NOVATEUR PUBLICATIONS
INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING RESEARCH AND TECHNOLOGY [IJIERT]
ISSN: 2394-3696 Website: ijiert.org
VOLUME 7, ISSUE 9, Sep.-2020

# HIV/AIDS ON THE INCREASE IN PAKISTAN: A STERN WARNING FROM THE BOX-JENKINS ARIMA APPROACH

DR. SMARTSON. P. NYONI
ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe

MR. THABANI NYONI
Department of Economics, University of Zimbabwe, Harare, Zimbabwe

## ABSTRACT
Pakistan is now a “highly concentrated” epidemic country. This simply implies higher incidence of HIV among high risk groups in the country, for example, injecting drug users and commercial sex workers. There is a silent and yet serious escalation of new HIV infections especially in adults in Pakistan. Using annual time series data on the number of adults (ages 15 and above) newly infected with HIV in Pakistan from 1990 – 2018, the study predicts the annual number of adults who will be newly infected with HIV over the period 2019 – 2030. The study employs the Box-Jenkins ARIMA technique. The diagnostic ADF tests show that the F series under consideration is an I (1) variable. Based on the AIC, the study presents the ARIMA (1, 1, 0) model as the optimal model. The diagnostic tests further reveal that the presented model is stable and acceptable. The results of the study indicate that the number of new HIV infections in adults in Pakistan is likely to continue on an upwards trajectory, over the period 2019 – 2030, from approximately 20718 in 2019 to about 28532 new infections by 2030.

## INTRODUCTION
HIV/AIDS is a growing public health threat endangering lives of millions of people around the world (Ministry of National Health Services, Regulations & Coordination, 2017). HIV/AIDS is becoming a prevalent disease in Pakistan, and its death toll has been steadily increasing each year since 1987 when the first case was reported in the country (Mujeeb & Hashmi, 1988; UNAIDS, 2014). The number of HIV/AIDS infections in Pakistan has been increasing at an alarming rate; especially from 2005 to 2015, the number of reported infections in Pakistan increased from 8360 to 45990 cases, the highest global average increase of 17.6% in history (Khan & Khan, 2012). The death toll has also been on the rise: the number of deaths caused by HIV/AIDS in Pakistan increased from 350 to a staggering number of 1480 between 2005 and 2015, which translates to an average increase of 14.2% per year (Ghani, 2016). As if this is not enough, the Antiretroviral Therapy (ART), a drug regimen that helps in the treatment of the HIV/AIDS (Bhatti et al., 2016), is neither easily available nor easily accessible in Pakistan (Hussain et al., 2018); only a very small percentage of individuals suffering from HIV/AIDS receives ART drugs for treatment (Yusufzai, 2011). The HIV epidemic in Pakistan is largely driven by people who inject drugs and is fast expanding in other key populations including men who have sex with men, “hijra” or transgender, male and female sex workers. The country has witnessed isolated episodes of HIV outbreaks from time to time including the most recent outbreak of HIV in chronic renal failure patients undergoing dialysis (Ministry of National Health Services, Regulations & Coordination, 2017).
The HIV epidemic in Pakistan is highly concentrated in adults (Siddiqui et al., 2020). For example, in 2018, an estimated 160000 people were living with HIV in Pakistan, with 2.2% being children aged below 15 years (UNAIDS, 2018). HIV/AIDS prevention is never easy (He et al., 2018) especially in a country like Pakistan where there is a weak domestic political commitment to HIV and a tepid response (Abdullah & Shaikh, 2015; Mansoor, 2017). There is high disease stigma (BBC News, 2011), no stated commitment by heads of state regarding HIV (Sheikh, 2017) and budgetary commitment to address the disease remains low (Khalid & Fox, 2019). The main goal of this research is to predict the number of adults newly infected with HIV in Pakistan over the period 2019 – 2030. This study will go a long way in stimulating national HIV response in the country.

## LITERATURE REVIEW

In a retrospective study carried out at a Sexually Transmitted Infections (STIs) clinic, District Headquarter (DHQ) hospital, Faisalabad, in Pakistan, Maan et al. (2014), tried to find out the prevalence of HIV and related risk factors. Between March 2010 and December 2012, a total of 31040 subjects were either interviewed or their medical records were reviewed. The study found out an overall HIV prevalence of 557 per 100000. Khan et al. (2017) explored the association between knowledge about HIV and discriminatory attitudes towards people living with HIV in Pakistan. The study was based on secondary data analysis of data from the Pakistan Demographic and Health Survey conducted in 2012-13. Ever-married women and men aged 15-49 who had heard about AIDS were included in this study. A chi-square test and multinomial regression analysis were performed. The study found a statistically significant inverse relationship between knowledge about HIV and discriminatory attitudes towards people living with HIV.

