

Research Article

Applications of Robust Regression Techniques: An Econometric Approach

Dost Muhammad Khan ¹, Anum Yaqoob,¹ Seema Zubair,² Muhammad Azam Khan,³ Zubair Ahmad ⁴, and Osama Abdulaziz Alamri⁵

¹Department of Statistics, Abdul Wali Khan University, Mardan, Pakistan

²Department of Statistics, Mathematics and Computer Science, The University of Agriculture, Peshawar, Pakistan

³Department of Economics, Abdul Wali Khan University, Mardan, Pakistan

⁴Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran

⁵Statistics Department, University of Tabuk, Tabuk, Saudi Arabia

Correspondence should be addressed to Zubair Ahmad; zubair@stu.yazd.ac.ir

Received 14 April 2021; Accepted 21 May 2021; Published 29 May 2021

Academic Editor: Amer Al-Omari

Copyright © 2021 Dost Muhammad Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Consistent estimation techniques need to be implemented to obtain robust empirical outcomes which help policymakers formulating public policies. Therefore, we implement the least squares (LS) and the high breakdown robust least trimmed squares (LTS) regression techniques, while using econometric regression model based on a growth equation for the two countries, namely, India and Pakistan. We used secondary annual time series data which covers a long period of 41 years. The adequacy of the time series econometric model was checked through cointegration analysis and found that there is no spurious regression. Classical and robust procedures were employed for the estimation of the parameters. The empirical results reveal that the overall fit of the model improves in case of LTS technique, while the significance of the predictors changes significantly in cases of both countries due to the removal of outliers from the data. Thus, empirical findings exhibit that the results, obtained through LTS, are better than LS techniques.

1. Introduction

The rise in the productive capacity of goods and services over certain period of time of a particular economy is termed the economic growth. Conventionally, it is measured in terms of gross domestic product (GDP). However, real GDP is the most commonly used measure of the economic growth of a country's economy in the literature. A rise in the GDP of a nation is referred to as economic growth which changes yearly with the recession and expansion of the economy. Economic growth is one of the key aims of development policy in almost every country. The economic growth accelerates society in a positive and productive direction and is influenced by a variety of factors, including human capital, physical capital, institutional factors, interest rate, exports, government expenditures, population growth, savings, and inflation [1, 2].

Several prior studies have been carried out to investigate the relationship of economic growth measured by GDP with FDI inflows, gross savings, and population growth. However, there are conflicting views regarding the impact of FDI and population growth on economic growth, e.g., analyzing the relationship between FDI and economic growth of India. For example, Ray [3] found a positive relationship between the two. Keshava [4] also found a positive relationship, but the relatively insignificant impact of FDI on GDP of India. Saqib et al. (2013) investigated the impact of FDI along with four other variables, debt, trade, inflation, and domestic investment, on the GDP of Pakistan for the period of 1981–2010. Their findings indicate that FDI, debt, trade, and inflation hurt the GDP while domestic investment has a positive impact. Zeb et al. [5] studied the impact of FDI along with three other variables: trade openness, political instability, and terrorist attacks on the economic growth of

Pakistan over the period of 1972–2012. They applied the least square method to investigate the impact of these variables on the GDP of Pakistan. The results unveil the positive and significant effect of FDI on the economic growth of Pakistan. Likewise, Mehmood [6] explored the impact of thirteen selected factors, including FDI inflow and gross saving on the GDP of Pakistan and Bangladesh. His findings are gross saving that has a positive and significant impact on the GDP of Pakistan, whereas FDI is the insignificant indicator of GDP. A review of related literature shows that the mixed impact of FDI on the economic growth depends on several factors, such as domestic infrastructure, highly educated labor force, trade regime, and FDI policies of the host country [7]. Azam and Ahmed [8] found that FDI inflow contributes to the economic growth in ten Commonwealth of Independent States during 1993–2011.

Both Pakistan and India are facing the same problem of rapid population growth. According to the World Population Data Sheet (2016), Pakistan is the sixth most populous country of the world, whereas India is the second-most populous country. To assess the impact of population growth on the economic growth of a nation, some researchers find a negative impact of the population growth that gives rise to the unemployment problem, lack of health and educational facilities, and reduction in the household savings which in turn lowers the national savings, while others argue that the high population growth is the real power of a nation as it gives rise to high labor force which helps the country by giving high output. For example, Afzal [9] discovered the negative impact of population growth on economic growth, whereas Ali et al. [10] noted the positive impact of population growth on the economic growth of Pakistan, which contradicts the findings made by Afzal [9]. Kothare [11] argues that India is the fastest-growing economy in the world because of its rapid population growth which has a positive impact on the GDP of the country. Similarly, Koduru and Tatavarthi [12] highlighted the positive impact of population growth on the economic growth of India. Using the pooled mean group approach, Olayungbo and Quadri [13] found that output per labor, population, trade openness, remittances, and FDI have a positive impact on economic growth in 20 sub-Saharan African countries during 2000–2015, while the inflation rate has a negative impact on growth. Azam et al. [14] observed that population growth, life expectancy, and investment have significant and positive impacts on economic growth in India from 1980 to 2018, whereas the inflation rate variable has a negative link with economic growth. Azam and Feng [15] found that official development assistance and inflation rate had negative, while FDI inflows, trade, and human capital by gross secondary school enrollment (%) had positive impacts on economic growth for 37 developing countries over 1985 to 2018.

