Evaluating the Impact of Education on Earnings in Nepal

Hari Gopal Risal
Assistant Professor – Finance and Account
Kathmandu University School of Management
Lalitpur, Nepal
Dr. C.P. Gupta
Professor
Department of Financial Studies
South Campus, University of Delhi
Delhi, India

Abstract:
This study will contribute to the literature of return to education and help policy maker while allocating the budget in education since education level has caused significant increment in earnings in context of Nepal. This paper has applied a Semi-log multivariate regression model using data from Nepal Labor Force Survey to provide empirical evidence in the entire analysis on this issue. Not just academic qualifications, but also there are several other factors which cause increase in income. We explore this fact to identify the impact of experience, age, training, gender and other variables in their earnings. The empirical evidence of this study has shown a direct relationship between the education and earnings (i.e. Increase in education by 1 level causes 16% increment in their wages on average). Apart from the academic qualification, their experiences in the similar field have statistically significant relationship (i.e. 1 level increase in experience causes 11.57% increase in log monthly earnings). The training dummy is found to be statistically insignificant to cause increment in earnings which is also an interesting finding of the study. Due to having the low R-squared value after introducing dummies, we can suggest that there may be other qualitative factors like ability of individual and the contextual fact that may cause the increment in earnings, which can be the future scope of this study. Two major implications of this empirical study are: firstly, this will help policymakers while allocating the budget in education sector and secondly, encourage people towards higher education.
Keywords: secondary data; multivariate regression; inferential analysis; monthly earnings and education.

Introduction
'Education is the most powerful weapon'- Nelson Mandela; it is considered as a most powerful weapon in the world and gives a symptom of civilization and development. It is also considered as a most needed assets and basic requirement for most of the prominent jobs and to assume any role in the corporations. We can observe a massive investment in the education sector to make country more civilized and for the overall development. In Nepal also we can see mushrooming of schools (i.e in private sector) from kinder garden to higher secondary level, colleges and

universities which has increased the access to people throughout the country. Apart from the formal education, people require many other skill based trainings and ability to perform in the real world. It is considered that the level of education has direct relationship with their earnings and valued as the most important determinants of their wage and other facilities. Numerous existing studies in both developed and developing countries have shown that better-educated individuals earn higher wages, experience less unemployment, and work in better occupations (higher wages, greater job security, etc.) than their less-educated counterparts (Card, 1999). But it is very hard to claim that the higher earning is an outcome of higher education and the reverse may be true. It can also be argued that people with higher earnings choose to get more schooling. Human capital earning function (HCEF) in determination of wage is the remarkable contribution in of Mincer (1974). This study tries to identify the causality between education and the earnings in context of Nepal. The study have been divided into three different parts as: the first part is introduction which gives a general overview of the research and rationale for choosing this particular area and some of the reviews and their findings; the second part is the methodology and data analysis; and finally the findings and discussions have been presented. One major contribution of this study will be in terms of its data which varies from other countries in terms of availability and choice of schooling, date of joining school is not consistent in the country. Another is the model which is unique in its nature and validated using different diagnostic tests for the robustness. This study will not just contribute to the existing evidences but also contributes at individual level for the human capital development. It can be further argued that higher education reduces the cost of training and motivating people for better productivity which ultimately leads to the higher earnings and civilization. Another argument can be to help the policy makers in understanding the causality between education and earnings and decide on the type of education to avail and make necessary investments. If we were to check the budget allocated by government in education sector, it seems still a very low. Hence, the outcome of this research will be helpful for government to increase in education budget by many times which ultimately causes the increase in GDP and overall development of country. Just to present the fact a comparative table has been presented for last five years.

