Research Article

Journal of Economic Research & Reviews

Mozambique Consumer Price Index Estimation through Time Series Models

Filipe Mahaluça[1,2*], Fernando Nhampossa[3], Valentim Macuacua[1,2], Abú Bacar[1], Gilberto Solomone[1,4], Manogil Mambo[1], José Manjate[1] and Alfeu Vilanculos[1,4]

[1]*Instituto Superior de Contabilidade e Auditoria de Moçambique (ISCAM)/ Higher Institute of Accounting and Audit of Mozambique (ISCAM), Mozambique.*

[2]*Universidade São Tomás de Moçambique (USTM) – Escola de Pós-Graduação/ São Tomás University of Mozambique (USTM) - School of Postgraduate, Mozambique.*

[3]*Escola Superior de Ciências Náuticas (ESCN)/ Higher School of Nautical Sciences*

[4]*Universidade Eduardo Mondlane-Faculdade de Economia (UEM)/ Eduardo Mondlane University-Faculty of Economics (UEM), Mozambique.*

[*]Corresponding author:
Filipe Mahaluça, Instituto Superior de Contabilidade e Auditoria de Moçambique (ISCAM)/ Higher Institute of Accounting and Audit of Mozambique (ISCAM), Mozambique.

Submitted: 16 Feb 2022; **Accepted**: 22 Feb 2022; **Published**: 12 April 2022

Citation: : *Filipe Mahaluça, Fernando Nhampossa, Valentim Macuacua, Abú Bacar, Gilberto Solomone, Manogil Mambo, José Manjate, Alfeu Vilanculos. (2022). Mozambique Consumer Price Index Estimation through Time Series Models, J Eco Res & Rev, 2(2), 70-80.*

Abstract

Background: *The Consumer Price Index (CPI) is the main indicator for determining the prices level of a certain economy. The above-mentioned rate refers to the average prices on an indicated period and it is calculated mainly on goods consumed basis by selected families through weighting criteria and therefore, identifying the inflation damages on waged workers as well as their life standards thus, providing precise information on their earnings improvement or compensating their previous or equivalent buying empowerment. The application of rates is based on statistics models whereby; the temporary series is identified as observation of different periods that relevant variable values stand to foresee the future hypothesis confirmation.*

Objective: *To Estimate the Consumer Price Index in Mozambique using Time Series Models.*

Methods: *The following study based on monthly CPI data of Mozambique dating January 2011 to July 2020 out of 115 observed analyses. The data was processed through free open codes Gretl and 4.02-R Version Statistic Packages. The data hereby used was provided by the Mozambican National Statistics Institute (INE). The study applied the Box and Jenkins approach claiming that each temporary series value is based on its previous values due to the temporarily correlation between the series values that generally exist. The method consists of adjusting the Auto-regressive integrated Moving Average Method (ARIMA) to a data group under four interactive stages cycles*

Results: *From 103 observations taken into account, the highest registered index was 158,25 corresponding to December 2016. On the other hand, it is notable that 101,62 is the smallest rate corresponding to January 2011period. The 114,68 medium rate breaks apart the price series in half resulting 50% of rates below such price and consequently, the remaining 50% above the same price. The similar parameter indicates all distribution format ends. Positive numbers indicate the widest end on the right while the negative numbers indicate the widest end on the left. Consequently, the positive asymmetry (1, 5217) indicates that the distribution has a wide end on the right. For kurtosis, the series is platykurtic since its result is less than 3 (2,6904 <3).*

Conclusion: *The CPI in Mozambique showed an increasing trend in the period analyzed, between January 2011 and July 2020, with a sharp rise in 2016 returning to the original level and the normal growth trend of the series. In the estimation stage, it was possible to adequately select (02) two ARIMA models (p,d,q) that presented good fit to the data, and among* *them, the model selected as the most appropriate for predictions according to performance measures, was ARIMA (0.1.0) with constant, because it presented lower values of absolute mean percentage error (AMPS) and mean quadratic error (REQM). According to the selected model, the previsions for the period from August 2020 to July 2021 indicate a slight growth in consumer price index.*

Keywords: Consumer Price Index; ARIMA; Seasonality; Series; Auto-regressive

Introduction

The Consumer Price Index (CPI) is the main indicator for determining the prices level of a certain economy [1]. When analyzing the inflation, the CPI is mostly the selected resource. Taking the United States of America (USA) as an example, the CPI is worked out by the Bureau of Labor Statistics (BLS) while in Mozambique is under the National Statistics Institute (INE). The CPI enhances the public institutions on laying policies over consumer prices, update public prices as well as on taxation and therefore used as an indicator of contractual adjustor within both public and private institutions such as Syndicates, Universities, Ministries and so forth.

The above mentioned rate refers to the average prices on an indicated period and it is calculated mainly on goods consumed basis by selected families through weighting criteria and therefore, identifying the inflation damages on waged workers as well as their life standards thus, providing precise information on their earnings improvement or compensating their previous or equivalent buying empowerment [2].The estimation of CPI as a unique rate for measuring the general pricing level, is done through a collection of various price commodities and services integrated in an economy type basis with various pondered categories allowing therefore, a monthly temporarily and sequentially organized issuing of data results organized in a time serial form [3]. The CPI is applied to determine the inflation mainly for the BLS case of the USA used to calculate the cost of goods percentage variation acquired in a based period whereby the standard expenditure changes are incorporated taking into account the updating of BLS market basket base [4].

