Macrothink Institute™

Journal of Public Administration and Governance
ISSN 2161-7104
2022, Vol. 12, No. 2

Impact of Welfare and Unemployment on the Crimes in the United States: Co-integration Analysis

Mahmoud MOURAD

Professor of Econometrics and Quantitative methods at Lebanese University

Faculty of economic sciences and business administration-branch (5)

Received: April 12, 2022 Accepted: May 23, 2022 Online published: June 9, 2022

doi:10.5296/jpag.v12i2.19957 URL: https://doi.org/10.5296/jpag.v12i2.19957

Abstract

The paper carries out a depth study of the total annual crimes by 100000 people in the United States, considered as the dependent variable, and two explanatory variables that are: welfare expenditure by 100000 people and unemployment rate during the period 1980-1919. The paper is oriented to a specific point: impact the unemployment and welfare variables on the total crimes (violent and property) in the United States of America. Much attention is given to the stationarity analysis and then to the cointegration analysis using the ARDL / Bounds testing Methodology proposed by Pesaran, Shin and Smith (PSS). Two long-term equilibrium relationships were identified and an interpretation of the results was performed. Both relationships have been validated and the Error Correction Model (ECM) is built. The results show that the adjustment speeds towards equilibrium are 11.18 % in the case of Unrestricted intercept and no trend, and 16.95 % in the case of Unrestricted intercept and trend.

Keywords: welfare, unemployment rate, total crimes, descriptive analysis, unit roots, co-integration

1. Introduction

When we study the links between crime rates in the United States of America with a number of socio-economic variables, we achieve very important goals. We first facilitate scientific measurements of the effects of short-term dynamics and the impact on the long-run equilibrium relationship. The Error Correction Model (ECM) strengthens the decision-making process. This study will reveal the short and long run impacts of each explanatory variable on the total crimes. This represents a work of great importance as it highlights the influence of the unemployment rate on crime rates. For example, an increase in unemployment produces some increase in crime rates. Also, a reduction in spending on Welfare contributes to a positive climate of disobedience and poverty. In addition, the transition to the long-term impact of each explanatory variable on the annual number of crimes will be present whether there is agreement on a possible co-integration between welfare, unemployment and annual crime that are available over the period 1980-2019 (40 years). For some clarity, it would be useful to define each of the three variables:

- **Welfare variable** is composed of: Family and children, Unemployment, Unemployment trust, Housing, Social exclusion not elsewhere classified (n.e.c.), Research and development (R & D), and Social protection n.e.c. This variable, designated by (X_{1t}) represents the spending in millions of USD per 100,000 people.

- **United States Unemployment Rate:** Unemployed percentage of the labor force, it is denoted by (X_{2t} **).**

- **Annual number of crimes per 100000 people:** First, *there are* seven types of crimes that are M*urder* (C_{1t})*, Forcible Rape* (C_{2t})*, Robbery* (C_{3t})*, Aggravated Assault* (C_{4t})*, Burglary* (C_{5t})*, Larceny Theft* (C_{6t} *) and Vehicle Theft* (C_{7t})*.* After respecting the definition proposed by the Bureau of Justice Statistics (BJS)[1], three variables will be used: The total number of Violent crimes ($Y_{1t} = C_{1t} + C_{2t} + C_{3t} + C_{4t}$) and the total number of Property crimes ($Y_{2t} = C_{5t} + C_{6t} + C_{7t}$) and the crime "Index" that is defined as the total of all crimes, i.e., the sum of violent and property crimes ($Y_t = Y_{1t} + Y_{2t}$) with more weight given to the property crime. Second, we calculated the annual total of crimes per 100,000 people.

In this paper, the co-integration approach is used, which is of great importance for the decision makers as it can help them to carefully read the future through an adequate policy based on the results confirmed by the proposed model. Indeed, the fundamental point of the co-integration domain is that the Engle-Granger (EG) two-stage procedure was the rapid induction of analytic techniques in econometric for testing common trends in multivariate time series. Then it was the methodology of Johansen that pushed the research to go further in the co-integration analysis. In fact, the two approaches are different for carrying out a long-run equilibrium between the variables. First of all, it may happen that the use of the (EG) technique could lead to a co-integration relationship which is different from that proposed by the Johansen approach. With the (EG) procedure, the starting point is both of variables is integrated in the same order, let us say I(1), we then begin with an estimate of the relationship of co-integration. The work will be completed by the test of the stationarity of the corresponding residues. While with the Johansen methodology, based on the maximum likelihood method for estimate, we can have several co-integration relationships and consequently the co-integration relationship is not unique. Moreover, this research would be very beneficial for all politicians who govern their country in the light of knowledge and in order to improve the quality of life for their peoples particularly to ensure a social well-being for all individuals. We previously announced that the benefit of this research is the achievement of a coherent measure of the short and long-term impacts of each of the two selected variables (X_{1t}), (X_{2t}) on the variable (Y_t). The approach of the ARDL model with co-integration will be used according to the methodology according to the ARDL / Bounds Testing Methodology proposed by (Pesaran, Shin and Smith (PSS), 2001). The advantage of this approach is that it does not impose the same order of integration I(1), and

[1] www.usgovernmentspending.com and U.S. Bureau of Justice Statistics (BJS)

it is applicable on time series which are a mixture of I(0) and I(1) but none of it is integrated of order two, i. e. I(2). In the following, we will focus on the impact of (X_{1t}) and (X_{2t}) on (Y_t) and this will provide us with important economic information that will crystallize through the short and long run relationships among the three variables. Testing for unit root in all variables, three tests will be used. These tests were referred by Augmented Dickey–Fuller (ADF) proposed by (Dickey & Fuller, 1979), (Dickey & Fuller, 1981), PP test carried out by (Phillips & Perron, 1988) and KPSS test built by (Kwiatkowski, Phillips, P., & Shin, 1992).

The research is divided into the following six sections: The first section deals with a general introduction to the topic that allows for a general understanding of the problem and its purpose. In the second section, a literature review is devoted and a panoramic view will be described in third section to *understand the evolution of the three-time series using the graphic illustration and the elementary statistics*. In the fourth section, an in-depth study of stationarity will be performed using the tests of unit roots ADF, PP and KPSS. The (PSS) methodology will be used in the fifth section to elaborate a co-integration analysis between Y_t X_{1t} and X_{2t} and by consequence, a conventional Error Correction Model (ECM) will be performed. Finally, in the sixth section offers a conclusion.

