Income inequality and an innovative approach to its measures

Nikola Dugandzija
Riot Investment Society

January 29, 2025

This article is a product of the author's contributions within Riot Investment Society, as part of our commitment to research and knowledge-sharing in finance and data analysis.

Abstract

The hypothesis is that there exists a relationship between income inequality and difference between median and mean income. The research aims to show that income inequality could be shown just through ration between median and mean income, and corresponding gross domestic product (GDP). Through the deep analysis of relationship between the most common measure of income inequality, Gini coefficient, and the above-mentioned ratio. Using statistical methods of recognition and comparison, the level of strength of the relationship between Gini coefficient and the ration (between median and mean income, and GDP) is shown.

Contents

1 Introduction

Income inequality has been highlighted as a defining issue of the 21st century's society. Affecting developed, developing and undeveloped countries, it tends to have several related socio-economic effects. The widening gap between the rich and the poor has implications for economic stability, social cohesion, and political systems worldwide. In the basis of those issues lies a complex interplay of factors—globalization, technological advancements, and education disparities, tax policies, labor market changes, and shifts in wealth distribution.

This paper has a purpose of exploring the underlying causes of income inequality by analyzing measures of income inequality and defining innovative approach to the problem. By examining both geographical and empirical data, the research will shed light on how income inequality shapes the economic landscape and how the improper analysis of socio-economic factors could mislead the whole process.

As not all the countries have necessary data (to support this research) available, the research paper has been focused on 147 countries for which all the data were accessible. In some cases, data such as Gini coefficient, GDP, median or average income were not available, so countries were excluded towards the accuracy and truthfulness of the final result.

2 Literature Review

The study of income inequality has long been at the forefront of economic research, with significant emphasis on measures that reflect not only the disparity in income distribution but also its implications for societal well-being. Traditional measures, such as the Gini coefficient, have provided a cornerstone for analyzing income inequality, but their limitations have sparked discussions on alternative approaches that better account for the distribution of incomes and societal impacts.

Median and Mean Income: Robustness and Limitations

The debate between mean and median income as indicators of economic well-being is central to this research. ***Chiripanhura*** (2011) highlights that median income provides a better indication of economic well-being for the "typical" household compared to mean income, which is heavily influenced by outliers at the upper end of the income distribution. Similarly, ***Birdsall*** and ***Meyer*** (2014) emphasize that the median, as a "distribution-aware" metric, is a superior indicator of material well-being in low- and middle-income countries. They argue that the median better reflects the conditions of the majority population compared to GDP per capita or survey-based mean income, which can exaggerate the perceived well-being due to skewed distributions.

The Relationship Between GDP, Median, and Mean Income

The relationship between income metrics and GDP has also been widely explored. ***Hazuchová*** and ***Stávková*** (2017) discuss the shortcomings of GDP per capita as an indicator of living standards, advocating for multidimensional approaches that incorporate economic, social, and environmental factors. Moreover, ***Kreinovich*** et al. (2014) argue that while both median and mean incomes provide insights into inequality, their combined use can help derive more robust metrics to understand income distribution comprehensively.

Extensions to Income Inequality Measures

Building on these foundations, several scholars have proposed new methods to enhance the analysis of inequality. For example, ***Kämpke*** (2011) discusses the political and analytical implications of using medians versus means, emphasizing that the median can obscure certain dynamics of inequality, such as the effects of income redistribution. This aligns with the current research aim to explore how the ratio between median and mean income, when normalized by GDP, can serve as an innovative and intuitive measure of income inequality.

Existing Gaps and Contributions of This Study

Despite these advancements, existing measures like the Gini coefficient and income ratios often lack intuitive links to broader economic metrics such as GDP. Moreover, as noted by ***Birdsall*** and ***Meyer*** (2014), existing inequality measures do not fully capture shifts in income distributions over time, especially in contexts where data availability and comparability pose challenges. This study aims to address these gaps by proposing the MMG index, a novel metric derived from the relationship between median and mean income, normalized by GDP. By linking this index to the Gini coefficient through statistical analysis, the research seeks to validate its applicability and relevance in income inequality studies.

