2023/2024

Econometric Analysis of Unemployment:

The Impact of Education Level and Investment Rate

yosr sellami

ECONOMETRICS

PART 1: INTRODUCTION
PART 2: DATA AND METHODOLOGY
PART 3: MODEL 1 UNEMPLOYMENT ~ EDUCATION LEVEL
PART 4: MODEL 2 UNEMPLOYMENT ~ INVESTMENT RATE
PART 5: MODEL 3 UNEMPLOYMENT ~ EDUCATION LEVEL + INVESTMENT RATE (MULTIPLE LINEAR REGRESSION)
PART 6: CONCLUSION

PART1: INTRODUCTION

Unemployment is a critical socio-economic issue that affects individuals, families, and entire economies. It is a multifaceted problem that not only reduces income levels and standards of living but also leads to a wide range of social and economic challenges, such as increased poverty, decreased consumer spending, and lower economic growth. Understanding the determinants of unemployment is therefore essential for policymakers to develop effective strategies to mitigate joblessness and its adverse effects.

This study aims to explore the relationship between unemployment, education level, and investment rate using econometric methods. Education level and investment rate are two fundamental factors that influence employment outcomes. Education is often viewed as a means to enhance an individual's skills and productivity, potentially leading to better job prospects. However, the relationship between education and unemployment is complex and can vary depending on various factors such as the quality of education, the relevance of skills to the job market, and the overall economic environment.

Investment rate, on the other hand, is a critical determinant of economic growth and job creation. Higher investment rates typically lead to the expansion of businesses, the creation of new enterprises, and the development of infrastructure, all of which contribute to higher employment levels. By analyzing the interplay between these variables, this study seeks to provide insights into how educational attainment and investment levels influence unemployment rates.

Objectives :

To investigate the impact of education level on unemployment: This involves examining how different levels of educational attainment correlate with unemployment rates. The study aims to determine whether higher education levels lead to lower or higher unemployment, and to understand the underlying reasons for this relationship.

To analyze how investment rate affects unemployment: This involves evaluating the extent to which investment in the economy influences employment outcomes. The study will assess whether higher investment rates are associated with lower unemployment and will explore the mechanisms through which investment affects job creation.

To compare the explanatory power of simple and multiple linear regression models in predicting unemployment: By employing both simple and multiple linear regression models, the study aims to identify the most robust predictors of unemployment. Simple linear regression models will be used to analyze the impact of each independent variable (education level and investment rate) separately, while multiple linear regression models will consider the combined effects of these variables.

To offer policy recommendations based on the findings: The study will provide actionable insights for policymakers to design interventions that can effectively reduce unemployment. This includes recommendations on improving the alignment between education and labor market needs, as well as strategies to enhance investment in the economy.

This research is significant because it addresses a fundamental economic issue with far-reaching implications. By shedding light on the determinants of unemployment, the study aims to contribute to the development of more informed and effective economic policies. The findings of this research can help governments and other stakeholders to create targeted interventions that address the root causes of unemployment, thereby promoting economic stability and improving the quality of life for individuals and communities.

In summary, this study will utilize econometric analysis to explore the relationships between unemployment, education level, and investment rate. Through this analysis, we aim to uncover the key factors driving unemployment and provide evidence-based recommendations for reducing joblessness. This research is crucial for developing strategies that can foster economic growth, enhance employment opportunities, and ultimately improve socio-economic outcomes.

PART 2: DATA AND METHODOLOGY

1. DATA

The dataset used in this analysis includes information sourced from an Excel file containing observations on unemployment rates, education levels, and investment rates. Key variables include:

- Unemployment: Percentage of the workforce that is unemployed.
- Education level: A measure or index representing the average educational attainment of the population.
- Investment rate: Percentage of GDP invested in various sectors of the economy.

2. METHODOLOGY

To analyze the relationships between these variables, two econometric models were employed: simple linear regression and multiple linear regression.

3. SIMPLE LINEAR REGRESSION MODEL

Model: $Y=\beta0+\beta1X+\epsilon Y + \epsilon$

In a simple linear regression model:

- Y represents the dependent variable (in this case, unemployment rate).
- X represents the independent variable (either education level or investment rate).
- $\beta0$ is the intercept, representing the expected value of Y when X is zero.
- $\beta1$ is the slope coefficient, indicating the change in Y for a one-unit change in X.
- ϵ is the error term, capturing the variability in Y that cannot be explained by X.

Objective: The objective of simple linear regression is to estimate the relationship between Y and X, assess the statistical significance of the relationship, and evaluate how well X explains the variability in Y (measured by R^2).

