IPS
Investigación en problemáticas sociales

Artículo original

KUZNETS HYPOTHESIS VERIFICATION: COLOMBIA, 1880 – 2018

Sergio Rodríguez Cifuentes
Estudiante Promoción 2020. Gimnasio Campestre

Juan Zambrano Barón
Asesor investigación de grado.
Gimnasio Campestre

Correspondencia para los autores:
jzambrano@campestre.edu.co,

Recibido: 15 de septiembre de 2021
Aceptado: 8 de octubre de 2021

KUZNETS HYPOTHESIS VERIFICATION

Fotografía: Camilo Constain. Diseño: Gimnasio Campestre

Source https://www.nytimes.com/2020/07/11/world/americas/coronavirus-latin-america-inequality.html

SUMMARY

This article explores the use of the Kuznets Hypothesis in the Colombian scenario for the period 1880-2018 in order to establish a relation between changes in the Gross Domestic Product per Capita of Colombia and its impact on how income inequality behaves nationwide using the Gini Coefficient as a reference. This research approaches inequality through an econometric scope by applying mathematical modeling techniques, integration, discrete comparisons with Multidimensional Poverty, alternative hypotheses, and economic history to comprehend and relate its behavior with macroeconomic variables. The results point out that there might be considered multiple Kuznets Curves for the stated period, which would discard the validity of the Kuznets Hypothesis, thus suggesting that the behavior of inequality is cyclic and therefore may correspond to Kuznets Waves. The research further links Multidimensional Poverty reductions with inequality reductions and economic growth and recalculates the Gini Coefficient using alternative methodologies to verify the accuracy of the indexes.

Key words: Kuznets Hypothesis, Inequality, Gini Coefficient, Gross Domestic Product per Capita, Poverty, Kuznets Waves, macroeconomics, Mathematical Modeling.

RESUMEN

Este artículo explora el uso de la Hipótesis de Kuznets en el escenario colombiano para el período 1880- 2018 a fin de establecer una relación entre los cambios en el Producto Interno Bruto per Cápita de Colombia y su impacto en cómo la desigualdad de ingresos se comporta a nivel nacional utilizando como referencia el Coeficiente de Gini. Esta investigación aborda la desigualdad a través de una perspectiva econométrica mediante la aplicación de técnicas de modelado matemático, integración, comparaciones discretas con la Pobreza Multidimensional, hipótesis alternativas e historia económica para comprender y relacionar su comportamiento con variables macroeconómicas. Los resultados indican que podrían ser consideradas múltiples curvas de Kuznets para el período indicado, lo cual descartaría la validez de la Hipótesis de Kuznets, sugiriendo que el comportamiento de la desigualdad es cíclico y por lo tanto puede corresponder a las Olas de Kuznets. Adicionalmente, la investigación vincula reducciones de la Pobreza Multidimensional con reducciones de la desigualdad y crecimiento económico, y recalcula el Coeficiente de Gini utilizando metodologías alternativas para verificar la exactitud de los índices.

Palabras clave: Hipótesis de Kuznets, Desigualdad, Coeficiente de Gini, Producto Interno Bruto per Cápita, Pobreza, Olas de Kuznets, Macroeconomía, Modelación Matemática.Bruto per Cápita, Pobreza, Olas de Kuznets, Macroeconomía, Modelación Matemática.

Source: https://freepik.es

INTRODUCTION

Market economies are complex systems in which multiple factors shape the outcomes of entire societies. Understood as one of those systems, Colombia is a country of immense economic and social inequalities among its citizens. Despite recent reductions on inequity indicators, there is still evidence of population living with less than the minimum wage yet coexisting with one percent that concentrates a large share of the national wealth. Although inequality could be a natural phenomenon, it is also the result of several interactions that widen differences among citizens, such as income gaps, levels of education, and governmental willingness, among other factors that hinder an exhaustive study.

Despite the previous, the use of mathematical and statistical approaches to analyze inequalities offers new perspectives on sociological and economic issues: focusing on spatial asymmetries and their behavior, and using chronological data, namely the Gross Domestic Product (GDP) per Capita, the Gini Coefficient (GC) of income distribution and the Kuznets Hypothesis (KH), it is possible to identify historical and future trends on the behavior income inequality may present by relating variations of income disparity with the Colombian development process.

To approach the query previously stated, the main purpose of this investigation will be to verify if the Kuznets Hypothesis is valid to Colombia by using mathematical modeling approaches, relating historical data of the Gross Domestic Product per capita and the Gini Coefficient of income distribution since 1880 until 2018 through the consideration of an alternative econometric model capable of incorporating several macroeconomic variables. To achieve so, this article understands and analyzes poverty through a holistic approach transcending the traditional notions of income and monetary-based poverty to determine the relationship between income inequality and the Gross Domestic Product per capita of Colombia since during the mentioned time frame and to use econometric techniques as tools for examining the short- and long-term effects of the macroeconomic variables entailed.

Source: https://freepik.es

THEORETICAL FRAMEWORK

2.1. Income

Orthodox approximations to the concept of income tend to understand it as the flow of a service in a given amount of time (Wallace Hewett, 1925). It ranges to over revenue earned from jobs (wages, pensions, incentives, income from self-employed employees, etc.) but also, investments (such as returns on savings plans and equity dividends), government grants, pension income – state, business, or personal pension –, and property income. It can also include contributions from other families (such as childcare or remittances), and home development.

Simon Kuznets (1934) was pioneer in incorporating the income concept to a national level and integrated a yearly time frame, which he defined as the sum of all items manufactured and all direct services provided during the year added to their market value, while subtracting the amount of that portion of the national stock of products invested (both as raw materials and as capital equipment) in the production of that total, so the balance forms the net income of the national economy during the year. He developed the notion that national income can be estimated as the sum of a series of incomes (labor incomes, property incomes, and entrepreneurial incomes) minus any government subsidy or tax to these incomes. This statistic became the basis of the Gross Domestic Product (GDP).

2.1.1. GDP & GDP per Capita.

The Gross Domestic Product (GDP) calculates the monetary value of finished products and services — that is, those purchased by the end-user — produced in a country for a fixed period. It counts all production produced within a country's borders. The GDP consists of products and services manufactured for sale on the market and incorporates certain non-market activities, like services offered by the government. It is usually expressed in a determined currency of a certain year in order to have reference to its value. In most cases this currency is United States Purchasing Power Parity (PPP) Dollars (Callen, 2008). It can be calculated by adding the consumption (C), investment (I), government spending (G), and net exports (NX) of a nation:

$$GDP = C + I + G + NX$$

The **Gross Domestic Product per Capita** (GDP per Capita) is a key predictor of economic success and is widely used as a general measure of average living conditions or economic well-being, including certain recognized deficiencies in a given population. It calculates the amount of overall economic production relative to the population of a country (OECD, 2019). It can be obtained by the division of the GDP of a population by the population size (***n***):

$$GDP\ per\ Capita = \frac{GDP}{n}$$

2.2. Poverty.

2.2.1. Multidimensional Poverty.

There is a growing consensus that poverty is a political construction, and hence its calculation should not be reduced to any single factor, such as wages, given that depravations can have several manifestations or dimensions. **Multidimensional Poverty Indexes** (MPI) depart from the axiom that poverty is the lack of well-being, and accordingly there has to be a concept of well-being and welfare in order to conceptualize poverty. The approach considers that income is insufficient to measure poverty because income and money are means to obtain resources to attain objectives, such as education, healthcare, and sanitation, among others. In this sense, indicators of resources can be useful proxies in order to measure poverty, and furthermore, to consider dimensions that income

cannot denote, as for instance, which capabilities and resources require intervention or further access in the future for a given population (Seth & Alkire, 2014).

Multidimensional Poverty (MP) is calculated using micro-data. This refers to the unit-level data comprising the responses given by each unit of analysis (such as the individual or household). It can be adapted to fit different criteria, and it can be calculated as the product H x A, the ratio of headcount or number of persons identified as poor (H) compounded by the average proportion of weighted suffering experienced by poor people (A), which is called the intensity of poverty. The result of this operation is called the adjusted headcount (M0).

