Escuela Internacional de Ciencias Económicas y Administrativas

Economía y Finanzas Internacionales

Trabajo de grado para optar por el título de:
Economista con Énfasis en Finanzas Internacionales

A Cointegration Analysis of the Okun Law in Colombia: An Approach using Non-linear Smooth Transition Autoregressive Models

Presented by: Saulo Samuel Linares Vázquez & Andrés Felipe Roncancio Rodríguez

Thesis Director:
Martha Alicia Misas Arango

Chía, May 2021

A Cointegration Analysis of the Okun Law in Colombia: An Approach using Non-linear Smooth Transition Autoregressive Models

Saulo Samuel Linares Vázquez & Andrés Felipe Roncancio Rodríguez

May 18 2021

Acknowledgments

We thank our director Martha Alicia Misas Arango for the unconditional commitment and valuable contributions, as well as the moral support throughout this work. We are grateful with our families for all the sacrifice they have made throughout this educational process. Also, we are thankful for the useful comments and suggestions from the professors Juan Felipe Bernal and Jaime Tenjo in the preparation process for this research. Thanks to Geovanni Portilla and Daniel Mayor for their helpful guidance in the consolidation of the data. Finally, we manifest our apreciation to Mauricio Senior and Tatiana Rodriguez for their valuable revision regarding the redaction in english of the present document.

Abstract

The intended research estimated the Okun's relationship for Colombia with a STAR procedure, in which the variable capable of enabling the transition was the long-run disequilibrium term. The study postulates two different regimes that characterize Colombian economy with the nonlinearity approach of the Okun's Law: A High regime and a Low regime. The periods in which each state of nature took place where determined and the economic conjuncture that lead to them was explored. The results suggested an improvement of the non-linear analysis of Okun's Law, as well as that employment indeed is linked with economic growth. This could contribute the construction of a correct design, implementation and effectiveness of future policies in the country.

Key Words: Okun's Law, Smooth Transition Regression, Cointegration, Disequilibrium, Error Correction Mechanism model, States of Nature (Regimes).

Jel Classification: C24, C01, E24, E23

1. Introduction

Currently, Colombia is facing its highest unemployment rates in decades and the worst economic growth figure. Even when the Covid-19 pandemic is enough justification for this drastic change, Colombia had suffered for several years an historical unemployment problem. For this reason, it is relevant to determine which factors have an impact in the relationship between economic growth and unemployment in normal circumstances, in order to create well-structured policies that counter the current conjuncture. From 2002 to 2019, Unemployment Rate (UR) in Colombia reflected a mean of 11%. In fact, changes between quarters in this rate did not reach 3% and had a mean of 0%[1]. This reflects a clear stagnation of the country around the problem of unemployment. Despite the improvement in the overall tendency of the series in the more recent years, the government has not managed to consistently maintain the UR under two digits. Later, from 2014 to 2018, for at least eight months, UR was maintained under 10%. Nevertheless, in all these periods only in one quarter UR reached 8% -Quarter 4 of 2015-, and the series bounced back to all quarters in 2019 being over 10%. On top of that, the pandemic has reverted all efforts to overcome this unsettled unemployment problematic.

In order to verify the cost of the unemployment problem for this economy, it is of great importance to determine the magnitude of the relationship between the change in the UR and economic growth. Acknowledging, accelerating product growth is a priority for the country at the moment and this analysis might help the construction of a single policy that could contribute to both objectives. This relationship will be analyzed through Okun's Law.

Okun's Law -OL- was initially defined by Arthur Okun back in 1962, where the author in his research described a linear relationship between unemployment and economic growth.[2] This relationship implies a link between the labor and goods market growth, and it is often seen as a key empirical regularity. A compelling critique of Okun's relationship founded in the literature is based on the linearity assumption. Indeed, many studies suggest that the relationship is characterized by non-linearities and asymmetries. In a relationship as central and profoundly ingrained as the Okun Law, if the possibility to contemplate any asymmetries exists, these will be likely to arise from the functioning of the economy itself. To better understand the possible implication of asymmetries, it is necessary to define that the unemployment rate is a common macroeconomic indicator of the business cycle that provides evidence about the state of the labor market and is a variable of essential relevance to policy makers. Therefore, the analysis of the subsequent relationship would imply that monetary policies in each nation needs to take into consideration the state of the macroeconomy we rely on.

1 All data has been obtained from the National Administrative Department of Statistics (DANE) of Colombia.
2 See Okun (1962).

Moreover, the investigation conducted by Lasso et al,. (2014) postulates that if the unemployment rate reveals an asymmetric pattern these would be crucial to analyze for a monetary policy because it would reflect a rapid surge when rising and a gradual decrease when diminishing and is totally linked to government expenditure and income. For that reason, there is a growing attention in being able to create a proper analysis of the OL.

The main purpose of the study is to provide an analysis of OL applied in Colombia, in order to check whether the existing relationship between the change in unemployment and economic growth follows a linear or non-linear behavior. It should be noted that the analysis has is rooted from the study realized by Christopoulos et. al (2019), using a Smooth Transition Autoregressive Model -STAR-, stated that is possible to suggest that the effectiveness of stabilization policy in the real economy would depend on which regime the Okun relationship is found. For Colombia, the econometric model implemented captures a smooth transition between regimes, which may be more reasonable compared to abrupt transitions.

A STAR model was used, since it allows to capture the effect of a smooth transition of the economic growth between two states of nature -a high and a low regime-, considering the speed of change and the optimal threshold level. The high regime was found to represent an event in which economic growth incremented without it being supported by its fundamentals. That is, without the unemployment rate having decreased a magnitude that could represent this increase in economic growth. The low regime represents an event in which unemployment rate decreases, without having an impact -or at least an immediate effect- on economic growth. An analysis of the transition between regimes will be presented in the results section.

The present document is divided into six sections, this introduction being the first. The second section presents a detailed review of the existing literature on nonlinear OL modeling. The data used in this document is explained in the third section, while in the fourth the methodology used for the study is explained in detail. The fifth section presents the results obtained. Finally, the last section will have the conclusions.

2. Literature Review

The first relevant element worth discussing, is the adaptation of OL that will be implemented in this study. Originally, Okun (1962) described two empirical relationships between unemployment and economic growth, both linear and represented with simple equations. The author observed that higher labor participation was required in order to being able to produce more goods and services. For this reason, the unemployment rate -UR- represents the percentage of people involved in the labor market. Thus, if an economy is growing, more labor is being used and the unemployment rate goes down. This is the first relationship developed by Okun in which changes or “growths” are analyzed. Indeed, the correspondence between these variables has been expanded by economists to consider factors that they believed have been left out by the author. The second relationship analyzed by the author represents how much the economy could produce in a condition of full employment, which was related with the potential output. These did not study the change in the unemployment rate, but analyzed the unemployment rate in levels contrasted with the gap between the potential and actual output. The last gave the estimation the possibility to estimate the level of unemployment that an economy must reach to guarantee a gap equal to cero. Nonetheless, the problem with this equation relays in both potential output and the full employment being unobservable variables.

Okun was aware of the simplicity of the postulated equations and noted that this was potentially problematic. According to Knotek (2007), this is the reason behind many economists proposing variations -such as dynamic equations- on Okun's original relationships and often still calling them "Okun's Law". However, Okun did not formally named the relationship as a "Law", he stablished a coefficient to evaluate the magnitude of the impact in unemployment when the level of output changes in an economy -Okun's coefficient-. In fact, what is known by economists as the OL, is a "rule of thumb" that started from Okun's observations, but that it is known to be really different if the period of time or if the economy of study changes. Okun proposed a strict coefficient that has been proven not being that strict for several economies. The truth is that the use of a dynamic version of Okun's observations is justified given that, as long as the intended investigation analyze the variations of output given changes in the unemployment rate, the OL relationship is still being analyzed.

Given all of the above, it is pertinent to implement a dynamic difference model to represent OL, in which changes in the unemployment rate and output growth will be used, represented by the first difference, and analyzing past changes in both variables. The usage of first difference of the unemployment rate and of the output is justified on the first relationship described by Okun. Contrary to the common perception that the variables must be estimated in levels to have "Okun's law", it was worked by its author from the beginning using the first difference.

To further understand the OL applicability, the nature of the relationship between economic growth and UR most be considered. Recently, abundant studies have explored the asymmetric relationship between the change in unemployment rate and economic growth for different time periods and for different countries. There appears to be a consensus in the literature concerning the empirical validity of the OL with a negative coefficient. However, discrepancy as to the actual linear assumption exists. There are several studies of asymmetric relationships that tend to use varying model specifications and estimation methods, as well as different country samples. These studies explain the cause of asymmetries and nonlinearities from different perspectives. Indeed, Harris et al.,(2001) formulates asymmetric responses due to heterogeneity in the location in terms of job creation and job destruction facing shocks. Mayes et al.,(2002) attributes asymmetry to the behavior of the labor market due to mismatching between jobs and unemployment in different regions and sectors. Campbell et al.,(2000) focuses on aggregate asymmetries in job creation and destruction attributable to microeconomic asymmetries in adjustment costs, in which the positive external shocks caused a higher response for firing workers than for hiring new ones.

Furthermore, Ekner (2018) found that asymmetric dynamics should be classified in two regimes: a recession regime and an expansion regime. Is important to notice that a key factor for the mentioned research is that the main source of non-linearity stems from economic downturns, where the model switches from an explosive recession regime to explain the observed steep increases in the unemployment rate, before it subsequently returns the to the stationary expansion regime. The paper published by Huang et al,.(2008) also estimated, for Canada, two regimes in output gap around a threshold. It also showed evidence that the asymmetry specification implies that expansions and contractions in economic growth are not associated with the same absolute changes in the unemployment rate. For that reason, a correct identification of the non-linear characteristics of the relationship between the change in unemployment rate and economic growth is crucial for the correct design, the implementation and the effectiveness of economic policies aimed at improving the labor market operation.

One interesting aspect of the analysis carried out by Ekner (ídem) is that the relationship of OL between measures of economic growth and the change in unemployment rate is defined as a linear cointegration correspondence between non-linear short-run dynamics. This linear cointegration and short run dynamics of the relationship are modeled by means of the STAR methodology in which the dynamics depend non-linearly on a transition variable. This extension facilitates asymmetry of the two resulting regimes and the transition between them may be smooth rather than discrete.

The way that nonlinearities and asymmetries are analyzed may be particularly important to better understand the relationship between the change in unemployment and economic growth, given the vast amount of studies that raised awareness of the presence of these asymmetries. The study aims to continue improving that potential problem of the simplicity of the original relationship stablished by Okun. If there is evidence for considering asymmetries, these should had been studied from a non-linear perspective given that the relationship cannot be accurately explained by any means with a lineal regression. Moreover, these asymmetries might be the reason, among others, behind Okun's coefficient shifting when analyzing different economies. The grade of "relationship" between these variables is not always the same, and this might influence the impact on these two variables that a certain policy could have, hence, justifying the nonlinear approach.