In their study, Zaman et al. [16] claim that “a literature search shows that robust regression techniques are rarely used in applied econometrics.” Therefore, the central purpose of this study is to implement robust regression techniques, while using data from two developing countries, namely, India and Pakistan, over the period ranging from

1975 to 2015. The robust regression technique encompasses the least squares (LS) method and the robust LTS technique [17]. The simultaneous use of both techniques is expected to highlight the key differences in the development paths of both countries comparing the output of both countries. The chosen set of variables is FDI, gross savings and population growth, and real GDP per capita. To the best of the authors’ knowledge, this is a pioneer study on the application of robust regression methods using data from the two developing Asian countries. The main contribution of this article is to capitalize on the weaknesses of the classical estimation methods in the presence of outlying observations in the data that cause ill-estimation by providing alternative methods that are robust and less sensitive towards outliers and encompass the weaknesses of the classical estimation methods.

The rest of the paper is organized as follows: Section 2 deals with empirical methodology covering regression analysis, outliers’ management, robust regression, robust regression through least trimmed squares (LTS), and data description. Section 3 presents spurious regression and cointegration analysis, Section 4 includes model specification, and Section 5 presents empirical findings. Finally, Section 6 discusses the conclusion of the study.

2. Empirical Methodology and Data

This section includes the description of the data collected, the variables under consideration, the appropriate models used in this study, and statistical analysis for its estimation that have been carried out to meet the desired objectives.

2.1. Data Sources. As mentioned earlier, the time series data on the selected variables are taken from World Development Indicators [18] published by the World Bank for the period of 41 years from 1975 to 2015. The variables that are considered here for this study are FDI inflow, annual population growth, gross savings, and GDP per capita of Pakistan and India. To make the interpretation simpler, the data are transformed by log transformation.

2.2. Regression analysis. Regression analysis is one of the important and commonly used statistical tools for investigating the relationship between a dependent and one or more independent variables, with wide applications in the field of finance, economics, medicine, and psychology. A regression model is generally defined as

$$Y = X\beta + \varepsilon, \quad (1)$$

where the dependent variable Y and the vector of true residuals ε are $n \times 1$ and the design matrix X is $n \times p$. Write $\hat{\beta}$ for an estimate of β , and

$$e = Y - X\hat{\beta}, \quad (2)$$

for the corresponding fitted residuals.

The regression analysis commonly makes use of the least squares method for estimation of model parameters under some assumptions to be satisfied, such as the normality of

errors with zero mean and constant variance, i.e., $\varepsilon \sim N(0, \delta^2)$. The least square principle is to estimate the parameters by minimizing the sum of squared residuals (difference between the actual and the fitted values of the dependent variable). Thus, the least squares (LS) method uses the function $\rho(e) = e^2$, which is extremely sensitive to outliers, particularly those occurring on high leverage cases. The least square method gives misleading results when the assumption of normality is dissatisfied or outliers happen in the data as outliers drag the least square fit towards itself. Because of the extreme sensitivity of least square, a single outlier in a large sample is sufficient to deviate the regression fit totally as its breakdown point is $1/n$ which tends to zero with the increase in sample size n [19].

2.2.1. Outliers Management. Outliers being inconsistent observations and largely deviated from the majority of the observations in data need proper handling as they pose serious threat to the regression model and its estimated coefficients and, as a result, give misleading outcomes. Two types of outliers can happen in the regression dataset. One with extremely large values in the response is referred to as vertical outliers, whereas observations with extremely large values in the explanatory variable are called leverage points. Outliers being influential to the classical regression require methods that are insensitive to it. Two approaches are commonly used to cope with this problem, popularly known as diagnostic approach and robust procedures [19].

The diagnostic approaches try to identify the unusual observations through diagnostic statistics and remove it from the data and then classical procedures, for example, least square estimation procedure is then applied to the remaining clean dataset. This approach is suitable for simple data or when there are one or two outliers, but it becomes inefficient to detect outliers in case of multiple outliers present in a multidimensional dataset. Therefore, an appropriate procedure to deal with outliers is robust procedures that not only detect multiple outliers in complex data but also give efficient results.

2.2.2. Robust Regression. A robust regression is an iterative procedure that is designed to overcome the problem of outliers and influential observations in the data and minimize their impact over the regression coefficients [16]. Most of the regression techniques, named as robust, do not have this property. The main objective of robust estimation is to obtain reliable estimates/inferences for unknown parameters in the presence of outliers. The robust procedure replaces the sum of squared residuals of the OLS with some other function that is being less influenced by the unusual observations. These procedures first fit a regression to the data and then identify the outliers as those observations having large residuals. Robust techniques have three desiring properties, namely, efficiency, breakdown point, and bounded influence. The breakdown point is the smallest fraction of the unusual observations that an estimator can tolerate before giving an incorrect result. It is always a value between zero and 0.5. It measures the degree of robustness, the robustness of an estimator increases as the

breakdown point increases. For example, OLS has a breakdown point of 0% which represents that even a single outlier is sufficient to distort the OLS estimators. The robust techniques have 50% of breakdown point which is considered as the highest breakdown point. The property of bounded influence measures the resistance of the estimator against bad observations. It encounters the tendency of the least squares to allow the leverage points to exhibit greater influence.

