Can education be a solution to the unemployment?

Muhammad Ramzan *, Muhammad Abbas, Saba Aslam, Umar Farooq

Air University Islamabad, Multan Campus, Pakistan

Abstract: This study is conducted to determine the impact of education and some economic factors on unemployment in the context of Pakistan taking data from 1972 to 2012. Unemployment is regressed by wages, education, GDP, inflation, money supply and population in a multivariate regression equation. Tests like VAR, VECM, ARDL, CUSUM and Ramsey also confirmed the relationship among unemployment and macroeconomic variables both in short and long run. Increase in population and education reduces the unemployment. Unemployment has a negative relationship with wages. Unemployment in the past affects the present and future unemployment as well so the Institutional policies have to be made carefully to solve the problem of unemployment.

Key words: Education; Unemployment; Wages; Inflation; Population; GDP; ARDL Test

1. Introduction

Unemployment is a situation that arises when there are people willing to work and pursuing it at the prevailing market rate but cannot acquire the aforesaid (Economic Survey of Pakistan, C12, 2012). It is a socioeconomic phenomenon that brings various harms to society. Unemployment can be defined and measured through narrow and wider approaches. Narrow unemployment approach refers exclusively to those people who are unemployed and seeking the work but unable to obtain it. Broader term includes all the eligible workers irrespective of their efforts to search for employment (Eita & Ashipala, 2010).

Keeping all the unemployment states in mind as defined by O’Sullivan and Sheffrin (2007), one can safely say that countries are either developed or developing; all of them face the problem of unemployment to the varying degrees. Since, the structure of every economy is different due to its resources and unique circumstances, thus in different countries, same economic variables produce diverse effect on unemployment. Pakistan is one of many developing countries that are facing the problem of high population growth alongside enlarged labor force. But there are lesser employment prospects while other factors of economy are also evolving which gives rise to socioeconomic disorders like unemployment. Observing the whole unemployment trend with a labor growing not only in numbers but in skills as well and different requirements in market is something that can help to understand; just wages or education attainment is not enough to reduce unemployment or only growth or inflation impact are not enough to understand the stimulants of unemployment. Rest of the article is comprised of: section 2 includes literature review, section 3 represents the research methodology, section 4 entails analysis and section 5 articulates the conclusion and recommendations.

2. Literature review

Unemployment has been studied for almost over a century by scholars of different origins with the help of multiple economic variables. Economic Growth, Inflation, Education, population, and wages have been the elements of utmost importance that contributed in unemployment regulation and study over a long period. Different institutional policies impacts along with wages policy change were also studied. Studies conducted by Irfan, (1982), Bassanini and Duval (2006) and Diamond, (2011) gave a blend of results where institutions have significant power to regulate the unemployment either positively or negatively. Some studies’ results show that wages increase was cause of increase in unemployment (Kooros, 2006; Fatima & Sultana, 2009) or its increase was not much significant (Phillips, 1958).

In some developed and developing countries unemployment was studied with macroeconomic variables such as GDP, inflation, exchange rate, etc. The variables proved significant impact on unemployment increase or decrease according to the economic structure of country (Aktar & Ozturk 2009; Doğan, 2012; Sheikh & Iqbal 1992; Mensah, Awunyo-Vitor & Asare-Menako, 2013; Roberts, 1997; Choudhry, Marelli & Signorelli, 2012). Education nexus with unemployment was also interesting study that yielded positive correlation between education levels and employment for a mixture of European countries at one time (Weber, 2002). Then in some studies, it had no significant effect on developing and developed economies alike (Hujer, Rodrigues & Wolf, 2009; Das, 1981; Amimo, 2012).
Some studies were amalgamation of two dimensions like education and wages, education and macroeconomic variables and wages and macroeconomic variables to determine their combined effect on unemployment. Kooros (2006), Goh and Wong (2010) and Muhammad, Sa’idu, Nwokobia and Yakubu (2013) conducted the macroeconomic variables and wages impact studies and found that variables like inflation and growth have effect on wages along with unemployment. Das (1981) conducted a multi country study with educated unemployed labor force and some major economic variables like GDP and GNP to understand the returns of education and employment opportunities to be availed. He found various socioeconomic and management issues like proper regulation.