4. MULTIPLE LINEAR REGRESSION MODEL

Model: $Y=\beta0+\beta1X1+\beta2X2+\ldots+\beta kXk+\epsilon Y$

In a multiple linear regression model:

- Y again represents the dependent variable (unemployment rate).
- X1,X2,…, Xk represent multiple independent variables (education level and investment rate in this case).
- $\beta0$ is the intercept, indicating the expected value of Y when all X variables are zero.
- $\beta1,\beta2,\ldots,\beta k$ the coefficients associated with X1,X2,…, Xk, showing the change in Y for a one-unit change in each X, holding other variables constant.
- ϵ is the error term, representing unobserved factors influencing Y.

➢ OBJECTIVE:

Multiple linear regression extends the simple model by including multiple predictors simultaneously. The goal is to understand how each independent variable contributes to explaining Y, control for confounding effects (if any), and improve the overall predictive power of the model (measured by R^2).

These models are essential tools in econometric analysis, used to quantify relationships between variables, make predictions, and test hypotheses about causal relationships. Simple linear regression provides a basic framework, while multiple linear regression allows for more complex analyses by incorporating multiple predictors simultaneously.

➢ CORRELATION ANALYSIS

Before fitting the regression models, a correlation analysis was conducted to understand the relationships between the variables. This step helps in identifying potential multicollinearity issues and understanding the strength and direction of the relationships.

```
> # Compute correlations between variables
> correlation_matrix <- cor(data)
>
> # Print correlation matrix
> print(correlation_matrix)
                 Unemployment Education level investment rate
Unemployment       1.0000000      0.61625849     -0.48565882
Education level    0.6162585      1.00000000     -0.06770509
investment rate   -0.4856588     -0.06770509      1.00000000
```

5. SUMMARY:

- The correlation matrix provides insights into the linear relationships between Unemployment, Education level, and investment rate.
- The positive correlation between Unemployment and Education level suggests a nuanced relationship that warrants further investigation into the reasons behind higher unemployment rates among more educated populations.
- The negative correlation between Unemployment and investment rate supports economic theory that higher levels of investment can contribute to lower unemployment rates.
- The weak correlation between Education level and investment rate indicates that these two variables may operate somewhat independently within the dataset.

PART 3: MODEL 1 UNEMPLOYMENT ~ EDUCATION LEVEL

➢ Summary Interpretation:

```
# Fit the linear models
> lm_model <- lm(Unemployment ~ `Education level`, data = data)
> summary(lm_model)

Call:
lm(formula = Unemployment ~ `Education level`, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-23.067  -6.843  -1.179   4.259  24.765

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        12.9530     3.9528   3.277  0.00272 **
`Education level`   0.4082     0.1597   2.557  0.01607 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.81 on 29 degrees of freedom
  (1 observation effacée parce que manquante)
Multiple R-squared:  0.1839,   Adjusted R-squared:  0.1558
F-statistic: 6.536 on 1 and 29 DF,  p-value: 0.01607
```

➢ Interpretation:

- **Coefficients:**
  - **Intercept:** The intercept (12.9530) represents the estimated unemployment rate when the education level is zero. It's statistically significant (p = 0.00272).
  - **Education level:** The coefficient (0.4082) indicates that for every unit increase in education level, the unemployment rate increases by 0.4082 percentage points. It's statistically significant (p = 0.01607).
- **Model Fit:**
  - **R-squared:** 0.1839 suggests that 18.39% of the variability in unemployment can be explained by variations in education level.
  - **Adjusted R-squared:** 0.1558 adjusts for the number of predictors and indicates a moderate fit of the model.
  - **F-statistic:** 6.536 with a p-value of 0.01607 indicates that the model is statistically significant, meaning education level has a significant impact on unemployment rates in this model.

PART 4: MODEL 2 UNEMPLOYMENT ~ INVESTMENT RATE

➢ Summary Interpretation:

```
> lm1_model <- lm(Unemployment ~ `investment rate`, data = data)
> summary(lm1_model)

Call:
lm(formula = Unemployment ~ `investment rate`, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-16.509  -7.478  -2.781   4.602  22.874

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)         33.118      4.320   7.667 5.06e-08 ***
`investment rate`   -1.067      0.384  -2.778   0.0102 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.87 on 25 degrees of freedom
  (5 observations effacées parce que manquantes)
Multiple R-squared:  0.2359,   Adjusted R-squared:  0.2053
F-statistic: 7.717 on 1 and 25 DF,  p-value: 0.01023
```

➢ Interpretation:

- **Coefficients:**
  - **Intercept:** The intercept (33.118) represents the estimated unemployment rate when the investment rate is zero. It's statistically significant ($p < 0.001$).
  - **Investment rate:** The coefficient (-1.067) indicates that for every percentage point increase in the investment rate, the unemployment rate is expected to decrease by 1.067 percentage points. It's statistically significant ($p = 0.0102$).
- **Model Fit:**
  - **R-squared:** 0.2359 suggests that 23.59% of the variability in unemployment can be explained by variations in the investment rate.
  - **Adjusted R-squared:** 0.2053 adjusts for the number of predictors and indicates a moderate fit of the model.
  - **F-statistic:** 7.717 with a p-value of 0.01023 indicates that the model is statistically significant, meaning investment rate has a significant impact on unemployment rates in this model.

