Test scores and economic performance, a brief literature overview

Niels Vermeer

CPB Achtergronddocument bij CPB Policy brief 05, 2011, Nederlandse onderwijsprestaties in perspectief, M. van der Steeg, N. Vermeer en D. Lanser

Summary

We briefly review the literature on returns to education. The literature consists of studies at both the micro level, which study the relationship between schooling, test scores and wages, as macro level, which study the relationship between schooling, international test scores and economic growth. The micro studies find returns of around 12 percent to a standard deviation higher test scores. Macro studies generally find considerably higher returns, but estimates vary more widely.

1 Introduction

The consensus in the literature is that returns to education are substantial. A large strand of research on returns to education has focused on the returns to years of education. The micro literature on wage returns to a year of education finds returns to years of education that vary between 5 and 10 percent. The macro literature finds GDP returns that vary between 7.5 and 12.5 percent.[1] (see van Elk et al., 2011).

More recently, the literature on returns to education also started to focus on the returns to skills, as proxied by test scores. The fundamental difference is that test scores can be seen as an output measure, whereas years of education are basically seen as an input measure. The value added of looking at test scores is that test scores allow to take into account quality differences of a year of schooling across schools or countries.

This discussion note briefly summarizes the recent literature on the returns to education with a particular focus on returns to skills. We start with the micro-literature on the private wage returns to skills, and then discuss the macro literature which looks at the effect of skills on GDP.

Van Elk et al (2011) provide a more detailed discussion of the literature on the returns to years of education.

2 Micro literature

We discuss three papers which shed light on the wage returns to skills (or test scores). These are Lazear (2003), Mulligan (1999) and Murnane et al. (2000).

Lazear (2003) discusses incentives for teachers for improving their output, the skills of their pupils. The author also performs some empirical research on US data. The dataset he uses is the National Education Longitudinal Study (NELS) of 1988. It provides test scores of pupils from the eighth grade (fourteen year olds). In subsequent years the earnings of the former students are reported. This study ends in 1999 so the oldest individuals are 27 years old at that time. The mean of the age is 25. Wage is the dependent variable and is regressed on the test scores and some observables like age and whether or not the parent went to college. The test scores are defined as the sum of the standardized scores on reading, math, history and science tests at the beginning of the survey. The effect of the change in test scores on wages is positive and significant. The estimated magnitude of the effect of a standard deviation increase in test score on earnings is 12 percent.

Lazear's estimate of 12 percent could well be a lower bound because of the relatively young age at which wages are measured. The earnings are recorded when the former students have an age of around 25 years. The reason is that individuals with greater cognitive skill are likely to engage in more training early in life, and thus to have steeper earnings age profiles than those with less cognitive skill. Thus studies using younger workers probably understate the lifetime labor market returns to cognitive skill. Another reason is that those with very high expected earnings at older higher ages may still be investing in skills when they are in their mid-twenties.

Mulligan (1999) discusses a model for human capital. He also analyses a micro data set National Longitudinal Study of the Youth (NLSY) and looks at the relation between wages, schooling and test scores. The dependent variable is hourly wages. This variable is regressed on various levels of schooling, years of schooling and test scores in one large model. The normalized test scores are obtained from the Armed Forces Qualification Test (AFQT) and are administered by the NLSY. AFQT consists of reading and math skills. In this regression the test scores and the years of schooling have a positive significant effect. The conclusion is a standard deviation increase in test scores increases the hourly wage by 11 percent. It should be stressed that this study only measures the direct effect of skills on wages since years of schooling is controlled for in the regression analyses. The total effect is expected to be larger, since it also includes the positive indirect effect of skills on wages that runs via its positive effects on years of education. Murnane et al. (2000) further address this issue of direct and indirect wage effects of skills.

Murnane et al. (2000) distinguish between the indirect and the direct effect of math skills on wages. The direct effect consists of wage differentials due to cognitive skills keeping educational attainments equal. The indirect effect is the effect of skills on wages which runs via the effect of skills on educational attainment. The authors use two different data sets, the National Longitudinal Survey of the High School Class of 1972 (NLS72) and High School and Beyond (HS&B). The age at which the earnings are recorded is 31 years in NLS72 and 27 years in HS&B.[2] In their first model they perform a regression of earnings on some background characteristics

[2] An important caveat of HS&B is that earnings are recorded annually and not hourly. Consequently it is difficult to distinguish between higher skilled working more hours and having a larger labour productivity.

with and without math score, but without schooling. So here the coefficient is a measure for the total effect of skills on wages. In a second model they include the amount of schooling on top of the parameters of the first model, which provides a measure of the direct wage effects of skills. For males they estimate a total effect on earnings of 11-15 percent and for females the total effect is estimated at 9-12 percent, dependent on the database used. For males two/thirds are direct effects. For females less than 20 percent are direct effects in NLS72 and somewhat more than 50 percent in HS&B. Taking together this leads roughly to a half/half division of direct and indirect effects for both genders in total.

To sum up, the findings of the micro studies discussed above point to a total wage return to a one standard deviation higher test score of around 12 percent. This average may well be a lower bound of the real return, for reasons mentioned earlier.[3]

These findings would imply the following for the decline in the Dutch math scores from PISA 2003 to PISA 2009. The decline of 12 points corresponds to a decline of 0,12 standard deviation.[4] Consequently the decrease in mathematics scores would lead to a decrease of 1.5 percent in future earnings or GDP in the long run.

3 Macro literature

On the macro level it has also been most common to look at schooling attainment. Some studies are Delafuente and Domenech (2006), Cohen and Soto (2007), Teulings and Van Rens (2008) and Coe et al. (2009). These studies find social returns roughly between 7.5% and 12.5% of raising the average educational attainment by one year. So the private returns and the social returns to a year of education roughly coincide.