3. Research methodology

There are various studies that have proposed the impact of macroeconomic variables on wages, education and on unemployment. Several of them used an ordinary least square method (Muhammad et al., 2013; Valadkhani, 2003; Eita & Ashipala, 2010; Sodipe & Ogunrinola, 2011). Here similar approach is adopted by producing an OLS equation. Considering the functional form:

$$Un = f(W, EDU, GDP, INF, M2, POP)$$  \hspace{1cm} (3.1)

Where:

- $Un$ = Unemployment Rate
- $W$ = Wages
- $EDU$ = Education
- $GDP$ = Gross Domestic Product
- $INF$ = Inflation
- $M2$ = Money Supply
- $POP$ = Population

The Specified model form is given below:

$$Un = \beta_0 + \beta_1(W) + \beta_2(EDU) + \beta_3(GDP) + \beta_4(INF) + \beta_5(M2) + \beta_6(POP) + \varepsilon$$  \hspace{1cm} (3.2)

Here $\beta_0$ is intercept, $\beta_{1,2,3,5,6}$ are the slope determinants of variables and $\varepsilon$ is the error term of regression equation that captures noise for it.

Different models are suggested for determination of relationship among variables. Some of those models are being used step by step for assuring the existence and understanding of relationship.

Step 1- unit root test

Unit root test is conducted so that the stationarity of data can be confirmed. Dickey and Fuller (1979) introduced the idea about different conditions of time series data and gave assumption that in stationary data, error terms are distributed normally. Gujarati (2003) also mentioned if there are some chances of correlation among the residuals of independent variables then augmented form of it is used to confirm the stationarity and usual t-test is valid. In the absence of trend and intercept, F-stats are valid. If the p value of t or F-stats is less than 5% then the null hypothesis of unit root is rejected and alternative is accepted.

Step 2- cointegration test

If the series are stationary at same level, for instance at first difference then they individually are stationary $I(1)$ and all are cointegrated level $I(1)$. After their cointegration confirmation, unrestricted VAR model is applied. It is to accommodate the maximum lag in order to get maximum cointegrating relations calculation without the biased impositions of deductive fundamental relations (Johansen, 1991). Vector Error Correction Model is used to measure the disequilibrium effects of variables. By considering all the differenced data as well, it integrates all long and short run measurements hence reducing any chances of valuable information loss ((Engle & Granger, 1987; Asteriou & Hall, 2007).

Step 3- autoregressive distributed lag

Pesaran and Shin (1997) examined the outcomes of different co integration theories and tests along with the traditional ARDL approach. They found that it does not only explain the stationary time series with random walk but also the ones that are differentiated for the sake of stationarity. ARDL has features that give consistent results for the time series data irrespective of their integrated level of “I(0)” or “I(1)”. Wald test determines the long run relationship among the model variables with the coefficients determination while keeping any kind of restrictions developed by the estimator (Asteriou & Hall, 2007).

Step 4- ordinary least squares and parameter stability test

Ordinary Least Squares gives insight in the intensity of association among variables for the long duration where the best and unbiased results for a linear model are needed to be calculated efficiently (Asteriou & Hall, 2007; Brooks, 2008). CUSUM is the type of recursive tests that are applied on residuals of OLS. It is the test that provides the visual imprinting of parameters stability If the value of cumulative sum of squares of residuals is within the positive and negative standard error range then the model is considered stable (Brooks, 2008).