In a review article, Hussain et al. (2018) uncovered the causes and consequences of HIV/AIDS in Pakistan and also examined the role of the Pakistani government in controlling the menace. The study argued that Pakistan exhibits a low level of education regarding the topic of Sexually Transmitted Diseases (STDs) and HIV/AIDS. Consistently, Khalid & Fox (2019) explored how policy actors tasked with implementing HIV programs navigate the competing demands placed upon them by development targets and national policies, particularly in the current context of waning international investments towards HIV. The researchers interviewed 29 key informants comprising health experts in donor organizations and government employees in HIV programs in Pakistan. Themes were also identified inductively through an iterative process and findings were triangulated with various data sources and existing literature. The study found out that in Pakistan there is low and heterogeneous political commitment for HIV, which has led to low HIV treatment coverage in the country.

In a recent study, Siddiqui et al (2020) analyzed pediatric HIV in Pakistan. A matched case-control study was conducted with 406 cases recruited. Conditional logistic regression was used to investigate the association of a priori defined risk factors with HIV infection. Furthermore, global positioning system coordinates of participants' addresses were collected to investigate concordance between the genetic and spatial epidemiology. At the time of preparing this paper, the final results of the study by Siddiqui et al. (2020) had not been released but their study is expected to basically provide information on the likely routes of infection and drivers of the HIV outbreak among children in Pakistan. Studies that forecast new HIV infections in adults in Pakistan are scanty. It is this information gap that this paper seeks to fill.

## METHODODOLOGY

### The Box – Jenkins (1970) Methodology

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018c). This approach will be used to analyze the F series under consideration.

### The Moving Average (MA) model

Given:

$$F_t = \sum_{i=1}^{q} \alpha_i \mu_{t-i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

where $\mu_t$ is a purely random process with mean zero and varience $\sigma^2$. Equation [1] is reffered to as a Moving Average (MA) process of order q, commonly denoted as MA (q). F is the annual number of adults newly infected with HIV in Pakistan at time t, $\alpha_0 \ldots \alpha_q$ are estimation parameters, $\mu_t$ is the current error term while $\mu_{t-1} \ldots \mu_{t-q}$ are previous error terms.

**The Autoregressive (AR) model**

Given:

$$F_t = \sum_{i=1}^{p} \beta_i F_{t-i} + \mu_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [2]$$

Where $\beta_1 \ldots \beta_p$ are estimation parameters, $F_{t-1} \ldots F_{t-p}$ are previous period values of the F series and $\mu_t$ is as previously defined. Equation [2] is an Autoregressive (AR) process of order p, and is usually denoted as AR (p).

**The Autoregressive Moving Average (ARMA) model**

An ARMA (p, q) process is just a mere combination of AR (p) and MA (q) processes. Thus, by combining equations [1] and [2]; an ARMA (p, q) process may be specified as shown below:

$$F_t = \sum_{i=1}^{p} \beta_i F_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \ldots \ldots \ldots \ldots \ldots \ldots [3]$$

**The Autoregressive Integrated Moving Average (ARIMA) model**

A stochastic process $F_t$ is referred to as an Autoregressive Integrated Moving Average (ARIMA) [p, d, q] process if it is integrated of order "d" [I (d)] and the "d" times differenced process has an ARMA (p, q) representation. If the sequence $\Delta^d F_t$ satisfies an ARMA (p, q) process; then the sequence of $F_t$ also satisfies the ARIMA (p, d, q) process such that:

$$\Delta^d F_t = \sum_{i=1}^{p} \beta_i \Delta^d F_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \ldots \ldots \ldots \ldots \ldots \ldots [4]$$

where $\Delta$ is the difference operator, vector $\beta \in \mathbb{R}^p$ and $\alpha \in \mathbb{R}^q$.

**Data Collection**

This study is based on annual observations (that is, from 1990 – 2018) on the number of new HIV infections in adults (ages 15 years and above) [denoted as F] in Pakistan. Out-of-sample forecasts will cover the period 2019 – 2030. All the data was gathered from the World Bank online database.

