Estimating the social returns to education

John Norell

October 2019

Abstract

This term paper describes how one could estimate how average incomes in society are affected by increases in average years of education. According to the signaling theory the total social returns to education is smaller than the sum of the private returns while others suggest positive externalities of education makes the social returns to education to be larger than the sum of private returns. Using the constant composition approach to deal with the potential bias from imperfect substitutability of differently skilled workers and with demographic structure as instrument, the social returns to education can be estimated which can help us understand whether investments in education on net enriches or impoverishes society.

1 Introduction

Many researchers have found that there is a positive causal effect of educational attainment on earnings. The mechanism of this effect is however not well understood. There are two competing theories: the *human capital theory* and the *signaling theory* (Caplan, 2018). The human capital theory states that additional schooling endows students with skills that makes them more productive in the labor force which in turn results in higher earnings in the labor market. The signaling theory instead propose that education don't increase a person's human capital but only reflects their inherent human capital. Education then serves a (costly and hence hard to fake) signal to employers of a person's inherent ability and productivity[1]. The signaling theory can help explain why much of what is taught in schools don't appear to correlate with the tasks that the students will perform once they get into jobs. Educational attainment signals not only intelligence to

[1]Seminal work on signaling include Spence (1973), and Arrow (1973) among others

employers, but also aspects of conscientiousness which are important for performance on the job.

For the individual's decision of how much education to get, it doesn't matter why education yield higher earnings on the labor market. From a societal perspective however, it matters greatly if education increases human capital, if the wage premium arise due to signaling or if it's a combination of the two. If the effect is due to increases in human capital only, then increases of average education in the population will come with increases in national income that are proportional to the effect of education on income for the individual. But if the effect is a mix of human capital increases and signaling effects then the effect of national education on national income will be less than proportional to the private returns to education, and in the extreme case where education is only signaling, then there will be no effect at all of national education on national income. Instead, increased national education will result in credential inflation where employers raise their requirements on educational attainment not because the job tasks become more complex but simply because education is used for screening.

The signaling theory suggests that it's not years of education in absolute terms that determine your wage but how much education you have in relation to other people in the labor market. Individual's investments in education then produces a *negative externality* as they then diminish the value of everyone else's education by increasing the average educational level in society. Other researchers have argued that there are positive externalities from education as well. A couple of channels have been proposed as for why a positive externality arise. The sharing of knowledge and skills through formal and informal interaction may generate externalities across workers (Lucas, 1988) or societies with more skilled workers generate more ideas and grow faster (Romer, 1990).

From a policy perspective, it's of importance if the social returns to education are larger, smaller or equal to the sum of private returns to education (i.e. the net effect of the positive and negative externalities of education). If the social returns are larger than the private returns, then there might be reason for subsidising investments in education, while if they are smaller, subsidising education will result in credential inflation and malemployment.

2 Previous research

The private returns to education we here define as the effect on own wage (in percent) or income of an additional year of education. Social or societal returns to education equals the private return plus potential externalities on the incomes or wages of other people. If the private returns are equal to the social returns, then national income increase in the same proportion to national education as own income increase to own education.

There is a large literature on the private returns to education. The estimates differ by educational level (for example between high school or college) and by country. In the USA, an additional year of education raises earnings by around 5-10% (Caplan, 2018). In a comprehensive study of 50 countries, (Psacharopoulos and Patrinos, 2004) the effect is estimated to 7.4% in high-income countries, 10.7% in middle-income countries and 10.9% in low-income countries. For the world as a whole, the estimate is 9.7%. Caplan (2018) notes that the study – as almost all international studies on the private returns to education – fail to control for ability bias, why the estimated effect is likely biased upwards and that the true effects are at least 25% smaller (Caplan, 2018, p313).

On an aggregate level, the effect of education on aggregate income is far less clear. Pritchett (2006), Bosworth and Collins (2003) and Kreuger and Lindahl (2001) find no significant effects of national education on national income. Lange and Topel (2006) review the literature and conclude that the empirical evidence on the existance of any positive or negative externalities is weak but argue the social returns are positive and probably in the magnitude of the private returns. The answers provided by researchers on the social returns to education are varied and almost all researchers note that their answers are highly uncertain (Caplan, 2018, p115).