There are various robust regression techniques. The first step in this respect came from Edgeworth [20] who proposed the least absolute deviation (LAD) by minimizing the sum of absolute residuals instead of minimizing the sum of squared residuals. LAD is preferable over OLS in providing protection against vertical outliers but is worse in the case of high leverage points with a breakdown point of $1/n$. Another popular approach is based on Huber's M-estimator [21] which minimizes a symmetrical objective function of residuals instead of squared residuals. M-estimator is robust against outliers in location and is more efficient than LAD. There are some other robust procedures, including least median square by Rousseeuw [22], which minimize the median of squared residuals, having 50% breakdown point, but low efficiency. To overcome the issue of low efficiency and to maintain high robustness, Rousseeuw [22] introduced least trimmed square estimators (LTS) which minimize sum of the smallest $h = [(n + p + 1)/2]$ squared residuals. Other robust estimators include S-estimator and MM-estimators. [23].

2.2.3. Robust Regression through Least Trimmed Squares (LTS). Least trimmed squares (LTS) is a highly robust and comparatively efficient estimator among all the robust estimators available in the literature and is obtained by minimizing the trimmed sum of the squared residuals [19]. LTS is a modified form of the LS estimator which corresponds to the more central values by ignoring the extreme observations in the ordered data. Consider the model in equation (1):

$$Y = X\beta + \varepsilon. \quad (3)$$

High breakdown admits the possibility that a large fraction of the data may have been replaced by arbitrary values. The high breakdown approach is exemplified by the least trimmed squares (LTS) criterion. Here, we write

$$e_1 = y_1 - x_1\hat{\beta}. \quad (4)$$

Defining e_i as the corresponding residual, write $e_{(i)}$ for the i^{th} order statistic of the e_i . That is,

$$|e_{(1)}| \leq |e_{(2)}| \leq \dots \leq |e_{(n)}|. \quad (5)$$

Then, LTS is defined by the criterion

$$\hat{\beta} = \min \sum_{i=1}^h e_{(i)}^2, \quad (6)$$

where h is a coverage parameter, commonly chosen to be $h = [(n + p + 1)/2]$ that determine the robustness of LTS. With the choice $h = n$, LTS specializes in OLS.

LTS is more efficient than least median squares (LMS), but its computational procedure is more complicated as compared to LMS. It has a greater convergence rate [24, 25]. The LTS procedure first identifies the outliers as the points with extreme positive or negative residuals. It then proceeds with the OLS for improved accuracy to classify a data point as outlier, standardized residuals are calculated, and data point with standardized LTS residual in its absolute value greater than 2.5 is considered as outlier. It helps in detecting the outliers efficiently. This goal is hard to achieve, otherwise, in high dimension data. It may be possible for LTS to detect too many data points as outliers, but removing a large proportion of data points as outlier may result in a regression that does not completely reflect the desired relationship [16]. In their study, Roozbeh and Arashi [26] noted that the LTS estimator is a highly robust regression estimator, while it is well known that the method of least squares is very sensitive to outliers.

3. Spurious Regression and Cointegration Analysis

The concept of cointegration introduced by Granger [27] has turned out to be extremely important in the analysis of nonstationary economic time series. To illustrate this problem, consider a simple regression model

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad (7)$$

where Y_t is a dependent variable, X_t is a single independent regressor, and ε_t is a white noise term with mean zero sequences.

If both $X_t \sim I(1)$ and $Y_t \sim I(1)$, by differencing one time, it is called integrated Y_t and X_t of order 1, denoted by $I(1)$. Then, generally $Y_t - \beta X_t \sim I(1)$ as well. However, there is one important exception. If $\varepsilon_t \sim I(0)$, then $Y_t - \beta X_t \sim I(0)$, that is, the linear combination of Y_t and X_t has the same statistical properties as $I(0)$ variable. In this case, variables Y_t and X_t are called cointegrated.

From regression models, the observed value of the t-statistic of the coefficient estimates is calculated under the assumption that the true value of the coefficient is equal to zero; despite this fact, researchers have found that the null hypothesis (H_0) of zero was rejected much more frequently than standard theory predictions. These results indicated that many of the significant relationships between nonstationary economic variables and existing econometric models will be spurious.

Researchers dealing with time series variables often suggest a simple solution to the problem of spurious regression. If the relationships between economic variables are specified in the first differences, the analytical complications due to nonstationary variables can be avoided because the differenced variables become stationary even if the original variables are not.

From Figures 1 and 2, it can be seen that there is nonstationary stochastic trend in all the variables of the India and Pakistan GDP datasets. Thus, augmented Dickey-Fuller (ADF) unit root test for testing the null hypothesis of nonstationarity was applied.

Since the variables at first difference become stationary as given in Table 1, so the variables are cointegrated and there is no spurious or nonsense regression. Therefore, all these variables can be used in the multiple regression model defined in (7).