Table 1.1

Comparison of education budget with GDP and national budget

Headings	2010/11	2011/12	2012/13	2013/14	2014/15
GDP in NPR	13.75 trillion	15.40 trillion	17.06 trillion	19.67 trillion	20.75 trillion
National Budget	3.38 trillion	3.85 trillion	4.05 trillion	5.17 trillion	6.18 trillion
Edu. Budget	57.8 billion	63.9 billion	63.00 billion	80.5 billion	86.0 billion

Source: Red Book 2010/11 to 2014/15

Most of the empirical evidences have said the education and wages do have the direct relationship. For instance, Parajuli (1999), have also found that the returns to education in Nepal are 9.7%. Similarly, Lamichhane (2009) have found the estimated rate of returns to education is very high among persons with disabilities, ranging from 19.4 to 33.2% controlling for endogeneity bias arising from schooling decisions as well as sample selection bias due to endogenous labor participation. Our findings shows that the one level increase in education causes 16.64 % increase in their monthly earnings on an average. Hence, policy maker based on this findings can improve on their focus and common people can also benefit and continue with their education for both the learning and monetary gains.

Research Methodology and Data Analysis

2.1 DATA

The data which have been considered for this study have been extracted from the Nepal Labor Force Survey 2008 (II). NLFS II is the second round of a multi topic national labor force survey carried out by Central Bureau of Statistics (CBS) from January to December 2008 that covered nationally representative sample of 16000 households from 800 Primary Sampling Units (PSU) equally distributed equally distributed between urban and rural areas The nature of data used here is cross section and research interest is to check whether empirical evidences hold true in context of Nepal or not. Total number of observations is 5322 and seven regress or have been used to check their impact on log monthly earnings of individuals. Some relevant data have been captured from the red book and relevant information has been captured from different websites.

Table 2.1 Summary Statistics of Data

Variable	Observation	Mean	Std. Dev.	Min	Max
logME	5322	8.577571	.8168075	4.60517	16.1181
Education	5322	3.581924	1.816356	1	7
Experience	5322	2.803457	1.05179	1	4
AGE2	5322	1182.345	778.9195	256	4900
MARITALSTATUS	5322	.7643743	.4244291	0	1
TRAINING	5322	.2157084	.4113515	0	1
AREA	5322	.7136415	.4521014	0	1
GENDER	5322	.7375047	.4400317	0	1

2.2 METHOD

Semi-log model has been used for the regression analysis and diagnostic approach has been used to justify the appropriateness of the model. Log monthly wage is the dependent variable and level of education, work experience and age are the independent variables. Training, gender, area and marital status have been added further to address of Endogeniety caused by omitted variables which also improves the R-squared and solve the issue related with heteroskedasticity and autocorrelation. In addition, a comparative analysis has been performed using those dummy variables.

2.3 THE MODEL

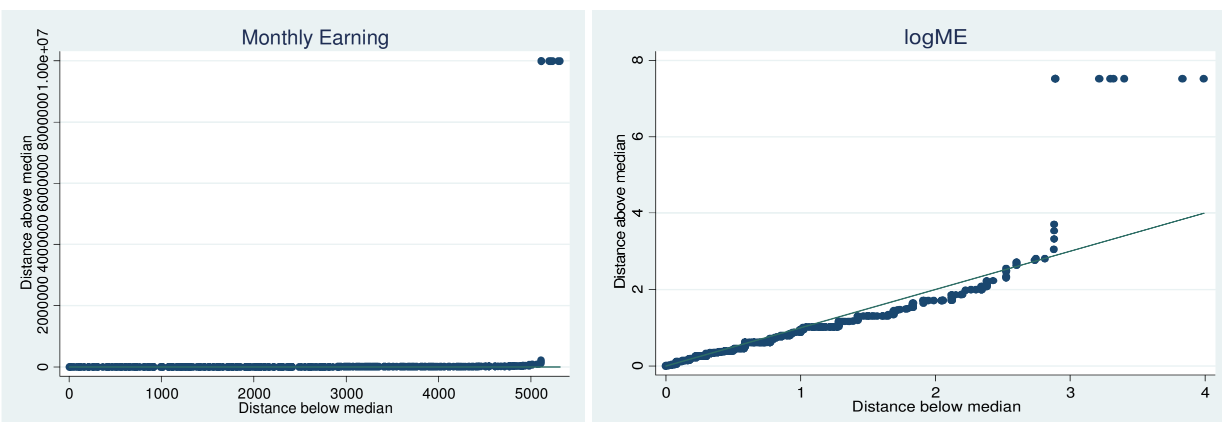
The Classical linear regression model (CLRM) is considered best linear unbiased estimate (BLUE) only when it is free from all biases stated in their assumptions. After passing all diagnonostic tests for multicollinearity, heteroskedasticity and specification biases, final and the robust model used for the purpose of study is as below:

$$\begin{aligned}\boldsymbol{Log\ Monthly\ Earning} &= \boldsymbol{\beta o + \beta_1\ Education + \beta_2\ Experience + \beta_3 AGE^2 + \beta_4 MSD_1} \\ &\quad \boldsymbol{+ \beta_5 GENDERD_2 + \beta_6 AREAD_3 + \beta_7 TRAININGD_4 + \varepsilon}\end{aligned}$$

Variable Explanation:

Log monthly earnings are the variable of research interest (dependent variable). The impact caused by several repressors on the log monthly earnings is captured using the above model. The wage and salaries payments in Nepal are practiced as monthly payments rather than weekly or some other patterns. The rationale behind using log monthly earning in place of monthly earning in the model can be explained with the help of symmetry plot and the histogram for skewness and kurtosis as:

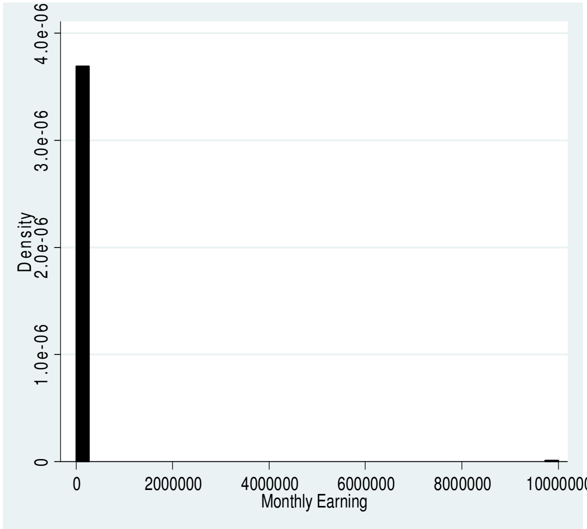
Symmetry Plot: Monthly Earning and Log Monthly Earning

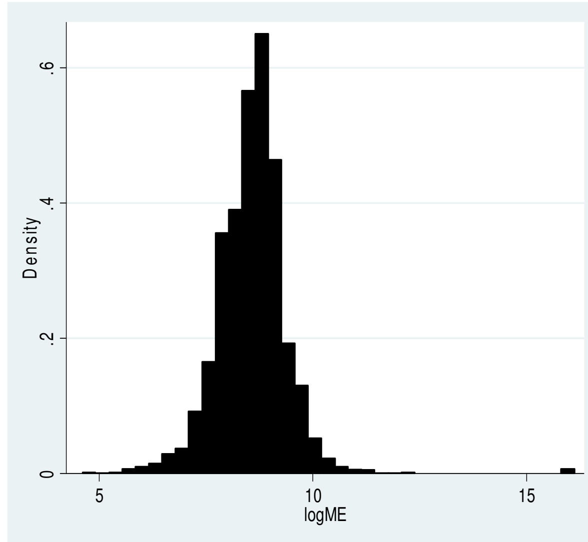


Above symmetry plot suggests for the use of log monthly earning in place of monthly earnings, which is flat and the log monthly earning seems to better explain the linear relations. This is further confirmed after checking the skewness and kurtosis below.

Histogram for Skewness and Kurtosis: Monthly Earning and Log Monthly Earning

As suggested by Damodar Gujarati in his book 'Econometrics by example' following histogram have been plotted to explain the use of log monthly earnings in place of monthly earnings. The histogram for monthly earning is highly skewed and do not express the normality of the distribution. Hence, the regressand used in the model Log Monthly Earning is justifiable.





EDUCATION

The main focus of this study is to establish the causal relationship between education and the log monthly earnings and to support the empirical evidences in context of Nepal. The education variable has been defined as education level starting from 1 to 7.

