

## Adopting some statistical methods in studying Iraqis immigration behavior

Anwaar Dhiaa Abdul Kareem<sup>1</sup>, Sabah Haseeb Hasan<sup>2</sup>

<sup>1</sup>Department of Mathematics, College of Education for Pure Sciences, University of Kirkuk

<sup>2</sup>Department of Statistics, College of Administration and Economics, University of Kirkuk

---

### ABSTRACT

This article dealt with the topic of adopting statistical methods in studying the behavior of the Iraqi immigration, especially the multiple regression method in determining the most important factors that affect this phenomenon. The results showed that statistical methods have a great ability to characterize the phenomenon of migration by defining a multiple regression model for the impact of some factors such as gross domestic product, gross domestic product per capita and the size of the population in Iraq on the migration. Parameters of the model have been estimated using the ordinary least squares method and two stage least squares method. As it was estimated the amount of the impact of each of the factors included in the model on the migration, as well as the direction of the impact of each factor, whether these effects are negative or positive. All tests and statistical criteria indicated the significance of the effects of the factors included in the model and the high quality of the estimated regression model. Therefore, the responsible authorities in Iraq must rely on statistical methods to develop plans and solutions to control the phenomenon of migration and the factors that affect it.

**Keywords:** Multiple regression, Two stage least squares, Unit root, Stationary tests, Migration

### *Corresponding Author:*

**Sabah Haseeb Hasan**

Department of Statistics, College of Administration and Economics

University of Kirkuk, Iraq

E-mail: [sabahsaqi@uokirkuk.edu.iq](mailto:sabahsaqi@uokirkuk.edu.iq)

---

### 1. Introduction

All countries of the world are searching for solutions that can contribute to reducing the negative effects of the problem of individual migration, by setting economic development policies that contribute to providing the largest possible number of employment opportunities for individuals, and achieving high growth rates, as this achieves the continuation of increasing the income of individuals and improving their economic and living standards, which helps increase the level of well-being of members of society.

The research problem is that it seeks to study the behavior of Iraqis' migration out of the country, and determine the impact of the most important factors that affect the number of immigrants through the use of some statistical methods represented by the multiple regression model. Parameters of the model are usually estimated in several methods, including the ordinary least squares method and the two stage least squares method.

The importance of the article lies in the fact that it seeks to determine the best statistical methods by which to determine the effects of the most important factors affecting the number of Iraqi immigrants through multiple regression methods, which is one of the topics that have not been adequately addressed in research and studies, which makes the research an addition scientific can be used to contribute to finding solutions that are useful in dealing with the effects of the phenomenon, which causes a lot of damage to the national economy.

The article aims to achieve a major objective which is to identify the best statistical methods by which to determine the effects of the most important factors that affect the number of immigrants from Iraq, by achieving a set of the following sub-objectives:

- 1- Knowledge of concepts and terminology related to immigration.
- 2- Study the statistical method used in the research
- 4- The practical application of real data regarding Iraqi immigration.
- 3- Reaching conclusions and recommendations through which the study objectives can be achieved.

A set of appropriate approaches will be used for the aims of the research and to verify its purpose, as the research will use the descriptive approach by which to describe the phenomenon of migration, and the research will use the analytical approach to real data on the migration of Iraqis for the period (2004-2018) through the use of statistical programs such as SPSS and Eviews to obtain useful results and recommendations.

## **2. Immigration: Causes and negative effects**

Immigration is defined as the move of an individual or group from their permanent place of residence to another place with the aim of effecting a change in their social or economic status or both and seeking towards what is best guaranteed. Migration may be for a specific period of time or with the intent to reside and may be individual or in the form of a group. Usually, migration is for political, religious, economic, and recreational reasons, or voluntary migration to seek living from one place to another or from one country to another. Migration may be internal within the country's borders or external outside the country's borders [1].

Migration may be legal and it is done according to the national procedures and laws of countries or illegal and means not respecting the requirements necessary to cross country borders [2]. Therefore, immigration today is considered one of the global phenomena and is of interest to all countries of the world.

