

# Factors that Contribute Positively to Improve the Performance of Innovation for Saudi Arabia

Dr. Ahmad Alkhaleefah

College of Engineering; Production, Manufacturing Management, Prince Sultan University, Saudi Arabia

**Abstract** - This paper is an attempt to examine the determinants of innovation proxies by the annual score of Global Innovation Index (GII) in Saudi Arabia over the period from 2011 to 2018 by using two statistical approaches: Box–Jenkins (1970) approach and multiple regression approach. Moreover, based on the estimation results from the two mentioned approaches, we presented the forecasted values for GI score over the period from 2019 to 2023. The results proposed that infrastructure, market sophistication, and creative outputs are important factors that contributed positively towards improve the annual score of GI in Saudi Arabia, and this score will increase in the coming years.

**Keywords:** Box–Jenkins approach; Innovation; GI; Saudi Arabia; Time series forecasting;

## I. INTRODUCTION

Innovation has always played an important role in the economic and social development of countries, it helps address social challenges, advances sustainable development, improve productivity, job growth and welfare of peoples, additionally, and it is the foundation of competitiveness between countries. To support these claims, innovation must be measured.

Over time, various definitions and characterizations of innovation have been proposed. In this paper, we will use GI as a measure of innovation performance. The GI is an annual ranking of countries by their capacity for innovation. It is published by Cornell University, INSEAD, and the World Intellectual Property Organization, in partnership with other organizations and institutions. The GI provides a key tool and a rich database of detailed metrics for economies. The index was started to be published in 2007.

The GI relies on two sub-indices - the Innovation Input Sub-Index and the Innovation Output Sub-Index - each based on key pillars. Five input pillars capture elements of the national economy that enable innovative activities: (1) Institutions, (2) Human capital and research, (3) Infrastructure, (4) Market sophistication, and (5) Business sophistication. Two output pillars capture actual evidence of innovation

outputs: (6) Knowledge and technology outputs and (7) Creative outputs.

This study aims to investigate the pillars that contribute positively on the innovation performance in Saudi Arabia. According to report of Global innovation index 2018, Saudi Arabia is ranked 61st in the GI 2018, dropping 6 positions from last year. It is ranked 42nd among the 47 high-income countries, and is ranked 8th among the 19 countries in Northern Africa and Western Asia. The important of this report comes from a strategic position of Saudi Arabia because it is the 2nd country in the World that producing oil. In addition, Saudi Arabia now trying to set up a vision until the year 2030 based on national transforming programs to diversify the income of the country and to encourage the knowledge and innovation based economy systems.

The contribution of this study is in two-folds. The paper discusses the innovation performance of Saudi Arabia for many years using two statistical approaches "Box–Jenkins approach and Multiple Regression approach". Moreover, this study deals with a group of comprehensive factors. Therefore, any conclusion from this study can make the reader to feel confident that the recommendation is based on so many factors that have been considered on this analysis.

This paper is organized as follows: Section 2 provides the literature review. Section 3 presents the empirical study. Finally, Section 4 proposes the concluding remarks.

## II. LITERATURE REVIEW

In the last years, the emphasis on encouraging innovation has now increased especially in developing countries. At the global level, the innovation is considered an important factor to enhance economic growth and to find solutions for social challenges. Accordingly, benchmarking innovation performance is becoming a greater priority. In addition, there is a growing body of empirical literature that examines the determinants of Innovation performance. The literature on this issue can be divided into two strands.

One strand of literature focuses mainly on identifying the factors that contributed positively towards improves the innovation performance for specific country or group of

countries. For instance, **Alkhaleefah (2018)** studied how to improve the international innovation score in Saudi Arabia taking advantage of the wealth of information produced by GII over the period from 2011 to 2017 by using multiple regression technique. This study proposed that the infrastructure, market sophistication, business sophistication, and creative output are the most significant factors. **Raghupathi and Raghupathi (2017)** investigated the impact of economic indicators such as (GDP, gross national income, labor cost, R&D expenditure, real minimum wage, tax revenue, and education enrollment) on innovation at a country level for the OECD-member countries during the period from 2000 to 2010 using Cognos technique. Innovation indicators included the ratio of patents owned by foreign residents and the number of patent applications in each industry in the technology sector. The study found that countries with low GDP rely on foreign collaboration for innovation, and education enrollment encourages innovation. **Nour (2016)** provided an overview of the characteristics, strengths, and weaknesses of national systems of innovation (NSI) in the Arab countries by using both the descriptive and comparative approaches. This study examined the causes of poor NSI in the Arab countries from a national perspective and investigated the hypothesis that NSI in the Arab countries were suffering from serious weaknesses comparing with the other World countries in terms of education, research and development (R&D), human resources (HR), and information and communication technology (ICT). **WUNSCH-VINCENT and GUADAGNO (2015)** compared the innovation performance in 11 developing countries during the period from 2011 to 2014. The degree of heterogeneity among these countries was significant, they range from relatively small European and Western Asian countries such as Georgia, the Republic of Moldova, and Jordan to important global players such as China and India. This study analyzed how specific developing nations improved their local efforts and their capacity to innovate, in addition, took a look at the policy strategies, strengths and weaknesses for each country. The study referred to the significant role of education and research systems for the innovation outperformers, and noted one commonality among these outperformer's countries is their relatively stronger performance in production of knowledge