After establishing how OL is going to be evaluated, it is still relevant to further explore what can be gained by this study. According to Nebot et al,.(2019) OL is used to measure the cost to society of excessive unemployment. The authors estimated Okun's relationship and confirmed the existence of two regimes for each European country analyzed. As can be seen, the same analysis of OL with the two regimes is relevant for Colombia. Flórez et. al (2018), using a Vector Error Correction Model (VECM), also found two regimes during the period from 1984 to 2016, concluding that the change of regimens was due to important institutional chances on the labor market in the late nineties. Considering that this investigation is the most recent work on a nonlinear approach to OL in Colombia, the intended purpose of the research is to contribute the economic literature regarding OL using a smooth transition approach. Moreover, the use of more recent data could be useful in future policy-making processes in the country.

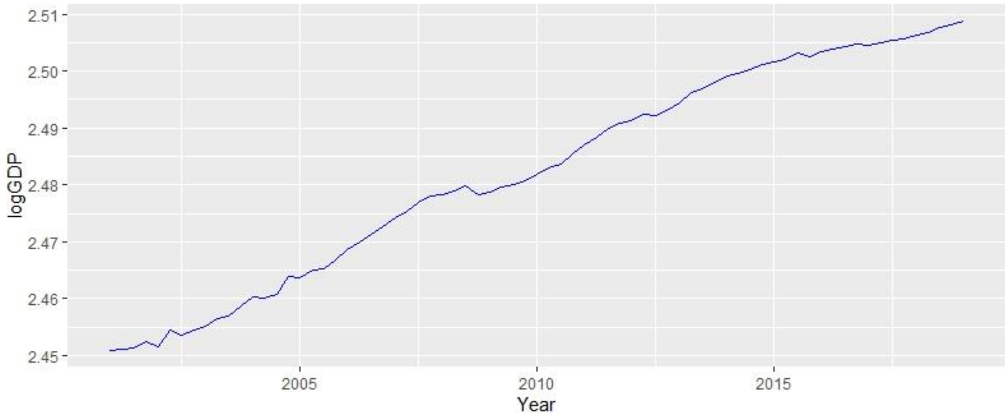
3. Data

In order to study the OL behavior in Colombia, a database was built which comprises a period that begins in the first quarter of 2002 until the last quarter of 2019. This section will describe the specifications of each variable.[3]

Figure 1. Figure 1 displays the logGDP time series, starting from the first quarter of 2002 through the last quarter of 2019. The Gross Domestic Product -GDP- data was collected from the National Administrative Department of Statistics (DANE) of Colombia. Data from the year 2020 was not consider due to the values from the last quarter not being public at the time the study started and given the shock from the Covid-19 pandemic in the first quarter. A logarithmic transformation was applied with reescalation purposes.[4]

Figure 1.[5]

Quarterly Economic Growth (Logarithmic Scale)

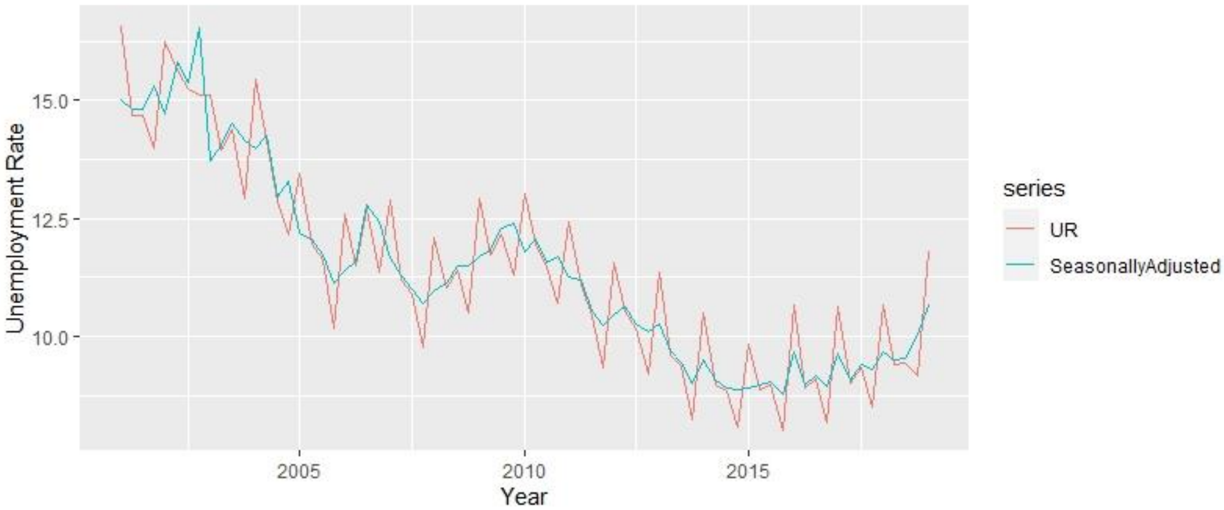


[3] Descriptive statistics of the data in Appendix A.1.
[4] Reescalation is not the only reason behind the transformation. It is important to remember that in the model the first difference is going to be implemented, and differences of logarithms are a close approximation to a change rate. Then, when performing the analysis with logGDP, growth rates are being analyzed.
[5] Own elaboration.

Figure 2. Figure 2 displays both UR and the seasonally adjusted unemployment rate series, starting from the first trimester of 2002 through the las trimester of 2019. The UR data was collected from the National Administrative Department of Statistics (DANE) of Colombia. The X11 seasonal adjustment technique was implemented and the first difference conducted in order to obtain the change in the unemployment rate seasonal adjusted.[6]

Figure 2[7]

Quarterly Unemployment Rate – Seasonally Adjusted



3.1 Variable Transformations

For economic growth, a logarithmic transformation was applied in order to rescale the series. In addition, the first difference was applied:

$$\Delta logGDP_t = log(GDP_t) - \log(GDP_{t-1}) \quad (1)$$

The Unemployment Rate (UR) exhibits a clear seasonal behavior during the period analyzed. The most commonly used seasonal adjustment techniques are those in the X11 family. The X11 methodology involves applying symmetric moving averages to a time series in order to estimate the trend, seasonal and irregular components.[8] The X11 technique was applied to the UR_t and the first difference was conducted in order to obtain the change in the unemployment rate seasonal adjusted:

[6] Regarding data prior to the year 2001, the unemployment rate series presented various problems. The methodology for data collection shifted two times. Before this year, the surveys were different and were taken only from Colombia's principal cities. The sample of the output, an indicator taken from the whole country, is then different from the one of UR and made an analysis of OL at a country level impossible to perform.

[7] Own elaboration

[8] See Macaulay, F. (1931).

$$\Delta SUR_t = SUR_t - SUR_{t-1} \quad (2)$$

Both variables were studied under the first difference regarding the presence of unit root and the stationarity condition required for the STAR model.

4. Methodology

4.1 Smooth Transition Auto Regressive Model (STAR)

The recent development of econometrics has made possible to evaluate economic variables from different approaches in order to better explain certain behaviors. The STAR -Smooth Transition Autoregressive- models introduced by Tong (1990), Teräsvirta, et al. (1994) helped to incorporate non-linear explanation into basic regression models through the implementation of a function that represents the smooth transition between two regimes of the dependent variable. These resulted in the following form, given $\forall\, t = 1, \cdots, T$:

$$y_t = (\phi_{0,1} + \phi_{1,1} y_{t-1} + \cdots + \phi_{p,1} y_{t-p})\big(1 - G(S_t, \gamma, c)\big) + \\ (\phi_{0,2} + \phi_{1,2} y_{t-1} + \cdots + \phi_{p,2} y_{t-p}) G(S_t, \gamma, c) + \varepsilon_t \quad (3)$$

In which $z_t = (1, \tilde{z}_t')'$; $\tilde{z}_t' = (y_{t-1}, \cdots, y_{t-p})$ and G is the transition function, continuous and bounded on the interval (0,1). In general, $\tilde{x}_t'$ can include not only lags of the dependent variable but also relevant exogenous variables in the explanation of y_t. In this way, we have that $\tilde{z}_t' = (y_{t-1}, \cdots, y_{t-p}, x_{1t}, \cdots, x_{kt})$. That is, $\tilde{\boldsymbol{z}}_{\boldsymbol{t}}'$ is made up of $m = (p + k)$ elements. Thus, the model can be rewritten as follows:

$$y_t = \phi_1' z_t \big(1 - G(s_t, \gamma, c)\big) + \phi_2' z_t G(s_t; \gamma, c) + \varepsilon_t \qquad (4)$$

where

$$\phi_i = \left(\phi_{0,i}, \phi_{1,i} \cdots, \phi_{p,i}, \phi_{p+1,i}, \cdots, \phi_{p+k,i}\right)' \quad \forall i = 1,2 \qquad (5)$$

As can be seen in the equation (3), the function **G** depends on the variables $(\boldsymbol{s_t}, \boldsymbol{\gamma}, \boldsymbol{c})$, where $\boldsymbol{s_t}$ is the variable in charge of making the transition, normally known as the delay variable; $\boldsymbol{\gamma}$ is the one that determines the rate of change, finally $\boldsymbol{c}$ is that optimal level of $\boldsymbol{s_t}$ that generates a change in the states of nature. Dijk et al. (2002) mention that recent econometric developments suggest that the most popular functional forms are logistic and exponential.

The LSTAR model is determined by considering the logistic function as a transition function in equation (6):

$$G(s_t;\gamma,c) = 1/\big(1+exp(-\gamma(s_t-c))\big) \qquad (6)$$

The ESTAR model is determined by considering the exponential as a transition function in equation (7):

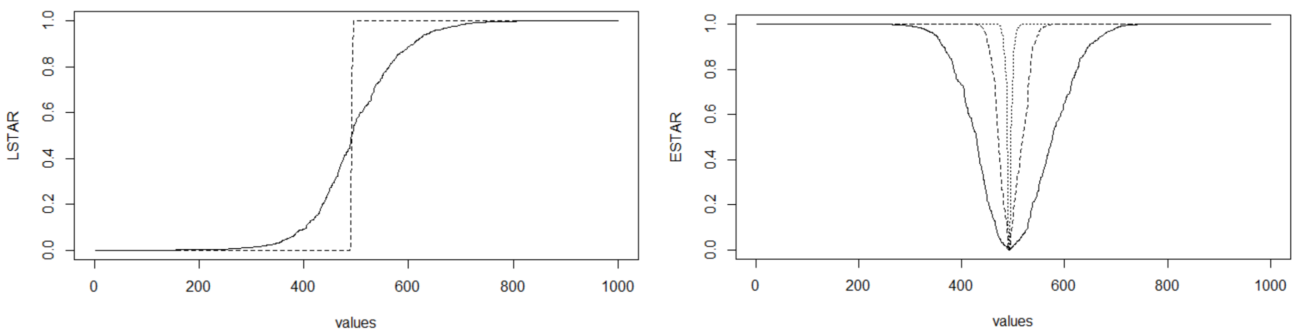
$$G(s_t;\gamma,c) = 1-exp(-\gamma(s_t-c)^2) \qquad (7)$$

The STAR models characterize asymmetric behaviors in the economy. The LSTAR model allows describing a process whose dynamics are opposite in each of the two states, which are determined by the fact that the transition variable is above or below the associated threshold. Although, in the ESTAR model, the dynamic behavior of the process is similar for both large and small values of the transition variable against the threshold and different in the mean values. It is noteworthy that in both models the transition between states is smooth.[9]

Figure 3. Simulation of the transition function. For both models, is crucial to notice that the parameter $\gamma > 0$ measures the smoothness of the change in the transition function. The change between the two regimes are related with large or small values of the value of the transition variable against the parameter c. The simulation was created using different values for the parameter γ, the LSTAR model was simulated using γ: 5,100 and the ESTAR model with γ: 1,10,100.