The application of rates is based on statistics models whereby, according to Pinheiro (2017), the temporary series is identified as observation of different periods that relevant variable values stand to foresee the future hypothesis confirmation [5].

The temporary series is data laid on time basis since the general economically correlated can be represented on a single chain. Various authors claim that the temporarily series model analysis is eventually applied in the study of such data category analysis aiming at, (i) series generator research, (ii) long and short term future provision analysis; (iii) series outstanding description with graphic tendencies variation existence, Cycles and seasonal variations of data research [6,7]. Generally, stationary temporary series applies the Autoregressive Integrated Moving Average Method (ARIMA) developed by George Box and Gwilym Jenkins [8]. The ARIMA method includes autoregressive parameters and those of moving averages and explicitly includes differences in the model formulation [9]. When analyzing the series through ARIMA model this must be stationary, i.e., it must have an average, variable and constant autocorrelation throughout the time.

Generally, there are four economic provision approaches based on temporary series namely, unique regressive equation models, simultaneous regressive equation models, ARIMA and Autoregressive Vector (ARV). Alternatively, to Box-Jenkins ARV is applied. Temporary series are not always stationary. Nevertheless, if it is the case, some techniques have to be used in order to turn them so before ARIMA is applied. The majority of temporary series standards can be described in two basic class components namely, seasonal and tendency. The first represents a general linear component and in most cases represent a non-linear altering throughout the time and does not repeatedly appear along the research analysis. The second is similar to the first but differs only because it cyclically appears in time systematic intervals [10]. The two temporary series components coexist in real problems. Formally, the (Zt, $t=1, 2, ...,n$) temporary series observations can be decomposed into the following model : $Zt = Tt + St + at$ (1) whereby Tt and St respectively represent the tendency and the seasonally while at is a random of zero average and constant [11].

Considering the aforementioned, the present study aims to Estimate the Consumer Price Index in Mozambique using Time Series Models.

Literature Revision

Consumer Price Index (CPI)

A precise evaluation of goods and services as well as the inflation is a fundamental key factor in any economy concerns. The economies have millions of market available goods and services; some are daily launched, the already existent are improved while others are launched to the market. In response to several factors the prices of goods and services constantly change [12]. The variable price effect and the inflation are measured through the consumer price Index. The CPI is one of any country's important economic measures as it displays economic effects namely, economic and productivity growth, the government taxes, budget deficit and debts, monetary policy real financial incomes, real salaries, average medium earnings and poverty taxes.

The process of price fixing is fundamental in micro economy. It affects different economic activities of any nation with economic growth, public welfare and Gross Domestic Product (GDP) [14]. Although its relevance, still there is no consensus as to which variables affect the pricing behavior [15]. The internal prices are both determined by internal costs and worldwide prices [16].

The CPI is projected to measure the cost of market basic goods and services representing specific period medium consumption standards. A rate based on an historic market basis and steady is known as the Laspeyres rate. When estimating the CPI the Biases must be taken into account no matter the model applied. The CPI of the USA is criticized due to the rise of the life standard cost as ivies replacement establishing a difference between the superlative index with the Laspeyres [17].

CPI Estimation

Within mid-1995, the North American Senate Finance Commission appointed a Consultative Commission to undertake CPI studies. The commission realized that the CPI strongly varies 1.1 % on life cost standard a year plus 0.8-1.6 % plausible figures [18]. The biases seem to be small but when combined with time results in significant implications. There was also a budget super indexation which resulted in more than US $ 3M a cumulative debt [12].

Strasburg, Aka and Pieretti (2008) studied the relationship between the CPI and its detailed variable theories known as Structural Time Series (STS) in order to account the CPI non observable components (tendency, cycle, seasonal and irregular). On one hand, the results show that the CPI is both positively related with external prices and negatively with work productivity as the theory claims. On the other hand, salaries are less relevant when dealing with the economy of Strasburg. Nevertheless, the work cost unity best explains the Consumer Price Index [19].

Lee (2012) conducted a CPI study in four Korean cities namely, Seoul, Busan, Daegu and Gwangju. The data hereby used in the studies was extracted from the National Statistical Office between 1998 to 2011 through auto regressive error model which is the temporary series model. The detailed variables used for consumer prices analysis are the following: Economic coincidence rate, the USA exchange rate, the producer price rate, the petroleum importing unitary price, the crude importing quantities and the international current account [20]. The importing price rates, the unemployment rate and the currency volume were also applied in this concerned analysis. The result of this analysis shows that regional CPI lies within 46% to 52% of auto regressive error model.

Jackson et al (2018) conducted a monthly consumer model index study in Serra Leon through Box- Jenkins seasonal methodology [21]. This study showed that there was a high disposal of food commodities in the local market and this factor directly contributed to prices decreasing. Relative foreign exchange rate stability was also observed due to the same factor plus the country's high interest on exporting mineral commodities in global and steady market. Positive prevision risks were also identified especially when the government financed the internal debt via the central bank contributed to inflation pressure threat due to domestic economy excess currency disposal.

