same way across schools. Another benefit of multilevel modelling is the decomposition of the variance components between levels.[7]

The model to be estimated is the following:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \gamma_1 \overline{X}_{\bullet j} + \gamma_2 K_j + \varepsilon_{ij}$$

With

$$\beta_{0j} = c + V_j$$

$$\beta_{1j} = \beta + \mu_j$$

X_{ij} is a vector of student characteristics (student i attending school j), $\overline{X}_{\bullet j}$ is a vector of peer effects (school aggregates of student characteristics), and K_j is a vector of pure school

[7] For a full review of multilevel modeling techniques, see Raudenbusch and Bryk (2002).

characteristics (e.g. funding, school environment, etc). ε_{ij} are the residuals of the model, they follow a normal distribution, with zero mean and a constant variance of σ^2, $\varepsilon_{ij} \sim \mathrm{N}(0,\sigma^2)$. Henceforth, Student level is called level one, and school level is labelled level two. When the intercept and the regression coefficient on X_{ij} are replaced by their values, the equation becomes:

$$Y_{ij} = c + \beta X_{ij} + \gamma_1 \overline{X}_{\bullet j} + \gamma_2 K_j + V_j + \mu_j X_{ij} + \varepsilon_{ij}$$

Note that, the intercept is divided into two elements: c is the overall intercept, which is constant for all schools and equal to the average of the intercepts β_{0j}, and a random part V_j, denoting school j departure from the overall intercept, which can also be seen as a unique effect of school j on the average intercept (Raudenbush and Bryk, 2002). V_j can be considered as comprising the unobserved school characteristics. V_j is assumed to have a zero mean and a variance of τ_0^2. $V_j \sim \mathrm{N}(0,\tau_0^2)$.

The slope on student variables is divided into two elements: β is the overall regression coefficient, equal to the average of regression coefficients β_{1j}, and a random part μ_j, denoting school j departure from the overall regression coefficient, which can also be seen as a unique effect of school j on the slope of X (Raudenbush and Bryk, 2002). μ_j is assumed to have zero mean and a variance of τ_1^2. $\mu_j \sim \mathrm{N}(0,\tau_1^2)$.

Notice that V_j and μ_j are treated as random errors following normal distributions. The variances on V_j and μ_j are also called between school variances. These should not be confused with the variances computed in the descriptive statistics. V_j and μ_j represent the dispersion of school specific effects on the intercept and on the regression coefficients while the former are directly computed from the data without any estimation.

As known from the theoretical literature (mainly Epple and Romano, 1998), stratification transforms a continuum of individual characteristics into a continuum of tuition levels. Tuition enters the utility function, and utility maximization determines the school to be chosen, and the level of educational quality. In conclusion, student characteristics determine

school characteristics, and both are correlated. From an econometric perspective, this correlation may lead to endogeneity problems if key variables are omitted. For instance, omitted school variables will be absorbed by the error term, and the latter will be correlated with the included student characteristics, giving rise to endogeneity problems that bias the estimation results. The model to be estimated must satisfy the following properties:

a) The independent variables at each level are not correlated with the random effects (error terms) on the other level - $\text{cov}(X_{ij}, \mu_j) = 0$, $\text{cov}(X_{ij}, V_j) = 0$, $\text{cov}(\overline{X}_{\bullet j}, \varepsilon_{ij}) = 0$ and $\text{cov}(K_j, \varepsilon_{ij}) = 0$. In other words, any unobservable student characteristics relegated to the error term should not be correlated with the observable school characteristics $\overline{X}_{\bullet j}$, and K_j. Similarly, any unobservable school characteristics relegated to the error terms should not be correlated with the observable student characteristics X_{ij}.
b) The level one independent variables are not correlated with level one error terms. $\text{cov}(X_{ij}, \varepsilon_{ij}) = 0$. Any unobservable student characteristics relegated to the error term should not be correlated with the observable student characteristics X_{ij}.
c) The level two independent variables are not correlated with level two error terms - $\text{cov}(\overline{X}_{\bullet j}, V_j) = 0$, $\text{cov}(\overline{X}_{\bullet j}, \mu_j) = 0$, $\text{cov}(K_j, V_j) = 0$ and $\text{cov}(K_j, \mu_j) = 0$. Any unobservable school characteristics relegated to the error terms should not be correlated with the observable school characteristics $\overline{X}_{\bullet j}$, and K_j.
d) Each level one error term ε_{ij} is independent and normally distributed with a mean of 0 and a constant variance of σ^2. $\varepsilon_{ij} \sim \text{N}(0, \sigma^2)$.
e) Each level two random effect (error term) is normally distributed with a mean of 0 and a variance τ^2. $V_j \sim \text{N}(0, \tau_0^2)$ and $\mu_j \sim \text{N}(0, \tau_1^2)$. These error terms are independent among the level two schools.
f) The error terms at level 1 and 2 are independent. $\text{cov}(\varepsilon_{ij}, V_j) = 0$ and $\text{cov}(\varepsilon_{ij}, \mu_j) = 0$.

