DA4 Term Project

Mukhammed Zainidinov

Introduction

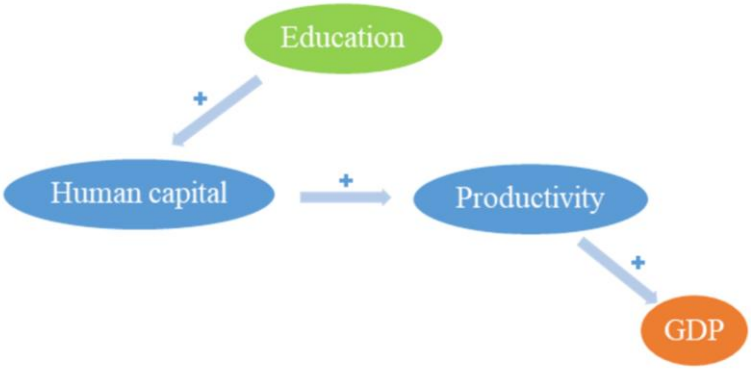
In the modern world, when the employment structure of the population is changing, the share of mental labor in production is increasing, demand for low-skilled labor is decreasing, and the issue of finding additional sources of the country's economic growth is becoming acute. One of such sources is the level of education of the population. Researches by many economists have concluded that education produces indirect effects in the form of externals that may have monetary and non-monetary expressions. Non-monetary externals can take a variety of forms, both higher levels of social tolerance and greater participation in the process of democratization of society, and solidarity at the national and international levels.

The results of the studies show that there is a clear connection between the achievement of higher education by citizens and their participation in public life, socially significant affairs, but the social and cultural benefits of higher education not only contribute to improving the social situation, they also have a direct economic effect.

A study of the social role of higher education and its contribution to economic growth concluded that the latter is generated by an increase in the use of labor and capital resources, as well as an improvement in their quality based on progress in technology and education. In this regard, much attention has recently been paid to the study of human capital as a primary lever of economic growth.

In this paper, I try to analyze the causal effect of education on economic growth using panel data for 136 countries over the period 1990-2017. A mechanism through which education affects economic growth is depicted in Figure 1. Education improves human capital. According to the human capital augmented Cobb-Douglas production function, human capital is one of the factors of production. More skilled human capital is expected to be more productive, hence output increases that is gross domestic product (GDP).

Figure 1. Directed acyclic graph



Data

Outcome variable is GDP per capita PPP, which is a good measurement of income when we compare different economies. It takes into account cost of living in a country. For example, one in Austria cannot purchase for $100 as many things as in Kyrgyzstan. That's why it allows to estimate the purchasing power of money. Data on GDP per capita PPP at constant 2017 international dollar are downloaded from the World Development Indicators database for the period 1960-2020. For the analysis, I will use log values of GDP per capita PPP to estimate percentage differences, i.e. year-to-year growth rate.

Causal variable is education. To measure the level of education quantitatively, I will use data on average number of years of total schooling across all education levels for the population aged over 25 estimated by Lee-Lee (2016), Barro-Lee (2018), and UNDP HDR (2018). The data can be downloaded from the Our World in Data database for the period 1870-2017. Through the analysis, I will use the variable in level.

Figure 2 depicts a strong positive correlation between average years of schooling and income in 2017. The correlation coefficient is 0.809. In 2017, countries where people achieved more years of education tend to have higher income per capita.

