SCHOOL OF ECONOMICS AND MANAGEMENT

To what extent may a decline in schooling caused by the Covid-19 pandemic affect Kenya's GDP?

A Scenario Modeling Approach using Multiple Linear Regression

By

Fanny af Petersens

Bachelor's thesis, June 2021

Bachelor's Programme in Economics

School of Economics and Management

Department of Economics

Lund University

Supervised by Gunes Gokmen

Abstract

This study focuses on the relationship between education and economic growth, specifically the effect a potential decline in schooling caused by the ongoing Covid-19 pandemic would have on the economic growth of a country. Using an inductive approach, the study focuses on the Sub-Saharan African countries, especially Kenya, since there is anecdotal evidence that a significant number of students have not returned to school after the closures caused by the Covid-19 pandemic. The study models a hypothetical scenario with a drop in schooling of one year and analyzes how it would differ from a base scenario without any drop in education. As education is recognized as an important factor of economic growth, the potential impact of the pandemic on educational attainment is an important subject to analyze. Two regressions based on the Solow model with human capital were used, the first for predicting human capital with the independent variables of population and average years of schooling. The result was used as a basis for the second regression to predict future GDP levels. The results suggest that a one-year drop in average years of schooling would cause a long-run expanding negative effect on GDP. This is cause for concern and increased focus on education, especially in emerging economies such as many of the Sub-Saharan countries. The study further provides thoughts for how future research regarding this subject could be conducted, and why it is needed.

Keywords: Schooling, Scenario Modeling, Human Capital, GDP, The Solow Model, Kenya, Sub-Saharan Africa, Economic Growth, Covid-19

Acknowledgements

I wish to express my deepest gratitude to my supervisor Gunes Gokmen for his guidance throughout the process of writing this bachelor’s thesis. His constructive feedback has helped immensely, while allowing me to choose the direction of the study, for it to remain completely my own work. Finally, I thank my parents for their continuous emotional support and encouragement throughout my study.

Table of Contents

List of Tables

List of acronyms and abbreviations

GDP	Gross Domestic Product
PWT	Penn World Table
SSA	Sub-Saharan Africa
UIS	UNESCO Institute for Statistics
WEF	World Economic Forum

1. Introduction

Economic growth, usually defined as the level and growth of a country’s gross domestic product (GDP), has long been a thoroughly discussed topic, specifically concerning which variables generate growth. Solow (1956) produced a standard neoclassical growth model of Cobb-Douglas type, dependent on capital accumulation as well as labour growth. This model has been augmented several times, and one of the most well-known is by Mankiw, Romer and Weil (1992). Their version included human capital, in order to decrease the coefficient bias in the other explanatory variables. Human capital is defined as the knowledge, skills and abilities that contribute to the well being of individuals and countries. Education is a vital basis of human capital, since people with higher education tend to have higher incomes, thereby contributing to a higher consumption that in turn benefits the country’s total level of GDP. (Keeley, 2007, p.3) The significance of education in generating economic growth has also been discussed by Elizabeth King (2011), the director for education of the World Bank, who recognizes education in emerging economies as a prerequisite for sustaining a high level of economic growth. Hence, studying education’s role in economic growth is fundamental to understanding the effect it has on GDP.

During the ongoing Covid-19 pandemic, schools around the world have been kept closed for a long period of time. This may have a negative impact on future economic growth, especially in emerging economies that lack sufficient technological infrastructure to continue education online. This temporary negative effect may become even more serious in many Sub-Saharan African countries where the school system is subject to fees. Due to the economic impact of the pandemic, many families may not be able to afford sending their children back to school when they reopen (Mlaba, 2021). Even in countries that provide free primary and secondary education, the cost of schooling is still significant, including direct costs for school uniforms, books, food and other essentials. In addition, many families depend on their children to contribute to the family's direct livelihood, even if their future earning potential would be negatively affected by them being denied education. In 2009, an estimated 34.4% of Kenya’s children between the ages of 7 and 14 were engaged in work (World Bank, 2021b), although due to the extensive share of the population living in rural areas, the actual statistics of child-labour in the farming sector could be higher.

Kenya was the last East-African country to fully reopen schools in January, after being closed for nine months. However, reports showed that thousands of students did not return back to school. (Mukoya, 2021). There is anecdotal evidence that this drop in return to schooling could be explained by child marriages, parents' loss of income as well as the children needing to work in order to help provide for their families (Nation, 2021).