Step 5- ramsey reset test

Ramsey (1969) intended to develop a test for the correct specification of model because an abnormality identified in residuals of a regression model could be due to different errors. Errors like non linearity of equation, error terms of variables affecting each other and the variables can be the causes where mean deviates from desired value.
3.1 Data collection

The annual data on unemployment, wages, education, GDP, inflation, money supply and population is included in the study from 1972 to 2012 in the form of different proxies. Data on GDP, inflation and money supply is collected from World Bank database in their growth rate forms to make it part of study. Data on unemployment, wages, education, population is collected from various issues of Economic Survey of Pakistan and Statistical Year Book of Pakistan. Unemployment is calculated by dividing the unemployed labor force by total labor force. Growth rate of population, high school enrollment and average amount of wages paid to labor force are taken as proxies for population, education and wages respectively.

4. Analysis

Unit root test is applied so that the stationarity of data can be tested and it can be clear if there are chances of any spurious results and take precautions against it.

Table 1: ADF Unit Root Tests Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Trend &amp; Intercept</th>
<th>Level (0)</th>
<th>Conclusion</th>
<th>Trend &amp; Intercept</th>
<th>Level (1)</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN</td>
<td>-1.959592</td>
<td>0</td>
<td>Non-Stationary</td>
<td>-5.947999***</td>
<td>0</td>
<td>Stationary I(0)</td>
</tr>
<tr>
<td>W</td>
<td>-7.059093***</td>
<td>1</td>
<td>Stationary I(0)</td>
<td>-7.127187***</td>
<td>3</td>
<td>Stationary I(1)</td>
</tr>
<tr>
<td>EDU</td>
<td>-4.643326***</td>
<td>0</td>
<td>Stationary I(0)</td>
<td>-8.772968***</td>
<td>0</td>
<td>Stationary I(0)</td>
</tr>
<tr>
<td>GDP</td>
<td>-5.776768***</td>
<td>0</td>
<td>Stationary I(0)</td>
<td>-11.28132***</td>
<td>0</td>
<td>Stationary I(1)</td>
</tr>
<tr>
<td>INF</td>
<td>-4.231914***</td>
<td>0</td>
<td>Stationary I(0)</td>
<td>-7.389978***</td>
<td>1</td>
<td>Stationary I(1)</td>
</tr>
<tr>
<td>M2</td>
<td>-5.005157***</td>
<td>0</td>
<td>Stationary I(0)</td>
<td>-7.620886***</td>
<td>0</td>
<td>Stationary I(1)</td>
</tr>
<tr>
<td>POP</td>
<td>-5.706149***</td>
<td>0</td>
<td>Stationary I(0)</td>
<td>-11.33775***</td>
<td>0</td>
<td>Stationary I(1)</td>
</tr>
</tbody>
</table>

*** Mackinnon (1996) critical values for rejection of null hypothesis of a unit root and significance at 1% level.

Table 2: VAR Lag Order Selection Criterion Test Results

<table>
<thead>
<tr>
<th>Order</th>
<th>LL</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>SC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>16.86974</td>
<td>NA</td>
<td>1.44e-09</td>
<td>-0.493487</td>
<td>-0.197933</td>
<td>-0.386624</td>
</tr>
<tr>
<td>1</td>
<td>99.53666</td>
<td>132.2671*</td>
<td>2.78e-10*</td>
<td>0.197933</td>
<td>0.386624</td>
<td>1.321929*</td>
</tr>
</tbody>
</table>

* indicates lag order selected by the criterion; LR: sequential modified LR test statistic; FPE: Final prediction error; AIC: Akaike information criterion; SC: Schwarz information criterion; HQ: Hannan-Quinn information criterion.

Table 2 presents the results of unrestricted VAR model. With the objective of assuring the optimum no. of lags, five commands of VAR model are used. Lag is selected at order one in the light of various VAR model works mentioned by Sharp (2010) that considered the Likelihood ratio, Akaike Information Criteria, and Hannan-Quinn Information Criteria as most appropriate in many cases. Also Brooks (2008) and Gujarati (2003) inferred likelihood ratio as the best measurement and Sharp (2010) himself is more in favor of Akaike Information Criteria. Now the model is VAR (1) because all the selection criterions except Schwarz information criterion indicate utmost fitting values at first lag. Once the lag orders are selected, cointegration test of Johansen is applied for the maximum no. of relationship determination in the long run.

