WHAT IS THE MOST EFFICIENT WAY TO IMPROVE AN EDUCATION SYSTEM?

CONOR PARLE

Senior Sophister

In this econometrics investigation, Conor Parle attempts to identify the factors which explain variation in educational achievement across OECD countries. By constructing a regression model, he finds interesting results in relation to student socioeconomic background, teacher development, and teacher-pupil ratios.

Introduction

This paper aims to answer the question on what is the best way for a country to improve its education system by examining data from the OECD member states. The idea for this project stems from research on the concept of mathematics education, and in particular, the recent results of the Program for International Student Assessment (PISA) survey, which takes place every 3 years. The 2012 results were recently released by the OECD to much intrigue and discussion.

PISA, started in 2000, is a survey undertaken by the OECD every 3 years. It assesses an education system on three main areas: maths, science, and reading in order to get an overall picture to describe the relative quality of a country's education system. Standardised tests are used to measure the ability of students aged between 15 years and 3 months, and 16 years and 2 months during the time period in question. These marks are then converted into a coefficient which is released every three years. Raw scores from the assessments are scaled to an OECD index, with the overall general average being 500, and the standard deviation being 100. The scaling is done using the Rasch Model of Item Response Theory (OECD, 2013).

This paper aims to account for the differences in standards of education between countries, and will attempt to discover what policies to improve standards have the greatest effect on overall outcomes.

Literary Review

There have been a number of articles exploring the differences in education systems between countries and there are many different theories as to which factors cause the greatest differentials in overall outcomes. Saleem, Nasseem, Ibrahim, Hussain and Azeem examine several schools of thought regarding the differences in education standards in their 2012 paper on "Determinants of School Effectiveness: A study at Punjab level". They show how the top systems often have highly skilled and motivated teaching staff and a high level of community involvement. However, they equally give heed to the idea that perhaps one of the biggest determinants of education standards is the socioeconomic climate, with lower-class students regularly performing worse than higher-income students, suggesting that educational equity is hampered by income inequality. This idea is also suggested by Raymond (1968) in a paper on the quality of schools in West Virginia. He argues that exogenous conditions such as the location of a school, and the surrounding economic climate have the greatest impact on the overall quality of a system.

A deeper inspection on such financial factors is made in the paper "Does Money Matter in Education" (Schanker Institute, 2012). In this paper there is a thorough examination of the financial factors which impact on the quality of an education system. One finding is that money matters when it comes to education, but wise and careful spending is necessary if a system is to reap the benefits of it. There is also emphasis on the idea that schemes that cost a significant amount of money, such as reducing class size and higher teacher salaries, are positively associated with overall student outcomes. Smaller class sizes in particular can go a long way to close the divide in standards that are caused by socioeconomic factors. This idea is backed up in the paper "Improving Educational Outcomes for Poor Children" (Jacob and Ludwig 2009).

Jacob and Ludwig extend the discussion on the financial side of improving educational standards, but also reinforce the idea of the undeniable impact that socioeconomic factors appear to have on the system. They contend that good social policy is just as important as education spending, while arguing that training, and research and development are perhaps the most important components of teaching spending.

In a 2007 paper, McKinsey extrapolate data from PISA to compare the education standards in the OECD. They use data from Singapore, which is ranked 27th out of the 30 OECD countries in terms of education spending, yet regularly comes out near the top of rankings to reinforce the idea of ‘it’s not how much you spend, but how you spend it’. They divide the process of developing a good system into three main steps: firstly, getting the right people to become teachers; secondly developing them into efficient instructors; and finally, ensuring that the system is able to deliver the best instruction for each child. This stresses the idea that by increasing the quality of the system, education standards can be raised, and further points to the benefits of increasing spending on research, development, and training as a means to improve a system.

Overall, previous research points to several factors which could explain what makes up the largest portion of the difference in cross-country education standards. This paper will focus on whether social factors, such as the allocation of government spending, or the fundamentals of the education system itself (e.g., class sizes, methods of teaching, etc.), are significant determinants of educational performance.