This paper seeks to examine the effect a drop in education could have on the future economic growth of Kenya by means of scenario modeling through an inductive approach. The research question is composed as "How will a potential drop in average years of schooling by one year affect the economic growth?". The hypothesis is therefore proposed as H_1: A drop in schooling would cause a significant disruption in the future economic growth of the country. If this is confirmed it would encourage policy changes for the national government, as well as for external aid organizations to target aid packages towards education. However, if the hypothesis is proven to be incorrect, that could be an indication that education is an overstated factor of economic growth in developing countries, and other variables could be prioritized in order to achieve a higher growth.

The results suggest that a one-year drop in average years of schooling causes a negative effect on the growth rate of GDP. The scenario with a simulated decline in schooling resulted in a lower growth rate than the base scenario without the decline. This indicates that a drop in schooling will cause a long-lasting negative effect on GDP, which is cause for concern.

The limitations of the study are many, principally due to a severe lack of data. However, the purpose of the study is not to accurately predict the future GDP of Kenya, but rather to analyze the effect a decline in schooling could have on the country's output. The ongoing Covid-19 pandemic is the causal factor underlying the simulated decline in school attendance in this paper, and emphasizing the word "ongoing", the pandemic's final effects on the education system have not yet been identified. Given the uncertainty, a scenario modeling approach is preferable to any attempt at forecasting.

The overall structure of the study takes the form of six chapters, following this introduction with a historical context of Kenya's education system as well as the economy in chapter two. Chapter three includes a presentation of the theoretical framework as well as previous research. This is followed by the empirical strategy in chapter four, explaining the methodology as well as the different variables in the analysis. In chapter five the study's results are presented and explained, while chapter six forms the conclusion.

2. Historical context

Compulsory school attendance in Kenya was extended from 8 to 12 years in 2013. Free primary education was introduced in 2003, which resulted in an increase in enrollment in primary school of more than 900 000 students from the year before. (UIS, 2021). In 2017, the World Economic Forum ranked Kenya as the highest quality education system on the continent (Leopold, 2017). Although significant progress has been made in the last two decades, there still is a lot of room for improvement. The net attendance rate of students in upper secondary education decreased from 81.5% in 2008 to 79.6% in 2014, and the literacy rate of the population aged 25-64 was at a level of 79.6% in 2018 (World Bank, 2021a).

Kenya's economy is heavily dependent on the agricultural sector, which accounts for 70% of the rural workforce and 40% of the overall workforce, as well as about 25% of annual GDP (International Trade Administration, 2020). 72% of the population live in rural areas (UIS, 2021), where it is not uncommon for children to work at neighbouring or their own family farms (Murray et al., 2010, p.23). Due to a lack of reporting, the data on child labour is close to impossible to fully apprehend, but it can be assumed that child labour is fairly common. As mentioned in the previous chapter, an estimated one third of all children between 7 and 14 in Kenya were engaged in labour in 2009 (World Bank, 2021b).

The Covid-19 pandemic has had a significant negative impact on economies worldwide, causing layoffs in many businesses. The unemployment rate in Kenya tripled due to the pandemic (World Bank, 2020), which helps explain the reports and anecdotal evidence that many children may have to forego schooling in order to help support their families.

3. Theoretical Framework and Previous Research

3.1 Theoretical Framework

The basis for this paper's analysis will be the Solow (1956) model for economic growth. The neoclassical growth model was developed by Robert Solow in 1956, and to this day constitutes a core model in the field. The model argues that output growth depends on two variables: capital accumulation and population or labour force growth. The standard production function of the neoclassical growth model is assumed to be characterized by constant returns to scale, meaning that if any factor grows by 1%, the overall output will also increase by 1%. The function appears as

$$Y = K^{\alpha} L^{1-\alpha} \tag{3.1}$$

where *Y* represents gross domestic product, *K* is the stock of capital and *L* is the labour force.

The labour force is assumed to grow exogenously at rate *n*. It is assumed that a constant fraction of output is invested, which constitutes the savings rate *s*. Furthermore, *k* is defined as the capital per unit of labour (K/L), and similarly *y* being the output per unit of labour (Y/L).

The Solow model is based on the assumption that countries with the same population growth, savings rate and capital depreciation will converge to the same steady state level. However, the model is fairly outdated in many aspects, and has attracted some critique based on the model not including other important factors, as well as possibly being better suited for explaining growth in a high-income country rather than a low-income one (Todaro & Smith, 2012, p.146). These concerns have been raised due to several obstacles in low-income countries that can hinder growth which are not explained by the three variables in the model, for example corruption, land resources and economic policies, among others.

The basic Solow model has been augmented many times, with several well-known variations including technological progress as a residual factor, determined exogenously. However, due to insufficient data for the sampled countries, this paper will disregard technological progress as an explanatory variable.

