The Impact of Education on GDP per Capita

Helena Weatherly

Kendy Lopez

Carol Tierra

Abstract

Education plays an important role in a country's economic growth. This paper looks at the impact of education on the gdp per capita of different countries across the globe. The data comes from the World Bank and from the Barro and Lee database for the year of 2015. We took the average of total years of schooling to determine how the quantity of education affects GDP per capita. First, we look only at the relationship of education and GDP per capita on a simple regression model. We then add other factors such as the unemployment rate, foreign direct investment, gross savings, and manufacturing to see how they also affect GDP per capita across these countries. Our models found a significant relationship between years of total education and gdp per capita.

I. Introduction

A nation with high economic growth allows its population to enjoy a higher standard of living, and there are different policies that a nation can employ in order to achieve high economic growth. A government's job is to take actions that influence the nation's economy, and to understand how it is performing. There are several actions that the government can take to accomplish a higher economic growth, but they must measure the impact of each action on the economy. Questions arise when deciding which policy to implement: what are the consequences? Is it helping domestic producers, foreign producers? Is the policy worth it?

Governments looking for ways to expand their productivity and economic power take different elements into consideration. Some of the elements that can contribute to promoting economic growth include investing in foreign direct investment, cutting taxes to consumers or businesses, and raising the employment rate in the labor force. There might be other factors such as regulations and restrictions of trade, how attractive the country is to invest, and the unemployment rate, among others.

Looking at the economic growth rate of different countries, it is evident that those with lower education suffer from low GDP per capita growth. We are interested in looking at the relationship between education and the GDP per capita.

Our objective is to investigate how the number of years of total schooling in the workforce affects the economic growth of a nation. The results of this research can impact how governments see the importance of education in their nations and therefore employ policies accordingly that can help their economy grow faster. The hypothesis of this research is that countries with a more educated workforce will display higher levels of economic prosperity. As mentioned before, education is a factor that can affect the income of each country because it depends on economic factors related to increasing human capital which can be evident through changes in employment rates. To measure the education level of each country, we will look at the average years of total schooling in a certain age interval for one year. Thus, the independent variable will be the average years of total schooling between the ages 15 to 64 in 2015 and since we want to investigate the effect of education on the output of each country, GDP per capita of 127 countries in 2015 will be the dependent variable.

II. Literature Review

Deme and Mahmoud (2020) look at the effect of quantity and quality of education on per capita Real GDP among 26 low-income and 34 low- and middle-income African countries over the period from 2003-2016. The study analyzes how economic growth is impacted by primary and secondary education as well as the quality of education offered in these African countries. They adopt the augmented Solow model of Mankiw, Romer, and Weil (1992), where they show the relationship of output, stock of physical capital, stock of human capital, quantity of labor and level of technology, with the exception that their empirical model uses the level of human capital instead of the rate of human-capital growth. Throughout the study, they found a statistically significant correlation between growth in primary and secondary school enrollments and growth in per capita real GDP, and a statistically insignificant correlation between growth in the quality of education and economic growth. Another interesting finding is that primary-school enrollment has the strongest impact on economic growth, followed by secondary school enrollment. These results suggest that policy makers should focus on increasing school enrollment to increase economic growth.

Asiedu and Nandwa (2007) examines the effect of foreign aid in education on economic growth in low and middle income countries. The study looked at data from 1990 to 2004 and analyzed how foreign aid in education at the primary, secondary, and higher education level affected the economic growth of a country. They argued in accordance with the Solow model that increases in human capital were the reason education aid had a positive effect on economic growth and that the aid led to more investment into the education sector of receiving countries. The researchers found that for developing countries, aid to education did not have a statistically significant impact on economic growth. They also concluded that there was no statistically significant relationship for primary, secondary, or higher education aid on the economic growth overall. Interestingly though, the researchers found a statistically significant correlation when separating their regressions by income. They found that in low income countries, primary education aid was the most effective, while in middle income countries, higher education aid was the most effective. This research suggests that educational aid might not always be the best answer and how just the quality of education alone doesn't determine economic prosperity.