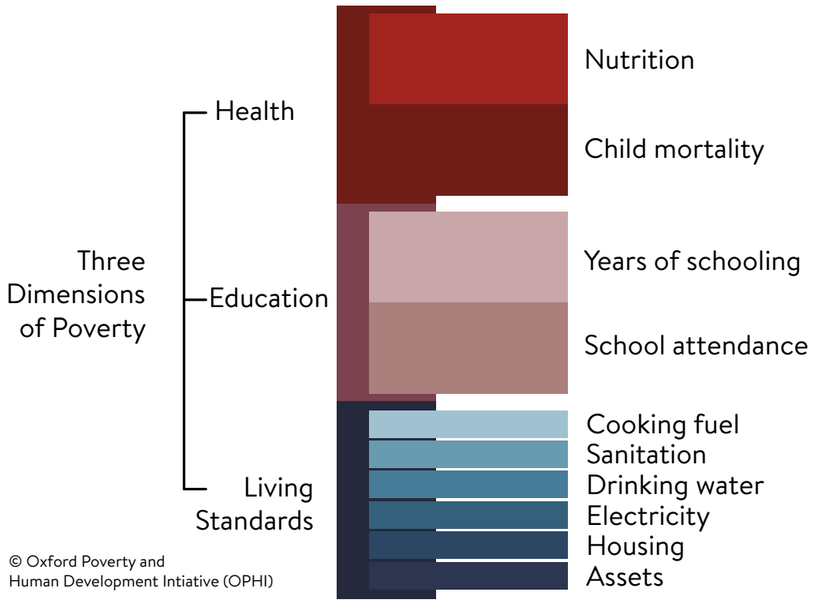


Figure 1. Indicators used by the Global MPI based on the AF Methodology. Taken from: OPHI (2018). Global Multidimensional Poverty Index 2018: The Most Detailed Picture to Date of the World's Poorest People. Oxford Poverty and Human Development Initiative, University of Oxford.

A person is identified as multidimensionally poor (or 'MPI poor') if they are deprived in at least one third of the weighted indicators shown above; in other words, the global cutoff for poverty (k) is 33.33% If a person is deprived in 20-33.3% of the weighted indicators they are considered 'Vulnerable to Poverty', and if they are deprived in 50% or more (i.e., *k=50%*), they are identified as suffering 'Severe Poverty' (Seth & Alkire, 2014).

2.3 Inequality

Inequality mainly outlines the difference among individuals or groups with respect to a specific parameter, which tends to be the uneven distribution of income and opportunities between people, demographic groups, or nations. Economic disparity is typically based on wages, wealth, and consumption differences, while several other associated inequalities, such as educational or health inequalities, may occur. Income Inequality can thus be understood as the disparities that individuals and/or households present in societies with respect to others. In other words, inequality can be treated simply as differences between the incomes of individuals or sub-groups in a larger population, hence, there is no inequality if the income is distributed equally among the population (De Maio, 2007).

Barro (2000) points out that the accumulation of income and capital in the hands of few individuals can be a positive process and result in new enterprises and higher investment in education, especially in developed countries. However, from the point of view of economic sciences, high inequality can induce under-optimal use of capital. There are also several studies that link inequality with inefficiencies in the social mobility of a country. Nations with low social mobility usually have high income inequality indicators as the flow of capital remains mostly unchanged through time (Yang, 2017).

2.3.1. Lorenz Curve.

The **Lorenz Curve** (LC) is a graphical representation of the cumulative income distribution of a population. It conveys for the bottom x_1 % of households, what percentage y_1 % of the total income they have. The percentage of households is plotted on the x–axis, the percentage of income on the y–axis, both cumulative. In the scenario that $\boldsymbol{x}_1 = \boldsymbol{y}_1$, the Lorenz Curve is a straight line which indicates perfect income equality. Accordingly, any departure from this 45-degree line will represent inequality (Lubrano, 2017).

There are several methodologies to obtain the Lorenz Curve for a given income distribution. All of them follow the fact that the Lorenz Curve $L(x)$ is the fraction of total income earned by the poorest fraction x, $0 \leq x \leq 1$, of the population, the number 100x is a percentile, and $L(0) = 0, L(1) = 1, L'(x) \geq 0, L''(x) \geq 0$ (Lubrano, 2017).

One of the techniques consists of building an approximation with a quadratic or higher degree polynomial function using data of income quantiles, such as quartiles, quintiles, or deciles of income, by adding them so that the quantiles show the cumulative income distribution, and performing a regression to obtain a continuous L(x) complying with the aforementioned parameters, aiming for an R2 near or equal to 1. A second approach corresponds to construct an appropriately smooth approximation to the underlying distribution function that matches the known data, which yields higher accuracy, yet requiring significantly more information to input (Liberati & Bellù, 2015)

2.3.2. Gini Coefficient.

The **Gini Coefficient** (GC) is a summary statistic that measures how equitably income is distributed in a population. It alludes to the difference between the actual income distribution and the ideal income distribution (perfect equality). Thus, it can be calculated referring to the Lorenz Curve, because if incomes were equally distributed, $L(x)$ would follow a 45º diagonal, which can be represented with the function $y=x$. As the degree of inequality increases, the curvature of the Lorenz Curve increases, and thus the area between the Curve and the 45º line becomes larger. Therefore, the Gini Coefficient can be defined as the ratio of the area between the Lorenz Curve and the line divided by the whole area below the 45° line (Kennedy, et al., 1996). This difference of areas between the ideal and the actual distribution can be expressed using the definite integral

$$G = 2 \int_{2}^{1} [x - L(x) \ dx]$$

in which x represents complete income equality, and $L(x)$ the actual income distribution. The integral is multiplied by a factor of 2, since the result is equal to dividing by 1/2, which arises from the area of the right triangle formed under the line of perfect equality in [0,1].

2.4 Hypotheses Relating Inequality, Macroeconomic Variables, and Income

Several hypotheses emerged in the aim of understanding how certain variables are interconnected and the impact of different phenomena in the overall state of economy. In this context, a handful of authors proposed theories seeking to establish a relationship between the behavior of income inequality and the economic development of a country.

Piketty Hypothesis

French economist Thomas Piketty establishes that, whenever the difference between the return on capital (r) and the rate of growth of production (g) increases, the share of capital in national income increases. Moreover, as capital income appears to be more unevenly allocated than labor income, an increase in capital share is likely to lead to an increase in total income (and, over time, in wealth) inequality (Piketty, 2014).

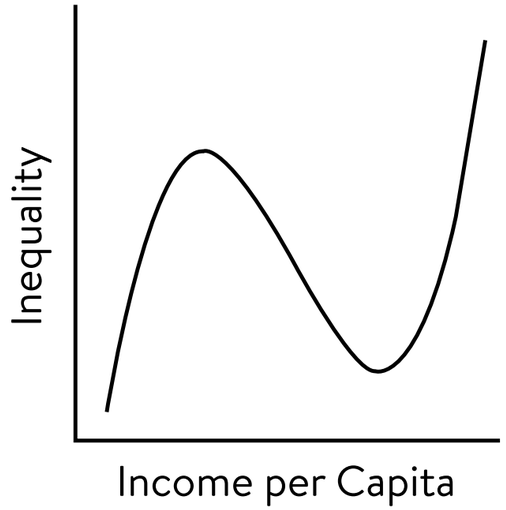


Figure 2. Visual representation of Piketty's *r&g* Hypothesis. Made by author.

Source: https://www.eluniverso.com/noticias/mundo/paro-nacional-colombia-lleva-tres-semanas-con-protestas-en-contra-del-presidente-ivan-duque-nota/

Kuznets Hypothesis

The Kuznets Hypothesis (1955), proposed by Nobel laureate Simon Kuznets, states that as a nation begins to experience economic growth, inequality will rise until a level-off point, and then it should start to reduce once the economy reaches an advanced stage of development. The implications of this would be that inequality is a natural consequence resultant from the societal transformations that occur as a nation industrializes, but also that the relationship can be constantly modeled as in the shape of an inverted U.