3 Methodology

To explore the relationship between income inequality and the proposed MMG coefficient, the following methodology was employed:

Data Collection

Data was collected for four key variables:

- **GDP per capita**, sourced from the World Bank database (latest available data).

- **Median and Mean Income**, sourced from the World Population Review website, which provides the most recent data, primarily for 2023. For some countries, the most recent data were from as early as 2015.
- **Gini Coefficient**, sourced from UNU-WIDER's database, which offers the latest available Gini indices for various countries. The years of availability vary depending on the country's data collection practices.

Data Cleaning and Organization

The collected data was preprocessed using **Python** and **Excel**. The following steps were undertaken:

1. **Handling missing data:** Countries with missing values for any of the four variables were excluded from the analysis to ensure that high-quality analysis could be done.
2. **Data synchronization:** Where necessary, the data were adjusted to ensure compatibility between different years of measurement, prioritizing the most recent data for each country, as the closest source to the actual data at present.
3. **MMG Calculation:** The proposed MMG coefficient was calculated using the formula:

$$\text{MMG} = \frac{\text{Mean Annual Income} - \text{Median Annual Income}}{\text{GDP per Capita}}$$

   This calculation was performed in both Python and Excel to ensure accuracy and reproducibility.

Statistical Analysis

To evaluate the strength of the relationship between the MMG coefficient and the Gini Coefficient, **Spearman's Rank Correlation Test** was applied:

1. Countries were first ranked according to their GDP per capita and MMG values.
2. Only countries with complete data for both rankings were included in the analysis.
3. Spearman's rank correlation coefficient was calculated following the standard formula to determine the degree of association between the two variables:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

   where:

- d_i: the difference between the ranks of each pair of observations (e.g., the rank of MMG and the rank of GDP per capita),
- n: the total number of observations.

Spearman's correlation coefficient r_s ranges between -1 and 1:

- $r_s = 1$: indicates a perfect positive monotonic relationship,
- $r_s = -1$: indicates a perfect negative monotonic relationship,
- $r_s = 0$: indicates no monotonic correlation between the variables.

Data Visualization

All visualizations, including scatter plots and correlation diagrams, were generated using **Python**. This provided a clear graphical representation of the relationships between MMG, Gini Coefficient, GDP per capita, and other variables.

Limitations

This study acknowledges certain limitations that may influence its findings:

- **Variability in the years of data collection across sources:** The inconsistency in data collection periods poses challenges for ensuring uniform comparisons. Two potential solutions to address this limitation include:
  - Utilizing predictive models to estimate the required values for a fixed reference year.
  - Defining a specific time period in the past and incorporating all relevant calculations by using their statistical first moments (e.g., means).
- **Exclusion of countries with incomplete data:** The omission of countries with missing data may affect the representativeness and generalizability of the results. This limitation could be mitigated by leveraging historical data from a broader range of countries to apply regression analyses for predicting the missing variables required in the study.

4 Results

MMG ratio (mean-median-GDP ratio)

The main aim of this research, as well as its scientific contribution, lays in the idea of *MM-GDP ratio*. MM-GDP ratio is meant to be a new measure of inequality. In the further analysis, it will be detonated as *MMG*. It is defined by formula:

$$\text{MMG} = \frac{\text{Mean Annual Income} - \text{Median Annual Income}}{\text{GDP per Capita}}$$

Theoretically, the range of the MMG index extends from $-\infty$ to $+\infty$, reflecting the possibility of extreme variations in the difference between mean and median incomes relative to GDP per capita. However, in practical applications, the index predominantly takes positive values, as mean income is almost always greater than median income due to the positive skewness in income distributions observed in most economies.

This characteristic makes the MMG index a valuable tool for capturing the degree of inequality, as it highlights deviations from the median, which represents the typical household, in relation to GDP. While extreme negative values are rare, they would theoretically correspond to distributions where the median exceeds the mean, an uncommon scenario in real-world datasets.

Lower values correspond to countries with greater equality, while by its increasing, countries with greater inequality are shown.