A couple of papers have employed quasi-experimental approaches to estimate the social returns to education. Exploiting exogenous variation in educational attainment produced by reforms on compulsory schooling in the USA, Acemoglu and Angrist (2000) find positive externalities of increases in average education. Due to issues with their empirical specification (not taking into account the problem which may arise due to imperfect substitutability, further discussed in section 5), Ciccone and Peri (2006) find that their findings are biased upwards. Using the same data but with a different estimation approach to correct for this bias, Ciccone and Peri find no positive spillover effects. La (2014) take advantage of exogenous changes in schooling from reforms on compulsory schooling in China and find no significant returns to education on the aggregate.

Björklund (2000) surveys the swedish literature on the private returns to education on earnings. Three techniques have been used to deal with possible ability bias (people with higher expected future earnings tend to choose to consume more education): controlling for ability (either with IQ-scores or grades), twin studies and IV-estimates. These different approaches reach similar estimates, around 4-5% and that the effect is fairly linear meaning the effect is the same for most schooling years. The paper Björklund mentions which uses exogenous variation in schooling as an instrument to estimate the returns to schooling is Meghir and Palme (1999). In the study, the authors take advantage of a school reform in Sweden in the 1960s which lengthened the compulsory schooling from seven or eight years to nine years. By exploiting the gradual implementation of the policy in different municipalities, they estimated the effect of education on earnings to be slightly smaller than what the other studies find (3.6%, see p31).

To my knowledge, no one has applied quasi-experimental approaches to estimate the societal returns to education in Sweden. Most economists assume that the social returns to education are at least as large as the private returns.

3 Data

In order for us to estimate the social returns to education on aggregate level (at regional level), we need data on average *wages* or *incomes* and *years of education* in regions at different years (for example at years 1985, 1995, 2005 and 2015). Wages are probably preferred as these are easier compare and interpret, though positive spillovers could potentially emerge on incomes and not wages. When calculating average wage, we restrict the sample to only full-time employed so as to account for changes hours worked. For the empirical strategy we need also information on demographic structure in regions as well (age and sex for inhabitants as well as how years of education changed during the period of study for each of the subgroups, for example men 55-60 years). To analyse the validity of our instrument we also need data (for each region and year) on migration in- and outflows of regions (during the time period), and participation rate, labor size force and population size (at each year). Statistics Sweden have individual data which can be used to calculate the data points we need. Statistics are available from Statistics Sweden for approximately 1985 and forward.

4 Empirical Strategy

The question of interest for our study is how national income is affected by a change in average years of education in the population. If externalities of education are realised more locally, then we can conduct the analysis on city or region level instead. The theories about positive spillovers generally say that the spillover of human capital arise from the sharing of ideas and innovation within workplaces or between firms. We can hence expect these effects will be visible using region or city as our level of analysis. The negative externality from signaling can be thought of as credential inflation. If applicants for jobs mainly come from withing the same city or region as the job position or firm is located, then the negative externality will be visible on a city or regional level as well. Under the assumption that the externalities will be visible on regional or city level, its preferable as this gives us much more observations to work with. To account for the risk that city level is too small, I have chosen region as the level of analysis for this study.

Our goal is then to estimate how the average income in a region is affected by a change in average years of education in that region. For a number of reasons, it is not appropriate to simply regress average income on average education for regions. Firstly, we possibly have unobserved variables that affect both education and income at the aggregate level. For example, there might be region characteristics that lead to both higher incomes and to highly educated individuals moving there. Second, there is likely a problem of reverse causality. It's reasonable to suspect rich regions to invest more in education and/or attract individuals with higher education. Thirdly, we have a more subtle problem that may arise from imperfect substitutability (complementarity) of more and less educated labor. To illustrate the issue, consider a model with only two types of labor, educated and uneducated. The standard neoclassical model predicts that as more individuals become educated, uneducated workers will become more productive if two two types of labor are complements in the production. Wages of educated workers will naturally decrease while wages of uneducated will increase. Depending on the relative share of educated and uneducated workers in the labor force, total wages are likely to increase even though there are no real externalities of human capital (Moretti, 2004). A formal model and in-depth investigation of this bias and its size is provided in Ciccone and Peri (2006).