Defined Level of Education
Illiterate = 0;
Literate (Formal/Informal), Class 1, 2 and 3 = 1;
Class 4,5, 6 and 7 = 2;
Class 8,9 and 10 = 3;
SLC pass= 4;
Intermediate (10+2) = 5;
Professional and Bachelor’s degree = 6 and
Master’s and above 7

EXPERIENCE

The experience is also not in terms of years of work rather it is defined as four different level as:

Level “1” is assigned for work experience of less than 1 year; level “2” for 1 to 5 years of work experience; level “3” for 5 to 10 years of work experience and level “4” for experience above 10 years. This study shows the relationship between one level upgrade in experience level and their log monthly earnings.

AGE^2

The rationale for using AGE^2 in the model and not the AGE only is due to two reasons; the first one is As a priori of life cycle hypothesis, age may have positive correlation with earnings up to certain life and then starts falling its earning capacity due to their old age and here data considered are from 16 to 70 years; and the second one is AGE^2 explains the model better (has a higher predictive power) than the AGE only.Although data have been available beyond 16 years to 70 years, the permitted age for labor in the country is 16 and generally people after 70 years of age face several health issues and most of the people get retired from both the full-time as well as part-time and any other short term consulting jobs too. Hence, few exceptional cases have been ignored for the purpose.

DUMMIES:
As stated earlier, four dummies have been introduced with two basic purpose; first to improve the significance of the model and second to see the differences in earnings caused by gender, area, marital status and the impact of trainings.
MSD1 (Marital Status): 1 for married and 0 for all; it tests whether the married people earn more than that of the unmarried or divorced.
Gender D2 : 1 for Male and 0 for Female; the rationale behind introduction of this dummy is to check whether there exists difference in log monthly earnings between male and female (caused by the gender); if yes by how much male employee earn more than the female employee
AREA D3 : 1 for people in Urban and 0 for people in Rural; main interest behind this dummy is to see whether there exists any difference in monthly earnings of people who work in rural and the urban area; if yes by how much the people in urban earn more than that of the people in rural.
TRAININGD4: 1 for Trained and 0 for not trained; it tests whether the training cause significant impact in the log monthly earnings.
"ε"An idiosyncratic error term to account for unobserved heterogeneity.
It is the disturbance, or error term, or a random (stochastic) variable that has well-defined probabilistic properties. The disturbance term "ε" u may well represent all those factors that affect log monthly returns but are not taken into account explicitly this is the reason why we call them unexplained. Whatever is left unexplained in our model can be the scope for future research and explain more by reducing the **"ε".**

2.4 ROBUSTNESS OF THE MODEL
This part starts with an Ordinary least square method and continues with several approaches for diagnosis and robustness of the model.

Ordinary Least Square (OLS):

$$\boldsymbol{Log\ Monthly\ Earning = \beta o + \beta_1\ Education + \beta_2\ Experience + \beta_3\ AGE + \beta_4 AGE^2 + \varepsilon}$$

Where,

β_o = Constant

β_1,β_2,β_3, and β_4 are the coefficient of the independent variables Education, Experience and Age.

ε = An idiosyncratic error term to account for unobserved heterogeneity.

Table 2.2 Outcome of OLS

```
. reg logME Education Experience AGE AGE2

      Source |       SS       df       MS              Number of obs =    5322
-------------+------------------------------           F(  4,  5317) =  407.42
       Model |  832.826811     4  208.206703           Prob > F      =  0.0000
    Residual |  2717.20901  5317  .511041754           R-squared     =  0.2346
-------------+------------------------------           Adj R-squared =  0.2340
       Total |  3550.03582  5321  .667174557           Root MSE      =  .71487

------------------------------------------------------------------------------
       logME |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   Education |   .1633849   .0055086    29.66   0.000     .1525857     .174184
  Experience |   .1067142   .0118218     9.03   0.000     .0835386    .1298898
         AGE |   .0428462   .0057953     7.39   0.000     .0314851    .0542073
        AGE2 |  -.0004108   .0000762    -5.39   0.000    -.0005602   -.0002615
       _cons |   6.777515   .0933004    72.64   0.000     6.594608    6.960423
------------------------------------------------------------------------------
```

From the above table, we can see that the all selected variables are significant at 5%. As per 2-t rule of thumb, all individual t-values are above 2 indicates it to be statistically significant. Further to this, all P>I t I = 0 shows it to be highly significant any 5% and 1% both. The only problem here with the model can be observed with the R-squared which is only 0.2267 which is further tested and improved.