Figure 3[10]

s_t vs $G(s_t;\gamma,c)$ Simulation size (1000)



A key relevant aspect is that when $\gamma \to 0$ or ∞ the change between the two regimes is immediate and the smoothness is not perceived. In fact, the LSTAR model becomes an AR (p) model since each value of G is constant. For intermediate values of γ, the degree of autoregressive decay depends on the value of s_t. As $s_t \to$ - ∞, $G \to 0$, and as $s_t \to$ + ∞, $G \to 1$. The ESTAR model as $s_t \to c$, $G \to 0$, and as $s_t \to \pm$ ∞, $G \to 1$.

[9] See Misas, M. (2020).

[10] Own elaboration. This simulation was performed for illustrative purposes. In both, the threshold used was 0 and the delay variable was defined for each model as a uniform distribution from (-5,5). For the estimation of the LSTAR, the simulation was run with the dotted line represented a gamma of 100 and the straight one of 5. For the estimation of the ESTAR, the simulation was run with a gamma of 100, 10 and 1, inside out respectively.

4.2 Stationarity Tests

STAR models origin from the assumption that the dependent variable follows a stationary stochastic process, that is, with a constant mean, finite variance that is not a function of time, and transitory effects of the disturbances on the level of the variable. To check the stationarity of the variables and following the methodology proposed by Dickey et al., (1979), it is necessary to perform statistical tests that suggest the existence of a unit root through the following ADF -Augmented Dickey Fuller- regression models:

$\boldsymbol{Model\ 1(\tau)}$:

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=1}^{p} \boldsymbol{\phi}_i \Delta Y_{t-i} + \varepsilon_t \qquad (8)$$

$\boldsymbol{Model\ 2(\tau\mu)}$:

$$\Delta Y_t = \mu + \gamma Y_{t-1} + \sum_{i=1}^{p} \boldsymbol{\phi}_i \Delta Y_{t-i} + \varepsilon_t \qquad (9)$$

$\boldsymbol{Model\ 3(\tau\tau)}$:

$$\Delta Y_t = \mu + \delta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \boldsymbol{\phi}_i \Delta Y_{t-i} + \varepsilon_t \qquad (10)$$

Finally, is important to consider that a reduce number of lags could lead to residuals that might not meet the white noise requirement. Consequently, the estimation of the parameter γ and its standard deviation is inappropriate, which would affect the result of the unit root test. On the other hand, a large number of lags reduce the power of the test to reject the unit root null hypothesis, which is basically obtained given thet the inclusion of more lags also increases the number of parameters to be estimated and there is a clear loss of degrees of freedom. For that reason, the methodology proposed to analyze the time series for the investigation is starting with a high -plausible- number of lags and go down through significance tests of type "t" or F.[11] The mathematical notion is, initiate with $P = P^*$, if the coefficient of certain lag is not significant, consider $P = P^* - 1$. Continue in this way until the last lag is significant and white noise is generated in the residuals, verifying using the Ljung-Box Test.[12] Lastly, after selecting the appropriate lag to run the test, the test statistic should be contrasted with the simulated critical values for the Dickey-Fuller distribution (showed in the Appendix A.2, A.4).

The analysis of deterministic components is also of great relevance. Initially, the specification of the model to be used - $\boldsymbol{\tau}, \boldsymbol{\tau\mu}, \boldsymbol{\tau\tau}$- will depend on the data generating process of the variables. For that reason, the introduction of the deterministic components will be determined by the significance of the parameters after

[11] Hendry, D. (2005). General-to-specific Modeling: An Overview and Selected Bibliography
[12] Ljung-Box Test present a null hypothesis of independence in the residuals. (White noise).

the estimation of the auxiliary regression. That is, for example, if the auxiliary regression reflects significance in the trend component, then it is sufficient evidence that the stationarity test must include the term. The same intuition work for the inclusion of the intercept term.

4.3 Cointegration Analysis

A common critique for many economists relies on the usage of linear regressions to find the relationship between several time series processes. However, a formal cointegration analysis should be consider for the purpose of the investigation. The concept of cointegration was introduced by Engle and Granger (1987) illustrating that when dealing with time series that are first-order integrated $I(1)$ there is the possibility that some cointegrating vector exists to form a stationary linear combination of the variables. First, consider a set of $\boldsymbol{n}$ economic variables that contemplate a long-run equilibrium relationship when $\beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_n x_{nt} \sim I(0)$

Letting β and x_t denote the vectors (β_1, β_2, ..., β_n) and (x_{1t}, x_{2t}, ..., x_{nt}) ′ the arrangement is in long-run equilibrium when $\beta x_t{}' \sim I(0)$. The disequilibrium error is ε_t, so that $\varepsilon_t = \beta x_t$, the term refers to any long-run disequilibrium relationship among nonstationary variables. Enders (1995) stipulates that the disequilibrium relationship may be causal, behavioral, or just a reduced form relationship among correspondingly trending variables. One important remark is that the linear combination of the equilibrium theories involving nonstationary variables entail the presence of a combination of the variables that is stationary. For continue exploring the possibility of cointegration of the OL, the Engle-Granger methodology should be used. First, pretest the variables for their order of integration, verifying that both are $I(1)$.

$$Y_t = \beta_0 + \beta X_t + \varepsilon_t \qquad (11)$$

Second, consider the autoregression of the residuals -disequilibrium-, $\Delta\widehat{\varepsilon_t} = \alpha_1 \varepsilon_{t-1} + \sum_{i=1}^{T} \alpha_{t-i} \Delta\widehat{\varepsilon_{t-1}} + \mu$. Then compute the Engle-Yoo test on $\alpha_1 = 0$ using the critical values described in the Appendix (A.3). The rejection of the null hypothesis of no cointegration implies that the disequilibrium term is $I(0)$ variable. Now on the intention is to apply the Error Correction Mechanism model -ECM-. In this model, the short-term dynamics of the variables in the system are influenced by the deviation from equilibrium, and using the Granger representation theorem stating that for any set of $I(1)$ variables, error correction and cointegration are equivalent representations. The possible ECM is presented:

$$\Delta Y_t = \beta_0 + \boldsymbol{\phi}_{\mathbf{1,1}} \Delta Y_{t-1} + \cdots + \boldsymbol{\phi}_{\boldsymbol{p},\mathbf{1}} \Delta Y_{t-p} + \boldsymbol{\phi}_{\mathbf{2,1}} \Delta X_{t-1} + \cdots + \boldsymbol{\phi}_{\boldsymbol{p},\mathbf{2}} \Delta X_{t-p} - \alpha \varepsilon_{t-1} + \upsilon_t \quad (12)$$

Finally, the selection of the number of lags included in the model will similarly follow the stationary test selection methodology in function of the parsimony principle. Additionally, is important to notice that a primary feature of cointegrated variables is that their time paths are influenced by the extent of any deviation from long-run equilibrium; if the system is to return to long-run equilibrium, the movements of at least some of the variables must respond to the estimated magnitude of the disequilibrium, which is defined as the speed of adjustment of the ECM. The speed of adjustment requires to be estimated significantly and with negative coefficient.

4.4 Linearity Tests

Considering that STAR models begin from the premise that there is a parameter that behaves non-linearly within the model, it is necessary to construct a linearity test to corroborate whether this characteristic is present in the relationship between the dependent variable and the delay variable. Since the non-linear model can only be identified under the alternative hypothesis, of non-linearity, according to Teräsvirta (1994) it is necessary to approximate the transition function in equation (6) through a Taylor expansion of order 3 around $\gamma = 0$, which generates the following auxiliary regression:

$$y_t = \boldsymbol{\phi}_0' z_t + \sum_{j=1}^{3} \boldsymbol{\phi}_j' z_t s_t^j + \varepsilon_t^* \qquad (13)$$

Through which the existence or not of linearity is verified by means of a statistical test type F that contrasts the null hypothesis of linear behavior of the relationship, H_0: $\boldsymbol{\phi}_1 = \boldsymbol{\phi}_2 = \boldsymbol{\phi}_3 = 0$ versus a non-linear dynamic in the alternative hypothesis, that is, H_a: $\exists\, j \ / \ \boldsymbol{\phi}_j \neq 0$.[13] Knowing that for the investigation it was concluded that $s_t \in z_t$, the test should be applied following these steps:

- Estimate the model under the null hypothesis using OLS and calculate its residuals $\tilde{\varepsilon}_t$ and its corresponding sum of squared residuals, SSR_0.
- Estimate by OLS the regression of $\tilde{\varepsilon}_t = f(\tilde{z}_t, h_t^0)$ where $h_t^0 = (\tilde{z}_t' s_t, \tilde{z}_t' s_t^2, \tilde{z}_t' s_t^3 \ )'$ and calculate their sum of squared residuals, SSR_1.
- Calculate the F test statistic

Where the Teräsvirta F test statistic is computed as follows:

$$LM_2 = \frac{(SSR_0 - SSR_1)/(3m)}{SSR_1/(T-4m-1)} \sim F(3m, T-4m-1\,) \qquad (14)$$

It should be noted that the test of verification of non-linearity by Escribano et al,. (1999) follows the same scheme already presented. However, in this case a Taylor expansion of order 4 is performed around the

[13] See Misas, M. (2020).

parameter γ as an approximation of the transition function, resulting in the following test regression:

$$y_t = \boldsymbol{\phi}_0' z_t + \sum_{j=1}^{4} \boldsymbol{\phi}_j' z_t s_t^j + \varepsilon_t^* \qquad (15)$$

The Escribano-Jordá F test statistic is computed as follows:

$$LM_3 = \frac{(SSR_0 - SSR_1)/(3m+1)}{SSR_1/(T-4(m+1))} \sim F(3m+1, T-4(m+1)) \qquad (16)$$

The rejection of the null hypothesis would indicate the presence of two states of nature in which the transition is caused by the chosen delay variable.

4.5 Transition Function Tests

Regarding the initial assumption that the relationship between the change in the unemployment rate and economic growth in Colombia has a non-linear behavior, it is necessary to establish the model -LSTAR or ESTAR- that best fits the nature of such relationship. Teräsvirta (1994) suggests a model selection process through the analysis of a nested F-type test that arises from equation (13). The selection of the transition function is based on the significance of $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$ and $\boldsymbol{\phi}_3$, considering again that $s_t \in z_t$.[14]

The significance of the cubic part is tested using the LM statistics. LM. H_{01}: $\boldsymbol{\phi}_3 = 0$ its rejection can be interpreted as rejection of the ESTAR family, test statistic:

$$LMR_1 = \frac{(SSR_0 - SSR_1)/(m)}{SSR_1/T-4m-1} \sim F(m, T-4m-1) \qquad (17)$$

Where SSR_0 is the sum of squared residuals of Equation (11) with j = 1,2

SSR_1 is the sum of squared residuals of Equation (11) with j = 1,2,3

If the cubic part is significant, the appropriate transition function is not the exponential. The rejection of the null hypothesis shows significance of the cubic part. The cubic function has a positive slope, this means that as s_t increases, the weighting of the explanatory variables also increases. Among the two possible models, the LSTAR fulfills these characteristics and would then be the indicated model.