In light of the life conditions in some countries, many people resort to emigration outside the country, and emigration outside the country has many negatives that affect the person and society in general, so that we can learn about the following the most important causes and negatives of immigration as follows:

### **2.1. Reasons for immigration**

There are many reasons that cause people to migrate and bear their suffering [3], and the reasons can be summarized in the following points:

- 1- Man resorts to immigration outside his country in order to search for suitable work.
- 2- Migration occurs, exposing the country to economic crises and unemployment.
- 3- The occurrence of wars and disasters, such as earthquakes, famines and conflicts in the country, which causes immigration, whether legal or illegal immigration.
- 4- Lack of appreciation of competencies, and this is why people search for other countries that value their different skills and competencies.
- 5- Low job opportunities and low wages for workers cause the resort to immigration.
- 6- The high population density that causes unemployment and the low economic level.

### **2.2. Negative effects of immigration**

- 1- One of the most important disadvantages of migration is the feeling of alienation and distance from the home country.
- 2- The effect on the families and dispersal of immigrants.
- 3- Incidence of mental and social disorders.
- 4- Countries lose much of the labor they need as a result of migration.
- 5- The loss of identity if the country to which the person emigrated is different from the customs and traditions in which he was brought up

With regard to the migration of scientists and scientific competencies, these are one of the pillars of national development and factors that drive them, and they are the force that influences the creation of creativity and participation in preparing material strength and moving the country from the state of influence to the state of influence. Remains in the event of underdevelopment and dependency as a result of its need for technological and scientific development [4].

Brain drain produces several negative effects on the reality of development in the countries from which these brains migrate. One of the most important negative effects of brain drain is [5]:

- 1-Brain drain represents a drain on an effective and influential segment and has a prominent role in society and in implementing development plans for countries from which brains migrate, as they are in great need of scientific competencies and trained manpower.
- 2- Brain drain is a loss in the field of education in all its stages, where scientific, technical and knowledge competencies are the main source of comparative advantage and the basis for excellence and competition among nations.

3- Brain drain widens the gap between developed and developing countries. Because these minds give the countries to which they migrate great benefits and direct economic return, while they constitute a loss for the countries from which these minds have been displaced.

### 2.3. Immigration in Iraq

The difficult living and security conditions that Iraqis live in have led to the emigration of large numbers of them abroad, as many others think about doing so because of the widespread unemployment and corruption and their feeling of frustration due to the lack of a spark of hope looming on the horizon to improve the conditions. And the past years recorded high rates of immigration, as large numbers of them left outside Iraq, and many of them died by drowning. Observers have warned of the danger of immigration, especially the brain drain, and its implications for the country, and it seems that the perceptions of individuals towards the situation of insecurity and lack of opportunities and economic in Iraq are among the most influencing factors in the decision of people to migrate outside the country.

On the other hand, the phenomenon of scientific talent migration is a very complex phenomenon, resulting from the interaction of multiple and interrelated factors between the economic, political, social, and scientific as well as related to the personal interest of the individual who always tries to achieve his goals, regardless of what the nation can lose. It has become one of the most important phenomena that affect the development of the national economy and the structural composition of the type of population, as well as being a major process to demolish the foundations of growth and development, and attempts to reduce the cultural, scientific and technological difference with the developed world due to the increasing numbers of immigrants from a highly qualified population that requires a serious and real position To deal with this bleeding of cognitive capital [6].

### 2.4. Previous studies

There are several previous studies that dealt with the issue of Iraqi immigration, including:

Raouf and Al-Askari, 2007, where the reasons for the migration of Iraqi talents were discussed through a questionnaire distributed to professors at the College of Education Al-Mustansiriya University to collect data on the causes of migration and ways to reduce it. Some simple statistical indicators were used in the data analysis [7].