Table 3: Cointegration Test Results

<table>
<thead>
<tr>
<th>Hypothesized No. of CE(s)</th>
<th>Eigenvalue</th>
<th>Trace Statistic</th>
<th>0.05 Critical Value</th>
<th>Prob.***</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0*</td>
<td>0.822644</td>
<td>196.0512</td>
<td>125.6154</td>
<td>0.0000</td>
</tr>
<tr>
<td>r = 1*</td>
<td>0.664317</td>
<td>128.5970</td>
<td>95.7366</td>
<td>0.0001</td>
</tr>
<tr>
<td>r = 2*</td>
<td>0.548797</td>
<td>86.02507</td>
<td>69.81889</td>
<td>0.0015</td>
</tr>
<tr>
<td>r = 3*</td>
<td>0.519074</td>
<td>54.98739</td>
<td>47.85613</td>
<td>0.0093</td>
</tr>
<tr>
<td>r = 4</td>
<td>0.365974</td>
<td>26.43777</td>
<td>29.79707</td>
<td>0.1162</td>
</tr>
<tr>
<td>r = 5</td>
<td>0.179035</td>
<td>8.666813</td>
<td>15.49471</td>
<td>0.3971</td>
</tr>
<tr>
<td>r = 6</td>
<td>0.024642</td>
<td>0.973086</td>
<td>3.841466</td>
<td>0.3239</td>
</tr>
</tbody>
</table>


Table 3 exhibits the results of Johansen cointegration test. Linear deterministic trend model is considered best as it depicts the significant relationships in a linear equation with pure random
walk. It rejects the null hypotheses of three pairs and accepts the alternate hypothesis of at most 4 relationships with one lag interval. This brings about notion that as a whole four variables share a long term affinity towards each other or at least one to another.

Once the relationship in long term is determined, there is need to consider not losing any information in the short term as well because it might contain some useful features that might be lost if the researchers are not careful in processing the differenced data as the original trend has its own disequilibrium (Asteriou & Hall, 2007; Gujarati, 2003).

Table 4 displays the results of Vector Error Correction Model so that any disequilibrium and short term relationship is measured and deliberated. Though sign of ECM is negative with 24% entailing that disequilibrium of one year is adjusted with this speed but statistically, it is insignificant that means no disequilibrium is adjusted here. Significance level of P-values shows only money supply and GDP have mildly significant effect at 5% on unemployment in the short run and it is .1% and .03% respectively.

Engle and Granger (1987) emphasized the significant results of bound testing even with the variables that are not cointegrated. So the Wald test is applied with the hypothesis of relationship and no relationship in long run among variables for the further confirmation as the variables are I(1), I(0).

Table 5 shows the highly significant results of bound testing. Its null hypothesis of no effect of variables on unemployment is rejected and alternate is accepted. The restriction applied in test is that all the coefficients' values are null hence ineffective to cause any changes in unemployment. F-stats value is reliable enough and for addition the value of Chi square also seconds the acceptance of variables affecting the regressed variable.

Ordinary Least Square is applied to measure the authenticity of every variable's effect on unemployment, overall goodness of fit of model and any signs of issues causing the false results and to rectify them.

Table 6 depicts that the results are significant for each of the variables of model. Education exhibits a negative impact on unemployment which means 1 unit increase in education will decrease the unemployment with .06 %. Similarly, population has negative effect on unemployment with highly significant value and implies approximately 3.17% decrease in unemployment with 1 unit increase in it. Wages share a positive relation with unemployment i.e. 1 unit increase in wages also increases unemployment with 0.014%. Overall impact of these variables' increase or decrease is .13% increase on unemployment. The value of $R^2$ points out that the model explains almost 77% variations in unemployment in help of model variables. F-stats value is 18.723 and highly significant being the p-value less than 1% rejects the null hypothesis of coefficients having zero effect on unemployment. Value of Durbin Watson-stats is 1.81 which is more close to 2 and indicates nonocurrence of autocorrelation that is coordinated with Asteriou and Hall (2007).