McGranahan and Paulson (2006) carried on a study on 31 Chicago demographic groups aiming at building the CPI basing on consumer expenditures data within 1982-2004 combined with CPI of specific commodities. From this study, they came to a conclusion that different group inflation experiences were highly correlated and generally similar to the magnitude of urban population's inflation [4]. They also realized that groups' accumulated inflation varies between 195% and 212%. Nevertheless, the group with an average of 11% high inflation rate bend was the one led by a 65 year old or more family member. Besides, the study also found out that variable inflation was higher in vulnerable population than the favored population. They also claimed that the standard inflation bend diminishes in accordance with educational background level. Less educated inflation rate stands on 3% more variable than all urban population inflation.

Furth (2017), studied the USA inflation and concluded that the North Americans lived a relatively higher life standard in mid-year 2000 rather than in the end of 1970. According to him, Personal Consumer Expenditures was 0.3% inflation tax rates less than the Consumer Price Index. He then adjusted the Personal Consumer Expenditures with the CPI which resulted in 7% average salaries rise between 1979 and 2007 whereby the family earning grew an average of 35% [22].

Evidences show that CPI estimation can be carried through various models. Mahmoud (2015) claim that when the biases occur favor criticisms on some rates since lots of them underestimate many families and life standard cost [15]. They also add that it is relevant to estimate the price index as it provides families aggregated micro economics behavior that enhance strategic policies setting for the development of any country.

Methods and Materials

Materials

The following study based on monthly CPI data of Mozambique dating January 2011 to July 2020 out of 115 observed analysis. The data was processed through free open codes Gretl and 4.02-R Version Statistic Packages.

Data Source

The data hereby used was provided by the Mozambican National Statistics Institute (INE)

Methods

The study applied the Box and Jenkins approach claiming that each temporary series value is based on its previous values due to the temporarily correlation between the series values that generally exist.

The method consists of adjusting the Auto-regressive integrated Moving Average Method (ARIMA) to a data group under four interactive stages cycles Morettin and Toloi 2006, namely:

i. Identification: An adequate model for the series in analysis is identified basing on partial autocorrelations and other criterions;

ii. Estimation: The parameter models are identified;
iii. Diagnostics: The adequate model is verified through residues analysis;
iv. Prevision: The final model is applied in the prevision of future series values;

In cases where the selected model is inadequate the cycle is repeated backing to the identification stage. The data analysis firstly started in graphic analysis in order to check non steady, seasonal and tendency possible traces [23].

Unique Root Test

The presence of unique root testing process was proceeded through Dicky-Fuller (ADF) based on the following model:

$$\Delta x_t = \alpha + \beta t + \eta x_{t-1} + \xi_t$$

$$\textit{where} \quad \eta = \sum_{i=1}^{p} \rho_i - 1$$

X_t represents a dependent variable, and the $\Delta x = xt - x_{t-1}$ difference operator. The parameters to be estimated are α, β, η. The Dickey &Fuller (1998) τ_τ, τ_μ e τ statistics correspond to the test to estimate the x_{t-1} equation (25) variable results. These statistics are specifically for a model that covers mobile parameter and a (τ_τ) tendency, a model including a (τ_μ) constant value and a (τ) model with no constant or moving value. The tested hypotheses in such models correspond to an invalid hypothesis and to a ($H_0: X_t$) it is neither $I(0)$ nor $\eta = 0$) non-stationary series against the alternative hypothesis stating that it is a non-integrated series, ie, it is a ($H_0: X_t$ é $I(0)$) series.

There were also applied Dickey's & Fuller's (1979-1981) ϕ_3 e ϕ_1 statistics testing if the variable tendency results and the x_{t-1} variable result and if the constant are in (1) zero equal respectively. (X_t) endogenous variable can be incorporated to equation (1) aiming at eliminating the presence of auto correlation between the error terms:

$$\Delta x_t = \alpha + \beta t + \eta x_{t-1} + \sum_{i=1}^{p-1} \lambda_i \Delta x_{t-i} + \xi_t$$

$$\textit{where} \quad \lambda_i = -\sum_{j=i+1}^{\rho} \rho_j$$

P, is the auto-regressive order or non-correlated number of resulting residues. In this case we are in the presence of Dicky-Fuller (ADF) test [24].

Seasonal state Kruskal-Wallis Test

The seasonal state was tested through Kruskal-Wallis test. Morettin and Toloi (2006) adopted the Kruskal-Wallis as alternative to identify the temporal seasonal series. Taking into account that the data is collected in equal time periods each K is seen as size sample, and therefore, the following is observed:

$$X_{ij},\ j=1,\ i=1,\ \ldots,\ nj \quad n = \sum_{i=1}^{k} n_{j.}$$

The X_{ij} observations are replaced by R_{ij} posts, and all N observations are placed in sequence order. R_{ij} is j-data sum up of associated posts:

$$R_{.j} = \sum_{i=1}^{nj} R_{ij} \quad j = 1, \ldots, k$$

The statistics test results from:

$$T_1 = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_{.j}^2}{nj} - 3(N+1)$$

The hypotheses taken into account are the following:
H_0: There is no seasonal case;
H_1: There is seasonal case.

The null hypothesis is rejected if the statistics test is more or equal to T_{1C} critic value (x^2 charted value) ($T_1 \geq T_{1C}$) for a meaningful level value. The T_{1C} critic value is p_h ($T_1 \geq T_{1C}$) = α where α is a relevant test number. For the case where n_j is big or , $k \geq 4$ under H_0, T_1 distribution can be approximated with x^2 variable with of k - 1 freed degrees [23].