Issa, 2016 conducted a study aimed at explaining the impact of external migration of Iraqi youth on the political, economic and social reality and how to find appropriate solutions to stop or limit immigration. And find out the reasons for it being a great loss for the country in all respects mentioned. The research used the historical and descriptive method for the purpose of reaching realistic scientific results regarding the research problem [6]

Raheem, 2017 search in the factors that affect the migration of Iraqi youth through a questionnaire distributed to a sample of young people in the Al-Qadisiyah Governorate - Iraq. A factor analysis method was used to determine the most important factors that affect migration [8].

Radhi, 2018, dealt with knowing the immigration phenomenon and its motives, applying this to the Iraqi situation, knowing the numbers of Iraqi immigrants illegally, the reasons behind the immigration, the authorities' position on monitoring and supporting migrants from Iraq and what can be provided to them. Some statistics have been presented on the numbers and percentages of immigrants from Iraq [9].

## 3. Statistical Methods in Analyzing the Immigration

Most researchers in their scientific studies on a particular phenomenon use a set of statistical methods to interpret or find logical and scientific solutions to it. Among the most important of these methods is the multiple regression method, which is one of the most important methods of applied statistics when studying economic, social and other life phenomena. Before that, we get to know the issue of the stationary of time series, its importance in analyzing the behavior of the phenomenon studied here. Therefore, these methods will be identified in the coming sections.

### 3.1. Time series stationary tests

When analyzing time series data, it is important to ensure the stationary of the time series of the variables. as the un stationary of the time series leads to the spurious regression problem which means that the relationship between the variables is not real. There are several statistical methods for checking the stationary of the time series, including unit root testing.

Unit root test is used to check the stationary of time series. The presence of the unit root means the time series instability over time. In other words, the mean and variance of the variable are not time independent. Several tests have been developed to detect the presence of the unit root, the most important of which is the augmented Dickey-Fuller test [10, 11]. As the following hypothesis is tested:

$H_0: \phi = 1$  ( Series has unit root)

$H_1: |\phi| < 1$  (Series not has unit root)

When hypothesis  $H_0$  is rejected, hypothesis  $H_1$  which states that there is no unit root is accepted, we conclude that the series is stationary.

### 3.2. Multiple regression methods

The use of a simple linear model is limited to analyzing the relationship between an approved variable and its relationship to one (independent) explanatory variable, but in reality there are studies that require the adoption of the dependent variable as a function of more than one explanatory variable, such studies are covered by the multiple regression model.

If we have a dependent variable  $Y$  is a linear function to  $(K)$  of the explanatory variables (independent)  $X_1, X_2, X_3, \dots, X_K$ , then the multiple linear regression model if  $n$  observations is available takes the following formula [12]:

$$Y_i = B_0 + B_1X_{1i} + B_2X_{2i} + \dots + B_kX_{ki} + u_i \quad i=1,2,3,\dots,n \quad (1)$$

Where,  $B_0, B_1, B_2, \dots, B_k$  are called model parameters,  $B_0$  is the intercept term.  $B_1, B_2, \dots, B_k$  denotes the change in  $Y$  with respect to the variables  $X_1, X_2, \dots, X_K$  respectively and  $u_i$  is the random error,

The above model can be rewritten as matrices as follows:

$$Y = BX + U \quad (2)$$

Where  $Y$  is a vector of the values of the dependent variable,  $X$  is a matrix of explanatory variables and  $U$  is a random error vector. When the assumptions of the general linear model are fulfilled, the ordinary least squares method (OLS) gives the best estimation of the model parameters as follows:

$$b = (X'X)^{-1}X'Y \quad (3)$$

Where  $b$  is the vector of the model parameter estimator and  $X'$  denotes the transpose of the matrix  $X$ .

The multiple linear regression model that was discussed is only a special case of general situation, according to which it was assumed that there is a single trend of causality, meaning that the set of explanatory variables  $X_1, X_2, \dots, X_k$  affects the dependent variable  $Y$  and is not affected by it. While the general condition of most relationships involves mutual dependence between the variables involved in the model, meaning that there are at least a number of variables that are determined simultaneously, they affect and are affected by each other. Therefore, the correct treatment of most phenomena requires formulating them as a nested set of relationships with the system of simultaneous equations.