TEST FOR MULTICOLLINEARITY

In this part, to test whether there exists and collinearity issues among the selected independent variables following two tests have been performed: Firstly, the correlation was checked and secondly, VIF and the Tolerance test have been performed.

Correlation among Coefficients

Table 2.3 Correlation Matrix

```
. estat vce, correlation

Correlation matrix of coefficients of regress model
```

e(V)	Educat~n	Experi~e	AGE	AGE2	_cons
Education	1.0000				
Experience	0.1155	1.0000			
AGE	-0.1694	-0.3333	1.0000		
AGE2	0.1443	0.2240	-0.9801	1.0000	
_cons	-0.0477	0.0812	-0.9312	0.9157	1.0000

The variable Education and Experience have positive correlation but very low which shows almost no correlation between them and it is found to be negative correlation between age and education, age and experience. This facts gives us a signal that there is no positive collenearity among the variables of the model. However, the negative high correlation between experience and AGE2 suggests for further check-ups using VIF and tolerance to confirm the collenearity issue.

VIF and Tolerance Level test

As per the definition of Wikipedia "The Variance inflation factor (VIF) quantifies the severity of multicollinearity in an ordinary least square (OLS) analysis. It provides an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of co linearity."

Table 2.4 Test for Multicollinearity

```
. estat vif

    Variable |       VIF       1/VIF
-------------+----------------------
         AGE |     39.36    0.025408
        AGE2 |     36.66    0.027274
  Experience |      1.61    0.621208
   Education |      1.04    0.959349
-------------+----------------------
    Mean VIF |     19.67
```

It is considered that the VIF above 10 means very high multocollinearity as a rule of thumb and 4 is also considered not very well. Here in our case, maximum VIF is far above the acceptable level in case of Age and it's squared. Hence, it is suggestive of dropping Age and consider the squared value only. The rationale behind dropping the Age and continuing with it squared value is due to retirement age policy and reduction in income after retirement. The tolerance level test is simply 1 divided by VIF and considered to be not collinear when it is close to 1. The result of VIF and tolerance level are the same.

Drop the Age and check for Multicollinearity:

Dropping the AGE from the model, we can regress and check whether it solves the multicollinearity issue or not and test for VIF and tolerance level.

Table 2.5 Test for Multicollinearity after dropping AGE

. reg logME Education Experience AGE2

Source	SS	df	MS
Model	804.892685	3	268.297562
Residual	2745.14313	5318	.516198408
Total	3550.03582	5321	.667174557

Number of obs =	5322
F(3, 5318) =	519.76
Prob > F =	0.0000
R-squared =	0.2267
Adj R-squared =	0.2263
Root MSE =	.71847

logME	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Education	.1702834	.0054563	31.21	0.000	.1595868	.1809801
Experience	.1358441	.011202	12.13	0.000	.1138836	.1578046
AGE2	.0001412	.0000152	9.30	0.000	.0001114	.000171
_cons	7.419834	.0341879	217.03	0.000	7.352812	7.486856

. estat vif

Variable	VIF	1/VIF
AGE2	1.44	0.693031
Experience	1.43	0.698834
Education	1.01	0.987687
Mean VIF	1.30	

In the revised model, maximum VIF is 1.44 which confirms no issue related to multicollinearity. This way, we can claim it to be free from the multicollinearity issue.

MODEL SPECIFICATION TEST

There may exist the specification error when any of the independent variables is correlated with the error term. There are several causes for this specification error or bias; popular types are due to incorrect functional form, omitted variables, inclusion of irrelevant variables, simultaneity and measurement error. Here, two popular types of specification tests have been performed; the Ramsey RESET test and the linktest.

Ramsey RESET Test

It tests whether non-linear combinations of the fitted value help explain the response variable. Hence, it suggests whether non-linear functional form have the explanatory power and tells about the omitted variable.