and technologies. **Ghazal and Zulkhibri (2015)** employed a panel data and negative binomial method to study the main determinants of innovation outputs which proxies by number of patent applications, trademarks and industrial designs in developing countries during the period from 1996 to 2010. The results proposed that foreign direct investment (FDI) could lead to encourage innovation outputs. The study also found the significant impact of institutions and governance factors for increasing innovation activities. **Blind (2012)** analyzed the impact of economic, social, and institutional regulations (i.e., R&D, imports, exports, human development indicator, technical cooperation indicator, competition legislation, price controls, product and service legislation, environmental laws and compliance, intellectual property rights, and the legal and regulatory framework) on innovation (patent intensity) in 21 OECD countries using panel data for the period between 1998 and 2004.

The second strand of the literature deals with analyzing the determinants of Innovation performance at firms' level in specific country. In the current context of international competition, firms can't survive without innovating. **Khiari and Rejeb (2015)** studied the main determinants of the innovation decision for 620 Tunisian companies in 2008 covering 24 governorates by using multilevel analysis and ordinal logistic regression. The dependent variable was expressed the number of type of innovation made by each firm in the survey (one type, two types, three types or four types of innovation). The exogenous variables were the activities of R&D of the company, cooperation with research centers', the qualifications of the employees, the foreign capital share and the turnover. The results showed that innovation differs from one governorate to another and the expending on research and development is a key determinant of innovation. Furthermore, another study was conducted by **Lasagni (2012)** to investigate the role of external relationships as key drivers of small business innovation for 490 small and medium-sized enterprises (SMEs) in six European countries in 2008 and 2009 by using a logistic regression model. The results showed that innovation performance is higher in SMEs that are proactive in strengthening their relationships with innovative suppliers, users, and customers.

TABLE 1  
Literature Review about the Determinants of Innovation Performance

| Ser. | Author             | Period of Study | Area of Study | The Used Statistical technique | The Main Results of The Study   |
|------|--------------------|-----------------|---------------|--------------------------------|---|
| 1    | Alkhaleefah (2018) | 2011-2017       | Saudi Arabia  | Time Series                    | The infrastructure, market sophistication, business sophistication, and creative output have significant impact on innovation |

|   |                                    |               |                                    |   |  |
|---|------------------------------------|---------------|------------------------------------|---|--|
| 2 | Raghupathiand Raghupathi (2017)    | 2000-2010     | OECD Countries                     | Cognos  | Countries with low GDP rely on foreign collaboration for innovation, and education enrollment encourages innovation  |
| 3 | Nour (2016)                        | -             | Arab countries                     | Descriptive and Comparative Approaches              | NSI in the Arab countries are suffering from serious weaknesses in terms of education, R&D, HR and ICT.  |
| 4 | WUNSCH-VINCENT and GUADAGNO (2015) | 2011- 2014    | 11 Developing Countries            | Descriptive and Comparative Approaches              | The education and research systems play significant role for enhancing the innovation  |
| 5 | Ghazal and Zulkhibri (2015)        | 1996- 2010    | Developing Countries               | Panel Data and Negative Binomial Method             | FDI could lead to encourage innovation outputs. In addition, institutions and governance factors have positive significant impact on innovation activities |
| 6 | Blind (2012)                       | 1998- 2004    | 21 OECD Countries                  | Panel Data Technique                                | The regulatory and legal framework conditions in OECD countries have a significant impact on innovative performance.                                       |
| 7 | Khiari and Rejeb (2015)            | 2008          | 620 Tunisian Companies             | Multilevel Analysis and Ordinal Logistic Regression | The expending on research and development is a key determinant of innovation.  |
| 8 | Lasagni (2012)                     | 2008 and 2009 | 490 SMEs in Six European Countries | Logistic Regression Model                           | Innovation performance is higher in SMEs that are proactive in strengthening their relationships with innovative suppliers, users, and customers.          |

### III. EMPIRICAL STUDY

#### 3.1 Data Description and Normality Test

This paper aims to identify the factors that contributed positively towards improve the international score of Saudi Arabia. It is used annual data of Saudi Arabia over the period from 2011 to 2018. The dataset is collected from the last reports of GII.