Table 1: Ten countries with the least inequality (based on MMG ratio)

Rank	Country	Median Income	Mean Income	GDP per Capita (PPP)	MMG ratio
1	Turkmenistan	$706	$963	$15,206	0.016901223
2	Uzbekistan	$591	$752	$7,308	0.022030651
3	Kazakhstan	$3,661	$4,313	$27,517	0.023694443
4	Slovakia	$9,037	$9,505	$19,382	0.024146115
5	Iraq	$1,951	$2,293	$11,362	0.030100334
6	Trinidad and Tobago	$2,868	$3,711	$27,334	0.030840711
7	Algeria	$2,612	$2,996	$12,019	0.031949413
8	Iceland	$20,630	$22,988	$72,010	0.032745452
9	Moldova	$3,133	$3,590	$13,626	0.033538823
10	Georgia	$2,273	$2,814	$15,655	0.034557649

Artificial examples of application

To provide a deeper understanding of the MMG coefficient, its distribution, and its implications, three artificial examples have been created. These examples represent hypothetical countries—country X, country Y, and country Z—each with distinct income distributions. The following visualization illustrates the data for these artificial countries, highlighting how MMG values vary under different conditions.

Table 2: Income Disparity Comparison

Country	Population (million)	Mean Income ($)	Median Income ($)	GDP per Capita ($)	MMG
Country X	1	100,000	30,000	50,000	1.4
Country Y	1	52,000	50,000	50,000	0.04
Country Z	1	70,000	45,000	50,000	0.5

The next visualization illustrates the differences in income distributions and MMG ratios across three hypothetical countries: Country X, Country Y, and Country Z.

· **Mean Income:** This chart shows the average income levels for each country, highlighting significant variations between them.

· **Median Income:** This chart displays the median income, providing insight into the central income level within each population.

· **GDP per Capita:** This chart shows the GDP per capita, allowing us to see the economic output relative to the population size for each country.

· **MMG Ratio:** The MMG ratio chart reveals the degree of income inequality in each country, with higher values indicating greater inequality. Country X shows a notably higher MMG ratio compared to the other two countries, suggesting more pronounced income inequality.

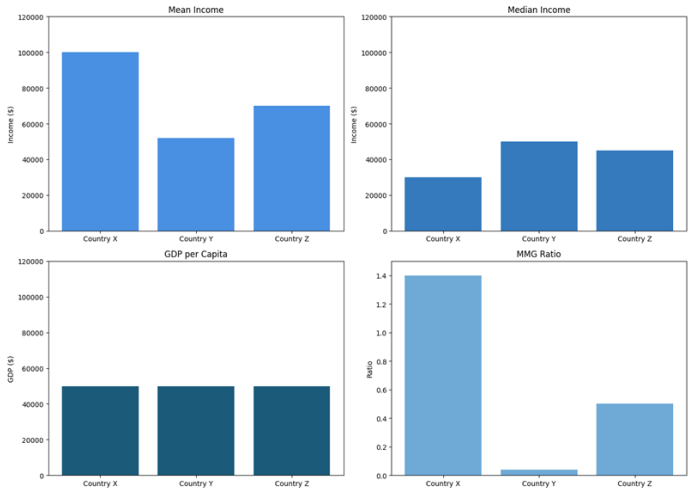

Figure 1: Graphical Representation of Income and MMG Distributions among Artificial Countries

By comparing MMG values from real countries with those of the artificial examples, it becomes possible to identify patterns and deviations in income inequality. This comparative approach offers valuable insights, allowing researchers to see where actual countries align or differ from theoretical models, thus reinforcing MMG's utility in analyzing income distribution on a broader scale.