Empirical Approach

In order to hypothesise the potential determinants which are outlined above, the following regression model is proposed:

$$(1)\quad Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$$

Where:

Y_i = PISA - the average PISA score of a country over the three areas of study in a given year (maths, science, and reading).
X_1 = GINI - the GINI coefficient for a country in a given year based off World Bank data, which is a measure of income inequality in a country.
X_2 = CLSIZE - the average pupil to teacher ratio for a country in a given year.
X_3 = SALARY - the average salary received by an upper secondary teacher after 15 years of service and training, in dollars per year adjusted for PPP.
X_4 = RDTRAINS - the portion of educational spending spent on core educational facilities – e.g., research and development and training.
X_5 =THIRDLEVEL - the percentage of students who go on to tertiary "Type A" education. This acts as a proxy for overall expectations of educational attainment.
X_6 = CONTACTHOURS - the number of compulsory contact hours in a year between teachers and students in upper secondary school.
ε_i = an error term consisting of statistic residuals.

Expectations

We would expect there to be a negative relationship between PISA and GINI. This is due to the widespread theory that the overall socioeconomic climate has a major effect on academic performance, as theorised in Jacob and Ludwig (2009). A high GINI coefficient implies high income inequality, and so the higher the GINI coefficient, the lower the overall PISA score.

Similarly, a negative relationship between PISA and CLSIZE is expected as theorised by the Schalker Institute (2012). They found that reducing class size would be an efficient way to increase education standards.

The expected relationship between SALARY and PISA is positive, due to the theory discussed in Saleem et al., (2012) and by the McKinsey Institute (2007). Here, higher teacher salaries were shown to increase teacher motivation, and further attract the most qualified people into the teaching profession.

Furthermore, a positive correlation between PISA and RDTRAIN is expected, again, due to the idea of the quality over quantity approach when it comes to spending.

Both the Schalker Institute (2012) and McKinsey (2007) theorised that investing in research and training methods would have a positive impact on educational achievements. Adding to this, a positive correlation is expected between THIRDLEVEL and PISA, mainly due to the idea that a high third level attainment rate generally suggests a positive communal outlook on education, and this in turn is said to impact positively on the success of an education system (Saleem et al., 2012).

Finally, the effect of CONTACTHOURS is ambiguous given several opposing arguments in the existing literature. Some contend a degree of saturation when it comes to contact hours (McKinsey, 2007), suggesting that the quality of these hours and class size is more important. Conversely, a positive relationship is perhaps more intuitive.

Potential Issues

There are a number of potential problems with the model. Firstly, the international financial crisis of 2008 may cause a number of figures to be greatly skewed. It is for this reason that I may run the regression with figures excluding those collected after 2009, alongside the regression with the full data.

The model may also omit a relevant variable or may be under-specified (Wooldridge, 2009). This may occur, for example, if variations in PISA scores are merely due to differences in the ability of students sitting the test in a given year. Such differences in ability are not observed and end up within the error term, causing bias.

Furthermore, there is also the potential threat of heteroskedasticity. This is the idea that the variance of the unobservable error ε conditional on our explanatory variables is not constant given our observed variables (Wooldridge, 2009. pg.264). Some omitted variables potentially included in the error term may not be constant throughout the different inputs into the model. Should this exist, then the usual OLS t-statistics will not have t-distributions, and furthermore, the F-statistics will no longer be F-distributed, causing problems with the ability to test the significance of the model.

Data

The majority of the data present comes from the OECD Publication "Education at a Glance" (OECD, 2003, 2006, 2009, 2012), which is a yearly in depth publication of indicators relating to education around the world. All data except PISA scores and GINI coefficients are taken from this source. The historical PISA scores are taken from the OECD educational GPS website. Finally, all data regarding GINI coefficients was extrapolated from the World Bank's historical economic indicators page.

China is excluded from this study due to recent revelations that they may be cheating the PISA system by only choosing schools in affluent regions, where parents spend a significant amount on private tuition (Stout, 2013) and preparing students greatly via rote-learning to specifically take the PISA tests (Ringmer, 2013).

Overall, the data can be summarised in Table 1. There are 78 observations overall, dropping sets with missing data, which is a relatively decent size for an international study.