In the opposite case, the significance of $\boldsymbol{\phi}_2$ is tested, excluding $\boldsymbol{\phi}_3$ from the regression H_{02}: $\boldsymbol{\phi}_2 = 0/\boldsymbol{\phi}_3 = 0$, the fact of not finding evidence to reject the null hypothesis is taken as evidence in favor of LSTAR and a strong rejection could show an ESTAR.

[14] See Misas, M. (2020).

$$LMR_2 = \frac{(SSR_0 - SSR_1)\big/_{(m)}}{SSR_1\big/_{T - 3m - 1}} \sim F(m, T - 3m - 1) \qquad (18)$$

Where SSR_0 is the sum of squared residuals of Equation (11) with j = 1

SSR_1 is the sum of squared residuals of Equation (11) with j = 1,2

The rejection of the null hypothesis shows significance of the quadratic part. The quadratic function does not have a strictly positive slope, this means that as s_t increases, the weighting of the explanatory variables does not necessarily surge. Among the two possible ones, the ESTAR model meets these characteristics and would be the one indicated.

In the case that $\boldsymbol{\phi}_2 = \boldsymbol{\phi}_3 = 0$, the significance of $\boldsymbol{\phi}_1 . H_{03}$: $\boldsymbol{\phi}_1 = 0 / \boldsymbol{\phi}_2 = \boldsymbol{\phi}_3 = 0$. The existence of evidence for rejecting H_{03} after accepting H_{02} supports the decision for the LSTAR model. Otherwise, after rejecting H_{02} it would suggest that an ESTAR model could be selected.

$$LMR_3 = \frac{(SSR_0 - SSR_1)\big/_{(m)}}{SSR_1\big/_{T-2m-1}} \sim F(m, T - 2m - 1) \qquad (19)$$

Where SSR_0 is the sum of squared residuals of Equation (#) with j = 0, the linear part.

SSR_1 is the sum of squared residuals of Equation (#) with j = 1

If $\boldsymbol{\phi}_1 \neq 0 | \boldsymbol{\phi}_2 = \boldsymbol{\phi}_3 = 0$ holds, there is evidence for not rejecting a logistic transition function. Finally, the rejection of the null hypothesis shows the significance of the linear part. The linear function has a positive slope, this means that as the variable delay increases, the weighting of the explanatory variables also increases. Among the two possible models, the LSTAR fulfills these characteristics and would be indicated model.

4.6 Estimation

Once it is decided that the relationship proposed in the regression model is non-linear and the transition function or type of model is selected in function of the transition variable that supports the non-linearity, the estimation procedure should start. Since the model is non-linear, the estimation must be made from a non-linear optimization algorithm whose objective function is defined as:

$$\lambda = \underbrace{min}_{\phi_1, \phi_2, \gamma, c} \sum_{t=1}^{T} \left\{ \left(y_t - \phi_1' z_t \left(1 - G(s_t, \gamma, c)\right) - \phi_2' z_t G(s_t; \gamma, c) \right)^2 \right\} \qquad (20)$$

For solving the minimization problem, the iterative mechanism requires the selection of initial values for the referent parameters in the objective function. The non-linear regression model presented in Equation (3) can be transformed into a linear model in parameters, if $\boldsymbol{\gamma}$ and $\boldsymbol{c}$ are determined. Basically, the idea is to established a grid of possible values for these parameters, from which the $(\boldsymbol{\gamma}^{\boldsymbol{n}}, \boldsymbol{c}^{\boldsymbol{n}})\ \ \boldsymbol{n} = \mathbf{1}, \cdots, \boldsymbol{N}$ are randomly sampled $\boldsymbol{N}$ times, and the corresponding $\boldsymbol{\phi}_{\mathbf{1}}^{\boldsymbol{n}}$ and $\boldsymbol{\phi}_{\mathbf{2}}^{\boldsymbol{n}}$ are estimated conditionally to the regarding value of the transition function through time. According to the procedure mentioned previously, the following $\boldsymbol{N}$ initial vectors are constructed $(\boldsymbol{\gamma}^{\boldsymbol{n}}, \boldsymbol{c}^{\boldsymbol{n}}, \boldsymbol{\phi}_{\mathbf{1}}^{\boldsymbol{n}}, \boldsymbol{\phi}_{\mathbf{2}}^{\boldsymbol{n}})\ \ \boldsymbol{n} = \mathbf{1}, \cdots, \boldsymbol{N}$.[15] The estimation of the $(\boldsymbol{\gamma}^{\boldsymbol{n}}, \boldsymbol{c}^{\boldsymbol{n}}, \boldsymbol{\phi}_{\mathbf{1}}^{\boldsymbol{n}}, \boldsymbol{\phi}_{\mathbf{2}}^{\boldsymbol{n}})\ \ \boldsymbol{n} = \mathbf{1}, \cdots, \boldsymbol{N}$ parameters is carried out considering in equation (3) the following transformation of the transition function:

- *LSTAR:* $1\Big/\left(1 + exp\left(-{}^{\gamma}\!\Big/_{\widehat{STD}(s_t)}(s_t - c)\right)\right)$ *(21)*
- *ESTAR:* $1 - exp\left(-{}^{\gamma}\!\Big/_{\widehat{STD}(s_t)}(s_t - c)^2\right)$ *(22)*

Finally, the sum of the estimated nonlinear quadratic residuals is obtained $\lambda^n = \sum_{t=1}^{T} \dot{\varepsilon}_t^2 \ \ n = 1, \cdots, N$ and the selection of that estimated vector will lead to the minimum value of $\lambda^n \ \ n = 1, \cdots, N$.

4.7 Regimes Intuition

The STAR models characterize asymmetric behaviors in the economy. This type of models allows describing a process whose dynamics is recognized in each of the two states, which depend by the fact that the delay variable is above or below the associated threshold. Although, in the ESTAR model, the dynamic behavior of the process is similar for both large and small values of the delay variable versus the threshold and dissimilar in the mean values. In other words, if the delay variable $\boldsymbol{s_t} > \boldsymbol{c}$ and $\boldsymbol{s_t} < \boldsymbol{c}$ the transition function will be closer to 1. Alternatively, if the delay variable $\boldsymbol{s_t} \rightarrow \boldsymbol{c}$ then de transition function will be closer to 0.

Then again, the LSTAR model is capable of characterizing asymmetric behavior as well. As an example, suppose that $\boldsymbol{s_t}$ measures the phase of the business cycle. Whenever the delay variable $\boldsymbol{s_t} > \boldsymbol{c}$ the transition function will be closer to 1. Otherwise, if the delay variable $\boldsymbol{s_t} < \boldsymbol{c}$ the transition function is closer to 0. Teräsvirta (1994) formulates that the model can describe processes whose dynamic properties are different in expansions from what they are in recessions, and the transition from one extreme regime to the other is smooth. The regimes characterization is helpful for the intended research of the OL relationship, in which is important to prioritize the economic intuition of the regimes.

[15] See Misas, M. (2020).

5. Results

5.1 Stationarity Tests

Following section 4.2 in order to establish the order of integration of the variables, the data generating process of each time series was individually analyzed. The inclusion of the deterministic components was determined by the analysis of the estimated auxiliary regression described in equations (8), (9) and (10). In this way, the test $\tau\mu$ was selected regarding the needed inclusion of the intercept term for each of the variables. The lag specification followed the fulfilment of the white noise condition and the procedure described in the section 4.2.

Table 1. The results of the stationary test for logarithmic economic growth $logGDP_t$ and the seasonal adjusted unemployment rate SUR_t showed that both are non-stationary at levels. Consequently, the value obtained by the test statistic in absolute value is less than the critical value proposed by the test. Thus, there is not sufficient evidence to reject the null hypothesis of unit root existence. For that reason, a second test was conducted on the first difference for each variable. The explained and the explanatory variables adopted the stationary behavior after applying the first difference, demonstrating that both variables are integrated of order one $I(1)$.

Table 1.

Unit Root Existence Tests

Variable	***Test ADF***	***Augmented Lags***	***Ljung-Box*** $Lags = 16$
$logGDP_t$	$\tau\mu = -1.6618$ $Critical\ value\ (5\%): -2.89$	2 $P - Value: 0.0607$	$P - Value$: 0.1623
SUR_t	$\tau\mu = -1.7082$ $Critical\ value\ (5\%): -2.89$	1 $P - Value: 0.0889$	$P - Value$: 0.5923
$\Delta logGDP_t$	$\tau = -4.95$ $Critical\ value\ (5\%): -1.95$	1 $P - Value: \sim 0.0001$	$P - Value$: 0.2414
$\Delta \mathrm{SUR}_t$	$\tau = -11.96$ $Critical\ value\ (5\%): -1.95$	0 $P - Value: 0.00002$	$P - Value$: 0.6327

5.2 Cointegration Analysis

At this point, the long-term equation was proceeded to be estimated following section 4.3, where the OL equation is presented in levels. In this, a trend component and an intercept were included. The following long-run equation was estimated:

$$logGDP = \beta_0 + \delta t + \beta UR_t + \varepsilon_t \qquad (23)$$

Table 2. Estimation of Okun's Law Long Run Equation. The estimated variables presented significance at 5% level.

Table 2

Estimation of the Okun's Law Long-run Equation

Variable	*Coefficient*	*Standard Error*	*t value*	*p value*
$Intercept$	*11.872**	*0.036*	*325.169*	*~0,0001*
SUR_t	*-0.019**	*0.002*	*-7.936*	*~0,0001*
δt	*0.008**	*0.0002*	*36.682*	*~0,0001*
**: Significance (5%)*				

After computing the estimation, the residuals -long-term disequilibrium- where extracted. Consequently, the test explained in the previous 4.3 section was constructed. The results for the test are presented in the following table 3.

Table 3. Engle-Granger Cointegration Test (Long Run Equation Residuals). Due to the fact that the test presented a test statistic (-3.95) higher than the critical value (-3.46) and was determined that the long-term disequilibrium does not present a unit root. Consequently, it is derived from a stationary process in mean and variance, and therefore there are evidence of long-term movements for both variables -cointegrated variables-
.

Table 3

Engle-Granger Cointegration Test (Long Run Equation Residuals)

Variable	***Engle – Yoo Test***
$\widehat{\varepsilon_t}$	$Test: -4.00$ $Critical\ value\ (5\%): -3.39$ $Sample \approx 100$

As described in the 4.3 section the equation (11) was used for the application of the error correction model. The ECM was estimated as follows:

$$\Delta logGDP_t = \beta_0 + \phi_{11}\Delta logGDP_{t-1} + \phi_{12}\Delta logGDP_{t-2} + \phi_{13}\Delta logGDP_{t-3} \\ +\phi_{21}\Delta SUR_{t-1} + \phi_{22}\Delta SUR_{t-2} + \phi_{23}\Delta SUR_{t-3} - \alpha\widehat{\varepsilon_{t-1}} + \upsilon_t \tag{24}$$

Table 4. In the following table the ECM estimation is presented. The speed of adjustment was estimated correctly, obtaining a significant negative coefficient that determines the long-term adjustment of the OL relationship.

Table 4.

