https://doi.org/10.25130/tjaes.17.55.3.33

Does Education Affect Income Distribution in Developing Countries?

Researcher: Ruba Rafea M. Al-Wazzan
College of Administration and Economics
University of Mosul

Prof. Dr. Mufeed D. Y. Almula-Dhanoon
College of Administration and Economics
University of Mosul

Abstract:

The research aims to measure the impact of the level of education on income distribution in developing countries. Panel data for 2003-2017 were collected for a 24 developing countries, using ARDL model. The estimated model took into account the impact of population growth, unemployment rate, and per capita GDP, in addition to the average years of schooling as explanatory variables for income shares held by, the highest 20%, the middle 40%, and lowest 40% within three models. It was found that education is an important determinant of income distribution in developing countries in the long run, as it contributes to reducing the share of the rich class, and increases the share of the middle and poor classes. As for the short-run, it has not been proven that there is a specific effect of education on income distribution.

Keywords: Education, Income Distribution, ARDEL, Panel Data, Developing Countries.

هل يؤثر التعليم على توزيع الدخل في البلدان النامية؟

الباحثة: ربى رافع محمد الوزان
كلية الإدارة والاقتصاد
جامعة الموصل
Ruba.bap303@student.uomosul.edu.iq

أ.د. مفيد ذنون يونس الملا ذنون
كلية الإدارة والاقتصاد
جامعة الموصل
dr.mufeed@uomosul.edu.iq

المستخلص:

يهدف البحث الى قياس تأثير مستوى التعليم على توزيع الدخل في الدول النامية. تم تجميع بيانات (Panel) للمدة ٢٠٠٣-٢٠١٧ لمجموعة مختارة من ٢٤ دولة نامية، وباستخدام نموذج (ARDL). النموذج المقدر اخذ بالاعتبار تأثير كل من النمو السكاني، ومعدل البطالة، وحصة الفرد من الناتج، بالإضافة الى متوسط سنوات الدراسة كمتغيرات تفسيرية لحصص الدخل، اعلى ٢٠%، واوسط ٤٠%، وأدنى ٤٠%، ضمن ثلاثة نماذج. وجد ان التعليم يعد محددا مهما لحصص الدخل الموزعة في الدول النامية في الاجل الطويل. حيث يسهم في تخفيض حصة الفئات الغنية من الدخل، ويزيد حصة الفئات المتوسطة والفقيرة. اما في الاجل القصير فلم يثبت وجود تأثير محدد للتعليم على توزيع الدخل.

الكلمات المفتاحية: التعليم، توزيع الدخل، نموذج ARDL، بيانات Panel، البلدان النامية.

1. **Introduction:** Developing countries have been interested in developing education, and increasing the average years of schooling as an important step in economic development. In September 2000, world leaders committed themselves to achieving the Millennium Development Goals by

2015. Achieving Universal Primary Education (UPE) is one of these goals. Education for All (EFA), a related international initiative launched in 1990, goes further, it calls for the world's commitment to provide education for all citizens. Achieving this was considered crucial to the success of the Millennium Development Goals. This importance was confirmed by the Sustainable Development Goals, which were issued in 2015 and included seventeen goals, of which education ranked as the fourth goal.

Theoretical literature focuses on the developmental effects of education, but gives less attention to the distributional effects of education. Education does not only affect increasing income, but also contributes to its redistribution. So, there is a need to study the distributional effects of education. So, the question raised by the research is to what extent does education contribute to the distribution of income? There is a wide debate among economists on the importance of education on income distribution. Some economists believes that education improves income distribution. Other economists suggest that education in developing countries is deepening inequality in income distribution. this issue produces a problematic search is trying to investigate.

The key research question is: Does education affect income distribution in developing countries? And what is the nature and extent of this effect?

There is very little empirical research that studies the effect of education on income distribution. Therefore, this research is an important knowledge addition through the experimental conclusions that will be reached.

The research hypothesis is: Education improves income distribution in the long-run, and the poorest class in society is the most benefited from it. But education not expected to affect income distribution in the short-run, because its impact requires a long time.

After introduction, section 2 will discuss the literature, methodology and data will displayed in Section 3, Section 4 review estimation results, last section contain the conclusion and policy implications.

2. **Literature Review:** There is a wide debate among economists on the importance of education on income distribution. Some economists believes that education improves income distribution. When education is free and available to all with the same quality, it will be a lifeline from poverty.

because it provides wide opportunities for the poor to increase their skills and knowledge that enable them to obtain higher-income jobs. Thus, education will be one of the important factors in improving income distribution (Todaro M., and, S. Smith, 2015, 404-405; Abdullah A., et al., 2011, 3; Ozturk I., 20084; Kumba D., 2009, 6). Empirically, (Lee J. and H. Lee, 2018), found that the expansion of education contributes to reducing inequality in education, and thus inequality in income. (Roy P. and Z. Husain, 2019) analysis revealed that education reduced inequality in India in the 1970s, and from the 1990s onwards. (Abdullah A., et al., 2011) concluded that education affects both sides of the income distribution, decreasing the income share of high-income earners and increasing the share of low income, but it has no effect on the middle-class share. Findings by (Shahabadi A., 2018) shows that, in selected Islamic countries, primary and secondary school enrollment has a significant negative effect on income inequality, and that college enrollment has a significant positive effect on income inequality. The results of (Jeng R. et al., 2019) confirm a slight negative relationship between education and income inequality.