Figure 2. Correlation between years of schooling and income

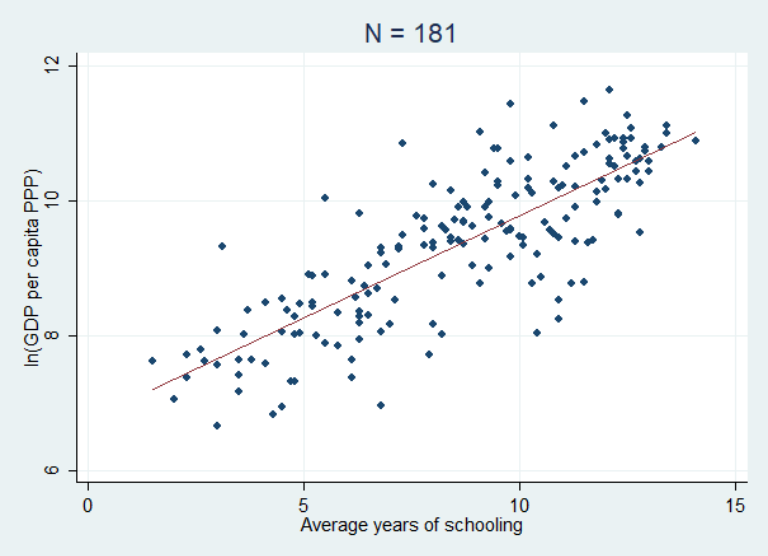
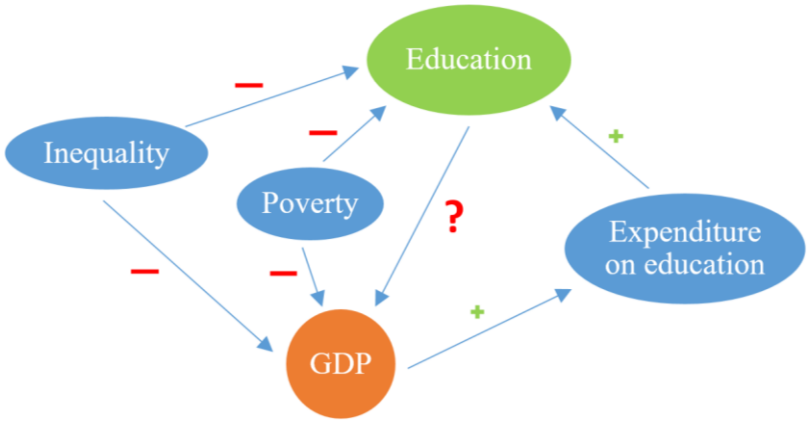


Figure 3 identifies potential confounders graphically. Inequality seems to be a common cause variable. When there is a high inequality, low income households may not afford studying at higher institutions, GDP growth is expected to be slow. Poverty similar to inequality is expected to be a common cause variable. Poor people cannot attend schools and GDP per capita is expected to be low when share of poor people is large. Expenditure on education represents the mechanism of reverse causality. The more income is earned, the more money is expected to be spent on getting education, enhancing school environment, teacher's salary, more scholarships for the poor, etc. This in turn increases average years of schooling.

Figure 3. Directed acyclic graph to visualize potential confounders



Due to lack of data on poverty and expenditure on education for panel data analysis, as a common cause variable I will include only income inequality in level. Data retrieved from the World Inequality Database cover the years between 1990 and 2017. It measures share of bottom 50% for adult population based on pre-tax income. It would be ideal if I used data on GINI coefficient to estimate the inequality. Because of missing values for some periods and some countries, I cannot employ it for the panel data analysis.

Now let's look at relationship of common cause variable with outcome and causal variables. Figure 4 illustrates that equality and income are moderately positively correlated with the coefficient of 0.415. In 2017, countries with more equally distributed income tend to have higher GDP per capita.

Figure 4. Correlation between equality and income

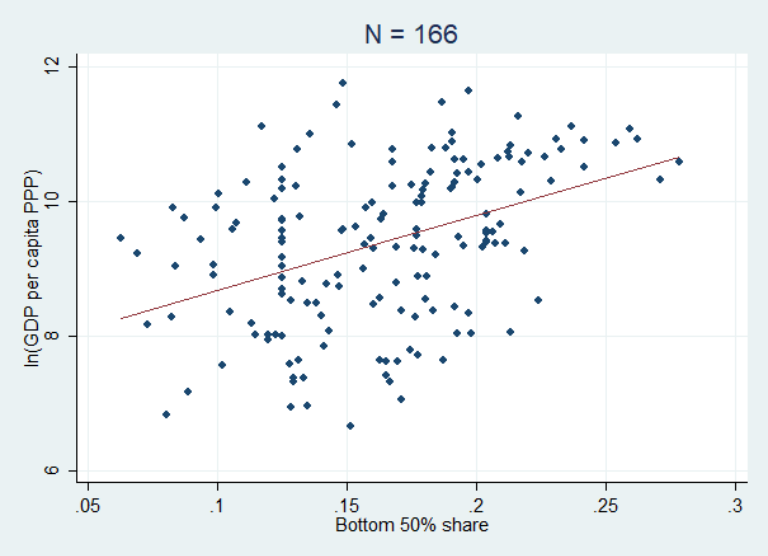
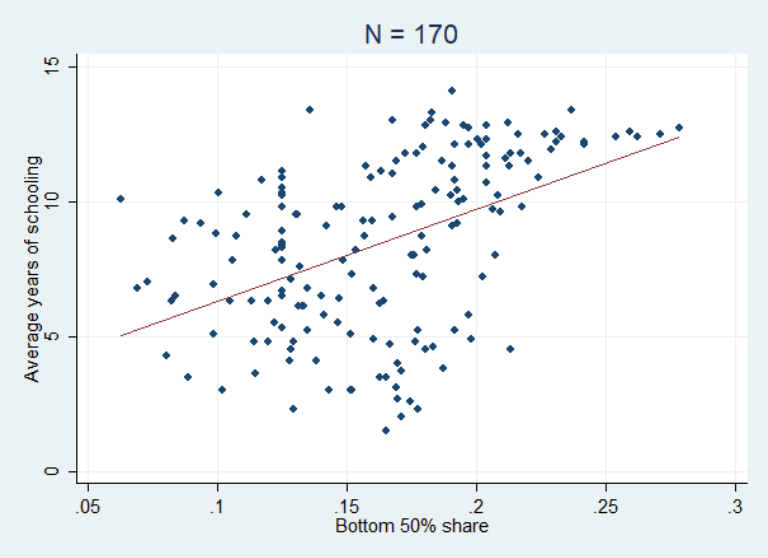