The two-stage least squares method (2SLS) is used to estimate model parameters in the case of a system of simultaneous equations that are in a state of exactly identified. Also, this method aims to eliminate the problem of correlation, that is, the relationship between dependent variables that are as an explanatory variable within the system and the limit of random error within the equation itself. This method indicates that the method used is in two stages [13].

The first stage: determining the endogenous variable in the fair whose parameters are to be estimated, and then finding the reduced form for this variable. And the use of the least squares method in the process of the parameters of the reduced form and thus find the estimated values of endogenous variables in the system.

The second stage: that the OLS method is used again in estimating the parameters of the structural form after replacing the estimated  $Y$  values in the first stage with the actual  $Y$  values. In order to estimate the 2SLS method, the parameters of all system equations are estimated using the OLS method.

### 4. Applied framework

Before applying multiple regression methods to the phenomenon of immigration, it is necessary to know the behavior of some of the variables used in this section. Where data were recorded for the period (2004-2018) on

the rate of Iraqi immigration as an dependent variable (MIG), gross domestic product growth rate (GDP), population growth rate (POP) and GDP per capita (CAPIT) as explanatory (independent) variables that may It has an impact on immigration. It is worth noting that the data was obtained from World Bank bulletins on Iraq. In the following paragraphs, the evolution of data will be studied during the period 2004-2018. To check the stability of time series, unit root test was performed using augmented Dickey-Fuller (ADF) test [14].

#### *Immigration rate*

It is noted from Figure 1 that represents the immigration rate data for the period 2004-2018, that the highest rate of migration was in the year 2013 and the lowest rate was in the year 2008 with a general rate of (-0.172). To demonstrate the stationarity of the migration data series , (ADF) test performed. The test results shown in Table 1 was series stationary at its level. In other words, it is stable without taking the difference, which in turn means that it is integrated of zero degree.

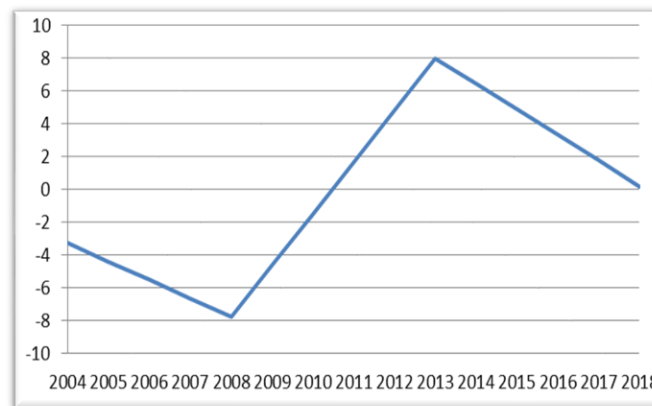


Figure 1. Evolution of immigration rate for the period (2004-2018)

#### *GDP rate*

With regard to the series of gross beekeeping products, Figure 2 shows the movement of the series during the period 2004-2018. It is noted that the highest rate of GDP was in the year 2004 and the lowest rate was in the year 2017, with an average rate of 8,860. With respect to the series stationary, the (ADF) test showed the stationarity of the series at its level, which means that it is zero integration.

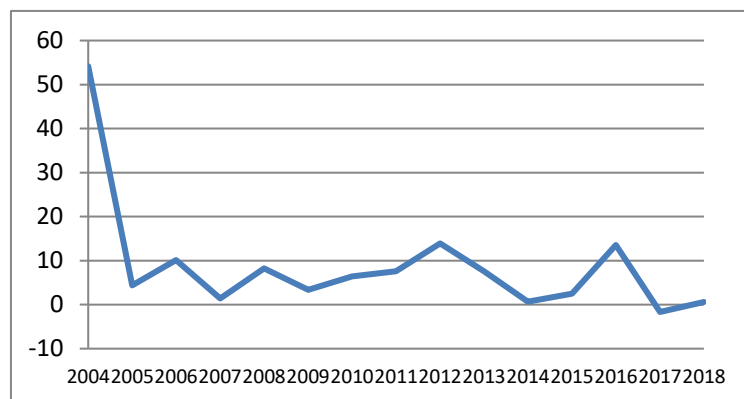


Figure 2. Evolution of the GDP rate for the period (2004-2018)