Other economists suggest that educational systems in developing countries are deepening inequality. The reason is that there is a positive correlation between the level of education and the level of income. Whereas the level of income that can be earned depends on the number of years of education, a great deal of income inequality will be generated if a large proportion of students belong to the upper- and middle-income groups, and are enrolled in secondary and university education. If the poor cannot enroll in secondary and university education for financial or other reasons, the education system will perpetuate and possibly increase the inequality. The same result can be obtained if we take into account the differences between females and males in access to education in developing countries (Todaro M., and, S. Smith, 2015, 406; Abdullah A., et al, 2011, 4; Kumba D., 2009, 7; Petcu C., 2014, 4). Empirically (Castelló-Climent A. and R. Doménech, 2014) findings show that while human capital inequality has declined significantly around the world, inequality in income distribution has not changed. Thus, improvements in education are not a sufficient condition for reducing income inequality, even though education significantly improves the living standards of people in the lower classes of income.

3. **Data and Methodology:** When both the dependent variable and the explanatory variables in one period are affected by their value in previous periods, it is necessary to include those variables in the model, so, we will have a dynamic model, in this case we are dealing with Lagged Time Models. An example of these models is the Auto-Regressive Distributed Lag Model (ARDL). In this paper, we have four explanatory variables with a dependent variable each time, so we will have the model of the order ARDL (p, q1, q2, q3, q4). It is important to note here that in the analysis of Panel data, the lag lengths for all explanatory variables must be of the same order. So, our model will be:

$$\Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \Delta Y_{t-i} + \sum_{i=0}^{q_1} \beta_{1,i} \Delta X_{1,t-i} + \sum_{i=0}^{q_2} \beta_{2,i} \Delta X_{2,t-i} + \sum_{i=0}^{q_3} \beta_{3,i} \Delta X_{3,t-i} + \sum_{i=0}^{q_4} \beta_{4,i} \Delta X_{4,t-i} + \sum_{j=1}^{4} \delta_j X_{j,t-1} + \lambda Y_{t-1} + U_t \quad \ldots\ldots (1)$$

Since, N is the number of cross-sections and T is the length of time or number of periods, the total number of views will be N×T.

Y_t = The dependent variable representing one of the income shares ($Y_{1,t}$ represents the highest 20%, $Y_{2,t}$ represents the middle segment (40%), $Y_{3,t}$ represents less than 40%).

X_1= unemployment rate (the percentage of unemployment in the total labor force).

X_2= average years of schooling.

X_3= rate of population growth.

X_4= per capita GDP.

Panel data for 2003-2017 were collected for a selected group of 24 developing countries (Argentina, Belarus, Bolivia, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Estonia, Honduras, Kazakhstan, Kyrgyzstan, Latvia, Moldova, Panama, Peru, Poland, Paraguay, Romania, El Salvador, Thailand, Turkey, Ukraine, and Paraguay), from two sources: world development indicators published by world bank, and, average years of schooling is collected from Our World in Data website. (https://ourworldindata.org/global-education#years-of-schooling)

4. **Estimation Results:**

4-1. Stability Test: The first step in building the regression model is to test the stability of the model variables using unit root test. Stability means that the arithmetic mean and variance of the variable are constant. In the case of instability of the variables, the regression results will be spurious. There are many kinds of unit root tests, but the most common one in studies of Panel data is the Levin, Lin & Chu (LLC) test. Table (1) shows the results of unit root (LLC) test for the variables of the three models.

Table (1): Results of the Unit Root Test (LLC)

	Max. Lag	At Level		1st Difference	
		Individual Intercept	Indiv. Inter. & Trend	Individual Intercept	Indiv. Inter. & Trend
Y_1	3	-5.816***	-8.825***	—	—
		(0.000)	(0.000)		
Y_2	3	-5.586***	-6.583***	—	—
		(0.000)	(0.000)		
Y_3	3	-5.309***	-9.438***	—	—
		(0.000)	(0.000)		
X_1	2	-5.815***	-3.621***	—	—
		(0.000)	(0.000)		
X_2	2	-5.655***	-8.001***	—	—
		(0.000)	(0.000)		
X_3	2	-4.621***	-16.161***	—	—
		(0.000)	(0.000)		
X_4	1	-11.649***	-11.670***	—	—
		(0.000)	(0.000)		
*** significant at 1% level					

Source: authors' work/EViews-10 program outputs.

Table (1) indicate that the three dependent variables, as well as the four explanatory variables were all stationary, i.e., they were integrated at the level I(0).

4-2. Estimation of the Models: We will estimate three models, income hare held by highest 20% model, income share held by middle 40% model, and income share held by lowest 40% model

4-2-1. Income Share Held by Highest 20% Model: This model estimates the impact of, unemployment rate, average years of schooling, population growth, per capita GDP, on the income share held by highest 20% for the period (2003-2017).

Determination of the Optimal Lag Length Using (VAR) Model:

Table (2) presents the criteria for selecting the optimal lag length of Income share held by highest 20% model according to (VAR) analysis.

Table (2): Determining the Optimal Lag Length for the Highest 20% Model

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1812.736	NA	1720.432	21.63972	21.73269	21.67745
1	-674.6868	2194.809	0.003027	8.389129	8.946980*	8.615532
2	-632.7164	78.44474	0.002475	8.187100	9.209826	8.602172*
3	-606.8421	46.82021	0.002455*	8.176691*	9.664293	8.780433
4	-599.2292	13.32243	0.003030	8.383682	10.33616	9.176092
5	-578.9682	34.25074	0.003225	8.440098	10.85745	9.421178
6	-568.4916	17.08687	0.003868	8.612996	11.49522	9.782744
7	-555.2902	20.74506	0.004507	8.753455	12.10056	10.11187
8	-521.8266	50.59386*	0.004143	8.652697	12.46468	10.19978

* Denotes the optimal Lag length of the variable
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Source: authors' work/ EViews-10 program outputs.