Data Transformation

After applying the statistics tests and realised the non-stationary temporal series in analysis there was a need for transforming the original data into stationary temporal series as most of temporal series statistics analysis procedures are claimed to be statics [25]. According to Morettin and Toloi (2006) riginal data can be transformed aiming at stabilizing the variance. In case where the series shows no variance it must therefore, be transformed in order to make it balance [23]. Such transformations are also aimed at distributing more symmetric data and make them close to normal distribution where the original series was submitted into logarithm transformation in order to stabilize the variance.

Identification

At this stage, a corelogram graphic analysis was conducted in order to identify possible standard or behaviour of the model to be estimated. Only the (0,1,0) ARIMA model was identified due to the standard provided by corelograms since no relevant discrepancy in the correspondent I(1) series corelograms was observed and therefore enhanced the selection of the best model. Autocorrelation functions (ACF) and partial autocorrelations functions (PACF) are applied to build up corelograms:

$$ACF:\tau_k=\frac{\sum_{i=k+1}^{T}\left(X_t-\bar{X}\right)\left(X_{t-k}-\bar{X}\right)}{\sum_{t=1}^{T}\left(X_t-\bar{X}\right)}$$

$$PACF:\phi_k=\begin{cases}\tau=1,k=1\\ \dfrac{\tau_k-\sum_{j=1}^{k-1}\phi_{k-1,j}\tau_{k-j}}{1-\sum_{j=1}^{k-1}\phi_{k-1,j}\tau_{k-j}},k>1\end{cases}$$

Next, Q Ljung-Box and p values were applied to test if k autocorrelations and partial autocorrelations are statistically different from zero when checking its statistics values.

Q Statistics test that the first k autocorrelations hypothesis are different from zero ($H_0: \tau_k = 0$) and calculated through the following equation with *m-p-q free co*

$$Q(m)=n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k}{(n-k)}$$

With approximately m - p - q chi-square free degrees.

Estimation

The parameters model estimation based on maximum verossimilarity model (MVM0 and its equation is the following:

$$L(\xi/X,x_0)=(2\pi\delta^2)^{-\frac{T}{2}}\exp\left\{-\frac{1}{2\delta^2}\sum_{t=1}^{T}(X_t-\phi_0-\phi_1X_{t-1})^2\right\}$$

Generally, statistics packages applied in seasonal series include an algorithm for parameters estimation.

Verification

Verifying estimated models implies adjusting very well the series data and if the residues are purely casual and normal based on residues graphic analysis and the Box-Pierce, Ljung and Box tests. The Box-Pierce test verifies if an autocorrelation group is significantly different from zero through statistics with the following equation:

$$Q=n\sum_{k=1}^{m}r_k^2$$

Practically, the number of autocorrelation data is typically selected between15-30. If the adjusted model is appropriate this will result in an approximately m - p - q chi-square free degrees. According to Ramos (2010), Box-Pierse test has not a good performance in small or moderated samples in a way that it stays away from the chi-square distribution and suggested a replacement with Ljung and Box (equation 7) [26]. The null hypothesis for residues is rejected to $Q(m)$ high values.

If the model is correct the residues must be independent and identically distributed with . Shapiro-Wilk and Jarque-Bera tests are applied to verify if the normal procedures were correctly observed. Other residues usual verifications were covered such as the histogram with a normal topped curve and the normal probabilities graphic (QQ-plot).

Arch-Lm Test

Verifying the heterocedasticity model the ARCH-LM test was applied basing on the following equation:

$$X_t=\beta_1+\beta_2X_{t-1}+\mu_t$$

Relevant specific procedures were used to detect the presence of ARCH, namely:

a). The equation is estimated through OLS and results X_t;

b). residues are saved and the $\varepsilon_t^2 = \omega_1 + \omega_2\varepsilon_{t-1}^2 + \delta$ equation is estimated (11) where:

ε_t^2 = error variance (h_t^2) in t (time);

ε_{t-1}^2 = error variance (h_t^2) in t-1 (time);

δ = error

The most important issue is to observe if "ω_2" is statistically significant having the following hypothesis alternatives:

$H_0:\omega_2=0 \Rightarrow$ *homocedasticity* hypothesis

$H_1:\omega_2\neq 0 \Rightarrow$ *heterocedasticity* hypothesis

The test is based in qui-square distribution where"p" free degrees obtained from Lagrange multiplier, i.e.:

$$R^2\times n\sim\chi_p^2$$

Where:

P = ε^2 break down number;

n = Observation number;

R^2 = Result of an estimated equation.

Models Comparisons

MacClain & Humphreys (1996) suggested the following best prediction model:

Absolute Medium Error (AME) $=\frac{1}{h+1}\sum_{t=n}^{n+h}\left|X_t-\hat{X}_t\right|$

Percentual Absolute Medium Error (PAME)

$$=\frac{100}{h+1}\sum_{t=n}^{n+h}\left|\frac{\left(X_t-\hat{X}_t\right)}{X_t}\right|$$

Medium Square Root Error

$$= \sqrt{\frac{1}{h+1}\sum_{t=n}^{n+h}\left(X_t - \hat{X}_t\right)^2}$$

h is the prevision periodic number, n is the sample size, X_t is the CPI observed in t period and $\hat{X}_t$ the CPI for t period. The smaller the error is the better is the prediction model capacity.