#### *Population rate*

As for the population chain, as in Figure 3, the movement of the series during the period 2004-2018, it is noted that the highest rate of population was in the year 2013 and the lowest rate was in the year 2008, with a general average of 2.736. With respect to the series stationary, the (ADF) test showed the stationarity of the series at its level, which means that it is zero integration.

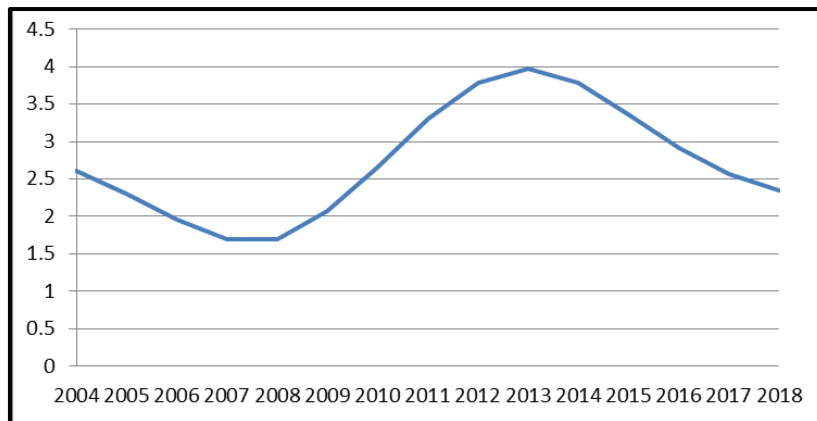


Figure 3. Evolution of population rate for the period (2004-2018)

*GDP Per capita*

With regard to the chain of GDP per capita, Figure 4 shows the movement of the series during the period 2004-2018. It is noted that the highest average GDP per capita was in the year 2009 and the lowest rate was in the year 2018 with a general average of 4.621. To demonstrate the stationarity of the GDP per capita data series, the (ADF) test showed the stationarity of the series at its level, which means that it is zero integration.

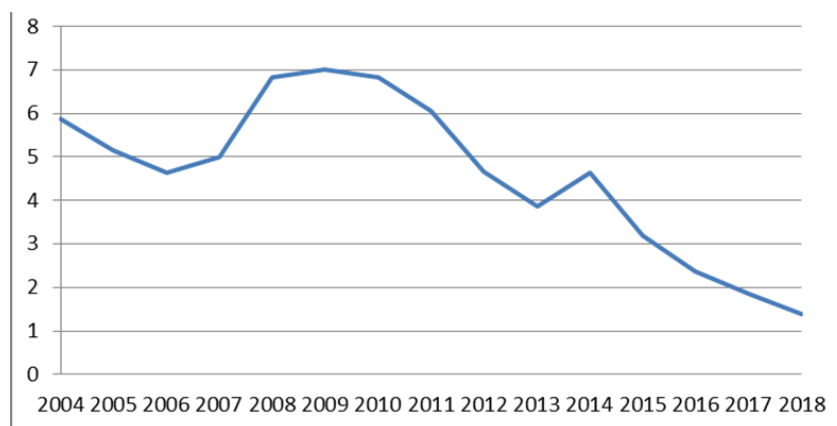


Figure 4. Evolution of GDP per capita for the period (2004-2018)

Table 1 shows the results of the Dickey-Fuller test for the stationarity of the variables included in the study during the study period (2004-2018). It is clear from the table that all variables are stationary at their level, whether at the level of significance of 1% for the GDP, POP, 5% for the MIG and 10%. For the CAPIT variable.