Table 2.6 Ramsey Reset Test for Model Specification

```
. estat ovtest, rhs
(note:  Experience^3 dropped because of collinearity)

Ramsey RESET test using powers of the independent variables
       Ho:  model has no omitted variables
                F(8, 5310) =     16.15
                  Prob > F =      0.0000
```

Above test results Prob>F= 0.0000 shows that the null hypothesis is rejected which means the model has explanatory power and passes the specification test. However, it is not the single tool to check for model specification. Hence, we proceed further to other tests before the confirmation.

Link Test

This is another very popular test to check the model specification. The Probability > F = 0.0000 and p>I t I = 0.0000 suggests that the model is appropriately specified.

Table 2.7 Link Test for Model Specification

```
. linktest

      Source |       SS       df       MS              Number of obs =    5322
-------------+------------------------------           F(  2,  5319) =  813.45
       Model |  831.505679     2  415.752839           Prob > F      =  0.0000
    Residual |  2718.53014  5319  .511097977           R-squared     =  0.2342
-------------+------------------------------           Adj R-squared =  0.2339
       Total |  3550.03582  5321  .667174557           Root MSE      =  .71491

------------------------------------------------------------------------------
       logME |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _hat |  -5.813012   .9444941    -6.15   0.000    -7.664608   -3.961416
      _hatsq |    .394423   .0546598     7.22   0.000     .2872675    .5015786
       _cons |   29.35988   4.074486     7.21   0.000     21.37221    37.34754
------------------------------------------------------------------------------
```

Based on the above results (i.e.; All Significant at 5% and 1%) of Ramsey RESET test and link test, we get the clean cheat from model specification test and continue further with our model for regression. However, a significant value of t implies that the model may have omitted variable bias. Hence, introduction of dummies will help to make the model robust.

INTRODUCING DUMMY VARIABLES

As of now, the model as per theory seems to be appropriate, valid and no omitted variables. However, the low R-squared value observed so far is not very encouraging. Based on theory, we are now introducing some extra variable as a dummy which may contribute in the r-squared and also check whether there exists any difference in monthly earnings between male and female (Gender Dummy); is there any difference in earnings of people in Urban and Rural (Area Dummy); whether married people earn more than unmarried and divorced (Marital status Dummy) and impact of trainings on earnings (Training Dummy). After addition, the new model looks like:

$$\begin{aligned} \boldsymbol{Log\ Monthly\ Earning} \\ = \boldsymbol{\beta o} + \boldsymbol{\beta_1\ Education} + \boldsymbol{\beta_2\ Experience} + \boldsymbol{\beta_3 AGE^2} + \boldsymbol{\beta_4 MSD_1} \\ + \boldsymbol{\beta_5 GENDERD_2} + \boldsymbol{\beta_6 AREAD_3} + \boldsymbol{\beta_7 TRAININGD_4} + \boldsymbol{\varepsilon} \end{aligned}$$

Table 2.8 OLS with added dummies

```
. reg logME Education Experience AGE2 MARITALSTATUS TRAINING AREA GENDER
```

Source	SS	df	MS
Model	887.290981	7	126.755854
Residual	2662.74484	5314	.501081076
Total	3550.03582	5321	.667174557

Number of obs	= 5322
F(7, 5314)	= 252.96
Prob > F	= 0.0000
R-squared	= 0.2499
Adj R-squared	= 0.2490
Root MSE	= .70787

logME	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Education	.16638	.0057227	29.07	0.000	.1551613	.1775988
Experience	.1157055	.0113282	10.21	0.000	.0934976	.1379134
AGE2	.0001094	.0000155	7.07	0.000	.0000791	.0001397
MARITALSTATUS	.140537	.0256152	5.49	0.000	.0903207	.1907533
TRAINING	.0079531	.0246037	0.32	0.747	-.0402803	.0561866
AREA	.1291996	.0222512	5.81	0.000	.0855781	.1728212
GENDER	.2107579	.0224665	9.38	0.000	.1667144	.2548014
_cons	7.171126	.0390372	183.70	0.000	7.094598	7.247655

Above results (i.e., the P>I t I =0.0000) shows all variables to be significant at both the 5% and 1% level of significance except training. Introduction of dummies have also introduced the R-squared to 0.2499 from the 0.2267. The reason for training dummy not found significant maybe due to the definition of training. It is not stated well in the data about the duration of the training, its quality and other qualitative characteristics.