Variable	Obs	Mean	Std. Dev	Min	Max
Pisa	78	497.9744	32.33308	389	554
ginicoeff	78	.3131477	0.0582474	.232	.508
clsize	78	23.3	4.234429	15.6	35.8
salary	78	38881.63	15494.3	10263	101775
rdtrain	78	3.747179	.6239129	2.6	5.3
thirdlevel	78	53.97436	16.29833	26	96
contact	78	816.8333	200.8963	466	1197

Table 1: Data Summary

Results

The overall results of the study can be found in Table 2. The GINI coefficient was multiplied by 100, to get the figures in percent as opposed to in a proportion. This has led to a smaller slope coefficient of -3.36 as opposed to 336, making the results easier to interpret.

The regression equation theorised in equation 1 can now be written as:

$$(2) \quad PISA = 540.96 - 3.36\beta_1 + 1.8\,\beta_2 + .0073\beta_3 - .8653\beta_4 + .349\beta_5 - .03\beta_6$$

Standard errors are shown in Table 2:

pisa	coef	SE	t	p-value
gini100	-3.36	0.49	-6.85	0.00
clsize	1.81	0.58	3.13	0.00
salary	0.00	0.00	4.50	0.00
rdtrain	-0.87	3.88	-0.22	0.82
thirdlevel	0.35	0.17	2.11	0.04
contact	-0.03	0.01	-2.29	0.03
_cons	540.96	24.35	22.22	0.00

Table 2: Regression Output

Interpreting the Results

Overall, all variables are found to be statistically significant when tested against the null hypothesis of $\beta_n = 0$ using a t-test at the 5 per cent significance level, apart from RETRAIN, which has a p-value of 0.824, leading us to fail to reject our null hypothesis. This is rather interesting, although it may be due to the manner of data collection and perhaps it suggests that the overall standard of education may be more related to spending per student as opposed to the portion of overall expenditure on research and development.

Furthermore, the F-statistic for this regression is 21.64, meaning that the p-value is 0, implying that this model is statistically significant at predicting PISA scores. We have an R2 value of 0.6465, which means that this model explains approximately 64.65 per cent of the variation in PISA scores.

It is interesting to note the direction of the signs of the dependent variables. The highly significant nature of the GINI coefficient suggests a high degree of correlation between PISA scores and income inequality. This is in line with the theory as discussed in Jacob and Ludwig (2009). For an increase in GINI coefficient of 0.01, i.e., 1 per cent, there will be a fall in PISA score of 3.36, ceteris paribus. This suggests a strong link between the two, and highlights the role good social policy has to play in establishing a good education system.

The positive coefficient on class size is, however, surprising, as it suggests a positive relationship between student to teacher ratio and PISA scores. This supports the argument which was initially made by McKinsey (2007) that lower teacher to student ratios may not be such a good thing all round, as it causes people who are hiring to be less selective about the calibre of people who enter the profession.

The positive coefficient on third level and salary, however, is unsurprising. The positive relationship between third level and PISA is intuitive. There may in fact be a degree of reverse causality, however, as higher test scores usually imply more people going to university. Nonetheless, we used THIRDLEVEL to be a proxy for the overall attitudes to education, and the positive relationship is once again in line with theory as outlined in papers such as Saleem et al. (2012). Similarly, the positive relationship between PISA and teachers' salaries is in line with the theory in McKinsey (2007), where it was theorised that high teacher salaries lead to higher motivation of teachers and in turn leads to a higher calibre of individual entering the profession.

Finally, the negative nature of the relationship between contact hours and PISA may not be too surprising. It may seem counter intuitive, but could also highlight diminishing marginal returns to contact hours, suggesting that perhaps only a smaller amount is necessary. Countries such as Finland are a leading example of having a low number of contact hours and high PISA results.

Omitted Variables in the Model

We can check as to whether any variables have been left out by running a Ramsay RESET test for omitted variables. The results in Table 3 suggests that there will be some degree of omitted variable bias in the model. This could be as a result of omitting variables such as IT standards or natural ability of a country's students.