Moretti (2004) deals with this issue by looking at the effect of an increase in the share

of educated on wages of educated and uneducated workers separately. Ciccone and Peri (2006) on the other hand develop the constant composition approach to deal with the same problem. This method also has the advantage – compared to the conventional Mincerian approach employed by Acemoglu and Angrist (2000) and Moretti (2004) – is that it doesn't require us to estimate the individual returns to education in order to estimate the social returns. This means that we don't need individual data nor an instrument for exogenous changes in years of education at the individual level.

4.1 The Constant Composition Approach

The problem is that as education increase in the region, the composition of skill change which will yield a positive bias when trying to estimate the social returns to education. The constant composition approach in essence mean that we modify the outcome variable – average wage in region – such that the composition of skill stays constant across the years we study.

With information on wages w_x for different educational levels x as well as labor-force shares l_x, we are able to calculate the log-change in the average holding skill-composition constant $\Delta \log \hat{w}_r$ for a specific period for a region r with the equation

$$\Delta \log \hat{w}_r = \log(\sum_x \bar{l}_x \tilde{w}_x) - \log(\sum_x \bar{l}_x \bar{w}_x) \tag{1}$$

where bars denote beginning-of-period values and tildes end-of-period values[2]. We then regress the log-change in constant composition average wages on (exogenous) changes in the supply of human capital. If we get positive and significant coefficients, this indicate a positive effect of increases in education on average income in society.

Control variables used by Ciccone and Peri (2006) to consider including are log-change in region employment to capture aggregate scale-effects and the change in average years of potential experience in the regions.

We have now accounted for the possibility of complementarity between skilled and unskilled workers by using the constant composition approach. Next we need to account for the endogeneity of aggregate human capital. We do this using an instrument.

[2]Ciccone and Peri (2006), p 382

4.2 Using demographic composition as instrument variable

In order to identify the true societal returns to education, we need to deal with the endogeneity problems discussed in the beginning of this section. This is done using an instrumental variable based on differences in demographic structure of regions. The idea is that much of the increase in average years of education in the work force come from younger cohorts entering in the labor market which is more educated than the general work force and especially the cohort exiting the work force into retirement. To the extent that the relative population shares of different cohorts vary across regions, this will be a source of exogenous change in the supply of educated labor. The instrument is defined as in Moretti[3].

$$\text{IV} = \sum_m S_{mr} \Delta e_m \tag{2}$$

where S_{mr} is the share of age-gender group m (for example men aged 55-60) in region r at the start of the period and Δe_m is the national change in years of education for age group m between the years we study. The weights S_{mr} are estimated based on the whole population, not only the labor force as the age structure of the labor force may be endogenous. Equations (3), (4) and (5) represent the first-stage, the second-stage and the reduced-form regression equations respectively.

$$\Delta \log \hat{e}_r = \alpha_0 + \alpha_1 IV_r \alpha_2 + e_r \tag{3}$$

$$\Delta \log \hat{w}_r = \beta_0 + \beta_1 \Delta \log \hat{e}_r + C_r \beta_2 + \epsilon_r \tag{4}$$

$$\Delta \log \hat{w}_r = \gamma_0 + \gamma_1 \Delta IV_r \gamma_2 + \varepsilon_r \tag{5}$$

where C_r is a vector of control variables and e_r, ϵ_r and ε_r are errors. What we estimate here is how a change in education between two points in time affect wage changes in that same time frame. An important question then is how large this time period should be. If there is a time lag for the effect of education of new cohorts on average wage are visible, then we want the time period to be more than a few years. Moretti (2004) use a ten year period (1980 to 1990) while Ciccone and Peri use a twenty year time period (1960 to 1980, and 1970 to 1990). By gathering data for every ten years from 1960 and forward, we will be able to employ both approaches. This way

[3]Moretti (2006), p 189

we can also get a sense of whether the societal returns to education have changed over time.

4.3 Instrument validity

In order for us to interpret the estimated effect estimated as a causal one, a couple of assumptions need to be met (Angrist and Pischke, 2008). First, the instrument need to have a clear effect on the endogenous variable (average years of education), *instrument relevance*. This is the first stage. If the instrument explain little variation in our endogenous variable (*weak instrument*), then our estimates could be inconsistent or have large standard errors. To check for risk of weak instruments, we compute the F-statistic that correspond to the hypothesis that the coefficient for our instruments are zero in the first-stage regression. Ideally we want the F-statistic to be above 10 as a rule of thumb.