**Diagnostic Tests & Model Evaluation**

**Stationarity Tests: Graphical Analysis**

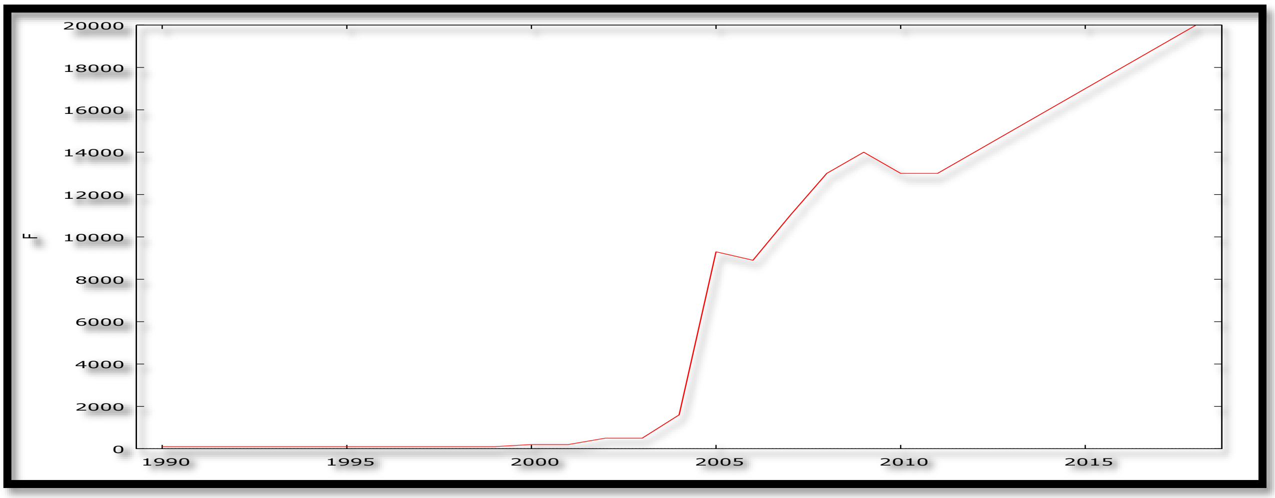


Figure 1

### 3.7.2 The Correlogram in Levels

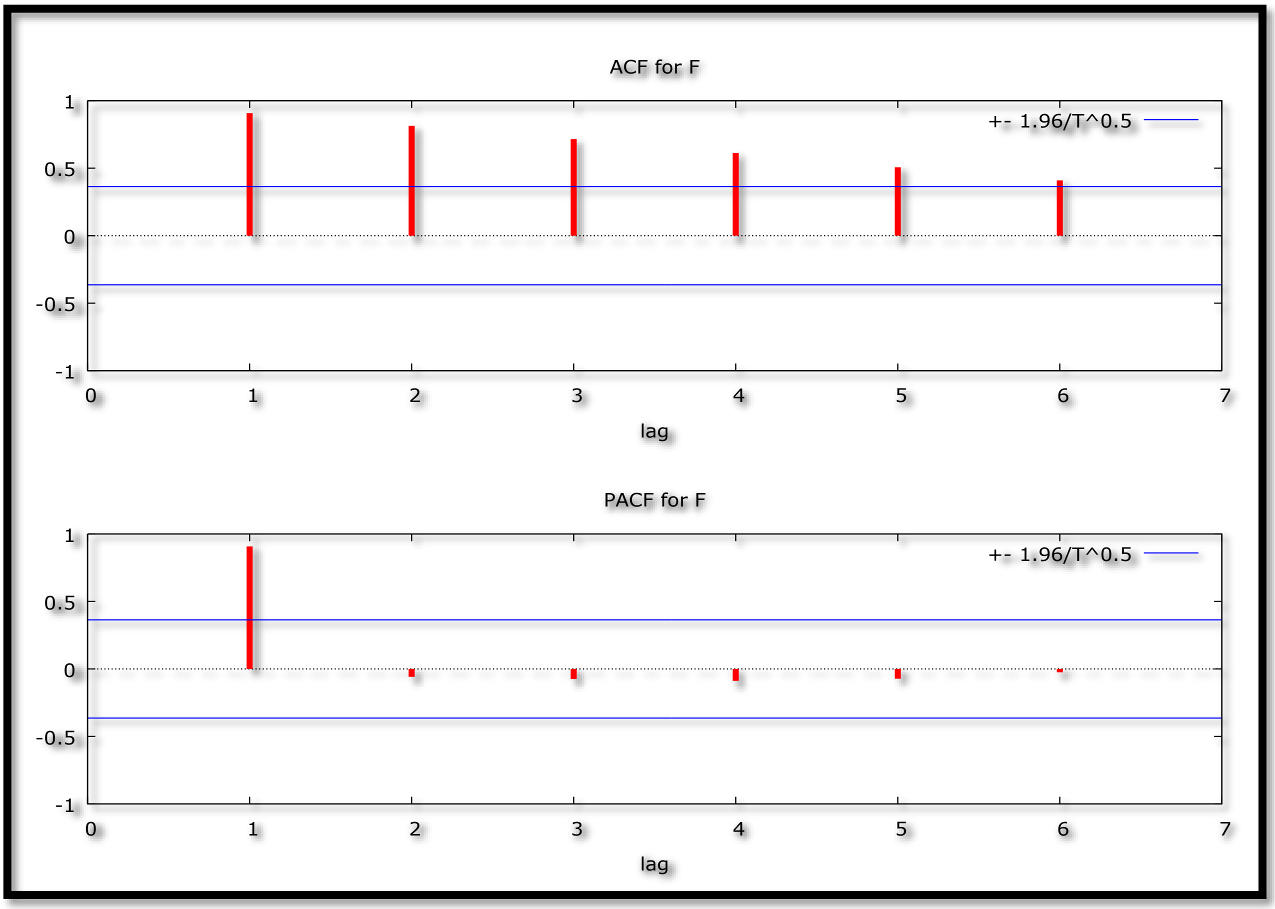


Figure 2: Correlogram in Levels

### The ADF Test in Levels

Table 1: with intercept

| Variable | ADF Statistic | Probability | Critical Values | | Conclusion |
|---|---|---|---|---|---|
| F | 0.497776 | 0.9835 | -3.689194 | @1% | Non-stationary |
| | | | -2.971853 | @5% | Non-stationary |
| | | | -2.625121 | @10% | Non-stationary |