**TABLE 2**  
Descriptive Statistics of Variables and Test of Normality

| Variable                       | Abbreviation | Mean   | Max.   | Min.   | SD    | Skewness | Kurtosis | Jarque-Bera | p-value |
|--------------------------------|--------------|--------|--------|--------|-------|----------|----------|-------------|---------|
| GII                            | y            | 38.434 | 41.600 | 34.270 | 2.684 | -0.212   | 1.663    | 0.655       | 0.721   |
| Institutions                   | X1           | 59.038 | 67.500 | 51.900 | 5.261 | 0.063    | 2.170    | 0.235       | 0.889   |
| Human Capital & Research       | X2           | 42.413 | 47.700 | 35.600 | 4.135 | -0.255   | 1.889    | 0.498       | 0.779   |
| Infrastructure                 | X3           | 45.288 | 53.300 | 27.800 | 8.286 | -1.217   | 3.480    | 2.053       | 0.358   |
| Market Sophistication          | X4           | 51.713 | 59.000 | 47.500 | 3.522 | 1.029    | 3.394    | 1.463       | 0.481   |
| Business Sophistication        | X5           | 37.338 | 47.500 | 31.300 | 5.101 | 0.913    | 3.045    | 1.112       | 0.573   |
| Knowledge & Technology Outputs | X6           | 21.675 | 25.700 | 15.300 | 3.634 | -0.523   | 2.137    | 0.614       | 0.736   |
| Creative Outputs               | X7           | 37.688 | 48.200 | 23.400 | 8.686 | -0.437   | 1.896    | 0.661       | 0.719   |

Table 2 presents summary of descriptive statistics for the dependent and independent variables over the period from 2011 to 2018. We can note that the data not have large variation and distributed normally because the probability value (p-value) of **Jarque-Bera (1980)** test greater than 0.05.

From table 2, we notice that:

- The GII score of Saudi Arabia is reached to (38.434) on average over the period from 2011 to 2018.
- Saudi Arabia has the highest score in **Institutions** pillar, which it scores on average (59.038) over the period from 2011 to 2018.
- Saudi Arabia has the lowest score in **Knowledge & Technology Outputs** pillar, which it scores on average (21.675) over the period from 2011 to 2018.

### 3.2 Correlation Matrix

**TABLE 3**  
Person's Correlation Matrix

|    | y             | X1            | X2            | X3     | X4     | X5     | X6    | X7 |
|----|---------------|---------------|---------------|--------|--------|--------|-------|----|
| y  | 1             |               |               |        |        |        |       |    |
| X1 | 0.370         | 1             |               |        |        |        |       |    |
| X2 | <b>-0.787</b> | -0.536        | 1             |        |        |        |       |    |
| X3 | -0.013        | <b>-0.783</b> | 0.354         | 1      |        |        |       |    |
| X4 | 0.389         | 0.078         | <b>-0.755</b> | -0.197 | 1      |        |       |    |
| X5 | 0.259         | <b>0.711</b>  | -0.186        | -0.597 | -0.134 | 1      |       |    |
| X6 | <b>0.532</b>  | -0.315        | -0.565        | 0.394  | 0.583  | -0.621 | 1     |    |
| X7 | <b>0.952</b>  | 0.576         | <b>-0.770</b> | -0.282 | 0.312  | 0.466  | 0.318 | 1  |

Table 3 presents Person's correlations between the dependent and independent variables. We find that Creative Outputs“X7”, Human Capital & Research“X2”, and Knowledge & Technology Outputs “X6” have high correlations with the GII“y”. Moreover, there are some independent variables are high correlated with each other. Although the correlation matrix is an important tool to understand the relationships between the variables but it is not sufficient to describe the effect of independent variables on the dependent variable, so we need to build a regression model.

### 3.3 Box-Jenkins Approach

Box–Jenkins (1970) approach for time series analysis is common and widely applied in many applications. This approach is based on ARIMA models. Figure 1 summaries the main stages of Box–Jenkins approach. These iterative stages of modeling involve; identification, estimation, diagnostic checking, and forecasting. For a lot of in-depth exposition of Box–Jenkins time series analysis, see for example: **Chatfield (2016)** and **Abonazel and Abd-Elftah (2019)**.

The four stages of modeling in the Box-Jenkins iterative approach:

1. **Model Identification:** In this stage, we will check the stationary of the variables, identify the seasonality in the dependent series, and use the plots of the autocorrelation and partial autocorrelation functions for the dependent time series to decide which autoregressive or moving average component should be used in the model.
2. **Parameter Estimation:** we will use computation algorithms to reach the coefficients that best fit the selected ARIMA model. The most common methods used in estimation are maximum likelihood estimation or non-linear least-squares estimation.
3. **Model Checking:** In this step, we will test whether the estimated model conforms to the specifications of a stationary univariate process. In particular, the residuals should be independent of each other and constant in mean and variance over time. We can plot autocorrelation and partial autocorrelation of the residuals to identify misspecification. If the estimation is inadequate, we must return to step one and try to build a better mode

4. **Forecasting:** when the selected ARIMA model conforms to the specifications of a stationary univariate process, then we can use this model for forecasting.