Table (2) indicate that the optimal lag length is, (8) according to (LR) criterion, (3) according to (FPE, AIC) criteria, one according to (SC) criterion, and two according to (HQ) criterion. Majority of the criteria indicated that the optimal lag length is (3). So, the optimal lag length for the purpose of estimating the ARDL model is at time (t-3).

Model Estimation (ARDL): using the optimal lag length (3) for all the model variables, it was found that the best order of the model was the first difference for all variables, meaning that the model would be of the type ARDL (1,1,1,1,1) with Individual Intercept and Trend. This model achieves the lowest value for the Akike Information criterion (AIC) of (2.78) among the other models, as is evident from Figure (1).

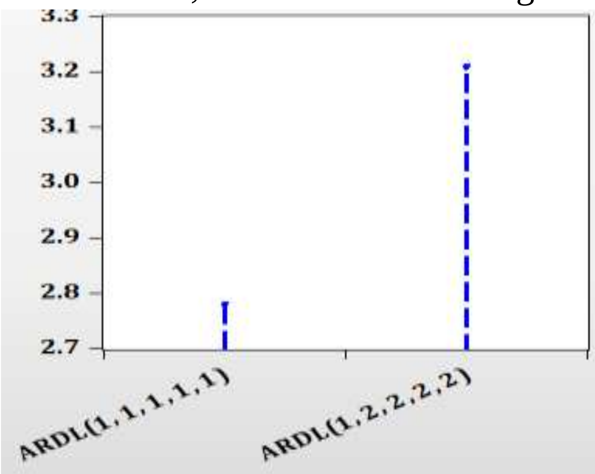

Figure (1): Akaiki Values for ARDL Models with Different Lags for the Income Share Held by Highest 20% Model

Source: authors' work/EViews-10 program outputs.

So, ARDL (1,1,1,1,1) model to be estimated will take the following form:

$$\Delta Y_{1,t} = \alpha_0 + \beta_{1,0}\Delta X_{1,t} + \beta_{2,0}\Delta X_{2,t} + \beta_{3,0}\Delta X_{3,t} + \beta_{4,0}\Delta X_{4,t} + \lambda Y_{1,t-1} + \delta_1 X_{1,t-1}$$

$$+\delta_2 X_{2,t-1} + \delta_3 X_{3,t-1} + \delta_4 X_{4,t-1} + \Phi@Ttrend + U_{2,t} \qquad \ldots\ldots(2)$$

Table (3) displays the model estimation results.

Table (3): Results of ARDL (1, 1, 1, 1, 1) Estimates of the Highest 20% Model

Dependent Variable: ΔY_1				
Method: ARDL				
Sample: 2004 2017				
Included observations: 336				
Maximum dependent lags: 1 (Automatic selection)				
Dynamic regressors (1 Lag, automatic): X_1 X_2 X_3 X_4				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Long Run Equation				
$X_{1,t}$	0.0843	0.0189	4.4659**	0.000
$X_{2,t}$	-0.8190	0.2045	-4.0058**	0.000
$X_{3,t}$	1.4456	0.2726	5.3027**	0.000
$X_{4,t}$	0.0007	0.0180	0.0378$^{n.s}$	0.970
Short-run Equation				
CointEq01	-0.9589	0.0950	-9.6277**	0.000
$\Delta X_{1,t}$	0.3502	0.3737	0.9372$^{n.s}$	0.350
$\Delta X_{2,t}$	0.7324	0.6987	1.0483$^{n.s}$	0.296
$\Delta X_{3,t}$	-0.8421	9.5927	-0.0878$^{n.s}$	0.930
$\Delta X_{4,t}$	0.0082	0.0152	0.5423$^{n.s}$	0.588
C	53.131	5.7405	9.2553**	0.000
@Trend	-0.3076	0.0689	-4.4652**	0.000
Mean dependent var	-0.3205	S.D. dependent var		1.7413
S.E. of regression	1.0783	Akaike info criterion		2.7864
Sum squared resid	218.5899	Schwarz criterion		4.6431
Log likelihood	-329.5590	Hannan-Quinn criter.		3.5247
R-squared	0.9894			
Adjusted R-squared	0.9892			
F-statistic	6000.612**			
Prob(F-statistic)	0.000			
** Significant at 1% level * Significant at 5% level n.s not significant				

Source: authors' work/ EViews-10 program outputs.

Short-run and Long-run Equations:

The long-run equation takes the following form:

$$\Delta Y_{1,t} = 0.0843X_{1,t} - 0.819X_{2,t} + 1.14456X_{3,t} + 0.0007X_{4,t}$$

$$\Rightarrow Y_{1,t} = Y_{1,t-1} + 0.0843X_{1,t} - 0.819X_{2,t} + 1.14456X_{3,t} + 0.0007X_{4,t} \quad \ldots\ldots\ldots(3)$$

While the short-run equation takes the following form:

$$\Delta Y_{1,t} = -0.9589CointEq01 + 0.3502\Delta X_{1,t} + 0.7324\Delta X_{2,t} - 0.8421\Delta X_{3,t}$$

$$+0.0082\Delta X_{4,t} + 53.1307 - 0.30762@Trend \quad \ldots\ldots\ldots(4)$$

The results of short-run and long-run estimation can be summarized as follows: Education did not have a significant impact on the income share of the rich class in the short-run, but in the long run, education contributes to reducing the gap between the poor and wealthy classes, as it was found that the average number of years of schooling has a significant negative impact on the income share held by highest 20% at (1%), as increasing the number of years of schooling by one year leads to a decrease in the income share for this class by (0.82%). This is evidence that investment in human capital has important distributional effects.