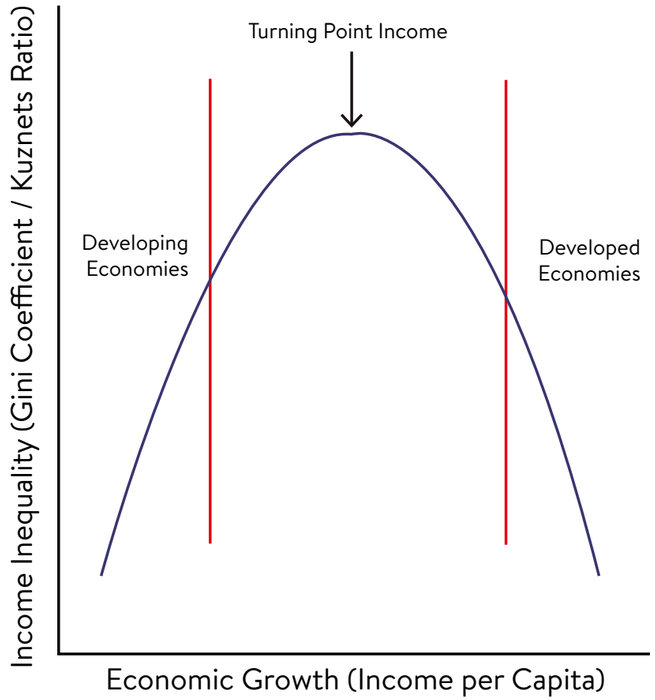


Figure 3. Visual representation of Kuznets Hypothesis. Made by author.

Kuznets' proof of his hypothesis is sustained over empirical analysis, made during mid-20th century, on income inequality and economic growth of developed countries. The reasoning behind this behavior,

according to the author, was industrialization. Under Kuznets' view, the advent of manufacturing leads to a situation in which households tend to move from the weak agricultural market, defined by a comparatively low-income inequality, to the wealthier industrial sector, where income can be less equally shared. This implies that inequality – according to this model – is a direct consequence of Economic Development.

The curve marks two moments for a country in terms of economic activity, social and geographical distribution. It matches the reasoning of W. A. Lewis' (1954) *Dual Economy Model*, in which inequality responds to development as shifts in the workforce take place as a result from the transition between an agricultural-based to a manufacturing-based economy.

Milanovic Kuznets Waves Hypothesis

In *"Global Inequality: A New Approach for the Age of Globalization"*, American economist Branco Milanovic (2016) illustrates the steady rise-and-fall trend of inequalities over the years. In the "pre-industrial" period, he argues, the patterns of inequality have essentially repeated the *Malthusian Cycles*[1], since they have taken place under conditions of quasi-stationary average income: demographic shifts have almost entirely caused the rise and fall of inequality (Duprat, 2018)

Milanovic claims that in the modern period, from the beginning of the Industrial Revolution to today, intervals of increasing and diminishing inequalities have responded to three prevailing economic forces: technology, openness, and politics ("TOP"). What he saw as the first wave of Kuznets trends in industrialized countries lasted from the beginning of the Industrial Revolution to around the 1980s. The

1.*Malthusian Cycles* are political-demographic cycles that were typical for complex premodern societies. Due to a number of mechanisms, within the premodern social systems (and some would argue even in the 21st century), population growth tended to produce a set of imbalances and strains, eventually resulting in political-demographic collapses and substantial population decline ("SAGE Reference – The SAGE Encyclopedia of Political Behavior", 2021)

change from a rural economy to an industry-based one contributed to inequalities before the maximum income disparity happened at the end of the 19th century or the beginning of the 20th century. After that, he notes that inequality behaves in cycles, driven by TOP.

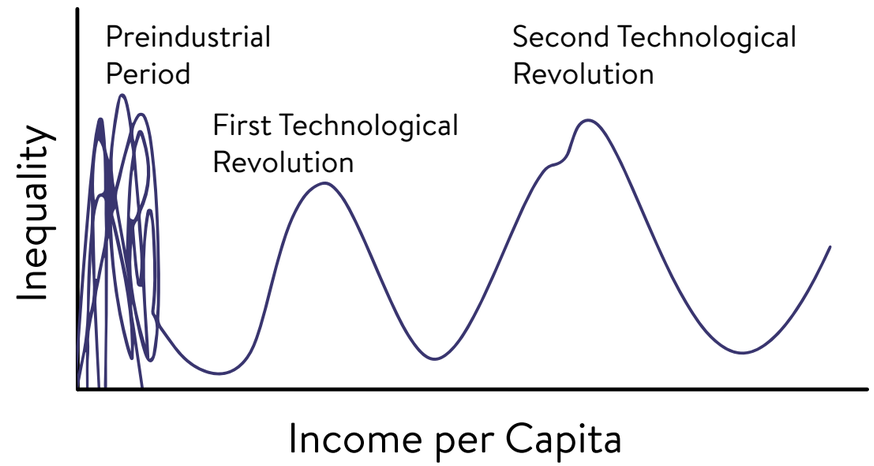
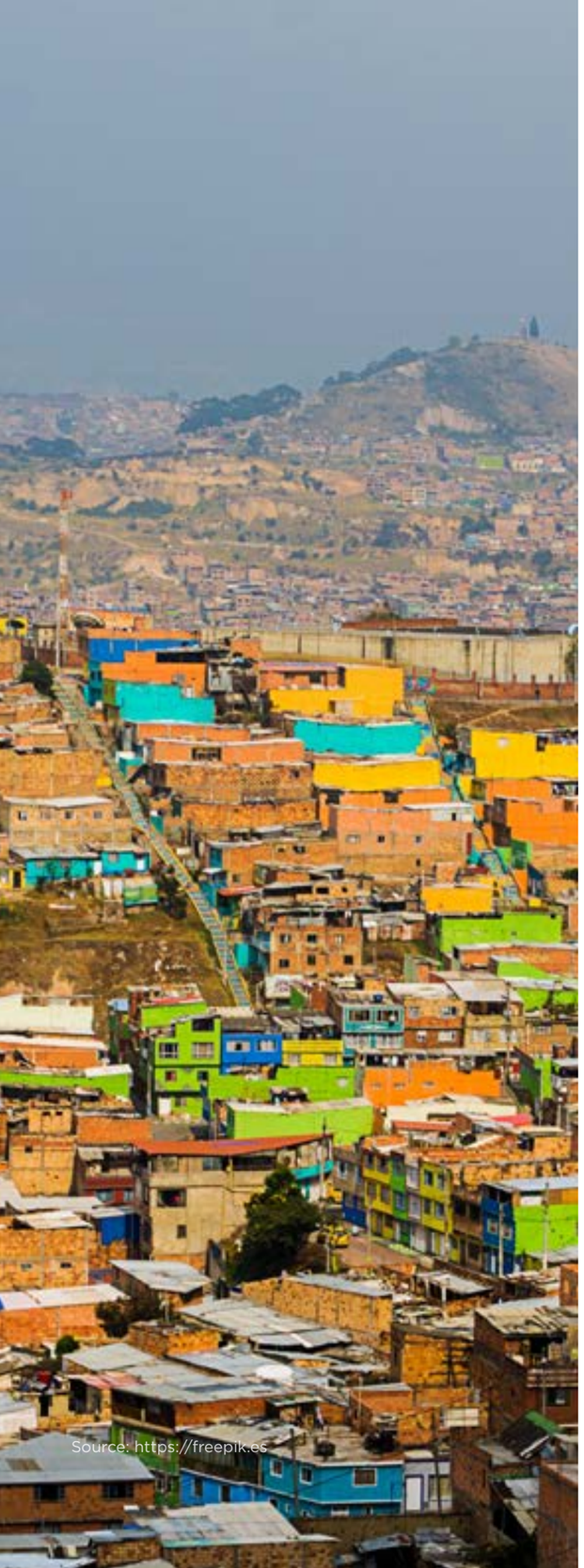


Figure 4. Visual representation of Milanovic's Kuznets's Waves Hypothesis. *Made by author.*

Milanovic (2016) suggests that modern societies have a system that is very close to what has been seen in the past, which is why there will be more curves that will cause inequality to behave as waves, driven by technological progress. This reasoning, plus the link between *Malthusian Cycles* and inequality in preindustrial societies explain a shift in the patterns of income distribution variations that will make inequality to present itself in waves in the long run. Milanovic uses the same methodology as Kuznets to prove the existence of Kuznets Waves in modern societies, which is to plot income per capita and inequality, yet arguing that the waves must be discerned analytically, and cannot be modeled using a sole function (Wardhana, 2020).