2. Literature Review

The studies of crime and its causes have become an important area of interest for researchers in various countries of the world. Different economic and social aspects were studied. A recent paper for (Mourad & Mourad, 2019) carried out a depth study of the crimes in the United States, considering the impact of several spending variables as health Care, Education, Protection, etc, on the crimes using the ARDL/Bounds testing Methodology proposed by (Pesaran, Shin, & Smith, 2001). For each type of crimes, Violent, Property and Total crimes, a long-term equilibrium relationship was validated. The Error Correction Model (ECM) was built and the adjustment speed towards equilibrium was calculated. An important study carried out by (Mourad M. , 2019,b) dealt with the exchange rate of the EURO against the US dollar (EURO to USD) considering the highest daily value of the exchange rate (HIGH) and the opening daily price (OPEN). The error correction model (ECM) showed high efficiency in forecasts (ex-post forecasts) over a very long run covering three hundred and seventy-three days. The results also revealed that the Chow predictive and cumulative sum (CUSUM) tests roughly supported the parameters stability of the estimated Restricted Error Correction (REC) Model. A study by (Mourad & Hadadah, 2019) is of great importance for the GCC countries as it focuses on the long-term equilibrium relationship between the annual rates of change for the gross national expenditure, the gross domestic product (GDP), and oil prices (dollars per barrel). Using Pedroni’s procedure, (Mourad M. , 2018,a) studied the long run equilibrium between the real GDP in GCC countries and six determinant variables. To estimate the long-run equilibrium individually and aggregately, two different regressions are done. so the group mean panel "Fully Modified Ordinary Least Squares" (FMOLS) and group mean panel "Dynamic Ordinary Least Squares" (DOLS) estimators. A paper carried by (Mourad M. , 2018,b) investigated the impact of exports and imports per capita on the nominal GDP per capita in the top ten economies in the world. The paper performed a rich analysis of the panel unit root tests using, firstly, the most six tests of the first generation assuming the independence of individual time series in panel and, secondly, dealing with cross-sectional dependence using the Sectionally Augmented Dickey-Fuller (CADF) test of the second generation. The findings revealed that all variables are integrated at order one. Finally, two regressions were executed, so the group mean panel Fully Modified OLS (FMOLS) and group mean panel Dynamic OLS (DOLS) estimators. The long-run equilibrium among total crime rates, inequality variable and a set of control variables in China was performed by (Zhu & Zilian, 2017) using the Johansen-Jesulius procedure for co-integration. In this research, the inequality variable is chosen as a main independent variable which was measured by dividing the per capita available income in the urban community to the per capita net income in rural community. They showed a very significant positive relationship i.e. an increase in inequality leads to an increase in crime rates. Using the (ARDL) bounds approach, the real broad money in Pakistan was analysed by (Ghumro, Zaini, & Karim, 2016) examining the short and long run relationships. The existence of the long run relationship among income inequality, crime, poverty and inflation for Pakistan was confirmed by (Ahad, 2016). Using Brazil data, a co-integration analysis based on the Johansen procedure was performed by (Santos & Kassouf, 2013) leading to a long-run relationship among crimes, economic activity, and police performance in São Paulo city. The results of the research done by (Cheong & Wu, 2014) have empirically shown that there is a positive effect of the inequality between intra-provincial regions in China and the crime rates while the education level has revealed a negative effect on crime rates, that is, a high level of education leads to a reduction of the crimes. Following common practice in co-integration studies, the long-run equilibrium between the violent crime rate in USA as a dependent variable and the percentage share of income of the top 10% of income earners (explanatory variables) was tested by (Chintrakarn & Herzer, 2012) considering Pedroni's procedure. The KSS nonlinear unit root test proposed by Kapetanios Shin and Snell (2003) was carried out by (Baharom, Habibullah, & Royfaizal, 2008) to test the unit root of the violent crime of each state in United States and the average of violent crime in United States. Baharom, Habibullah,and Noor (2013) discovered a negative long run relationship between income and crime; positive long run relationship between inflation and crime; unemployment and crime; as well as lending rate and crime. Saridakis (2004) used the Johansen's procedure to estimate in the United States, the dynamic relationships between the overall violent crime, murder, rape and assault as dependent time series and a set of weakly exogenous variables as the overall prison population, alcohol consumption expenditures, duration of unemployment, black males, Gini index and chain-type price index, all variables are taken at the national-level. Kuziemko and Levitt (2004) demonstrated that the prison for the drug traffickers will reflect a reduction in the crimes in United States. Levitt (2004) distinguished between the factors that lead to an increase of the crimes in United States and the factors that decrease the crime rates. Alison (2002) used the log-linear approach to estimate the relationship that links the United States National Crime Rate as a dependent variable and the explanatory variables that are a number of national economic and social characteristics (deterrence and demographic variables).

3. Descriptive Statistical Analysis

In this section, an analysis of each time series will be carried out to understand its evolution during the period 1980-2019 (40 years) exploring the political and economic events that left their traces in the studied variable. The interested in this research will be allowed to focus on the different phases that have marked the variable in question. More clearly, the basic descriptive statistics will be used to measure the basic features of each variable helping to provide an important insight into the extent of change over the period under review. Among these statistics are the Min, the Max, the median, the mean, and the standard deviation $(S.D)$. The coefficient of variation $\left(CV = \frac{S.D}{Mean}\right)$ is also used to measure the relative dispersion, or the relative standard deviation, which is the ratio between the standard deviation and the mean. If the mean is representative of the data, then the CV value will be low. Practically a CV less than 25% may indicate consistency in the data which fluctuates close to the mean. On the other hand, a large CV may indicate variability in the data that deviates from the mean. In addition, it is beneficial to perform a point-by-point cumulated calculation of the mean, the standard deviation and therefore of the CV coefficient. If there is a clear increase or decrease evolution of CV, it is doubtful that there is a mean or standard deviation that varies over time, and the non-stationarity of the time series in question seems to be accepted. Hence the use of appropriate unit root tests which will be the objective of the next section. Also the Compound Annual Growth Rate (CAGR) will be used. It is a specific term for the geometric progression ratio that provides a constant rate of evolution over the time period. For a time series X_t, the CAGR, between the first and the end observations, will be calculated using the following formula: $CAGR(1,T) = \left(\frac{X_T}{X_1}\right)^{\frac{1}{T-1}} - 1$. However, if a break in the trend occurred at some point, the use of CAGR masks an interesting analysis. Indeed, this statistic loses its meaning if the measure concerns the beginning of an increasing trend and the end of a second decreasing trend. Since the welfare variable graph shows two different trends, the first is increasing and covers the period 1980-2010 while the second is decreasing over the period 2010-2019, the CAGR statistic has been divided into two according to the two suggested periods. This division will allow a more rigorous measurement. The graph of the variable Y_t reveals a decreasing linear trend especially after the year 1991 accompanied by an increasing trend in the Welfare variable and this represents an important indicator on the relationship between the two variables. In fact, according to Levitt (2004), the behavior towards crime of people aged 65 and above is completely different from that of young people aged 15 to 19. This is well revealed by an arrest rate " fiftieth the level of 15-19 year-olds ". Levitt cites four factors that caused the decline in crimes in United States since 1991: First, an increase in the number of police. This factor is already stated by Marvell and Moody (1996) by using the Granger causality approach to conclude that an increase in the police force will have a negative impact on the number of crimes. Second, the rising number of prisoners with an important adult incarceration rate. In fact, by 2000, there were more than two million individuals incarcerated according to Kuziemko and Levitt (2004) demonstrating that an increase in drug prisoners reduced the crimes. Third, the setback in drug dealing, and fourth, the legalization of abortion. Notice that the graphs of the raw series (Figure (1)) and all other figures are stored in the appendix. The series X_{1t} associated to the unemployment rate shows an evolution with a cyclical trend of different amplitude. In fact, a succession of trend breaks is well observed. Over the whole period of the study, the minimum and maximum are 3.67% in 2019 and 9.7%

Macrothink Institute™
Journal of Public Administration and Governance
ISSN 2161-7104
2022, Vol. 12, No. 2

in 1982 respectively. Investigating the graph of Welfare variable X_{2t}), we observe that the trend increased sharply in the late 2000s due to the Great Recession. Additionally, the September 11, 2001 terrorist attack on the World Trade Center towers and the Pentagon, which was the deadliest in American history, disrupted the economy and stunted growth, causing growth in social expenditure especially for three consecutive years followed by a decline until the year 2006 then a significant change with increasing exponential trend emerged to reach its maximum in 2010. After this year, a break in the trend is observed giving rise to an exponential decrease until the year 2019. It is notable that over the period 2010-2019, both variables X_{1t} and X_{2t} fall sharply, but at different rates, and this is evident from the value of the CAGR, which is -0.101 for X_{1t} and -0.048 for X_{2t}, i.e. the unemployment rate has declined by almost twice the rate of decline in the welfare variable. For the variables X_{1t} and Y_t, the minimum values correspond to the year 2019 (the end of the series) but for X_{2t}, the minimum was in the year 1980 (the beginning of the period). The maximum values for the variables Y_t, X_{1t} and X_{2t} are observed on the years 1980, 1982 and 2010 respectively. In Table (1), the basic statistics are calculated. The relatively small values of the standard deviations (s.d) permits to calculate the statistic $(Mean \mp 2\ (s.d))$, and by consequence the minimum and maximum values are within this interval for all variables. The values of the coefficient of variation are all above 25% and the very high value (43.5%) for the variable X_{2t} indicates the presence of significant variability over time. This is in accordance with the CAGR values that provides a good information on this situation from its two almost opposite values (5.61% before 2010 against -4.8% after 2010.