There was no significant effect of unemployment rate, and population growth rate on the income share held by the richest 20% in the short-run. While in the long run, it was found that unemployment rate has a significant and positive impact on the income share held by rich class at (1%), as increasing the unemployment rate by (1%) leads to an increase income share held by this class by (0.08%). On the other hand, population growth rate has a significant and positive effect on the income share of the rich class at (1%), in the long run, as an increase in the population growth rate by (1%) leads to a decrease in the income share held by this class by (1.14%). Per capita GDP, did not show significant effect, neither in the long run nor in the short-run, on the income share held by richest class.

CointEq01 represents the coefficient of co-integration between the model variables in the long run, although it is significant at (1%), but it does not indicate the existence of a long-run co-integration relationship between the variables, because all the variables were originally static at I(0).

The value of the intercept is significant at (1%), indicates that in the absence of the influence of the explanatory variables, the average income shares held by highest 20% will be (53%).

The significance of secular trend (@Trend) at (1%), indicates that explanatory variables have a general decreasing trend in the short-run, meaning that they contribute to reducing the income share held by highest 20% by (0.31%) in the short-run, however, this decline is not guaranteed to be achieved in the long run.

Estimating the Overall Model: It can be seen from the results of table (3) that the estimated ARDL (1, 1, 1, 1, 1) model takes the following formula:

$$\Delta Y_{1,t} = 0.0843X_{1,t} - 0.819X_{2,t} + 1.14456X_{3,t} + 0.0007X_{4,t} - 0.9589CointEq01$$
$$+0.3502\Delta X_{1,t} + 0.7324\Delta X_{2,t} - 0.8421\Delta X_{3,t} + 0.0082\Delta X_{4,t} + 53.1307$$
$$-0.30762@Trend \quad \ldots\ldots\ldots(5)$$

$\bar{R}^2$ indicate that explanatory variables explain 99% of the changes in the income share held by highest 20%, the value of (F) test reflect the significancy of ARDL (1, 1, 1, 1, 1) model at (1%).

4-2-2. Income Share Held by Middle 40% Model: This model estimates the impact of, unemployment rate, average years of schooling, population growth, per capita GDP, on the income share held by middle 40% for the period (2003-2017).

Determining the Optimal Lag Length Using VAR Model: Table (4) presents the criteria for choosing the optimal lag length for the Income share held by middle 40% model in developing countries for the period (2003-2017) according to (VAR) analysis.

Table (4): Determining the Optimal Lag Length for the Middle 40%

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1733.370	NA	668.8038	20.69488	20.78785	20.73261
1	-618.7522	2149.619	0.001555	7.723240	8.281091*	7.949643
2	-568.0194	94.82208	0.001146*	7.416897*	8.439623	7.831969*
3	-547.8528	36.49184	0.001216	7.474438	8.962040	8.078180
4	-538.9459	15.58706	0.001478	7.666023	9.618500	8.458433
5	-520.0075	32.01488	0.001599	7.738185	10.15554	8.719265
6	-508.7267	18.39849	0.001899	7.901509	10.78374	9.071257
7	-496.2545	19.59916	0.002232	8.050649	11.39775	9.409067
8	-458.2261	57.49531*	0.001943	7.895549	11.70753	9.442636

*Denotes the optimal Lag length of the variable
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Source: authors' work/EViews-10 program outputs.

The results show that the optimal lag length is (8) according to (LR) criterion, two according to (FPE, AIC, HQ) criteria, and one according to (SC) criterion. because the majority of criteria indicated that the optimal length lag is (2), accordingly, the optimal lag length for the purpose of estimating the ARDL model is at time (t-2).

Estimation ARDL Model:

Using the optimal lag length 2 for the model, it was found that best rank for the model was the first difference for all variables, so, the model will be of type ARDL (1,1,1,1,1), with Individual Intercept and Trend, this model achieves the lowest value for the Akaiki Information criterion (AIC) (2.269), among other models, as shown in Figure (2).

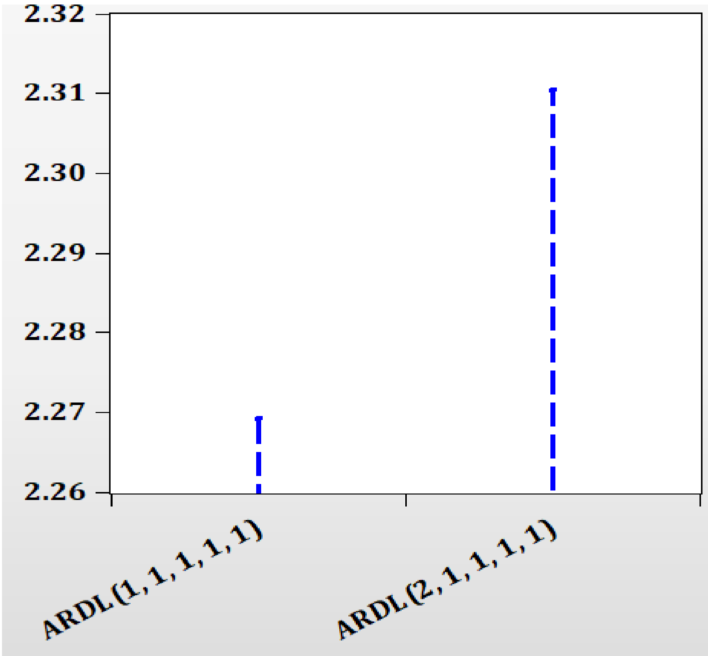

Figure (2): Akaiki Values for ARDL Models with Different Lags for the Income Share Held by Middle 40% Model

Source: authors' work/ EViews-10 program outputs.