Error Correction Model Estimation

Variable	*Coefficient*	*Standard Error*	*t value*	*p value*	*Ljung-Box* $Lags = 16$
$Intercept$	*0.005**	*0.002*	*2.278*	*0.026*	*p value: 0.820*
$\Delta logGDP_{t-1}$	*-0.070*	*0.124*	*-0.567*	*0.573*	
$\Delta logGDP_{t-2}$	*0.229*	*0.122*	*1.866*	*0.066*	
$\Delta logGDP_{t-3}$	*0.259**	*0.116*	*2.235*	*0.028*	
ΔSUR_{t-1}	*-0.005**	*0.002*	*-2.366*	*0.021*	
ΔSUR_{t-2}	*0.0001*	*0.002*	*0.079*	*0.937*	
ΔSUR_{t-3}	*-0.0002*	*0.002*	*-0.130*	*0.897*	
$\widehat{\varepsilon_{t-1}}$	*-0.112**	*0.059*	*-1.875*	*0.045*	
**: Significance (5%)*					

The number of significant lags was determined using the parsimony principle and following the significance of the lagged variables in first difference, regarding the "General-to-specific" methodology.[16] Moreover, the seasonal adjusted unemployment rate presented significance only the first lag. However, the inclusion of the other two lagged terms was necessary because economic growth reflected significance for the third lag. Therefore, the third lagged period is the maximum lag that allows parsimony for the fulfillment of the white noise condition in the model.

The ECM expresses that the speed of the adjustment, towards the long-run equilibrium, must be significantly negative; otherwise, an ECM specification should be rejected. The estimation of the long-run disequilibrium lagged one period produced a negative magnitude and significance: the speed of adjustment of the ECM model is estimated at (-0.1122639). The slope coefficient of the disequilibrium implies that, if in the preceding period the economic growth of Colombia was 1% higher than the predicted by the long-run equilibrium relationship, there will be an adjustment to reduce the economic growth by -0.1122639 during this period to restore the long-run equilibrium relationship between the economic growth and the change in the unemployment rate.

[16] Hendry, D. (2005). General-to-specific Modeling: An Overview and Selected Bibliography

5.3 Linearity Tests

The linearity tests are intended to identify the existence of a linear or non-linear economic relationship between the economic growth and the long-run disequilibrium. Thus, the regressions of equations (13) and (15) were carried out, evaluating the corresponding null hypothesis using the type F test proposed in equation (14) and following the procedure described in the 4.4 section. The value of the statistic turned out to be greater than the critical value corresponding to the F test with $(T - 4m - 1)$ degrees of freedom, where T is the total number of observations and m is the number of parameters estimated by the model.

Table 5. The two tests showed a clear rejection of the null hypothesis of the linearity in the ECM model only when the delay variable is the disequilibrium. The Escribano-Jordá test reflected nonlinearities using other variables as the delay; however, the common rejection stands out for the disequilibrium term.

Table 5.

Teräsvirta & Escribano-Jordá Test

S_t	LM_2		LM_3	
	$p - Value$	$Choice$	$p - Value$	$Choice$
$\Delta logGDP_{t-1}$	*0.407*	*Linear*	*0.033*	*Non-Linear*
$\Delta logGDP_{t-2}$	*0.273*	*Linear*	*0.082*	*Linear*
$\Delta logGDP_{t-3}$	*0.242*	*Linear*	*0.0001*	*Non-Linear*
ΔSUR_{t-1}	*0.222*	*Linear*	*0.012*	*Non-Linear*
ΔSUR_{t-2}	*0.690*	*Linear*	*0.329*	*Linear*
ΔSUR_{t-3}	*0.856*	*Linear*	*0.194*	*Linear*
$\widehat{\varepsilon_{t-1}}$	*0.012*	*Non-Linear*	*0.004*	*Non-Linear*

The basic intuition is that $I(1)$ time series with a long-run equilibrium relationship cannot move too far apart from the equilibrium, because economic forces will serve the purpose to reestablish the equilibrium relationship. Despite the fact that Teräsvirta Test provides evidence to accept the one lagged period disequilibrium term as the delay variable, the justification could arise from an economic perspective.

The delay variable guarantees the nonlinearity approach of the model making a smooth transition between regimes. Theoretically, OL is a linear relationship where the disequilibrium term represents the deviation of the variables from the long-run equilibrium. This is the driven factor that breaks the long-run tendency of the Okun's linear relationship at levels. The Okun's linearity assumption in the ECM model is invalidated by the disequilibrium term lagged one period, which is capable to allow the transition between two different states of nature of the short and long run dynamics of economic growth.

5.4 Transition Function Tests

Once the existence of Non-linearity has been verified, the type of STAR model or transition function that best represents the data generating process is determined following the described methodology specified in the 4.5 section. The objective is to define the model that best fits the data generating process. Thus, making use of the Table 4, the tests suggested by Teräsvirta (1994) were used for the identification, carrying out the series of F-type tests described above. This was made considering the regression of equation (13).

Table 6. At this stage, it was found that for the disequilibrium term ($\widehat{\varepsilon_{t-1}}$) the model that best fits is an LSTAR, since the p-value obtained from the LMR1 test (0.00894) conduce primarily to the rejection of the null hypothesis, which is evidence that the cubic part is significant and the appropriate transition function is not the exponential. In other words, the first test presents significance for the cubic function, which has a positive slope, this means that as disequilibrium term ($\widehat{\varepsilon_{t-1}}$) increases, the weighting of the explanatory variables also increases. Continually, the LMR2 test (0.077542) provides evidence of no rejection of the null hypothesis and shows not significance of the quadratic part. The existence of evidence for no rejecting the null hypothesis supports the decision for the LSTAR model. Among the two possible models, the LSTAR fulfills the correct characteristics of the data, then being the indicated model.

Table 6

Tests on model type: Teräsvirta

	LMR_1		LMR_2		LMR_3	
S_t	$p - value$	*Choice*	$p - value$	*Choice*	$p - value$	*Choice*
$\widehat{\varepsilon_{t-1}}$	*0.008*	*Rejection of ESTAR*	*0.077*	*Evidence to support LSTAR*	*0.690*	*Could be ESTAR*

The decision of selecting the LSTAR model using the disequilibrium delay variable ε_{t-1} is quite interesting, since it allows to verify asymmetries for the Okun's Law ECM model, in which the smooth transition between the regimes is permitted due to the long-run disequilibrium of Colombian economic growth and the change of the unemployment rate at level. The delay variable indicates the rate of adjustment of OL long-run relationship to its equilibrium level in the way that economic growth adjusts when there is an imbalance in the change of unemployment rate.

5.5 Estimation

In view of the results of the previous tests established, which followed the fulfillment of the requirements to perform the estimation using the optimization algorithm, the following estimation was made based on equation (17), yielding the results shown in Table 7.

Table 7. Results of the estimation of the LSTAR model. The table shows the results obtained from the estimation of the LSTAR model through the non-linear optimization. Thus, ϕ_0 represents the intercept, ϕ_1 the first difference dependent variable lagged one period, ϕ_2 the variable laged two periods, and ϕ_3 the first difference dependent variable lagged three periods. ϕ_4 is reflecting the impact of the first difference of the change of Unemployment Rate Seasonally Adjusted, and so on until ϕ_6 which is the first difference of the change of unemployment rate lagged three periods. Finally, α corresponds to the impact of the disequilibrium $\widehat{\varepsilon_{t-1}}$. The non-linear part is comprised of the same number of parameters, which correspond to the previously mentioned variables given in the same order multiplied by $G(s_t; \gamma, c)$. Lastly, the optimal values of γ and threshold c reached by the algorithm.

Table 7.[17]

Estimation

	Variable	***Parameter***	***Coefficient***	***Standard Deviation***
Linear Part	$Intercept^*$	ϕ_{L0}	*0.005*	*0.0008*
	$\Delta logGDP_{t-1}{}^*$	ϕ_{L1}	*-0.126*	*0.001*
	$\Delta logGDP_{t-2}{}^*$	ϕ_{L2}	*0.070*	*0.001*
	$\Delta logGDP_{t-3}{}^*$	ϕ_{L3}	*0.132*	*0.004*
	$\Delta SUR_{t-1}{}^*$	ϕ_{L4}	*-0.005*	*0.001*
	$\Delta SUR_{t-2}{}^*$	ϕ_{L5}	*0.0009*	*0.001*
	$\Delta SUR_{t-3}{}^*$	ϕ_{L6}	*-0.002*	*0.001*
	$\widehat{\varepsilon_{t-1}}^*$	α_L	*-0.190*	*0.001*
Non-Linear Part	$Intercept^*$	ϕ_{NL0}	*0.010*	*0.005*
	$\Delta logGDP_{t-1}{}^*$	ϕ_{NL1}	*0.334*	*0.001*
	$\Delta logGDP_{t-2}{}^*$	ϕ_{NL2}	*0.431*	*0.019*
	$\Delta logGDP_{t-3}{}^*$	ϕ_{NL3}	*0.266*	*0.022*
	$\Delta SUR_{t-1}{}^*$	ϕ_{NL4}	*0.001*	*0.007*
	$\Delta SUR_{t-2}{}^*$	ϕ_{NL5}	*-0.002*	*0.001*
	$\Delta SUR_{t-3}{}^*$	ϕ_{NL6}	*0.005*	*0.001*
	$\widehat{\varepsilon_{t-1}}^*$	α_{NL}	*-1.024*	*0.117*
	$Speed\ of\ the\ transition^*$	γ	*7.383*	*1.185*
	$Threshold^*$	c	*0.00899521*	*0.001*
*Observations: 72 – Significance *: (5%)*				
Objective Functions: 0.00340134				
Iterations: Convergence: 166				
Normality Test p value[18] *Shapiro-Wilk: 0.9696 - Kolmogorov-Smirnov: 0.083768 - Anderson-Darling: 0.5893*				

[17] Dijk (1999) formulated that the estimation of the gradient and the hessian matrix can be performed using any conventional nonlinear optimization procedure. In fact, the optimization algorithm estimates are consistent and asymptotically normal ($\sqrt{T}\left(\hat{\theta}-\theta_0\right) \rightarrow N(0, C)$), where θ_0 denotes the true parameter values and the asymptotic covariance-matrix C of $\hat{\theta}$ can be estimated consistently as $\hat{A}_T^{-1}\hat{B}_T\,\hat{A}_T^{-1}$, where $\hat{A}_T$ is the Hessian evaluated at $\hat{\theta}$.