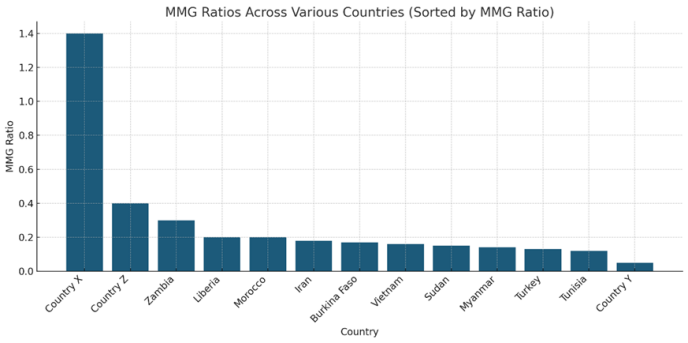

Figure 2: MMG Ratios Across Various Countries

Distribution of MMG

The distribution of the MMG was analyzed using a histogram with a Kernel Density Estimation (KDE) curve, providing a visual insight into the distribution of MMG values within the dataset. The histogram displays the frequency of different MMG values. The KDE curve represents a smooth estimate of the data's distribution on the given range. The analysis revealed right-skewness, indicating that most MMG values tend to cluster in the lower range, while a smaller number of countries exhibit relatively high MMG values. In terms of inequality, it would lead to a conclusion of smaller number of countries with more moderate inequality. Theoretically, the MMG ranges from to , but empirically, its values are in the range from 0.017 to 0.696, highlighting a significant variation among the observed countries.

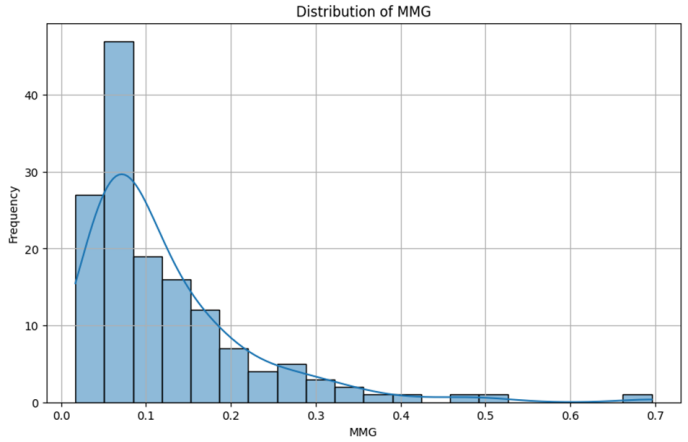

Figure 3: Distribution of MMG

Outlier detection

Outlier detection was performed using a box plot analysis of the MMG values. A box plot provides a visual method for identifying outliers by displaying the spread of the data through the interquartile range (IQR), with the central box representing the middle 50% of the data, and whiskers extending to 1.5 times the IQR. Any data points that fall outside these whiskers are considered potential outliers. This method allows for the identification of countries or regions with MMG values that are significantly higher or lower than the majority of observations. The analysis revealed 9 outliers, suggesting that certain countries or time periods exhibit unusual levels of MMG compared to the overall trend. Identifying these outliers is crucial as they can indicate unique economic conditions or measurement anomalies that may require further investigation.

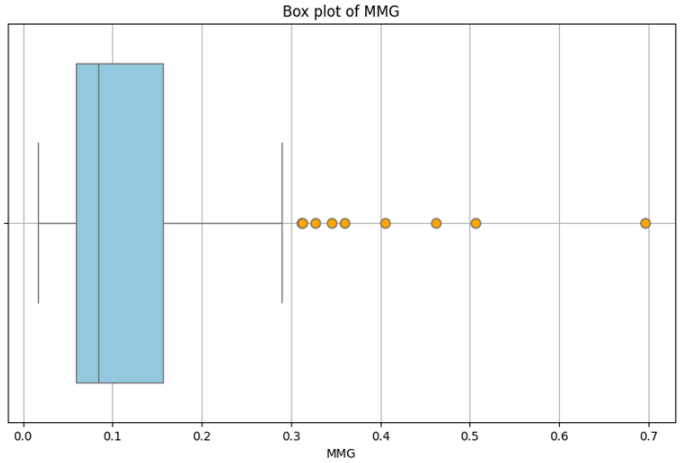

Figure 4: Boxplot of MMG

Regression analysis

To evaluate the relationship between the MMG and the Gini coefficient, a linear regression analysis was conducted. The results revealed a statistically significant positive correlation between these two indicators of income inequality. The regression model suggests that an increase in the MMG is associated with an increase in the Gini coefficient, indicating that countries with higher MMG tend to experience greater income inequality. The coefficient for the MMG is 43.14, implying that for each unit increase in MMG, the Gini coefficient increases by approximately 43.14 points. The model's value of 0.324 indicates that MMG explains about 32.4% of the variance in the Gini coefficient, supporting the hypothesis that MMG can serve as a complementary measure of income inequality. This relationship underscores the potential utility of MMG in providing additional insights alongside traditional measures like the Gini coefficient.