Table 1. Descriptive Statistics of the Variables

Period: 1980- 2019								
Variables	**Min**	**Max**	**Median**	**Mean**	**STD.DEV**	**CV**	**CAGR**	
Violent crimes per 100000 people Y_t	2489.30	5949.90	4214.50	4405.76	1107.88	25.15%	**-0.0221**	
United States Unemployment Rate X_{1t}	3.67	9.70	5.80	6.20	1.67	26.96%	trend break	
United States Welfare per 100000 people X_{2t}	42.13	216.83	84.08	101.67	44.20	43.47%	**0.0561**	**-0.0480**

It remains to comment on the accumulated calculation, carried out point by point, of the means, medians and coefficients of variation (Figure 2). The Median and Mean values appear to be very linearly related. This is a good sign and may indicate a normal distribution of the time series in question. The estimation was performed by the Cochrane-Orcutt method, see (Mourad, 2017, p.319). However, an autocorrelation of the errors close to unity was observed. This implies that the transformed equations used in the Cochrane-Orcutt's procedure render the variables in first difference and implicitly a unit root is adopted and consequently the accumulated means and medians are considered integrated of order one $(I(1))$ and therefore the non-stationarity of each variable is due to the variance assumed to be in dependence with time. Generally, if a process is stationary in the wide sense then it fluctuates around a constant mean and consequently the series generated by accumulated means will be highly auto-correlated. For this reason, one must be vigilant of not making a quick decision of the non-stationarity of the series itself. This information encourages the stationarity analysis of the variables Y_t, X_{1t} and X_{2t} to be carried out using the most frequent tests in practice, such as the ADF, PP and KPSS tests. Therefore, the estimation by Cochrane-Orcutt's procedure was applied to these statistics in annual increments and the Durbin-Watson statistic (DW) was considered.

For Y_t time series:

$$\Delta\widehat{Mean} = \underbrace{-26.134}_{-1.75} + \underbrace{0.1311}_{2.58}\Delta Median;\ \ R^2 = 78.86\,\%, DW = 1.115;\ \hat{\rho} = \underbrace{0.81}_{9.05}$$

For X_{1t} time series:

$$\Delta\widehat{Mean} = \underbrace{-0.0594}_{-3.61} + \underbrace{0.1745}_{5.33}\Delta Median;\ \ R^2 = 74.27\,\%, DW = 0.525;\ \hat{\rho} = \underbrace{0.418}_{6.43}$$

For X_{2t} time series:

$$\Delta\widehat{Mean} = \underbrace{1.4794}_{4.35} - \underbrace{0.1274}_{-1.19}\Delta Median;\ R^2 = 52.53\,\%, DW = 1.894;\ \hat{\rho} = \underbrace{0.674}_{5.33}$$

It appears that the autocorrelation of the errors still exists in the models of the increments of mean and median associated to the variables Y_t and X_{1t} as shown by the Durbin-Watson statistics. However, the first error autocorrelation has receded from one and it seems that the residuals are stationary but do not behave like white noise. Figure (2) shows that the mean and median graphs of these two variables are almost similar and therefore it is normal that their cross-correlation is close to unity. The graphs of the accumulated values of CV carried out point by point, reveal very interesting information. For Y_t, $(CV = 9.6\%)$ over the first twenty observations, then an almost linear rise to reach its final value of $(CV = 25.15\,\%)$. The CVs of X_{1t} show fluctuations generated by an almost periodic and increasing trend due to the evolutions of the unemployment rates themselves. Finally, the CVs associated to the variable X_{2t} indicate much greater changes compared to the CVs linked to changes in Y_t and X_{1t}. But how can we interpret this large growth in CV? If the size of the observations used in the calculation of the mean and standard deviation increases, then a strong change occurs in either the mean or the standard deviation. This means that the mean or/and the standard deviation varies with time and consequently there is a strong chance that the studied series is not stationary in level.

4. Unit Root Tests

In the econometric literature, unit root tests started seriously with Dickey and Fuller (1979,81) with an AR(1) model without intercept, with intercept or with intercept and linear trend. To ensure non-autocorrelated residues, the Augmented Dicky-Fuller ADF tests are being developed. The null hypothesis informs about the integration at order (d), noting I(d), of the concerned time series against the alternative that it is I (0). Practically, it is enough to choose the order p for the ADF equation leading errors that behave like white noise. The choice of p requires to fixe previously the maximum order p_{max}. Under the unit root null hypothesis, the asymptotic distribution of ADF statistics are not the standard t −distribution and so the conventional critical values are proposed. According to Phillips and Perron (1988) procedure (PP), the null and alternative hypotheses are maintained as in the ADF case and the critical values remain those tabulated by Dicky-Fuller (1979,1981- Tables I-III). The strategy PP has not adopted the ADF path and thus it does not impose restrictions on the residuals, which can be auto correlated and even heteroscedastic. A remarkable advantage of the Phillips-Perron test is its non-parametric nature. An inconvenience of PP test arises from the asymptotic theory itself which makes its poor power in the case of small samples. In this context, (Phillips, 1987a) treated a model without intercept imposing four conditions on the residuals, which allowed him to obtain the Weighted Variance Estimator, whose value changes with the lag truncation number (ℓ) having the expression performed by Schwert (1989) and (Newey & West, 1987) that is $\ell = 12\left(\frac{T}{100}\right)^{\frac{1}{4}}$ or $4\left(\frac{T}{100}\right)^{\frac{1}{4}}$, so by interpolation we choose, $(p_{max} = 6)$ for ADF test and $(0 \leq \ell \leq 12)$ for PP and KPSS tests. To complete what Phillips had started, the PP procedure was able to generalise this to both intercept and linear trend cases. Thus, the unit root test was completed in the context of a new strategy. Unlike other methods used to test the stationarity of time series, the KPSS method aims to test the null hypothesis that the associated time series is stationary against the alternative hypothesis that it is not. According to the KPSS test, a time series can be decomposed into three components which are deterministic trend, a random walk and a stationary error. Two cases will be considered. The first case considers the null hypothesis that the time series is stationary around a linear trend $\left(trend - stationary \Rightarrow \hat{\eta}_\tau\right)$. The second case consists in assuming the null hypothesis that the time series is stationary around a level $\left(stationary\ in\ level \Rightarrow \hat{\eta}_\mu\right)$. In the two cases, the lag truncation number is used to estimate the long-run variance based on the partial sum process of the residuals due of the regression of the dependent on an intercept and time trend or on an intercept only. The upper tail critical values for $\hat{\eta}_\mu$ and $\hat{\eta}_\tau$ are given at (10%, 5%, 2.5% and 1%)[2].

- **Dickey-Fuller Unit Root Tests:** Let's start with the Dickey-fuller method to decide about the stationarity of our variables. After estimating three AR(6) models with intercept, we calculated the first twelve autocorrelations of the residuals of both models. The Ljung-Box (LB) statistics confirmed the validity of the residuals as a white noise because for the variables Y_t , X_{1t} and X_{2t} and for degrees of freedom $(1 \leq df \leq 12)$, the significance levels belong to the intervals $[0.231; 0.823]$, $[0.510; 0.994]$ and $[0.736; 0.996]$ respectively. Figure (3) reinforces the idea of the adequacy of the proposed AR (6) models since the autocorrelation errors lie within the Bartlett interval, i.e. the null hypothesis $(H_0: \rho_h = 0, h = 1, \dots 12)$ is accepted at the

[2] Table (1) in (Kwiatkowski et al. 1992).