There were also applied the Akaike (AIC) and Schwarz information criterions observing the minimum values criterion and are highly penalizing. According to Gujarati (2006), the model that shows less statistics values has the best performance [27].

Analysis and Discussion of Results

Exploratory series Analysis

There were initially used January 2011-July 2019 CPI series data whereas the August 2019 up to July 2020 values used to evaluate the adjusted model validity. Lately, the model was applied to predict July 2021 period. The whole analysis was again applied to analyse complete series from January 2011 up to July 2020.

Table 1 displays the main descriptive measures series monthly obtained from a correspondent CPI, for a specific period:

Table 1: CPI monthly descriptive measures

Descriptive Statistics	Value
Medium	116,50
Average	114,68
Standard bend	11,008
Variable Result	0,094489
Maximum	158,25
Minimum	101,62
kurtosis	2,6904
Asymmetry	1,5217
Author (2020)	

On one hand, from table 1, we can realise that the medium series data was 116,50 with 10,008 standard bend. We can also see that the variable result (medium and standard bend relationship) is approximately 0,094. From 103 observations taken into account, the highest registered index was 158,25 corresponding to December 2016. On the other hand, it is notable that 101,62 is the smallest rate corresponding to January 2011period. The 114,68 medium rate breaks apart the price series in half resulting 50% of rates below such price and consequently, the remaining 50% above the same price.

The similar parameter indicates all distribution format ends. Positive numbers indicate the widest end on the right while the negative numbers indicate the widest end on the left. Consequently, the positive asymmetry (1, 5217) indicates that the distribution has a wide end on the right. For kurtosis, the series is platykurtic since its result is less than 3 (2,6904 <3).

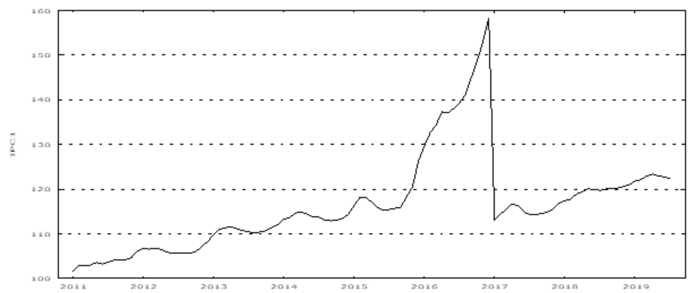

Figure 1: January 2011– July 2019 CPI evolution

Series Figure 1 shows that there is a CPI crescent tendency with a normal series intense increase in December 2016. The crescent tendency is a strong indicator that the series is not stationary. The other tool for checking the series stationary is through observing the auto correlation function (ACF). Graphic 2 shows that the auto correlation function presents a smooth decreasing breaking, suggesting therefore, that the series is a non-stationary one.

For the seasonal aspect, table 1 does not show equal tendencies in specific months every year therefore indicating non seasonal series. The other way of analyzing could be via ACF and PACF where if the (12) period delayed multiples are significant then the seasonal case is present. Nevertheless, such case does not take place since the (12) delay multiple is significant in ACF and PACF suggesting that the series presents no seasonality to be confirmed through formal tests to be laid ahead. Aiming at checking the data logarithm transformation the relationship between amplitude and medium was analysed. According to Morettin and Toloi (2006) when the relation produces a significant statistically result there is a need of series transformation. The result has presented a crescent linear behaviour between the medium and the amplitude suggesting therefore a transformation in the series [23]. The hypothesis test for leaning adjusted result with 5% significant level the null hypothesis was rejected (P-value=0,0139103), i.e., the leaning result is statistically zero equal.

Besides, the variant result is a reasonable indication of data favourable transformation. When the transformed data variant result is smaller than the genuine variant data then, the transformation is valid and, if not there is no use in applying it. For such case, the logarithm data variant result is less than the genuine data ($CV_{Log(IPC)} = 1{,}8775\% < CV_{(IPC)} = 9{,}4489\%$) and this situation reinforces the need of series transformation resulting therefore in variant stabilization.

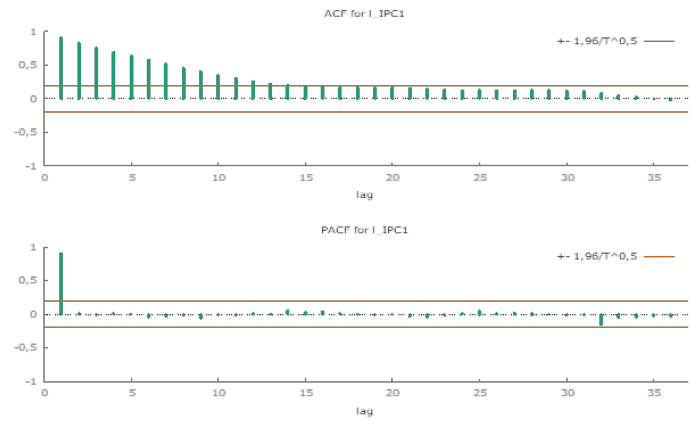

Figure 2: CPI series ACF & ACP logarithm transformation

Figure 2 displays the auto correlated function with a slow decline which suggests a non-stationary series.

