Proceedings of the 5th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2026.38565

Income Inequality in Hong Kong S.A.R.: An Analysis Using the Gini Coefficient, Lorenz Curve, and Simpson's Paradox

Yanzhao Wang

Shanghai Concord Bilingual School, Shanghai, China
202122060005@concordschool.com.cn

Abstract. This research aims to investigate the income inequality problem in Hong Kong. Apart from practical solutions, basic theories are also included. Simpson's Paradox, the Lorenz Curve, and the Gini Coefficient are the three main tools for investigating the degree of change in income inequality and adjusting for errors caused by insufficient data analysis. These researchers would use a combination of theory and real-life data to conduct an overall investigation and analysis of the income inequality problem in Hong Kong, China. The solutions suggested by the government and the other possible methods and policies could both be stated in the final part of the research. Investigating income inequality is very important since it could help scholars to develop a more mature and well-developed understanding of economic phenomena and the effective use of different economic models and theories. At the end of the research, a conclusion could be reached: Although income inequality and the Gini Coefficient are still fluctuating, noticeable improvement could still be recognised. Governments could impose policies such as progressive taxation or social welfare to improve the situation.

Keywords: income inequality, Simpson's paradox, Lorenz curve

1. Introduction

As a crucial and unavoidable economic crisis, income inequality has long been a problem for countries or regions that have implemented a market or mixed economy. Hong Kong, which is a city well known for its free financial market and highly market-dominated society, also faced income inequality within various social groups in the 20th and 21st centuries. Based on deeper and detailed investigation, a conclusion could be reached that income inequality is an issue that has a wide range of causes. These include not only average wages and salaries but also broader and more specific factors across society, such as age, population size, and government decisions.

The Lorenz curve and Simpson's paradox have become some of the most useful tools for illustrating income inequality and considering different factors. The Lorenz curve uses a graph to illustrate how the degree of income inequality in a country or region varies over time, including the distribution of different economic variables. The Lorenz curve is a specific type of concentration curve, and these curves do not intersect regardless of changes in the welfare function of different groups and individuals [1]. Besides that, defining an egalitarian line requires researchers to find the straight line that passes through at least two points with the same x and y values, such as (0,0) and (1,1). The Gini coefficient is calculated by dividing the area between the curve and the egalitarian line by the entire area below the egalitarian line [2]. Beyond the Lorenz Curve, multiple scholars have also studied Simpson's paradox. This paradox occurs when a population or group is divided into subgroups, leading to a changed or even reversed correlation between two variables [3]. However, simply listing theories will not directly link to the real-life problem of income inequality in Hong Kong without current data from authoritative sources.

Therefore, this research will combine all the following research results and theories listed above with recent data on Gini coefficients and other related information on income inequality released by the Census and Statistics Department [4] and the Office of the Government Economist [5]. By using real research results and data, such as income gaps between groups and Gini Coefficients, this research can draw specific and accurate conclusions about the current state of Hong Kong's economy.

The rest of the paper is structured as follows: Section 2 provides a more detailed and logical analysis of the theory and use of Simpson's Paradox, Lorenz's Curve, and the Gini Coefficient. Policy recommendations will be provided at the end of the article. A more specific investigation of the situation in Hong Kong and relative suggestions on policymaking

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.38565

will be provided in the 3rd section. The final section will provide an overall conclusion about the research results and potential weaknesses.

2. Lorenz curve and Simpson's paradox

2.1. Lorenz curve

The Lorenz curve is a graph introduced in the 20th century that illustrates wealth distribution in modern society. This paper uses the Lorenz curve to determine the distribution of different indicators and variables. Usually, the Lorenz curve is used to investigate the degree of economic development. According to research done by Ginevičius, R et al, the Gini coefficient must be changed in this manner because the expansion is more even when the variation in development is smaller, and the Lorenz curve is closer to the square of the square. The Gini coefficient is the data that could be deduced from the graph by calculating the ratio between the area between the curve and the optimum line and the entire area below the optimum line. See equation (1) and Figure 1.