Mankiw, Romer and Weil (1992) augmented the Solow model with the component human capital. They argued that the basic model correctly showed predictions of the direction in output growth, but that the predicted effect of savings and population growth on output were too large. Including human capital in the model reduces the bias towards the coefficients of savings rates and population growth.

The production function, adding human capital is proposed as

$$Y(t) = K^{\alpha} H^{1-\alpha} \tag{3.2}$$

with human capital being dependent on factors such as years of education (u) and the quality of said education (ψ), as well as the population, as presented in equation (3.3).

$$H = e^{\psi u} L \tag{3.3}$$

Capital is accumulated by investing in production. Alterations in capital accumulation are distinguished as the savings rate (s_K) of total output, subtracted by the rate of depreciation (δ) on capital as shown in equation (3.4).

$$\dot{K} = s_K Y - \delta K \tag{3.4}$$

The calculations using the Solow model are not solely focused on the levels of GDP, but also the growth rate. The function for calculating the average growth rate for a period of time is proposed as

$$Average\ growth = \left(\frac{value\ of\ the\ last\ year\ in\ the\ period}{value\ of\ the\ first\ year\ in\ the\ period}\right)^{1/number\ of\ years} - 1 \tag{3.5}$$

where the rate is tracked over time as an indicator of the direction of a country’s economy.

3.2 Previous Research

There seems to be general agreement that there is a positive correlation between education and GDP growth. However, the precise effect of education on growth is debated, as opinions on which variables should be used to identify education vary.

Most SSA countries, with the exception of conflict areas, have experienced a long period of positive development in terms of schooling (World Bank, 2021a). A major drop in schooling is unprecedented, and therefore research on the effects of such a decline has not been a priority.

A study conducted on an online math platform in the United States (Chetty et al., 2020) concluded that children in high-income areas temporarily showed a loss of learning after converting to online based teaching, but soon recovered. Meanwhile, children in lower-income areas remained below baseline levels throughout the school year. Although the study only includes a small proportion of overall learning, it raises serious concerns about a potential increase in inequality in the long term, which would be especially damaging for children in lower-income areas. Compared to the United States, with 1.2% of the population living in extreme poverty ($1.90 a day, 2011 PPP) in 2015, Kenya's corresponding number at 37.1% in the same year indicates that a significant drop in learning in the African country could have highly alarming negative consequences (World Bank, 2021b).

3.2.1 Ebola and Covid-19

Certain similarities, albeit on a significantly larger scale, exist between the impact of the Ebola epidemic in western SSA in 2014 and the current Covid-19 pandemic. The countries that were extensively affected by the epidemic were primarily Sierra Leone, Liberia and Guinea, and projections have shown that the Ebola epidemic caused an estimated loss of $2.2 billion in the GDP of the three countries (CDC, 2019).

The schools in Sierra Leone, Liberia and Guinea were closed for six to eight months during the Ebola epidemic, which caused a severe but temporary disruption in learning in the West-African countries (Chavez, 2015). The West-African countries have previously suffered from conflict and instability, and the closure of schools caused by the epidemic severely impacted the educational system that likely will have effects on the countries' GDP. However, since economic growth is analyzed over a longer period of time, and the effect of temporary school closures on human capital will be shown primarily when the students that were affected enter the labour force, the final impact on GDP has not yet been demonstrated.

4. Empirical strategy

4.1 Scenario Modeling

This paper will be conducted as a scenario-based modeling of a single-year drop in schooling in Kenya, primarily due to school closures as a result of the Covid-19 pandemic, assuming that all students go back to school when reopening. This scenario (Scenario 1) will be compared with the base scenario of the pandemic not occurring, and no drop in school attendance (Scenario 0).

Scenario modeling is a way of analyzing possible scenarios that could take place in the future and forecasting different possible outcomes. This type of modeling is typically used in situations with a high degree of uncertainty, and where key data are difficult to collect for reasonably accurate projections. Key benefits of scenario analysis include being able to analyze many possible futures, to proactively plan for all or a selected few, and to make investments to prevent a worst-case outcome. The main difficulties of scenario analysis are to identify the relevant possible futures, and to maintain flexibility in the face of unforeseen outcomes. (Corporate Finance Institute, 2021).

4.2 Data collection

The dataset that has been used as foundation for the analysis has been a combination of Penn World Table and World Bank Databank for the years 2000 to 2019. These can be expected to be reliable and unbiased, although errors may appear due to difficulties in collecting the data. Many of the countries in SSA have difficulties regarding their collection of data, as a large portion of the population live either in remote locations or in urban clusters with poor living conditions. The collection of data is often logistically difficult and expensive in these countries, which accounts for its lack of availability (Runde, 2017).