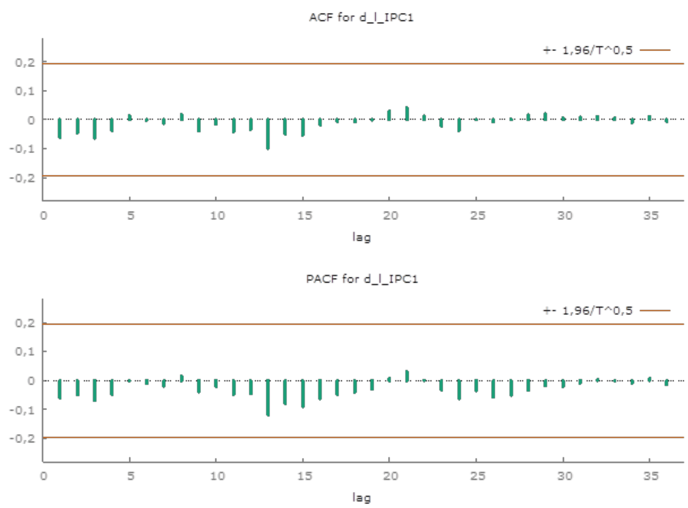

Figure 3: CPI logarithm transformed first differences of ACF & ACP series.

The CPI with logarithmic transformation was submitted to first simple differences (figure 3) aiming at removing the tendency. The auto correlation functions show a sharp decline with non-significant break-downs suggesting that the logarithmic transformation series is stationary in the first simple differences.

Stationary and Seasonal Verification Through Formal Tests

After the graphic analysis the next step is testing statistically the previously mentioned perceptions since the non-stationary and seasonal rates graphically obtained are not enough as they only indicate a possible series behavior. Therefore, the logarithmic transformation series was submitted to a and the risen Dickey-Fuller (-2,25211) statistics value shows less module value against 5% significant level of critic value meaning that the series in analysis is not stationary (see table 2).

A test of unique root for the first differences was conducted aiming at finding out the series integration. The Dickey-Fuller statistics added value (-10,5617) is highly superior in module against 5% critic value (see table 2); this means that the series is stationary in its first simple differences.

Table 2: LOG (CPI) Unique Root Test

LOG(CPI)	LOG(CPI)
Dickey-Fuller Statistics	Critic Value (5%)
-2,25211	-2.88
Augmented Dickey-Fuller Test	DLOG(IPC)
Dickey-Fuller Statistics	Critic Value (5%)
-10,5617	-2.88
Source: Authors (2020)	

The Kruskal-Wallis non parameter test was conducted in order to verify the seasonal aspect and the obtained a p-value=0,9991 (see table 3) was obtained. Therefore, the non-determined seasonal existence null hypothesis is not rejected and does not even show a stock full season since it does not show delayed 12 multiples out of reliable intervals when a differentiated series correlogram (see figure 3) indicating that there is no need of any added value to the seasonal model component.

Table 3: LOG (CPI) Seasonal Test

Kruskal-Wallis	LOG(IPC)
Chi-Squared Statistics	P-Value
1,7753	0,9991
Source: Authors (2020)	

dentification, Estimation and Diagnosis

According to figure4, there is no significant delay in both FAC and FACP. These results suggest possible absence of auto-regressive terms and moving averages in the candidate models.

Therefore, two (02) Autoregressive Integrated Moving Average (ARIMA) models (p,d,q) were tested, to highlight: ARIMA (0.1.0) with constant and ARIMA (0.1.0) without constant, which were estimated and analyzed for the significance of the parameters and adequacy of the model, having been shown to be able to represent the series of the Consumer Price Index.

Table 4: Estimated Models, ARIMA (p,d,q) for LOG (CPI) series

ARIMA (0,1,0) without constant				
Parameter	Estimates	Standard Error	Z Value	Pr(>|z|)
-	-	-	-	-
ARIMA (0,1,0) with constant				
Parameter	Estimates	Standard Error	Z Value	Pr(>|z|)
Constant	0,00181765	0,00347918	0,5224	0,6014
*Codes meaning: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*				

Source: Authors (2020)

After estimation, the residues were analyzed for the characteristics of independence, through the Ljung-Box and Box-Pierce tests, normality, through the Shapiro-Wilk and Jarque-Bera tests and the ARCH test to verify that the squares of the residues depend on their past values. The results are given in table 5 and indicate the non-rejection of the null hypothesis of independence of the residues for the Ljung Box and Box Pierce tests at the level of 5% significance, that is, the residues are not correlated over time. After the estimation, the residues were analyzed

Table 5: Diagnostic test of model residues

Model	Ljung-Box	ARCH	Shapiro-WilK	Jarque-Bera
ARIMA (0,1,0) withou constant	0,9997	0,99917	< 2.2e-16	< 2.2e-16
ARIMA (0,1,0) with constant	0,9997	0,99861	< 2.2e-16	< 2.2e-16

Source: Authors (2020)

Through the Shapiro-WilK and Jarque-Bera test for the significance level of 5%, there is evidence that the residues do not follow normal distribution. Therefore, for sample sizes that are large enough, the violation of the premise of normality is practically inconsequential, because by the central limit theorem, the test statistics follow the appropriate distributions, even in the absence of normality of errors [28].