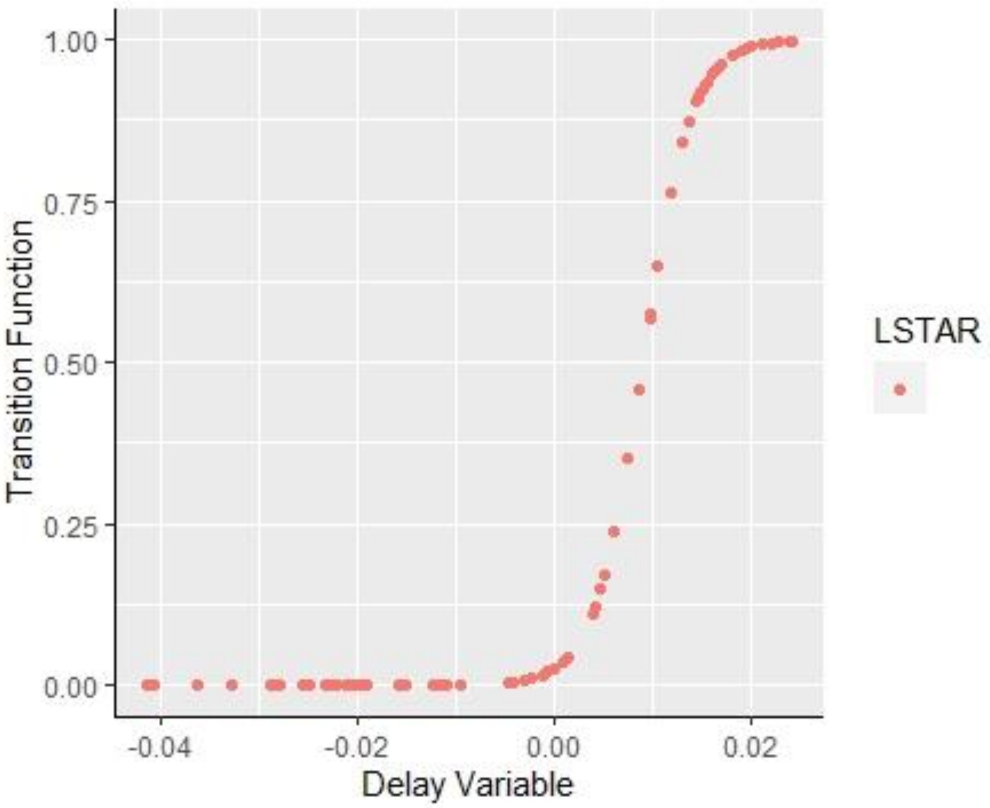
[18] The normality tests utilized were Shapiro-Wilk, Kolmogorov-Smirnoff and Anderson Darling. All the test reviews the null hypothesis of normality with different procedures and degrees of freedom. The computation of the tests provided not sufficient evidence to reject the normality null hypothesis. Given that, the significance of the variables was conducted using the normality distribution and the standard deviation of the variables.

According to the values obtained, all parameters turned out to be significant at 5% confidence level. The significance of each variable was obtained for both linear and nonlinear part. The threshold and the speed of the transition variable achieved significance. It is relevant to notice that the threshold takes a strictly positive value and the optimal value is within the limits of the minimum and maximum corresponding to the delay variable $\boldsymbol{\varepsilon_{t-1}}$, (minimum: -0.0413473, maximum: 0.02413), which is appropriate for the model.

Figure 4. Figure 4 represents the graph of the transition of the delay variable (long-run disequilibrium of the OL ECM model) between two regimes. A curious insight is that positive deviation from the long-run equilibrium are associated with the higher regime and negative deviations from the long-run equilibrium with the lower regime.

Figure 4.[19]

Transition Function vs Delay Variable

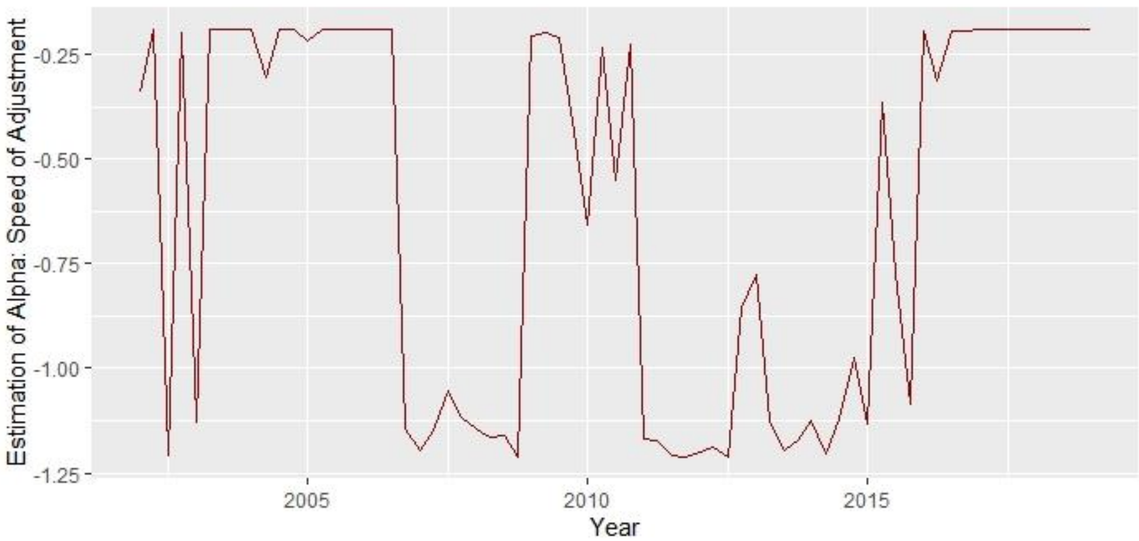


[19] Own elaboration.

Figure 5. The estimated speed of adjustment (α) associated with the disequilibrium term of the LSTAR model vary over time. The parameter relies on the transition function in which the long-term deviations from equilibrium affect the economic growth in Colombia. The estimated parameters reflect the effect of the explanatory variables on economic growth. For that reason, the cause-effect relationship varies additionally with fluctuations between the two different regimes, that way reflecting the change of the parameters over time.

Figure 5[20]

Speed of Adjustment (α)



For each period the magnitude of the disequilibrium estimated coefficient varies because it depends on the transition function. Basically, the estimated coefficient of the disequilibrium implies the speed of adjustment to reduce the economic growth to restore the long-run equilibrium relationship between the economic growth and the change in the unemployment rate.

The estimation of the disequilibrium coefficients permits the evaluation of the regimes over time. Each position in the graphic demonstrate the type of regime in the economy. For better understanding the regimes transition and the relevance for the Colombian economy, the next section will provide a further analysis of the changes in the state of nature.

[20] Own elaboration.

5.6 Regimes intuition

Each time the realization for the disequilibrium variable for a certain quarter is greater than the value of the threshold, the state of nature corresponds to a high regime. Meaning that, the realization of economic growth is higher than the economic growth produced by its fundamentals -Unemployment Rate-. Thus, economic growth presented a deviation from the behavior it should have had given its fundamental variables. What generated this deviation is the existence of an economic event that affected economic growth without disturbing unemployment -or at least having a change which magnitudes do not correspond to the change in economic growth-. When there is economic growth without the unemployment rate falling, the realization of the economic growth is higher than the estimated by its fundamental variables. The economic event generated growth despite the fact that its fundamentals did not reflected it.

On the other hand, when the realization of the disequilibrium variable for a certain quarter is lower than the value of the threshold, the state of nature corresponds to a low regime. A reduction in Unemployment Rate represents an increase of the estimated economic growth. However, an economic event in the economy that only affected Unemployment Rate implies the realization of the economic growth variable being lower than the value estimated by its fundamental variables. There should have being economic growth, but the event prevented it. In addition, a decrease in the realization of the economic growth given an event that does not affect the Unemployment Rate is a negative change in economic growth that is not supported by its fundamentals and is represented by a disequilibrium below the threshold -a low regime-.

Table 8. Corresponding state of nature for each quarter. The table displays the value of the disequilibrium estimated for each quarter. If this value is above the threshold -0.00899521-, the state of nature corresponds to a High regime, and if the value of the disequilibrium is below the threshold, the state of nature corresponds to a Low regime.

Table 8.

Regimes for each quarter

Quarter	Delay	State of nature	Quarter	Delay	State of nature
2002Q1	0,0047	Transition	2011Q1	0,0166	High regime
2002Q2	-0,0111	Low regime	2011Q2	0,0168	High regime
2002Q3	0,0211	High regime	2011Q3	0,0222	High regime
2002Q4	-0,0031	Low regime	2011Q4	0,0240	High regime
2003Q1	0,0149	High regime	2012Q1	0,0194	High regime
2003Q2	-0,0248	Low regime	2012Q2	0,0181	High regime
2003Q3	-0,0212	Low regime	2012Q3	0,0241	High regime
2003Q4	-0,0152	Low regime	2012Q4	0,0105	Transition
2004Q1	-0,0123	Low regime	2013Q1	0,0097	Transition
2004Q2	0,0039	Transition	2013Q2	0,0147	High regime
2004Q3	-0,0096	Transition	2013Q3	0,0190	High regime
2004Q4	-0,0362	Low regime	2013Q4	0,0168	High regime
2005Q1	0,0000	Low regime	2014Q1	0,0147	High regime
2005Q2	-0,0252	Low regime	2014Q2	0,0201	High regime
2005Q3	-0,0255	Low regime	2014Q3	0,0146	High regime
2005Q4	-0,0414	Low regime	2014Q4	0,0119	Transition
2006Q1	-0,0407	Low regime	2015Q1	0,0151	High regime
2006Q2	-0,0198	Low regime	2015Q2	0,0051	Transition
2006Q3	-0,0108	Low regime	2015Q3	0,0097	Transition
2006Q4	0,0154	High regime	2015Q4	0,0137	High regime
2007Q1	0,0191	High regime	2016Q1	-0,0048	Low regime
2007Q2	0,0155	High regime	2016Q2	0,0041	Transition
2007Q3	0,0131	High regime	2016Q3	-0,0042	Low regime
2007Q4	0,0145	High regime	2016Q4	-0,0045	Low regime
2008Q1	0,0155	High regime	2017Q1	-0,0120	Low regime
2008Q2	0,0162	High regime	2017Q2	-0,0191	Low regime
2008Q3	0,0161	High regime	2017Q3	-0,0222	Low regime
2008Q4	0,0227	High regime	2017Q4	-0,0233	Low regime
2009Q1	-0,0013	Low regime	2018Q1	-0,0288	Low regime
2009Q2	-0,0025	Low regime	2018Q2	-0,0327	Low regime
2009Q3	-0,0007	Low regime	2018Q3	-0,0278	Low regime
2009Q4	0,0061	Transition	2018Q4	-0,0284	Low regime
2010Q1	0,0085	Transition	2019Q1	-0,0225	Low regime
2010Q2	0,0013	Low regime	2019Q2	-0,0204	Low regime
2010Q3	0,0075	Transition	2019Q3	-0,0157	Low regime
2010Q4	0,0009	Low regime	2019Q4	-0,0120	Low regime

Low regime from 2002Q1 to 2006Q3

From 2002 to 2006, a new government was at the head of the country. Giraldo (2010) stipulated how this government had a plan with a mixture of strategies regarding policy making. A monetary strategy and a Keynesian one with a strong social orientation from the 1991 Political Constitution. The first came in the form of investment agreements with 55 countries and sealed agreement with 29 others on double taxation, on top of an agreement signed with the International Monetary Fund (IMF). For this last agreement, it stands out the compromise in the goal of flexibility or liberalization of the markets, the reduction of the role of the State and the strict control of inflation. However, these policies where not anything new to Colombia or Latin America.