Note that in Mostafa (2009), endogeneity problems were assessed by applying the Hausman test on several variants of the aforementioned model. Furthermore, homoscedasticity and the independence of the error terms were also assessed using residual scatterplots and Q-Q plots.

The major finding is that, when peer effects are omitted, the model does not pass the Hausman test in the five selected countries. This confirmed that peer effects are a major product of stratification and that their omission leads to correlations between the error term and the included student characteristics and, hence, to biased results.

The most reliable model is the one that controls for the following three vectors: student characteristics, peer effects, and school variables. In what follows, only the results from the aforementioned general model are interpreted, since it passed the Hausman, the homoscedasticity, and the independence tests.[8] Note that the inclusion of all these factors as independent variables minimizes the risk of endogeneity and leads to robust and consistent results. Moreover, the sensitivity of the model with imputed data was tested against several regressions estimated without imputations, and with different imputation methods. In all cases, the various regressions generated results of similar magnitude and statistical significance, confirming that they are not driven by imputation techniques.

Estimation Procedures

Two estimation procedures are possible: maximum likelihood methods and Bayesian techniques. Since all the selected countries have large student and school samples (more than 4000 students and about 200 schools) and since most schools have about 35 students, Bayesian techniques are not needed.

In order to estimate the model with a maximum likelihood technique, one has to construct a likelihood function. Let us consider Y to be a vector representing the dependent variable. U is a vector of all random effects in the model at level 2 and β is a vector containing all parameters (including the fixed coefficients and variance and covariance components). The probability distribution of the dependent variable at level 1, given the random effects and parameters, can be written as $f(Y|U,\beta)$. The distribution of the random effects, given the parameters, can be written as $p(U|\beta)$. The likelihood function, given the parameters, is the following:

$$L(Y|\beta) = \int f(Y|U,\beta)p(U|\beta)dU$$

[8] See Mostafa (2009) chapter 4 for more details on endogeneity problems, for the estimation results of the different variants of the model, and for the results on the Hausman test.

The second step consists of maximizing the likelihood function. In my model, this is done using a restricted maximum likelihood (RML) and a Newton-Raphson algorithm. The RML technique is described in Patterson and Thompson (1971) and is known to generate unbiased estimates of the variance and covariance components.

2. The Results

Before interpreting the results, it is useful to start with a statistical definition of inequalities. Inequalities do not exist when - in a regression analysis - all student and school variables have insignificant effects on performance scores. In other words, the variables that may explain differences in performance scores are still hiding in the unobserved component (the error term). If most student and school characteristics are controlled for, and are found to have no effect, the only possible variable that may explain differences in achievements would be student ability (e.g., IQ). This situation is a perfect meritocracy, where the surrounding environment of a student - whether at home or at school - does not affect his or her achievements. Of course, this situation does not exist, but it is useful to consider it as a benchmark against which countries are compared.

Table 8: The variance components

	Germany	**Finland**	**The UK**	**Italy**	**Japan**
Total variance	3659.08	3875.95	4187.87	3950.43	4431.33
Within variance σ^2	2436.41	3647.92	3771.91	2771.36	3661.98
Between variance τ_0^2	1222.67	228.02	415.96	1179.06	769.35
% Between/Total	33.41	5.88	9.93	29.85	17.36

Table 9: The between school variances on student variables' random effects

	Between school variance τ_1^2				
Variable	**Germany**	**Finland**	**The UK**	**Italy**	**Japan**
ESCS	11.24	56.08	116.03	57.36	48.07
COMPHOME	21.38	18.01	108.11	43.44	27.04
INTMAT	29.05	59.18	166.34	65.07	3.53
ANXMAT	21.90	35.61	126.59	84.14	29.52
DISCLIM	17.06	41.06	86.28	55.97	47.18
ETR	116.21	1070.54	1126.49	1487.15	

