

STEPWISE REGRESSION MODEL FOR STUDY THE UNEMPLOYMENT RATE IN IRAQ

Ali Sadiq Mohommed BAGER ^{a}, Bahr Kadhim MOHAMMED ^b,
Hajem Ati Daham ^a, Sadiq AL-HUSSEINY ^a*

^a AL- Muthanna University, Iraq

^b AL-University of AL-Qadisiyah, Iraq

ABSTRACT

One of the most important problems facing the Iraqi economy is the problem of unemployment, which has risen in recent years. This paper will focus on determining the most important macroeconomic factors, which affect the unemployment rate in Iraq, using the stepwise regression as one of the most widely used methods for variables selection. The results were compared with those obtained from the ordinary least squares (OLS) method in order to obtain the most possible model which illustrate the examined phenomenon. We find that unemployment in Iraq seems to be affected by economic growth, volume of inflation, average number of foreign workers, average worker age and size of the population.

KEYWORDS: *Ordinary Least Squares (OLS), stepwise regression, unemployment rate, variable selection.*

1. INTRODUCTION

World nations, and therefore the Arab nation in specific, are influenced by asymmetric workforce markets and unemployment issues. The value of this is relying on the characteristics of the nation's economics and its growth. Iraq as one of the Arabic countries has witnessed an increment of unemployment as a reason for the financial crisis, which also has affected the level of the country living style. In 2014, where oil prices fell to 65%, Iraq had encountered an extreme financial and economic crisis that has started to raise and worsen for implicit and explicit reasons such as the lack of the financial planning and corruptions. It ought to mention that Iraq depends totally on the oil resources instead of searching for alternative resources such as industry, agriculture, tourism, etc. In addition to the reduction of the money liquidity of the country, the decline of oil prices has a significant impact on the society by raising the unemployment and poverty rates and reducing the services levels introduced by the government in different sectors as in education, healthcare and infrastructure as well (Jaba et al., 2010).

Wang and Chen (2016) applied Venn Diagrams and some experiments to demonstrate three problems of the Stepwise Regression (SR): the wrong degree of freedom, capitalization when sampling and the case when the error R^2 does not optimized. The authors are also used the method of all possible subsets regression for selecting the variables of a model and the method is tested on a real education data. La-Siudem and Jaroszewicz (2014) introduced an approach of correlation based on indexing and selection of stepwise regression for linked data sets. According to the authors, the method is tested to be efficient for large data (Bager et al., 2017, <https://mpra.ub.uni-muenchen.de/id/eprint/81390>).

* Corresponding author. E-mail address: nader.ali62@yahoo.com

In this paper, we will study some of factors that led to the high rate of unemployment in Iraq, by using the stepwise regression method, where the dependent variable is the unemployment rate and a set of major economic indicators are considered as explanatory variables. For data analysis, an efficient algorithm was used in programming (NCSS). This paper is organized as follows: after the introduction, in Section 2 we will present the methodology of stepwise regression. In Section 3 we will illustrate the sample study, while In Section 4 we will discuss the analysis of the study results. In the end, a brief conclusion is included in Section 5.

2. METHODOLOGY OF THE STEPWISE REGRESSION MODEL

Many approaches have been developed for identifying variables and important factors for the linear models. The stepwise regression is one of these common approaches, which was firstly introduced by Efroymson in 1960, and illustrated later by Draper and Smith in 1981. Wilkinson (1979) in this approach, variables are identified for regression models to find the appropriate model for a group of subsets models which allow variables to be included or removed at each separate step. In this case the less important and significant variables is removed from the model and alternative variable is included based on the last add variables. The forward selection method is one of the common stepwise selection methods to obtain the best model. Two cutoff values which are the F_{IN} and F_{OUT} are needed for the stepwise regression where some analysts assume that $F_{IN} = F_{OUT}$, which is not important (Thompson, 1995). As a reason we assume that $F_{IN} > F_{OUT}$.

$$R^2 = \frac{1}{SST} (SSR(x_1, x_2, \dots, x_n)) \quad (1)$$

$$= 1 - \frac{SSE(x_1, x_2, \dots, x_n)}{SST}$$

$$Mse = \frac{1}{n - p} (SSE(x_1, x_2, \dots, x_n)) \quad (2)$$

$$F = \frac{SSR}{Mse} \quad (3)$$

3. SAMPLE STUDY

In this paper, we aim to identify the determinants of the unemployment rate in Iraq, using macroeconomic data. Although the purpose of this paper is to test the performance of the stepwise regression model. The specific economic context in Iraq, described in the previous sections, is taken into account when selecting some specific variables, such as public expenditures.

Having these limitations in mind, the set of variables affecting the rates of unemployment in Iraq includes:

- (a) Rates of unemployment (y): rates of unemployment in Iraq.
- (b) Economic growth (X_1): expressed as gross domestic product, is a cumulative measure of the final goods and services produced in an economy over time. This variable is measured as a million Iraqi dinars.
- (c) Rate of inflation (X_2): is the rate at which for change the general level of prices for goods and services and the measured of this variable was taken as a percent rate.

(d) Volume of investment (X_3): is the act of committing money or capital to an endeavor (a business, project, real estate, etc.) with the expectation of obtaining an additional income or profit. This variable is measured as a million Iraqi dinar.

(e) Government expenditures (X_4): It is the expenditure that the state spends to expand production such as capital transfers of projects, expenditures for physical infrastructure such as roads and transportation, and institutional and legal infrastructure, which create a suitable investment climate. This variable measured as of a million Iraqi dinars.

(f) Size of the population (X_5): The total population of the study area, this variable measured as a population.

(g) Average income Workers (X_6): The median wage for workers.

(h) Average number of foreign worker (X_7): Is a worker who works in a country other.