5% level. After fixing the maximum lag $(p_{max} = 6)$, the $(\text{optimal lag} = p_{opt})$ must then be determined. The (p_{opt}) announced by the AIC criterion will undoubtedly affect the $(t_{\hat{\psi}})$ statistic in the $(ADF_i, i = 1,2,3)$ equations. We will now try to detect the change in the statistic $(t_{\hat{\psi}})$ when the value of (p_{max}) changes in the equations. We put in Table (2) the results related to the optimal order associated to each one of the three equations $(ADF_i, i = 1,2,3)$, which we symbolized this option with $(i\ from\ j; i \leq j)$, i.e. (p_{opt}) from (p_{max}). To facilitate the reading of the information required about this, it is sufficient to see the figure (4). As we have seen, we will reject the null hypothesis versus the alternative hypothesis $(H_0: \psi = 0\ \ vs\ H_a: \psi < 0)$ if $(t_{\hat{\psi}} \leq \tau_\mu)$. For the variable $\mathbf{Y_t}$, the τ_τ statistic confirms the rejection of the null hypothesis and it appears that the variable $\mathbf{Y_t}$ is stationary around an intercept or around a linear trend. To facilitate the ADF's comments, the associated graphs were drawn for each variable (Figures 4,5) and the graphs in each row are respectively tuned to the variables Y_t, X_{1t} and X_{2t}. Two strategies are adopted to inspect all the information: The first one called "DBM" was inspired from (Dickey, Bell, & Miller, 1986) while the second one (DJSR) was carried out by (Dolado, Jenkinson, & Sosvilla-Rivero, 1990). Following the organigram of each strategy, we arrive at the following conclusions: The DBM strategy suggests series TS, I(0), TS for Y_t, X_{1t} and X_{2t} respectively while we retain the situation I(1), I(0) and I(1) according to the strategy (DJSR).

- **Phillips-Perron Unit Root Tests:** To calculate the statistics associated to PP_2 (Model with intercept) and PP_3 (Model with trend), it is necessary to determine the lag truncation number (ℓ) used in the estimation of the long-run variance of the errors which has been chosen in the domain $(0 \leq \ell \leq 6)$. For each value of (ℓ), the weighted variance estimator of errors was done and the statistics resulting from PP_2 and PP_3 are obtained and the results are presented in graphs. Figure (6) shows clearly that the null hypothesis of the unit root is not rejected for each of the three variables. On the other hand, for the variables in first differences, figure (7) supports their stationarity especially around an intercept. There is stability in the rejection of the null hypothesis for the lag truncation number $(\ell \geq 6)$ especially for ΔX_{1t} and ΔX_{2t}, perhaps the value $(\ell = 6)$ agrees with $(p_{max} = 6)$ in the ADF cases.

- **KPSS stationarity test:** We will now use the KPSS technique to test the stationarity of our variables. Unlike most unit root tests, the null hypothesis suggests stationarity around a linear trend $(\hat{\eta}_\tau\ \text{statistic})$or around a constant $(\hat{\eta}_\mu\ \text{statistic})$. Just as it is possible for a DF or PP tests not to reject the unit root null hypothesis, it is also possible for a KPSS test not to reject the second order stationarity. The decision moves towards rejecting the null hypothesis, at the significance level (5%), if we get $\hat{\eta}_\mu$ greater than the specified critical value. Similarly we progress with the $\hat{\eta}_\tau$. As we did with the PP test, we need to specify the lag truncation number using Bartlett window, so we will choose $(0 \le \ell \le 6)$ which will allow a comparison between the two tests.

The outcome of the tests reveals the following (Figure 8): for the unemployment rate series, we cannot reject the null hypothesis of level or trend stationarity at usual critical levels with a decision that favors more the $I(0)$ series. However, for the two series, total crimes and welfare expenditures, we find that we can reject the hypothesis of level stationarity ($\hat{\eta}_\mu$ is very close to the critical value at level 5%) but the hypothesis of trend stationarity (TS series) is accepted as soon as $(\ell \ge 3)$. We find the same decision due to the DBM strategy which suggested TS, I(0), TS for Y_t, X_{1t} and X_{2t} respectively. Finally, the important information about this section of the unit root tests is that no variables need a second difference to become stationary. Regardless if the order of integration I(0) or I(1), we can use the procedure of co-integration according to the ARDL / Bounds Testing Methodology proposed by Pesaran, Shin and Smith (PSS) (2001).

Table 2. Dickey-Fuller Unit Root Tests

Variables	p-1	Lags	τ	Lags	τ_μ	Lags	τ_τ
Y_t	0	0 from 0	-3.744**	0 from 0	0.317	0 from 0	-1.338
	1	1 from 1	-1.720	1 from 1	-0.372	1 from 1	-2.726
	2	2 from 2	-1.987*	2 from 2	0.354	2 from 2	-2.275
	3	2 from 3	-1.987*	2 from 3	0.354	1 from 3	-2.726
	4	2 from 4	-1.987*	2 from 4	0.354	4 from 4	-5.374**
	5	1 from 5	-1.720	1 from 5	-0.372	5 from 5	-5.087**
X_{1t}	0	0 from 0	-0.938	0 from 0	-1.524	0 from 0	-1.734
	1	1 from 1	-1.138	1 from 1	-3.089*	1 from 1	-3.245
	2	2 from 2	-1.601	1 from 2	-3.089*	1 from 2	-3.245
	3	2 from 3	-1.601	1 from 3	-3.089*	3 from 3	-2.901
	4	2 from 4	-1.601	1 from 4	-3.089*	1 from 4	-3.245
	5	2 from 5	-1.601	1 from 5	-3.089*	1 from 5	-3.245
X_{2t}	0	0 from 0	0.604	0 from 0	-1.220	0 from 0	-1.622
	1	1 from 1	-0.075	1 from 1	-1.664	1 from 1	-3.551*
	2	2 from 2	0.318	2 from 2	-1.314	1 from 2	-3.551*
	3	2 from 3	0.318	2 from 3	-1.314	1 from 3	-3.551*
	4	2 from 4	0.318	2 from 4	-1.314	1 from 4	-3.551*
	5	2 from 5	0.318	2 from 5	-1.314	1 from 5	-3.551*

Note: 5%(*) and 1%(**)

5. Conventional Error Correction Model (ECM) and Empirical Results

Before carrying out the empirical study, and to ensure clarity about the implementation of the PSS procedure, we briefly mention the following steps as quoted by Mourad (2019a, p.65-71) to arrive at the estimate of the ARDL model:

Step 1: Based on the unit-root analysis, assuming that all the variables are $I(0)$ or $I(1)$ and without variables integrated of order two in the model. The findings of the unit-root study respond well to this step.

Step 2: Since there are no constraints on the parameters associated with the variables in level, the Unrestricted Error Correction Model (UECM) will be considered:

$$\Delta Y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \Delta Y_{t-j} + \sum_{j=0}^{q_1} \alpha_{1j} \Delta X_{1,t-j} + \sum_{j=0}^{q_2} \alpha_{2j} \Delta X_{2,t-j} + \beta_1 Y_{t-1} + \beta_2 X_{1,t-1} + \beta_3 X_{2,t-1} + \varepsilon_t \qquad (1)$$

Using an automatic criterion such as the AIC or BIC per example, the orders (p) and ($q_i, i = 1,2$) will be determined taking into account that the residues are not correlated according to Ljung-Box statistic. For this purpose, we first determined the optimal order for p because there is a strong autocorrelation between the past and the present of the dependent variable in question. Determining the order of each of the explanatory variables will be done sequentially by introducing the variable that is most correlated with the dependent variable and we retain only the significant parameters.

Step 3: The long-run equilibrium relationship between the variables in levels will be tested:

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0 \quad \text{no long} - \text{run relationship}$$
$$H_1: \beta_1 \neq \beta_2 \neq \beta_3 \neq 0 \quad \text{long} - \text{run relationship}$$

The bounds testing statistic that is symbolized by $F_Y(Y|X_1, X_2)$ is calculated. This statistic is subject to a non-standard distribution related by the order integration I (1) or I (0), by the number of estimated parameters in the model and by the presence or absence of constraints on the intercept and trend. According to the PSS procedure, there are two sets of critical values: The first set is associated to the minimum values assuming all variables are $I(0)$ and, therefore, there is no co-integration. The second set is composed of large values and assuming all variables are $I(1)$ and, therefore, there is co-integration. If the calculated F is outside of the bound specified by PSS procedure, then three findings can be concluded:

- If the calculated F is greater than the upper critical value of the bound, then co-integration will be accepted.
- If the calculated F is smaller than the minimum critical value of the bound, then co-integration will be rejected.
- Finally, if the F statistic is included in the bound, then the decision will be inconclusive, indicating that the F statistic depends on the order of integration I (0) or I (1).