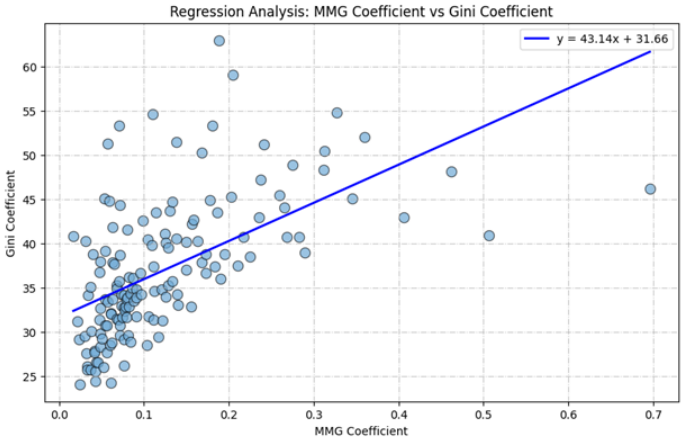

Figure 5: Regression Analysis: MMG Ratio vs Gini Coefficient

Correlation analysis

The Spearman's rank correlation test was applied to evaluate the accuracy of the model in predicting MMG coefficients across various countries. This non-parametric test measures the strength and direction of the association between the MMG coefficient and the Gini coefficient, focusing on rank-order relationships rather than absolute values.

The Spearman correlation coefficient r_s is calculated using the formula:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 0.654$$

where:

- d_i: the difference between the ranks of each pair of observations (e.g., the rank of MMG and the rank of Gini coefficient),
- n: the total number of observations.

With a resulting correlation coefficient of **0.654**, the test outcomes suggest a moderate positive correlation between MMG and Gini coefficients, supporting the hypothesis that countries with higher MMG values tend to exhibit greater income inequality as represented by the Gini coefficient. This finding reinforces the reliability of MMG as a complementary metric to traditional inequality measures, providing additional insights into income distribution patterns.

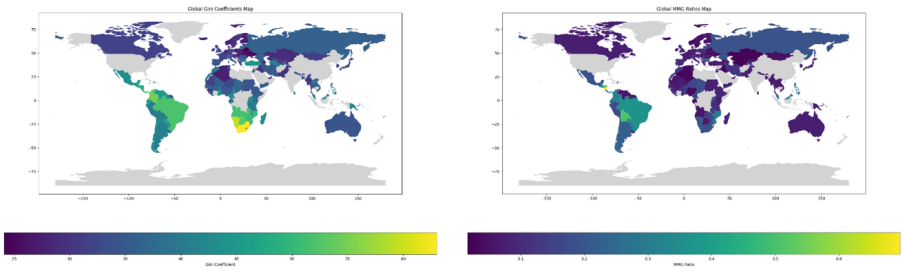

Figure 6: Gini Coefficient Map and MMG Map (respectively)

5 Discussion and Conclusion

The findings of this analysis underscore the potential of the MMG (Mean-Median-GDP ratio) as a valuable measure of income inequality, complementing the well-established Gini coefficient. The observed positive relationship between MMG and Gini coefficient suggests that MMG effectively captures structural aspects of the economy that influence income distribution, similar to those reflected by the Gini coefficient. This connection implies that MMG is not only a statistical indicator but also an economic one, revealing nuances in how income is spread across populations.