So, ARDL (1,1,1,1,1) model to be estimated will take the following form:

$$\Delta Y_{2,t} = \alpha_0 + \beta_{1,0}\Delta X_{1,t} + \beta_{2,0}\Delta X_{2,t} + \beta_{3,0}\Delta X_{3,t} + \beta_{4,0}\Delta X_{4,t} + \lambda Y_{2,t-1} + \delta_1 X_{1,t-1} + \delta_2 X_{2,t-1} + \delta_3 X_{3,t-1} + \delta_4 X_{4,t-1} + +\Phi @Trend + U_{3,t} \quad \ldots\ldots\ldots(6)$$

Table (5) presents the model estimation results.

Table (5): Results of ARDL (1, 1, 1, 1, 1) Estimates of the Middle 40% Model

Dependent Variable: ΔY_2				
Method: ARDL				
Sample: 2004 2017				
Included observations: 336				
Maximum dependent lags: 1 (Automatic selection)				
Dynamic regressors (1 Lag, automatic): X_1 X_2 X_3 X_4				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Long-run Equation				
$X_{1,t}$	-0.0650	0.0243	-2.6722**	0.008
$X_{2,t}$	1.0081	0.0957	10.5300**	0.000
$X_{3,t}$	-2.5111	0.1405	-17.8766**	0.000
$X_{4,t}$	-0.0156	0.0066	-2.3537*	0.020
Short-run Equation				
CointEq01	-0.9299	0.1104	-8.4218**	0.000
$\Delta X_{1,t}$	-0.0998	0.2267	$-0.4403^{n.s}$	0.660
$\Delta X_{2,t}$	-0.9277	0.6317	$-1.4686^{n.s}$	0.144
$\Delta X_{3,t}$	-0.7531	8.0889	$-0.0931^{n.s}$	0.926
$\Delta X_{4,t}$	0.0103	0.0107	$0.9618^{n.s}$	0.337
C	15.792	1.7525	9.0113**	0.000
@Trend	0.1391	0.0550	2.5313*	0.012
Mean dependent var	0.1661	S.D. dependent var		1.0560
S.E. of regression	0.6536	Akaike info criterion		1.8372
Sum squared resid	80.3016	Schwarz criterion		3.6939
Log likelihood	-158.6914	Hannan-Quinn criter.		2.5754
R-squared	0.9884			
Adjusted R-squared	0.9879			
F-statistic	1765.346**			
Prob(F-statistic)	0.000			
**Significant at 1% level * Significant at 5% level n.s not significant				

Source: authors' work/EViews-10 program outputs.

Short-run and Long-run Equations:

The long-run equation takes the following form:

$$\Delta Y_{2,t} = -0.065X_{1,t} + 1.0081X_{2,t} - 2.5111X_{3,t} - 0.0156X_{4,t}$$

$$\Rightarrow Y_{2,t} = Y_{2,t-1} - 0.065X_{1,t} + 1.0081X_{2,t} - 2.5111X_{3,t} - 0.0156X_{4,t} \quad \dots\dots\dots(7)$$

While the short-run equation takes the following form:

$$\Delta Y_{2,t} = -0.9299CointEq01 - 0.0998\Delta X_{1,t} - 0.9277\Delta X_{2,t} - 0.7531\Delta X_{3,t}$$
$$+0.0103\Delta X_{4,t} + 15.792 + 0.1391@Trend \quad \dots\dots\dots(8)$$

The results of short-run and long-run estimation can be summarized as follows: Education has no significant effect on income share of middle class in the short-run in developing countries. In the long-run, the average years of schooling contributes to achieving equality, as it was found that the average years of schooling has a significant positive effect on the income share held by middle 40% at (1%), as the increase in the average years of schooling by one year leads to an increase the income share held by this class by (1%). The middle-income group is predominantly employees, and their incomes are closely related to their level of education.

Unemployment rate, population growth rate, and per capita GDP, do not have a significant effect on the income share of the middle class in the short-run. But in the long-run, we can discern a significant negative impact of unemployment on the income share held by middle 40% at (1%). An increase in unemployment rate by (1%) leads to a decrease in the income share of this class by (0.07%). As specified this class is mostly from the employees, so, their share in income get hurt with unemployment. Also, in the long-run, population growth affects adversely the income share of this class. As the increase in the population growth rate by (1%) leads to a decrease in the income share of this class by (2.51%). On the other hand, per capita GDP has a significant and adverse effect on the Income share held by middle 40% at (5%). As an increase in per capita GDP by (1%) leads to a decrease in the income share of this class by (2%). This confirms the idea that the fruits of growth are distributed in disadvantage of the poor in developing countries.

(CointEq01) represents the co-integration coefficient between the variables in the long-run, although it is significant at (1%), it does not indicate the existence of a long-run co-integration relationship between the variables, because all variables were originally static at I(0).

The intercept value (15,792) indicates that in the absence of the influence of all explanatory variables, the income share of the middle 40% in developing countries will be (16%).

The significancy of secular trend value (@Trend) of (0.1391) indicates that the explanatory variables have an increasing secular trend in the short-run, meaning that they contribute to increasing the Income share held by middle 40% by about (0.14%). This increase is not guaranteed in the long-run.