4. Model Specification

The multivariate regression model within the framework on economic growth equation which is also used by many prior studies including the studies by Hasan [28], Adenola and Saibu [29], Peter and Bakari [30], and Azam [31, 32] and to be used for the two datasets is expressed as follows:

$$\text{Log GDP}_t = \beta_0 + \beta_1 \text{Log FDI}_t + \beta_2 \text{Log PG}_t + \beta_3 \text{Log GS}_t + \varepsilon_t, \quad (8)$$

where LogGDP is the gross domestic product per capita, LogFDI is the foreign direct investment, net inflows (% GDP), LogPG is the population growth (annual %), and LogGS is the gross savings (% GDP). The GDP is specified as the dependent variable and the remaining three variables are the explanatory variables and ε is the error term. The GDP per capita (constant 2010 US \$) has been used as a proxy for economic growth. The data are in log form to avoid any nonlinearity problem in the data.

5. Empirical Results and Discussion

To apply the robust regression techniques on the two developing Asian countries, namely, Pakistan and India, the classical least squares and robust LTS regressions are being used. Since the data used in the model are time series and the error terms of time series data often suffer from autocorrelation, therefore, Newey-West HAC (heteroscedasticity and autocorrelation consistent) estimation is considered for correcting the OLS standard errors in case of unknown autocorrelation and heteroscedasticity of the errors. The Newey-West estimation procedure gives the same estimates of the regression parameters as the OLS, but different values of the standard errors result in different t-statistic and p value for testing the null hypothesis. Moreover, HAC estimation is valid in the case of large samples and gives better results than OLS, since the sample size of 41 observations used in this analysis is reasonably large; therefore, HAC estimation procedure is implemented here. The regression estimates of both LS and LTS are presented in Tables 2 and 3. It is evident from Table 2 that the LS results suggest that population growth affects economic growth negatively and significantly, i.e., a rise in the population growth is associated with the decrease in economic growth. Results also reveal that the FDI inflow and gross savings have a positive, but insignificant impact on the GDP of Pakistan. The overall regression model is highly significant. The predictors explain 79.55% of the variation in the GDP per capita. Figure 3 shows the standardized residuals versus fits in the case of Pakistan, whereas Figure 4 shows the standardized residuals versus robust distance for Pakistan data.

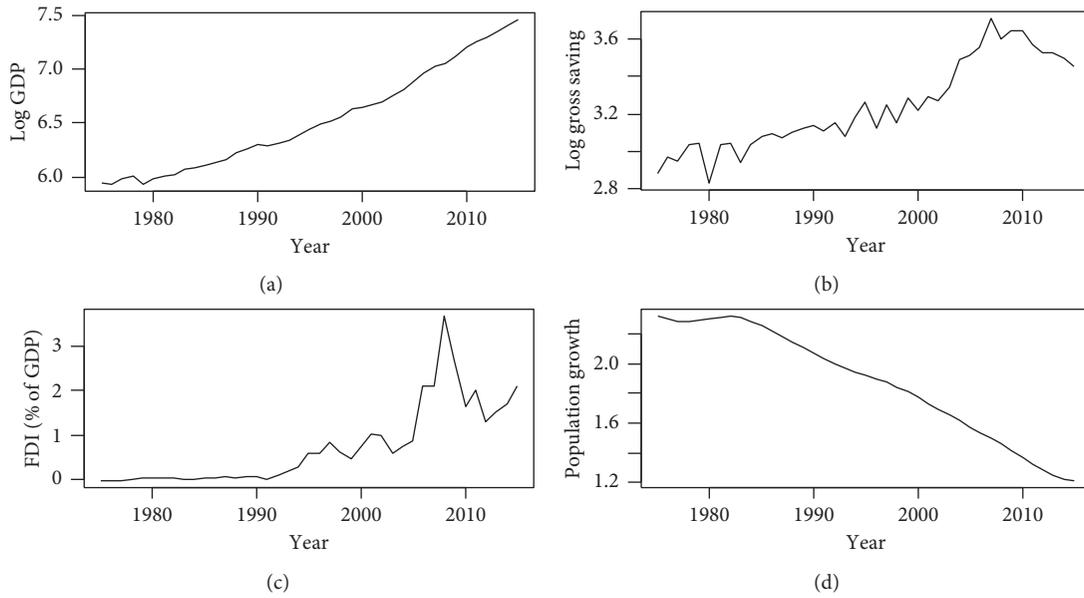


FIGURE 1: Time series plot of (a) log of GDP, (b) log of gross saving, (c) foreign direct investment, and (d) population growth for Indian data.

