# نموذج إحصائى لقياس العوامل المؤثرة فى هجرة الكفاءات العلمية المصرية على التنمية الإقتصادية

**د/ ممدوح عبد العليم سعد موافى**
**أستاذ مساعد بقسم الإحصاء والرياضة والتأمين**
**كلية التجارة – جامعة عين شمس**

## مستخلص

**يهدف البحث إلى قياس العوامل المؤثرة فى هجرة الكفاءات العلمية للخارج على النمو الإقتصادى المصرى وذلك فى الفترة من عام ١٩٩٠ الى عام ٢٠١٩ ، ثم التنبؤ بالناتج المحلى الإجمالى كمؤشر للتنمية الإقتصادية نتيجة أثر هذه العوامل وذلك فى الفترة من عام ٢٠٢٠ إلى عام ٢٠٣٠ ولقد تم إستخدام كل من نموذج الإنحدار المتعدد بطريقة المربعات الصغرى (OLS) Ordinary Least Squares، ونموذج الإنحدار النشط (RR) Robust Regression، ونموذج إنحدار المربعات الجزئية الصغرى Partial least squares(PLS)، ونموذج الدمج بينهما Robust Partial Least squares والمقارنة بينهم من حيث القدرة التفسيرية والتنبؤية، وتوصلت النتائج إلى أن أفضل نموذج للتنبؤ بإجمالى الناتج المحلى ($GDP_{it}$) هو نموذج الدمج Robust Partial Least squares(RPLS)حيث إرتفعت قدرته التفسيرية والتنبؤية عن النماذج الأخرى، ووجد أن أهم العوامل المؤثرة على إجمالى الناتج المحلى ($GDP_{it}$) هى مؤشر هجرة الكفاءات العلمية ($MSC_{it}$) ويقاس بنسبة إجمالى عدد المهاجرين من التعليم العالى الجامعى وفوق الجامعى إلى إجمالى عدد الخريجيين ،ومعدل البطالة($UNEMP_{it}$) ،وحجم القوى العاملة ($EMP_{it}$)، ومؤشر رأس المال البشرى($HC_{it}$)، ونسبة الإنفاق على البحث والتطوير ($ERD_{it}$)،وعدد الباحثون العاملون فى البحث والتطوير( $RRD_{it}$)،ونسبة الإنفاق العام على التعليم كنسبة من الدخل القومى ( $ES_{it}$).**

**الكلمات المفتاحية : هجرة الكفاءات العلمية ، التنمية الإقتصادية ، الإزدواج الخطى ، القيم الشاذة ، نموذج إنحدار المربعات الصغرى العادية ، نموذج الإنحدار النشط ، نموذج إنحدار المربعات الصغرى الجزئية، نموذج إنحدار المربعات الصغرى الجزئية النشط ،السلاسل الزمنية**

# A Statistical Model to Measure the Factors Affecting on the Migration of Egyptian Scientific Competencies on Economic Development

**Mamdouh Abdel Alim Saad Mowafy**

Statistics , Mathematics and Insurance Department
Faculty of Business , Ain Shams University , Egypt.
E-mail: dr.mamdouhmowafy @ bus.asu.edu.eg

## Abstract

The research aims to measure the factors affecting on the migration of scientific competencies abroad on The Egyptian economic growth from 1990 to 2019.Then, forecasting GDP as an indicator of economic development as a result of the impact of these factors from 2020 to 2030. The Research used ordinary least squares (OLS) model, robust regression (RR), Partial Least Squares (PLS) model and the hybrid model between robust regression and partial least squares . After that it was compared between them in terms of explanatory and predictive ability. The results found that the best model for forecasting of gross domestic product ($GDP_{it}$) is robust partial least squares (RPLS) model, where its explanatory and predictive ability is higher than other models. It was found that the most important factors affecting of gross domestic product ($GDP_{it}$) is the Migration of Scientific Competence Index ($MSC_{it}$) and it is measured by the total number of immigrants from higher university and university education to the total number of graduates, the unemployment rate ($UNEMP_{it}$),The size of the employees ($EMP_{it}$), the Human Capital Index ($HC_{it}$), the rate of expenditure on Research and Development ($ERD_{it}$), the number of researchers working in research and development ($RRD_{it}$), the proportion of public spending on education as a proportion of national income ($ES_{it}$).

**Keywords** :- Migration of scientific competencies, Economic development, Multicollinearity, Outliers, Ordinary least squares model, Robust regression model , Partial least squares model, Robust partial least squares model, Time series .

## 1. Introduction

The phenomenon of the migration of scientific competencies is one of the most important issues that concern the countries of the world in general and the developing world in particular, it constitutes a negative and dangerous phenomenon on these countries in terms of scientific and material as well as affecting their development plans, in addition to depriving these countries of benefiting from the experiences and qualifications of their competencies to work on their qualification over many years. The first beneficiary is the developed countries, which are based through various means and methods to attract these competencies , and the real problem in this phenomenon is not to leave these experiences and competencies their natural positions in their homeland and transfer them to advanced scientific institutions where this transition can benefit in the development of science and push its wheel forward and bring new experiences to the country, But the problem that imposes itself is the difficulty of their desire or ability to leave their new home countries and return to their home countries after receiving scientific and technical skills that can drive development in their home country. (Murad, 2015). [10]

Although the availability of highly skilled professionals in high-tech areas and scientific research is vital for the growth of an economy, the migration of human capital is inevitable. (Regional Report on Arab Labor Migration, 2006) [13] ,therefore, developing countries have been able to do little but see their most distinguished citizens go abroad to study and not return to motherland, so they decide to migrate to the

west where they can find good salaries, better working conditions, stable political systems and more distinctive lifestyles. The phenomenon of the migration of competencies is a global phenomenon linked to human development and human nature in moving from one place to another in search of a safer place to settle or gain from it. So there is a direct relationship between science and creativity, also between stability and prosperity. There is no time without the migration of scientists, whether intellectuals or worldly scientists, and it is not just one country without another, (David & et al , 2008)[5].

The phenomenon of migration or depletion of scientific minds and competencies is a major concern for the exporting countries of these minds and competencies where they are most needed to get them out of their crises and contribute to raising their economic growth rates. There is still controversy about the impact of brain drain on economic growth whether it negatively affects according to the vision of traditional literature or positive according to recent trends that the migration of minds can represent a motivation for individuals to increase investment in human capital in the hope of obtaining opportunities for migration and achieving high income levels, and this effect is called the gain of minds, and consequently the consequences of increased investment in human capital and higher economic growth. In light of this controversy over the possible consequences of brain drain or competencies on the economies of developing exporting countries, where they are most needed to achieve higher levels of economic growth, it is necessary to know how the migration of these minds can affect the rates of economic growth?.