Estimating the Overall Model: The results of Table (5) show that the estimated ARDL (1, 1, 1, 1, 1) model takes the following form:

$$\Delta Y_{2,t} = -0.065X_{1,t} + 1.0081X_{2,t} - 2.5111X_{3,t} - 0.0156X_{4,t} - 0.9299CointEq01$$
$$-0.0998\Delta X_{1,t} - 0.9277\Delta X_{2,t} - 0.7531\Delta X_{3,t} + 0.0103\Delta X_{4,t} + 15.792$$
$$+0.1391@Trend \qquad \dots\dots\dots(9)$$

$\bar{R}^2$ indicates that the explanatory variables explain (99%) of the changes in the Income share held by middle 40%, (F) test value of (4930.346) reflects the significance of the ARDL model (1, 1, 1, 1, 1) at (1%).

4-2-3. Income Share Held by Lowest 40% Model: In this model, the impact of (unemployment rate, average years of education, population growth rate, per capita GDP) on the Income share held by lowest 40% in developing countries for the period (2003-2017) is estimated.

Determining the Optimal Lag Length Using VAR Model: Table (6) presents the criteria for choosing the optimal lag length for the income share held by lowest 40% in developing countries according to (VAR) analysis.

Table (6): Determining the Optimal Lag Length for the Lowest 40% Model

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1696.046	NA	428.8738	20.25055	20.34352	20.28828
1	-557.4998	2195.768	0.000750	6.994046	7.551896*	7.220449
2	-516.3038	76.99746	0.000619	6.801235	7.823961	7.216307*
3	-486.5085	53.91521	0.000586*	6.744149*	8.231751	7.347890
4	-479.4948	12.27392	0.000728	6.958272	8.910749	7.750682
5	-455.5022	40.55891	0.000742	6.970265	9.387618	7.951344
6	-445.1746	16.84391	0.000891	7.144936	10.02716	8.314684
7	-429.9078	23.99065	0.001013	7.260807	10.60791	8.619225
8	-401.2372	43.34717*	0.000986	7.217110	11.02909	8.764197

* Denotes the optimal Lag length of the variable
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Source: authors' work/EViews-10 program outputs.

Table (6) shows that the optimal lag length is (8) according to (LR) criterion, (3) according to (FPE, AIC), one according to (SC), and two according to (HQ). the majority of criteria indicate that the optimal lag length is (3) so, we will estimate the ARDL model with lag length (t-3).

Estimation (ARDL) Model:

using the optimal lag length (3) for all the variables, it was found that the best rank for the model was the second difference for all variables, meaning that the model would be of the type ARDL (2,2,2,2,2). This model achieves the lowest value of the Akaiki Information Standard (AIC), which is (0.91) among the other models, as shown in Figure (3).

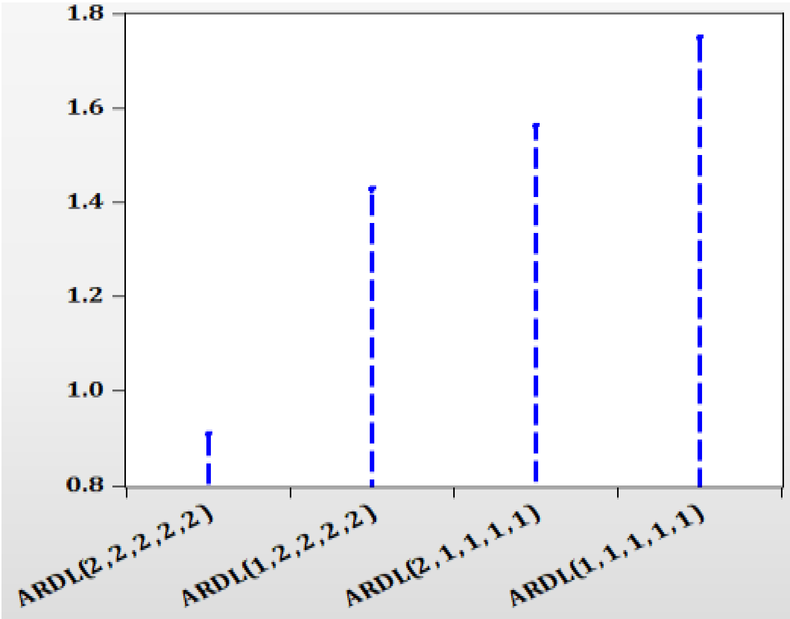

Figure (3): Akaiki Values for ARDL Models with Different Lags for the Income Share Held by Lowest 40% Model

Source: authors' work/ EViews-10 program outputs.

So, ARDL (2,2,2,2,2) model to be estimated will take the following form:

$$\Delta Y_{3,t} = \beta_{1,0}\Delta X_{1,t} + \beta_{2,0}\Delta X_{2,t} + \beta_{3,0}\Delta X_{3,t} + \beta_{4,0}\Delta X_{4,t} + \beta_{1,1}\Delta X_{1,t-1} + \beta_{2,1}\Delta X_{2,t-1}$$
$$+\beta_{1,1}\Delta X_{1,t-1} + \beta_{2,1}\Delta X_{2,t-1} + \beta_{3,1}\Delta X_{3,t-1} + \beta_{4,1}\Delta X_{4,t-1} + \alpha_1\Delta Y_{3,t-1} + \lambda Y_{3,t-1}$$
$$+\delta_1 X_{1,t-1} + \delta_2 X_{2,t-1} + \delta_3 X_{3,t-1} + \delta_4 X_{4,t-1} + U_{4,t} \qquad \ldots\ldots(10)$$

Table (7) displays the model estimation results.