$$G = \frac{S1.}{S1+S2} \tag{1}$$

As the distribution irregularity decreases, the area between the curve and the optimum line decreases, which means that the curve bends more towards the optimum line. As a result, the Gini coefficient decreases, suggesting improved income equality. Conversely, as the curve bends away from the straight optimum line, the distribution gets more irregular, suggesting a rise in the Gini coefficient and inequality [6]. It's also worth noting that the Lorenz curve is a specific type of a kind of graph called the concentration curve. According to research done by Kakwani, N.C, the Lorenz curve concept has been modified and extended to analyze the relationships between the distributions of different economic variables. The Lorenz curve is just one example of these curves—more precisely, the income concentration curve. The concentration curve is another name for the generalized Lorenz curve. Moreover, Kakwani, N.C. and Podder, N had done another research together in 1973, in which one of the sections about the features of the Lorenz Curve stated that the parametric relationship between a and b is represented graphically by the Lorenz curve. A unit square is used to depict the curve's graph. The egalitarian line is the straight line that connects the points (0, 0) and (1, 1). Along this line, a=b, indicating that each unit receives the same revenue. These two studies both explained the position that the Lorenz Curve holds in economics and its exact function and features.

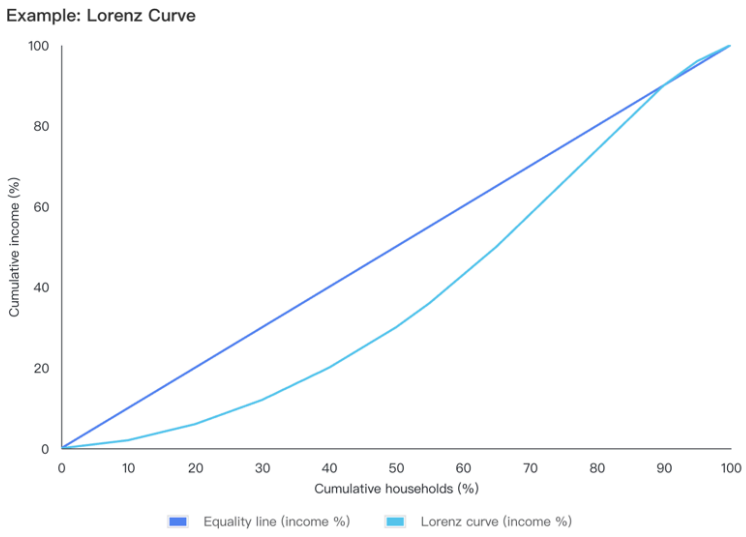


Figure 1. Lorenz curve(picture credit: original)

2.2. Simpson's paradox

Statisticians often recognise Simpson's Paradox as a common mistake that scholars will make while doing data analysis. This paradox refers to situations where people lack critical thinking and rely entirely on continuous, rigorous data from various platforms and websites that seem authoritative and reasonable. However, simply combining all results and data

DOI: 10.54254/2754-1169/2026.38565

could lead to a misleading result. The conclusion may therefore become inaccurate [7]. Other research has also led to similar conclusions that Simpson's Paradox occurs without notice. According to research done by Stanford University, a relationship between two variables may exist, vanish, or reverse when a population is split up into smaller groups. This phenomenon is called Simpson's Paradox. In a population, for example, two variables may be positively connected, but in each subpopulation they may be independent or even negatively correlated [3]. Related research on real-life samples and cases has also shown results that support the theory. One well-known example of Simpson's paradox is a study that compares open surgery with percutaneous nephrolithotomy for the treatment of kidney stones. Table 1 is plotted by the size of stones, and the success rates of each kind of surgery are provided. The paradox is that open surgery is associated with better success rates for tiny stones (93.1% vs. 86.7%) and large stones (73.0% vs. 68.8%), yet percutaneous nephrolithotomy appears to be more successful when the stone size is not taken into account.