The following presentation shows the development of the numbers of immigrants and immigrants who have acquired the status of immigrant by the immigrants with scientific qualifications according to the main undergraduate and postgraduate scientific groups during the period (2012-2018).

**Table 1**. Evolution of number of immigrants and those who acquired immigrants feature with scientific qualifications according to main scientific groups during the period (2012 - 2018)

| Source / Year | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|
| Medical Sciences (MS) | ٧٦ | 60 | 39 | 42 | 26 | 21 | 17 |
| Engineering Scien . (ES) | ٦١ | 55 | 67 | 49 | 57 | 27 | 44 |
| Agricultural Scien. (AS) | ٥ | 2 | 9 | 7 | 3 | 2 | 4 |
| Basic Sciences (BS) | ٨ | 10 | 10 | 8 | 6 | 6 | 7 |
| Humanities (Hu.) | ١٦٧ | 120 | 135 | 117 | 117 | 118 | 110 |
| **Total** | **367** | **347** | **260** | **223** | **209** | **174** | **182** |

*Source: Central Authority for Public Mobilization and Statistics, annual bulletin for those who have received approval for immigration abroad and Egyptians with another nationality in 2017 & 2018.

It is clear from the previous table that the number of immigrants working in the field of humanities (Hu.) came in first rank where it was 167 immigrants in 2012 and reached to 110 immigrants, equivalent to 61 % of the total number of immigrants in terms of education in Egypt in 2018 ,followed by the number of immigrants working in engineering sciences (ES), which reached 61 immigrants in 2012 to 44 immigrants in 2018, which is equivalent to 24 % of the total number of immigrants in terms of Educational. The

number of immigrants working in the field of medical sciences (MS) came in third rank, where it was 76 immigrants in 2012 and reach to 17 immigrants in 2018, which is approximately 9 % of the total number of immigrants in terms of education in 2018. As for the number of immigrants working in basic sciences (BS), in 2012 to 8 immigrants in 2012 and then increased to 7 in 2018, equivalent to about 4 % of the total number of immigrants in terms of education in 2018, and in the last rank the number of immigrants working in agricultural sciences (AS) reached to 5 immigrants in 2012 to 4 immigrants in 2018, equivalent to 2 % of the total number of immigrants according to education in 2018, as shown in the following figure: -

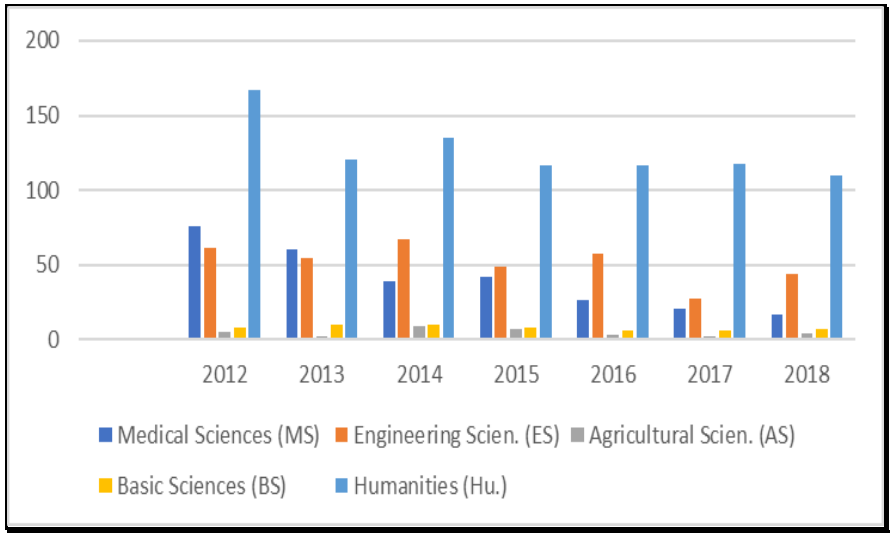



**Figure1**.Numbers of immigrants according to academic qualifications from (2012-2018).

As for the numbers of immigrants who gained the status of immigrants according to the country of immigrants during the period (2012-2018) annually appears in the table (2) which shows that the most attractive countries for the number of immigrants Italy, the United

States, Canada and Australia, where the United States of America came in first rank, where there were 112 immigrants in 2012 to reach 140 immigrants in 2018, equivalent to 37% of the total number of immigrants abroad.

**Table 2.** Total number of immigrants who acquired the status of immigrants according to country from (2012 – 2018) .

| Source | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|
| **USA** | 112 | 114 | 107 | ٧٩ | 95 | 100 | 140 |
| **Canada** | 256 | 166 | ١٧٤ | 145 | 106 | 76 | 89 |
| **Australia** | 166 | 26 | ٣١ | 21 | 24 | 11 | 10 |
| **Italy** | 80 | 114 | ١٨٨ | 155 | 135 | 154 | 133 |
| **France** | 4 | 3 | – | 1 | – | 1 | 1 |
| **England** | 2 | 3 | ٢ | 1 | – | 2 | – |
| **Germany** | – | – | ١ | 1 | 1 | 1 | 1 |
| **New Zealand** | 1 | 1 | ٢ | 1 | 4 | 2 | 1 |
| **Other Countries**** | 5 | 3 | – | 5 | 2 | 1 | – |
| **Total** | **510** | **430** | **٥٠٥** | **409** | **367** | **348** | **383** |

*Source: Central Authority for Public Mobilization and Statistics, annual bulletin for those who have received approval for immigration abroad and Egyptians with another nationality in 2017 and 2018.

**Other countries: Switzerland, Netherlands, Sweden, Cyprus, Russia, Romania, Portugal, Bulgaria, Austria, Denmark.

Italy is in the second rank which had 80 immigrants in 2012 and increased to 133 immigrants in 2018, accounted for 35 % of the total number of immigrants, Canada is in the third rank, with 256 immigrants in 2012, reached to 89 immigrants in 2018, equivalent to 23 % of the total number of immigrants. Australia is in the fourth rank with 166 immigrants in 2012 and reached to 10 immigrants in 2018, equivalent to 3%, and France, England, Germany and New

Zealand, as well as many other countries, where very few immigrants were concentrated, only 2% of the total number of migrants during a period (2012-2018).The following figure shows the continuing brain drain to abroad.