Table (7): ARDL (2, 2, 2, 2, 2) Estimation Results of the Lowest 40% Model

Dependent Variable: ΔY_3				
Method: ARDL				
Sample: 2005 2017				
Included observations: 312				
Maximum dependent lags: 2 (Automatic selection)				
Dynamic regressors (2 Lag, automatic): X_1 X_2 X_3 X_4				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Long-run Equation				
$X_{1,t}$	0.0887	0.0191	4.6405**	0.000
$X_{2,t}$	2.5851	0.0198	130.64**	0.000
$X_{3,t}$	3.3288	0.1141	29.173**	0.000
$X_{4,t}$	-0.1549	0.0310	-5.0004**	0.000
Short-run Equation				
CointEq01	-0.1943	0.0792	-2.4546*	0.016
$\Delta Y_{3,t-1}$	-0.2165	0.1119	-1.9350$^{n.s}$	0.055
$\Delta X_{1,t}$	0.3162	0.3038	1.0407$^{n.s}$	0.300
$\Delta X_{1,t-1}$	-0.0484	0.1694	-0.2857$^{n.s}$	0.776
$\Delta X_{2,t}$	-0.1808	0.7957	-0.2272$^{n.s}$	0.821
$\Delta X_{2,t-1}$	-0.0490	0.4684	-0.1047$^{n.s}$	0.917
$\Delta X_{3,t}$	30.894	10.8589	2.8450**	0.005
$\Delta X_{3,t-1}$	-27.008	8.8234	-3.0610**	0.003
$\Delta X_{4,t}$	0.0405	0.0228	1.7772$^{n.s}$	0.078
$\Delta X_{4,t-1}$	0.0208	0.0234	0.8882$^{n.s}$	0.376
Mean dependent var	0.1554	S.D. dependent var		0.6686
S.E. of regression	0.4202	Akaike info criterion		0.7853
Sum squared resid	20.482	Schwarz criterion		3.4192
Log likelihood	102.646	Hannan-Quinn criter.		1.8326
R-squared	0.9859			
Adjusted R-squared	0.9857			
F-statistic	1232.375**			
Prob(F-statistic)	0.000			
** Significant at 1% level * Significant at 5% level n.s not significant				

Source: authors' work/ EViews-10 program outputs.

Short-run and Long-run Equations:

The long-run equation takes the following form:

$$\Delta Y_{3,t} = 0.0887X_{1,t} + 2.5851X_{2,t} + 3.3288X_{3,t} - 0.1549X_{4,t}$$
$$\Rightarrow Y_{3,t} = Y_{3,t-1} + 0.0887X_{1,t} + 2.5851X_{2,t} + 3.3288X_{3,t} - 0.1549X_{4,t} \quad \ldots\ldots.(11)$$

While the short-run equation takes the following form:

$$\Delta Y_{3,t} = -0.1943CointEq01 - 0.2165\Delta Y_{3,t-1} + 0.3162\Delta X_{1,t} - 0.0484\Delta X_{1,t-1}$$
$$-0.1808\Delta X_{2,t} - 0.049\Delta X_{2,t-1} + 30.894\Delta X_{3,t} - 27.008\Delta X_{3,t-1} + 0.0405\Delta X_{4,t}$$
$$+0.0208\Delta X_{4,t-1} \quad \ldots\ldots\ldots.(12)$$

The results of short-run and long-run estimation can be summarized as follows: In the short-run education has no significant effect on income share held by the lowest 40%. In the long-run, we found a significant positive effect of education on Income share of this class at (1%). an increase in the average years of education by one year leads to an increase income share of this class by (2.59%). So, with more education this class will be able to access jobs with greater financial returns and their income share will increase.

While it has not proven a significant impact of unemployment on the income share held by the lowest 40% in the short-run. In the long-run, unemployment rate has positive effect on the income shar held by the poor class. An increase in unemployment rate by (1%) leads to an increase in the income share of this class by (0.09%).

In the short-run, it is noted that the population growth has positive effect on the income share of the lowest 40%. The annual increase in three consecutive years by (1%) leads to an increase income share held by the poor class by 4% (31%-27% = 4%). In the long-run, population growth has a significant positive effect on the Income share held by lowest 40% at (1%), so, an increase in the population growth rate by (1%) leads to an increase in the income share of this class by (3.33%).

Per capita GDP did not appear as significant explainer of the poor class share of income in the short-run. In the long-run, average per capita GDP contributes to inequality, as it was found that it affects significantly and adversely the income share of the poor class at (5%), so, an increase in per capita GDP by (1%) decrease the income share of this class by (0.15%). This result supports the view that the benefits of economic development do not go to the poor class in developing countries.

Although co-integration coefficient (CointEq01) is significant at (1%), it does not indicate the existence of a long-run co-integration relationship between the model variables, because all the variables were originally static at I(0).

Overall Model Rating:

The results of table (8) show that the ARDL (2, 2, 2, 2, 2) estimation model takes the following form:

$$\begin{aligned}\Delta Y_{3,t} &= 0.0887X_{1,t} + 2.5851X_{2,t} + 3.3288X_{3,t} - 0.1549X_{4,t} - 0.1943\mathrm{CointEq01}\\ &-0.2165\Delta Y_{3,t-1} + 0.3162\Delta X_{1,t} - 0.0484\Delta X_{1,t-1} - 0.1808\Delta X_{2,t} - 0.049\Delta X_{2,t-1}\\ &+30.894\Delta X_{3,t} - 27.008\Delta X_{3,t-1} + 0.0405\Delta X_{4,t} + 0.0208\Delta X_{4,t-1} \quad \ldots (13)\end{aligned}$$

$\bar{R}^2$ shows that the explanatory variables explain (99%) of the changes in the income share held by the lowest 40%. F-test reflects that ARDL (2, 2, 2, 2, 2) is significant at 1%.