Table 1. Treatment of kidney stones

Stone diameter	Open surgery	Percutaneous nephrolithotomy	RR
Stone diameter < 2 cm	81/87 (93.1%)	234/270 (86.7%)	RR = 1.07
Stone diameter ≥ 2 cm	192/263 (73.0%)	55/80 (68.8%)	RR = 1.06
All stones (aggregate)	273/350 (78.0%)	289/350 (82.6%)	RR = 0.94

As a similar example, the hospital epidemiology field investigates the difference in the infection rate between patients in eight Dutch hospitals, some of which are low-incidence hospitals, and the rest are all high-incidence hospitals. If researchers only focus on the aggregate data at the bottom of Table 2, a conclusion that using Antibiotic Prophylaxis could effectively reduce the infection rate could be easily made. However, if data are broken down into disaggregated parts, then the result would be reversed. The infection rate of every group of patients that takes the medicine in each kind of hospital is higher than the infection rate of those who didn't take the antibiotics. This is because the proportion of patients receiving antibiotic prophylaxis varies between low-incidence and high-incidence hospitals [8]. Therefore, when two groups of data with completely different sizes are combined, the final result will unquestionably become unreliable.

Table 2. Infection rate

Hospital type	Antibiotic prophylaxis yes	Antibiotic prophylaxis no	RR
Low-incidence hospitals	20/1113 (1.8%)	5/720 (0.7%)	RR = 2.59
High-incidence hospitals	22/166 (13.3%)	99/1520 (6.5%)	RR = 2.03
All hospitals (aggregate)	42/1279 (3.3%)	104/2240 (4.6%)	RR = 0.71

For more recent cases, the COVID-19 pandemic could also be a typical example to discuss. Early data released in 2020 indicated that the fatality rate caused by the virus in Italy was way higher than that of China. However, this crude analysis was quickly overturned as further research suggested that the fatality rate within every age group in China was actually higher than that of Italy [9]. See Figure 2.

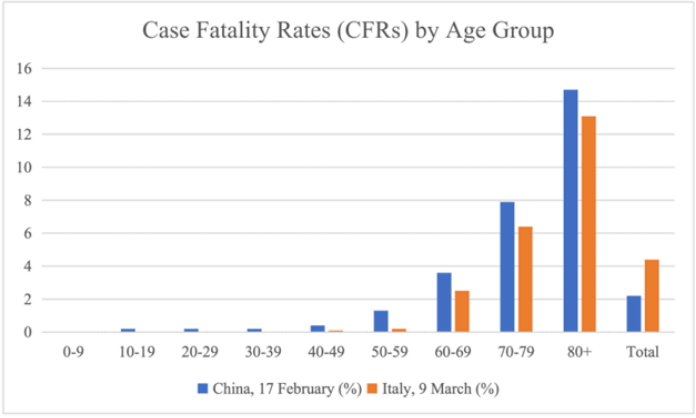


Figure 2. COVID-19 fatality rate(picture credit: original)

DOI: 10.54254/2754-1169/2026.38565

3. Implications for Hong Kong

3.1. Simpson's paradox in real cases

As a modern international city with a market economy, Hong Kong continues to face income inequality challenges in the 20th and 21st centuries. Different social groups have income gaps, which result in huge differences in people's living standards. Local government and other official academic organizations have published related data analyses and research reports on the Hong Kong economy. Within those reports, the Lorenz curve and Gini coefficient are widely used. Other data may show signs of Simpson's Paradox in the details. The Census and Statistics Department, which is the most direct and authoritative source, published a report in 2021 about the entire Hong Kong population. Inside it, the entire population is separated into ten decile groups according to their income range from lowest to highest [4]. See Table 3.

Table 3. Change in household income

Decile group	Measure	2011	2016	2021
1st (lowest)	All domestic households	2 290	3 290	2 860
1st (lowest)	Economically active households	7 000	9 250	8 310
Decile group	Ratio category	Period		Ratio
1st (lowest)	All-household ratio	2016/2011		1.44
1st (lowest)	All-household ratio	2021/2016		0.87
1st (lowest)	Active-household ratio	2021/2011		1.25
1st (lowest)	Active-household ratio	2016/2011		1.32

Table 3 clearly shows the variation in different household incomes. The ratio change on the right side of the table directly indicates the degree and the direction of the income change, whether it has improved or reduced. Overall, income in every household group seems to have improved from 2011 to 2021 as the ratio numbers are all larger than 1. However, when the aggregated data is broken down into detail, it is obvious that the all-household ratio from 2016 to 2021 in the 1st decile group had decreased. However, the overall aggregate data from 2011- 2021 perfectly covered this fact. This could be classified as Simpson's paradox since the overall, summative result could cover the details within the case.