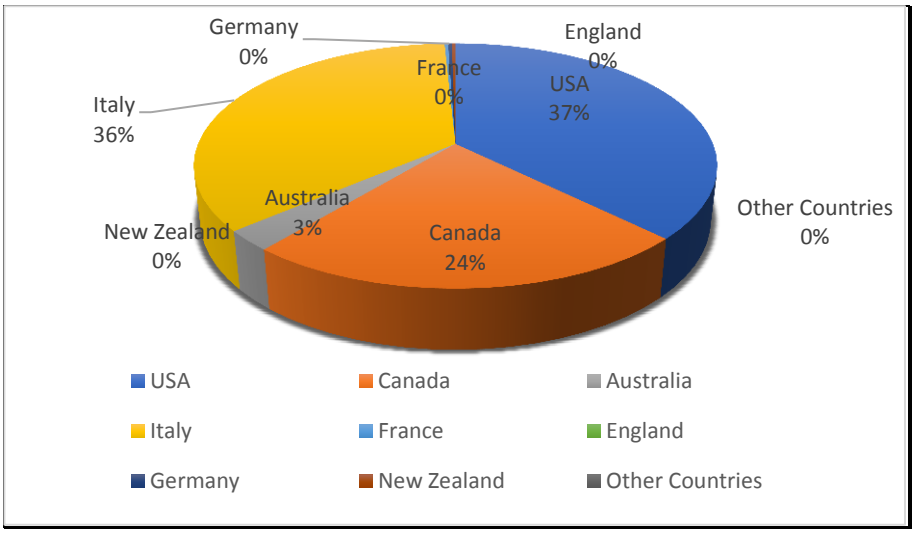



**Figure 2.** The numbers of immigrants according to the state in (2012 - 2018)

From above we note that the developing countries, especially Egypt, lose human capital, which is the key to economic and social development, and therefore the current study aims to shed light on the reasons for the migration of Egyptian minds abroad in their various dimensions, which pose a great danger to Egypt's progress and growth. This helps in identifying the problems facing these scientists and experts and trying to remove them by taking appropriate mechanisms to support the efforts of Egyptian immigrant skilled people in the development of Egypt. The current research focused on the period from 1990 to 2019 and try to predict GDP as a result of the impact of these factors from 2020 to 2030. the

researcher will make a comparison between ordinary least squares (OLS) model , partial least squares, robust regression model and the hybrid model between , robust regression and partial least squares, and determine which has the highest explanatory and predictive ability for GDP from 2020 to 2030, the study also aims to test the validity of the following hypotheses:-

- There is no positive significance relationship between gross domestic product ($GDP_{it}$) as a dependent variable and the volume of employment ($EMP_{it}$), Human Capital Index ($HC_{it}$), the rate of expenditure on research and development ($ERD_{it}$), the number of researchers in research and development ($RRD_{it}$) , and the proportion of public spending on education as a proportion of national income ($ES_{it}$) as independent variables.
- There is no inverse significance relationship between gross domestic product ($GDP_{it}$) as a dependent variable and both of the migration scientific competence Index ($MSC_{it}$) and the unemployment rate ($Unemp_{it}$) as independent variables.
- The hybrid model which called robust partial least squares model is the better for predicting to GDP than ordinary least squares model (OLS), robust regression model and Partial least squares model.
- An increase in gross domestic product (GDP) is expected as a result of factors affecting the migration of Egyptian scientific competencies abroad.

The study will be based on obtaining its data through economic journals, statistical bulletins and reports issued by the following entities Central Agency for Public Mobilization and Statistics, World Bank Organization).

## 2. Literature Review

Many Arab and foreign studies have presented the factors affecting the phenomenon of the migration of scientific competencies abroad and their causes, whether the causes are repellent or attractive, including:

Kazlauskiene & Rinkevicius, (2006) [9] aimed at identifying the pushing and attracting causes of competencies in Lithuania and recognizing the factors causing the migration of highly skilled people in Lithuania to obtain a good evaluation of the problem of 'competencies migration' and solve it in correct ways. It has found that the pushing factors of migration play a much greater role for highly skilled migration than attracting factors, and that migration is always for rich countries, but this does not mean that migratory competencies are always from poor countries. The study also found that the poor social and economic situation in Lithuania is one of the most important pushing of competencies.

Raouf & Elaskry, (2007) [12] aim at identifying the reasons that led to the migration of minds from the point of view of scholars in

the scientific and human departments in the faculty of education at Al-Mustansiriyah University in Iraq and reaching proposals to reduce this phenomenon . The study found that is necessary from the government to interest the immigrants, research centers, public libraries, provide resources and the universities have to put plans to stop the bleeding of brain drain.

Al-Nadwi, (2007) [2] aimed to identify the phenomenon of the migration of Arab minds, identify the pushing and attract factors for them and know the dangers of the migration of Arab minds, It also found that there are different motives for this phenomenon political, economic, social, intellectual and others, but economic factors were a priority in the direct impact on Arab minds and competencies, especially since the people most affected by this factor are the best prepared and most competent people.

Tawfik, (2007) [21] aimed to identify immigrants from the North Africa countries to the European countries in terms of the reasons that led them to emigrate, their economic and social situation and ways to support their links to their original society. In addition to the role that they can play in communication between their countries and countries of emigration at all levels. Also the study found that build bridges of communication with its immigrants abroad and the research and diplomatic activity seeks to invest their presence in the countries of the migration to strengthen the relationships with these countries and emphasize their role in the development process in their home country.

Al-Badrani, (2009) [1] aimed to clarify and estimate the extent of the losses achieved due to the drain of skilled Arab minds and to diagnose and understand the reasons for the migration of Arab competencies to developed countries. Identifying successful ways and means in reducing the volume of the flow of Arab competencies in order to benefit from them in strengthening and building the economies of Arab countries. Also the study has reached some mechanisms that contribute to reducing the migration of Arab competencies and achieving the benefit from their energy.

Tessema, (2009) [22] aimed to measure the deeper into the causes of brain migration from less developed countries, through the adoption of Eritrea as a case study, the study also highlighted the impact of brain drain. The results of this study have found that the deterioration of economic conditions, lack of good governance and weak causes of brain drain in less developed countries such as Eritrea. The study suggested some measures that could be taken by governments of least developed countries that contribute to reducing this phenomenon.

Al-Mansouri and Al-Daikh, (2010)[3] aimed to identify the numbers of Arab qualified people migrating to Western countries by looking at statistics and reports that show this, and trying to identify the causes of the loss of this important economic resource. Also it has found that economic factors have been and continue to be a priority in the direct impact on Arab minds and competencies, especially since the people most affected by this factor are the best prepared

and most efficient people to run the production, education and teaching sectors in the Arab world.