4.2.1 Sample

The sample for this study is composed of 29 SSA countries. Many of these countries suffer from similar difficulties affecting economic growth, such as a high level of corruption, conflicts between different ethnic groups, and weak systems for registering and protecting property rights. Most SSA countries also share a background as colonies of European powers, and of having achieved independence in the latter half of the 20th century. North Africa and SSA are generally separated from each other when discussing Africa, since the countries in SSA are very different from their northern counterparts. North Africa is comparably more developed and primarily Arab, having more similarities with the Middle East than the Sub-Saharan part of the continent (BBC, 2004). Because of these differences between the two regions, only countries in SSA have been chosen. Some SSA countries have been excluded from the analysis due to lack of data in the exogenous variables, thereby misrepresenting the extent of their impact on the endogenous variables of human capital and GDP. This sample has been composed in order to correctly estimate the coefficients of the various explanatory variables, for the purpose of simulating the effect on GDP for Kenya.

4.2.2 Dependent variable

The dependent variable for this study is the GDP, Gross Domestic Product. The GDP is often used as a reference point for the health of economies, as growth in GDP generally indicates that workers and businesses are better off (Callen, 2020). It is common to calculate the GDP per capita, meaning per citizen, in order to accurately compare countries. This is often preferred as a country with a larger population presumably has a higher GDP than a country with fewer citizens, although the average citizen is not necessarily “richer”.

In this paper the term “output” will often be used as a reference to the GDP, since one of the approaches to measuring GDP is the total production, or output, of a country.

4.2.3 Research variables

Time spent in education is the variable that will be researched for the education’s effect on the dependent variable GDP. Average years of schooling will represent this variable, as many countries have a set duration of compulsory schooling, but the actual time students spend in education does not necessarily correspond with the compulsory education. The data on average years of schooling is important for this study, since analyzing how GDP would be affected by a one-year decrease in education will be the main focus point of Scenario 1.

The quality of education, which is commonly used as a complement to the variable time spent in schooling, could have been relevant to the analysis of education’s effect on GDP. However, as the quality of education is most commonly determined by harmonized test scores, and data on these are non-existent for the countries in the sample, the variable has been removed.

The variable for time spent in education will together with the population constitute the exogenous variables for an endogenous variable of human capital. When predicting GDP, firstly a forecast for expected human capital will be regressed, which in turn will lay the foundation for calculating expected GDP.

4.2.4 Control variables

The basic Solow model is dependent on labour force and capital accumulation. Again, because of insufficient data, population is used in this study as a substitution for the labour force, as they presumably grow at the same rate. The model states that key components of economic growth are savings and investment. An increase in these factors raises the capital stock, which in turn raises employment and national output. The model argues that capital accumulation is composed of the factors of savings and depreciation rates. However, the full measure of capital stock has been used as a variable, as the study does not consider any drastic changes in the savings or depreciation rates.

4.3 Multiple linear regression analysis

A version of the Solow model augmented with human capital will be used for the analysis, where human capital will depend on the variable of average years of schooling, as well as the population.

The analysis will be conducted using a multiple linear regression of the natural logarithm of the variables. The logarithm is used since the parameters in the Cobb-Douglas type function in the Solow-model will be estimated as coefficients when the variables are logged. This is explained as

$$log(X^{\alpha}) = \alpha * log(X) \tag{4.1}$$

where α is the coefficient of the logged variable X. However, the variable of average years of schooling is not logged, as it is expressed in the manner of e^u, and when logging this exponential, the equation will present as

$$log(e^{u}) = u \tag{4.2}$$

and therefore, the original values of this variable will be used for the analysis.

The regression equations are thereby produced as following

$$ln(human\ capital) = \alpha + \beta_1 ln(population) + \beta_2(average\ years\ of\ schooling) \qquad (4.3)$$

$$ln(GDP) = \alpha + \beta_1 ln(capital\ stock) + \beta_2 ln(human\ capital) \qquad (4.4)$$

4.4 Panel Data and Heterogeneity

Panel data, or longitudinal data, is cross-sectional data collected over a period of time. Panel data can be characterized as balanced or unbalanced, depending on the amount of data distributed over individuals or time periods. A balanced collection of panel data contains the same amount of data for all individuals over the same time periods, whereas unbalanced datasets have missing values for some of the individuals at some time observations.

Panel data methods can be divided into two categories, homogeneous and heterogeneous. Homogeneous models assume that the parameters are common across the countries, where heterogeneous models allow for parameters to differ across countries, as is preferable in this study. Heterogeneous models can be used with either random or fixed effects.