Table 1 and figure 1 consistently show that F is not stationary in levels. This is attributed to the fact that the series under consideration has a clear upward trend; hence its mean cannot be constant over time.

### The Correlogram (at First Differences)

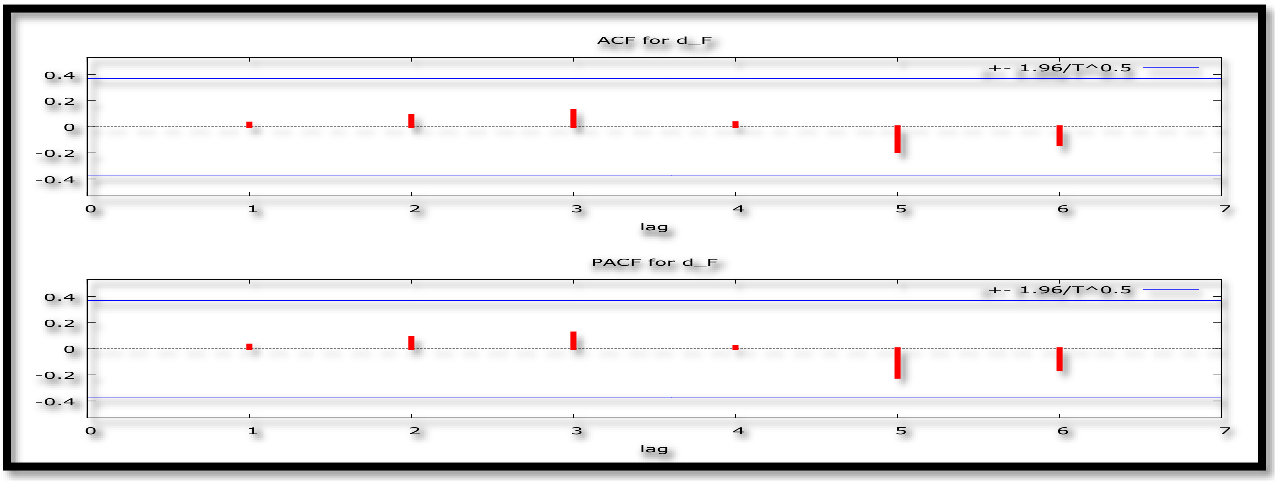


Figure 3: Correlogram (at First Differences)

**The ADF Test (at First Differences)**

Table 2: with intercept

| Variable | ADF Statistic | Probability | Critical Values | | Conclusion |
|---|---|---|---|---|---|
| ΔF | -4.876355 | 0.0006 | -3.699871 | @1% | Stationary |
| | | | -2.976263 | @5% | Stationary |
| | | | -1.627420 | @10% | Stationary |

Figure 3 and table 2 consistently indicate that F is an I (1) variable.

**Evaluation of ARIMA models (with a constant)**

Table 3: Evaluation of ARIMA Models (with a constant)

| Model | Akaike Information Criteria (AIC) |
|---|---|
| ARIMA (1, 1, 1) | 497.2953 |
| ARIMA (1, 1, 0) | **495.4317** |
| ARIMA (2, 1, 0) | 497.2255 |
| ARIMA (2, 1, 1) | 499.0814 |
| ARIMA (0, 1, 1) | 495.4348 |
| ARIMA (0, 1, 2) | 497.2587 |
| ARIMA (1, 1, 2) | 498.9234 |
| ARIMA (3, 1, 0) | 498.8315 |

A model with a lower AIC value is better than the one with a higher AIC value (Nyoni, 2018b) In this research paper, only the AIC is used to select the optimal model. Therefore, the ARIMA (1, 1, 0) model is finally selected.