Table 1. Stationary test results of the variables for the period 2004-2018

Variable	t-Statistic	Prob.	Difference
MIG	-2.356	0.025	Level
GDP	-10.934	0.0000	Level
POP	-11.888	0.0000	Level
CAPIT	-2.488	0.0657	Level

After studying the behavior of the variables, we try to study the effect of the explanatory variables on the migration variable through the following multiple linear regression model:

$$MIG = B_0 + B_1GDP + B_2CAPIT + B_3POP + u$$

Whereas, MIG denotes the dependent variable and represents the percentage of Iraqi immigration. As for the POP, CAPIT, and GDP variables, they refer to the explained variables, which represent the gross domestic product, per capita GDP, and the proportion rate respectively.

By adopting the period data (2004-2018) concerning the variables, the results of estimating the parameters in the OLS method and some tests and statistical criteria were as in Table 2 as follows:

Table 2. OLS parameter estimation and test of significance

Variable	B	Std. Error	T	Sig.
Bo	-12.394	1.081	-11.468	0.000
GDP	-0.034	0.016	-2.062	0.064
CAPIT	-0.76	0.125	-6.083	0.000
POP	5.859	0.286	20.514	0.000

It is noted from Table 2 that the parameters of the GDP and CAPIT parameters take a negative sign, and it is consistent with what the economic theory says, as the increase of these variables leads to a decrease in migration, and we also note from the values of the t- statistic that the effect of these variables is significant in migration. While the parameter of the variable POP Its sign was positive, which means that the increase in the population has a positive impact on immigration. Also, we note from the value of t statistic that the effect of this variable was significant in migration.

As for Table 3, it provides a summary of some of the tests and statistical criteria related to the significance of the variables and the quality of the estimated model. As all indicators were good and in favor of the quality of the estimated model and a high degree of significance according to statistical concepts.

Table 3. Criteria of goodness of the estimated model by OLS method

R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Std. of the Estimates	F	Sig.
0.990	0.981	0.976	0.7863	188.857	0.000

The model that was estimated by the ordinary least squares OLS method may produce biased and inconsistent results because there is a correlation between the explained variables (GDP) and random error in the equation as a result of the GDP variable being affected in turn by the migration variable (MIG), meaning that it is an endogenous variable, and to solve the problem of having endogenous variables determined, Setting the search for an approximate indicator of the interpreted variable, GDP, is associated with it and at the same time not associated with random errors. This approximate indicator is known as Instrumental Variable and symbol (IV). Therefore, the model is reassessed by the two-stage least squares method (2SLS), which gives a solution to the indogeneity problem. The instrumental variables that we use in this research are represented by the life expectancy variable (LIFEXP), unemployment variable (UNEMPL) and the previous year GDP variable (LAGGDP). Accordingly, the description of the variables of the previous model will be as in Table 4 as follows:

Table 4. Description of the variables in the model

Variable	Type of Variable
MIG	Dependent
DGP	Predictor
CAPIT	Predictor
POP	Predictor
LIFEXP	Instrumental
UNEMPL	Instrumental
LAGGDP	Instrumental

Therefore, by re-estimating the previous model using the two stage least squares method (2SLS), the results were as in Table 5 as follows:

Table 5. 2SLS parameter estimation and test significance

Variable	Coefficients	Std. Error	t	Sig.
----------	--------------	------------	---	------

Bo	-10.055	3.148	-3.194	0.000
GDP	-0.059	0.034	-1.716	0.114
CAPIT	-0.931	0.241	-3.861	0.000
POP	5.376	0.805	6.682	0.000

It is noted from Table 5 that the parameters of the GDP and CAPIT variables also take a negative sign, which is consistent with what the economic theory says, as the increase of these variables leads to a decrease in migration, and we also note from the values of t-test that the effect of these variables is significant in migration. While the parameter of POP variable was positive, which means that the increase in population has a positive impact on immigration. We also note from the value of t-test that the effect of this variable was significant in migration.