Figure 5 shows a moderate positive correlation between equality and average years of schooling in 2017. The correlation coefficient is 0.468. In countries where income equally distributed, on average, people tend to achieve more years of education.

Figure 5. Correlation between equality and years of schooling



For analysis, I only took the years between 1990 and 2017 because yearly data on education starts only in 1990 and ends in 2017. Number of missing values was around 1000 for each variable. First, I dropped countries where there is no single observation for one of the three variables for the whole period. There were still 144 missing values for GDP per capita and 299 for education variable. Then I decided not to include countries where there are more than or 10 missing values in one of the three variables. In the end there left 136 countries for the period 1990-2017 (28 years). There are 53 missing values in GDP per capita (1.39% of total observations) and 20 missing values in education (0.53% of total observations). Number of observations is 3808, while number of non-missing (in all three variables) observations is 3735. Share of missing values is low, so I can proceed with this data.

Analysis

Through the analysis, I don't weight on population because the variables is expressed in per capita terms. First, I run regression with fixed effect. Column 1 in Table 1 shows statistically significant regression results of simple fixed effect model. Per capita income of a country is, on average, higher by 16.44% compared to its mean when years of education is higher by 1 year compared to its mean. When I add a common cause variable into regression (column 1 in Table 1), the coefficient does not change much. However, these numbers cannot tell us a causal effect of education on per capita income because there is a trend in time series and the results can be spurous.

To get unbiased estimation, variables should be detrended. For this purpose I include dummy variables for each year. When I add year dummies into regression, regression coefficients are not statistically different from 0 in both models with and without confounder (columns 3 and 4 in Table 1). This means that there is no statistically significant effect of education on GDP per capita.

Then I tried to estimate the causal effect using first difference model. I tried various regression models with and without confounder, with and without lags, with and without leads, with and without aggregate trend, with and without country linier trend. All regression results were statistically insignificant that's why I decided not to put the results here (code is available).

Table 1. The effect of education on income. Fixed effect regressions

VARIABLES	(1) lngdppc	(2) lngdppc	(3) lngdppc	(4) lngdppc
Average years of schooling	0.1644**	0.1640**	0.0042	0.0011
	(0.014)	(0.014)	(0.022)	(0.022)
Bottom 50% share		-1.0465		-1.5396*
		(0.781)		(0.662)
Constant	7.9295**	8.1033**	8.9239**	9.2010**
	(0.104)	(0.144)	(0.138)	(0.182)
Year dummies	No	No	Yes	Yes
Observations	3,735	3,735	3,735	3,735
R-squared	0.427	0.431	0.544	0.553
Number of c	136	136	136	136

Conclusion

By running different regression models, I couldn't find statistically significant effect of average years of education on GDP per capita. It can be because of that the models do not capture all potential confounders, or average years of schooling might not be a good measure of level of education, or income might not be perfectly measured by GDP per capita PPP.

References

Lee, Jong-Wha and Hanol Lee, 2016, "Human Capital in the Long Run," Journal of Development Economics, vol. 122, pp. 147-169.

Barro, Robert and Jong-Wha Lee, 2013, "A New Data Set of Educational Attainment in the World, 1950-2010." Journal of Development Economics, vol 104, pp.184-198.

United Nations Development Programme, Human Development Report (2018 Statistical Update).