**Residual & Stability Tests**

**Correlogram of the Residuals of the ARIMA (1, 1, 0) Model**

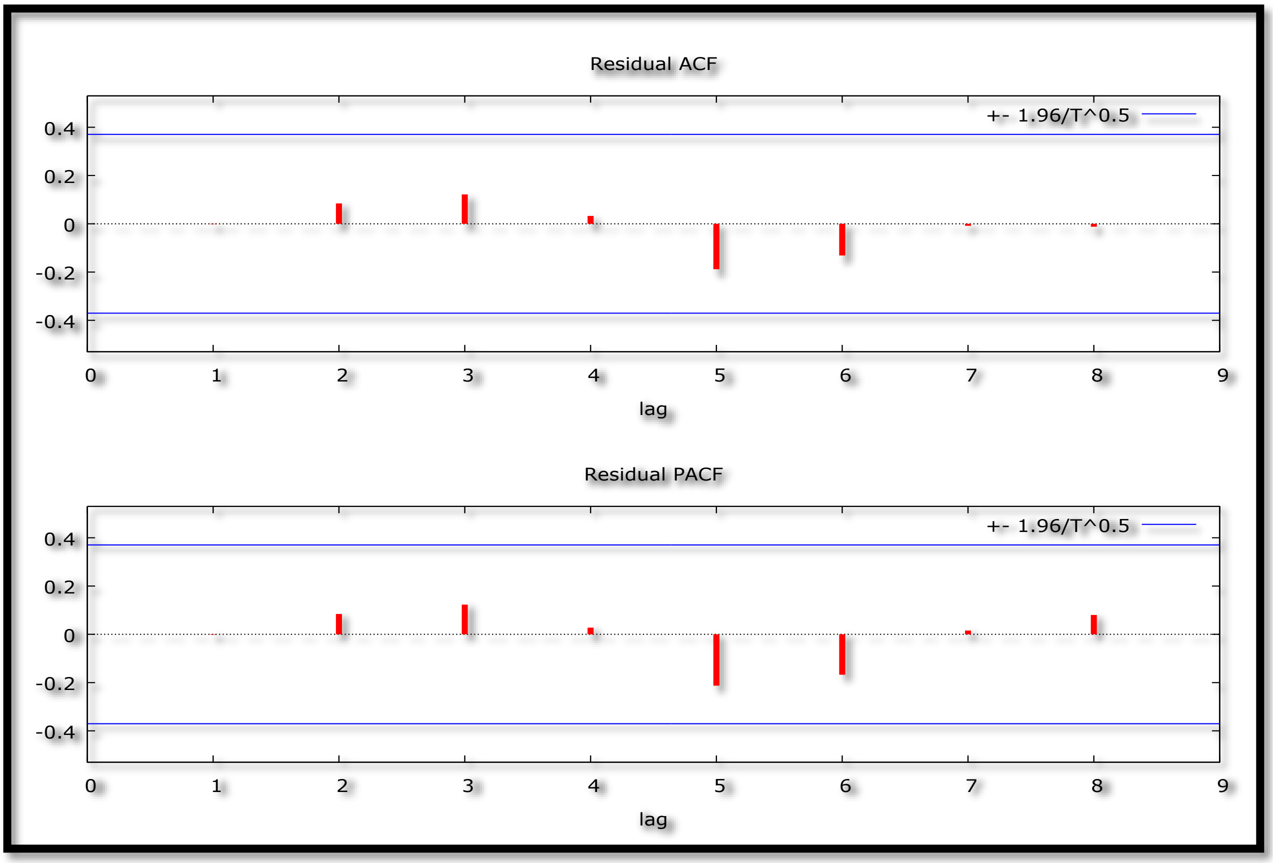


Figure 4: Correlogram of the Residuals

Figure 4 indicates that the estimated optimal model is adequate since ACF and PACF lags are quite short and within the bands. This implies that the “no autocorrelation” assumption is not violated in this paper.

**Stability Test of the ARIMA (1, 1, 0) Model**

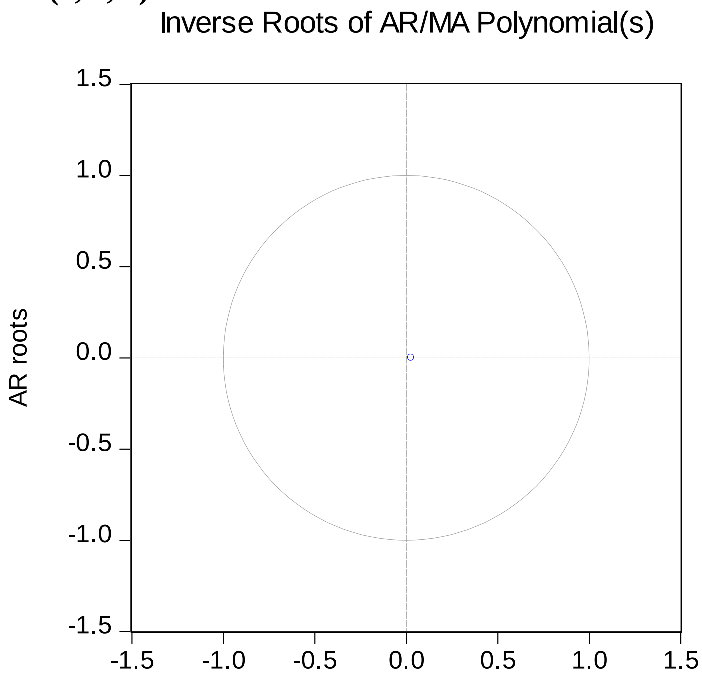


Figure 5: Inverse Roots

Since all the AR roots lie inside the unit circle, it implies that the estimated ARIMA process is (covariance) stationary; thus confirming that the ARIMA (1, 1, 0) model is stable.