Heteroskedasticity, Autocorrelation and Multicollinearity are the issues that are matter of much concern when someone is applying the regression. As these issues might produce biased, inefficient results those are of no use to economists because they neither present the true picture nor can forecast the potential changes.

Table 7 displays the results of three diagnostic tests. First autocorrelation is tested and here the language Multiplier test results infer null hypothesis cannot be rejected because the P-value for both F-stats and LM stat are quite high. Secondly, heteroscedasticity of model is tested. It establishes that value of chisq in ARCH test is high enough to safely accept the null hypothesis of homoscedasticity. Third slot confronts the idea of multicollinearity because any precise linear relationship among independent variables will lead to inefficiency. Results indicate that though multicollinearity exists but none of the relationship is that strong to be even considered near perfect.
multicollinearity of 90% so we can firmly say that data outcomes are reliable.

### Table 7: Diagnostic Tests Results

<table>
<thead>
<tr>
<th>F-statistic</th>
<th>0.222882</th>
<th>Prob. F(1,33)</th>
<th>0.6400</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs*R-squared</td>
<td>0.275056</td>
<td>Prob.Chi-Square(1)</td>
<td>0.6000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-statistic</th>
<th>0.646464</th>
<th>Prob. F(1,33)</th>
<th>0.4264</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs*R-squared</td>
<td>0.669105</td>
<td>Prob.Chi-Square(1)</td>
<td>0.4134</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation</th>
<th>UN</th>
<th>W</th>
<th>EDU</th>
<th>GDP</th>
<th>INF</th>
<th>M2</th>
<th>POP</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>W</td>
<td>-0.1066</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EDU</td>
<td>-0.2126</td>
<td>0.3017</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GDP</td>
<td>-0.1615</td>
<td>0.1022</td>
<td>0.1566</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>INF</td>
<td>-0.1257</td>
<td>-0.0244</td>
<td>0.0445</td>
<td>-0.2179</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M2</td>
<td>0.0104</td>
<td>0.0921</td>
<td>0.0017</td>
<td>0.2337</td>
<td>-0.1998</td>
<td>1.0000</td>
<td>-</td>
</tr>
<tr>
<td>POP</td>
<td>-0.8341</td>
<td>0.2677</td>
<td>0.1041</td>
<td>0.2878</td>
<td>-0.0478</td>
<td>0.0432</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Table 8 presents the value of all F and t stats and the likelihood ratio. Some consider that value of F-stat is enough to confirm the correct specification. Though the critical value of F-stat is not available but probability of both F-stats and likelihood ratio suggests that the null hypothesis of correct specification cannot be rejected and model is accurate enough.

Once the model is tested with correct specification, parameter stability also needs to be tested.

Fig. 1 illustrates that the residuals values are normalized as they lie within the range between negative and positive values suggesting that model is stable. Even small transitions in data can be seen. Residuals are $\hat{f}(0)$, hence hinting at the normal distribution of residuals eradicating any chances of inefficient results entitles it safe for logical conclusion.

### 5. Conclusion

The study is conducted to determine the effect of wages, education, and macroeconomic variables’ impact on unemployment in Pakistan from 1972 to 2012. ADF test showed that only the unemployment has unit root which means that the shocks in previous periods affect not only the specific time interval but also the future rate of it. Result is seconded by vector error correction model that shows insignificant disequilibrium adjustments. It can be said that it is like the unemployment scarring (Arulamplam, Booth, & Taylor, 2001) where previous and current unemployment has influence on future. Once there is unemployment, it needs to be contained and more employment opportunities have to be created for future labor because previous and new unemployed labor will be hard to handle when amassed together. Results of Wald test and regression confirms the impact of wages, education, and population on unemployment in the long run deciding the significance of study to be reliable.