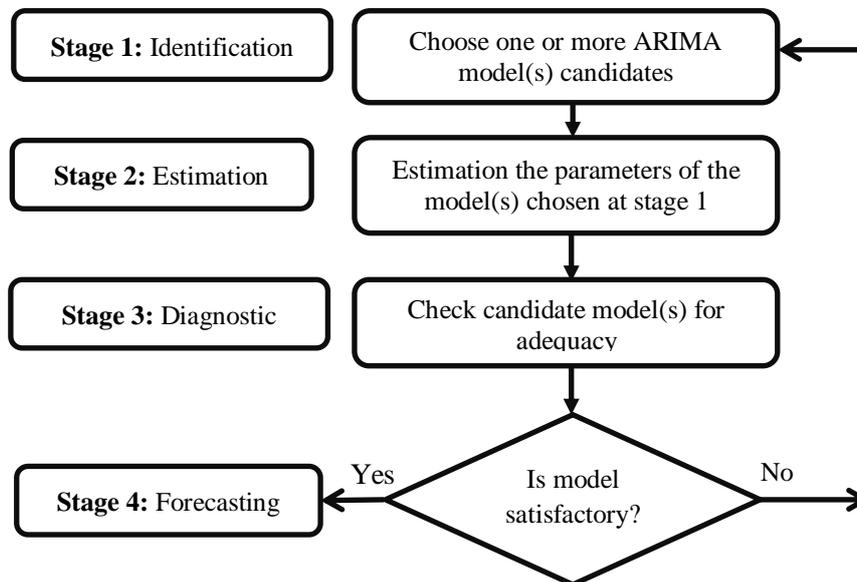


Figure 1: Stages in the Box-Jenkins Iterative Approach  
Source: Abonazel and Abd-Elftah (2019)

Applying this approach on  $y$ -data, the first stage is check the stationary of  $y$ . Figure 2 shows that  $y$  is stationary because the  $p$ -values of  $Q$ -test are greater than 0.05 and the bars of  $AC$  and  $PAC$  in the limits. So “ $d$ : the degree of differencing = 0” in ARIMA model. This means that the data is stationary.

| Autocorrelation | Partial Correlation | AC       | PAC    | Q-Stat | Prob  |
|-----------------|---------------------|----------|--------|--------|-------|
|                 |                     | 1 0.513  | 0.513  | 3.0029 | 0.083 |
|                 |                     | 2 -0.021 | -0.384 | 3.0086 | 0.222 |
|                 |                     | 3 -0.451 | -0.381 | 6.2616 | 0.100 |
|                 |                     | 4 -0.486 | -0.074 | 10.990 | 0.027 |
|                 |                     | 5 -0.241 | -0.010 | 12.542 | 0.028 |
|                 |                     | 6 0.019  | -0.119 | 12.556 | 0.051 |
|                 |                     | 7 0.168  | -0.066 | 14.812 | 0.038 |

Figure 2: ACF and PACF Plots for  $y$

As a next step, identify the value of the other two parameters “ $p$ : the order of autoregressive” and “ $q$ : the order of moving-average” of ARIMA model, so we check various ARIMA models to select the best model for the data.

TABLE 4  
Evaluation of Various ARIMA Models

| Model                | Goodness-of-fit measure |              |              |
|----------------------|-------------------------|--------------|--------------|
|                      | AIC                     | BIC          | MSE          |
| ARIMA (1,0,0)        | 4.884                   | 4.914        | 5.318        |
| ARIMA (2,0,0)        | 3.810                   | 3.850        | 1.098        |
| ARIMA (3,0,0)        | 3.111                   | 3.161        | 0.325        |
| <b>ARIMA (4,0,0)</b> | <b>1.075</b>            | <b>1.135</b> | <b>0.005</b> |
| ARIMA (4,0,1)        | 1.276                   | 1.346        | 0.006        |
| ARIMA (0,0,1)        | 4.794                   | 4.824        | 4.060        |

|               |       |       |       |
|---------------|-------|-------|-------|
| ARIMA (0,0,2) | 4.648 | 4.687 | 2.859 |
| ARIMA (0,0,3) | 4.564 | 4.614 | 2.241 |
| ARIMA (0,0,4) | 5.045 | 5.105 | 4.289 |
| ARIMA (1,0,4) | 4.978 | 5.078 | 5.968 |

From table 4, the best model is ARIMA (4,0,0), because it has the minimum values of goodness-of-fit measures: Mean Squared Error “MSE”, Akaike information criterion “AIC”, and Bayesian information criterion “BIC”. Modeling results of ARIMA (4,0,0) model have been estimated by maximum likelihood estimation method and are presented in the table 5.