5. **Conclusion and Policy Implication:** Most of the explanatory variables, including the level of education, do not, in the short-run, significantly affect the way income is distributed among the three income groups (on the 20%, the middle 40%, and the lowest 40%). The reason for this is that education takes a long time before its effect is reflected in reality. This is consistent with the research hypothesis.

In the long-run, education plays an important distributive role in developing countries. Increasing the average years of education by one year reduces the share of the rich class by (0.819%), increases the share of the middle class by (1.0081%), and increase the share of the poorest class by (2.5851%). This result suggests that the greatest beneficiary of education is the poorest classes in society, as more education enable them obtaining jobs that generate higher incomes, which increases their income share. This result confirms the validity of the research hypothesis.

Increasing per capita GDP leads to an increase in income maldistribution in developing countries. It increases the share of the rich class, while reducing the share of the middle class, and the share of the poor class. The share of the poor depends a lot on the elasticity of employment to output growth. The output growth using capital-intensive techniques enhances productivity, but lowers the flexibility of employment, which leads to the distribution of the fruits of development to the disadvantage of the poor. This reinforces the idea that economic growth in developing countries is not pro-poor and its fruits go to the rich owners and capitalists. (Nallari R. and B. Griffith, 2011, 272)

Unemployment in the long-run reduces the share of the middle class, while it positively affects the share of the rich and the poor. The reason for this is that the poor receive unemployment benefits from the government. They engage in work within the informal sector when they are in unemployment. while, the middle class does not want to work in the informal sector.

Population growth positively affects the share of the highest class, while it negatively affects the share of the middle class. the lowest class is positively affected by population growth, both in the short-run and long-run. The reason for the increase in the share of the poor is that children contribute to the labor force and achieve incomes that are added to the family income. At the same time, their costs to poor families are low due to their low enrollment rate, and the modesty of their living requirements.

There are some development policy considerations that can be drawn from these results.

❖ Increasing the average years of education should be given great attention, through long-run education programs.
❖ Changing development policies to direct the fruits of development towards the poor.
❖ Addressing unemployment is one of the main keys that effectively contribute to improving the share of the poor and middle classes in society. Therefore, policies aimed at addressing unemployment are pivotal in alleviating the large disparity in income distribution.
❖ Population growth is another variable that effectively affect income distribution. So, continuing population policies aimed at regulating fertility is one of the keys of income redistribution.
❖ Countries differ in their attitudes towards the way income is distributed. Therefore, if the goal is to improve the level of the poorest class, the appropriate policies should focus more on education, address unemployment, and target this class in directed the fruits of development, by adopting policies that increase employment flexibility with increased output growth. But if the goal is to improve the share of the middle class, the policies should extend, in addition to education, address unemployment, and employment flexibility, to address population growth that negatively affects the income share of this class.

References:

1. Abdullah A., et al., (2011), "Does education reduce income inequality? a meta-regression analysis," Journal of Economic Surveys; 29, pp. 301-316. https://www.hendrix.edu/uploadedFiles/Departments_and_Programs/Business_and_Economics/AMAES/Education%20and%20Inequality%20Meta%20Analysis%20JABBAR.pdf
2. Castelló-Climent A., R. Doménech, (2014), "Human Capital and Income Inequality: Some Facts and Some Puzzles," BBVA Working Paper, No. 12-28, pp. 1-39. https://www.bbvaresearch.com/wp-content/uploads/migrados/WP_1228_tcm348-430101.pdf
3. Digdowiseiso K., (2009), "Education inequality, economic growth, and income inequality: Evidence from Indonesia, 1996-2005," MPRA, pp. 1-19: https://mpra.ub.uni-muenchen.de/17792/1/Economic_Growth_and_Inequality-Kumba.pdf
4. Jeng R., et al, (2019), "Effect of Education on Income Inequality: A Cross-National Study," pp. 1-17: https://smartech.gatech.edu/bitstream/handle/1853/62056/Effect%20of%20Education%20on%20Income%20Inequality%20A%20Cross-National%20Study.pdf?sequence=1&isAllowed=y
5. Lee J. and H. Lee, (2018), "Human capital and income inequality," ADBI Working Paper, No. 810. https://www.adb.org/sites/default/files/publication/401466/adbi-wp810.pdf
6. Nallari R. and B. Griffith, (2011), Understanding Growth and Poverty, The World Bank.http://documents.worldbank.org/curated/en/644851468325258956/Understanding-growth-and-poverty-theory-policy-and-empirics
7. Our World in Data website. https://ourworldindata.org/global-education#years-of-schooling
8. Ozturk I., (2008), "The Role of Education in Economic Development: A Theoretical Perspective," Journal of Rural Development and Administration, Volume XXXIII, No. 1, pp. 39-47: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1137541_code568270.pdf?abstractid=1137541&mirid=1
9. Petcu C., (2014), "Does Educational Inequality Explain Income Inequality Across Countries?" Honors Projects. Paper 125. pp. 1-37: https://core.ac.uk/download/pdf/59230033.pdf
10. Roy P. and Z. Husain, (2019), "Education as a way to reducing inequality: Evidence from India," MPRA, pp. 1-28. https://mpra.ub.uni-muenchen.de/93907/
11. Shahabadi A., (2018), "The Effect of Education on Income Inequality in Selected Islamic Countries," IJAPS, Vol. 14, No. 2, pp. 61–78: https://www.researchgate.net/publication/326409551_The_Effect_of_Education_on_Income_Inequality_in_Selected_Islamic_Countries
12. Todaro M., and, S. Smith, (2015), Economic Development, 12th edition, REARSON.
13. World Bank, (2019), World Development Indicators, Online Data.