Table 6 shows a summary of some tests and statistical criteria related to the significance of the variables and the quality of the model estimated by the 2SLS method. And also notes that all indicators were also in favor of the quality of the estimated model and a high degree of significance.

Table 6. Criteria of goodness of the estimated model by 2SLS method

R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Std. of the Estimates	F	Sig.
0.976	0.952	0.939	1.032	73.293	0.000

## 5. Conclusion

In this article we studied the behavior of Iraqi immigration in the period 2004-2018. The multiple regression method has been adopted in determining the most important factors affecting this phenomenon. The study showed that multiple regression methods have a great ability to characterize the phenomenon of migration by specifying a model for the impact of some factors such as the gross domestic product, gross domestic product per capita, and the size of the population in Iraq on migration. Using the OLS estimation method and the 2SLS method, we obtained a high-quality regression model that determines the amount and direction of each factor's effect on migration, whether negative or positive, and the results are highly significant. Therefore, the responsible authorities in Iraq must rely on statistical methods to develop the necessary plans to control this phenomenon and the factors that affect it as it is a drain on the workforce and an obstacle to the development of the country.

## References

- [1] N. Marzouk, " Brain Drain and its Impact on Economic Development, Syrian Economic Sciences Association, "Twenty-third Economic Forum Tuesday, Damascus, pp. 1-16, 2010.
- [2] H. A. Al-Madhi, and N. A., Mandeel, " International Migration, a study in the framework of public international law, " *Journal of Tikrit University for Law*, vol. 1, no. 2, Part 1, pp. 168-202, 2017.
- [3] H. A. Salman, "The Economic Effects of International Migration on the Economies of the Gulf Cooperation Council Countries and Evaluating Their Policies, " *Kuwait Economic Journal*, vol. 18, no. 24, pp. 134-168, 2014.
- [4] J. K. Al-Hasnawi, "The Migration of Arab Competencies for the Period (1970-2010), " *Maysan Research Journal*, vol. 12, no. 24, pp. 104-123, 2016.
- [5] H. Elasrag, "Migration of Arab Talent, " Munich Personal RePEc Archive, MPRE Paper No. 74220, pp.1-13, 2016.
- [6] A. A. Issa, "The Impact of Foreign Migration of Youth on the Political, Economic and Social Reality of Iraq (A Study in Political Geography), " *Al-Mustansiriya Journal for Arab and International Studies*, vol. 13, no. 55, pp. 116-140, 2016.
- [7] I. A. Raouf, and K. Y Al-Askari, " Study of the brain drain phenomenon, the reasons for which are remedied from the viewpoint of men and women teachers in the College of Education - Al-Mustansiriya University, " *Journal of the College of Education*, no. 2, pp. 215-252, 2007.



- [8] S. H. Raheem, " Using Factor Analysis to Determine the Most Important Factors Affecting Iraqi Youth Migration: A Statistical Study of the Status of Youth Migration in Al-Qadisiyah Governorate, " *Al-Qadisiyah Journal of Computer Science and Mathematics*, vol. 9, no. 1, pp. 1-11, 2017.
- [9] M. M. Radhi, "The illegal immigration of Iraqis, numbers, Causes and Solutions, " Iraqi Council of Representatives, Research Department, Research Department, Fourth Election Course, First Legislative Year, First Legislative Chapter, 2018.
- [10] D. Dickey and W. Fuller, " Distribution of the estimators for Autoregressive Time Series with a Unit Root, " *Journal of the American Statistical Association*, vol. 74, no. 366 pp. 427-431, 1979.
- [11] D. Dickey and W. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, vol. 49, no. 4, pp. 1057-1072, 1981.
- [12] W. H. Greene, *Econometric Analysis*, Fifth Edition, Pearson Education, Inc, 2003.
- [13] J. M. Wooldridge, *Introductory Econometrics: A Modern Approaches*, 5<sup>TH</sup> Edition, South-Western, Cengage Learning, USA, 2013.
- [14] U. Ergün and A. Göksu, *Applied Econometrics with Eviews Applications*, IBU Publications, Sarajevo, 2013.