In the current literature there is much discussion about using test scores instead of average schooling attainment. The basic idea is that test scores are a direct proxy for human capital whereas average schooling attainment is more of an input measure for the production of the relevant skills.

Hanushek and Woessmann (2009) are an example of this approach. They construct a database of average test scores and economic growth between 1960 and 2000. Controlling for the level of GDP of 1960 they find a significant relationship between economic growth and average test scores. The inclusion of average years of schooling in a country in 1960 does not enter significantly in the growth equation. The effects are robust for the inclusion of instruments like openness and property rights.

Hanushek and Woessmann estimated the following equation: $average\ annual\ growth = c + \alpha_1 T + \alpha_2 GDP_{1960}$. In this T represents test scores (expressed in standard deviations), c represents a constant and GDP is initial GDP level per capita. Using steady state this becomes: $0 = \alpha_1 \Delta T + \alpha_2 \Delta GDP$. The equation can be converted into an equation with GDP growth (so with log(GDP)) instead of initial GDP level. The result is

[3] These reasons are: 1) returns are generally measured at a relatively young age at which the lifetime benefits of better skills are probably not fully visible yet. and 2) one of the three studies (Mulligan) has estimated a direct effect instead of the total effect of skills. Based on the results of Murnane et al. on the relative size of direct and indirect effects, the total effect may be up to twice as large as the estimated direct effect.

[4] One standard deviation amounts to 99 points in PISA 2003 and 96 points in PISA 2009.

the following equation: $\Delta logGDP = \frac{-\alpha_1}{\alpha_2 GDP_{1960}} \Delta T$. In this equation we can plug the numbers: $\alpha_1 = 1.26, \alpha_2 = -0.35$ and $GDP_{1960} = 5$, so : $\Delta logGDP = 0.72 \Delta T$. The first two parameters are found in column 7 of table 1 while the GDP level of 1960 per capita is found in table C1. Hence, a 0.12 standard deviation loss in the test score as witnessed in the Netherlands at math would yield a loss in GDP of 0.12 x 0.72 = 9%, which is highly compared to the earlier reported 1.5 percent based on the micro estimates.

Appleton et al. (2008) adjust the results of HW for the fact that the period of GDP growth used as variable to be explained in the regression largely dates before the test scores results. The motivation for this is to deal with the possibility of reverse causality. A larger income could imply more expenditure on education and therefore rising test scores. In general they find estimates which are about half of those of Hanushek and Woessmann.

Hanushek and Woessmann do not include physical capital in their growth regression. Breton (2011) also includes physical capital in similar growth regressions. As Breton mentions there is a high correlation between stocks of physical capital and human capital. He tries to remedy this by estimating reduced form equations and using the estimated value of the share of the physical capital compared to the national income. He compares these estimated values to the empirical value of 35% found by Bernanke and Gurkaynak (2001). The result is that average schooling attainment is more significant than average test scores, although the average test scores are still significant.

Breton finds that a one standard deviation increase in test scores corresponds to a 30 percent larger GDP. Based on the estimates presented in table 2, column 5 and column 8, the decrease in average test scores of 12 points which is witnessed in the Netherlands for math between 2003 and 2009 would correspond to a decline of 3.6 to 5% GDP in the long run steady state, which is also about half the size of the predicted effect based on the HW estimates. This is still quite large compared to the estimates from the micro literature.[5]

We conclude from the available literature on returns to skills that the loss of 0.12 standard deviation in math scores may cost the Netherlands a couple of percentage points of GDP in the long run. More research on this subject is needed and will probably shed more light on the difference between micro and macro estimates. This might also clarify further the difference in magnitudes of the various macro studies.

[5] Furthermore, an increase of one year in average schooling attainment would lead to an increase of 10% to 14% GDP in the long run, which is somewhat larger than the earlier macro studies. In those studies the upper bound was around 12.5%.

4 Literature

Appleton, S., P. Atherton and M. Bleaney, 2008, International School Test Scores and Economic Growth, Credit Research Paper, No. 08/04

Bernanke, B.S., R.S. Gurkaynak, 2001, Taking Mankiw, Romer and Weil seriously, NBER Macroeconomics Annual, 16, 11-57

Breton, T.R., 2011, The quality vs. the quantity of schooling: What drives economic growth?, *Economics of Education Review*.

Coe, Helpman en Hoffmaister, 2009, International r&d spillovers and institutions, *European Economic Review*, vol. 53, pag. 723–741.

Cohen en Soto, 2007, Growth and human capital: good data, good results, *Journal of Economic Growth*, vol. 12, pag. 51–76.

Delafuente, A. en R. Domenech, 2006, Human capital in growth regressions: how much difference does data quality make?, *Journal of the European Economic Association*, vol. 4, pag. 1–36.

Hanushek, E., L. Woessmann, 2009, Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, NBER Working Paper 14633

Lazear, E., 2003, Teacher incentives, *Swedish Economic Policy Review*, vol. 10, pag. 179–214.

Mulligan, C., 1999, Galton versus the human capital approach to inheritance, Journal of Political Economy, vol. 107, pag. S184–S224.

Murnane, R., J. Willett, Y. Duhaldeborde en J.H. Tyler, 2000, How important are the cognitive skills of teenagers in predicting subsequent earnings?, *Journal of Policy Analysis and Management*, vol. 19, pag. 547–568.

Teulings, C. en T. van Rens, 2008, Education, growth and income inequality, *Review of Economics and Statistics*, vol. 90, pag. 89–104.

Van Elk, R., D. Lanser , S. Van Veldhuizen, 2011, Onderwijsbeleid in Nederland: de kwantificering van effecten.