PART 5: MODEL 3 UNEMPLOYMENT ~ EDUCATION LEVEL + INVESTMENT RATE (MULTIPLE LINEAR REGRESSION)

➢ Summary Interpretation:

```
> multiple_lm_model <- lm(Unemployment ~ `Education level` + `inves
tment rate`, data = data)
> summary(multiple_lm_model)

Call:
lm(formula = Unemployment ~ `Education level` + `investment rate`,
    data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-14.7616  -5.0482  -0.8557   4.9691  19.8311

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        19.7899     4.4588   4.438 0.000173 ***
`Education level`   0.6153     0.1396   4.408 0.000187 ***
`investment rate`  -0.9796     0.2920  -3.355 0.002635 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.245 on 24 degrees of freedom
  (5 observations effacées parce que manquantes)
Multiple R-squared:  0.5778,   Adjusted R-squared:  0.5426
F-statistic: 16.42 on 2 and 24 DF,  p-value: 3.212e-05
```

➢ Interpretation:

- **Coefficients:**
  - **Intercept:** The intercept (19.7899) represents the estimated unemployment rate when both education level and investment rate are zero. It's statistically significant ($p < 0.001$).
  - **Education level:** The coefficient (0.6153) indicates that for every unit increase in education level, the unemployment rate is expected to increase by 0.6153 percentage points. It's statistically significant ($p = 0.000187$).
  - **Investment rate:** The coefficient (-0.9796) indicates that for every percentage point increase in the investment rate, the unemployment rate is expected to decrease by 0.9796 percentage points. It's statistically significant ($p = 0.002635$).
- **Model Fit:**
  - **R-squared:** 0.5778 suggests that 57.78% of the variability in unemployment can be explained by variations in both education level and investment rate.
  - **Adjusted R-squared:** 0.5426 adjusts for the number of predictors and indicates a strong fit of the model.
  - **F-statistic:** 16.42 with a very low p-value (3.212e-05) indicates that the model is statistically significant, meaning both education level and investment rate together have a significant impact on unemployment rates.

6. SUMMARY:

- **Education Level:** Higher levels of education are associated with higher unemployment rates, as indicated by positive coefficients in both simple and multiple regression models.
- **Investment Rate:** Higher investment rates are associated with lower unemployment rates, shown by negative coefficients in both simple and multiple regression models.
- **Multiple Regression:** Combining education level and investment rate improves the explanatory power of the model (higher R-squared and lower p-values), indicating that these factors collectively provide a more comprehensive understanding of unemployment trends.

PART 6: CONCLUSION

This econometric analysis explored the relationships between unemployment rates, education levels, and investment rates using linear regression models. Here are the key findings and implications derived from the analysis:

1. **Impact of Education Level:**
    - Higher education levels are associated with higher unemployment rates. The coefficient for education level in both simple and multiple regression models indicates that an increase in education level leads to a corresponding increase in unemployment. This unexpected finding suggests potential issues such as overqualification or mismatches between skills acquired through education and job market demands.
2. **Effect of Investment Rate:**
    - Higher investment rates are linked with lower unemployment rates. The negative coefficient for investment rate across all models suggests that increased investment stimulates economic growth, leading to job creation and reduced unemployment. This underscores the importance of fostering an environment conducive to investment to bolster economic activity and employment opportunities.
3. **Multiple Regression Insights:**
    - The multiple linear regression model combining both education level and investment rate provides a more comprehensive understanding. It reveals that while higher education levels contribute to unemployment, higher investment rates mitigate it. This dual perspective highlights the complex interplay between human capital development (through education) and economic policy (through investment).
4. **Limitations and Further Research:**
    - The analysis assumes linear relationships between variables, which may oversimplify real-world dynamics. Future research could explore nonlinear effects and interactions between education, investment, and other socio-economic factors.
    - Additionally, the dataset's scope and quality play a crucial role in the robustness of conclusions drawn. Further studies could benefit from larger, more diverse datasets spanning different regions and economic contexts.

In conclusion, this econometric analysis underscores the intricate dynamics influencing unemployment rates. By understanding and addressing the impacts of education and investment, policymakers can formulate targeted interventions to foster economic growth, reduce unemployment, and promote sustainable development