Figures 3 and Figure 4 show the change in the Gini coefficient of different countries and regions [5]. Among all countries and regions, Hong Kong's Gini coefficient change varied across the two graphs: the first showed a rise in the Gini coefficient, while the second showed that the value didn't change throughout the period. This is due to a consideration of household sizes and composition. Again, ignorance of disaggregated data or other detailed factors could lead to completely different results. Simpson's Paradox has again been proved to be a crucial and unavoidable part of statistics and Economics.

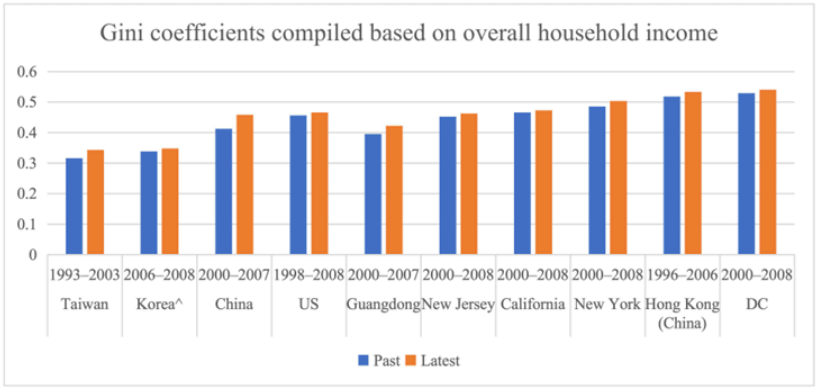


Figure 3. Gini coefficients compiled based on overall household income(picture credit: original)

DOI: 10.54254/2754-1169/2026.38565

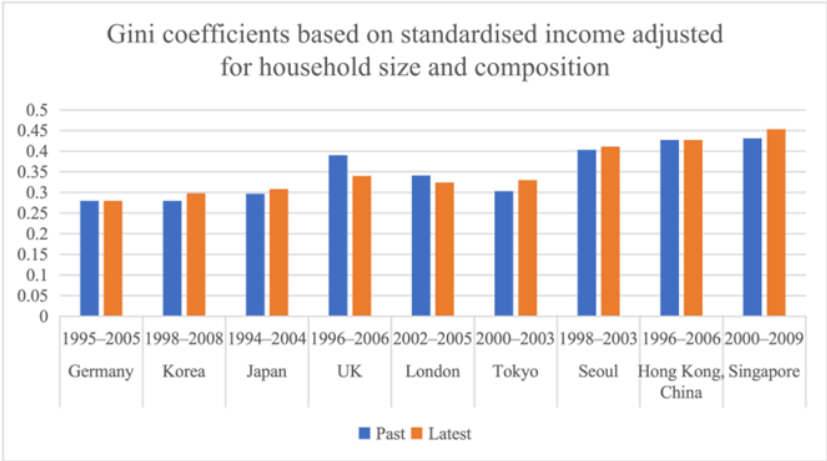


Figure 4. Gini coefficients based on standardised income adjusted for household size and composition(picture credit: original)

3.2. The Lorenz curve in Hong Kong

The Hong Kong government also used the Lorenz curve to illustrate and visualize the change in income inequality of the local society. The Census and Statistics Department provided a graph of the Lorenz curve at the end of their report. Figure 5 shows three curves that stay close together below the line of equality [4]. Although the difference may not be easy to recognize, there is indeed a variation in the degree of income inequality. By using integration and the variables on the graph to calculate the ratio of the area between the line of equality and the Lorenz curves, researchers could obtain the value of the Gini Coefficient. The closer the curve is to the line of equality, the smaller the income inequality in Hong Kong is.