Step 4: After determining the optimal values $(p, q_i, i = 1,2)$, the long-run linear relationship between the variables at levels will be estimated: $Y_t = \theta_0 + \sum_{j=1}^{2} \theta_j X_{j,t-1} + \eta_t$ and the then we maintain the estimated residues $\hat{\eta}_t$ symbolized by EC_t:

$$EC_t = Y_t - \hat{\theta}_0 - \sum_{j=1}^{2} \hat{\theta}_j X_{j,t-1} \qquad (2)$$

Step 5: The traditional error correction model:

$$\Delta Y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \Delta Y_{t-j} + \sum_{j=0}^{q_1} \alpha_{1j} X_{1,t-j} + \sum_{j=0}^{q_2} \alpha_{2j} X_{2,t-j} + \gamma EC_{t-1} + \varepsilon_t \qquad (3)$$

The parameter γ indicates the speed of adjustment to restore equilibrium if a deviation from equilibrium is produced. The sign of γ must be negative and with significance level 5% according to the Student distribution in order to ensure the dynamic adjustment towards equilibrium. In general, $-1 < \gamma < 0$, if its estimated value is close to (-1.0) then the return to equilibrium will be almost complete and immediate. In the final estimate of the Error Correction Model (ECM), the parameters with low significance $(|t| < 1.64)$ were removed. According to the step one, all variables are integrated I(0) or I(1). The step two was conducted in the proposed orders of the three automatic criteria that are presented in Table (3). After determining the optimal order for each variable, we retain only the parameters that are significantly different from zero. In Table (4), the results of the step three are presented and the null hypothesis of no co-integration is rejected at 5 % level of significance. The long-run linear relationship between the variables will be estimated and then we maintain the estimated residues $\hat{\eta}_t$ symbolized by (EC). The two long-run linear relationship are the following:

- **Case 1: Unrestricted intercept, no trend**

$$EC_{t-1} = Y_{t-1} - \left(\underbrace{5165.833}_{16.39} + \underbrace{242.4891}_{5.41} X_{1,t-1} \underbrace{-22.2613}_{-13.13} X_{2,t-1}\right) \; ; \; R^2 = 83.38\,\%$$

The ECM_1 model :

$$\Delta Y_t = \underbrace{-29.9677}_{-1.56} + \underbrace{0.6794}_{6.04} \Delta Y_{t-1} \underbrace{-59.6492}_{-2.44} \Delta X_{1,t-1} + \underbrace{32.8447}_{2.05} \Delta X_{1,t-3} + \underbrace{3.7408}_{2.30} \Delta X_{2,t-1} \underbrace{-0.1118}_{-3.18} EC_{t-1}$$

Since the coefficient of EC_{t-1} is sufficiently negative, the ECM_1 model is accepted, so it is composed of two parts. the first part describes the short-run impact of change in Y at lag (1), X_1 at lags (1,3) and X_2 at lag (1). The second part explains the long-run gravitation towards the equilibrium relationship between the variables. It is useful to know how the ECM model works: Suppose in period $t-1$ the system is in equilibrium, i.e. $EC_{t-1} = 0$. Assume that $X_{1,t}X_{1,t}$ increases by one and the other variables remain unchanged then $X_{1,t}$ returns to its previous level. Then $Y_t Y_t$ first (in period t) decreases by (59.65), but after the second period $Y_t Y_t$ begins to increase and converges to its initial level.

- **Case 2: Unrestricted intercept, Unrestricted trend:**

The ECM_2 model :

$$\Delta Y_t = \underbrace{111.514}_{2.156} \underbrace{-5.0372}_{-2.92} t + \underbrace{0.5323}_{4.20} \Delta Y_{t-1} + \underbrace{0.2061}_{1.72} \Delta Y_{t-4} \underbrace{-0.1695}_{-3.84} EC_{t-1}$$

The variable $Y_t Y_t$ doesn't suffer any short-run impact of change in $X_{1,t}X_{1,t}$ and $X_{2,t}X_{2,t}$. However, the appearance of a linear trend with a negative slope indicates that time is having a negative impact, i.e. there will be a decrease in total crimes from one year to the next. The impact of the variables $X_{1,t}X_{1,t}$ and $X_{2,t}$will only occur in the long term and the coefficient (-0.1695) of the EC_{t-1} confirms this situation.

Suppose that in the period t, $X_{1,t}X_{1,t}$ increases by one and the other variables remain unchanged then Y_tY_t first (in period t) stays the same (impact zero). But after the second period Y_tY_t increases by$[-0.1695(-5.0372-242.4891\times 1)=41.96]$ and it begins converge to its initial level.

Table 3. Sequential determination of the optimal orders of the ARDL model

Unrestricted Error Correction Model (UECM) **Dependent variable: Total crimes Y_t**									
Unrestricted intercept, no trend									
Variables	AIC	BIC	HQ	Optimal orders	Variables	AIC	BIC	HQ	Optimal orders
Y_t	3	3	3	3	Y_t	3	3	3	3
Y_t, X_{1t}	1	1	1	3 1	Y_t, X_{2t}	1	1	1	3 1
Y_t, X_{1t}, X_{2t}	1	1	1	3 1 1	Y_t, X_{2t}, X_{1t}	3	1	3	3 1 3
Sum of Squared Residuals: $SSR=191419.578$, $\bar{R}^2=58.76\%$, $R_U^2=0.6701$, $R_R^2=0.4835$ Durbin – Watson Statistic: $DW=2.2607$ $F(3,28)=\frac{(R_U^2-R_R^2)}{3}\times\frac{df}{1-R_U^2}=5.279$ with Significance Level 0.0052 **Note**: number of independent variables $(k=2)$.									
	Unrestricted intercept, Unrestricted trend								
Variables	AIC	BIC	HQ	Optimal orders	Variables	AIC	BIC	HQ	Optimal orders
Y_t	5	4	4	4	Y_t	5	4	4	4
Y_t, X_{1t}	1	1	1	4 1	Y_t, X_{2t}	1	1	1	4 1
Y_t, X_{1t}, X_{2t}	1	1	1	4 1 1	Y_t, X_{2t}, X_{1t}	1	1	1	4 1 1
Sum of Squared Residuals: $SSR=144922.878$, $\bar{R}^2=69.41\%$, $R_U^2=0.7481$, $R_R^2=0.3870$ Durbin – Watson Statistic: $DW=2.2684$ $F(3,28)=\frac{(R_U^2-R_R^2)}{3}\times\frac{df}{1-R_U^2}=13.383$ with Significance Level 0.0000 **Note**: number of independent variables $(k=2)$.									

Table 4. Estimate of UECM Model

ARDL(1,3,1) **Unrestricted intercept no trend**			**ARDL(4,0,0)** **Unrestricted intercept Unrestricted trend**		
Variables	Coefficients	T-stats	Variables	Coefficients	T-stats
Constant	595.6750	2.93	Constant	2021.6061	5.75
ΔY_{t-1}	0.6565	6.04	Trend	-33.0589	-4.22
$\Delta X_{1,t-1}$	-52.8827	-2.19	ΔY_{t-1}	0.4900	4.74
$\Delta X_{1,t-3}$	35.1385	1.89	ΔY_{t-4}	0.2942	2.96
$\Delta X_{2,t-1}$	3.4218	2.18	Y_{t-1}	-0.2989	-6.27
Y_{t-1}	-0.1160	-3.42	$X_{1,t-1}$	-3.7830	-0.16

$X_{1,t-1}$	36.8865	2.76		$X_{2,t-1}$	0.3922	0.26	
$X_{2,t-1}$	-3.3425	-3.90					
F-statistic	Significance level	Bound critical values Intercept and trend		F-statistic	Significance level	Bound critical values Intercept and trend	
		I(0)	I(1)			I(0)	I(1)
5.279	1% **5%** 10%	5.15 **3.79** 3.17	6.36 **4.85** 4.14	13.383	1% **5%** 10%	6.34 **4.87** 4.19	7.52 **5.85** 5.06
Decision : $F_Y(Y\|X_1,X_2) = 5.279 > 4.85$ The null hypothesis of no cointegration is rejected				**Decision** : $F_Y(Y\|X_1,X_2) = 13.383 > 4.87$ The null hypothesis of no cointegration is rejected			