Source: https://freepik.es

METHODOLOGICAL FRAMEWORK

To test the Kuznets Hypothesis (KH), this research will use GDP per Capita in the *x*–axis, while the *y*–axis will present the Gini Coefficient, as inequality is considered dependent of Income per Capita. In order to test the relationship among the two variables, the research will use a similar approach as proposed by Naguib (2017) and Pini (2014), which will consist of evaluating the Kuznets Hypothesis empirically, directly relating these two variables using a regression. The procedure will be to plot each income, in ascendent order, to its corresponding Gini Coefficient, because as Kuznets (1955) proposes, the base assumption is that income will progress as time passes, thereafter allowing to do not consider time as a variable in this exercise. Once plotted, a quadratic regression using the Least Squares Method (Angeles, 2010) will be performed in order to search for a Kuznets Curve (KC), which must fulfill two mathematical criteria: an R^2 close to 1, and a resultant function with a negative leading coefficient, to obtain an inverted U-curved parabola.

If the regression does not yield a quadratic polynomial with the abovementioned parameters, it will then result necessary to look for Kuznets Curves in different time intervals, which would allow to see if there might be Kuznets Waves (KW), as established by Branco Milanovic (2016). To test for the KW, the same methodology will be applied as for looking for the general Kuznets Curve. Regression will also be used to find the line of best fit that most accurately models the relationship between Income per Capita and Income Inequality for designed intervals.

The Data for testing the Kuznets Hypothesis, and to search for Kuznets Curves in Colombian Economic History between 1880 and 2018, will be taken from the *Maddison Project*, which is available, updated in 2020, in the research of Bolt & van Zanden (2020). This dataset meets two conditions: it is designed to represent accurately the foreign variations in GDP per Capita (making use of the state-of-the-art estimates) and it summarizes the details available on historical trends of growth and decline in the best possible way. Since the methodology is the same for the whole time series in the case of GC and GDP per Capita, it means that the data is consistently calculated and placed in comparable standards. Data was selected from the 1880s until 2018, because it is the longest interval with continuous records for both variables.

In order to include poverty and understand its relationship with the KH or with the pattern that inequality exhibits in relation to GDP per Capita, a similar methodology to Santos et al. (2019) will be used, which will consist of making discrete comparisons between the MPI and the Inequality levels and trends reported for the given years. The MPI figure data was retrieved from the *Socio-Economic Database for Latin America and the Caribbean (SEDLAC and The World Bank*). This database was selected because it contains the most extensive data regarding Multidimensional Poverty, tracing it back from 2001 until 2016. This poverty indicator uses the Alkire-Foster methodology: a household is considered MPI poor if they perform badly in 33% of the indicators or more (as theoretically presented before). SEDLAC considers eight criteria: house rooms, house location, house materials, water and sewage, restroom, education of the children, household head education, and education+earners.

Additionally, to verify the accuracy of the Gini Coefficients used, this research utilizes data from the Colombian Government *Departamento Nacional de Planeación (DNP)*, and recalculates it using Gerber's (2007) methodology by constructing the Lorenz Curves of six years (1976-1978 and 1994-1996) using quintic polynomial regression, and then finding the Gini Coefficient by taking twice the integral of the difference of areas between the line of perfect equality and the constructed functions.

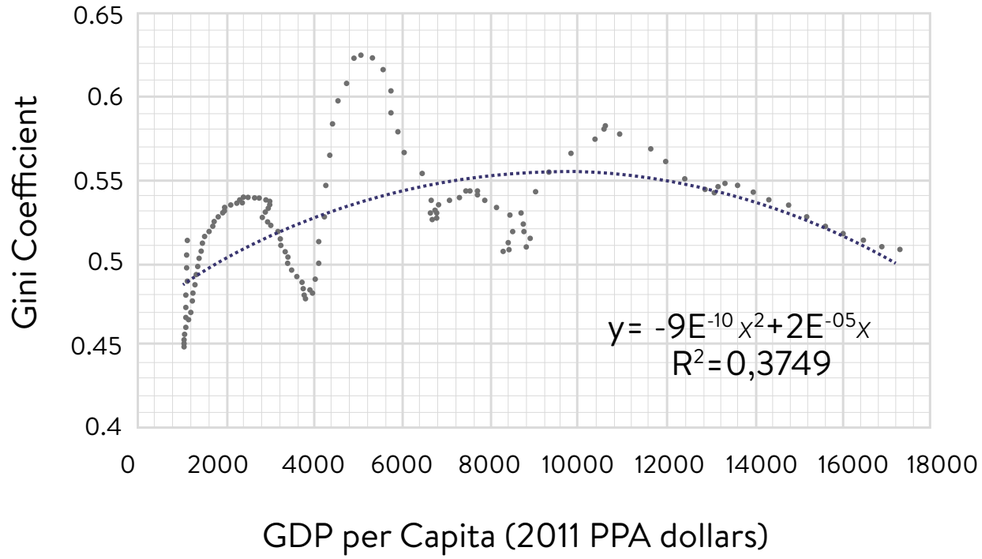
Source: https://freepik.es

RESULTS

Assessment of the Relationship between Income Inequality and GDP per Capita in the Long Run for Colombia (1880-2018).

Plotting the Colombian GDP per Capita in the *x*–axis and the corresponding Gini Coefficient in the *y*–axis of a cartesian plane yields the graph below.

KH: Relationship Between Income per Capita (GDP per Capita) and Income Inequality (Gini Coefficient) in Colombia 1880-2018



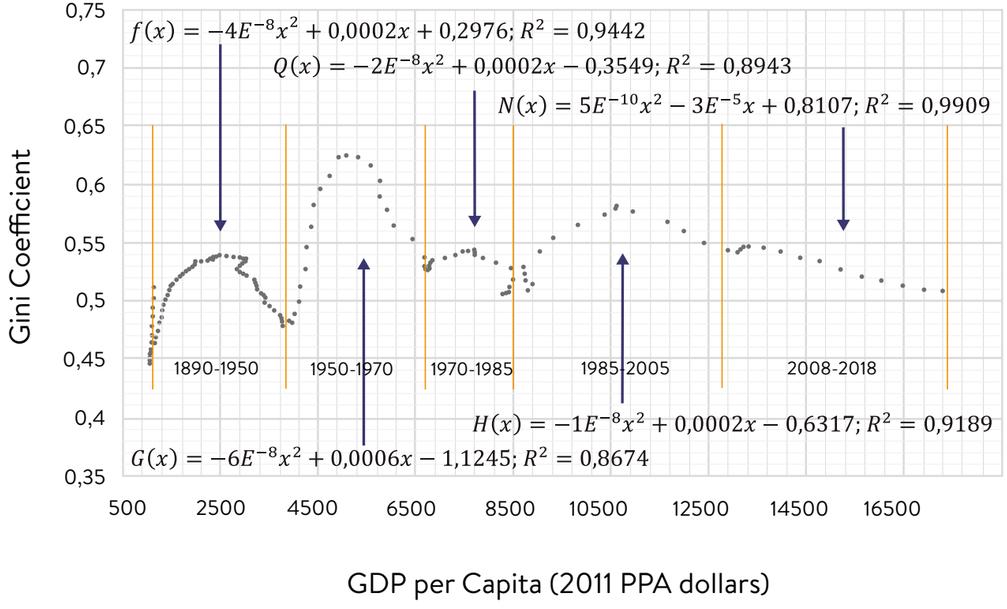
Applying a quadratic regression using the least squares method considering *x* income (GDP per Capita), and *K(x)* Income Inequality (Gini Coefficient) yields the function $K(x)= -9E^{(-10)} x^2 + 2E^{(-5)} x + 0{,}4709$ with an R^2 of *0,3749*.