A second assumption for the instrument to be valid is that the instrument affect our outcome variable only trough the endogenous variable. This means that differences in demographic structure of regions only affect future average wages through differences in educational attainment. By adding controls for log-change in region employment and the change in average years of potential experience in the regions we account for the possibility of demographic structure influencing the outcome variable through experience or aggregate scale-effects. It's possible for us to test the likelihood of this exclusion restriction by investigating if our instrument can predict changes during the same time period in other labor market outcomes such as net inflow of immigrants (domestic and international), changes in population size, labor force size and labor force participation rate. If the demographic structure can't explain changes in these variables, then it's reasonable to assume that it only affects wages through education.

Lastly, its important that the instrument don't share a common cause with the outcome Y (*instrument exogeneity* or *instrument independence*). This would mean that the age distribution is a function of anticipated future wages in the region. If our level of analysis would have been cities, this would have been a reason for concern. Moretti, who does just that, lessens the fear of the demographic structure being endogenous by using the demographic structure of 1970 instead of 1980 as instrument for changes in educational attainment during 1980-1990.

5 Discussion

In the proposed study, a constant composition approach is utilised together with instruments for demographic structure is used to estimate the returns to education on an aggregate level. To know if the total societal effect is larger or smaller than the sum of the private returns to education we need to compare the estimates from this study to estimates on the private returns. These have not been calculated here but in a survey of the literature by Björklund (2000), a reasonable estimate is around 4-5%. By making several estimation using data from different time periods we can get a better sense of the true effect as well as find out if there are any trends over time in the social returns to education, for example if they are increasing or decreasing over time.

If externalities arise locally, then the effect of increases in average education in a region on incomes in that region should be the same as the effects on national level. If this assumption don't hold, the effects estimated here may be biased up or downward.

An issue that need further analysis is the fact that spillovers from human capital can still exist in a situation when signaling stand for a major part of the returns to education. The development of models to take such dynamics into account such as Bidner (2010) can possibly help us develop better methods for measuring the relative size of signaling and human capital creation in education.

6 References

Acemoglu, D., Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. NBER macroeconomics annual, 15, 9-59.

Angrist, J. D., Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Arrow, K. J. (1973). Higher education as a filter. Journal of public economics, 2(3), 193-216.

Benhabib, J., Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. Journal of Monetary economics, 34(2), 143-173.

Bidner, C. (2010). A spillover-based theory of credentialism. UNSW Australian School of Business Research Paper, (2010).

Björklund, A. (2000). Education policy and returns to education. Swedish Economic Policy Review, 7(1), 71-106.

Bosworth, B., Collins, S. M. (2003). The empirics of growth: An update. Brookings papers on economic activity, 2003(2), 113-206.

Caplan, B. (2018). The case against education: Why the education system is a waste of time and money. Princeton University Press.

Ciccone, A., Peri, G. (2006). Identifying human-capital externalities: Theory with applications. The Review of Economic Studies, 73(2), 381-412.

Krueger, A. B., Lindahl, M. (2001). Education for growth: why and for whom?. Journal of economic literature, 39(4), 1101-1136.

La, Vincent (2014): Does Schooling Pay? Evidence from China. MPRA working paper

Lange, F., Topel, R. (2006). The social value of education and human capital. Handbook of the Economics of Education, 1, 459-509.

Lucas Jr, R. E. (1988). On the mechanics of economic development. Journal of monetary economics, 22(1), 3-42.

Meghir, C., Palme, M. (1999). Assessing the effect of schooling on earnings using a social experiment. Stockholm School of Economics Working Paper, (313).

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. Journal of econometrics, 121(1-2), 175-212.

Pritchett, L. (2006). Does learning to add up add up? The returns to schooling in

aggregate data. Handbook of the Economics of Education, 1, 635-695.

Psacharopoulos, G., Patrinos*, H. A. (2004). Returns to investment in education: a further update. Education economics, 12(2), 111-134.

Romer, P. M. (1990). Endogenous technological change. Journal of political Economy, 98(5, Part 2), S71-S102.

Spence, A. M. (1973). Job Market Signaling, The Quarterly Journal of Economics, 87(3), 355–74.