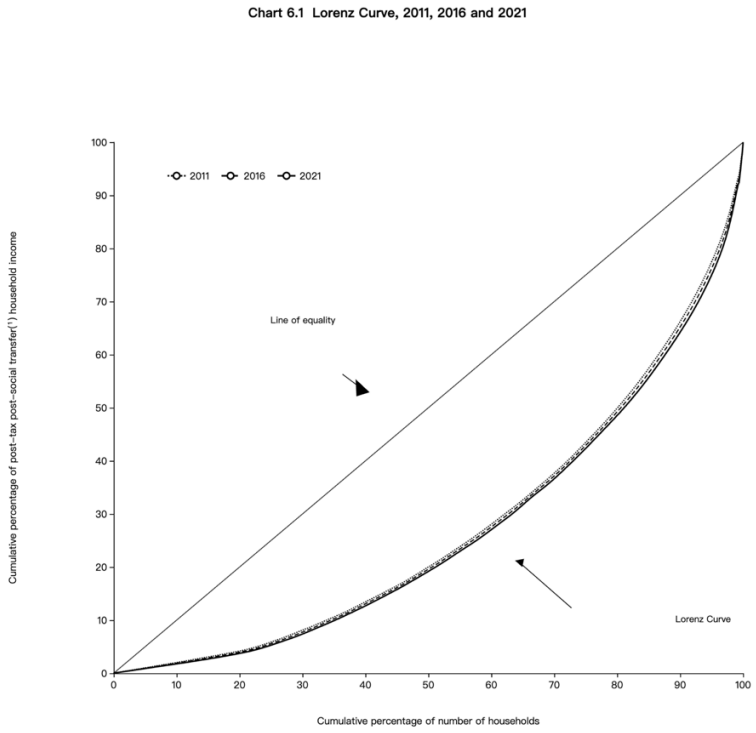


Figure 5. Lorenz curve of Hong Kong (picture credit: original)

DOI: 10.54254/2754-1169/2026.38565

3.3. Suggestions and future study

Although the income inequality problem in Hong Kong still exists, related policies and economic methods could effectively reduce the negative effects and improve the current situation. Among all policies, taxation and social welfare have always been two major useful tools to achieve the goal of reduced inequality. See Table 4.

Table 4. Taxation and welfare

Step	Item	Operation	Description
1	Original household income	Starting point	The starting point of the calculation
2	Taxation: Salaries tax, property tax, rates and Government rent	− (deduct)	Deducted from original household income
3	Post-tax household income	= (intermediate result)	Household income after deducting taxation
4	Social benefits: Education, housing and medical benefits	+ (add)	Social benefits provided in kind
5	Social benefits: Government's one-off relief measures	+ (add)	Social benefits provided in cash
6	Post-tax post-social transfer household income	= (final result)	Final household income after social transfers

It's easy to identify different types of social welfare that the Hong Kong S.A.R. government provides to support citizens in relative or absolute poverty [4]. As a result, the average income and living standard of those citizens may be improved to a large extent. Apart from that, taxes are also imposed, and some of these could limit the income of high-income households. These taxes include salary tax, property tax, rates, and government rent. Therefore, the overall income gap could be narrowed as the wealthy population would face more taxes and the groups in relative poverty could receive support from social welfare. This research believes that there are multiple ways to reduce inequality. One policy is progressive taxation, which refers to a taxation system that forces wealthy people or organisations to pay a higher proportion of taxes from their income. Apart from taxation, supply-side policy could also support the improvement of equality by providing education and training for individuals and constructing more roads, railways, stations, and airports in order to make transportation more convenient. The cost of production for firms will be lowered, which may therefore lead to a rise in average salary. Other policies are not limited to the policies mentioned above but also include ways such as strengthening trade unions, regulate merging, and encouraging foreign investment. For future studies, researchers could investigate more on the theory and usage of more complex formulas related to calculating the Gini Coefficient [10]

and other values related to income inequality, since this research used a relatively easier way of calculation due to a lack of professional mathematical knowledge.