6. Conclusion

We have performed a global analysis of stationarity by using elementary cumulative statistics such as means, medians, standard deviations and coefficients of variation, and three fundamental techniques of unit root tests or stationarity, ADF, PP and KPSS. The unemployment rate series appears to be stationary but the other two series, welfare expenditures and total crimes, are not stationary, and the tests reveal TS or I(1) behaviour for both of them. It is certain that the three series do not need to make a secondary difference to ensure stationarity. So the PSS technique is allowed to develop an Error Correction Model with two versions. The first deals with unrestricted intercept and no trend and the second involves unrestricted intercept and unrestricted trend. The long term relationships among the three variables can be interpreted using the first case. Indeed, if the unemployment rate increases by one and the welfare variable remains unchanged, then the number of crimes by 100000 people increases by about 243. This positive long run relationship between unemployment and crime has also been found by (Baharom et al. 2013). On the other hand, if the welfare expenditures per 100,000 people increase by one unit (one million USD) and the variable X_{1t} remains unchanged then the number of crimes decreases by 23. Focusing on the two long-run equilibrium relationships, if a deviation is made in the previous year then the system will adjust itself toward equilibrium by $11.18\ \%$ one year later for the first version and $16.95\ \%$ for the second. The estimated ECM will be facilitate a scientific measurement of the effect of short-term and the impact on the long-term of the unemployment rate and welfare variables in United States on the total annual crimes. This estimated ECM can help decision-makers with their decision-making tasks.

Recommendations

To reach our goal about a long-run equilibrium relationship between the dependent and explanatory variables, Pedroni’s methodology will be used and panel co-integration test will be performed considering the 51 states from Alabama to Wyoming.

References

Ahad, M. (2016). Nexus between income inequality, crime, inflation and poverty: new evidence from structural breaks for Pakistan. *International Journal of Economics and Empirical Research, 4*(3), 133-145.

Alison, O. (2002). The Economics of Crime: An Analysis of Crime Rates in America. *The Park Place Economist, 10*, 30-35.

Baharom, A. H., Habibullah, M. S., & Noor, Z. M. (2013). Crime and Its Socio-Macroeconomics Determinants: A Panel-Error-Correction. *Jurnal Ekonomi Malaysia, 47*(2), 13-24.

Baharom, A. H., Habibullah, M. S., & Royfaizal, R. C. (2008). Convergence of violent crime in the United States: Time series test of nonlinear. *MPRA Paper 11926, University Library of Munich, Germany*.

Cheong, T., & Wu, Y. (2014). Crime Rates and Inequality: A Study of Crime in Contemporary China. *Journal of the Asia Pacific Economy, 20*(2), 1-24.

Chintrakarn, P., & Herzer, D. (2012). More inequality, more crime? A panel co-integration analysis for the United States. *Economics Letters, 116*(3), 389-391.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association, 74*(366), 427-431.

Dickey, D., & Fuller, W. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica , 49*(4), 1057-1072.

Dickey, D., Bell, W., & Miller, R. (1986). Unit Roots in Time Series Models: Tests and Implications. *The American Statistician, 40*(1), 12-26.

Dolado, J., Jenkinson, T., & Sosvilla-Rivero, S. (1990). Cointegration and unit roots. *Journal of Economic Surveys, 4*, 249-273.

Ghumro, N., Zaini, M., & Karim, A. (2016). The effects of exchange Rate on Money demand: Evidence from Pakistan. *International Research Journal of Social Sciences, 5*(4), 11-20.

Habibullah, M. S., & Baharom, A. H. (2008). Crime and economic conditions in Malaysia: an ARDL bounds testing approach. *MPRA Paper, 11910*.

Kapetaniosa, G., Shinb, Y., & Snellb, A. (2003). Testingfor a unit root in the nonlinear STAR framework. *Journal of Econometrics, 112*, 359-379.

Kuziemko, I., & Levitt, S. (2004). An Empirical Analysis of Imprisoning Drug Offenders. *Journal of Public Economics, 88*, 2043-2066.

Kwiatkowski, D., Phillips, P., P., S., & Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics, 54*, 91-115.

Levitt, S. D. (2004). Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not. *Journal of Economic Perspectives, 18*(1), 163-190.

Marvell, T., & Moody, C. (1996). Specification Problems, Police Levels, and Crime Rates. *Criminology, 34*, 609-646.

Mourad, M. (2017). *Econometric from Theory to Practice. The Lebanese University, Department of Economic Studies, Distributed by the Lebanese University , Central Administration, Museum,* (Vol. 1). Beirut, Lebanon.: Lebanese University.

Mourad, M. (2018,a). Vital Economic Determinants and Real GDP in GCC Countries: Panel Cointegration Analysis. *Review of Economics and Business Administration, 2*(2), 141-168.

Mourad, M. (2018,b). Trade and GDP in World's Top Economies: Panel Co-integration Analysis. *Arabian Journal of Business and Management Review, 8*(5), 1-8.

Mourad, M. (2019 a, p. 65-71). *Econometric from Theory to Practice. Distributed by the Lebanese University , Central Administration, Museum,* (Vol. 2). Beyrouth: Lebanese University.

Mourad, M. (2019,b). Daily Exchange Rate EURO/USD: ARDL Methodology to Co-Integration. *Journal of Business & Financial Affairs, 8*(1), 1-6.

Mourad, M., & Hadadah, A. (2019). Impact of Oil Prices and GDP on National Expenditure in the GCC Countries:ARDL Technique for Co-integration. *Arabian Journal of Business and Management Review, 9*(2), 1-15.

Mourad, M., & Mourad, F. (2019). Socioeconomic Variables and Crimes: Co-integration Analysis for the United States. *Journal of Business & Financial Affairs (JBFA), 8*(2), 1-8.

Newey, W. K., & West, K. D. (1987). A Simple Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica, 55*, 703-708. Princeton Discussion Paper No. 92, 1985.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics, 16*(3), 289-326.

Phillips, P. (1987a). Time Series Regression with a Unit Root. *Econometrica, 55*, 277-301.

Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335-46.

Santos, M., & Kassouf, A. (2013). A cointegration analysis of crime, economic activity, and police performance in São Paulo city. *Journal of Applied Statistics, 40*(10), 2087-2109.

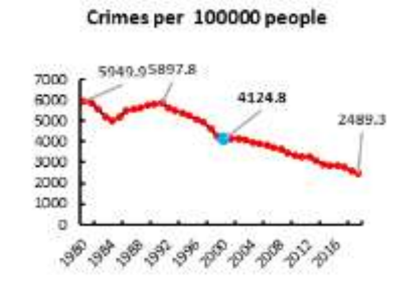
Saridakis, G. (2004). Violent Crime in the United States of America: A Time-Series Analysis Between 1960–2000. *European Journal of Law and Economics, 18*, 203-221.

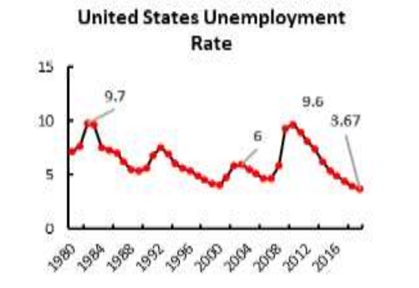
Schwert, G. W. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics, 7*, 147-160.

Zhu, J., & Zilian, L. (2017). Inequality and crime rates in china. *Frontiers of Economics in China, 12*(2), 309–339.