## FINDINGS

### Descriptive Statistics

Table 4: Descriptive Statistics

| Description | Statistic |
|---|---|
| Mean | 7075.9 |
| Median | 1600 |
| Minimum | 100 |
| Maximum | 20000 |

As shown in table 4 above, the mean is 7075.9. This means that, over the study period, the annual average number of new HIV infections in Pakistan is approximately 7076 adults. This is a serious warning signal for policy makers in Pakistan. The minimum number of new infections is 100 while the maximum, observed in 2018, is 20000.

### Results Presentation

Table 5: Main Results

**ARIMA (1, 1, 0) Model:**

Guided by equation [4], the chosen optimal model, the ARIMA (1, 1, 0) model can be expressed as follows:

$$\Delta F_t = 710.288 + 0.0275781\Delta F_{t-1} \ldots \ldots \ldots \ldots \ldots [5]$$

| Variable | Coefficient | Standard Error | z | p-value |
|---|---|---|---|---|
| constant | 710.288 | 293.465 | 2.420 | 0.0155** |
| $\beta_1$ | 0.0275781 | 0.189034 | 0.1459 | 0.8840 |

Table 5 shows the main results of the ARIMA (1, 1, 0) model.

**Forecast Graph**

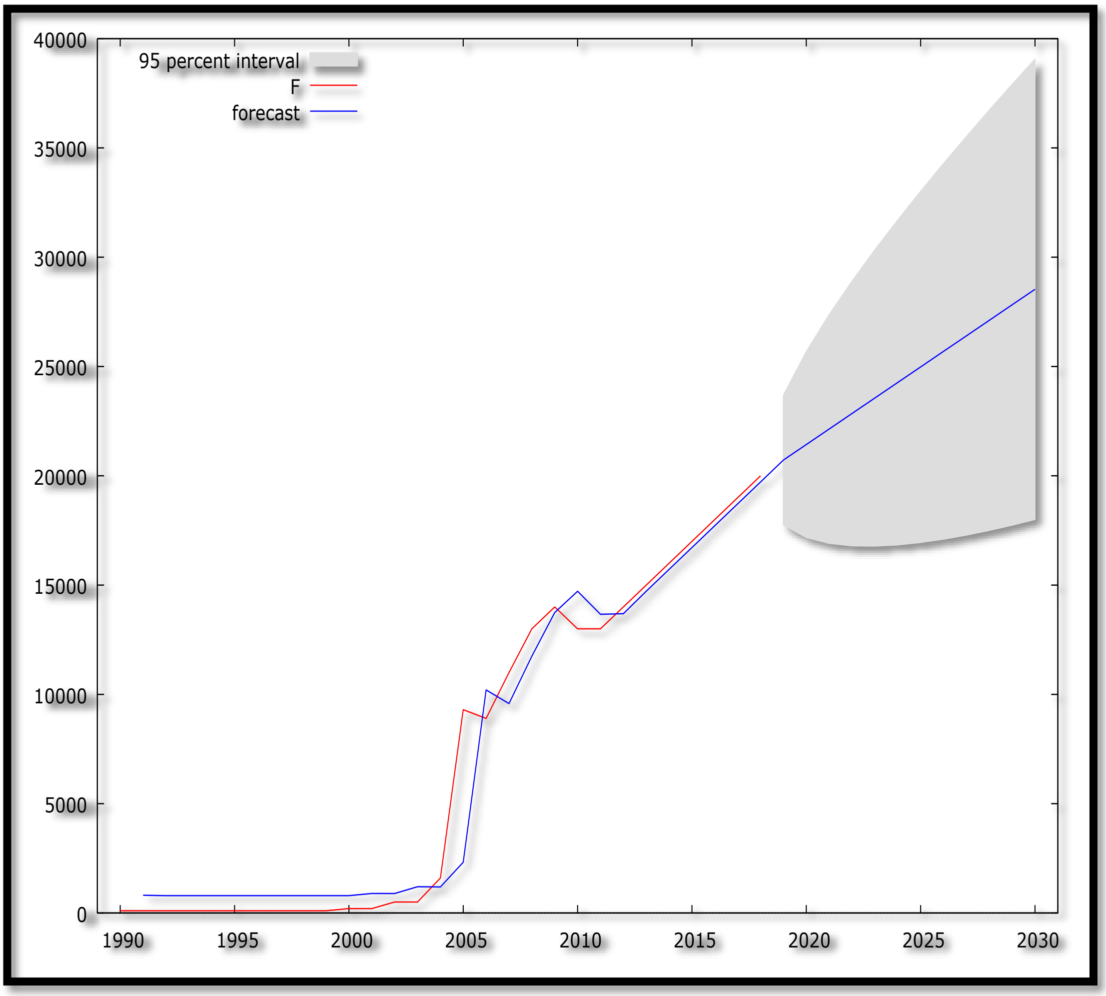


Figure 6: Forecast Graph – In & Out-of-Sample Forecasts

Figure 6 shows the in-and-out-of-sample forecasts of the F series. The out-of-sample forecasts cover the period 2018 – 2030.