Omamegbe, (2010) [11] has been uncovering the main causes of human capital flight from Africa to the United States of America, as well as suggesting ways to mitigate the adverse effects , or maximize the potential benefits of human capital migration from Africa to the United States, It has concluded that the governments of sending and receiving countries should work together to formulate a mutually beneficial policy, and that sending countries should focus on softwares development and create favorable for maintaining competencies at motherland.

Semela (2011) [20] aimed at identifying the causes of brain drain among academics in Ethiopian higher education institutions and also discussed the implications of this, It has found that the functional changes in these institutions, particularly the material return, are one of the most important reasons for the brain drain from these institutions.

Shumba & Mawere , 2012)[ [19] aimed at identifying the causes of brain drain, the impact of mass brain migration on higher education and on the quality of graduates in Zimbabwe and efforts to address the problem of brain migration. It has found that there are push and pull factors affecting brain drain in Zimbabwe, including push factors and attraction affecting brain drain in Zimbabwe, the main push factors include on the low of salaries and job satisfaction , poor funding, political instability, while the attraction factors included

high salaries, research and study opportunities that cannot be achieved at motherland.

Yuan Li, (2012) [26] aimed to analyze the factors of brain drain in both China and India, also discussed the costs of brain drain on the source country and the benefits to the host country and the attitudes and policies of the Chinese and Indian governments on brain drain . It found that there are complex and difficult challenges facing the government and decision makers to cope with the negative effects of brain drain, and the study recommended a number of mechanisms that help reduce the migration of competencies.

Vogo, (2014) [23] aimed to identify the reasons for the migration of Palestinian competencies to developed countries in order to determine the best ways to reduce this phenomenon, and to shed light on the motives that can positively affect the decision to return immigrant skilled people to the homeland, and has reached the political, economic and social conditions of one of the most important factors of the migration of Palestinian scientific competencies, as well as the attention to scientific research is one of the factors that contribute to reducing the migration of Palestinian competencies.

Murad, (2015)[10] aimed to know the causes of brain drain in 2014 and the impact on the rate of economic growth, and how to remove obstacles to these competencies to return from abroad and benefit from them. The study used the questionnaire in 2014, which distributed to university professors . The most important

reasons for the Egyptian competencies from the point of view of the sample of research were poor salaries and poor interest in scientific research, violation of academic freedoms and poor appreciation of the political systems governing the distinct minds, administrative bureaucracy and red tape. As for the attractive factors of competencies were the outstanding financial return of qualified people abroad and increased interest in scientific research, the quality of the education system abroad, and the scientific and technological leadership of developed countries.

The current study was to study the factors affecting the migration of scientific competencies abroad on the growth of the Egyptian economy from 1990 to 2019 and comparing between ordinary least squares ,robust regression ,partial least squares and robust partial least squares models, addition to select the best model and try to predict GDP as a result of the impact of these factors in the period from 2020 to 2030.

## 3. Methodology and Data Analysis

### 3.1 OLS Regression Model

The multiple linear regression model is one of the statistical models needed to study the relationship between the dependent variable (Y) and a number of independent variables (X) whose model takes the following form:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ..... + \beta_k X_{ik} + \varepsilon_i \qquad i = 1, 2, \ldots, n \qquad (1)$$

Where $Y_i$ is dependent variable, $X'S$ Represents independent variables , $(\beta_0, \beta_1, \beta_2, .....\beta_k)$ represent the regression of coefficients . $\varepsilon_i$ is random error for observation $i$ ، $i = 1,2,...,n$where $n$ represents the number of observations and can calculate an estimated $\hat{\beta}$, OLS using matrices from the following equation: -

$$\hat{\beta}_{LS} = (X^t X)^{-1} X^t Y \qquad (2)$$

Where $(X^t X)$ represents simple correlation coefficients matrix between each pair of independent variables, and . $(X^t Y)$ represents the vector of correlation coefficients between the dependent variable and each independent variable can be calculated and the estimated variance of the OLS squares can be calculated from the following equation:

$$MSE(\hat{\beta}_{LS}) = Cov(\hat{\beta}_{LS}) = \sigma^2 (X^t X)^{-1} \qquad (3)$$

In practical reality, we note that there are many problems in the model of linear regression, including the problem of multicollinearity between independent variables and the problem of outliers that are found in the data of the dependent and independent variables, and the problem of linear multicollinearity leads to an inflation in the variations of the estimates of the regression coefficients (Habshah and Marina, 2007) [6] . And subsequently in order to reduce these Variations we use the PLS model, which improves the accuracy of coefficient estimates using the OLS method (Samkar & Aplu, 2010) [18].

### 3.2 Robust Regression Model

Ronchetti, 1987[14] noted that the purpose of using the robust regression model is to select the model that fit the data, also keeping in mind that the errors may not follow the standard normal distribution. Several robust regression estimates have been proposed which represented in least trimmed squares (LTS ), robust M , robust MM, least absolute deviations (LAD), least median squares (LMS) , and robust S estimator. Rousseeuw, (1984 &1985) [15,16]. This study will concern on robust M , robust MM because they are the best estimators of robust regression model , we can show them as follows :

• (Huber,1972a) [8] presented robust M Regression estimator. It uses to calculate model estimates by reducing random error multiplication weights weighted by Huber function. the estimate calculates as follows: -

$$\widehat{\beta}_M = \min_{\beta} \sum_{i=1}^{n} \rho(r_i) \qquad (4)$$

Yohai, 1987,[25] also introduced robust MM, which is more widespread because it is the most resistant to outliers observations in the direction of both the dependent variable and the independent variables. The name of MM is due to the use of the M-estimator in more than one stage to calculate the final estimates, and the MM estimate is obtained in several stages . Finally it is getting the infinty least absolute as follows :-

$$L(\widehat{\beta}_{MM}) = \sum_{i=1}^{n} \rho_1 \left( \frac{r_i}{\widehat{\sigma}_{(0)}} \right) x_i = 0 \qquad (5)$$

### 3.3 Partial Least Squares Model

The need to use the partial least squares (PLS) method has emerged in many studies and social sciences .Wold, (1966) [24] was the first to find this method. The purpose of the PLS method is to predict the dependent variables $(X's)$ , and there are two cases in which the first partial least squares, which are characterized by PLS1 that used to predict the values of a single dependent variable $(Y)$ with a number of independent variables $(X's)$. The second case (PLS2) is the application of this method to predict the values of a number of dependent variables $(Y's)$ with a number of independent variables $(X's)$ and this method requires repeatedly to reach the final solution Harald & Luis (1986) [7]. The method of the partial least squares regression depends on the matrix of the common variance $COV(X,Y)$ . This method allows for the identification of factors, which are linear compositions of independent variables $(X's)$, these factors are known as latent variables, which in turn are the best model for dependent variables $(Y's)$, and the partial least squares method is similar to both the canonical correlation method , the principal component method and the discriminant analysis, but it has conditions when applied( Dante : 2006) [4]:-