The second strategy objective was maintaining a greater integration and social cohesion in the country, what made the Government to assume a more active role against poverty, inequality, income distribution and expansion of social and economic opportunities for the population. This produced new changes that lead to a more flexible labor market. A labor reform was approved that substantially reduced the costs of firms in the country and that aimed to generate 1 million new jobs -that ended generating only 120,000, that are still a lot of new hires-. One strong critic to this measure is that the informal jobs grew more than the actual formal jobs, which made the measure to have less impact in economic growth. On the other hand, according to this author, this strategy also materialized in the form of an increased social spending on education, health, housing, security, and spending on social cohesion and national defense -all of which are characterized as long-run measures-.

This last strategy generated a strong overall downward trend in the unemployment rate from the really high figures that existed in years prior 2002. However, this was not translated to economic growth immediately. The first strategy included a great opening to international markets and less state intervention. This transition made that the economic stability of the country stopped depending just of its internal affairs. To add on, the jobs that where created and all of the newly signed agreements where done through-out the four-year period. It took time for the effects of this measures to be notable in the economy. In fact, these effects contributed to a strong economic growth in the next period. As a result of all of the above, the unemployment rate decreased while the economic growth did not grow in the magnitude that its fundamental variables held. This represents a Low regime state of nature.

High regime from 2006Q4 to 2008Q4

This period presented the greatest economic growth in Colombia since 1978, first in 2006Q4 -7,97%- and then in 2007Q4 with 8,14%, due to the economic bases that where stablished in the prior period -among other important aspects as follows-. Both growth figures exceeded all analysts' expectations. There was no notable shock, event or decrease of the unemployment rate that sustained this growth. However, the sector that grew the most was construction and housing -14,36% in the overall sector and 17,75% in the housing sub-section compared to 2005-21. Analysts like Germán Umaña22 -in the newspaper Portafolio (2008) pointed out that this growth was un-called for. Additionally, for the next years there might occur a strong international crisis in the housing sector that they "hoped" it did not the affect severely the country. Indeed, the crisis came and confirmed that this unprecedented growth was from a bubble. Thus, the economy experimented an increase on its productivity that was not associated by any means with the labor market. An increment on the realization of the economic growth variable that was not supported by its fundamental variables, justifying a High regime state of nature. The crisis did not mean an immediate change of state of nature. It took a couple more periods to have its full effects in the Colombian economy.

Low regime from 2009Q1 to 2010Q4

In the quarters corresponding to the period between 2009 and 2011, a Low regime took place. In 2008, the whole world was shaken by the subprime crisis in United States. The retaliation of this great shock reached the Latin American countries a few periods later, justifying the change of regime in 2009Q1. Mesa et al,.(2008) developed how in 2008 the economic growth shifted from the figure of 2007 -the year closed with an economic growth of 2,5% compared with the 7,5% figure of the year before-. The authors justify this behavior in inflation and a lower spending disposition by households and firms. However, the economy kept growing -in a lower rate, but growing anyway-.

From the last quarter of 2008 to all the quarters of 2009 the change in the economic growth was negative, due to the aggravation of the previous conjuncture. The low regime took place because the crisis was financial, and its greatest problem was inflation -a variable that is not considered in OL-. The effects of this crisis did not incentivize a negative detriment to the unemployment rate. A contraction of the realization of the economic growth variable that is not supported by its fundamental variables, shows a state of nature that corresponds to a Low regime.

High regime from 2011Q1 to 2014Q4

The Federal Reserve (FED), the Central Bank of United States, to counter the effects of the subprime crisis, guaranteed liquidity facilities known as a whole as "Quantitative Easing". According to Leyva et al,.(2016), it consisted of three fundamental measures. The first where liquidity measures with which the FED gave loans

[21] All data was collected from the National Administrative Department of Statistics (DANE) of Colombia.
[22] Portafolio (2008). General Director of the "Centro de Investigaciones para el Desarrollo -CID-" of the "Universidad Nacional".

to different banks and financial entities preventing any bankruptcy. The second had the objective of decreasing any uncertainty in the market through making public any monetary policy -or expectative of one-. The final one consisted of lowering the long-term interest rate to incentivize investment, and the purchase of several financial assets in a large scale. This last one is the most relevant one because it introduced a large amount of dollars to the financial markets. This measures, jointly with the weakening of the dollar generated by the crisis, produced that the Colombian peso strengthened against the American dollar and represented an increase in foreign investment.

Additionally, between 2011Q1 and 2014Q3 the prices of oil where over 100 American dollars per barrel. Given that, Colombia is a country which its income strongly depends on oil, which is its first export good, this represented a boom in economic growth that was not related at all with the labor market. Although this increase of income also influenced the overall tendency of the unemployment rate to decrease, this reduction of the unemployment rate did not generate the expansion of the economic growth, and the magnitude of its change did not support the bigger increment of the economic growth. The above reflected a high regime as the state of nature of this period.

Low regime from 2015Q1 to 2019Q2

This Low regime is characterized by important alterations in the Colombian labor market. Indeed, since the late nineties, numerous aspects of the labor market where reformed. Initially, there was a clear improvement in matching efficiency between companies and workers. Arango et al. (2016) postulated that there existed certain indirect costs of search for employment that have been reduced due to the introduction of technology information systems, public and private employment agencies. These types of agencies are adscript to the Ministry of Labor, and their main objective is to improve the corresponding interaction between companies and workers. Secondly, there is evidence that major encouragements for the creation of Small and Medium Enterprises (SMEs) where introduced to the market, which according to Varela (2017) is one of the core sources of employment improvement. It is mentioned by Flórez et al. (2018) that the important institutional changes in the labor market is determinant of a Low regime and may produce a more flexible labor market.

However, these was not necessary translated to an immediate labor market improvement. Indeed, despite that the country created a network of entities responsible for bringing labor supply and demand closer together -which in 2018 reached almost 360 entities and 800 service points throughout the country-, Colombia still does not have consolidated information on all the supply of services. Accordingly, to the INEI (2018) this network of providers of the Public Employment Services, which includes public and private providers, registered 1,105,312 search engines and 1,892,418 vacancies. Regarding the intermediation offer financed with public authorities' resources, around $ 364 billion Colombian pesos were allocated in group orientation and training sessions for job seekers. In fact, they only managed to employ 362,000 people out of the 1.8 million that the search engines registered. Even though, a great effort was made to reduce the cost of the search for employment. Since 2010 the details of job applicants and vacancies, the count of the databases on beneficiaries per single consolidated user or the microdata of the beneficiaries of each of the activities contemplated within the intermediation services, in particular those financed with public resources, have not been well consolidated nor have had the expected impact on the labor market.

The positive effects of all of these aspects of the labor market that where reformed started to shown in 2015. It was the first time in this century that in a single year, for eleven consecutive months the unemployment rate was under 10%. In fact, the last half of 2014 showed this tendency, being reflected with its last quarter being the transition from a high to a low regime. According to Banco de la República (2019), the urban occupancy rate and the global participation rate from 2016 to 2019 showed a sustained decrease. These two have a direct implication on the unemployment rate, whose behavior has been diminishing during this period of time.

To add on, after the mentioned transition from the higher regime to the lower regime from (2015Q1 to 2019Q4) there existed a clear shock regarding the reductions of oil prices. As a matter of fact, the study conducted by Tapia (2018) showed that for a total of 23 cities analyzed in Colombia, occupations respond positively to changes in oil prices; while unemployment, informality and underemployment have an inverse relationship with the occupation rate. However, according to Basque Trade and Investment (2020), the National Government estimated that for every dollar that oil price dropped in a sustained way in a year, tax revenues are reduced by $400,000 million pesos annually. All this information suggests that the low regime is supported by an overall decreasing tendency of the unemployment rate, while economic events such as the oil price shock held back economic growth.

6. Conclusions

This research estimated Okun's Law (OL) for Colombia using the STAR model procedure, adapted to the long-run cointegration determined for the variables. The ECM model was elaborated in order to guarantee the stationarity requirement for economic growth and the unemployment rate. The linearity tests indicated that the usage of the disequilibrium term as the delay variable implies a non-linearity approach that should be consider. The selection of the delay variable for the disequilibrium was quite interesting for the analysis: the disequilibrium term characterizes deviations of the economic growth and unemployment rate from the long-run equilibrium relationship. Finally, Arthur Okun linearity assumption was annulled by the disequilibrium of the Okun's Law ECM model in Colombia.

To summarize our results: For all of the considered variables, the estimation of the parameters was found significative at a 5% significance level. Additionally, the model apprehended two different states of nature of the short and long run dynamics of the OL relationship. The threshold represents the value that the delay variable must over-pass in order for the state of nature to correspond to a High Regime, or the value it must stay below to correspond to a Low Regime. Initially, the High Regime is defined by a realization of the economic growth that is higher than the estimated by its fundamental variables. In other words, there was no proportional reduction in the unemployment rate or the unemployment rate increased but the change was not reflected negatively in the realization of the economic growth. Secondly, the Low Regime relies on the fact that the realization of the economic growth is lower than the value estimated by its fundamental variables. That means a reduction in the unemployment rate, which represents an increase of the estimated economic growth but was not translated to the realization of the economic growth. Otherwise, the effects of certain events -that did not affect unemployment rate- produced an estimation of the fundamentals higher than the realization of the economic growth.

Given all of the above, a specific regime was stablished for different periods of time. The first period corresponds to a Low regime from 2002Q1 to 2006Q3. In this period, a new government implemented strategies allowing the expansion of social and economic opportunities for the population mainly via job creation -among others-. This generated a fall of the unemployment rate that was not backed by economic growth given that the positive effects on this variable where perceived in later periods. This period was followed by a High Regime from 2006Q4 to 2008Q4 that was a result of the materialization of the previously mentioned positive effects and the bubble of the housing sector -both events increased economic growth without affecting the unemployment rate-. The subprime crisis resulted in inflation and a lower spending disposition by households and firms. This decreased economic growth, but neither these variables presented an immediate effect in the unemployment rate, resulting in a period of a Low regime from 2009Q1 to 2010Q4. Subsequently, from 2011Q1 to 2014Q4 a combination of high oil prices with a strengthened Colombian peso against the

American dollar, generated an economic growth that was not accompanied by a diminution in the unemployment rate, ensuing a High regime. Finally, a Low Regime was defined for the period between 2015Q1 and 2019Q4 as a result of alterations in the Colombian labor market that lower the unemployment rate and an oil price crisis that decreased -and for the most part deaccelerated- economic growth.

In conclusion, this research aims to continue improving the simplicity of the original linear relationship stablished by Okun. OL is frequently used to measure the cost to society of excessive unemployment. Certainly, generating an upward trend of economic growth is a priority for Colombia. A correct identification of the non-linear characteristics of the relationship, might help the construction of policies that could contribute to its actual conjuncture. This, in the light of the conclusion of the present document, relays on the fact that job generation indeed produces economic growth.[23]

---

[23] For information regarding job generation, its fundamental variables -such as education- and social inequality, consult the following (respectively):

Stewart, A., & Stanford, J. (2017). Regulating work in the gig economy: What are the options? The Economic and Labour Relations Review, 28(3), 420–437. doi:10.1177/1035304617722461

Borzaga, C., Salvatori, G., & Bodini, R. (2019). Social and Solidarity Economy and the Future of Work*. This paper draws on a work that was previously published by the ILO.