Table 10: The regression results

	Germany			Finland			The UK			Italy			Japan		
Variables	Coef.		SE	Coef.		SE	Coef.		SE	Coef.		SE	Coef.		SE
Intercept	120.2		15.4	132.8		25.1	302.0		18.9	181.5		16.3	600.8		31.8
Grade	36.4	***	1.3	43.3	***	2.5	16.0	***	1.5	34.6	***	1.3			
ESCS	8.2	***	1.1	25.0	***	1.3	21.8	***	1.2	1.6	**	0.9	1.8		1.7
COMPHOME	0.5		1.1	-2.6	**	1.2	7.8	***	1.2	8.4	***	0.8	3.1	**	1.3
INTMAT	5.3	***	0.9	15.0	***	1.2	-1.1		1.2	6.5	***	0.8	9.8	***	1.2
ANXMAT	-16.8	***	0.9	-31.4	***	1.1	-25.1	***	1.2	-21.7	***	0.9	-8.5	***	1.2
DISCLIM	0.9		0.8	0.9		1.1	12.4	***	0.9	1.7	**	0.8	2.6	**	1.3
ETR	-21.8	***	2.9	-45.0	***	8.0	-2.2		4.3	5.1		5.5			
DESCS	47.9	***	9.3	6.3		6.1	46.9	***	6.5	31.5	***	6.1	91.9	***	14.4
VARESCS	1.4		9.5	-0.4		9.4	22.1	***	7.1	-33.6	***	8.2	43.2	**	20.2
DCOMPH	41.6	***	14.6	-3.7		9.1	-2.4		9.2	50.0	***	8.5	17.0		14.9
DINTMAT	-23.8	**	10.5	-11.1		8.5	-22.4	***	7.2	-17.1	***	6.3	45.9	***	12.6
DANXMAT	-22.5	**	11.6	4.5		9.2	-14.4	*	7.6	-51.2	***	9.2	38.6	***	13.2
DDISCL	27.1	***	7.6	-1.8		5.1	13.0	***	4.7	14.7	***	5.2	29.9	***	7.0
DETR	25.0		21.3	51.4		51.4	-42.7	***	15.1	72.1	***	24.2			
COMPWEB	9.8		8.1	13.4		10.4	18.2	***	6.9	6.7		5.7	-1.9		9.2
MACTIV	6.8	*	4.0	1.9		2.7	-1.9		1.6	7.8	***	2.3	1.0		4.6
MSTREL	44.2		54.3	-111.7	**	56.4	-103.2	**	43.4	-83.0	*	46.7	-108	**	51.4
TCSHORT	-5.2	*	3.7	0.01		2.4	-1.3		1.8	3.2	*	2.5	-0.4		3.1
TCMORALE	6.5	*	3.4	2.6	*	2.1	-0.8		1.8	-1.5		2.4	-2.1		3.3
TEACBEHA	-11.3	**	4.8	-2.5		2.3	0.1		2.4	1.1		2.3	7.8		4.0
private	-2.5		11.3	-20.3	***	6.8	13.2	**	5.2	-38.3	***	8.9	-38.3	***	6.8
SCMATEDU	6.2	*	3.4	-0.5		2.3	1.2		1.8	5.5	***	2.2	-5.1		3.2
Academic	5.0		5.9	11.0	**	5.1	15.8	***	4.3	-5.4		4.1	-47.9		25.3

(***), (**), and (*) stand for significance at the level of 1%, 5%, and 10% respectively.

It is also worth noting that a thorough assessment of inequalities cannot be done unless the regression coefficients and the distribution of the variables are simultaneously analyzed. For instance, if in a particular country the social background of students is found to have a high and significant impact on performance scores; this does not necessarily imply that the country has a high level of inequalities. In fact, if the total dispersion of social backgrounds in the population is very limited then the level of inequalities in achievements will be low. In other words, the level of inequality is determined by the regression coefficient and the dispersion of a variable. The higher the coefficient and the more dispersion is there, the higher inequalities will be. In what follows, the results on key variables are interpreted.

Social Status and Social Peer Effects

ESCS is the most important student level variable affecting performance. In most countries, it is at the centre of educational policies, since one of the objectives is to ensure equality of opportunity in the access to education. Furthermore, ESCS is the most important dimension according to which stratification operates.

On the one hand, ESCS has a positive and statistically significant effect on performance scores across in all countries except in Japan. Finland has the highest value on the regression coefficient, with an increase of 25 points in performance scores caused by an increase of one unit of ESCS. Finland is followed by the UK, Germany, and Italy. On the other hand, the coefficients on school average ESCS (DESCS: representing social peer effects) are all positive and statistically significant at the level of 1% for all countries except for Finland which has an insignificant result. The highest value on this coefficient is for Japan, followed by Germany, the UK, and Italy.

The results for Finland seem to be counterintuitive for a comprehensive schooling system. However, when average ESCS is taken into account, a full picture emerges. Comprehensiveness in Finland is associated with high levels of homogeneity between schools. Therefore, the impact of school variables, including peer effects on performance scores, is expected to be small. As a consequence, the only factors that would explain the variation in performance scores are student characteristics, such as ESCS. Furthermore, this high value on the regression coefficient is not alarming since Finland has the lowest national dispersion of ESCS. In other words, even if the slope is steep, there are limited variations of ESCS to cause high inequalities in performance. Note that in Finland, students have the choice between different schools; however, this choice is irrelevant, since all schools are almost identical. Finland can be described as the most equal country, since its education system is conceived in a way that does not generate inequalities.