4.5 Multicollinearity

Multicollinearity, or simply collinearity, occurs when two or more independent variables in a regression model are correlated. This is a problem since the independent variables should be just that, independent. The intention when analyzing a regression model is to calculate the relationship between the dependent variable and each independent variable. The presented coefficients in a regression model explain the change in the dependent variable caused by a unit change in an independent variable, holding all other variables constant. If two independent variables are correlated this is not possible, as a shift in one variable will then entail a shift in the correlated variable. Multicollinearity weakens the statistical power of the model, and the p-values cannot be trusted as significant for the regression.

One way of testing for multicollinearity is using the Variance Inflation Factor (VIF), which identifies the correlation, as well the strength of it, between independent variables. No formal cutoff value exists for determining a VIF-value as too high, although typically cutoff points are suggested at 5 or 10 (Craney & Surles, 2002).

A VIF-value between 1-5 indicates moderate correlation between independent variables, but it is typically not critical enough to require making any changes. However, a value greater than 5 indicates a higher level of correlation between independent variables, and the coefficient estimates and p-values are likely unreliable. This would be a problem that needs to be addressed.

The variance inflation factors were computed for each variable in the regression models and are presented in Table 1. Looking at the variables there seems to be no issue with high VIF, since all variables show a value below 5, indicating no issue with multicollinearity.

Table 1. Variance inflation factor

	VIF	1/VIF
ln(Population)	1.000	0.999873
Mean education	1.000	0.999873
Mean VIF	1.000	.
ln(Capital stock)	1.955	.511487
ln(Human capital)	1.955	.511487
Mean VIF	1.955	.

4.6 Delimitations

The collected sample for this study consists exclusively of countries in Sub-Saharan Africa, which share several key historical and current socio-economic characteristics. Including countries from other regions could distort the results, since explanatory factors for changes in GDP in countries from other regions could be inapplicable for Kenya.

The conducted study has not considered the direct positive effects on productivity when children work instead of participating in education, as the statistics on these effects are next to impossible to analyze. Furthermore, the chosen variables are primarily collected based on those appearing in the Solow model augmented with human capital. Other factors that arguably affect GDP, for example corruption, land resources and net exports, have not been considered.

The actual effect of closing schools for one year will not necessarily correspond to the theoretical scenario developed in this paper. If the current generation in education loses one year of schooling, the overall average years of schooling in the society will not decrease by one year. The population will still include those who already have completed their full education. Thereby, only the average schooling of the part of the population still in school would decrease. Additionally, a drop in education would presumably not display an immediate effect on the GDP, as those affected are students currently enrolled in education, and do not yet contribute to the country's output as they are not yet engaged in labour. However, in order to facilitate the analysis and clarify the results, a simulated full one-year drop in the year of 2020 will be used for the study.

4.6.1 Limitations of the model

The most significant limitation of the produced model for this study is the lack of sufficient data. Furthermore, the model may be, as previously stated, better suited for analysis on high-income countries. Some of the variables needed for an accurate analysis are non-existent for the countries sampled for this study.

The model is used to explain long-term growth in a country, and forecasting the output for a country for a longer period of time can be highly inaccurate, as it is impossible to know which factors will grow at which rate. For this study, the forecasting for the exogenous variables have been estimated with the same growth rate as previous years.

5. Results

In Table 2 produced below the descriptive statistics for the variables used in the two regressions are presented.

Table 2. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP (in mil. 2017 USD)	586	65270.955	129526.6	2639.275	734094.38
Capital stock (in mil. 2017 USD)	586	184267.77	422926.07	446.213	2897706.3
Human capital (index)	586	1.55	.402	1.049	2.908
Population (in millions)	586	15.123	13.049	2.116	58.558
Mean education (years)	620	4.622	2.206	.832	10.03

The observations for average years of schooling *(Mean education)* differ from the other variables. This is due to the fact that there is unobserved data for certain time periods for some countries. However, this should not unduly affect the results, as these unobserved values are set as missing, and are therefore not included in the data analysis.

Since GDP is dependent on capital stock and human capital, with the latter in turn being dependent on population and average years of schooling, the first step was forecasting the exogenous variables in order to be able to predict the endogenous ones. The time period to be predicted was 2020 through 2030. After forecasting the exogenous variables, a Hausman-test had to be executed (Table 3) in order to determine if the panel data was to be regressed using random or fixed effects.

Table 3. Hausman (1978) specification test

	ln(GDP)	ln(Human capital)
	Coef.	Coef.
Chi-square test value	7.417	11.75
P-value	.025	.003

The Hausman-test analyzed the variables, to check for systematic differences in the coefficients. The null hypothesis was that the differences are not systematic, and a p-value of less than 0.05 rejects this hypothesis. The test indicated that fixed effects should be used for further analysis in order to achieve accurate results, as the differences in the coefficients are systematic.