**TABLE 5**  
The Results of ARIMA (4, 0, 0) Model

| Variable        | Coefficient | Std. Error     | t-Statistic | p-value |
|-----------------|-------------|----------------|-------------|---------|
| C               | 38.22800    | 0.093543       | 408.6677    | 0.0000  |
| AR(1)           | -0.442457   | 0.001525       | -290.1061   | 0.0000  |
| AR(2)           | 0.903050    | 0.062040       | 14.55593    | 0.0047  |
| AR(3)           | -0.451819   | 0.001403       | -322.1125   | 0.0000  |
| AR(4)           | -0.993697   | 0.000325       | -3053.784   | 0.0000  |
| Goodness-of-fit |             |                |             |         |
| R-squared       | 0.999       | Adj. R-squared | 0.999       |         |
| F-statistic     | 2188.263    | p-value of F   | 0.000       |         |

The coefficient estimates of all parameters and the model overall are statistically significant at 1% level of significance. The estimated regression equation is

$$\hat{y}_t = 38.23 - 0.44 y_{t-1} + 0.90 y_{t-2} - 0.45 y_{t-3} - 0.99 y_{t-4} \tag{1}$$

Where  $\hat{y}_t$  is the fitted value of GII in t, and  $y_{t-s}$ ;  $s = 1,2,3,4$  are the values of GII in different lags s.

According to Box-Jenkins approach, the diagnostic tests of the model are checking the normality and the stationary of the residuals. Figure 3 shows that the residuals are stationary because all p-values of Q-test are more than 0.05. While Figure 4 shows that the residuals are distributed normally because the p-value of Jarque-Bera is more than 0.05.

**TABLE 6**  
Actual GII, fitted GII, and residuals values of ARIMA (4, 0, 0) model

| year | GII     |         | Residuals of ARIMA (4, 0, 0) model |
|------|---------|---------|------------------------------------|
|      | Actual  | Fitted  |                                    |
| 2011 | 36.4000 | 36.4246 | -0.02456                           |
| 2012 | 39.3000 | 39.2548 | 0.04520                            |
| 2013 | 41.2000 | 41.1919 | 0.00810                            |
| 2014 | 41.6000 | 41.6473 | -0.04730                           |
| 2015 | 40.7000 | 40.7520 | -0.05203                           |
| 2016 | 37.8000 | 37.7713 | 0.02872                            |
| 2017 | 36.2000 | 36.1729 | 0.02709                            |
| 2018 | 34.2700 | 34.2711 | -0.00115                           |

|   | Autocorrelation | Partial Correlation | AC     | PAC    | Q-Stat | Prob  |
|---|-----------------|---------------------|--------|--------|--------|-------|
| 1 |                 |                     | -0.112 | -0.112 | 0.1441 | 0.704 |
| 2 |                 |                     | -0.101 | -0.115 | 0.2797 | 0.869 |
| 3 |                 |                     | -0.147 | -0.177 | 0.6248 | 0.891 |
| 4 |                 |                     | -0.259 | -0.333 | 1.9711 | 0.741 |
| 5 |                 |                     | 0.141  | 0.002  | 2.5044 | 0.776 |
| 6 |                 |                     | -0.106 | -0.228 | 2.9577 | 0.814 |
| 7 |                 |                     | 0.084  | -0.072 | 3.5286 | 0.832 |