Specific Time Intervals of the Relationship Between Income and Income Inequality in Colombia: Kuznets Hypothesis

The application of a quadratic regression to different time intervals looking for Kuznets Curves, which would exist if the relationship between income and income inequality exhibits the behavior of an inverted U-curve, drives to a second-degree polynomial with a downward concavity (a function of the form ax^2+bx+c in which $a<0$)

The quadratic functions resulting for each of the intervals can be observed in the graph below, further detailed in the table with their corresponding *R*–squared

KH: Relationship Between Income per Capita (GDP per Capita) and Income Inequality (Gini Coefficient) in Colombia 1880-2018



Interval	Length (year)	Function of Quadratic Regression	R-squared
1890-1950	60	$f(x) = -4E^{-8}x^2 + 0{,}0002x + 0{,}2976$	0,9442
1950-1970	20	$G(x) = -6E^{-8}x^2 + 0{,}0006x - 1{,}1245$	0,8674
1970-1985	15	$Q(x) = -2E^{-8}x^2 + 0{,}0002x - 0{,}3549$	0,8943
1985-2005	20	$H(x) = -1E^{-8}x^2 + 0{,}0002x - 0{,}6317$	0,9189
2008-2018	10	$N(x) = 5E^{-10}x^2 - 3E^{-5}x + 0{,}8107$	0,9909

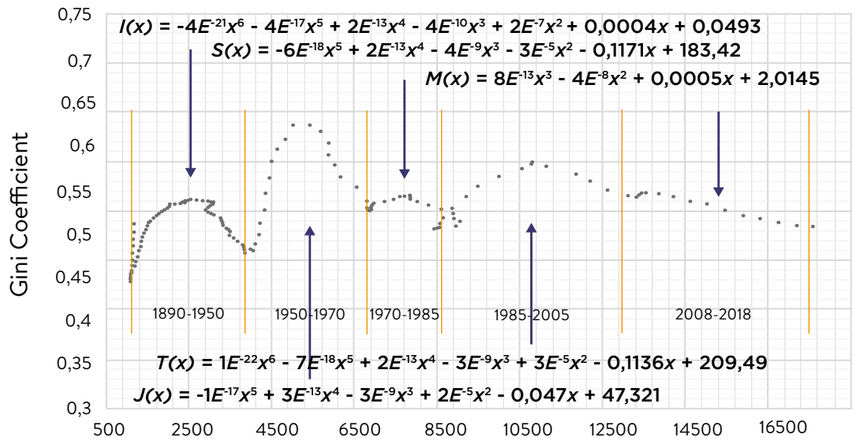
Table 1. Intervals created and regressions performed to look for Kuznets' Curves.

Specific Time Intervals of The Relationship Between Income and Income Inequality in Colombia: Best-Fit Lines

In this section we will identify the line of best fit for each of the previously selected intervals presented in Table 1, plus the additional interval that models the behavior of the inequality until 2018 in order to be able to make predictions about future inequality trends. The best-fit line will be the one that yields the closest R2 to 1.

The functions resultant for each of the intervals can be observed in the graph below, further listed in the table with their respective *R*–squared.

KC: Relationship Between Income per Capita (GDP per Capita) and Income Inequality (Gini Coefficient) in Colombia 1880-2018



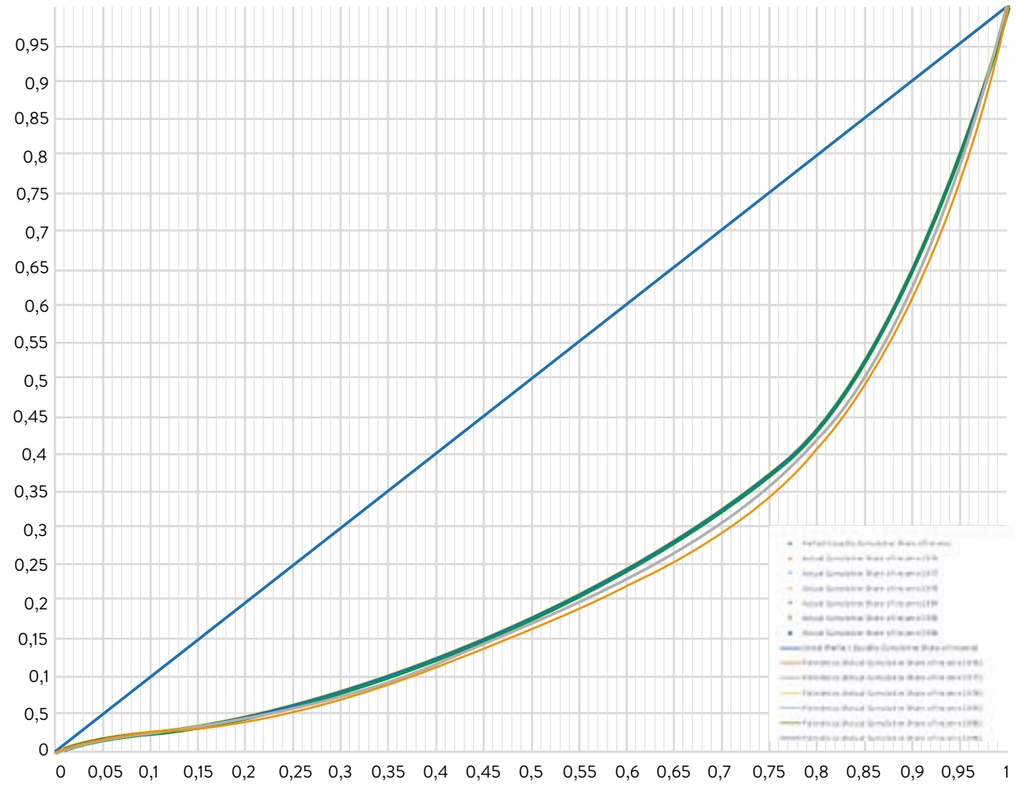
Interval	Length (year)	Function of Regression	R-squared
1890-1950	60	$I(x) = -4E^{-21}x^6 - 4E^{-17}x^5 + 2E^{-13}x^4 - 4E\text{-}10x^3 + 2E^{-7}x^2 + 0{,}0004x$	0,9892
1950-1970	20	$J(x) = -1E^{-17}x^5 + 3E^{-13}x^4 - 3E^{-9}x^3 + 2E^{-5}x^2 - 0{,}047x + 47{,}321$	0,9835
1970-1985	15	$S(x) = -6E^{-18}x^5 + 2E\text{-}^{13}x^4 - 4E^{-9}X^3 + 3E\text{-}^{5}x^2 + 0{,}1171x + 183{,}42$	0,9336
1985-2005	20	$T(x) = 1E^{-22}x^6 - 7E^{-18}x^5 + 2E^{-13}x^4 - 3E^{-9}x^3 + 3E^{-5}x^2 - 0{,}1136x + 209{,}49$	0,95
2008-2018	10	$M(x) = 8E^{-13}x^3 - 4E^{-8}x^2 + 0{,}0005x - 2{,}0145$	0,9981

Table 2. Intervals created and regressions performed to look for lines of best fit lines.

Construction of Lorenz Curves (LC) and Calculation of Gini Coefficients (GC)

The Lorenz Curves created can be observed in the graph below. The line of equality ($y=x$) is included in blue. In the x–axis of the graph can be found the cumulative share of the population, and in the y–axis the cumulative share of income, both of them in a 0 to 1 scale.

Constructed Lorenz Curves: Colombia, 1976-1978 & 1994-1996



The table below presents the Gini Coefficient extrapolated from each of the Lorenz Curves created, alongside the official GC calculated by the Colombian Government, and the simple percent error for the experimental measurements.

Year	GC From Integration of Constructed LC	Yearly GC Calculated by DNP	Error	% Error
1976	0,52305	0,52373	0,00129	0,129
1977	0,50905	0,50943	0,00074	0,074
1978	0,49713	0,49795	0,00164	0,164
1994	0,49492	0,49653	0,00324	0,324
1995	0,49675	0,49986	0,00622	0,622
1996	0,49390	0,49844	0,00912	0,912

Table 3. Gini Coefficients calculated by author compared with government Gini calculations for Colombia of 6 years.