**Predicted F– Out-of-Sample Forecasts Only**

Table 6: Predicted

| Year | Prediction | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| 2019 | 20718.3 | 1511.56 | (17755.7, 23680.9) |
| 2020 | 21428.8 | 2167.35 | (17180.9, 25676.7) |
| 2021 | 22139.1 | 2667.13 | (16911.6, 27366.6) |
| 2022 | 22849.4 | 3087.04 | (16798.9, 28899.9) |
| 2023 | 23559.7 | 3456.31 | (16785.4, 30333.9) |
| 2024 | 24269.9 | 3789.77 | (16842.1, 31697.8) |
| 2025 | 24980.2 | 4096.17 | (16951.9, 33008.6) |
| 2026 | 25690.5 | 4381.19 | (17103.5, 34277.5) |
| 2027 | 26400.8 | 4648.77 | (17289.4, 35512.2) |
| 2028 | 27111.1 | 4901.77 | (17503.8, 36718.4) |
| 2029 | 27821.4 | 5142.34 | (17742.6, 37900.2) |
| 2030 | 28531.7 | 5372.14 | (18002.5, 39060.9) |

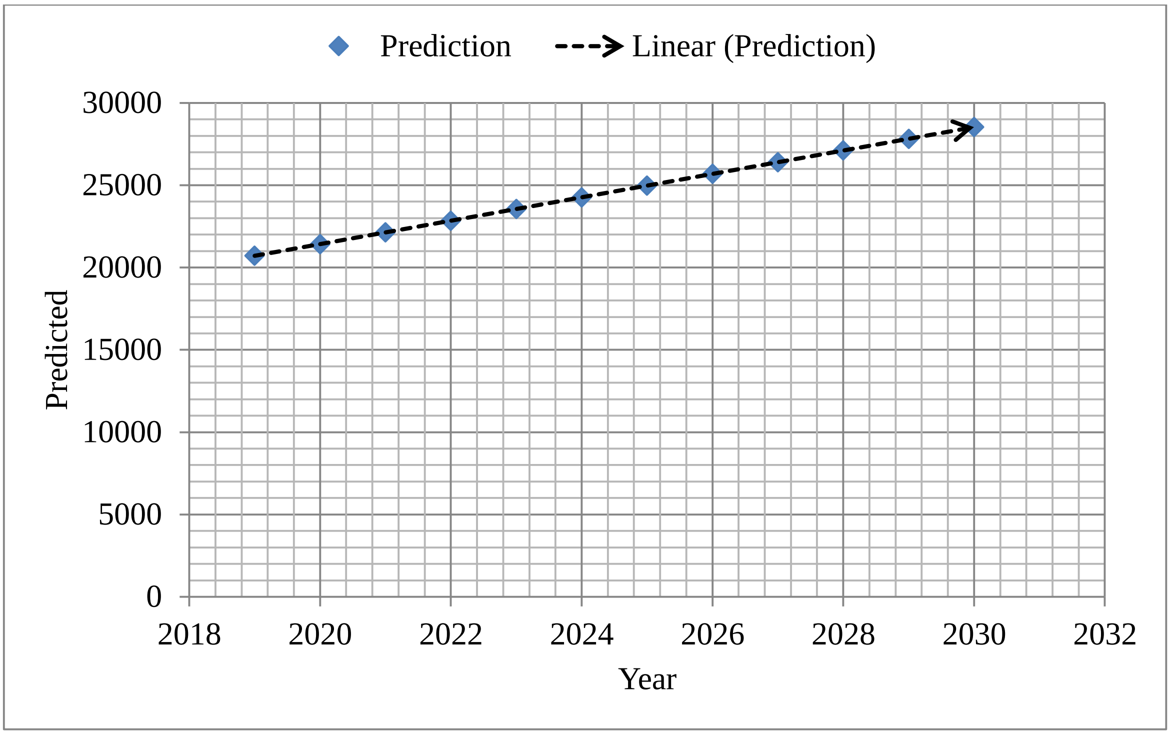


Figure 7: Graphical Analysis of Out-of-Sample Forecasts

Table 6 and figure 7 show the out-of-sample forecasts only. The number of new HIV infections in adults in Pakistan is forecasted to rise sharply from 20718 in 2019 to approximately 28531 by 2030. This is not a desirable health outcome for the country. The results of this study are line with scientific observations made by Mujeeb & Hashmi (1988), Khan & Khan (2012), UNAIDS (2014), Ghani (2016), the Ministry of National Health Services, Regulations and Coordination (2017) and Hussain et al. (2018) and is also consistent with previous empirical studies such as Maan et al. (2014), Khalid & Fox (2019) and Siddiqui et al. (2020).