After determining that fixed effects should be used for the analysis, the two separate regressions were drawn (Table 4 & 5). The variable for population showed significance on a 0.01 level, which is equivalent to being significant to a degree of 99%. The problem here is that the variable for average years of schooling seems to not be significant, but non-significance at a highest of 0.1 level does not render the model useless. As the variable is crucial for the analysis, the choice was made to prioritize the practical significance over the statistical. The non-significance could have been due to the missing values in the data, as well as no other variables being included to describe the human capital, apart from the population. The coefficients indicate the effect on the dependent variable that the independent variables have, for example an increase of 1 in the logarithm of population would increase the logarithm of human capital by 0.284. Apart from the education variable not being significant, the variables appear as expected, with positive coefficients towards the dependent variable. The value of the constant implies that if both the other variables are at zero, the value of the logged human capital would be at -0.293.

Table 4. Regression results for Human Capital

Dependent variable: Logarithm of Human Capital

ln(Human capital)	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln(Population)	.284	.013	21.08	0	.258	.311	***
Mean education	.004	.005	0.91	.365	-.005	.014	
Constant	-.293	.031	-9.48	0	-.354	-.233	***
Mean dependent var	0.383			SD dependent var	0.229		
R-squared	0.489			Number of obs	580.000		
F-test	262.505			Prob > F	0.000		
Akaike crit. (AIC)	-1512.042			Bayesian crit. (BIC)	-1498.953		

**** p<.01, ** p<.05, * p<.1*

The results for the second regression showed that both the independent variables were significant at a 0.01 level as well as having a positive relationship with the dependent variable (Table 5). The variable *ln(pr. Human capital)* constitutes the predicted values of the logged human capital that were calculated by the previous regression (Table 4) and then predicted to determine the predicted GDP. As the scenarios were constructed with the hypothetical shift in education, the predicted human capital from these calculations were used as a basis for the prediction of GDP, as the actual human capital was not viable in light of the simulated decline in the alternate scenario, of an independent variable in the first regression (Table 4).

Table 5. Regression results for GDP

Dependent variable: Logarithm of Gross Domestic Product

ln(GDP)	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln(Capital stock)	.243	.032	7.54	0	.179	.306	***
ln(pr. Human capital)	2.673	.21	12.71	0	2.26	3.086	***
Constant	6.371	.285	22.37	0	5.811	6.93	***
Mean dependent var		9.954		SD dependent var		1.304	
R-squared		0.598		Number of obs		580.000	
F-test		407.749		Prob > F		0.000	
Akaike crit. (AIC)		-160.054		Bayesian crit. (BIC)		-146.965	

**** p<.01, ** p<.05, * p<.1*

The value for *R-squared* displays how much of the variation in the dependent variable that is explained by the variation in the independent variables. *R-squared* in the regression results show that 48.9% and 59.8%, respectively, the variation in the dependent variables are explained by the variation in the independent ones.

After estimating the coefficients of the explanatory variables for the first regression (Table 4), a prediction for human capital was generated using an *xb* prediction. This functions by recalling the previous estimated coefficient *b* for variable *x*, and combining the information with the data currently in memory. Predicting the j_{th} observation for a variable *y* is defined as

$$\widehat{y_j} = x_j b + offset_j \quad (5.1)$$

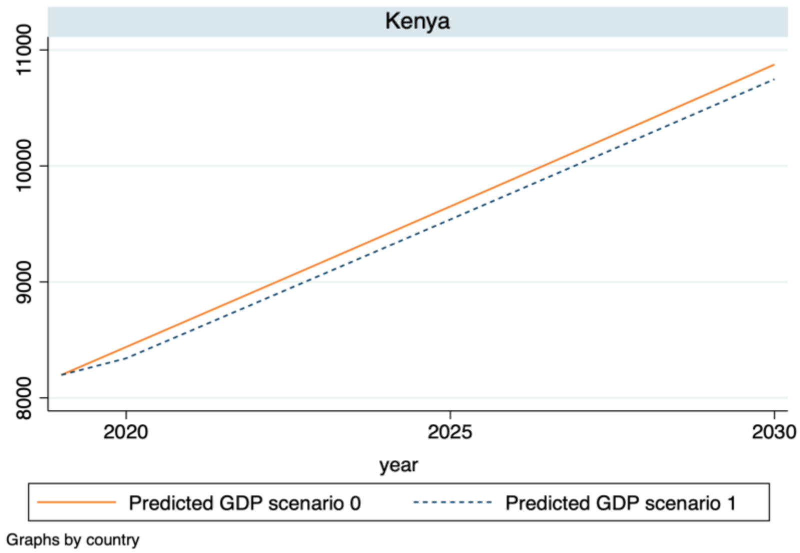
where the offset variable represents the size of each observational unit, constrained to a coefficient of 1.