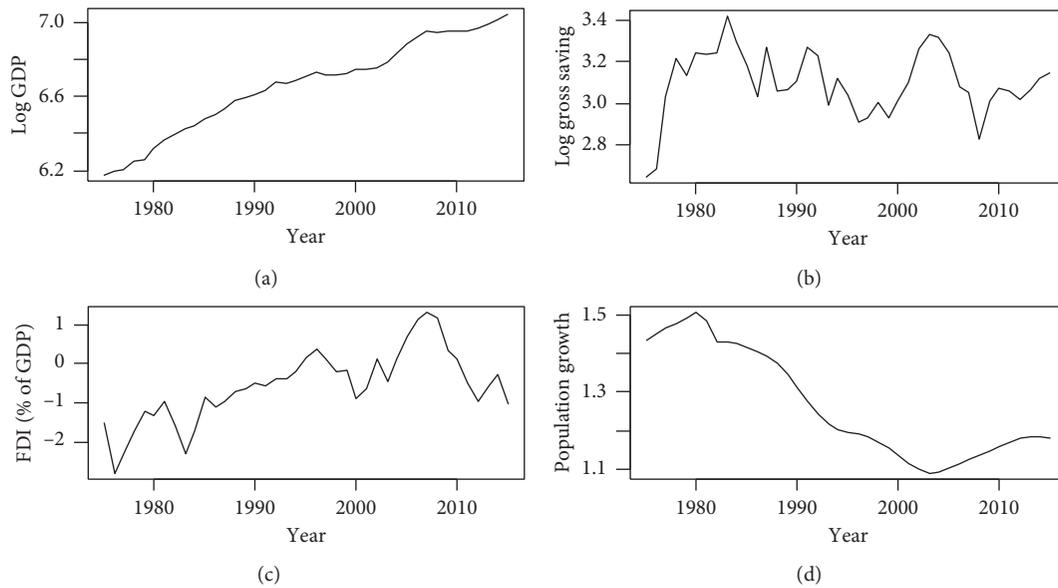


FIGURE 2: Time series plot of (a) log of GDP, (b) log of gross saving, (c) foreign direct investment, and (d) population growth for Pakistan data.

TABLE 1: Augmented Dickey–Fuller test and P values.

Variables	India		Pakistan	
	ADF	P value	ADF	P value
LogGDP_t	-6.093	0.001	-3.442	0.006
LogGS_t	-6.014	0.010	-4.942	0.01
LogFDI_t	-4.113	0.013	-5.199	0.01
LogPG_t	-5.083	0.010	-3.963	0.021

TABLE 2: Regression estimates (analysis of economic growth of Pakistan).

Var.	LS				Robust			
	Coeff.	S.E	<i>t</i> -val.	<i>P</i> val.	Coeff.	S.E	<i>t</i> -val.	<i>P</i> val.
Const.	3.5050*	0.2358	14.8610	0.0000	3.2338*	0.1029	31.428	0.0000
Log (FDI _{<i>t</i>})	0.0409	0.0445	0.9188	0.3641	0.1044*	0.0173	6.038	0.0000
Log (PG _{<i>t</i>})	-0.4112*	0.3854	3.6617	0.0008	-0.7934*	0.1155	6.866	0.0000
Log (GS _{<i>t</i>})	0.1286	0.1454	0.8850	0.3819	0.0770	0.0642	1.199	0.2400
	R^2 : 0.8108				R^2 : 0.9336			
	Adj. R^2 : 0.7955				Adj. R^2 : 0.927			
	<i>F</i> -statistic: 52.87				<i>F</i> -statistic: 140.70			
	Residual standard error: 0.0489				Residual standard error: 0.0249			

Note. * represents the level of significance at the 1% level.

TABLE 3: Regression estimates (analysis of economic growth of India).

Var.	LS				Robust			
	Coeff.	S.E	<i>t</i> -val.	<i>P</i> val.	Coeff.	S.E	<i>t</i> -val.	<i>P</i> val.
Const.	3.1756	0.0830	38.2339	0.0000	3.0939	0.0937	33.025	0.0000
Log (FDI _{<i>t</i>})	0.0086	0.0117	0.7347	0.4672	0.0055	0.0059	-0.9380	0.3548
Log (PG _{<i>t</i>})	-2.0359	0.0576	-35.3675	0.0000	2.0528	0.0664	-30.895	0.0000
Log (GS _{<i>t</i>})	0.1455	0.0542	2.6856	0.0108	0.2017	0.0572	3.5250	0.0012
	R^2 : 0.9944				R^2 : 0.9958			
	Adj. R^2 : 0.9939				Adj. R^2 : 0.9954			
	<i>F</i> -statistic: 2192				<i>F</i> -statistic: 2735			
	Residual standard error: 0.0158				Residual standard error: 0.01367			

Note. * represents the level of significance at the 1% level.

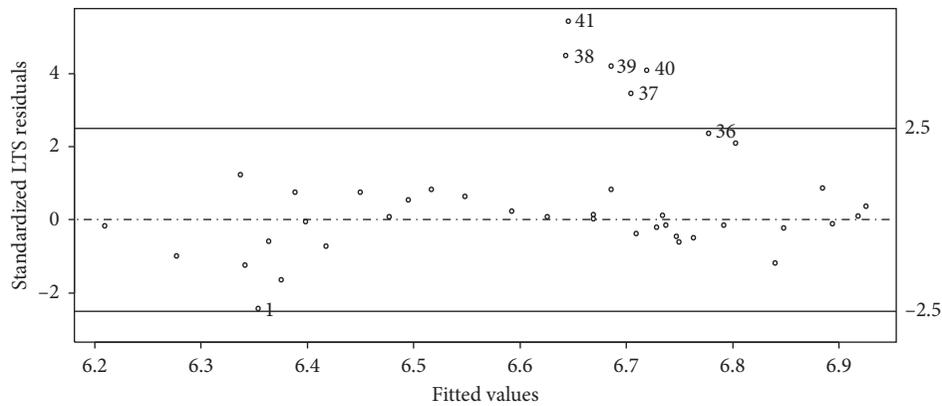


FIGURE 3: Standardized residuals vs fitted values.