Ramsey RESET Test	
F Statistic	3.09
p-value	0.03

Table 3: Ramsey RESET Test

To check if the financial crisis had a major impact on the results, the model was rerun with only the values from 2003, 2006 and 2009 included. The results are outlined in Table 4.

pisa	coef	SE	t	p-value
gini100_01	-3.36	0.60	-5.63	0.00
clsize_01	1.19	0.61	1.97	0.06
salary_01	0.00	0.00	5.50	0.00
rdtrain_01	-4.93	4.41	-1.12	0.269
thirdlevel_01	0.68	0.21	3.27	0.00
contact_01	-0.02	0.01	-1.42	0.162
_cons	529.41	27.50	19.25	0.00

Table 4: Pre-Crisis Model

It is notable that contact hours are no longer statistically significant. This may be due to the lack of data given the smaller amount of observations; this may also explain the lack of significance of class size. Similarly, it is interesting to note the effect of salary is higher, perhaps in line with the fall in wages in some countries post-recession. On the whole however, the model is quite similar. The signs on all variables remain the same, suggesting that there was not much of a difference in the pre- and post-crisis determinants of PISA scores, with the major notable change lying in the difference in salary coefficient.

Extensions to the Model

There are a number of possible extensions to the model. Firstly, it would be interesting to see the relationship between research and training spending per student and test scores as opposed to the portion of expenditure discussed earlier. Furthermore, more variables could be included given that there is evidence of causal variables having been omitted. For example, teachers' job satisfaction could be measured via a survey or even teachers' salary relative to the national average. Moreover, it would be interesting to explore a link between positive family life and any positive impact on education. Further study into the relationship between contact hours and PISA scores could be carried out to ascertain if the negative impact only really exists at higher contact hours.

Conclusion

The aim of this econometric investigation was to explore how education systems may be improved. In conclusion, it is clear that the theory surrounding socioeconomic factors having a major impact on education standards is somewhat valid. This is evident by the statistically significant relationshii between GINI coefficients and PISA scores, suggesting that it has a great degree of causality. One implication following from this is that perhaps one of the best ways to improve overall standards is social policy to try and close the income gap. In practice, more opportunities can be given to those who are worse off. Furthermore, the negative relationship between class size and PISA along with the positive relationship between teachers' salary suggests that one of the best ways to improve an education system is to concentrate on the teacher quality. Funding should therefore be targeted into the attraction of the best candidates to the job rather than hiring extra teachers to reduce pupil teacher ratios as this seems to not have as strong a positive effect as suggested.

References

Jacob, B, Ludwig, A. 2009. Improving Economic Outcomes for Poor Children, [on-line], http://www.irp.wisc.edu/publications/focus/pdfs/foc262j.pdf.

McKinsey Institute 2013, How the World's Best-Performing School Systems Come out on Top. [on-line], http://tinyurl.com/pj6phj.

OECD. 2003. Education at a glance 2003, charts and tables 2003. [on-line], http://tinyurl.com/orykyul.

OECD. 2006. Education at a glance 2006, charts and tables 2006. [on-line], http://tinyurl.com/po3lvvl.

OECD. 2009. Education at a glance 2009, charts and tables 2009. [on-line], http://www.oecd.org/education/skills-beyondschool/educationataglance2009indicators.html.

OECD. 2012. Education at a glance 2012, charts and tables 2012 [on-line], http://www.oecd.org/educ/EAG%202012_e-book_EN_200912.pdf.

OECD. 2013. Educational GPS 2013, [on-line], http://gpseducation.oecd.org.

OECD, "About Pisa", 2013 [on-line], http://www.oecd.org/pisa/aboutpisa.

Raymond, R. (1968). Determinants of the quality of private and public primary and secondary education, 3(4), pp4.

Ringmar, S. 2013, Here's the truth about Shanghai schools, They're terrible. [on-line], http://tinyurl.com/mcueo5h.

Saleem, F, Naseem, Z, Ibrahim, K, Hussain,A, Azeem, M. 2012. Determinants of School Effectiveness: A Study at Punjab Level, International Journal of Humanities and Social Sciences. [on-line], http://www.ijhssnet.com/journals/vol_2_no_14_special_issue_july_2012/28.pdf.

Stout, D. 2013. China is Cheating the World Student Rankings System. [on-line], http://tinyurl.com/lbthevy.

The Albert Shanker Institute. 2012. Does money matter in education? [on-line]. http://www.shankerinstitute.org/images/doesmoneymatter_final.pdf.

Wooldridge, J. 2009. Introductory Econometrics, a Modern Approach, fourth edition. Cengage Learning.

World Bank, GINI Coefficients, 2013, [on-line], http://data.worldbank.org/indicator/SI.POV.GINI.