The linear prediction from the fitted model was calculated by using the estimated coefficients. Firstly, a base scenario (Scenario 0) was predicted in order to compare the eventual difference to the scenario with a simulated drop in schooling (Scenario 1). After generating the predicted values for human capital, GDP was predicted using the same method. For Scenario 1, a new variable was constructed for mean education, decreasing the values of this variable by one unit, in this case representing one year. This new variable for mean education *(mean education_1)* replaced the original variable for education when predicting new values for the human capital, resulting in a decrease of human capital compared to the base scenario. A new prediction for the GDP was generated, and consequently, the results differed from the scenario without a decrease in the variable for education.

The final results in predicting the GDP showed that a drop in schooling has a direct and increasing negative effect on the output (Figure 1). The predicted output for the base scenario (Scenario 0) was higher than the alternate scenario after a drop in education in 2020 (Scenario 1).

Figure 1. Predicted GDP

Predicting GDP in millions 2017 USD



The predicted GDP values for 2018 and 2019 differ somewhat from the actual GDP values. The results can be explained by the fact that the predictions were based on linear forecasts of data for the independent variables for the years 2000 to 2019. A country's GDP is highly unlikely to appear as a linear function. However, it is necessary to again emphasize that the purpose of this analysis is not to predict actual GDP, but rather to analyze the effect a potential drop in schooling could have on long-term economic growth. In Figure 1, it appears that the simulated drop in schooling resulted in an immediate decrease of predicted GDP for scenario 1. This may not be representative of how the GDP would be immediately affected, as the effects of education presumably would not be displayed on the output until the affected students emerge onto the labour market, which is explained in detail in *4.6 Delimitations*.

Table 6. Predicted average growth

Predicted time period: 2020-2030

Scenario 0	2.6022 %
Scenario 1	2.49308 %

As presented in Table 6, the predicted growth differs between the scenarios, meaning that in theory Scenario 0 and 1 the GDP would grow by 2.60% and 2.49% per year, respectively. This indicates, as previously shown in Figure 1, that a drop in schooling will cause a long-lasting negative effect on GDP.

6. Conclusion

As stated in the previous chapter, the results showed that a drop in schooling causes a negative effect on the GDP, with an increasing negative effect in the long run.

However, the difference between the two scenarios studied in this paper (Scenario 0: no pandemic and correspondingly no drop in education; Scenario 1: a one-year drop in education) was not as extensive as expected. This is presumably due to the fact that changes in education, both in terms of length of participation in formal education and the quality of the education, take some time to affect the human capital.

The implied correlation between a drop in education and lower expected future GDP indicates that it is important for national governments to address the issue of a potential drop in school attendance as a result of the Covid-19 pandemic, or other similarly disrupting events. The results of the analysis also imply that aid organizations should reconsider their priorities to focus added resources on education when recipient nations are subject to major economic disruptions. As stated in the chapter on historical context in this paper, children often need to contribute to the family income in developing countries, and many families are not able to afford the direct costs of schooling. Getting children back to school and ensuring regular attendance by means of economic support for the direct and indirect costs of schooling may be an important part of any long-term recovery program for developing countries faced with epidemics or similar disrupting events.

The increasingly negative effects of a drop in education on GDP over time implies that the problems could be longer lasting, and further research is merited due to the apparent decrease in human capital.

The distance between the predicted values in the two scenarios appears to be increasing throughout the period. This effect may not be surprising, but it indicates that further research could be useful, with additional data collected in the sampled countries.

Further studies could include the variables of technological progress and quality of education, provided that sufficient data on these variables can be collected.

In addition, a scenario including a long-term decrease in school attendance would be interesting to analyze, since the economic impact of the current Covid-19 pandemic can reasonably be expected to last for several years, with an increase in unemployment and harsher living conditions for low-income households. This study did not include this expanded scenario, primarily due to insufficient data on attendance rates. However, the effect on GDP would most likely be much larger than the one caused by a one-year drop in schooling as presented in this study.

7. References

BBC. (2004). *How African is North Africa?* [online] BBC. Available at: http://news.bbc.co.uk/2/hi/africa/3421527.stm [Accessed 17 May 2021].

Callen, T. (2020). *Gross Domestic Product: An Economy's All*. [online] International Monetary Fund. Available at: https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm [Accessed 18 May 2021].

Centers for Disease Control and Prevention. (2019). *Cost of the Ebola Epidemic*. [online] Available at: https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cost-of-ebola.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fvhf%2Febola%2Foutbreaks%2F2014-west-africa%2Fcost-of-ebola.html [Accessed 25 May 2021]

Chavez, D. (2015). *Back to School After the Ebola Outbreak*. [online] World Bank. Available at: https://www.worldbank.org/en/news/feature/2015/05/01/back-to-school-after-ebola-outbreak [Accessed 10 May 2021].