- Extracted Factors for each of the changes $(Y)$ and variables $(X)$ they're derived from matrices $(Y'Y)$ and $(X'X)$ only Not from the arrays that came from the product of multiplying the variables $(Y)$ in $(X)$ That's not $(X'Y)$ or what else.
- The number of independent equations should not exceed the minimum number of dependent variables and for the method of partial least squares , it is an extended method of multiple linear regression but without the previous two conditions.The independent equations represented by the factors obtained from the matrices $(Y'X'XY)$ to get a regression model to predict values $(Y)$

The partial least squares method is called the endocrine or subjective variables sometimes because they are similar to the subjective variables of self analysis and this formula can be described as the following, and let us have the matrix of independent variables modified . Let's have a matrix of independent variables modified $(X's)$ multiplied by a random vector before $(U)$.

$$X'U = \begin{bmatrix} \sum X_{i1}U_i \\ \sum X_{i2}U_i \\ . \\ . \\ \sum X_{ip}U_i \end{bmatrix} \quad (6)$$

From this matrix you can get the coefficient matrix $(\beta_0, \beta_1, \beta_2, .....\beta_k)$ Which shows the type and shape of the relationship between each dependent variable and an independent variable.

The partial least square method is a linear combination of the least squares of the correlation matrix and the common contrast between the independent variables and the dependent variables, and the part on which the partial least square method in the correlation matrix and the co-variance matrix is cross block part that means correlations between independent and dependent variables, as this method provides factors scores in the form of linear groups between the original independent variables, so there will be no correlations between the variable factors used in regression model and can analysis two matrices $(X)$ and $(Y)$ from the following relationship (Saikat and Jun : 2008) [17] .

$$X = TP' + K \quad (7)$$

$$Y = UC' + R \quad (8)$$

Where $T$ is matrix of factors extracted from the matrix $(X)$ , $P$ is a loaded vector for the matrix $(X)$ , $U$ is a matrix of factors extracted from the matrix $(Y)$ , $C$ is a loaded vector of the matrix $(Y)$ . $R$ means a linear combination of matrix columns $(X)$ , after that is multiplied by the matrix $(X)$ to find linear compounds t as follows :

$$T = XW \quad (9)$$

where $W$ is a vector for random values or is a distinct vector of the matrix $(Y'X'XY)$, In the same way the linear composition of the matrix $(Y)$, which is represented by the vector $U$ as follows :-

$$U = YC \quad (10)$$

As $C$ is the first distinct vector and has a distinct root of the matrix $(Y'X'XY), (X'Y)$ is the contrast variance matrix between $(X)$ and $(Y)$

This method looks for a set of components called latent vectors which illustrate as much as possible the co-variance for $(X)$ and $(Y)$ Independent variables can be analyzed by the following equation:

$$X = TP' \quad (11)$$

T is a linear set of independent variables but in orthogonal factors form . Each column contains all the independent variables in $(X)$, but in the form of a linear set of weights , $P$ is a loaded vector, which is a linear set of orthogonal factors $t$ , And the original matrix of independent variables :

$$P = X't \quad (12)$$

**So that $t$ is a column of the T matrix**

$$TT' = 1 \quad (13)$$

$$PP' \neq 1 \quad (14)$$

After finding the first self-vector we will subtract it from $(X)$ and $(Y)$ . This procedure will repeat until $(X)$ becomes a zero matrix. Finally, the hybrid model that called the robust partial least Squares (RPLS) model will be used .

## 4. Results of Statistical Models.

This part includes an application of statistical models where we begin to estimate and evaluate the multiple linear regression model ( OLS) and then conduct diagnostic tests of the model and then we detect the extent of the problems of multicollinearity and outliers in the data of GDP function. To analysis the data will use EVIEWS 10 and SYSTAT 12 softwares . The following equation will illustrate gross domestic product GDP function and the impact of factors affecting on the migration of scientific competencies function and is measured by the total number of immigrants from undergraduate and postgraduate to the total number of graduates, Unemployment rate ($UNEMP_{it}$), volume of employment ($EMP_{it}$), Human Capital ($HC_{it}$) measured by total number of graduates to population, rate of expenditure on research and development ($ERD_{it}$) and number of researchers in research and development ($RRD_{it}$), education expenditure ratio as a percentage of national income education spending ($ES_{it}$), the GDP function takes the following form: -

$$GDP_{it} = MSC_{it} + UNEMP_{it} + EMP_{it} + HC_{it} + ERD_{it} + RRD_{it} + ES_{it} + \varepsilon_{it} \quad (15)$$

Table 3 shows the estimates of ordinary Least squares model results as follows:

**Table 3**. OLS model estimates for GDP function

| Variable | Coeff. | Std. Error | T – stat | Prob. | VIF |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| **C** | -2.3E+12 | 3.61E+11 | -6.35236 | 0.000 | - |
| **MSC** | -9.7E+12 | 3.17E+13 | -0.30488 | 0.763 | 2.25408 |
| **UNEMP** | -4.0E+10 | 1.25E+10 | -3.20688 | 0.004 | 2.48666 |
| **EMP** | 1.35E+08 | 11928912 | 11.32339 | 0.000 | 22.2860 |
| **HC** | 1.84E+14 | 3.64E+13 | 5.061101 | 0.000 | 1.44445 |
| **ERD** | 9.08E+08 | 2.48E+08 | 3.659592 | 0.001 | 11.9600 |
| **RRD** | 9.60E+11 | 2.24E+11 | 4.274738 | 0.000 | 1.92655 |
| **ES** | 3.99E+10 | 6.77E+10 | 0.589281 | 0.562 | 3.26617 |
| **R-squared** | 0.9945 | **Mean dep. var.** | | 2.22E+12 | |
| **Adjusted R- Squ.** | 0.9928 | **SD. dep. Var.** | | 8.36E+11 | |
| **S.E. of reg.** | 7.11E+10 | **Akaike inf. crit.** | | 5.30E+01 | |
| **Sum squ. resid.** | 1.11E+23 | **Schwarz crit.** | | 5.34E+01 | |
| **Log likelihood** | -787.540 | **Hannan-Qui.cr.** | | 5.32E+01 | |
| **F-statistic** | 569.847 | **Durb. -Wat .stat** | | 2.13E+00 | |
| **Prob (F-statistic)** | 0.000 | **Mean dep. var.** | | 2.22E+12 | |

the previous table shows that there is a significance relationship between the gross domestic product ($GDP_{it}$) and unemployment rate ($UNEMP_{it}$),size of the workforce ($EMP_{it}$), Human Capital Index ($HC_{it}$), rate of expenditure on research and development ($ERD_{it}$) and number of researchers in research and development ($RRD_{it}$) as independent variables.There is a significant model where the calculated (F) value was 569.847 and the value of P-Value less than 5% and the value of the determination coefficient was (0.993) which reflects the high explanatory ability of the model that means that the independent variables included in the model explain 99.3% of the changes in the dependent variable and that 0.7 % of the changes are due to random errors.