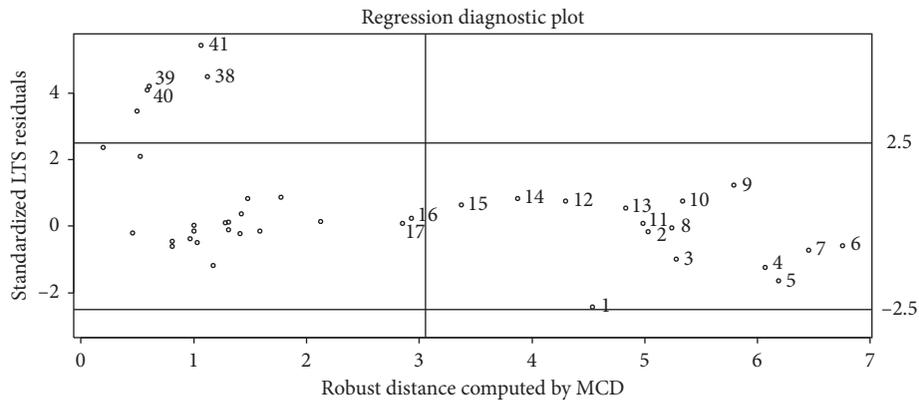


FIGURE 4: Standardized residuals vs robust distance.

Figure 5 shows the standardized residuals versus fitted values for India and Figure 6 shows the standardized residuals versus robust distances for India.

The robust analysis indicates that the population growth and foreign direct investment inflow both contribute significantly to the economic growth of Pakistan. However, FDI inflow is positively related, and population growth is negatively related to economic growth, whereas the impact of gross savings is insignificant.

Least trimmed squares reveals that there are 8 points with standardized residuals bigger than 2.5 standard deviations in absolute values. These values indicate the years 1975, 2009, 2010, 2011, 2012, 2013, 2014, and 2015 with standardized LTS residuals of -3.77 , 2.65 , 3.01 , 4.39 , 5.69 , 5.33 , 5.17 , and 6.89 , respectively, whereas the standardized residuals plot and regression diagnostic plot indicate that five points 2011, 2012, 2013, 2014, and 2015 are the influential observations, and removal of these 5 points has a large influence over the regression. The FDI inflow though insignificant with the classical method appears significant with the removal of outliers. The overall fit of the model improves as the R^2 and adjusted R^2 substantially increase to 93.36% and 92.7%, respectively.

The negative impact of the population growth reveals that rapid population growth is alarming and slows down the per capita GDP of Pakistan. It lowers the saving rate both at the domestic as well as national levels, whereas the significance of FDI reflects the truth that the economic development of Pakistan depends on the performance of FDI inflow up to some extent.

Likewise, Table 3 shows that the LS analysis of the economic growth of India suggests that population growth and gross savings are significant indicators for the economic growth of India. Population growth has a negative impact on economic growth. One percentage increase of the annual population growth reduces the GDP per capita of India by 2.03%, whereas the GDP per capita increases 0.14% with a one percent increase in the gross savings. The impact of FDI inflow on the GDP per capita is positive but insignificant. The R^2 and adj. R^2 are very high and the regression seems to be spurious, but as mentioned in Section 3, all the variables are cointegrated and a long-run equilibrium relationship exists among variables and thus the regression model is not spurious. Thus, the overall regression model is highly significant; 99.4% of the variation in GDP per capita is explained by its linear relationship with the predictor variables. The significance of the regression model is also indicated by the residual standard error that takes a minimum value.

Robust LTS regression technique and plots of standardized residuals and regression diagnostic reveal two outlying influential observations in the data with the standardized residuals of 3.4 and 3.9 in the absolute value corresponding to 1976 and 1979. Elimination of these two outliers improves the result. The gross savings is significant at 5% level of significance in the classical model that is significant at 1% with the robust regression. The impact of FDI inflow though insignificant and positive in the classical model has turned out to be negative with LTS procedure,

which indicates that the FDI inflow has an insignificant negative impact on the GDP per capita of India. Therefore, it results in a decrease of economic growth to some extent. Similarly, population growth lowers the GDP per capita. The improvement in the model fit by LTS is also indicated by the value of R^2 , F -value, and residuals standard error.

Empirical findings of the present study, regarding the relationship of FDI inflow and GDP per capita, do not verify the Ray [3] and Keshava's [4] findings, who explored the positive relationship between FDI and economic growth. Similarly, the results of this study regarding the impact of population growth on economic growth do not verify the findings of the studies conducted by Kothare [11] and Koduru and Tatavarthi [12] who argue that population growth positively affects economic growth. All these results suggest that the government of India should not rely on the FDI inflow and population growth for the improvement of the economic growth. The comparison of the outputs of both techniques represents that least square estimates are highly affected by outliers and give significantly different results from that of the LTS results which are in accordance with the findings of Zaman et al. [16], Al-Athari and Al-Amleh [33], and Onur and Cetin [34], who found that least squares method gives invalid estimates in the presence of a single outlying observation in the data, while the LTS give good estimates and are effected less as compared to LS estimates in the presence of outliers.