Chetty, R., Friedman, J.N., Hendren, N., Stepner, M. & The Opportunity Insights Team. (2020). *The Economic Impact of Covid-19: Evidence from a new public database built using private sector data*, NBER working paper series, Cambridge, Mass. National Bureau of Statistics.

Corporate Finance Institute. (2021). *What is Scenario Analysis?*. [online] Available at: https://corporatefinanceinstitute.com/resources/knowledge/modeling/scenario-analysis/ [Accessed 25 May 2021]

Craney, T. A. & Surles, J. G. (2002). *Model-Dependent Variance Inflation Factor Cutoff Values*, Quality Engineering, vol. 14, no. 3, p. 391-403

International Trade Administration. (2020). *Kenya: Country Commercial Guide.* [online] Available at: https://www.trade.gov/knowledge-product/kenya-agriculture#:~:text=The%20country's%20major%20agricultural%20exports,of%20national%20commercial%20agricultural%20output. [Accessed 10 May 2021]

Keeley, B. (2007). *Human Capital: How what you know shapes your life*, OECD Insights, OECD Publishing, Paris

King, E. (2011). *Education is Fundamental to Development and Growth*. [online] World Bank Blogs. Available at: https://blogs.worldbank.org/education/education-is-fundamental-to-development-and-growth [Accessed 7 May 2021].

Leopold, T., (2017). *Here's what's happening to African jobs - in 5 charts*. [online] World Economic Forum. Available at: https://www.weforum.org/agenda/2017/05/africa-jobs-skills-in-charts/ [Accessed 10 May 2021].

Mankiw, N.G., Romer, D. & Weil, D.N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, vol.107, no.2, p.407-437.

Mlaba, K. (2021). *Thousands of Children in Kenya Failed to Return to School This Week*. [online] Global Citizen. Available at: https://www.globalcitizen.org/en/content/children-missing-out-on-school-kenya-covid/ [Accessed 7 May 2021].

Mukoya, T. (2021). *Parents worry as crowded Kenyan schools reopen after coronavirus shutdown*. [online] Reuters. Available at: https://www.reuters.com/article/us-health-coronavirus-schools-kenya-idUSKBN29912E [Accessed 10 May 2021].

Murray, U., Hurst, P. & de Lang, A. (2010). *Child labour prevention in agriculture: Junior farmer field and life school - Facilitator's guide*, Food and Agriculture Organization of the United Nations, Rome: FAO.

Nation. (2021). *Thousands of learners fail to report back as schools open*. [online] Available at: https://nation.africa/kenya/news/education/thousands-of-learners-fail-to-report-back-as-schools-open-3247308 [Accessed 10 May 2021].

Penn World Table. (2021). *PWT 10.0 - Penn World Table version 10.0*. [online] Available for download at: https://www.rug.nl/ggdc/productivity/pwt/ [Accessed 16 May 2021]

Penn World Table. (2021). *PWT 10.0 - Additional data and programs: Labor Detail* [online] Available for download at: https://www.rug.nl/ggdc/productivity/pwt/ [Accessed 16 May 2021]

Runde, D. (2017). *The Data Revolution in Developing Countries Has a Long Way to Go*. [online] Forbes. Available at: https://www.forbes.com/sites/danielrunde/2017/02/25/the-data-revolution-in-developing-countries-has-a-long-way-to-go/?sh=6b5b60341bfc [Accessed 17 May 2021].

Solow, R. M. (1956). A contribution to the theory of economic growth, *Quarterly Journal of Economics*, vol. 70, no. 1, p.65-94

Todaro, M.P. & Smith, S.C. (2012). *Economic Development*. 11th edn, Boston: Pearson Addison Wesley.

UNESCO Institute for Statistics. (2021). *Data - Education - National Monitoring, Kenya*. [online] Available for download at: http://data.uis.unesco.org/# [Accessed 7 May 2021]

World Bank. (2020). *The Human Capital Index 2020 Update: Human Capital in the Time of COVID-19*. [online] World Bank, Washington, DC: World Bank. Available for download at: https://openknowledge.worldbank.org/handle/10986/34432 License: CC BY 3.0 IGO.

World Bank. (2021a). *Data Bank - Education Statistics.* [online] Available for download at: https://databank.worldbank.org/home [Accessed 18 May 2021]

World Bank. (2021b). *Data Bank - World Development Indicators.* [online] Available for download at: https://databank.worldbank.org/home [Accessed 18 May 2021]