VECM test confirms that GDP and money supply maintain a positive relation with unemployment in short run. Education is stationary and mostly it has been an increasing trend throughout the study period. Education shares negative correlation with unemployment in the long run. Population has an increasing trend in Pakistan and the stationarity of data in its pure form relays about its consistent growth. It is the variable that showed greatest prevailing correlation with unemployment. Population shares no short run relation with unemployment but it decreases the later in the long run with 3.14% as a whole. Wage in Pakistan has seen a consistent growth and it is stationary in its random walk. Wages have a negative correlation with unemployment in short run but results are insignificant in that matter. However, an increase in wages in the long run causes slight increase in unemployment.

All the tests conclude that relationship among wages, education and macroeconomic variables exists. Some might have more effect on unemployment than others and the directions are either positive or negative but it definitely is there. Stability test of Ramsey confirms the correct specification of model and Johansen confirms the integrating relation of variables in long run. VECM confirms its existence in short run for some but no disequilibrium adjustments are significant. Results of OLS reconfirm the impression with diagnostics tests supporting the fact that assumptions of classic linear regression are fulfilled.

The reason behind education decreasing the unemployment can be the demand for the skilled workers, high market competition, etc. where only the skilled and experienced can keep the firm grip on job as they have more chances of getting the job and
less probability of being laid off. This outcome is similar to the outcome of study conducted by Webber (2002) in European countries. Rise in population has been considered to cast a negative effect on employment by many but current study concludes that increase in population is good thing for economy as this reduces the unemployment. This result is in agreement with Ezekiel (1936) and this might be due to the migration of population in working areas, government taking steps to accommodate some new entrants in market, more labor means the less time interval it takes to fill a vacancy.

![CUSUM and CUSUMSQ Test](image)

**Fig. 1:** CUSUM and CUSUMSQ Test

Inflation has negative relation with unemployment in long run but it is insignificant one which is coherent with outcome of Li and Liu (2012). Wages increase is reason for unemployment increase which can be due to increase in cost of employment, gap in demand and supply of remunerations. This result is harmonized with the previous studies of Kooros (2006) and König and Möller (2009), where increase in wages caused decrease in employment. GDP growth has slightly positive significant impact in the short run only that can be because increase in GDP is somewhat superficial as there are various crises that are being dealt at the moment and they badly affect the economy's stability. Though the result in long run also shows the minor positive relation but that is insignificant. This outcome is consistent with the study of Aurangzeb and kashif (2013) and Aktar and Ozturk (2009), where unemployment and GDP has positive relationship with each other.

A panel study can be conducted on this topic to capture the impact of these variables on provincial and individual level that will provide with more observations and a dynamic view can be tested.
through it. Short run increase in unemployment due to money supply is one indicator that tells monetary policy makers need to be more careful in devising the strategies so that the problem will be solved rather than growing. Policy makers must be vigilant while dealing with the things involving social dilemmas as policies need to be revised because the percentage of labor force has increased greatly specially the youth. Youth unemployment can harm the society the most with them being unemployed and going astray. The gap in market is in the form of inadequate guidance for required education and skill in the field, less employment opportunities, even the lesser for self-employment causing the increase in unemployment. These problems should be taken in consideration while devising the strategies to cop unemployment.

Since the youth makes the most portion of population of Pakistan, it’s a good opportunity to train the youth for technical support work to secure them a safe job and utilize talent as part of efficient economic system. Students’ career counseling and some training programs should be started by government. Wages need to be considered more critically according to the demand and supply in economy by the labor unions as well so that imbalance will not cause unemployment. In order to address the scarring impact of unemployment, at least more part time and contract jobs can be encouraged to use as respite to reduce the complete or long run unemployment circle.

References


Li, C., & Liu, Z. (2012). Study on the Relationship among Chinese Unemployment Rate, Economic Growth and