Source: Nathalia Angarita. https://expansion.mx/mundo/2021/05/12/colombianos-toman-las-calles-tercera-jornada-paro-nacional

ANALYSIS OF RESULTS

The first plot presented in the results, which modelled Income Inequality (Gini Coefficient) as a function of Income per Capita (GDP per Capita), demonstrates that for the period 1880-2018 there is no clear Kuznets Curve (KC). Even before running the quadratic regression through the dataset, it is visually evident that the shape of the distribution is not homogeneous, and moreover, the way it presents fluctuations does not exhibit one sole peak of inequality, but rather a series of peaks after certain time passes, and for the most recent years, the behavior appears to be more of that of a straight line than of an inverted U-curve.

Conclusively, the quadratic regression confirms the previously stated. Although the function obtained complies with the parameter of the downward concavity required for having an inverted U-shaped graph, given that the leading coefficient (a) has a negative sign, the coefficient of determination (R^2) was of 0,3749, meaning that few of the behavior that income inequality presents in the long run could be attributed to changes in per Capita Income. Moreover, it implies that the quadratic model proposed does not provide enough statistical confidence to do predictions of the behavior of inequality in the future. The Kuznets Hypothesis is hence initially discarded for Colombia during the time interval 1880-2018, further suggested by a non-parabolic behavior.

Alternative models of regression, such as logarithmic, exponential, linear, or higher-degree polynomial regressions using the least squares method do not yield significantly more accurate functions, given the low values of the coefficient of determination, capable of describing the relationship between Income per Capita and Income Inequality for the 138 years considered in the research.

It is also important to note that Kuznets' Assumption that Income per Capita would rise as time passes does hold true in general for the time series, given that between 1880 and 2018 the GDP per Capita increased more than $16.000 United States 2011 PPP Dollars, with the consideration that the population also increased significantly from around four million

in the 1900s, to almost fifty million inhabitants in 2018 according to the *Maddison Project* data (Bolt & Van Zanden, 2020). The periods of time in which this does not hold, seem to correspond to exogenous macroeconomic shocks, including conflicts economic crashes, as the Latin American Debt Crisis (1980s) and the Great Depression (1930s), as Kuznets himself stated.

Furthermore, there is a difficulty in observing relationships relating income and inequality, gived that the use of income measures that are inclined toward formal market activities, as it is the case of official statistics, tends to leave out of consideration a large sector of the economy of developing nations that is classified as "informal". This holds true for Colombia, since informal market activities, that in most cases do not get registered in the official statistics like the GDP, has composed between 55% and 45% of national economic activity at peaks and lows of the 21st century; (according to the Colombian State Bank (Banco de la República)), thus implying that there is a large possibility that the data is skewed, or that misrepresents both the income distribution and the income inequality for Colombia ("Observatorio Laboral de la Universidad del Rosario", 2018). Alonso Cifuentes & González Terán (2017) further note that failing to consider the fluctuations in the extent of market participation in Colombia has large repercussions on the income data available, which is why there is a certain degree of uncertainty in research minding income in nations still transitioning into formal markets.

Source: https://www.freepik.es/

Analysis of the Defined Intervals

The first defined interval of inequality and economic growth, between 1890 and 1950, appears to exhibit a Kuznets Curve, given that the quadratic regression yields an R^2 of 0,9442, implying that most of the variation in inequality could be correlated to variations in economic growth. This period also matches to Colombia's first industrialization phase, since the confluence of technologies would imply an economic change affecting the income distribution swiftly as stated by the KH (Pombo, 2002). The line of best fit is a higher-order curve (six-degree polynomial), although presenting a restriction, since it might not predict that accurately values outside the time interval.

Similarly, the 1950s-1970s quadratic regression yielded an inverted U-curved polynomial, with an R^2 of 0,8674, meaning that there is a relatively high likelihood that income fluctuations induced a KC. On this time of Colombian economic history, the country was experiencing a third wave of industrialization, which would make sound the existence of a KC, given that arrival of new technology or switch of economic policy sparks a change in the macroeconomic structure in the country. In this particular phase, it can be seen how there is the shift that Kuznets elucidates between an agricultural -and land-oriented model toward a manufacturing-oriented model, through the Import Substitution approach to development (Pombo, 2002). The best-fit line is more accurate overall, but it does not match the end behavior of inequality, making it unreliable.

The regression of the third period selected, 1970-1985, yielded an inverted U-curved polynomial, with an R^2 of 0,8943, once again pointing out to a KC relationship. However, there was a recession intertwined to a boom –the one that corresponded to Coffee Prices–, but also to the debt crisis that the region experienced, which would imply that the behavior of inequality might be linked to the effect of the crisis in the disadvantaged, not implying the existence of a KC, but rather showing that this model assumes no endogenous and exogenous macroeconomic shocks (Pombo, 2002).

The fourth interval selected, between 1985 and 2005, yielded a quadratic regression with an R^2 of 0,9189, suggesting the existence of a KC. On this time of Colombian economic history, the country was experiencing a gradual aperture that came in hand with a series of neoliberal policies intended at shifting the role of the Colombian State in the economy and prioritizing markets which may have been a driver for inequality to exhibit Kuznets' pattern (Pombo, 2002).

The last developed interval spans for only 10 years, in which the KH does not hold since the regression yields a U-curved positive parabola, yet the function yields an R-square close to 1, and it appears to be consistent with the fact that the Government and the Colombian Central Bank have managed a tight monetary policy to have inflation under control (Pombo, 2002).

Analysis of Alternative Hypotheses

The previously elaborated historical comparisons, aiming to associate the inequality trend with the exhibited KCs might indicate that Milanovic Hypothesis does hold for the Colombian Case for the period 1880-2018. This is affirmed on the grounds that there was a corresponding historical explanation that would match inequality peaks to periods of industrialization and of economic transformation for Colombia. In this way, the KW Hypothesis could occur, as there is an observable graphical tendency for inequality to behave as a downward-opened parabola during multiple parts of Colombian economic history, more specifically during the time intervals elapsing between 1890 and 1950, 1950 and 1970, and 1985 and 2005; but also because, analytically, the existence of this peaks is reasonable in the sense that they appear to match to moments of change in the Colombian development process and the national political economy (Milanovic, 2017). These analytical findings are further backed-up by regressions confirming the existence of several parabolas with downward concavity, obtaining $R^2 > 0,85$ in all models.

The second model to consider is the *r&g* Hypothesis as proposed by Thomas Piketty (2014). This hypothesis seems unfitting for the Colombian case, not only because the determination coefficient of a third-degree polynomial regression using the least squares method for this time interval is low (R^2=0,4012), but because the overall trend of inequality does not correspond to Piketty's description of the points where it hits maximums and minimums, and the overall end-behavior. According to the *r&g* Hypothesis, in market-oriented economies – as the one of Colombia – inequality should keep rising infinitely as GDP per Capita increases; however, this assertion is inaccurate for Colombia since the trend of inequality in the last and previous decades (2008-2018 and 1985-2005) seem to indicate that income disparities are steadily declining after a peak in 1999, while the GDP per Capita during this time interval has been increasing constantly until 2018, with the notable exception of the Global Financial Crisis experienced between 2007 and 2009 (Mesa et al., 2008).

Discrete Comparisons

Discrete comparisons of the defined intervals also allow to verify the findings of previous research. Especially, it appears that although there is an eight-year difference to the proposed 1985-2005 interval in this research, findings on the KH for the 1977-2005 period hold. Bonilla-Mejía (2008) research of a KC between 2006 and 2008 is also backed-up by this paper, as the regressions suggest that the behavior of inequality between 2006 and 2018 cannot be modeled by a parabola.