Appiah (2017) examines the effect of education expenditure on per capita GDP in developing countries and investigates if the impact is different from that of Sub-Saharan African (SSA) countries. The paper argues that an increase in education expenditure will result in improved labor productivity due to the role that education has in increasing human capital and with it, productivity. The estimator implemented for this analysis is the system General Method of Moment (GMM) proposed by Blundell and Bond (1998), that measures the effect of an increase in education expenditure on per capita GDP. The empirical analyses use a panel data of 139 countries over the period 1975-2015. The research estimates a dynamic panel model of the impact of education expenditure on GDP per capita in developing countries, and contrasts the impact with that of SSA countries' average. The results of the paper confirm that if there is an expansion in education expenditure in developing countries, there is a strong positive effect on GDP per capita in developing countries. However, when comparing these results with those of SSA countries, the researcher found no significant difference in the impact of education expenditure on per capita GDP in developing countries and SSA countries. Note, however, that the impact is more significant in developing countries than that of SSA countries, which means that the level of human capital in SSA countries should be improved to make a significant impact on per capita GDP.

Our contribution to this literature is a cross-sectional analysis, unlike previous literature, that analyzes the relationship between average quantity of education in the workforce on gdp per capita in solely the year 2015. Our research is unique in that we are also utilizing new control variables that have significant impacts on the overall economic health of a country which therefore means we can zero in on the exact effects education may be having on gdp. These four control variables, the unemployment rate, foreign direct investment (FDI), savings rate, and the percentage of gdp made up by manufacturing have seldomly been used in the related literature we looked at, and we thought it important to include them in our model as they are key determinants of a country's economic state. Understanding a country's specific economic situation is crucial when conducting our analyses because as we saw in previous literature, specific circumstances can have widely varied impacts on the results. When controlling for these circumstances, we can measure whether there is truly a statistically significant correlation between average years of total schooling and gdp per capita.

III. Data

Our dataset consists of 127 observations and is obtained from two sources, the Barro and Lee dataset and the World Bank. We decided on a cross sectional analysis for our research and decided to look at the year 2015 so our data would be relatively current but also exempt from any effects of COVID-19. To obtain the average years of schooling for our 127 countries, we used the Barro and Lee dataset on education from 1950 to 2015. In the Barro and Lee dataset, each country's educational attainment data is separated according to 5 age groups: 15 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64. For our variable to accurately represent the entire workforce of a country, we averaged the values of all five age groups into one variable to represent the entire 15-64 age group for each country in 2015. The descriptions of our variables are displayed below in Table 1.

Table 1. Data Descriptions

Variable	*Description*	*Units*	*Source*
GDPperCapita	GDP per capita	(current US$)	World Bank
log_GDPperCapita	Log of GDP per Capita	(current US$)	World Bank
Avgyrsofsch	Average years of total	Years	Barro and Lee

	education (15-64)		
Unempt	Unemployment rate	Percentage (%)	World Bank
FDIinflows	Foreign direct investment, net inflows	% of GDP	World Bank
gsavings	Gross savings	% of GDP	World Bank
manufact	Manufacturing, value added	% of GDP	World Bank
developed	Dummy variable for developed and undeveloped countries	Binary (1 = developed, 0 = undeveloped)	World Bank

We decided to take the log of our dependent variable, *log_GDPperCapita*, for our models as the scatter plot of education to GDP per capita revealed a non-linear relationship. We chose the unemployment rate as one of our control variables because it is an important indicator of the economic health of a country in a given year. By controlling for a country's current economic state, we won't risk misattributing high or low gdp per capitas to education level when the country could just be in recession, for example. We also chose to control for the gross savings rate as this could also affect each country's GDP per capita. We chose foreign direct investment as a control variable so we could measure how much of each country's GDP could be attributed to foreign investment into the country. Lastly, we chose to measure and control for the contribution of the manufacturing sector to each country's GDP. By controlling for these four variables, we can more accurately measure the correlation between years of total schooling and gdp per capita without other factors that can affect gdp. Lastly, we used a dummy variable, *developed,* to separate countries into high-income and non high-income and measure any possible changes in the relationship between education and GDP per capita.