6. Summary and Conclusion

The current study has applied the least squares (LS) and the high breakdown robust least trimmed squares (LTS) regression techniques to estimate the impact of FDI inflow, annual population growth, and gross savings on the GDP per capita of Pakistan and India. Within the LS framework, FDI exerts an insignificant and positive impact on the economic growth of both Pakistan and India, whereas, after the application of LTS technique, the FDI enters positively and significantly in the economic growth model of Pakistan, but in the case of India, FDI has a negative insignificant impact on the economic growth due to the elimination of 5 and 2 outliers from the data of Pakistan and India, respectively. Population growth contributes to GDP per capita for both economies identically. Both techniques reveal that rapid population growth negatively influences the economic growth of both countries and hence is a serious problem for the economic growth of both economies, and it requires immediate attention. Gross savings have a positive and insignificant impact on the economic growth of Pakistan, whereas, for India, gross savings is the significant determinant of GDP per capita. Thus, sound economic policies that improve and encourage the FDI inflow and gross savings in Pakistan are required to be formulated and implemented.

The application of LTS technique reveals that the overall fit of the model improves, and the significance of the predictors changes significantly in both cases of Pakistan and India due to the removal of outliers from the data. Thus, empirical results suggest that, to avoid the impact of bad data

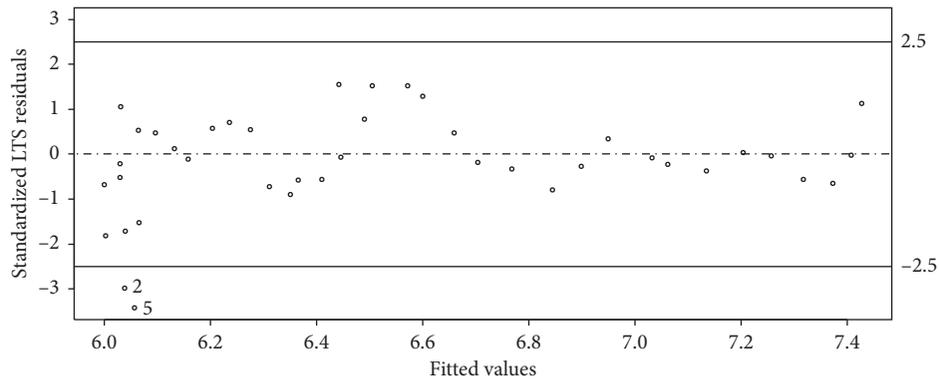


FIGURE 5: Standardized residuals vs fitted values.

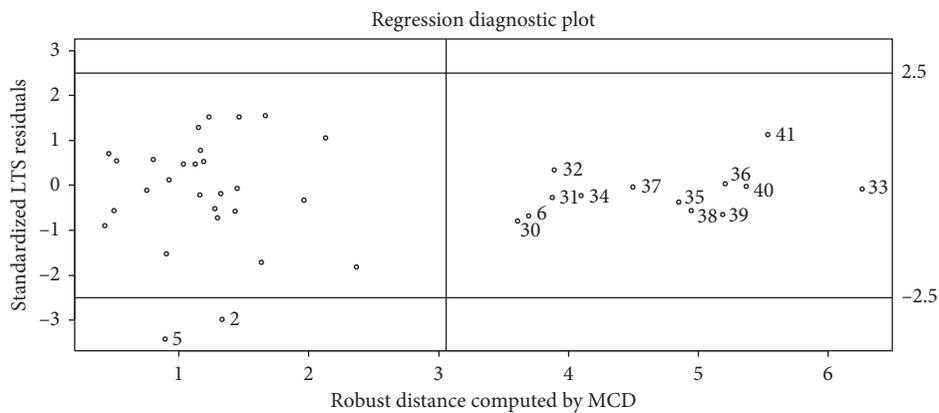


FIGURE 6: Standardized residuals vs robust distances.

points and to avoid misleading results, the robust technique is strongly recommended. Results obtained through the application of robust regression will largely help the policymakers.

Data Availability

The datasets are provided within the main body of the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] L. Alfaro, *Foreign Direct Investment and Growth: Does the Sector Matter*, Harvard Business School, Boston, MA, USA, 2003pp. 1–31, Retrieved from <https://www.grips.ac.jp/teacher/oono/hp/docu01/paper14.pdf>.
- [2] N. A. Burney and A. S. Aftab, "Sources of Pakistan's economic growth," *The Pakistan Development Review*, vol. 25, no. 4, pp. 573–589, 1986.
- [3] S. Ray, "Impact of foreign direct investment on economic growth in India: a Co-integration analysis," *Advances in Information Technology and Management*, vol. 2, no. 1, pp. 187–201, 2012.
- [4] D. S. Keshava, "The effect of FDI on India and Chinese economy: a comparative analysis," in *Proceedings of the Second Singapore International Conference on Finance* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1089964, December 2008.
- [5] A. A. Alamgir, S. A. Khan, D. M. Khan, and U. Khalil, "A new efficient redescending M-estimator: alamgir redescending M-estimator," *Research Journal of Recent Sciences*, vol. 2, no. 8, pp. 79–91, 2013.
- [6] S. Mehmood, "Effect of different factors on gross domestic product: a comparative study of Pakistan and Bangladesh," *Academy of Contemporary Research Journal*, vol. 1, no. 1, pp. 18–35, 2012.
- [7] N. Saqib, M. Masnoon, and N. Rafique, "Impact of foreign direct investment on economic growth of Pakistan," *Advances in Management and Applied Economics*, vol. 3, no. 1, pp. 35–45, 2013.
- [8] M. Azam and A. M. Ahmed, "Role of human capital and foreign direct investment in promoting economic growth," *International Journal of Social Economics*, vol. 42, no. 2, pp. 98–111, 2015.
- [9] M. Afzal, "Population growth and economic development in Pakistan," *The Open Demography Journal*, vol. 2, no. 1, pp. 1–7, 2009.
- [10] S. Ali, A. Ali, and A. Amin, "The impact of population growth on economic development in Pakistan," *Middle-East Journal of Scientific Research*, vol. 18, no. 4, pp. 483–491, 2013.
- [11] R. Kothare, "Does India's population growth has a positive effect on economic growth?" *Social Science*, vol. 410, pp. 2–14, 1999.
- [12] B. P. K. Koduru and A. Tatavarthi, *Effect of Population Growth Rate on Economic Development in India*, Retrieved from, 2016.