Appendix

Figure 1: **graphs of variables in level**





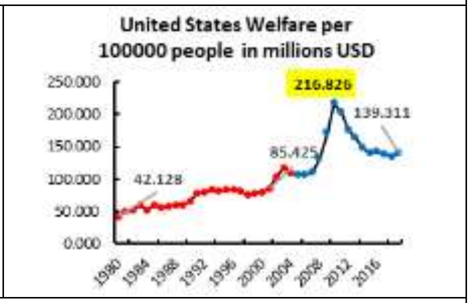
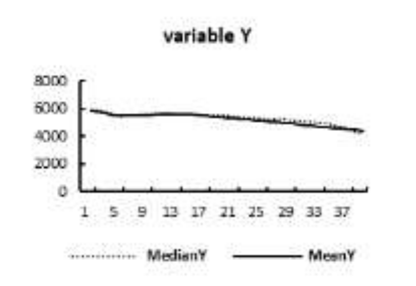
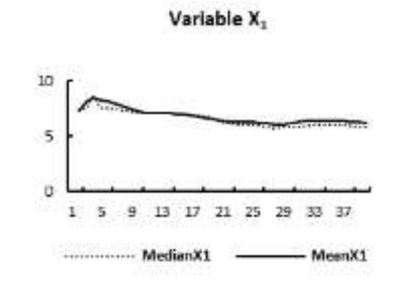
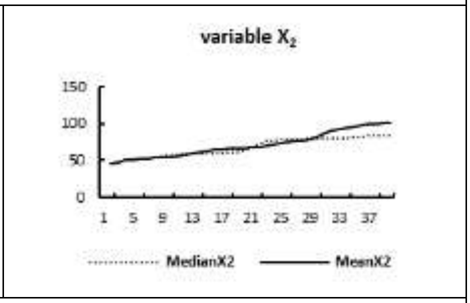


Figure 2: **Accumulated calculation carried out point by point**

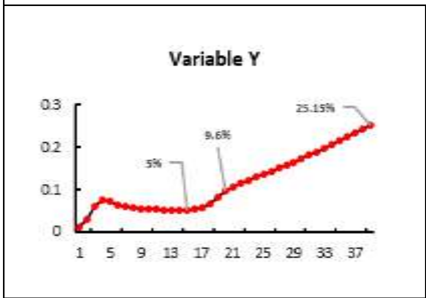
Evolution of the Mean and Median

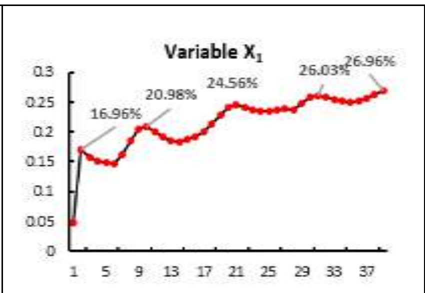






Evolution of CV





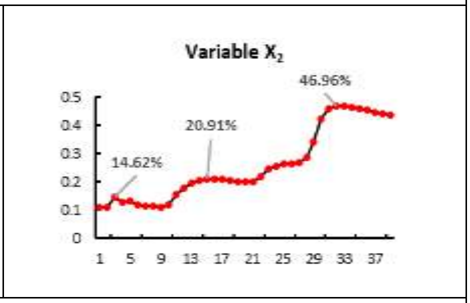
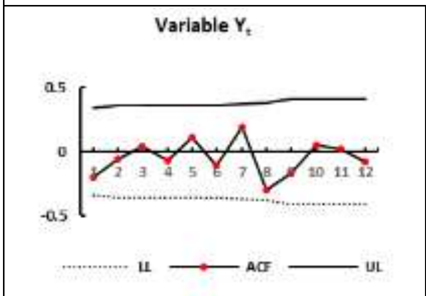
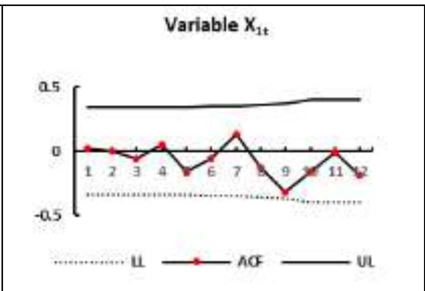


Figure 3: ACF for residues





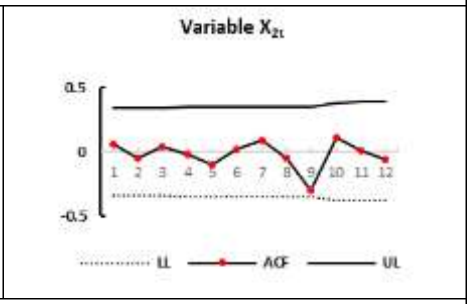
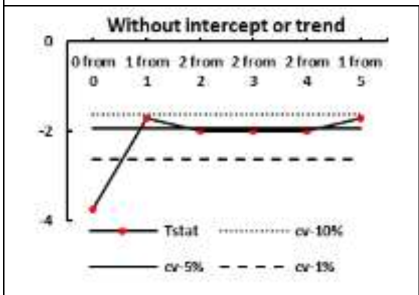
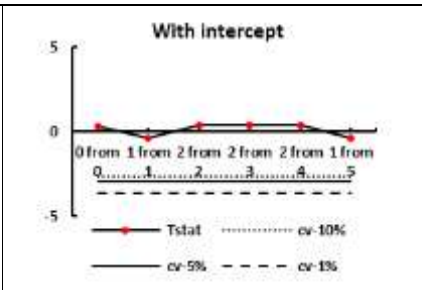
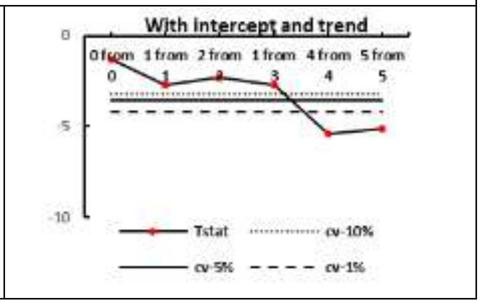


Figure 4: Dickey-Fuller Test responses to AIC criterion: Variables Y, X_1 and X_2







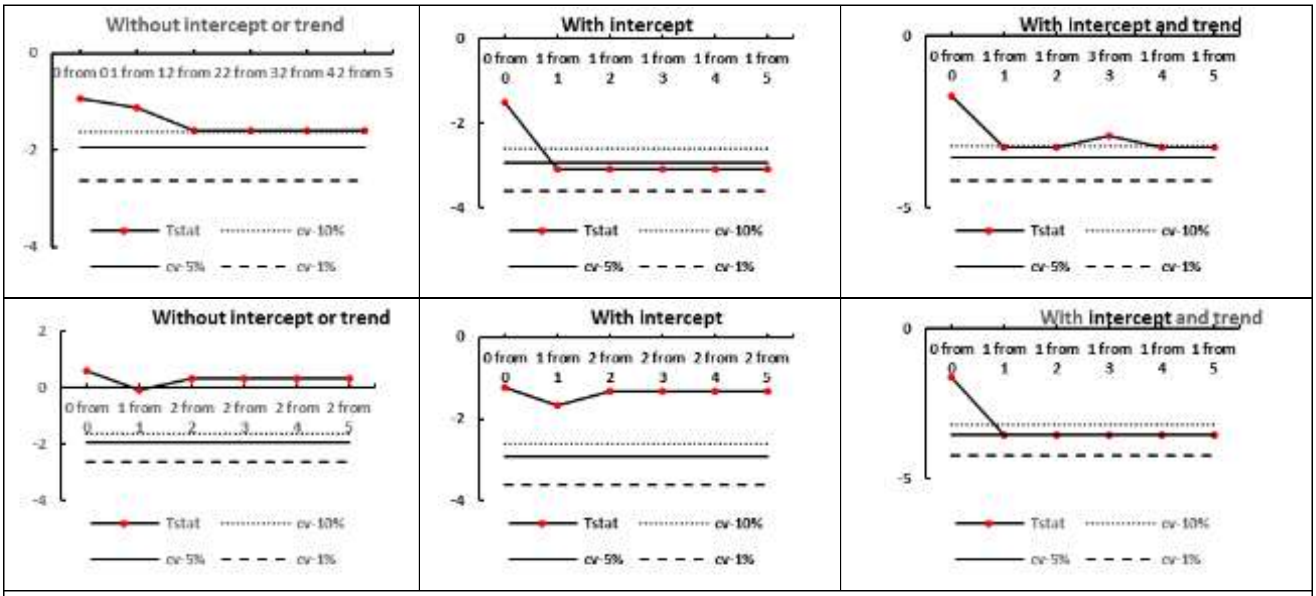


Figure 5: Dickey-Fuller Test responses to AIC criterion: Variables Y, X_1 and X_2 in first difference