The UK also has an important effect of ESCS on student performance scores. However, the case of the UK is different from that of Finland, since school variables have strong effects, too. This is perhaps the result of the unachieved comprehensiveness of education in the UK

and most particularly in England which dominates the sample here.[9] In fact, both student level ESCS and school level DESCS have significant and important effects on achievements. The UK - or more precisely England - resembles Germany on this aspect, even though the effects are more moderate.

In Germany, the high levels of stratification and heterogeneity between schools are translated into inequalities in performance scores. The ESCS of a student determines the school in which he or she is enrolled as well as a certain proportion of his performance. Hence, a student with a low level of ESCS is likely to be streamed into schools (probably vocational) where other students with similar levels of ESCS are enrolled. Since ESCS has an important effect on performance, low ESCS students will get lower results. And since average ESCS in a school also has an important effect on performance, students attending schools with low average ESCS are likely to have lower performances. In Germany, the selective school system is a generator of inequalities, since it allows ESCS to play fully through its direct household effect and through its indirect school effect. However, the German system cannot be understood unless the labour market is considered. Germany retains a strong apprenticeship system through which low ESCS students are shepherded into vocational tracks and educational inequalities are absorbed by the labour force. In addition to this, attending vocational schools is not regarded as a sign of failure and is not associated with a socially negative status. In fact, the vocational tracks probably mitigate inequalities in performance since apprentices get 3 to 4 years of additional mathematics and language learning allowing them to 'catch up'. Apprenticeships also mitigate inequality of opportunity in the system since they allow lower achieving students to gain skilled worker qualification which give them access to skilled jobs which are generally quite well paid.

Japan also has high levels of social stratification. However, inequalities operate differently than in Germany. On the one hand, a student's ESCS has an economically and statistically insignificant direct effect on his performance; while on the other hand, school average ESCS has a very important and significant effect on performance scores. These results reflect the expected impact of stratification. The large between-school disparities combined with the within-school homogeneity mean that student level variables will have limited effects while

[9] It should be noted here that the generalisations apply mostly to England since the English students dominate the sample. In fact, variation in student achievement (total dispersal of scores) is much higher in England and Northern Ireland than in Scotland and Wales when you break the figures down by country (see Green, 2008).

school level variables will have stronger impacts. In fact, students are stratified into schools according to a placement test undertaken at the end of the lower secondary phase. This assessment determines which student goes in which school; and as a result top performers are accepted in top performing high schools. The presence of such a strong effect of DESCS reflects the importance of stratification as a determinant of inequalities in achievements. In other words, the access to a school with a high social and economic status can be the source of success. However, this result should not be used to establish a complete causality going from DESCS towards performance scores. Since the sampled Japanese students have been together for only 3 months after being tracked into different high schools, the notion of peer effects in the case of Japan should be explained carefully. In fact, school average ESCS is the result of stratification according to achievements on the placement test undertaken at the end of the lower secondary phase. Thus, it is reasonable to acknowledge that the causality between DESCS and performance scores in the case of Japan might works in both directions. In other words, DESCS can be seen as peer effects affecting performance scores on the standardized PISA test as well as the result of performance scores on the placement test undertaken 3 months earlier. Furthermore, since between-school heterogeneity has expanded in the upper secondary phase, we can expect that the effect of school characteristics on achievements will remain strong with the passing of time.

Moreover, the Japanese education system has very particular features. The lower compulsory education phase is very comprehensive and very egalitarian while the upper phase is highly selective. This may help explain my findings. In the lower secondary phase, schools are very homogenous. Therefore the effects of school characteristics on achievements are expected to be weak while the effects of household characteristics (e.g. ESCS) are expected to be stronger. Hence, household characteristics determine achievements on the placement tests at the end of the lower secondary phase. Once students have been tracked into high schools according to their performance scores, the mechanism is reversed. The heterogeneity between high schools and the homogeneity of their intake imply that school variables will have a much stronger effect than household characteristics, which is the case in this analysis.

Italy has results very similar to Japan, even if they are quantitatively lower. It can also be described as a country with a high level of stratification, where the social status of a school determines the performance of students. A student's own ESCS has a very low effect on his performance, while the status of his school is much more important. Note that the impact of

social peer effects is significantly lower than that of Japan and even lower than that of the UK and Germany.[10]

Another interesting finding is that peer effects are non-linear in their means in three of the five countries (Varescs is significant in the UK, Italy and Japan). This finding emphasizes the fact that any future research dealing with stratification should consider the distribution of students within schools and not only average student characteristics. The regression coefficient on Varescs determines how performance scores react to changes in social diversity. Note that linearity in means was criticized in Hoxby and Weingarth (2005). Such findings may be of great interest to policy makers since they can provide better orientation for policies related to school choice (e.g. vouchers, diversification of educational supply, public school choice, etc). In the UK and Japan, an increase in the within-school dispersion of ESCS enhances performance while the reverse is true in Italy.