Our dataset originally measured educational attainment for 146 countries but we excluded the few countries that were missing data for our other control variables such as Cuba and Syria. This decreased our total observations down to 127 countries. Descriptive statistics of our variables are shown below.

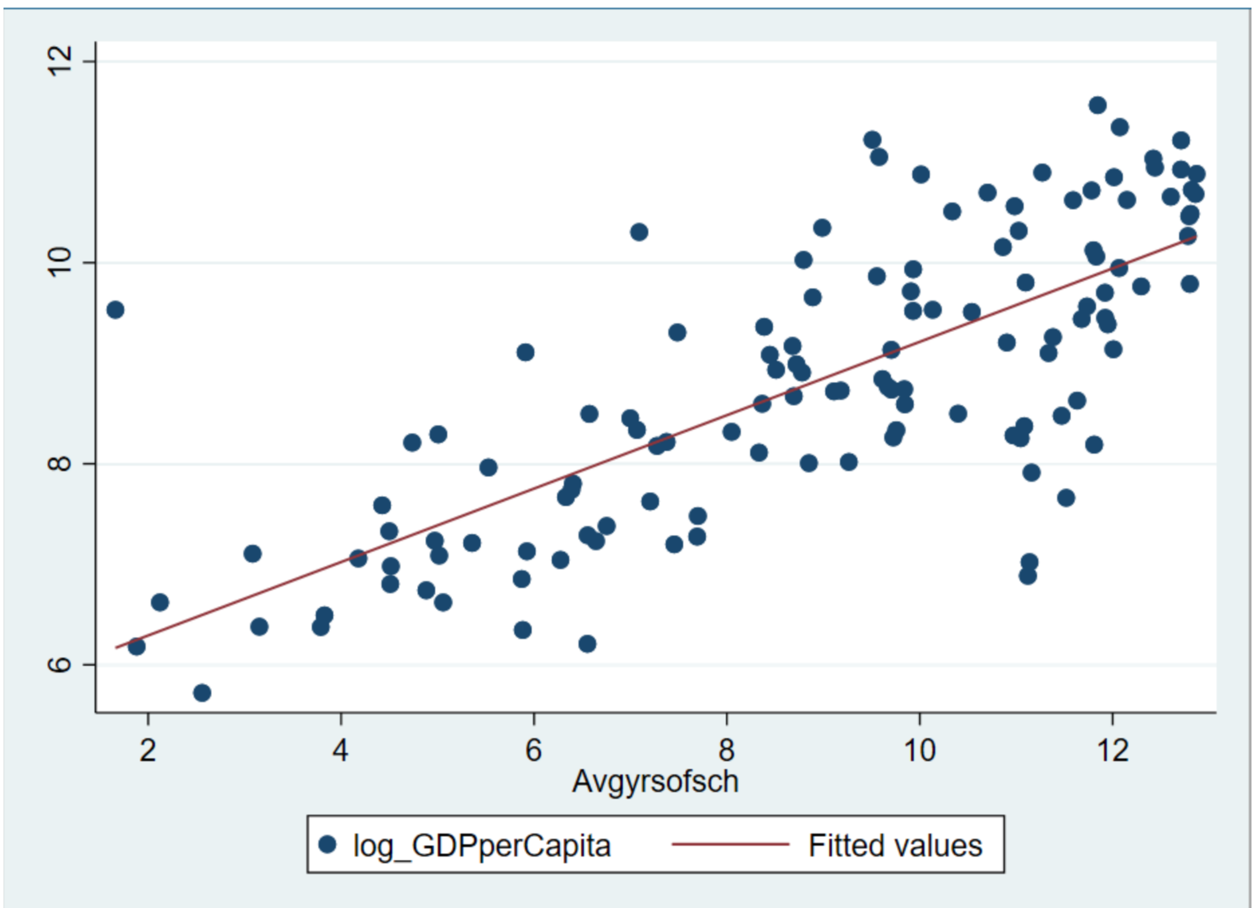
Table 2. Descriptive Statistics

Variable	*Obs.*	*Mean*	*Std. Dev.*	*Min*	*Max*
GDPperCapita	127	15977.49	20520.12	305.5	105462
log_GDPperCapita	127	8.82	1.43	5.72	11.57