DROP the TRAINING??? No contribution!!!

Table 2.9 OLS with added dummies after dropping the Training

```
. reg logME Education Experience AGE2 MARITALSTATUS AREA GENDER

      Source |       SS       df       MS              Number of obs =    5322
-------------+------------------------------           F(  6,  5315) =  295.16
       Model |  887.238623     6  147.873104           Prob > F      =  0.0000
    Residual |   2662.7972  5315   .50099665           R-squared     =  0.2499
-------------+------------------------------           Adj R-squared =  0.2491
       Total |  3550.03582  5321  .667174557           Root MSE      =  .70781

------------------------------------------------------------------------------
       logME |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   Education |   .1668104   .0055652    29.97   0.000     .1559003    .1777205
  Experience |   .1156348   .0113251    10.21   0.000     .0934329    .1378367
        AGE2 |   .0001092   .0000155     7.06   0.000     .0000789    .0001395
MARITALSTATUS|   .1399808   .0255552     5.48   0.000     .0898821    .1900794
        AREA |   .1295187   .0222274     5.83   0.000     .0859438    .1730936
      GENDER |   .2104463   .0224439     9.38   0.000      .166447    .2544455
       _cons |   7.172168   .0389007   184.37   0.000     7.095907    7.248429
------------------------------------------------------------------------------
```

Now, it is an interesting question for researcher on whether to drop the Training dummy or, not. As a priori, training must cause some impact on the skill and ability and should have a positive impact on the monthly earnings. Although it increases the F statistic substantially when dropped the training dummy does not contribute to improve the R-squared value. The R-squared value with and without training dummy model is same. Despite the statistical findings of training dummy being insignificant, it has been continued as per the theoretical understanding.

HETEROSKEDASTICITY TEST

This test is performed to know whether the variability of variables is equal or, unequal across the range of values. Our null hypothesis (Ho) here is that there exists Homoskedasticity (i.e., constant variances); which is also a basic requirement to run the OLS. For the purpose, we have considered the BP test and the white test.

Breusch Pagan Test

This test model was developed in 1979 by Trevor Breusch and Adrian Pagan and further contributed by cook and Weisberg in 1983. The purpose of BP test is to check whether the estimated variance of the residuals from a regression are dependent on the values of the independent variables. In that case, heteroskedasticity is present- Wikipedia. This tests fall under the category of chi^2 test.

Table 2.10 BP Test for Heteroskedasticity

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of logME

         chi2(1)      =    71.81
         Prob > chi2  =   0.0000
```

Based on above calculation, we reject the null hypothesis and confirm that there is no heteroskedasticity issue and we can continue with the model.

White Test

This is a major contribution of Halbert white in 1980 and his paper became one of the most cited one in the field of economics. The white test doesn't just test the heteroskedasticity but also tests for specification error.

Table 2.11 White Test for Heteroskedasticity and Specification Error

```
. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(31)     =     45.86
         Prob > chi2  =    0.0417

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	45.86	31	0.0417
Skewness	38.68	7	0.0000
Kurtosis	13.52	1	0.0002
Total	98.06	39	0.0000

Here the Prob> chi2 =0.0417 which is significant at 5% confirms that there is no issue of heteroskedasticity. Hence, both the BP test and White test passes the test for heteroskedasticity. Hence, the model considered for this study seems to be robust.

2.5 Regression with the Final Model

$$\textit{Log Monthly Earning} = \beta o + \beta_1 \textit{Education} + \beta_2 \textit{Experience} + \beta_3 AGE^2 + \beta_4 MSD_1 + \beta_5 GENDERD_2 + \beta_6 AREAD_3 + \beta_7 TRAININGD_4 + \varepsilon$$

All the required test have been performed to make the model robust and BLUE, the regression model gives us the following outcome