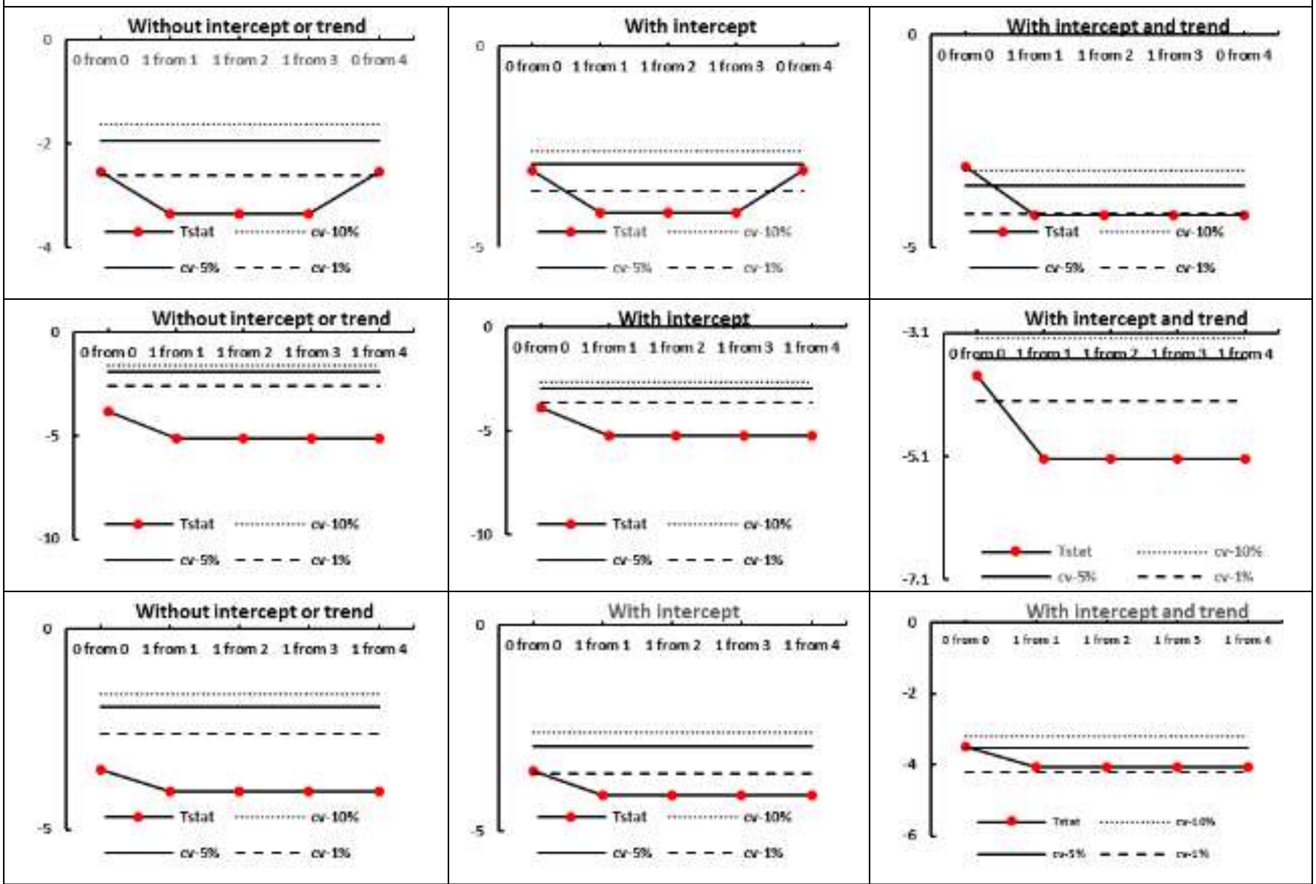


Figure 6: Phillips-Perron Unit Root Tests for the variables Y, X_1 and X_2 in level

Figure 7: Phillips-Perron Unit Root Tests for the variables Y, X_1 and X_2 in first difference

Journal of Public Administration and Governance
ISSN 2161-7104
2022, Vol. 12, No. 2

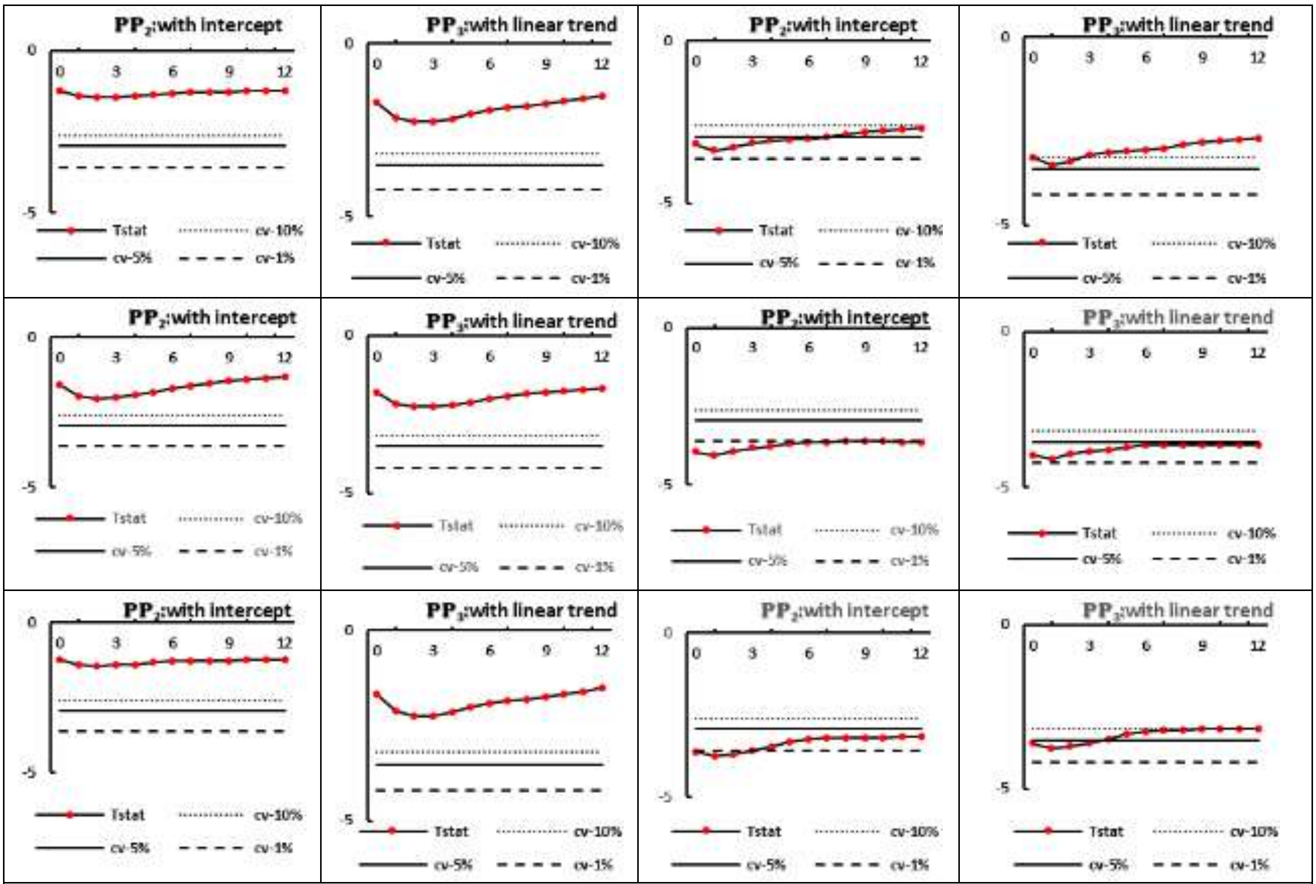


Figure 8: KPSS stationarity Tests for the Variables Y, X_1 and X_2 in level by column