(i) Average working age (X_8): is defined as those aged 15 to 60.

4. ANALYSIS OF STUDY RESULTS

In order to determine the variables which, influence the unemployment rate in Iraq, the stepwise regression model is composed by the response variable, which represents the rates of unemployment in Iraq, and the independent variables as listed above. It is coded using NCSS programming.

Table 1. Stepwise regression analysis (FIN, FOUT and Sig.F Change)

	Independent Variable	FIN	FOUT	Sig.F Change
1	0.000	0.023	0.023	X_1
2	0.501	0.098	0.087	X_2
3	0.020	0.012	0.130	X_3
4	0.654	0.847	0.135	X_4
5	0.047	0.158	0.172	X_5
6	0.989	0.890	0.083	X_6
7	0.003	0.095	0.099	X_7
8	0.001	0.018	0.023	X_8

Source: Produced by the author, using NCSS programming

Table 2. Stepwise regression coefficient

	Independent Variable	Stepwise Regression Coefficient	Standard Error	t-value	Pr> t
1	Intercept	758.2541			
2	X_1	0.11	0.02	5.76	0.000
3	X_2	0.01	0.19	0.05	0.521
4	X_3	0.20	0.07	2.87	0.006
5	X_4	0.07	1.12	0.07	0.801
6	X_5	-0.19	0.09	-2.12	0.04
7	X_6	0.03	0.69	0.06	0.780
8	X_7	0.46	0.08	5.63	0.001
9	X_8	0.36	0.15	2.40	0.021
10	R-Squared	0.8313			

Source: Produced by the author, using NCSS programming

The results presented in Table 1 and 2 show the relationship between the unemployment rate in Iraq and the following eight explanatory variables: economic growth (X_1), volume of investment (X_3), size of the population (X_5), average number of foreign worker (X_7), average worker age (X_8) are significant from a statistical point of view and have a positive relationship with rates of unemployment. This means that if there is an increase in these variables rate by 1%, the rates of unemployment will also increase.

Table 3. OLS regression coefficient

	Independent Variable	OLS	Standard Error	t-value	Pr> t
1	Constant	-502.406	307.650	-1.633	0.108
2	X_1	2.801	0.627	4.467	0.004
3	X_2	-3.204	11.179	-0.287	0.775
4	X_3	0.819	0.028	28.818	0.005
5	X_4	0.550	1.228	0.452	0.652
6	X_5	0.009	0.012	0.774	0.442
7	X_6	- 0.001	0.001	-1.106	0.273
8	X_7	8.433	3.710	2.273	0.021
9	X_8	-3.361	4.326	-0.777	0.440
10	R-Squared	61.257			

Source: Produced by the author, using NCSS programming

This also explains why the other rate of inflation (X_2), government expenditures (X_4), average income Workers (X_6) in the model have no significant effects on the unemployment rate.

The explanatory robust of the model represented by $R^2 = 83\%$ of the changes in the unemployment rate is explained by the changes in each of the variables and (17%) variables are due to other factors that may not be quantified within the estimated model.

While the OLS results showed the importance of the three variables: the economic growth (X_1), volume of investment (X_3), average number of foreign worker (X_7) and the value of $R^2 = 61\%$. This is an indication that the stepwise regression method is more efficient, and it shows better results as in Table 3.

5. CONCLUSIONS

The study showed that the stepwise regression method is one of the successful ways. The results of the Stepwise regression method results can be classified as follows:

(a) Through the results reached from this study. We conclude that the importance of the first variable which is the economic growth (X_1) as one of the most important variables affecting the unemployment rate in Iraq, so the significant value of this variable is less than 0.05. This is consistent with economic logic since that the economic behavior is primarily connected with the economic. The second variable is volume of inflation (X_3) with significant value less than 0.05. The average number of foreign workers (X_7) is in the third order, the reason is that many organizations prefer foreign workers for their skills and ability. This has led to increase unemployment for local workers. The average worker age (X_8) comes in the fourth, finally the size of the population (X_5), which has the less significant value amongst variables.

(b) As we can see from the results there are no significant effects of the following variables: rate of inflation (X_2), government expenditures (X_4), average income workers (X_6), because the value is more than 0.05.

By applying the stepwise regression method, we found that there are five variables with a significant impact on the unemployment rate in Iraq, while other variables have a weak and non-significant. Finally, stepwise regression method is recommended for other applications, since it provides better estimators than the ordinary least square method.

REFERENCES

- Bager, A., Roman, M., Algedih, M. & Mohammed, B. (2017). *Addressing multicollinearity in regression models: a ridge regression application*. Retrieved August 2, 2019, from <https://mpra.ub.uni-muenchen.de/id/eprint/81390>.
- Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis*, John Wiley and Sons. New York, 407.
- Efroymson, M.A. (1960). *Multiple regression analysis. Mathematical methods for digital computers*, 191-203.
- Jaba, E., Balan, C., Roman, M. & Roman, M. (2010). Statistical evaluation of spatial concentration of unemployment by gender. *Economic Computation and Economic Cybernetics Studies and Research*, 44(3), 79-92.
- La-Siudem, B.Z. & Jaroszewicz, S. (2014). Fast stepwise regression on linked data. In *Proceedings of the 1st Workshop on Linked Data for Knowledge Discovery (LD4KD) co-located with ECML/PKDD*, 14, 17-26.
- Thompson, B. (1995). *Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial*.
- Wang, K. & Chen, Z. (2016). Stepwise regression and all possible subsets regression in education. *Electronic International Journal of Education, Arts, and Science (EIJEAS)*, 2.
- Wilkinson, L. (1979). Tests of significance in stepwise regression. *Psychological Bulletin*, 86(1), 168.