Figure 3: ACF and PACF Plots of ARIMA (4,0,0) Residuals

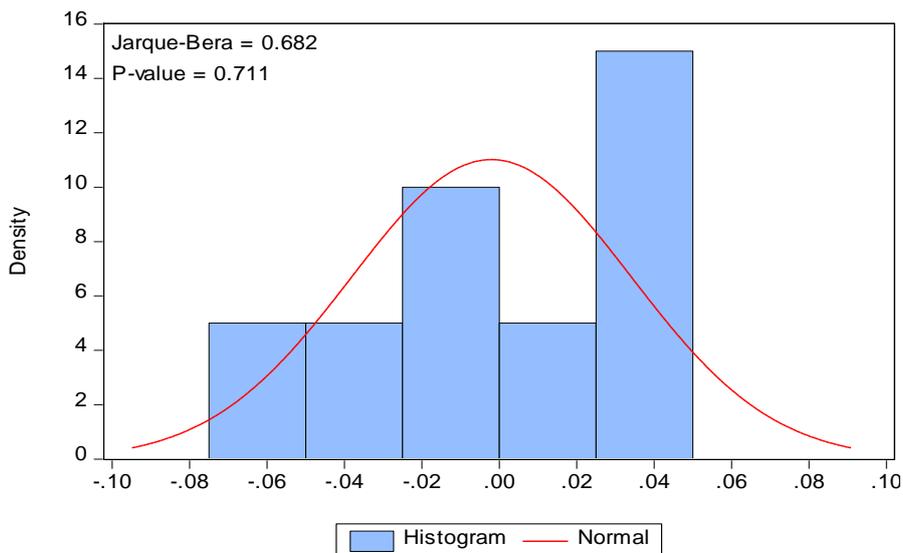


Figure 4: The Distribution of ARIMA (4,0,0) Residuals

### 3.4 Multiple Regression Approach

Since our data represent a time series from 2011 to 2018, so it is should make a unit root test of it before building the regression model to avoid the spurious regression problem. Three unit root tests are applied on the data. According to table 7, we can say that the data not have unit root process because p-values of Levin, Lin & Chu (0.0003), ADF (0.0271), and PP (0.005) tests are less than 0.05. This means that the multiple-regression model will be built without problems.

TABLE 7  
Unit Root Tests of the Variables

| Method  | Statistic | p-value | No. variables | Observations |
|---|-----------|---------|---------------|--------------|
| <b>Null: Unit root (assumes common unit root process)</b>     |           |         |               |              |
| Levin, Lin & Chu t*   | -3.46279  | 0.0003  | 8             | 53           |
| <b>Null: Unit root (assumes individual unit root process)</b> |           |         |               |              |
| ADF - Fisher Chi-square                                       | 28.5564   | 0.0271  | 8             | 53           |
| PP - Fisher Chi-square  | 34.0630   | 0.0053  | 8             | 56           |

Since the dataset is small (eight observations only), so it is should use a variable selection method to determinate the higher significant independent variables. Stepwise regression method has been used to make this. This method selected X3, X4, and X7 as shows in table 8. The results refer that the model is highly significant because the p-value of F-statistic (0.000) less than 0.05 and the selected independent variables explain 99% (R-squared value) of changes in y.

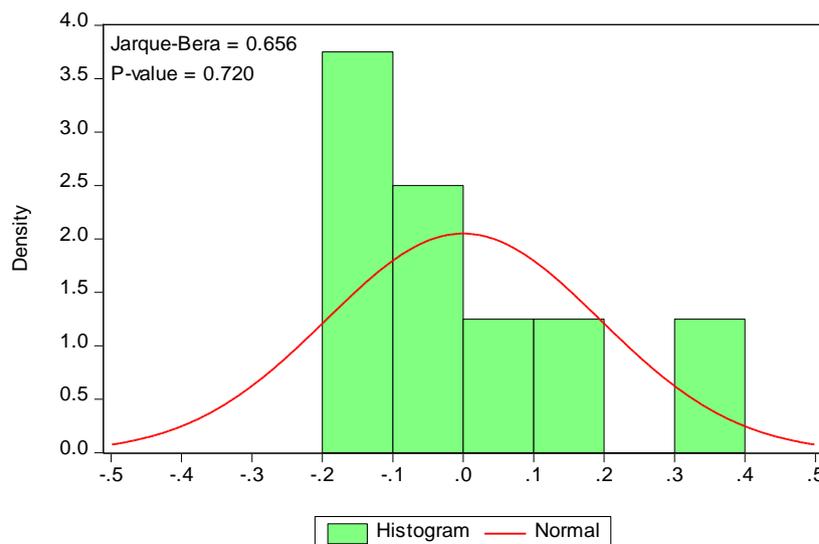
**TABLE 8**  
The Results of Regression Model

| Variable   | Coefficient | Std. Error | t-Statistic              | p-value      | VIF   |
|--|-------------|------------|--------------------------|--------------|-------|
| C  | 17.167      | 1.664      | 10.317                   | 0.001        | NA    |
| X3   | 0.095       | 0.012      | 7.715                    | 0.002        | 1.102 |
| X4   | 0.104       | 0.029      | 3.566                    | 0.024        | 1.123 |
| X7   | 0.307       | 0.012      | 25.277                   | 0.000        | 1.173 |
| Goodness-of-fit                                    |             |            |                          |              |       |
| <b>R-squared</b>                                   | 0.995       |            | <b>Adj. R-squared</b>    | 0.991        |       |
| <b>F-statistic</b>                                 | 252.351     |            | <b>p-value of F</b>      | 0.000        |       |
| Diagnostic tests                                   |             |            |                          |              |       |
| I. Serial Correlation LM Test: Breusch-Godfrey     |             |            |                          |              |       |
| F-statistic  | 0.330       |            | p-value of F(2,2)        | 0.752        |       |
| Obs*R-squared                                      | 1.985       |            | p-value of Chi-Square(2) | <b>0.371</b> |       |
| II. Heteroscedasticity Test: Breusch-Pagan-Godfrey |             |            |                          |              |       |
| F-statistic  | 0.394       |            | Prob. F(3,4)             | 0.765        |       |
| Obs*R-squared                                      | 1.825       |            | p-value of Chi-Square(3) | <b>0.610</b> |       |