Eslava, M., Haltiwanger, J., & Pinzón, A. (2020). Job creation in Colombia vs the U.S.:b"up or out dynamics" meets "the life cycle of plants". Working Paper 25550 http://www.nber.org/papers/w25550

Stiglitz, J., & Greenwald, B., (2014). Creating a Learning Society. A New Approach to Growth, Development, and Social Progress.

Voronkova, O., Hordei, O., Barusman, A.R.P., Ghani, E.K. (2019). Social Integration As A Direction For Humanization Of Economic Relations And Improvement Of Social Welfare. SocioEconomic Challenges, 3(4), 52-62. http://doi.org/10.21272/sec.3(4).52-62.2019.

7. References

Arango, L. E., & Flórez, L. A. (2020). Determinants of structural unemployment in Colombia: a search approach. Empirical Economics, 58(5), 2431–2464. https://doi.org/10.1007/s00181-018-1572-y

Attfield, C. L. F., & Silverstone, B. (1998). Okun's law, cointegration and gap variables. Journal of Macroeconomics, 20(3), 625–637. https://doi.org/10.1016/S0164-0704(98)00076-7

Basque Trade & Investment. (2020). Informe impacto de la Covid-19 en Colombia. Basque Trade & Investment, 11. https://www.semana.com/nacion/articulo/coronavirus-tiene-la-economia-colombiana-en-emergencia-cuales-son-las-opciones/661550

Banco de la República. (2019). Tendencia a la baja de la demanda laboral: pausa en las ciudades y continúa en las áreas rurales. Repositorio.Banrep.Gov.Co. https://publicaciones.banrepcultural

Campbell, J. R., & Fisher, J. D. M. (2000). Aggregate employment fluctuations with microeconomic asymmetries. American Economic Review, 90(5), 1323–1345. https://doi.org/10.1257/aer.90.5.1323

Campos, J., Ericsson, N. R., & Hendry, D. F. (2005). General-to-specific Modeling : An Overview and Selected Bibliography. International Finance Discussion Paper, 2005(838), 1–94. https://doi.org/10.17016/ifdp.2005.838

Christopoulos, D., McAdam, P., & Tzavalis, E. (2019). Exploring Okun's law asymmetry: an endogenous threshold LSTR approach. ECB Working Paper, 2345.

Dickey, D. A., y Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association, 74 (366a), 427–431.

Dijk, J. (1999). Smooth Transition Models: Extensions and Outlier Robust Inference.

Dijk, D. v., Teräsvirta, T., y Franses, P. H. (2002). Smooth transition autoregressive models—a survey of recent developments. Econometric reviews, 21 (1), 1–47.

Ekner, L. E. (2014). Cointegration and Regime Switching Dynamics in Macroeconomic Applications.

Enders, W. (1995). Applied Econometric Time Series. In Technometrics (Vol. 37, Issue 4, pp. 469–470). https://doi.org/10.1080/00401706.1995.10484400

Engle, Robert F.; Granger, Clive W. J. (1987). "Co-integration and error correction: Representation, estimation and testing" (PDF). Econometrica. 55 (2): 251–276. doi:10.2307/1913236. JSTOR 1913236.

Escribano, Á., & Jordá, O. (2001). Testing nonlinearity: Decision rules for selecting between logistic and exponential STAR models. Spanish Economic Review, 3(3), 193–209. https://doi.org/10.1007/PL00011442

Flórez, L. A., Pulido-Mahecha, K. L., & Ramos-Veloza, M. A. (2018). Okun´s law in Colombia: a non-linear cointegration. Borradores de Economía; No. 1039, 1039. http://repositorio.banrep.gov.co/handle/20.500.12134/7010

Giraldo Isaza, F. J. (2010). Balance económico de la administración Uribe primer periodo: 2002-2006. Apuntes Del Cenes, 27(43), 93. https://doi.org/10.19053/01203053.v27.n43.2007.213

Harris, R. (1995). Using cointegration analysis in econometric modelling. In Library.Wur.Nl. http://library.wur.nl/WebQuery/clc/933062%5Cnpapers://22880b5a-c3f7-4a61-9e33-ed4c97fcb3c7/Paper/p610

Harris, Richard, & Silverstone, B. (2001). Testing for asymmetry in Okun's law: a cross-country comparison. Economics Bulletin, 5(2), 1–13.

Hsing, Y. (1991). Unemployment and the GNP Gap: Okun's Law Revisited. Eastern Economic Journal, 17(4), 409–416.

Huang, H., and S. Lin. (2008). Smooth-Time Varying Okun's Coefficients. Economic Modelling 25: 363–375.

Instituto Nacional de Estadística e Información (INEI). (2018). Panorama General del Empleo Inclusivo. In Informe Nacional de Empleo Inclusivo INEL.

Knotek, B. E. S. (2006). How Useful is Okun's Law? 73–103.

Lasso-valderrama, B. F., & Flows, F. (2019). Forecasting the Colombian Unemployment Rate Using Labour.

Leyva-Uribe, B., Gómez-González, J. E., Valencia-Arana, O. M., & Villamizar-Villegas, M. (2016). Efectos del Quantitative Easing sobre los retornos accionarios en mercados emergentes. Borradores de Economía; No. 929. http://repositorio.banrep.gov.co/handle/20.500.12134/6240

Mayes, D. G., & Viren, M. (2002). Asymmetry and the problem of aggregation in the Euro area. Empirica, 29(1), 47–73. https://doi.org/10.1023/A:1014681414484

Mesa, C., Restrepo, D., & Aguirre, B. (2008). Crisis externa y desaceleración de la economía colombiana en 2008-2009: coyuntura y perspectivas. Perfil de Coyuntura Económica, 12, 31–67.

Misas, M. (2020). Working Paper of the Econometrics Research Group. Universidad de La Sabana.

Macaulay, F. (1931). The Smoothing of Time Series. [Pp. 172. New York: National Bureau of Economic Research, Incorporated.

Nebot, C., Beyaert, A., & García-Solanes, J. (2019). New insights into the nonlinearity of Okun's law. Economic Modelling, 82(December 2018), 202–210. https://doi.org/10.1016/j.econmod.2019.01.005

Okun, A. M. (1962). Potential GNP: its measurement and significance, Cowles Foundation Paper 190. In Cowles Foundation, Yale University: New Haven, CT, USA (pp. 98–104).

Portafolio. (n.d.). economia-colombiana-crecio-52-ciento-2007-acuerdo-dane- www.portafolio.co. https://www.portafolio.co/economia/finanzas/economia-colombiana-crecio-52-ciento-2007-acuerdo-dane-434100

Tapia, A. M. E. (2018). IMPACTOS DEL PRECIO DEL PETRÓLEO SOBRE LOS PRINCIPALES INDICADORES DE EMPLEO POR DEPARTAMENTOS EN COLOMBIA Universidad del Norte Seminario IEEC.

Teräsvirta, T., Tjøstheim, D., & Granger, C. W. J. (1994). Modelling Nonlinear Economic Time Series. Modelling Nonlinear Economic Time Series, April, 1–592. https://doi.org/10.1093/acprof:oso/9780199587148.001.0001

Tong, H. (1990). Non-linear time series: a dynamical system approach. Oxford University Press.

Varela Villegas, R. O. (2016). Colombia small- and medium-sized enterprise's 70 years of progress: what's next? Small Enterprise Research, 23(3), 302–315. https://doi.org/10.1080/13215906.2016.1269241

Yaya, O. O. S., & Shittu, O. I. (2016). Symmetric variants of logistic smooth transition autoregressive models: Monte Carlo evidences. Journal of Modern Applied Statistical Methods, 15(1), 711–737. https://doi.org/10.22237/jmasm/1462077240

8. Appendix

A.1 Descriptive Statistics[24]

Variable	Observations	Mean	Std. Dev.	Min	Max
GDP	76	164153.3	36238.6	108906.7	223097.7
Unemployment	76	11.41	2.14	8.01	16.56

A.2 Dickey-Fuller Critical Values

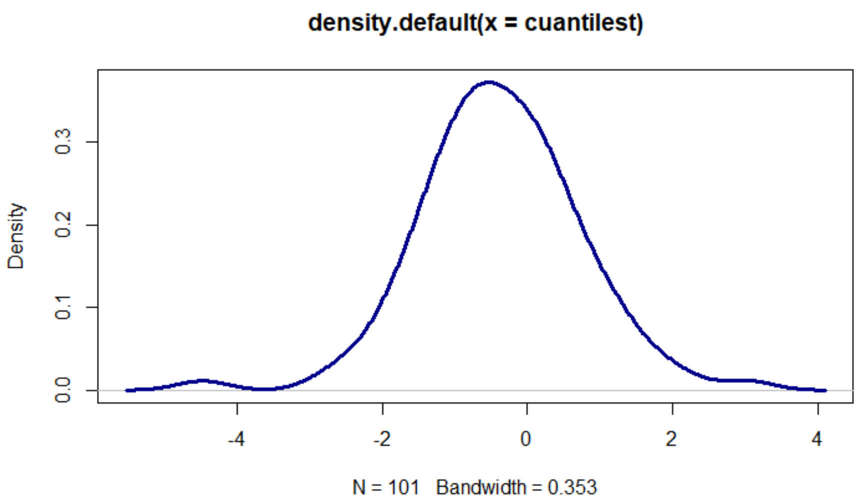
	$\tau\mu$	$\tau\tau$	τ
Sample Size	5%	5%	5%
50	-2.93	-3.5	-1.95
100	-2.89	-3.45	-1.95

A.3 Engle-Yoo Critical Values

Two Variables		
Sample Size	1%	5%
50	-4.12	-3.46
100	-4	-3.39

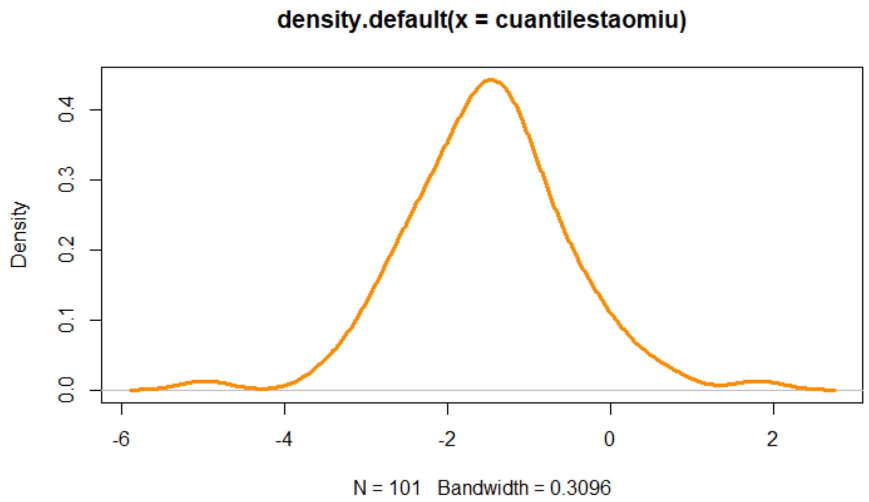
A.4 Simulated Dickey-Fuller Distributions

Model 1 (τ)



[24] Is important to clarify that the observations for the A.1 Appendix are 76 because the descriptive statistics where conducted for the original time series without transformations or any other further modifications.

Model 2 ($\boldsymbol{\tau\mu}$)



Model 3 ($\boldsymbol{\tau\tau}$)