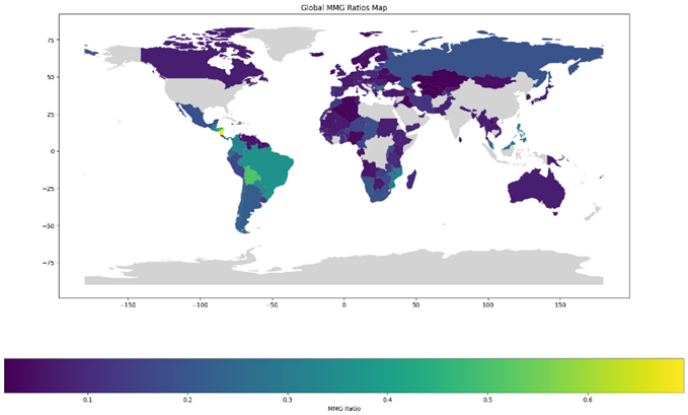

Figure 7: Global MMG Map

The MMG's reliance on the relationship between mean and median income, relative to GDP, provides additional insight into economic disparities that Gini coefficient alone might not fully capture. For instance, a high MMG value could indicate a larger gap between average income and the economic output of a country, pointing to disparities that may arise from unequal resource distribution, labor market segmentation, or other socio-economic factors.

In conclusion, the MMG serves as an innovative addition to inequality research, offering a nuanced perspective that enriches our understanding of income inequality. Its integration alongside traditional measures like the Gini coefficient could enhance the analytical framework available to policymakers and researchers, providing a more comprehensive view of economic inequality across different national contexts.

Future Research

While this study provides a novel approach to measuring income inequality through the MMG, several avenues for further research remain open.

One promising direction is the development of a predictive model that utilizes historical data on mean income, median income, GDP per capita, and the Gini coefficient to forecast future income inequality trends. By collecting and analyzing data over multiple time periods, MMG could be estimated over time and it could be assessed to its evolution. This approach would allow for the identification of patterns and trends in income inequality, providing a more dynamic and forward-looking perspective.

Future studies could explore:

- **Time-series analysis** of MMG using econometric or machine learning models to predict its trajectory based on macroeconomic variables.
- **Trend identification** in income inequality, examining whether MMG consistently signals rising or falling disparities across different economic conditions.
- **Policy implications**: investigating whether MMG trends can serve as early indicators for socioeconomic shifts, aiding policymakers in designing more effective interventions.

By incorporating predictive analytics, the MMG could evolve into a descriptive measure meant for equality comparisons and a forecasting tool that helps anticipate future economic disparities, thus contributing to academic research and policy-making efforts.

Appendix

Data link: https://en.wikipedia.org/wiki/List_of_countries_by_income_equality

Gini Coefficient Rankings

Rank	Country	Gini	Year	Rank	Country	Gini	Year
1	Slovakia	24.1	2021	75	Mali	35.7	2021
2	Slovenia	24.3	2021	76	Sierra Leone	35.7	2018
3	Belarus	24.4	2020	77	Russia	36	2020
4	Ukraine	25.6	2020	78	Vietnam	36.1	2022
5	Moldova	25.7	2021	79	Senegal	36.2	2021
6	Netherlands	25.7	2021	80	Lithuania	36.7	2021
7	United Arab Emirates	26	2018	81	Yemen	36.7	2014
8	Iceland	26.1	2017	82	Mauritius	36.8	2017
9	Czech Republic	26.2	2021	83	Dominican Republic	37	2022
10	Azerbaijan	26.6	2005	84	Burkina Faso	37.4	2021
11	Belgium	26.6	2021	85	Chad	37.4	2022
12	Algeria	27.6	2011	86	Burundi	37.5	2020
13	Finland	27.7	2021	87	Sri Lanka	37.7	2019
14	Norway	27.7	2019	88	Israel	37.9	2021
15	Armenia	27.9	2022	89	Togo	37.9	2021
16	Denmark	28.3	2021	90	Gabon	38	2017
17	Bhutan	28.5	2022	91	Malawi	38.5	2019
18	Poland	28.5	2021	92	Kenya	38.7	2021
19	Kyrgyzstan	28.8	2021	93	El Salvador	38.8	2022
20	Croatia	28.9	2021	94	Gambia	38.8	2020
21	Hungary	29.2	2021	95	Laos	38.8	2018
22	Kazakhstan	29.2	2021	96	Bulgaria	39	2021
23	Maldives	29.3	2019	97	Suriname	39.2	2022
24	Albania	29.4	2020	98	Morocco	39.5	2013
25	Iraq	29.5	2012	99	United States	39.8	2021
26	Guinea	29.6	2018	100	Micronesia	40.1	2013
27	Pakistan	29.6	2018	101	Jamaica	40.2	2021
28	Sweden	29.8	2021	102	Peru	40.3	2022
29	Ireland	30.1	2021	103	Trinidad and Tobago	40.3	1992
30	Austria	30.7	2021	104	Tanzania	40.5	2018
31	Fiji	30.7	2019	105	Uruguay	40.6	2022
32	Myanmar	30.7	2017	106	Argentina	40.7	2022
33	Uzbekistan	31.2	2022	107	Malaysia	40.7	2021
34	Cyprus	31.3	2021	108	Philippines	40.7	2021
35	Malta	31.4	2020	109	Turkmenistan	40.8	1998
36	Mongolia	31.4	2022	110	Bolivia	40.9	2021