It is also clear from table 3 that the estimated model has not autocorrelation, where Durbin-Watson statistic value was 2.13. from

the previous table, we also note that the VIF values for all independent variables are less than 10 except two variables the size of the employment ($EMP_{it}$) and the rate of expenditure on research and development ($ERD_{it}$), which means that there is multicollinearity between independent variables.

Table 4 shows the diagnostic tests of the model, where the Breusch – Godfrey LM Test shows that there is no problem of autocorrelation between the errors where the value of obs * R-squared was 0.41437 and that the value of Prob.Chi-Sq.(2) was 0.8129 which is greater than 5% that means accepting the null hypothesis which shows there not autocorrelation between the errors.

**Table 4** . Tests of Autocorrelation , Heteroskedasticity, and Normality test .

| | | | |
|---|---|---|---|
| **Breusch-Godfrey Serial Correlation LM Test :** | | | |
| **F-statistic** | 0.14006 | **Prob. F(2,20)** | 0.8702 |
| **Obs*R-squared** | 0.41437 | **Prob. Chi-Square(2)** | 0.8129 |
| **Heteroskedasticity Test: Breusch-Pagan-Godfrey** | | | |
| **F-statistic** | 0.9156 | **Prob. F(7,22)** | 0.5131 |
| **Obs*R-squared** | 6.7684 | **Prob. Chi-Square(7)** | 0.4534 |
| **Scaled expla.SS** | 3.8846 | **Prob. Chi-Square(7)** | 0.793 |
| **Normality Test** | | | |
| **Jarque – Bera** | 0.77292 | **Prob.obs(20)** | 0.67945 |

The Breusch Pagan – Godfrey test results also show that the estimated model does not have Heteroskedasticity between the errors, where obs*R-squared was 6.7684 and the Prob. value. Chi-Sq.(7) is equal to 0.4534 which is greater than 5%, which means

accepting the null hypothesis , which confirms the absence of a Heteroskedasticity , as well as the previous table shows the results of Jarque – Bera test , which shows the series of errors of the estimated model does not follow the normal distribution where the value of the test is JB = $0.77292$ and the value of p-value was equal to $0.67945$ which is Higher than $5\%$.

As for the existence of outliers values and how to detect them, through the following graphs and using the Box Plot, it was found that there are outliers in only four variables whish are Migration of the Scientific Competence Index ($MSC_{it}$), the Human Capital Index ($HC_{it}$), and the number of researchers in research and development. ($RRD_{it}$), and the proportion of public spending on education ($ES_{it}$) as a proportion of national income as shown in figures $4$, $8$, $9$, and $(10)$, which will lead to biased estimates and an inflation in the differences in variances estimates coefficients which requires the use of the estimates of robust regression models and the partial least squares regression in order to remove the impact of outliers and multicollinearity.

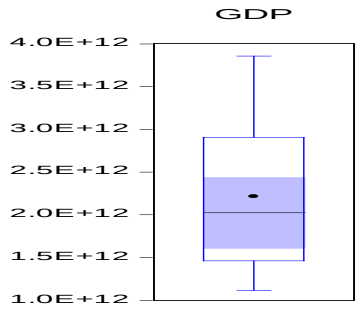



**Figure 3.** Gross Domestic Product

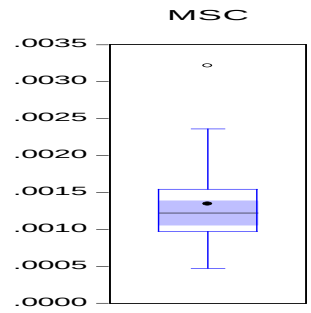



**Figure 4** . Migration of Scientific Competences

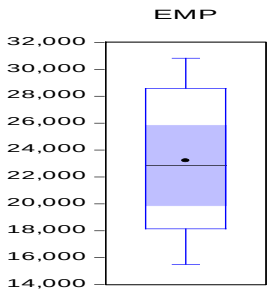



**Figure 5** . Employment

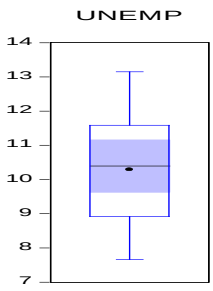



**Figure 6** . Unemployment

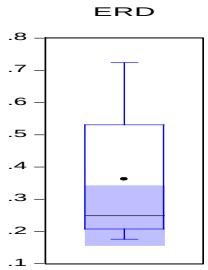



**Figure 7.** Expenditure on Research and Development

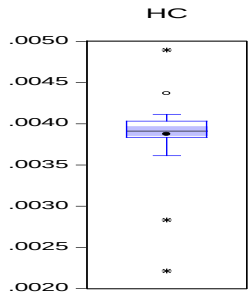
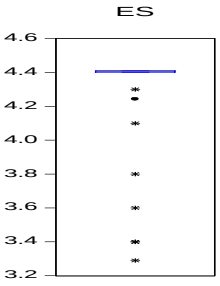



**Figure 8.** Human Capital Index



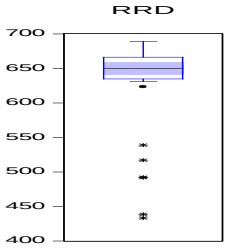




**Figure10.** Researchers in Research

| | |
|---|---|
| **Figure 9**. Education Spending | and Development. |

Table 5 shows a comparison of GDP function estimates using ordinary least squares (OLS) method, robust regression model(RR), partial Least squares(PLS) and the hybrid model which called RPLS:-