The estimated regression equation is

$$\hat{y}_t = 17.167 + 0.095 X_{3t} + 0.104 X_{4t} + 0.307 X_{7t} \tag{2}$$

Table 8 shows that our model not has multi collinearity problem between the independent variables because all the values of centered VIF are less than 5. In addition, it shows that there is no serial correlation problem between the error terms in our model since the p-value (0.371) of Breusch-Godfrey test (developed by Breusch, 1978 and Godfrey, 1978b) is greater than 0.05. Moreover, our model does not have heteroscedasticity problem because the p-value (0.610) of Breusch-Pagan-Godfrey test (developed by Breusch and Pagan, 1979 and Godfrey, 1978a) is greater than 0.05. Based on Figure 5 and Jarque-Bera test that the test statistic equal 0.656 (with p-value 0.720 > 0.05), so the values of residuals are distributed normally.



**Figure 5: The Distribution of the Residuals**

According to the results of the diagnostic tests and the residual analysis mentioned in table 8 and figure 5, we can conclude that the estimated regression model is suitable (fit) to this data, and we can use this model for forecasting GII from 2019 to 2023.

### 3.5 Forecasting Results

According to regression approach, we must forecast (using any forecasting method) the independent variables (X3, X4, and X7) firstly, and then using equation (2) for forecasting the values of GII based on the forecasted values of the independent variables. In this paper, we used Box–Jenkins approach again to forecast the independent variables (X3, X4, and X7) in the first step. The forecasted values of the independent variables are given in table 9.

**TABLE 9**  
Forecasted Values of Independent Variables Using Box–Jenkins Approach

| Independent variables | Model           | R-squared | Forecasted values |       |       |       |       |
|-----------------------|-----------------|-----------|-------------------|-------|-------|-------|-------|
|                       |                 |           | 2019              | 2020  | 2021  | 2022  | 2023  |
| X3                    | ARIMA (4, 0, 1) | 0.976     | 49.8              | 42.31 | 41.68 | 33.79 | 33.6  |
| X4                    | ARIMA (4, 0, 0) | 0.999     | 58.03             | 48.99 | 45.87 | 51.25 | 51.12 |
| X7                    | ARIMA (4, 0, 0) | 0.995     | 19.31             | 18.63 | 18.48 | 20.8  | 25.44 |

However, according to Box-Jenkins approach, we can use the final model given in equation (1) directly to forecast GII values from 2019 to 2023. The forecasted values of GII with two approaches (multiple regression and Box–Jenkins) are given in table 10.

**TABLE 10**  
Forecasted Values of GII with Two Approaches

| years                         | Forecasted values   |                      |         |
|-------------------------------|---------------------|----------------------|---------|
|                               | Regression Approach | Box–Jenkins Approach | Average |
| 2019                          | 33.89               | 35.93                | 34.91   |
| 2020                          | 32.02               | 36.96                | 34.49   |
| 2021                          | 31.59               | 40.52                | 36.05   |
| 2022                          | 32.11               | 41.13                | 36.62   |
| 2023                          | 33.50               | 41.86                | 37.68   |
| Goodness-of-fit measures      |                     |                      |         |
| RMSE: Root Mean Squared Error | 0.1820              | 0.0612               | ----    |
| MAE: Mean Absolute Error      | 0.1522              | 0.0586               | ----    |
| MAPE: Mean Abs. Percent Error | 0.3959              | 0.1599               | ----    |
| Theil Inequality Coefficient  | 0.0024              | 0.0008               | ----    |

Table 10 shows that all the values of the goodness-of-fit measures (RMSE, MAE, MAPE, and Theil coefficient) for Box–Jenkins approach results are less than the values of the goodness-of-fit measures for regression approach results.