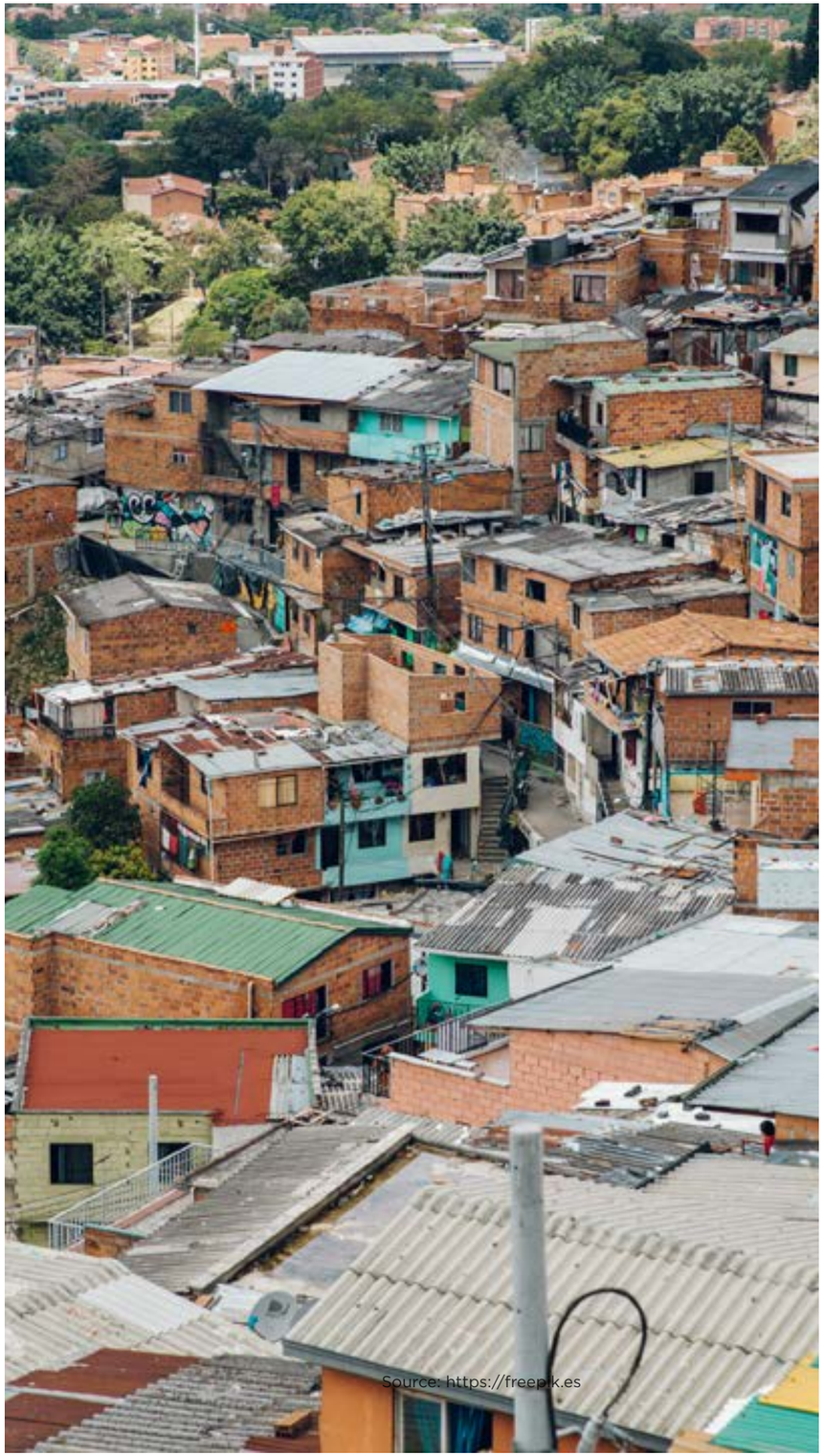
From 2001, the first year in which Multidimensional Poverty Data is available, until 2005, good economic situation of the main trading partners and the increase in prices of some strategic goods for the Colombian national economy (oil, coal, and coffee) helped to reduce poverty, while at the same time increasing GDP per Capita from $11.996 United States 2011 PPP Dollars in 2002 up to $13.115 United States 2011 PPP Dollars in 2005, reflecting an increment of almost USD $2.000 in three years (Mesa et al., 2005). The indicators of inequality also reduced, especially the GC, passing from 0,56 in 2002 to 0,542 in 2005, which in general tends to confirm Santos et al. (2019) findings, establishing that ceteris paribus, GDP Growth usually contributes to reduce Multidimensional Poverty, considering that inequality kept reducing and the macroeconomic panorama was favorable.

The aforementioned kept rule for most of the period 2001-2018, for which MPI data is accessible, but the 2008 financial crisis did affect significantly the general growth trend, having a unfavorable effect until 2011, which is why it is difficult theorizing beyond the fact that economic downturns affect inequality negatively, and specially the vulnerable to poverty, giving that predicting this type of events and their first and second order consequences is nearly impossible (Santos et al., 2019; Taleb, 2010). This shock definitely had an impact on income inequality, which made the Gini Coefficient increase from 0,545 to 0,546, breaking the overall decreasing trend that inequality was presenting during the first part of the 20th century. Although the GDP per Capita increased from $13.186 United States 2011 PPP Dollars in 2007 to $13.333 United States 2011 PPP Dollars in 2008.

From 2012 and until 2016, Public Social Spending largely increased, which means that State efforts might have played a role in this positive period for the economy, helping -along with an overall stable political panorama- to reduce inequality, and to perhaps "delay" the appearance of a future Kuznets Curve (Moreno & Liz, 2018).

Analysis of the Calculated Gini Coefficients

The methodology used can be considered precise as it can be seen on *Table 3* of the results, in which the percentage error in all cases does not exceed 1% between the value obtained experimentally and the value that the DNP reported for the same years for which the LCs were built (%Error<1%). This means that there is correspondence with the DNP data, as the preciseness suggests that the income quintiles portray fairly the whole income distribution, and it also implies that Gerber's methodology can produce fairly good estimates of the income inequality of a society, if the income quintiles are given.



Source: https://freepik.es

CONCLUSIONS

Mathematical models and conceptualization may provide useful insight to social realities. As it is observed in this research, the quantification of social issues as inequality and poverty can be executed with greater ease if econometric techniques are incorporated. However, identifying if mathematics *per se* can describe and explain realities could become an epistemological challenge. Accordingly, this article may only suggest certain findings that can be extrapolated from an analytical standpoint of the mathematics herein used to study the relationship of Economic Growth and Income per Capita with Income Inequality and Poverty.

The regression performed in this research, relating GDP per Capita with Changes in the Gini Coefficient seems to suggest that the Kuznets Hypothesis does not hold for Colombia in the timeframe considered (1880-2018). This mathematically coincides with the fact that inequality did not present an absolute maximum, and rather had different local maxima throughout the elapsed time. In this sense, the observed behavior apparently does not respond to the inequality hypothesis proposed by Piketty either, which is hinted by the fact that contrary to his reasoning, inequality in Colombia during the last decade observed has presented a trend to stabilize, and to decline at a low rate.

However, the general curve obtained for the whole 138 years analyzed, and especially the constructed sub-intervals (which in most scenarios exhibited the behavior of Kuznets Curves), could point out towards the existence of a cyclic behavior of inequality that may be associated with the appearance of an alternative econometric model, namely the Branco Milanovic *Kuznets Waves* Hypothesis. This opens up a new inquiry, as this hypothesis does not have a fully developed political economy, but specially because it is complex to find a function that corresponds to this model: the waves can present themselves with different slopes through time, and distinctions between noise and signal must be defined to assess which inequality spikes may be or may be not considered *"waves"*.

Additionally, the constructed Lorenz Curves, and the respective calculated Gini Coefficient from each of the built models for the income distributions of Colombia in the 1970s and the 1990s show congruent results with the indexes used for these years, confirming the reliability of the data, but also, they appear to suggest that income quintiles can be used to obtain accurate estimate levels of income inequality in a given population, while requiring significantly less data.