Avgyrsofsch	127	8.91	2.93	1.66	12.87
Unempt	127	7.66	5.65	.2	24.9
FDIinflows	127	6.94	17.54	-4.54	146
gsavings	127	22.56	9.72	-1	57
manufact	127	12.91	6.09	1	35
developed	127	.38	.49	0	1

Looking at the standard deviation, we can see a high variation in GDP per capita among countries. Countries like Luxembourg's have large GPDs with relatively small populations resulting in a very high GDP per capita. When taking the coefficient of variance, we see that there is more variability in foreign direct investment compared to the other variables.

Figure 1. Scatter Plot



The scatterplot shown above shows a strong relationship between the years of education and GDP per Capita. The correlation between both variables is 0.75. This suggests a strong positive correlation demonstrating that as education increases, per capita GDP also increases. The scatter plot and the line of best fit demonstrate that there is a positive linear relationship present, that the years of education has a positive effect on the per Capita GDP of a country.

The Gauss-Markov Classical Linear Model Assumptions are checked:

1. Model is linear in parameters- The data can be modeled with a simple regression with a dependent variable GDP per capita and an independent variable which is average years of education, where the parameters are linear. Therefore, assumption SLR.1 of the Gauss-Markov assumptions for simple regression can be met by the data.
2. Random Sampling- We have a random sample of size 127 since there are 127 countries in this model. The sources collected this data using random sampling, so it follows assumption SLR.2.
3. No perfect collinearity. We checked for collinearity between the independent variables and none exhibited linear relationships among each other. Furthermore none of them were constant since we can observe that the sample outcomes on the average years of education are not all the same value, so it follows SLR.3.
4. Homoskedasticity and Normality of Error- Additionally, assumptions SLR.4 and SLR.5 can be met with our data since the error u has an expectation of 0 and it has the same variance given any value. We will interpret the results to check for these conditions to be certain they are met.

IV. Results

Simple Regression:

Model 1:

In the simple regression model, we only show the relationship between the years of education and the log of GDP per capita.

The equation is shown below:

$$log\_GDPperCapita = \beta_0 + \beta_1 Avgyrsofsch + u$$

Model 1 in the Appendix displays the STATA output.

Estimated Equation:

$$\text{log\_GDPperCapita} = 5.5632 + 0.365{*}\text{Avgyrsofsh}$$

$N = 127$ $\qquad$ $R^2 = 0.56$

Multiple Regression:

Model 2:

Our first multiple regression model is between our dependent variable, log(gdp per capita), and all of our independent variables.
The equation is as follows:

$$log\_GDPperCapita = \beta_0 + \beta_1 Avgyrsofsch + \beta_2 Unempt + \beta_3 FDIinflows + \beta_4 gsavings + \beta_5 manufact + u$$

Model 2 in the Appendix displays the STATA output.
Estimated Equation:

log_GDPperCapita = 5.2 + .344*Avgyrsofsch + .001*Unempt + .006*FDIinflows + .029*gsavings - .011manufact

N = 127 R^2 = 0.602

Model 3:

Our second multiple regression model is Model 2 but excluding control variables unemployment rate and manufacturing.
The equation is as follows:

$$log\_GDPperCapita = \beta_0 + \beta_1 Avgyrsofsch + \beta_2 FDIinflows + \beta_3 gsavings + u$$

Model 3 in the Appendix displays the STATA output.
Estimated Equation:

log_GDPperCapita = 5.12 + .342*Avgyrsofsch + + .007*FDIinflows + .027*gsavings

N = 127 R^2 = 0.600

Model 4:

Our last multiple regression model adds the dummy variable *developed* to Model 3 and removes *FDIinflows*.
The equation is as follows:

$$log\_GDPperCapita = \beta_0 + \beta_1 Avgyrsofsch + \beta_2 gsavings + \delta developed + u$$

Model 4 in the Appendix displays the STATA output.
Estimated Equation:

log_GDPperCapita = 6.12 + .197*Avgyrsofsch + + .014*gsavings + 1.63*developed

N = 127 R^2 = 0.79

Combining the results from the simple model and the multiple regression models we obtain the following table:

Table 3: Estimation Results

Dependent Variable: log_GDPperCapita				
Independent Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Avgyrsofsch	0.365*** (0.029)	0.344*** (0.029)	0.342*** (0.029)	0.349*** (0.028)
Unempt	--	0.0012 (0.015)	--	--
FDIinflows	--	0.0065* (0.0048)	0.0071* (0.0047)	--
gsavings	--	0.0285*** (.0089)	0.027*** (0.0085)	0.026*** (0.0086)
manufact	--	-0.011 (0.014)	--	--
constant	5.56*** (0.27)	5.20*** (0.337)	5.12*** (0.299)	5.11*** (0.30)
Observations	127	127	127	127
R - squared	0.56	0.60	0.60	0.59

Significant at *10%, **5%, ***1%

Table 4: Statistical Inference

Independent Variable	t-value	p-value	95% Confidence Interval
Avgyrsofsch	11.78	0.000	(0.286, 0.402)
Unempt	0.08	0.934	(-0.028, 0.030)
FDIinflows	1.36	0.175	(-0.003, 0.016)
gsavings	3.2	0.002	(0.011, 0.046)
manufact	-0.81	0.419	(-0.039, 0.016)

In the table above, we see important characteristics of each independent variable. We include the t-value, p-value, and the confidence intervals of all the constant variables in model 2. Here we see the importance of the Avgyrsofsch and gsavings variables as they are shown to be significant at 1% level. The following variables seem not to be statistically significant: Unempt, FDIinflows, and manufact. We also do an F-statistics to test the significance of those variables. As shown in the table, and the F test, we conclude that they are not jointly significant and that they do not help explain the dependent variable. Therefore we drop these variables and only include Avgyrsofsch and gsavings to help us explain the dependent variable.

Unrestricted model: log_GDPperCapita = $\beta_0 + \beta_1$Avgyrsofsch + β_2Unempt+ β_3FDIinflows+ u
Restricted model: log_GDPperCapita = $\beta_0 + \beta_1$Avgyrsofsch+ u
We test the joint significance of the coefficient of Unempt and FDIinflows by performing an F-test, where our hypotheses are as follows: H_o: $\beta_2 = 0$, $\beta_3=0$ and H_a: Null is not true.

$$F = \frac{(R^2_{ur} - R^2_{r})/q}{(1-R^2_{ur})/n-k-1} = \frac{(0.599-0.594)/2}{(1-0.599)/140-3-1} = 0.8479.$$

At 5% significance level, CV = 3.00. Since F < CV, we fail to reject null hypothesis.
At 10% significance level, CV = 2.71. Since F < CV, we fail to reject null hypothesis.
At 1% significance level, CV = 4.61. Since F < CV, we fail to reject null hypothesis.
Unemployment rate and Foreign direct investment/net inflows are not jointly significant in the model.

V. Interpretation

In the simple regression model, the equation indicates that there is a positive relationship between the years of education and GDP per Capita. This aligns with our hypothesis as we stated that as education increases, the GDP per capita of a country increases. We can see that as education increases by one year, GDP per Capita increases 36.4%. The R^2 suggests that our independent variable, education, explains 56% of the variation on our GDP per Capita variable. With a strong correlation of 0.78 and the positive percentage change of effect of the independent variable, it supports our hypothesis that more education leads to a higher GDP per capita. We extend our model to a multiple regression model and add other control variables in order to better explain the GDP per Capita variable.