Table 2.12 OLS with the Robust Model

```
. reg logME Education Experience AGE2 MARITALSTATUS TRAINING AREA GENDER

      Source |       SS       df       MS              Number of obs =    5322
-------------+------------------------------           F(  7,  5314) =  252.96
       Model |  887.290981     7  126.755854           Prob > F      =  0.0000
    Residual |  2662.74484  5314  .501081076           R-squared     =  0.2499
-------------+------------------------------           Adj R-squared =  0.2490
       Total |  3550.03582  5321  .667174557           Root MSE      =  .70787
```

logME	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Education	.16638	.0057227	29.07	0.000*	.1551613	.1775988
Experience	.1157055	.0113282	10.21	0.000*	.0934976	.1379134
AGE^2	.0001094	.0000155	7.07	0.000*	.0000791	.0001397
MARITALSTATUS	.140537	.0256152	5.49	0.000*	.0903207	.1907533
TRAINING	.0079531	.0246037	0.32	0.747	-.0402803	.0561866
AREA	.1291996	.0222512	5.81	0.000*	.0855781	.1728212
GENDER	.2107579	.0224665	9.38	0.000*	.1667144	.2548014
_cons	7.171126	.0390372	183.70	0.000*	7.094598	7.247655

*significant at 1% level of significance.

3. Results And Discussions

3.1.1 Major Findings

The fitted model has successfully established the causality between education and the earnings. Following are the key findings of the study:

One level increase in education causes 16.64 % increase in their monthly earnings on an average.One additional level increase in Experience causes 11.57% increase in their monthly earnings on an average.Age has nothing much to do with their monthly earnings. It may be identified only if we categorize the age and continue the process. Female employee on an average earns 21% less than the male employee.People in rural on an average earn 13% less than the people in Urban.Trained employee get 0.8% more than non-trained on an average.Married people are earning 14% more than unmarried. It may be due to some qualitative factors like family commitments, extra time, experiences, trainings etc.

3.1.2 Conclusion and Discussions

Based on the large data set with a final model derived after conducting diagnostic test for their robustness, the findings suggest for a massive investment need in the education to foster the economic growth of the country. This study gives a contradictory findings in case of relationship between age and the earnings. Most of the empirical evidences have shown a direct relationship between the age and the education but the negligible coefficient of age shows no relationship with the earnings. Similarly, the lower earnings of female suggests for some training and female-friendly work policy or the differential wage system. People in rural area earning less than the urban is justifiable as per the higher expenditure requirements; comparatively higher school fee of their children, house rents and price discriminations. This study also captures the qualitative fact of earning behavior among married and unmarried or divorced people. It may be further studied controlling their work hours and commitments. Countries having similar level of development stage can also use the method used in this study to test the causality between education, experience and age with the earnings while forming policy related to wage and the education.

References:

Ashenfelter, O. and Card, D., Handbook of Labor Economics. Rotterdam: Elsevier.

Ashenfelter, O. Harmon, C. & Oosterbeek, H. (1999). A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias." Labour Economics, 6 (3), 453–470.

Becker, G.S. (1962). Investment in Human Capital: A Theoretical Analysis. The Journal of Political Economy, 70 (5), 9–49.

Blundell, R, Dearden, L and Sianesi B. (2004) Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS; THE INSTITUTE FOR FISCAL STUDIES, WP03/20.

Card, D. (1999). The Causal Effect of Education on Earnings.

Central Bureau of Statistics, National Planning Commission Secretariat (2008).

Chevalier, A. (2011). Subject Choice and Earning of UK Graduates. Economics of Education Review, this issue.

Colm Harmon, Economic Returns to Education: What We Know, What We Don't Know, and Where We Are Going – Some Brief Pointers; IZA Policy Paper No. 29

Dee, T.S. (2004). Are There Civic Returns to Education?" Journal of Public Economics 88 (9- 10), 1697–1720.

Gujarati, D. (2011) Econometrics by Example, Plagrave Macmillan

Lamichhane K and Sawada Y (2009), Disability and Returns to Education in a Developing Country, The University of Tokyo

Mincer, Jacob; Polachek, Solomon (March–April 1974). "Family investments in human capital: earnings of women". Journal of Political Economy

Nepal Labor Force Survey 2008 Statistical Report, Government of Nepal.

Parajuli, D. 1999. "External Efficiency and Equity in Nepalese Education." World Bank(mimeo).