The results of the ARCH test present p-values greater than 5% for the two models, so the series of residuals present constant variances because the squares do not depend on their past values. In this case, there is no need to estimate conditional heteroscedasticity model.

Comparison of Model Performance

Table 6: Comparison of the performance of the models within the sample

Model	AIC	EAMP	REQM
ARIMA (0,1,0) without constant	−392,3545	0,22026	0,035013
ARIMA (0,1,0) with constant	−390,6297	0,20901	0,034965

Source: Authors (2020)

According to the results of table 6, the two models were well adjusted. However, despite the ARIMA model (0.1.0) with constant presenting higher value of the Akaike Information Criterion (AIC) in relation to the non-constant model, in terms of comparison of performance within the sample, the model with constant provides better performance in relation to the other, because it has a lower value of the absolute mean percentage error (AMPE) and the mean quadratic error (REQM).

Forecasts show that the ARIMA model (0.1.0) with constant performance presents better performance, since its predictions are close to the actual values observed, which means that the model with constant has less forecast error in relation to the model without constant.

Comparison of The Performance of Out-Of-Sample Models

Table 7 shows the predicted CPI values from August 2019 to July 2020 produced based on the two adjusted models, ARIMA (0.1.0) without constant and ARIMA (0.1.0) with constant, in order to compare the predictive capacities outside the sample.

Table 7: Comparison of CPI forecast values from August 2019 to July 2020

		ARIMA (0,1,0) with constant			ARIMA (0,1,0) without constant		
Month	Real	Forecast	Confidence Intervals confidence 95%		Forecast	Confidence Intervals 95%	
			Lower	Higher		Lower	Higher
August/2019	122,46	122,5425	114,3872	131,2793	122,3201	114,1794	131,0410
September/2019	122,58	122,7655	111,3725	135,3240	122,3201	110,9683	134,8329
October/2019	122,96	122,9889	109,1593	138,5705	122,3201	108,5658	137,8169
November/2019	123,69	123,2126	107,3584	141,4081	122,3201	106,5807	140,3836
December/2019	125,27	123,4368	105,8193	143,9873	122,3201	104,8619	142,6847
January/2020	126,06	123,6613	104,4650	146,3851	122,3201	103,3319	144,7973
February/2020	126,52	123,8863	103,2500	148,6471	122,3201	101,9446	146,7677
March/2020	126,80	124,1117	102,1447	150,8028	122,3201	100,6701	148,6258
April/2020	127,46	124,3374	101,1285	152,8729	122,3201	99,4876	150,3924
May/2020	126,70	124,5637	100,1865	154,8723	122,3201	98,3819	152,0826
June/2020	126,00	124,7903	99,3075	156,8121	122,3201	97,3417	153,7079
July/2020	125,75	125,0173	98,4828	158,7010	122,3201	96,3581	155,2771

Source: Authors (2020)

A new adjustment of the model was set, with all observations (data from January 2011 to July 2020), with a view to predicting the CPI for the next 12 months whose actual observations are unknown. It should be noted that, like the first adjustment, the stages of verifying the presence of the components trend, seasonality, independence of residues, normality of residues and arch effect were respected.

Table 8: Estimated models, ARIMA (p,d,q) for the complete LOG series (CPI)

ARIMA (0,1,0) without constant				
Parameters	Estimates	Standard Error	Z Value	Pr(>|z|)
-	-	-	-	-
ARIMA (0,1,0) with constant				
Parameters	Estimates	Standard Error	Z Value	Pr(>|z|)
Constant	0,00186891	0,00311509	0,6000	0,5485

Source: Author (2020)

Table 9: Performance comparison of log full series (CPI) models

Modelo	AIC	EAMP	REQM
ARIMA (0,1,0) with constant	−451,0983	0,20700	0,033167
ARIMA (0,1,0) with constant	**−449,4609**	**0,19529**	**0,033114**

Source: Author (2020)

According to table 9, it is also verified that the performance measures of the model with constant remain lower in relation to the model without constant, which means that the new adjustment with constant remains better despite presenting higher value of the AIC.

Forecasting with Selected Model

Table 9: CPI forecast values from August 2020 to July 2021

Month	Forecast	95% confidence intervals	
		Lower	Higher
August/2020	125,9853	118,0344	134,4716
September/2020	126,2210	115,1048	138,4105
October/2020	126,4571	112,9554	141,5726
November/2020	126,6935	111,2071	144,3367
December/2020	126,9305	109,7137	146,8492
January/2021	127,1680	108,4002	149,1851
February/2021	127,4059	107,2224	151,3887
March/2021	127,6442	106,1514	153,4889
April/2021	127,8830	105,1670	155,5056
May/2021	128,1223	104,2550	157,4533
June/2021	128,3619	103,4045	159,3431
July/2021	128,6021	102,6067	161,1834

Conclusions

The CPI in Mozambique showed an increasing trend in the period analyzed, between January 2011 and July 2020, with a sharp rise in 2016 returning to the original level and the normal growth trend of the series.

In the estimation stage, it was possible to adequately select (02) two ARIMA models (p,d,q) that presented good fit to the data, and among them, the model selected as the most appropriate for predictions according to performance measures, was ARIMA (0.1.0) with constant, because it presented lower values of absolute mean percentage error (AMPS) and mean quadratic error (REQM).