- [13] D. O. Olayungbo and A. Quadri, "Remittances, financial development and economic growth in sub-Saharan African countries: evidence from a PMG-ARDL approach," *Financial Innovation*, vol. 5, no. 9, pp. 1–25, 2019.
- [14] M. Azam, H. N. Khan, and F. Khan, "Testing Malthusian's and Kremer's population theories in developing economy," *International Journal of Social Economics*, vol. 47, no. 4, pp. 523–538, 2020.
- [15] M. Azam and Y. Feng, "Does foreign aid stimulate economic growth in developing countries? Further evidence in both aggregate and disaggregated samples," *Quality and Quantity*, 2021.
- [16] A. Zaman, P. J. Rousseeuw, and M. Orhan, "Econometric applications of high-breakdown robust regression techniques," *Economics Letters*, vol. 71, no. 1, pp. 1–8, 2001.
- [17] D. M. Hawkins and D. M. Khan, "A procedure for robust fitting in nonlinear regression," *Computational Statistics and Data Analysis*, vol. 53, no. 12, pp. 4500–4507, 2009.
- [18] World Development Indicators, *The World Bank*, Available at <https://databank.worldbank.org/source/world-development-indicators>, National University of Singapore, Singapore, 2020.
- [19] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, New York, NY, USA, 1987.
- [20] P. K. Koduru and A. Tatavarthi, "Effect of population growth rate on economic development in India," *International Journal of Social Sciences Management and Entrepreneurship (IJSSME)*, vol. 3, no. 2, 2019.
- [21] P. J. Huber, "Robust regression: asymptotics, conjectures and Monte Carlo," *The Annals of Statistics*, vol. 1, no. 5, pp. 799–821, 1973.
- [22] P. J. Rousseeuw, "Least median of squares regression," *Journal of the American Statistical Association*, vol. 79, no. 338, pp. 871–880, 1984.
- [23] V. J. Yohai, "High breakdown point and high efficiency robust estimates for regression," *Annals of Statistics*, vol. 15, no. 20, pp. 642–656, 1987.
- [24] A. A. Alamgir, S. A. Khan, D. M. Khan, and U. Khalil, "A new efficient redescending M-estimator: alamgir redescending M-estimator," *Research Journal of Recent Sciences*, 2013.
- [25] D. M. Khan, S. Ihtesham, A. Ali, U. Khalil, S. A. Khan, and S. Manzoor, "AN efficient and high breakdown estimation procedure for nonlinear regression models," *Pakistan Journal of Statistics*, vol. 33, no. 3, 2017.
- [26] M. Roozbeh and M. Arashi, "Least-trimmed squares: asymptotic normality of robust estimator in semiparametric regression models," *Journal of Statistical Computation and Simulation*, vol. 87, no. 6, pp. 1130–1147, 2017.
- [27] C. W. J. Granger, "Some properties of time series data and their use in econometric model specification," *Journal of Econometrics*, vol. 16, no. 1, pp. 121–130, 1981.
- [28] M. S. Hasan, "The long-run relationship between population and per capita income growth in China," *Journal of Policy Modeling*, vol. 32, no. 3, pp. 355–372, 2010.
- [29] F. Adenola and O. M. Saibu, "Does population change matter for long run economic growth in Nigeria?" *International Journal of Development and Sustainability*, vol. 6, no. 12, pp. 1955–1965, 2017.
- [30] A. Peter and I. H. Bakari, "Impact of population growth on economic growth in Africa: a dynamic panel data approach (1980-2015)," *Pakistan Journal of Humanities and Social Sciences*, vol. 6, no. 4, pp. 412–427, 2018.
- [31] M. Azam, "Energy and economic growth in developing Asian economies," *Journal of the Asia Pacific Economy*, vol. 25, no. 3, pp. 447–471, 2020.
- [32] M. Azam, "Governance and economic growth: evidence from 14 Latin America and Caribbean countries," *Journal of the Knowledge Economy*, 2021, Springer- Emerging Sources Citation Index.
- [33] F. M. Al-Athari and M. A. Al-Amleh, "A comparison between least trimmed of squares and MM-estimation in linear regression using simulation technique," in *Proceedings of the International Arab Conference on Mathematics and Computations*, Zarqa University, Zarqa, Jordan, May 2016.
- [34] T. Onur and M. Cetin, "The comparing of S-estimator and M-estimators in linear regression," *Gazi University Journal of Science*, vol. 24, no. 4, pp. 747–752, 2011.