**Table 5** . Comparing of GDP function estimates using models ( OLS , Roust M , Robust MM , Partial LS, Robust Partial LS )

| Var. | Cons. | MSC | UNEMP | EMP | HC | ERD | RRD | ES |
|---|---|---|---|---|---|---|---|---|
| | | | | **OLS** | | | | |
| **Coff.** | −2E+12 | −10E+12 | −4E+10 | 1.4E+08 | 2E+14 | 9.1E+08 | 10E+11 | 4E+10 |
| **P.value** | 0.000 | 0.763 | 0.004 | 0.000 | 0.000 | 0.001 | 0.000 | 0.562 |
| | | | | **Robust M Reg .** | | | | |
| **Coff.** | −3E+12 | −2E+13 | −4E+10 | 2E+08 | 2E+14 | 1E+12 | 9E+08 | 6E+10 |
| **P.value** | 0.000 | 0.6433 | 0.0034 | 0.000 | 0.000 | 0.342 | 0.0001 | 0.387 |
| | | | | **Robust MM Reg .** | | | | |
| **Coff.** | −2.4E+12 | −2E+13 | −3E+10 | 1E+08 | 2E+14 | 1E+12 | 8.4E+08 | 7E+10 |
| **P.value** | 0.000 | 0.5665 | 0.0346 | 0.000 | 0.000 | 0.000 | 0.0034 | 0.320 |
| | | | | **Partial LS Reg .** | | | | |
| **Coff.** | −2E+12 | −10E+12 | −4E+10 | 2E+08 | 2E+14 | 10E+11 | 9E +08 | 4E+10 |
| **P.value** | 0.000 | 0.017 | 0.018 | 0.000 | 0.000 | 0.020 | 0.020 | 0.000 |
| | | | | **Robust Partial LS Reg .** | | | | |
| **Coff.** | −2E+12 | −10E+12 | −4E+10 | 1.4E+08 | 2E+14 | 10E+11 | 9E+08 | 4E+10 |
| **P.value** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

| | OLS | Robust M | Robust MM | PLS | Robust Rartial LS |
|---|---|---|---|---|---|
| **F-test** | 569.85 | 4676.91 | 4709.081 | 569.85 | 9.6E+08 |
| **P.value** | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **$R^2$** | 0.994 | 0.968 | 0.839721 | 0.995 | 0.999 |

The previous table shows a comparison of many of the estimates of regression models, and it has been shown that the best significance model with the highest explanatory ability is robust partial least squares model which shows as follows : -

• There is a significance relationship between gross domestic product ($GDP_{it}$), and Migration of Scientific Competence Index ($MSC_{it}$), Unemployment rate ($UNEMP._{it}$), the size of employment ($EMP_{it}$), Human Capital Index ($HC_{it}$), rate of expenditure on research and development ($ERD_{it}$), number of researchers in research and development ($RRD_{it}$), proportion of public education spending as a proportion of national income ($ES_{it}$) as follows:-

$$GDP_{it} = -2E+12-10E+12MSC-4E+10UNEMP+1.4E+08EMP$$
$$+2E+14HC_{it}+10E+11ERD_{it}+9\text{E+08 RRD}_{it}+4\text{E+10OP}_{it} \quad (16)$$

• There is a significant model where the calculated of (F) value was 9.6E+08 and the value of P-Value less than 5% and the value of determination coefficient was (0.999), which reflects the high explanatory abilities of the model and means that the independent variables explain 99.9% of the changes in the dependent variable and that 0.1% of the changes are due to random errors.

**Table 6**. Predictability ability measurement of regression models

( OLS ,Roust M , Robust MM , Partial LS, Robust Partial LS )

| MODEL | RMSE | MAE | MAPE |
|---|---|---|---|
| **OLS** | 6.٠٩E+10 | 4.٩١E+10 | 2.36 |
| **Robust M .** | 6.١٦E+10 | 4.7٣E+10 | 2.2٧ |
| **Robust MM .** | 6.٢٨E+10 | 4.٧٠E+10 | 2.2٥ |
| **Partial LS** | 6.192E+10 | 4.86E+10 | 2.33 |

| | | | |
|---|---|---|---|
| **Rob. Partial LS** | 1.56E+08 | 1.11E+08 | 0.01 |

• The previous table shows that the best model is robust rartial least squares model (RPLS), which has the highest predictive ability according to (RMSE, MAE, MAPE ) criteria that their values are 1.56E+08, 1.11E+08 and 0.01 respectively .So this model we can use it to forecast of GDP variable from 2020 to 2030.

Table 7. the predictive values of Egypt's GDP using time series from 2020 to 2030 as follows:

**Table 7**. Predictive values and confidence intervals for GDP

| Year | Forecasts | LCL | UCL |
|---|---|---|---|
| **2020** | 3.9566E+12 | 3.77E+12 | 4.15E+12 |
| **2021** | 4.14229E+12 | 3.87E+12 | 4.41E+12 |
| **2022** | 4.32798E+12 | 3.94E+12 | 4.72E+12 |
| **2023** | 4.51366E+12 | 3.97E+12 | 5.06E+12 |
| **2024** | 4.69935E+12 | 3.98E+12 | 5.42E+12 |
| **2025** | 4.88504E+12 | 3.96E+12 | 5.81E+12 |
| **2026** | 5.07073E+12 | 3.94E+12 | 6.21E+12 |
| **2027** | 5.25641E+12 | 3.89E+12 | 6.62E+12 |
| **2028** | 5.4421E+12 | 3.83E+12 | 7.05E+12 |
| **2029** | 5.62779E+12 | 3.76E+12 | 7.50E+12 |
| **2030** | 5.81348E+12 | 3.67E+12 | 7.95E+12 |

• It is clear from the previous table that using the estimated values of the robust partial least squares from 2000 to 2019, we can forecast Egypt's GDP from 2020 to 2030 using time series, which it explained

in Table 7 where that GDP is expected to increase from 3.9566E+12 pounds in 2020 to reach 5.81348E+12 pounds in 2030.

## 5. Conclusion

The current research studied the factors affecting on the migration of scientific competencies on GDP from 2000 to 2019. the aim of this study is reaching to the best model for forecasting GDP from 2020 to 2030. It compares between ordinary least squares model, robust regression model , partial Least squares model and robust partial least squares model in terms of explanatory and predictive abilities , the study was based on the test of four statistical hypotheses and the results were as follows:

(1) Rejection of the first hypothesis because the robust partial least squares model has shown that there is positive significance relationship between gross domestic product ($GDP_{it}$) as a dependent variable and the size of the employment ($EMP_{it}$), Human Capital Index ($HC_{it}$), rate of expenditure on research and development ($ERD_{it}$), the number of researchers in research development ($RRD_{it}$), and the ratio of public spending education ($ES_{it}$) as a proportion of national income as independent variables.