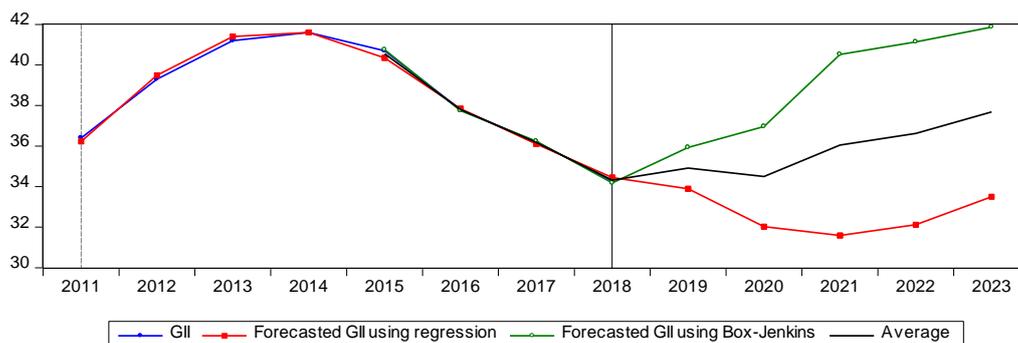


Figure 6: Time Series Plot for Actual and Forecasted Values with Two Approaches

Therefore, we can say that the forecasted values from ARIMA (4,0,0) that based on Box–Jenkins approach are more suitable for GII data than the forecasted values from regression approach. For further clarification figure 6 presents a comparison between the two results graphically.

#### IV. SUMMARY AND CONCLUSION

This paper investigated the factors that contributed positively towards improve the annual score of GII in Saudi Arabia over the period from 2011 to 2018 by using two statistical approaches: Box–Jenkins (1970) approach and multiple regression approach.

By using Box–Jenkins (1970) approach, we find that ARIMA (4,0,0) is the best model to describe GII data, because it has the minimum values of goodness-of-fit measures: MSE, AIC, and BIC. All coefficient estimates of our model ARIMA (4,0,0) and the model overall are statistically significant at 1% level of significance.

Based on multiple regression approach, we proposed using stepwise regression method as a variable selection method to determinate the higher significant independent variables among all independent variables mentioned in table 2, because the dataset is small (eight observations only). This method selected infrastructure, market sophistication, and creative outputs. The results referred that the model is highly significant and all selected independent variables explain 99% (R-squared value) of changes in GII. The results of the diagnostic tests and the residual analysis that mentioned in table 8 and figure 5 showed that the estimated regression model is suitable (fit) to this data, and we can use this model for forecasting GII from 2019 to 2023.

The results showed that the forecasted values from ARIMA (4,0,0) based on Box–Jenkins approach are more suitable for GII data than the forecasted values from regression approach.

#### REFERENCES

[1] Abonazel, M. R. and Abd-Elftah, A. I. (2019). Forecasting Egyptian GDP Using ARIMA Models. *Reports on Economics and Finance*, 5(1), 35-47.

[2] Alkhaleefah, A. (2018). What To Do To Improve The International Saudi Innovation Rank/Score. *Scopus*, 9, 435-442.

[3] Blind, K. (2012). The influence of regulations on innovation: A quantitative assessment for OECD countries. *Research Policy*, 41(2), 391-400.

[4] Breusch, T. S. (1978). Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers*, 17, 334–355.

[5] Breusch, T.S. and A.R. Pagan. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47, 1287- 1294.

[6] Chatfield, C. (2016). *The Analysis of Time Series: An Introduction*. CRC Press.

[7] Ghazal, R., and Zulkhibri, M. (2015). Determinants of innovation outputs in developing countries: Evidence from panel data negative binomial approach. *Journal of economic studies*, 42(2), 237-260.

[8] Godfrey, L. G. (1978a). Testing for multiplicative heteroscedasticity. *Journal of Econometrics*, 8, 227–236.

[9] Godfrey, L. G. (1978b). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*, 46, 1293–1301.

[10] Jarque, Carlos M.; Bera, Anil K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6 (3), 255–259.

[11] Khiari, M., & ben Rejeb, J. (2015). Determination of the Regional Impact on Innovation with an Ordinal Logit And a Multilevel Analysis. *Procedia-Social and Behavioral Sciences*, 195, 592-602.

[12] Lasagni, A. (2012). How can external relationships enhance innovation in SMEs? New Evidence for Europe. *Journal of small business management*, 50(2), 310-339.

[13] Nour, S. M. (2016). Overview of National Systems of Innovation in the Arab Countries. In *Economic Systems of Innovation in the Arab Region*, 91-165. Palgrave Macmillan, New York

[14] Raghupathi, V., and Raghupathi, W. (2017). Innovation at Country-Level: Association between Economic Development and Patents. *Journal of Innovation and Entrepreneurship*, 6(1), 1-20.

[15] WUNSCH-VINCENT, S. A. C. H. A., and GUADAGNO, F. (2015). Benchmarking Innovation Outperformance at the Global and Country Levels. *The Global Innovation Index 2015: Effective Innovation Policies for Development*, 65-79.

**Citation of this Article:**

Dr. Ahmad Alkhaleefah, "Factors that Contribute Positively to Improve the Performance of Innovation for Saudi Arabia"  
Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 4, Issue 6, pp  
26-36, June 2020. <https://doi.org/10.47001/IRJIET/2020.406003>

\*\*\*\*\*