Other Peer Effects and Student Aptitudes

From table 10 it is possible to see that the significance of the regression coefficients on peer effects varies between countries. For instance, in Finland neither of the variables has a significant effect, while the reverse is true for Italy. The UK, Germany and Japan are middle ranking. This finding indicates clearly that school homogeneity in Finland causes peer effects and other school variables to be less significant. In other words, school environment would not be a source of inequality. Note that in this analysis only two states have dominant comprehensive systems at the age of 15 (the UK, in parts, and Finland) while the rest of them have already streamed their students.

Student's attitudes and motivation are very important determinants of performance. The coefficient on INTMAT, ANXMAT, DISCLIM, and the corresponding peer effects are highly significant. All the coefficients on the student-level variables are of the expected sign. Interest in mathematics and the perception of discipline are positively related to performance scores, while anxiety in mathematics has a negative effect. On the other hand, the sign on the coefficients on peer effects vary between countries. For instance, interest peer effects have an unexpected negative impact on performance in Germany, the UK, and Italy; and anxiety peer effects have a positive impact on performance in Japan. This is perhaps due to the fact that a

[10] In the case of Italy we can speak about social peer effects in the upper secondary phase because students have been together for more than one year.

generalized feeling of anxiety and helplessness in mathematics (DANXMAT) may comfort students and reduce the direct effect of being anguished (ANXMAT). Hence, in some cases it may have a positive effect. Similarly, if interest in mathematics is a generalized feeling in a school (DINTMAT), that may dilute the direct effect of being interested in maths (INTMAT). In other words, if all students are interested in mathematics, this may reduce the special value of being interested in mathematics.

Funding and School Characteristics

Several proxies of school funding were retained: the proportion of computers connected to the web, the number of activities promoting mathematics, teacher shortages, and the quality of educational infrastructures. The level of significance and the value on the coefficients vary between countries. For instance, in Germany, the number of activities promoting mathematics, teacher shortages, and the quality of educational infrastructures have significant effects; while in Finland and Japan none of the coefficients is significant. In the UK, only COMPWEB has a significant effect, while in Italy, the number of activities promoting mathematics, teacher shortages, and the quality of educational infrastructures have significant effects. All the coefficients are of the expected sign. These results reflect the high between-school disparities in Germany and Italy and the homogeneity of schools in Finland. When it comes to school environment, only MSTREL (poor teacher student relations) has a significant and negative effect across all countries. Teacher morale has a positive and significant effect in Germany and Finland, while negative teacher behaviour has a negative and significant effect in Germany.

Private Schooling

Private education is also an important determinant of performance. It has a significant effect across all countries except in Germany. At the first sight, the results seem to be counter intuitive since the sign on the coefficient is negative except for the UK. However, these results can be explained. Recall that in the UK, 71% of private school enrolment is socially elitist and most of the schools are expensive and government independent. Thus, these schools have financial constraints and are expected to maximize a profit function. In other words, they are expected to have higher qualities and higher achievements than public schools in order to attract any students (positive sign on the regression coefficient). In contrast, private schools in Finland are government dependent, and are funded and controlled by the state. They are not

socially elitist (they are not selective and do not perceive tuition fees) and usually are attended by students who cannot follow the regular curricula in public schools. Hence, they are not expected to perform better than their public counterparts. Similarly, in Japan, the private sector was conceived in order to complement public school supply and is not highly elitist, while in Italy most private schools are catholic non-elitist and non-subsidized. The negative effect that private schooling has on performance scores is a clear indication that the apparent superiority of private schools is channelled through better peer quality or funding and not through structural differences between the two sectors. In other words, when peer effects and funding are controlled for in a regression analysis, the effect of private schooling becomes nonsignificant or even negative.

The Variance Components

Before interpreting the variance components, it is of interest to recall what each component stands for. The between-school variance is the dispersion of school specific effects around zero, or, in other words, the dispersion of the random parts of the intercept around zero. Note that these random parts V_j follow a normal distribution, with zero mean and a constant variance, $V_j \sim \mathrm{N}(0, \tau_0^2)$. The within-school variance is the dispersion of the model residuals around zero; note that these residuals ε_{ij} follow a normal distribution, with zero mean and a constant variance, $\varepsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$. The total variance is the sum of the two variances. The between-school variances on student variables are the dispersions of school specific effects, on these variables, around zero. These specific effects μ_j are assumed to follow a normal distribution with zero mean and a variance of τ_1^2. $\mu_j \sim \mathrm{N}(0, \tau_1^2)$.