4. Conclusion

In conclusion, Hong Kong's economic problem of income inequality among the entire population is not a recent temporary issue that appears suddenly but is actually a case that has started since the middle of the 20th century. There are a large variety of methods that could be used in applications to solve the issues, but the primary theory that supports the investigation of the features of income inequality behind all phenomena is limited in a few well-known models. This includes the Lorenz Curve, Simpson's Paradox, and the Gini coefficient. All the theories above require mathematical formulas to support them. For example, researchers often use the equation of integration to achieve the aim of obtaining the area between two lines in order to calculate the ratio between two areas. After all those steps, the value of the Gini coefficient could be obtained. However, data sometimes may "lie" to scholars at an imperceptible way. It may indicate parts of the facts but may mislead researchers to wrong conclusions. In this scenario, Simpson's Paradox was introduced to help to guide researchers in the right. This method supported people in different areas to maintain the ability of critical thinking and examine a subject from all angles. For example, the cases of Covid-19 and the ignorance of the fall in income of the first decile group in Hong Kong have both reflected the same pattern of the effect caused by the fallacy and paradox caused by the incorrect and partial usage of data. The Hong Kong S.A.R. government published its own Lorenz Curve graph, which indicates that income inequality fluctuated only slightly throughout the decade. Considering the difficulties faced by the low-income group, the government had imposed taxation and social welfare, which aim to

DOI: 10.54254/2754-1169/2026.38565

minimize income inequality. Apart from that, other solutions such as supply-side policy or monetary policy could also be applied to the economy, so this shows that practical methods and solutions are unlimited. Nevertheless, this research has various limitations, which include a lack of detailed and precise explanation of the application of formulas and other analysis of mathematical models. This is due to limited knowledge in this area. Therefore, future researchers could try to focus on that specific topic and dive deeper.

References

[1] Kakwani, N. C. (1977). Applications of Lorenz curves in economic analysis. *Econometrica, 45*(3), 719–727. https://doi.org/10.2307/1911684

[2] Kakwani, N. C., & Podder, N. (1973). On the estimation of Lorenz curves from grouped observations. *International Economic Review, 14*(2), 278–292. https://doi.org/10.2307/2525921

[3] Sprenger, J., & Weinberger, N. (2026). Simpson's paradox. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/paradox-simpson/

[4] Census and Statistics Department. (2023). *2021 population thematic report: Household income distribution in Hong Kong*. Hong Kong : HKSAR Census and Statistics Department. https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=600&pcode=B1120108

[5] Office of the Government Economist. (2025). *Hong Kong economic analysis platform*. Hong Kong: Financial Secretary's Office. https://www.hkeconomy.gov.hk/en/pdf/gini_comparison.pdf

[6] Ginevičius, R., Nazarko, J., Gedvilaitė, D., & Dacko-Pikiewicz, Z. (2021). Quantifying the economic development dynamics of a country based on the Lorenz curve. *E&M Economics and Management, 24*(1), 55–65. https://doi.org/10.15240/tul/001/2021-1-004

[7] Pearl, J. (2022). Comment: Understanding Simpson's paradox. In *Probabilistic and causal inference: The works of Judea Pearl* (pp. 399–412). Association for Computing Machinery. https://doi.org/10.1145/3501714.3501738

[8] Bonovas, S., & Piovani, D. (2023). Simpson's paradox in clinical research: A cautionary tale. *Journal of Clinical Medicine, 12*(4), 1633. https://doi.org/10.3390/jcm12041633

[9] von Kügelgen, J., Gresele, L., & Schölkopf, B. (2021). Simpson's paradox in COVID-19 case fatality rates: A mediation analysis of age-related causal effects. *IEEE Transactions on Artificial Intelligence, 2*(1), 18–27. https://doi.org/10.1109/TAI.2021.3073088

[10] Baydil, B., de la Pena, V. H., Zou, H., & Yao, H. (2025). Unbiased estimation of the Gini coefficient. *Statistics & Probability Letters, 222*, 110376.