Rank	Country	Gini	Year	Rank	Country	Gini	Year
37	South Korea	31.4	2016	111	Haiti	41.1	2012
38	France	31.5	2021	112	Djibouti	41.6	2017
39	Canada	31.7	2019	113	Papua New Guinea	41.9	2009
40	Germany	31.7	2019	114	Cameroon	42.2	2021
41	Estonia	31.8	2021	115	Madagascar	42.6	2012
42	Lebanon	31.8	2011	116	Uganda	42.7	2019
43	Mauritania	32	2019	117	Central African Republic	43	2021
44	Seychelles	32.1	2018	118	Chile	43	2022
45	United Kingdom	32.4	2021	119	Ghana	43.5	2016
46	Luxembourg	32.7	2021	120	Mexico	43.5	2022
47	Nepal	32.8	2010	121	Rwanda	43.7	2016
48	Greece	32.9	2021	122	South Sudan	44.1	2016
49	Japan	32.9	2013	123	Turkey	44.4	2021
50	Niger	32.9	2021	124	DR Congo	44.7	2020
51	Bosnia and Herzegovina	33	2011	125	Venezuela	44.8	2006
52	Serbia	33.1	2021	126	Lesotho	44.9	2017
53	Bangladesh	33.4	2022	127	Guyana	45.1	1998
54	North Macedonia	33.5	2019	128	Paraguay	45.1	2022
55	Jordan	33.7	2010	129	Comoros	45.3	2014
56	Switzerland	33.7	2020	130	Ecuador	45.5	2022
57	Tunisia	33.7	2021	131	Nicaragua	46.2	2014
58	Romania	33.9	2021	132	Costa Rica	47.2	2022
59	Spain	33.9	2021	133	Honduras	48.2	2019
60	Tajikistan	34	2015	134	Guatemala	48.3	2014
61	Georgia	34.2	2021	135	Panama	48.9	2023
62	Sudan	34.2	2014	136	Zimbabwe	50.3	2019
63	Australia	34.3	2018	137	Mozambique	50.5	2019
64	Latvia	34.3	2021	138	Saint Lucia	51.2	2016
65	Montenegro	34.3	2021	139	Angola	51.3	2018
66	Benin	34.4	2021	140	Zambia	51.5	2022
67	Portugal	34.6	2021	141	Brazil	52	2022
68	Iran	34.8	2022	142	Belize	53.3	1999
69	Italy	34.8	2021	143	Botswana	53.3	2015
70	Thailand	34.9	2021	144	Eswatini	54.6	2016
71	Ethiopia	35	2015	145	Colombia	54.8	2022
72	Nigeria	35.1	2018	146	Namibia	59.1	2015
73	Ivory Coast	35.3	2021	147	South Africa	63	2014
74	Liberia	35.3	2016				