Germany has the highest ratio of between-school variance over total variance, followed by Italy, Japan, the UK and Finland. The high level of between/total ratio in Germany, Italy, and Japan indicates that schools tend to have specific effects that diverge from the average effect (overall intercept). Note that these countries have already streamed their students into differentiated schools. The disparities in Italy may also be the reflection of important territorial differences between the North and the South in addition to stratification in the upper secondary phase. In Germany they reflect differences between general and vocational tracks. In Japan they reflect the hierarchical and stratified nature of the education system in the upper secondary phase which is also differentiated between general and vocational schools.

When it comes to the between-school dispersions on the effects of student level variables, the results are the following. Starting with ESCS, it is possible to see that the UK has the highest value, followed by Italy, Finland, Japan and Germany. High dispersions indicate that schools differ on the strength of the impact of ESCS on performance scores. In other words, some British schools have very low impact of ESCS, while others have a stronger one. It should be noted that these dispersions cannot be considered as measures of inequality; they only indicate that schools differ on their specific effects without mentioning how strong these effects are (one has to analyze the overall regression coefficient to establish the magnitude of this strength). More precisely, if most German schools have a very high impact of ESCS on performance scores; the dispersion of their specific effects is likely to be small (which is the case here). The UK also has the highest dispersion of school specific effects on COMPHOME, INTAMAT, ANXMAT and DISCLIM. On COMPHOME, the UK is followed by Italy, Japan, Germany and Finland. On INTMAT, the UK is followed by Italy, Finland, Germany and Japan. On ANXMAT, the UK is followed by Italy, Finland, Japan and Germany. On DISCLIM, the UK is followed by Italy, Japan, Finland and Germany. And finally, on ETR, Italy is followed by the UK, Finland and Germany.

Section Three: Policy Implications

So far I have analyzed the determinants of inequalities in achievements in the context of five countries differentiated by their contextual characteristics. In what follows, policy implications are derived from the regression results and a special emphasis is given to the comparison between the UK and Finland since the sampled students in both countries are in comprehensive schooling.

A number of policies were used in different countries to improve equity in the distribution of achievements. These include additional educational resources for particular schools based on their performances and social intakes. Such policies were used in France, where additional funds were provided to schools enrolling students at risk of school failure (zone d'éducation prioritaire), and in the England, where funds were provided for equalization purposes in favour of poor neighbourhoods' schools. Other policies consist of spending more on students

presenting specific characteristics, such as belonging to a disadvantaged social class. These policies include vouchers and conditional cash transfers.

On the one hand, policies designed to enhance the situation of individuals should be used in countries were ESCS and other student level variables have a large and significant impact on achievements and on the formation of inequalities (e.g. Germany and the UK). On the other hand, policies designed to enhance the situation of schools should be used in countries were the heterogeneity of schools is the main source of inequalities (e.g. Italy, Germany, and England in the UK).

Other types of policies that could enhance performance scores for unprivileged social groups are related to the geographical organization of educational supply. Different school choice policies have been used across the OECD countries, ranging from free choice to a strict application of catchment areas. However, a middle solution consists of the use of zoning policies through which district boundaries are fixed in a manner that maximizes achievements and enhances their distribution. This type of policy is supported by my findings. The nonlinearity of peer effects in their means (in the UK, Italy and Japan) suggests that student allocation is not a zero sum game, and that achievements can be enhanced through a better distribution of peers. Hence, the induced reallocation of students can be the tide that lifts all boats.

The choice of a particular policy depends on the context of each country. On the one hand, countries with strong direct effects of ESCS may be inclined to adopt compensation policies to enhance the situation of unprivileged students. On the other hand, countries with high coefficients on school social peer effects may favour policies that enhance the distribution of peers across schools or that enhance the situation of unprivileged schools through compensation. Countries like the UK, and Germany should do both, since inequalities in achievements originate from students' social backgrounds and schools' peer effects. In Italy, compensation policies should be used in order to achieve more territorial homogeneity. Moreover, the homogenization of schools may require the abolishment of practices such as academic and early selection and the reduction of regional and territorial differences in funding. However, such policies are concerned with the structure of the education system and can be implemented only through a long-term approach.

A final concern would be private schooling and public subsidies to private schools. As my results have shown, after controlling for student and school characteristics, private schooling does not have a positive impact on performance scores except in the UK. Hence, the advantages that private schools may offer are channelled through higher peer quality or higher funding and not through structural differences between public and private schools. Public subsidies to private schools have so far been used to maintain a choice outside the public system. However, such subsidies are subject for debate when private schools become the schooling institutions for the social elite. In this case, the question that can be asked is: why should public subsidies be maintained when the access to private schools is selective and does not favour equality of opportunities? Perhaps the most coherent answer is the one applied in Greece, where private schools exist but are not subsidized. In other words, a school system should offer equal opportunities to all students; yet it should maintain freedom of choice for those who have special tastes in education (religious, etc), without subsidizing these particular tastes. Moreover, private schools should be subsidized when they provide education to students with particular needs that the public sector cannot satisfy.