According to the selected model, the previsions for the period from August 2020 to July 2021 indicate a slight growth in consumer price index.

Conflict of Interests

The authors have not declared any conflict of interests.

References

1. Chen, M., & Hu, X. (2018). Linkage between consumer price index and purchasing power parity: Theoretic and empirical study. The Journal of International Trade & Economic Development, 27(7), 729-760.
2. Sherif, M., & Shaairi, N. A. (2013). Determinants of demand on family Takaful in Malaysia. Journal of Islamic Accounting and Business Research.
3. Anghel, M. G., Mirea, M., & Badiu, A. (2018). Analysis of the Main Aspects Regarding the Price Indices Applied in the Determination of Inflation. International Journal of Academic Research in Accounting, Finance and Management Sciences, 8(2), 28-36.
4. McGranahan, L., & Paulson, A. L. (2005). Constructing the Chicago Fed Income Based Economic Index-Consumer Price Index: Inflation Experiences by Demographic Group: 1983-2005, Working Paper 2005-20.
5. Pinheiro, M. C. (2017). Gestão de risco e modelos de VAR: comparação do poder preditivo de mensuração de risco para diferentes classes de ativos.
6. Dong, M., & Sun, J. (2020). Partial discharge detection on aerial covered conductors using time-series decomposition and long short-term memory network. Electric Power Systems Research, 184, 106318.
7. Canizo, M., Triguero, I., Conde, A., & Onieva, E. (2019). Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study. Neurocomputing, 363, 246-260.
8. WAHYUDI, S. T. (2017). The ARIMA Model for the Indonesia Stock Price. International Journal of Economics & Management, 11.
9. Nwankwo, S. C. (2014). Autoregressive integrated moving average (ARIMA) model for exchange rate (Naira to Dollar). Academic Journal of Interdisciplinary Studies, 3(4), 429.
10. Peter, Ď., & Silvia, P. (2012, September). ARIMA vs. ARIMAX–which approach is better to analyze and forecast macroeconomic time series. In Proceedings of 30th international conference mathematical methods in economics (Vol. 2, pp. 136-140).
11. Golyandina, N., & Shlemov, A. (2013). Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. arXiv preprint arXiv:1308.4022..
12. Farmer, D. L., & Stark, R. A. (2016). Should Patients Choose Their Care?. JAMA surgery, 151(5), 415-416.
13. Evans, O., Adeniji, S., Nwaogwugwu, I., Kelikume, I., Dakare, O., & Oke, O. (2018). The relative effect of monetary

and fiscal policy on economic development in Africa: a GMM approach to the St. Louis equation. Business and Economic Quarterly, 2(October), 3-23.
14. Ferreira, J. J., Fayolle, A., Fernandes, C., & Raposo, M. (2017). Effects of Schumpeterian and Kirznerian entrepreneurship on economic growth: Panel data evidence. Entrepreneurship & Regional Development, 29(1-2), 27-50.
15. Mahmoud, L. O. M. (2015). Consumer Price Index and Economic Growth: A Case Study of Mauritania 1990–2013. Asian Journal of Empirical Research, 5(2), 16-23.
16. Martin, C. (1997). Price formation in an open economy: theory and evidence for the United Kingdom, 1951–1991. The Economic Journal, 107(444), 1391-1404.
17. Bradley, R. (2001). Finite sample effects in the estimation of substitution bias in the consumer price index. Journal of Official Statistics, 17(3), 369.
18. Boskin, M. J., Dulberger, E. R., Gordon, R. J., Griliches, Z., & Jorgenson, D. W. (1997). The CPI commission: Findings and recommendations. The American Economic Review, 87(2), 78-83.
19. Aka, B. F., & Pieretti, P. (2008). Consumer Price Index dynamics in a small open economy: A structural time series model for Luxembourg. International Journal of Applied Economics, 5(1), 1-13.
20. Lee, H. J. (2012). Analysis of time series models for consumer price index. Journal of the Korean Data and Information Science Society, 23(3), 535-542.
21. Jackson, E. A., Sillah, A., & Tamuke, E. (2018). Modelling monthly headline consumer price index (HCPI) through seasonal Box-Jenkins methodology. International Journal of sciences, 7(01), 51-56.
22. Furth, S. (2017). Measuring inflation accurately. Backgrounder, (3213).
23. Morettin, P. A., & Toloi, C. M. C. (2006). Análise de séries temporais: Blücher. São Paulo.
24. Vasilev, A. (2015). Modeling real private consumption expenditure in Bulgaria after the currency board implementation (1997-2005). Zagreb International Review of Economics & Business, 18(1), 81-89.
25. Clements, M. P., & Hendry, D. F. (2001). Forecasting non-stationary economic time series. mit Press.
26. RAMOS, A. S. (2010). Previsões de séries temporais combinando modelos ARMA e Redes Neurais Artificiais.
27. Beck, P. S., Atzberger, C., Høgda, K. A., Johansen, B., & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. Remote sensing of Environment, 100(3), 321-334.
28. Mertler, C. A., Vannatta, R. A., & LaVenia, K. N. (2021). Advanced and multivariate statistical methods: Practical application and interpretation. Routledge.