According to the previous, it may be valuable to extend the model of the relationship of Income per Capita and Inequality into a multivariate model, capable of considering other macroeconomic variables to analyze how much they affect the overall observed trends, and to determine which factors tend to incise more on economic growth and on economic inequality. Furthermore, it appears that Multidimensional Poverty tends to reduce under the bases that income inequality declines whilst economy grows. This holistic characterization of poverty integrated to the utilized mathematical modeling component allowed to broaden the scope of this research by transcending the empirical characterization of the long-term effects of the considered variables and directing the research into also considering sociological, historical, and political explanations for the exhibited behavior.

Source: https://www.gaceta.unam.mx/llaman-a-combatir-la-desigualdad-social-en-al/

This investigation also distinguishes itself since the defined period had never been analyzed for such extension of time, but also to the extent that in order for the mathematical approach to make sense in the context of a social research like this one, there must be certain historical, sociological, political, or anthropological correspondence that may be associated with what is observed. This was the motivation to select Multidimensional Poverty instead of income-based approaches while doing the discrete comparisons, for it is more sensitive to the structural and intersectional depravations that poverty implies beyond what monetary-based models suggest.

Ultimately, there is much to do at the purpose of conciliating econometric and mathematical approaches with the humanities, not necessarily as methods for the later to use the former, but as an opportunity to pose enriching quandaries and generating new questions and insights. The econometric techniques allow to comprehend in what way inequalities in Colombia have been related to economic growth in the long run. However, it would be interesting to see how political philosophy would analyze inequalities, for instance, by examining if disparities would be considered justifiable under the *Rawlsian Difference Principle*[2]. Within such framework, this research constitutes a step toward the use of mathematical modeling to humanist ends

2. Rawls proposes two principles of justice, within which is defined the Difference Principle: Social and economic inequalities are to satisfy [...] they are to be to the greatest benefit of the least advantaged members of society. (Distributive Justice (Stanford Encyclopedia of Philosophy), 2021. Available [online] at: https://plato.stanford.edu/entries/justice-distributive/#Difference)

Source: https://forbes.co/2020/10/30/economia-y-finanzas/colombia-es-el-pais-mas-desigual-de-toda-america-latina/

REFERENCES

Research Papers:

Alonso Cifuentes, J., & González Terán, Ángela. (2017). Evaluación de la Hipótesis de Kuznets para Colombia con una aproximación de series de tiempo: 1977-2005. Apuntes Del Cenes, 36(63), 43-62. https://doi.org/10.1905/01203053.v36.n63.2017.4642

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. Journal of economic growth, 5(1), 5-32.

Bolt, J., & van Zanden, J. L. (2020). Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update. University of Groningen, Groningen Growth and Development Centre, Maddison Project Working Paper, (15).

Bonilla-Mejía, L. (2008). Diferencias regionales en la distribución del ingreso en Colombia. Documentos de Trabajo Sobre Economía Regional y Urbana; No. 108.

Callen, T. (2008). What is gross domestic product. Finance & Development, 45(4), 48-49

Cárdenas-Santamaría, M. (2007). Economic Growth in Colombia: a reversal of "fortune"? Revista Ensayos Sobre Política Económica; Vol. 25. No. 53. Enero, 2007. Pág.: 220-259.

Gerber, L. (2007). A quintile rule for the Gini coefficient. Mathematics Magazine, 80(2), 133-135

Kennedy, B. P, I. Kawachi, & D. Prothrow-Stith. (1996). "Income Distribution and Mortality: Cross Sectional Ecological Study of the Robin Hood Index in the United States." BMJ 312 (7037): 1004–7. https://doi.org/10.1136/bmj.312.7037.1004.

Kusnic, M. W., & DaVanzo, J. (1980). Income inequality and the definition of income: The case of Malaysia. Santa Monica: Rand.

Kuznets, S. (1934). National Income, 1929-1932. In National Income, 1929-1932 (pp. 1-12). NBER.

Kuznets, S. (1966). Modern economic growth: Rate, structure, and spread (Vol. 2). New Haven: Yale University Press.

LaboUR, O.L., (2018). Perfil actual de la informalidad laboral en Colombia: estructura y retos. Universidad del Rosario, Observatorio Laboral de la Universidad del Rosario, Informe, (6).

Mejía, L. B. (2008). Diferencias regionales en la distribución del ingreso en Colombia. Cartagena, Colombia: Centro de Estudios Económicos Regionales (CEER) del Banco de la República. https://bit.ly/2yRKYYr

Mesa Callejas, Ramón Javier, & López Gonzáles, Mauricio (2005). Coyuntura y política económica colombiana 2005-2006: análisis y perspectivas. Perfil de Coyuntura Económica, (6),5-32. ISSN: 1657-4214. Available at: https://www.redalyc.org/articulo.oa?id=861/86100601

Naguib, C. (2017). The Relationship between Inequality and GDP Growth: An Empirical Approach Schweizerische Zeitschrift Für Volkswirtschaft Und Statistik, 153, 183–225. https://doi.org/10.1007/BF03399507

Pini, D. (2014). The Kuznets Curve and Inequality: A Comparafive Study of 27 OECD Countries Across 30 years of Neo-Liberal Economic Policies. Johns Hopkins University Faculty of Economics, https://bit.ly/2SmMM2l

Pombo, C. (2002). Industrialization and industrial policy in Colombia: a tale of economic development (No. 002842). Universidad del Rosario.

Santos et al. (2019). Growth and Poverty Revisited from a Multidimensional Perspective. The Journal of Development Studies, 55(2), 260–277. https://doi.org/10.1080/00220388.2017.1393520

Seth, S. and Alkire, S. (2014). "Measuring and Decomposing Inequality among the Multidimensionally Poor Using Ordinal Data: A Counting Approach." OPHI Working Papers 68, University of Oxford.

Yang, L. (2017). The relationship between poverty and inequality: Concepts and measurement.

Yang, L., & Vizard, P. (2017). Multidimensional poverty and income inequality in the EU.

Books:

Duprat, M. (2018). The Dynamics Of Inequality: Is There A General Pattern? [Ebook] (1st ed., pp. 4-18). Paris, France: Societe Generale. Retrieved from https://www.societegenerale.com/sites/default/files/2018/the-dynamics-of-inequality.pdf

Economic growth. (2021). In Encyclopædia Britannica. Retrieved from https://school.eb.com/levels/high/article/economic-growth/106198

Gross domestic product (GDP). (2021). In Encyclopædia Britannica. Retrieved from https://school.eb.com/levels/high/article/gross-domestic-product/389903

Milanovic, B. (2016). Global Inequality: A New Approach for the Age of Globalization. Cambridge, Massachusetts; London, England: Harvard University Press. Retrieved March 2, 2020, from http://www.jstor.org/stable/j.ctvjghwk4

Miller, S. M. and Roby, P. (1970). The Future of Inequality. New York: Basic Books, 1970, 255 pp., $7.95, L.C. 74-94291. American Behavioral Scientist, 14(1), 143–143. https://doi.org/10.1177/000276427001400178

Rawls, J. (2009). A theory of justice. Harvard university press.

Piketty, T., & Goldhammer, A. (2014). Capital in the twenty-first century (1st ed., pp. 95-187). Belknap Press: An Imprint of Harvard University Press.

Taleb, N. N. (2010). The black swan: the impact of the highly improbable. 2nd ed., Random trade pbk. ed. New York: Random House Trade Paperbacks.

Book Chapters:

Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., and Ballon, P. (2015). Multidimensional Poverty Measurement and Analysis, Oxford: Oxford University Press, ch. 1.

Lubrano, M. (2017). The econometrics of inequality and poverty Chapter 4: Lorenz curves, the Gini coefficient and parametric distributions.

Odekon, M. (Ed.) (2006). Encyclopedia of world poverty. Sage Publications, Inc., https://www.doi.org/10.4135/9781412939607

Lectures:

Milanovic, B. (2017). Inequality in the Age of Globalization. In Lecture in Honor of Anthony A. Atkinson. Brussels: Annual Research Conference. https://ec.europa.eu/info/sites/default/files/economy-finance/00.keynote_milanovic.pdf

Source: NathaliaAngarita https://www.rtve.es/noticias/20210505/estallido-social-colombia-reforma-fiscal-pandemia-policia-dispara-protestas/2088745.shtml