

Name: Nana Osei Sarpong

Student Number: 11371873

Course: Development Economics (ECON 822)

Assignment 4

The presence of negative income values makes the computation of Gini coefficients and other measures of inequality problematic. According to Ostasiewicz et al. (2018), this causes the Gini coefficient to lose the property of being bounded between zero and one, making it difficult to interpret. De Battisti et al. (2019) explore different ways of dealing with such a problem to get a bounded Gini coefficient. Common ways of dealing with this problem is to omit the negative values or assume they are equal to zero. By so doing, we understate the level of inequality and bias our sample.

I deal with negative values in a simple, yet intuitive manner. All members are given a transfer worth the lowest possible value in the range of income categories. Subsequently, all income sources are summed up to give the total income. This ensures the average and the distribution of income are not tampered with.

Table 1: Inequality Measures in Total Income

Measure of Interest	Computational Results
Coefficient of Variation	0.121121
Gini Coefficient	0.040726
Income Share of Top 20%	0.228565
Income Share of Bottom 20%	0.186927
Kuznets Ratio	1.222750

From Table 1, it is observed that the coefficient of variation is quite small. Typically, a coefficient of variation which is less than 0.3 is considered to be small. The standard deviation of income per unit of average income is small, signifying that income does not vary a lot across the population and hence inequality is low.

The Gini coefficient also lends itself to a similar conclusion. The Gini coefficient is very close to zero, hence, income is quite equally distributed.

The top 20% of the population hold 22.86% of total income while the bottom 20% of the population hold 18.69% of total income. This is remarkably close, signaling low income inequality.

The Kuznets ratio is 1.222750 which means that the income held by the top 20% is only 22.28% more than that of the bottom 20%. This further cements our earlier stated position of low income inequality in the population.

Table 2: Decomposition of Inequality by Different Income Sources

	Agriculture	Non-Agriculture	Remittances	Other
Weight of Income Source, wk	0.260558	0.262409	0.240403	0.236629
Relative Correlation, Rk	0.824859	0.697861	0.410879	0.352881
Gini of Income Source, Gk	0.093264	0.087205	0.037623	0.011925
Relative Gini, Gk/G	2.290027	2.141258	0.923813	0.292798
Share in Total Gini, k/G	0.492181	0.392118	0.091251	0.024449

Table 2 decomposes inequality by different income sources. All income sources are approximately of the same level of importance. The Non-Agricultural source of income, however, is the income source with the highest share in total income at 26.24%.

The relative correlation measure (Rk) tells us that people with high incomes more frequently have high agricultural and non-agricultural incomes. Remittances and other sources of income are less prevalent among the rich.

All income sources show a very low Gini coefficient, signaling equal distribution among households. Remittances and other income sources have the most equal distribution of incomes across households.

The relative Gini tells us which income sources require attention if we further want to reduce inequality. Agricultural and non-agricultural income sources are 1.29 and 1.14 times more than the overall Gini coefficient. The share in total Gini further confirms that these are the main sources of inequality. Over 80% of the income inequality among households emanates from the agricultural and non-agricultural income sources.

References

De Battisti, A., Porro, F., & Vernizzi, A. (2019). The Gini coefficient and the case of negative values. *Electronic Journal of Applied Statistical Analysis*, 12(1), 85.

doi:10.1285/i20705948v12n1p85

Ostasiewicz, K., & Vernizzi, A. (2018). Decomposition of the Gini index in the presence of observations with negative values. *Mathematical Economics*, 14, 37-50.

doi:10.15611/me.2018.14.04

Appendix I: ChatGPT prompt

<https://chat.openai.com/share/6e7d1f7d-a186-418d-b56a-66c73e274a58>

how do we arrange data in ascending order in python based on a column in a data frame

I want to also create a column that assigns a rank starting with 1 at the lowest up till the last observation

what if you want to sort the data based on a data series (only one column)

how do we calculate the coefficient of variation

based on data in a data frame

how do we calculate the covariance between two columns

how do I add a number to every observation in a particular column in a dataframe

how do we multiply the observations in two different columns and add the result

how do I calculate the income of the, say bottom and top 10 percent of the total population