## CONCLUSION

The study shows that the ARIMA (1, 1, 0) model is not only stable but also the most suitable model to forecast the annual number of new HIV infections in adults in the country over the period 2019 – 2030. The study recommends that the government of Pakistan should vigorously scale up HIV prevention and treatment access; with special emphasis on behavior change interventions such as increased condom use, reduction of sexual partners and elimination of cultural and religious stigma related to HIV. Educational programmes are pivotal in the fight against HIV/AIDS especially when considering the fact that in Pakistan there is high prevalence of cultural stigma related to HIV/AIDS. Health policy makers in Pakistan ought to strengthen HIV, TB, and Sexual & Reproductive Health programme linkages throughout the country. There is also a need for up scaling of medical male circumcision in Pakistan as an additional HIV prevention strategy. Last but not least, there is need for more political commitment in the country if new HIV infections are to be significantly averted any time soon.

## REFERENCES

1) Abdullah, M. A., & Shaikh, B. T. (2015). Review of HIV Response in Pakistan Using a System Thinking Framework, Global Health Action, 8 (1): 1 – 3.
2) BBC News (2011). Pakistan Battles Against Hidden HIV/AIDS Menace, BBC News, accessed 01/08/2020.
3) Bhatti, A. B., Usman, M., & Kandi, V. (2016). Current Scenario of HIV/AIDS, Treatment Options, and Major Challenges with Compliance to Antiretroviral Therapy, Cureus, 8 (3): 1 – 15.

4) Ghani, A. (2016). HIV/AIDS Cases on the Rise in Pakistan, The Express Tribune, accessed on 01/08/2020.
5) He, G., et al. (2018). Using the Baidu Search Index to Predict the Incidence of HIV/AIDS in China, Scientific Reports, 8: 1 – 10.
6) Hussain, A., et al. (2018). HIV/AIDS – A Growing Epidemic in Pakistan, Journal of Evolution of Medical and Dental Sciences, 7 (8): 1057 – 1062.
7) Khalid, H., & Fox, A. M. (2019). Political and Governance Challenges to Achieving Global HIV Goals with Injecting Drug Users: The Case of Pakistan, International Journal of Health Policy and Management, 8 (5): 261 – 271.
8) Khan, A. A., & Khan, A. (2012). Sexually Transmitted Infection Care in Pakistan: The Providers' Perspective, Journal of Pakistan Medical Association, 62 (9): 941 – 945.
9) Khan, R., Bilal, A., & Siddiqui, S. H. (2017). Knowledge About HIV and Discriminatory Attitudes Towards People Living with HIV in Pakistan, USAID, DHS Working Paper No. 134.
10) Maan, M. A., Hussain F., & Jamil, M (2014). Prevalence and Risk Factors of HIV in Faisalabad, Pakistan – A Retrospective Study, Pakistan Journal of Medical Sciences, 30 (1): 32 – 35.
11) Mansoor, H. (2017). UN Expresses Concern Over HIV Prevalence in Karachi, DAWN, accessed 01/08/2020.
12) Ministry of National Health Services, Regulations & Coordination (2017). AIDS Epidemic Modelling Exercise for Pakistan 2017, Ministry of National Health Services, Regulations & Coordination, Government of Pakistan.
13) Mujeeb, A., & Hashmi, M. R. (1988). A Study of HIV-antibody in Sera of Blood Donors and People at Risk, Journal of Pakistan Medical Association, 38 (8): 221 – 222.
14) Nyoni, T. (2018a). Modeling and Forecasting Naira/USD Exchange Rate in Nigeria: A Box-Jenkins ARIMA Approach, MPRA Paper No. 88622, University Library of Munich, Munich, Germany.
15) Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, University Library of Munich, Munich, Germany.
16) Sheikh, S. S. (2017). A Village in Punjab Suffers as the Government Dithers, DAWN, accessed on 01/08/2020.
17) Siddiqui, A. R., et al. (2020). Investigation of an Extensive Outbreak of HIV Infection Among Children in Sindh, Pakistan: Protocol for a Matched Case-Control Study, BMJ Open, 10: 1 – 8.
18) UNAIDS (2014). Global AIDS Response Progress Report, UNAIDS, Islamabad.
19) UNAIDS (2018). Country Factsheets – Pakistan, UNAIDS, Geneva.
20) Yusufzai, A. (2011). AIDS Drugs Access Fraught with Hazard for Pakistan's Tribal Regions, The Guardian News and Media, accessed on 01/08/2020.