(2) Rejection of the second hypothesis because the robust partial least squares model has shown that there is negative significance

relationship between gross domestic product ($GDP_{it}$) as a dependent variable and Migration of Scientific Competence Index ($MSC_{it}$) and Unemployment rate ($UNEMP_{it}$) as independent variables.

(3) Acceptance of the third hypothesis which shows that robust partial least squares model is better for predicting of GDP than ordinary least squares model (OLS), robust regression model and partial least squares model.

(4) Acceptance of the fourth hypothesis because the results showed an increase in GDP during the forecast period from 2020 to 2030.

(5) The results also showed a negative impact of the migration of scientific competencies on the gross domestic product, which leads to making the migration of scientific competencies a pushing factor for individuals and the state in general to make investment in human capital, and this confirms that there is a positive relationship between human capital and gross domestic product (GDP) as shown by the results of the study.

(6) Policy makers should develop programs and policies that will make the phenomenon of the migration of scientific competencies a gain rather than a drain, which leads to the polarization of these minds, such as providing them with incentives to stay at motherland , both by reviewing their income level, by increasing spending on

research and development and education in general, in addition to increasing the number of researchers in research centers as they have a positive impact on GDP as indicated by the study results.

(7) The state should provide a specific support or assistance to students returning from abroad after receiving education or training in an attempt to attract high-productivity individuals to return homeland, in addition to providing them with jobs with appropriate returns, which will help them to make the decision to return home permanently.

(8) The state must communicate with immigrant scientific competencies to benefit from their experiences, for example through the universities adopting the idea of inviting these scientific competencies to give lectures and participate in conferences and workshops , and benefit them in the fields of research, education and training, such as these scholars Magdi Yaacoub, Ahmed Zweil and Farouk El Baz, who benefited the country with everything their learned abroad.

**References**

[1] Al- Badrani, A. A. (2009). **Arab Competence Migration, Causes and Results (Iraq Model),** Unpublished Master's Thesis, Faculty of Management and Economics, Open Arab Academy, Denmark, Kopenhagen.

[2] Al-Nadawi, K. A. (2007), **Brain Drain**, is available on the website http://www.iraqwriters.com/np/views.asp?ID=67

[3] Al-Mansouri, M. and Al- Daikh, A. A. (2010). **The migration of Arab minds, its causes and economic effects, the Foundation for Arab Thought** http://www.arabthrough.org/node/305#.

[4] Dante, M. P. (2006). **An overview of partial least squares**, The Paul Merge School of Business University of California, Irvine, pp1-15.

[5] David, Z. and others (2008 ). **Redefining the brain drain : china`s diaspora option** , Science Technology Society,28 May. http:// sts. Sagepub.com/contentl,13/1/1.

[6] Habshah ,M., and Marina, z. (2007), **A Simulation study on ridge regression estimators in Presence of outliers and multicollinearity** , Journal Teknologi, Universiti Teknologi Malaysia ,47, 59-74.

[7] Harald , M. and Luis, I . (1986) . **Partial least squares regression on design variables as an alternative to analysis of variance** , Analytica Chimica Act, 191, pp133-148.

[8] Huber, P.J.(1972a ) , **Robust statistics** , Ann. Math. Statist., 43, 1041-1067 .

[9] Kazlauskiene,A. and Rinkevicius ,L .(2006).**Lithuanian Brain drain causes : push and pull factors** , Engineering Economics , Vol.46, No.1 .

[10] Murad, H .A. (2015). **Reasons for the migration of Egyptian minds (field study)** , master's thesis, Faculty of Education, Damietta University.

[11] Omamegbe,M.E. (2010). **African brain drain : a Narrative inquiry on human capital flight from sub saharan Africa to the united states** , A Doctor`s degree of philosophy, The university of texas at Dallas.

[12] Raouf, I. K. and Al-Askari, K. Y. S., (2007) . **Study of the Phenomenon of minds (causes - treatment) from the point of view of male and female students in the Faculty of Education Of The University of Al-Mustansiriyah**, Journal of the Faculty of Education, Al-Mustansiriyah University, Volume 1, Issue 2.

[13] **Regional Report on International Labour Migration** (2006). Department of Population and Immigration Policies, Social Sector, Arab League Cairo. pp 1057-1072 .

[14] Ronchetti , E.M. (1987). **Robust** $C(\alpha)$ **- type tests for linear models**, Sankhya. A,49 , 1-16 .

[15] Rousseeuw,P.J.(1984), **Least median of squares regression**, Journal of the American Statistical Association, 79, 871- 880 .

[16] Rousseeuw , P. J. (1985) , **Multivariate estimation with high breakdown point in Mathematical Statistics and Applica-tions** , B ( ed . by W. W . Wertz ) , Dordrecht, Netherlands : Reidel Publishing Company , 283-297 .

[17] Saikat , M. and Jun, Y. (2008) . **Principle Component Analysis and Partial least squares : Two dimension reduction techniques for regression**, Causality Actuarial Society , pp79-90

[18] Samkar , H., and Alpu, O.(2010 ), **Ridge regression based on some robust estimators ,** Journal of Modern Applied Statistical Methods ,9,2, Article 17.

[19] Shumba, A. and Mawere,D.(2012). **The causes and impact of the brain drain in institutions of higher learning in Zimbabwe**, International Migration, Vol.50 , No.4.

[20] Semela, T. (2011) . **Vulnerability to brain drain among Academics in institutions of higher learning in Ethiopia,** Asian Social Science , Vol.7 ,No.1 .

[21] Tawfik , R., .(2007) . **Migration of North Africans to Europe (Analysis of Causes and Motives) A seminar entitled Arab Expatriates from North Africa to the European country**. Arab League, (Arab Expatriates Management) in cooperation with the Egyptian-African Studies Program at Cairo University, April.

[22] Tessema, M.(2009) . **Causes challenges and prospects of brain drain: the Case of Eritrea**,International Organization Migration(IOM),Vol.48 ,No.3 .

[23] Vogo, M. Z. (2014). **Human development strategies and their role in reducing the phenomenon of scientific competencies in Palestine (Study of the Gaza Strip)**, Journal of the Islamic University for Islamic and Administrative Studies, Volume 22, Issue 1, Islamic University, Gaza.2.

[24] Wold , H. (1996). **La regression PLS Paris : Technique iterative least squares** , In P. R. Kisnaiaah (Ed). Multivariate Analysis ,New York : Academic Press . pp391-420.

[25] Yohai, V. J. (1987), **High breakdown point and high efficiency robust estimates for regression** , Ann. Statistics,15,642 – 656.

[26] Yuan L.(2012) . **Brain drain of China and India , A Master`s of education**, Department of International Enviroment and Development Studies, Norwegian University .