When comparing the UK and Finland, the two comprehensive systems in the model, it is possible to deduce that the large disparities between British schools (mainly in terms of peer quality) lead to the high significance on school variables. Furthermore, the social inequalities that exist on the population level also mean that student characteristics will generate inequalities. Even though some characteristics of the Finnish system (low school choice due to low population density, limited private education, limited social inequalities and established traditions of social democracy that favours equality, etc) cannot be reproduced in the UK, some lessons can still be drawn. Firstly, despite the fact that social inequalities at the population level cannot be dealt with in the short term, compensation in favour of less privileged students can reduce the dispersion of ESCS and hence its effect on achievements. Secondly, school homogeneity can be increased by phasing out or at least by limiting the numbers of grammar schools and differentiated schooling alternatives (see David Willetts's speech on grammar schools – The Daily Telegraph 16 May 2007); this should lead to reduced effect and significance of school variables. Homogeneity can also be enhanced through wider compensation policies. Note that in the UK an increase of the number of computers connected to the web of 2.57 units compensates a decrease in DESCS of 1 unit and a decrease of ESCS of 2.14 units (achievements held fixed). Finally, the privileged charitable tax status granted to private schools, while helping to maintain freedom of choice outside the public system, does

not encourage homogeneity of educational supply and constitutes an indirect subsidy to schools that are selective and socially elitist.

The regression results also show that some student attitudes (interest and anxiety in mathematics) have strong effects on performance scores. Students with high levels of interest in mathematics are likely to invest more time and effort in learning, and are more able to control their learning strategies and to monitor their progress. Since interest and anxiety depend on previous successes and failures, it is essential for any educational policy to be able to teach students how to overcome their failures and to assess the efficiency of their learning strategies. This can be done through different activities, undertaken by teachers and schools and aimed at fostering efficiency in learning. Moreover, schools should be able to cater for the special needs of students by organizing instruction according to their types.

Concluding Remarks

On the one hand, the research here sheds light on the mechanisms of stratification and inequalities in attainments. On the other hand, the comparative analysis allowed for a better understanding of the functioning of these mechanisms under different schooling systems.

The results showed clearly that comprehensiveness-driven school homogeneity is a source of equality since it dilutes the impact of schools on performance scores. Moreover, the trade-off between equity and efficiency (high average achievements) does not necessarily exist since Finland combines high levels of achievements with high levels of equity in their distribution. The rest of the countries have higher levels of inequalities than Finland for different reasons. Early selection and the high levels of social disparities in Germany mean that inequalities are transmitted through school and household characteristics. This is also the case in the UK even though inequalities are more moderate. The UK also retains the highest level of elitist private schooling. In Italy, household characteristics have limited effects and inequalities are transmitted through school characteristics. This finding reflects school heterogeneity in terms of their funding and peer quality levels. The case of Japan is probably the most ambiguous because Japanese students have been tracked into differentiated high schools three months before the PISA tests. Though, what is possible to say is that increased school heterogeneity will definitely lead to higher school-generated inequalities.

In general, I can conclude that the delayed selection of students (Finland) is associated with limited and delayed inequalities. Other major findings include the following. Firstly, Private schooling is found to have a negative effect on performance scores in all countries expect in the UK, indicating that the apparent superiority of private schools is the result of better peer quality and funding. Secondly, social peer effects are non-linear in their means in three of the selected countries indicating that the distribution of peers within schools also affects their performances. Finally, it is important to note that this analysis can be further extended through the inclusion of country-level data that accounts for the non-school macro characteristics of each country. Furthermore, inequalities can also be treated as a dynamic concept if the necessary data is available.

References

Anderson, P. (1979) *Lineages of the Absolutist State*. London: Verso.

Arnott, R. and Rowse, J. (1987) 'Peer Group Effects and Educational Attainment.' *Journal of Public Economics.* Vol 32, pp. 287–306.

Barzel, Y. (1973) 'Private Schools and Public School Finance.' *The Journal of Political Economy*. Vol 81, No 1, pp. 174-186.

Boli, J. (1989) *New Citizens for a New Society: the Institutional Origins of Mass Schooling in Sweden*. Oxford: Pergamon Press.

De Bartolome, C. (1990) 'Equilibrium and Inefficiency in a Community Model with Peer Group Effects.' *Journal of Political Economy.* Vol 98, pp. 110-133.

Dore, R. and Sako, M. (1989) *How the Japanese Learn to Work*. London: Routledge.

Epple, D., Filimon, R. and Romer, T. (1993) 'Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes.' *Regional Science and Urban Economics.* Vol 23, pp. 585-610.

Epple, D. and Platt, G. (1998) ‘Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income.’ *Journal of Urban Economics.* Vol 43, pp. 23-51.