In the multiple regression model, we added control variables Unempt, FDIinflows, gsavings, and manufact to our equation. The resulting STATA table and estimated equation indicate that when we hold unemployment rate, foreign direct investment, gross savings rate, and % of manufacturing fixed, we find a positive relationship between years of schooling and GDP per capita. Our equation shows a one year increase in total schooling resulted in a 34.4% increase in gdp per capita. Our third model took out the control variables we found insignificant, unkempt and manufact, and we again found a positive relationship between education and gdp per capita when holding foreign direct investment and gross savings fixed. A one year increase in total schooling results in a 34.2% increase in gdp per capita. In our final model, after holding gross savings and developed vs. non-developed countries fixed, we saw a positive, though slightly weaker, relationship between total schooling and GDP per capita. A one year increase in total schooling results in a 19.7% increase in gdp per capita. However, the R^2 increased significantly from 0.600 to 0.79, meaning our independent variables now explained 79% of the variation on gdp per capita. The results from our models align with our hypothesis of a positive correlation between education level and gdp per capita.

VI. Conclusions

After running different models, our study reveals that our initial hypothesis, that a more educated workforce leads to a prosperous economy, is correct. In the simple regression model we saw that there is a strong positive relationship between years of education and the GDP per capita. We see that as education increases by one year, GDP per capita increases by 36%. However, the R-squared indicates that only 56% of the independent variable explained the variability in the GDP per capita. Therefore, we decided to include other control variables such as unemployment rate, gross savings rate, foreign direct investment, and the % of gdp made up by the manufacturing sector. These multiple regression models again revealed a positive relationship between years of total schooling and gdp per capita. We can see that when education increases by one year, GDP per capita increases by 34.4%. However, the R-squared only slightly increased to 0.602.

After we added new variables and performed F-statistics tests, and looked at the t-values and p-values of the variable we see that only education and savings rate are statistically significant as well as our dummy variable. This supports our hypothesis that a positive significant relationship exists between years of total schooling and the economic prosperity of a country. It also states that higher education and higher savings rates tend to increase the economic growth of a country. Countries with better education and high saving rates tend to have a high GDP per capita. This conclusion is important to further research as a strong relationship between education and economic output of a country can influence policymakers to put more emphasis on education expenditure. It can also influence policymakers to reform the standards and quality of their countries' education systems if it is shown to have a significant impact on the country's economic health. Overall, this research can have lasting impacts on the development of the educational system and human capital investment as a whole.

Resources

Appiah, Elizabeth. "The Effect of Education Expenditure on Per Capita GDP in Developing Countries." *International Journal of Economics and Finance*, vol. 8, no. 10, 2017, pp. 136-144., https://doi.org/10.5539/ijef.v9n10p136

Asiedu, Elizabeth, and Boaz Nandwa. "On the Impact of Foreign Aid in Education on Growth: How Relevant Is the Heterogeneity of Aid Flows and the Heterogeneity of Aid Recipients?" *Review of World Economics / Weltwirtschaftliches Archiv* 143, no. 4 (2007): 631–49. http://www.jstor.org/stable/40441156.

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. http://www.barrolee.com/

Deme, Mamit, and Ali M. Mahmoud. "Effect of Quantity and Quality of Education on per Capita Real-GDP Growth: Evidence from Low- and Middle-Income African Countries." *Applied Economics*, vol. 52, no. 57, 2020, pp. 6248–6264., https://www.tandfonline.com/doi/pdf/10.1080/00036846.2020.1789058?needAccess=true

World Bank Open Data. https://data.worldbank.org/

Appendix

Simple Regression:

Model 1:

```
. reg log_GDPperCapita Avgyrsofsch

      Source |       SS           df       MS      Number of obs   =       127
-------------+----------------------------------   F(1, 125)       =    160.31
       Model |  144.112617         1  144.112617   Prob > F        =    0.0000
    Residual |  112.367894       125  .89894315    R-squared       =    0.5619
-------------+----------------------------------   Adj R-squared   =    0.5584
       Total |  256.480511       126  2.03555961   Root MSE        =    .94813

------------------------------------------------------------------------------
log_GDPper~a | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
 Avgyrsofsch |    .364993    .028827    12.66   0.000     .3079408    .4220453
       _cons |   5.563195   .2703035    20.58   0.000      5.02823    6.098159
------------------------------------------------------------------------------
```

Multiple Regression Models:

Model 2:

```
. regress log_GDPperCapita Avgyrsofsch Unempt FDIinflows gsavings manufact

      Source |       SS           df       MS      Number of obs   =       127
-------------+----------------------------------   F(5, 121)       =     36.63
       Model |  154.441512         5  30.8883024   Prob > F        =    0.0000
    Residual |  102.038999       121  .843297512   R-squared       =    0.6022
-------------+----------------------------------   Adj R-squared   =    0.5857
       Total |  256.480511       126  2.03555961   Root MSE        =    .91831

------------------------------------------------------------------------------
log_GDPper~a | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
 Avgyrsofsch |   .3440916   .0291989    11.78   0.000     .2862847    .4018985
      Unempt |   .0012213   .0147556     0.08   0.934    -.0279912    .0304339
  FDIinflows |   .0065241   .0047874     1.36   0.175    -.0029539    .0160021
    gsavings |   .0285426   .0089065     3.20   0.002     .0109098    .0461755
    manufact |  -.0114148   .0140823    -0.81   0.419    -.0392944    .0164648
       _cons |   5.198268   .3367075    15.44   0.000     4.531667    5.864869
------------------------------------------------------------------------------
```

Model 3:

excluding unemployment and manufacturing

```
. regress log_GDPperCapita Avgyrsofsch FDIinflows gsavings
```

Source	SS	df	MS
Model	153.883935	3	51.294645
Residual	102.596576	123	.834118504
Total	256.480511	126	2.03555961

Number of obs	=	127
F(3, 123)	=	61.50
Prob > F	=	0.0000
R-squared	=	0.6000
Adj R-squared	=	0.5902
Root MSE	=	.9133

log_GDPper~a	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Avgyrsofsch	.3418683	.0286591	11.93	0.000	.2851393	.3985973
FDIinflows	.0070636	.0047092	1.50	0.136	-.002258	.0163851
gsavings	.0267501	.0085177	3.14	0.002	.0098898	.0436104
_cons	5.116812	.2986421	17.13	0.000	4.525668	5.707956

Excluding unemployment, manufacturing and FDI

```
. regress log_GDPperCapita Avgyrsofsch gsavings
```

Source	SS	df	MS
Model	152.007292	2	76.0036459
Residual	104.473219	124	.84252596
Total	256.480511	126	2.03555961

Number of obs	=	127
F(2, 124)	=	90.21
Prob > F	=	0.0000
R-squared	=	0.5927
Adj R-squared	=	0.5861
Root MSE	=	.91789

log_GDPper~a	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Avgyrsofsch	.3492092	.0283801	12.30	0.000	.293037	.4053814
gsavings	.0261782	.0085519	3.06	0.003	.0092515	.0431049
_cons	5.113288	.3001341	17.04	0.000	4.519239	5.707338

Model 4:

```
. regress log_GDPperCapita Avgyrsofsch gsavings developed

      Source |       SS           df       MS      Number of obs   =       127
-------------+----------------------------------   F(3, 123)       =    154.02
       Model |  202.559893         3  67.5199643   Prob > F        =    0.0000
    Residual |  53.9206181       123  .438379009   R-squared       =    0.7898
-------------+----------------------------------   Adj R-squared   =    0.7846
       Total |  256.480511       126  2.03555961   Root MSE        =     .6621

------------------------------------------------------------------------------
log_GDPper~a | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
 Avgyrsofsch |   .1970763   .0248954     7.92   0.000     .1477974    .2463552
    gsavings |   .0140441   .0062714     2.24   0.027     .0016303    .0264579
   developed |   1.634321   .1521915    10.74   0.000     1.333067    1.935574
       _cons |   6.124979   .2361058    25.94   0.000     5.657622    6.592335
------------------------------------------------------------------------------
```