Epple, D. and Romano, R. (1998) ‘Competition between Public and Private Schools, Vouchers, and Peer Group Effects.’ *The American Economic Review*. Vol 88, No 1, pp 33-62.

Epple, D. and Romano, R. (2006) ‘Admission, Tuition, and Financial Aid Policies in the Market for Higher Education.’ *Econometrica*, Vol 74, N0 4, pp. 885-928.

Fernandez, R. and Rogerson, R. (1996) ‘Income Distribution, Communities, and the Quality of Public Education.’ *The Quarterly Journal of Economics*. Vol 111, No 1, pp. 135-164.

Green A. (1997) *Education, Globalization, and the Nation State*. Basingstoke: Palgrave Macmillan.

Green, A. (1999) ‘East Asian Skills Formation Systems and the Challenge of Globalization.’ *Journal of Education and Work*, Vol 21, No 3, pp. 253 – 279.

Green, A. (2008) ‘Le Modèle de l’École Unique, l’Égalité et la Chouette de Minerva.’ *Revue Française de Pédagogie*, No 164, pp. 15-26.

Green, A., Janmaat, G. and Han, C. (2009) *Regimes of Social Cohesion.* LLAKES Research Paper No 1. London: Institute of Education.

Green, A., Preston, J. and Janmaat, J. (2006) *Education, Equality and Social Cohesion: A Comparative Analysis*. London: Palgrave Macmillan.

Gill, J. (2008) *Bayesian Methods: a Social and Behavioural Sciences Approach* (Second Edition). London: Chapman and Hall/CRC.

Hanushek, E. and Welch, F. (2006) *Handbook of the Economics of Education*. London: Elsevier.

Hoxby, C. and Weingarth, G. (2005) *Taking Race out of the Equation: School Reassignment and the Structure of Peer Effect.* Working paper, Harvard University.

Ichikawa, S. (1989) 'Japanese Education in American Eyes: a Response to William K Cummings.' *Comparative Education*, Vol 25, No 3, pp. 303 – 307.

Mons, N. (2004). *De l'École Unifiée aux Écoles Plurielles : Évaluation Internationale des Politiques de Différenciation et de Diversification de l'Offre Éducative*. Thèse de Doctorat, Université de Bourgogne.

Marsden, D. and Ryan, P. (1995) 'Work, Labour Markets and Vocational Preparation: Anglo-German Comparisons of Training in Intermediate Skills' in Bash, L. and Green, A. (eds) *Youth, Education and Work*. London: Kogan Page, pp. 67-79.

Mostafa, T. (2009) *The Anatomy of Inequalities in Attainments: an International Investigation on Stratification and Choice*. PhD Thesis, Université de la Méditerranée.

Nechyba, T. (1997) 'Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting.' *Economic Theory*. Vol 10, No 2, pp. 277-304.

Nechyba, T. (2003) 'Centralization, Fiscal Federalism, and Private School Attendance.' *International Economic Review*. Vol 44, No 1, pp. 179-204.

OECD (2003a) *PISA Data Analysis Manual for SAS Users*. Paris: OECD.

OECD (2003b) *PISA Technical Report*. Paris: OECD.

OECD (2003c) *Learning for Tomorrow's World*. Paris: OECD.

OECD (2006) *Science Competencies for Tomorrow's World*. Paris: OECD.

Patterson, H. and Thompson, R. (1971) 'Recovery of Inter-Block Information when Block Sizes are Unequal.' *Biometrika*, Vol 58, pp. 545-554.

Raudenbusch, S. and Bryk, A. (2002) *Hierarchical Linear Models* (Second edition). London: Sage Publications.

Robert, C. and Casella, G. (2004) *Monte Carlo Statistical Methods* (second edition). New York: Springer-Verlag.

Rose-Ackerman, S. (1979) 'Market Models of Local Government: Exit, Voting, and the Land Market.' *Journal of Urban Economics.* Vol6. pp. 319-337.

Schafer, J. (1997) *Analysis of Incomplete Multivariate Data.* London: Chapman and Hall.

Stiglitz, J. (1974) 'The Demand for Education in Public and Private School Systems.' *Journal of Public Economics.* Vol 3. No 4, pp. 349-385.

Westhoff, F. (1977) 'Existence of Equilibrium in Economies with a Local Public Good.' *Journal of Economic Theory.* Vol 14, pp. 84-112.

White, M. (1987) *The Japanese Educational Challenge*. London: Macmillan.

For more information, please contact
Jeremy Tayler: J.Tayler@ioe.ac.uk

LLAKES Centre
Institute of Education
20 Bedford Way
WC1H 0AL
London
UK

National Institute of Economic and Social Research

University of BRISTOL

UNIVERSITY OF Southampton School of Education

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL