Income Inequality in China*

Wei Zhang Yitian Li Nan An

28 February 2022

Abstract

Regional disparity data in China is pulled from the Department of Comprehensive Statistics of the National Bureau of Statistics of China. It is used to analyze the revolution of inequality in China in terms of urban-rural disparity and provinces' disparity, within the time from 1978 to 2016. In this report, we will use visualized data to explore the real situation of inequality in China and apply the time series analysis to predict further economic growth in China. In the end, we discuss the underlying economical and political reasons that cause the disparity.

1 Introduction

Between 1978 and 2015, China moved from a poor, underdeveloped country to the world's leading emerging economy. Despite the decline in its share of the world population, China's share of world GDP increased from less than 3 percent in 1978 to about 20 percent by 2015. According to official statistics, real national income per adult grew more than eightfold between 1978 and 2015(Piketty and Yang (2019)). However, China's income inequality has been one of the worst problems in the whole society, and it poses serious challenges to national unity and political stability. There is a large bias of people's income level, living level, diet level, and educational level in urban and rural and different provinces of China. The underlying factors that cause inequality include geographical location, government policy, and history.

In this essay, we collect the income-per-capita data in China within the time from 1978 to 2016, from the Department of Comprehensive Statistics of the National Bureau of Statistics of China. Then, we make data visualization to analyze the evolution of inequality in China in terms of urban-rural disparity and provinces' disparity. To make sure the consistency of exploring the data, we mainly utilized the dplyr package to process the original dataset. Filter function to ensure that we just have the numeric data left in the dataset. Rename guarantees the convenience of calling the variable. In addition, we apply the time series analysis to contribute to the prediction for further economic growth.

From the figures, we can tell that per capita household income in the East and urban areas was constantly higher than those in the Central and Western regions and rural areas, with widening gaps over time. The gap between the East and Central regions grew relatively faster in the 1980s and 1990s, whereas the gap between the East and Western regions widened more quickly after the 2000s.

Then, we collect the data of each provinces' income per capita in 2020. In order to make it visualized, we plot China's map by province, where the region with deeper colour represents the residence over there are wealthier, vice versa. The rapid growth of export-oriented manufacturing in the coastal region was accompanied by intensive regional development disparities between coastal and inland regions (Liao and Wei 2016).

Within the implementation of the open-door policy in 1978, one critical adjustment of the development policy in China has been the emphasis on the importance of regional comparative advantages. Some regions are

*Code and data are available at: https://github.com/macoyo2/Income-Inequality-in-China.

allowed and even encouraged with the mentality "get rich faster" and others to follow. This encouragement proved efficient in stimulating China's economic development and intensifying regional inequality.

The key challenge ahead is to develop a nationwide social security program that ensures a minimum standard of living and provides basic social services for people in the poorer regions. Policies aimed at facilitating growth catch-up have been helpful and should be continued (Poonpatpibul 2019). They should have access to good healthcare services and obtain better support from stronger social safety nets. Better education is also important to help young people in these regions have better job opportunities.

2 Data

The data set will be processed and analyzed in R (R Core Team 2020) primarily using the tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2021) packages. Figures and tables will be created with ggplot2 (Wickham 2016). The packages readxl (Wickham and Bryan 2019), ggpubr (Kassambara 2020), astsa (Stoffer 2021) and lubridate (Grolemund and Wickham 2011) are used to generate the R markdown report.

2.1 Data Collection Process

This dataset is made up of a mixture of data resources. The majority of data, other than Gini, is compiled by the Department of Comprehensive Statistics of the National Bureau of Statistics of China. It is a Chinese-English version book with full accurate reliable and authoritative data. It reflects the economic and social development of China as a whole and its 31 provinces, autonomous regions, and municipalities under the direct jurisdiction of the central government between 1949 and 2020. In regard to the Gini coefficient, it is from the article of (Ravallion and Chen 2005). Chen writes the paper regarding China's progress against poverty. The algorithm of Gini is pretty clear in his paper.
The publisher Zhang combined two resources, creating and merging the worksheet. He extracted the rural and urban economic data and made some sorting and computation. Finally, the result yielded would be exactly our utilized dataset.

2.1.1 Drawback

Although the data is from the official government department, the detailed data collection process is not stated and clarified. Therefore, we reasonably doubt the precision of data. For example, there may exist some bias and missing not at random data to satisfy the particular purposes. But the National Bureau of Statistics of China will not allow data published is NA, they may fill the missing part by their willingness. In addition, before the 21st century, the technology of the Chinese government is not developed, however, the number shown in the dataset is extremely accurate, which is abnormal and raises the rational suspending inference.

2.2 Data Cleaning

Via viewing the summary of the dataset, the data is pretty fantastic and possesses the goodness property. For example, there do not exist missing values and all types of data are the same. But still, this possesses two main problems. One is that the dataset has the description sentences in the bottom rows, which have different traits from all other observations. Since in the top 39 observations, the types are all numeric and meanwhile, we will analyze the numeric data in further exploration. Hence, we would eliminate the last three rows for the next stage process. Another is that the names of variables are so long, which makes the calling inconvenient.

2.2.1 Filter

For our dataset, it demonstrates the economic aspects data from the year 1978 to 2016, which is exactly the number of 39. This means we have in total 39 observations. There are 3 lines below, which is the description of data resources and the Gini coefficients. Hence, we have to cut them off to sustain the further exploration by filter function with the package of tidyverse. Otherwise, the description rows would affect the overall data analysis, for example, we are not capable of creating the plots based on the mixture of categorical data and numerical data.

2.2.2 Rename

To make the reader convenient to understand, the names of variables are all specific and long-winged. But at the same time, it makes the operation process complicated to call. Thus, we determine to change the names of parameters, making them brief and short, but remain almost the same meaning to simplify the calling process. For example, we change the `Per capita Disposable Income of Urban Households` to `Urban income`, which is clearly briefer.

2.3 Results

2.3.1 Important variables

- `Year`: This variable represents the meaning of year in the dataset. Specifically, years from 1978 to 2016.
- `Urban income`: This variable represents the per capita disposable income of urban households. Notify that the unit is based on the American dollar.
- `Rural income`: This variable represents the per capita disposable income of rural households. Note that the unit is based on the American dollar.
- `Urban-rural disparity`: This variable clarifies the multiple gaps between rural income and urban income. The algorithm is that Urban income divides the rural income, the result, yielded, is exactly the times. Notify that the larger results are, the more huge disparity would be.
- `Gini`: This variable possesses the full names of the Gini coefficient, which is collected by the National Bureau of Statistics of China. the larger Gini is, the difference between the Urban income and Rural income would be larger.

Table 1: Variables Summary

Variable	type	mean
`Year`	character	DNE
`Urban income`	numeric	8689.3
`Rural income`	numeric	2985.3
`Urban-rural disparity`	numeric	2.666
`Gini`	numeric	0.4037

As Table 1 shown, all variables are numeric types, in which they fully correspond to their definition. Since the year means is a time variable, the mean does not possess the actual meaning. In addition, there exists a tremendous difference between the per capita disposable income of urban households and per capita disposable income of rural households. In addition, the urban income is on average 2.66 times of rural income, which is no doubt that this phenomenon is drawn further apart from the initial wish of the Chinese government. Specifically, we would continue to make further analyses.

2.3.2 Plot Analysis

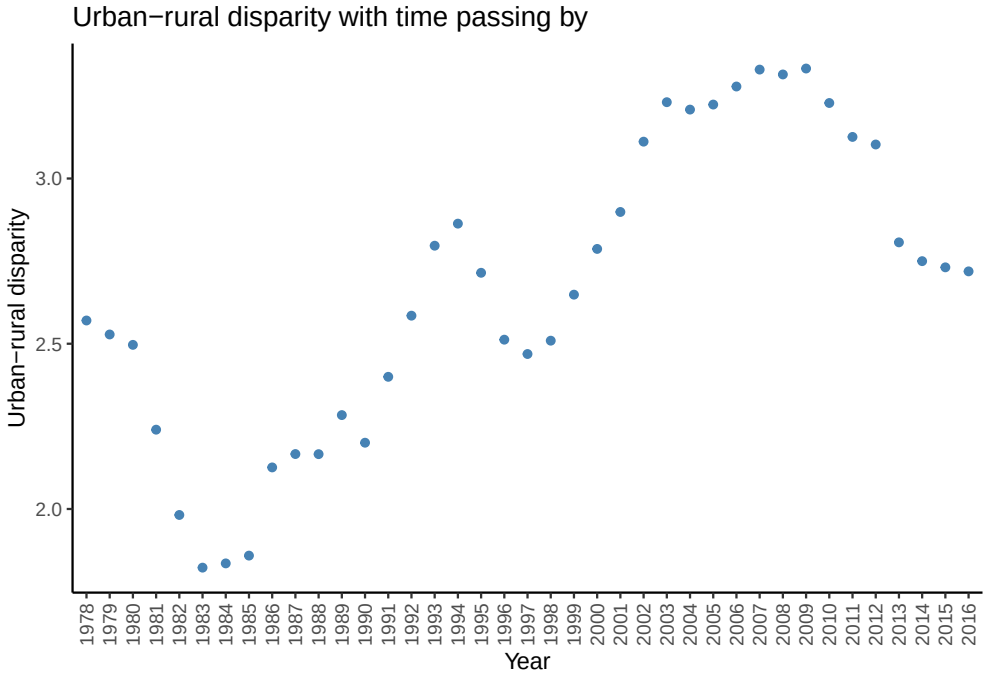
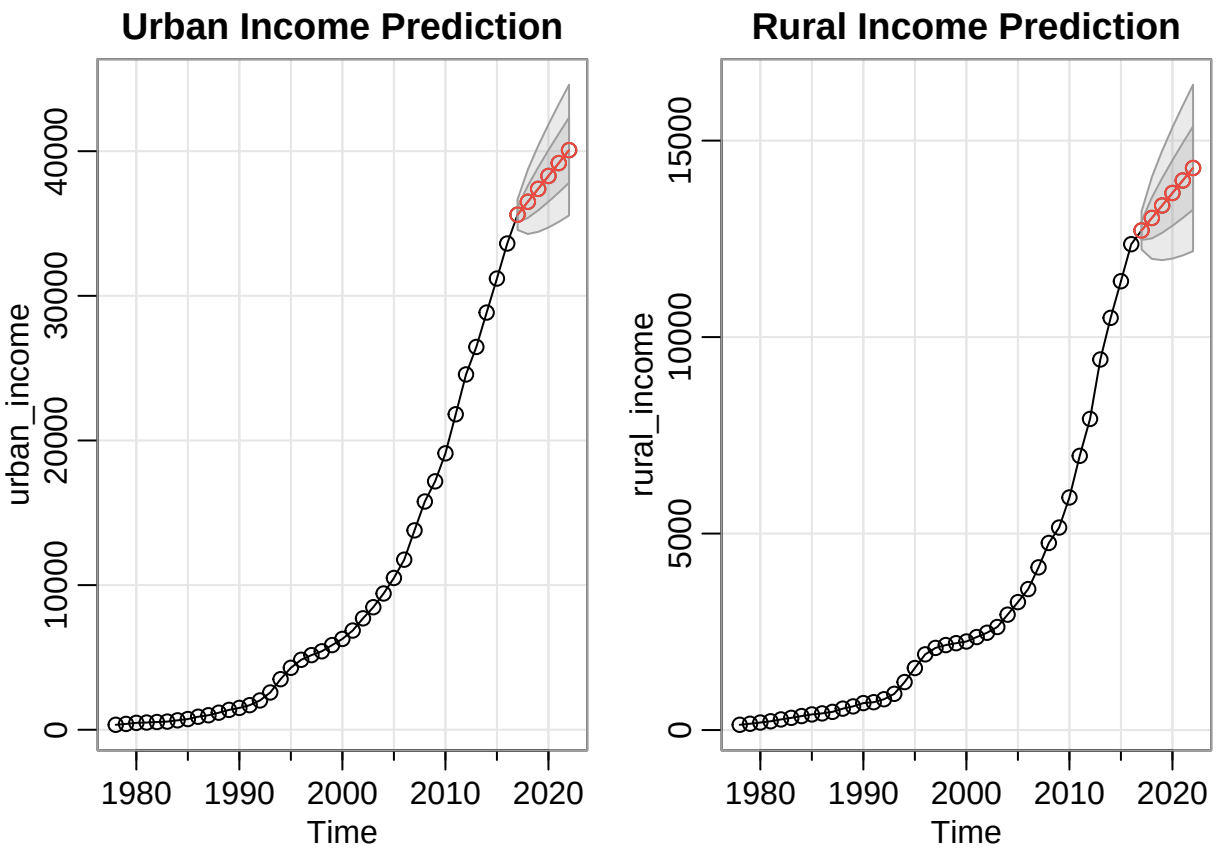


Figure 1: Urban-rural disparity in China from 1978 to 2016

In Figure 1, the x-axis represents the year, which is a period from 1978 to 2016. Y-axis represents the urban-rural disparity, in which the number is from the result based on $\frac{Urban}{Rural}$.

With the distribution of points, the urban-rural disparity altogether takes on the increasing trend but is fluctuating. Although the disparity in recent years has become milder and milder, still the difference is over two times. This shows that there exists a huge difference between the economy of urban areas and rural areas.

Below are two plots based on the time series model.



Notify that the plot on the left-hand side is the urban income with respect to years after the reform and opening-up campaign, while the plot on the right-hand side would be concerning the rural income.

We utilize the ARIMA(0,1,1) model for both `rural income` and `urban income`. Then, we apply the model on predicting the future 10 years of income. With the comparison of the two plots, they both present a homologous increasing tendency. The outcome yielded shows that the rural areas are a little more rapid than urban areas. It seems like the difference between the two areas in China has been minimized. However, is that true? Next, we would attempt to explore the answer via the direct comparison of income rather than the coefficient.

With respect to Figure 2, the x-axis represents the year after 1977. For example, x = 1 means that year is 1978. Since we have 39 observations in total, the x-axis demonstrates the year from 1978 to 2016. Besides, the y-axis represents the income per capita and the unit is the American dollar.
This plot derives that although the initial per capita income of urban and rural areas isn't that far off. But with time passing by, the disparity becomes larger and larger, where the disposable income of urban households per capital is much greater than that of rural areas. This circumstance raises people's concerns that the performance is completely contradicted with the initial targets of common prosperity, although the increasing trend is truly inspiring, which is associated with consistent economic growth. The government is necessary to take the condition into account before the intensification of contradictions of society.

2.3.3 Regional Inequality

Looking into the income inequality problem in China, the regional disparity among East, Middle, and West China cannot be avoided to analyze, because it poses serious challenges to national unity and political stability. The rapid growth of export-oriented manufacturing in the coastal region was accompanied by intensive regional development disparities between coastal and inland regions (Liao and Wei 2016). Then, we collect the data of each provinces' income per capita in 2020. In order to make it visualized, we plot

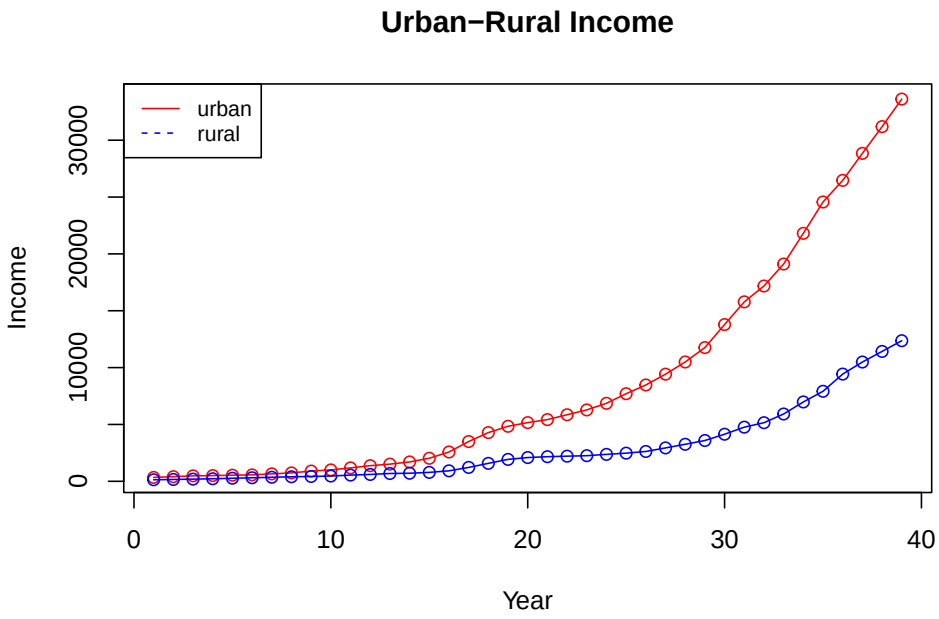


Figure 2: Urban-Rural Income in China from 1978 to 2016

China's map by province, where the region with deeper color represents the residence over there are wealthier, vice versa.

From the above map of GDP per capita in China by province, we can see the large differences of color between the East and the West, which shows the large bias of people's income level, living level, diet level, and educational level in different regions of China. The underlying factors that cause the inequality would include geographical location, government policy, historical reasons.

Firstly, seeing the geographical location in the map, we find that lower-income provinces are typically located in land-lock areas, where they share borders with low-income neighboring economies, but the higher-income provinces are typically in coastal areas, have lower transportation costs when doing international trades overseas. Looking deeper, many lower-income provinces are even situated in mountainous or desert areas, such as Xizhang and Xinjiang, which are not conducive for urbanization and manufacturing activities. Furthermore, it is costly to develop infrastructure and increase their connectivity to other regions (Poonpatpibul 2019). Coastal cities, on the other hand, are suitable for urbanization and infrastructure development, including seaports and close to much richer neighboring countries, hence having more exposure to manufacturing and trade. For example, Guangdong, Fujian, Zhejiang, and Jiangsu, have become manufacturing hubs, benefiting from China's opening-up to international trade as well as the rapid growth in national incomes.

Secondly, the role of local governments in leveraging geographical advantages has also been important. While the central government has been strategic in enabling the populous coastal areas to develop quickly to lift up the country's overall development, local officials in the coastal areas have also been very driven in attracting foreign direct investment, and facilitating the establishment of local enterprises that process manufactured goods and re-export them (Poonpatpibul 2019), which allows their income per capita higher.

Thirdly, labor and technology factors are determined. Once regional disparities start to develop, further divergent paths tend to kick in. As coastal areas boomed, these areas developed faster and have more opportunities than other areas. More and more skilled workers and entrepreneurs are attracted to come here, which are beneficial for the further fast development.

To see the evolution of regional disparity in China much more clearly, we plot a graph of 3 lines that represents

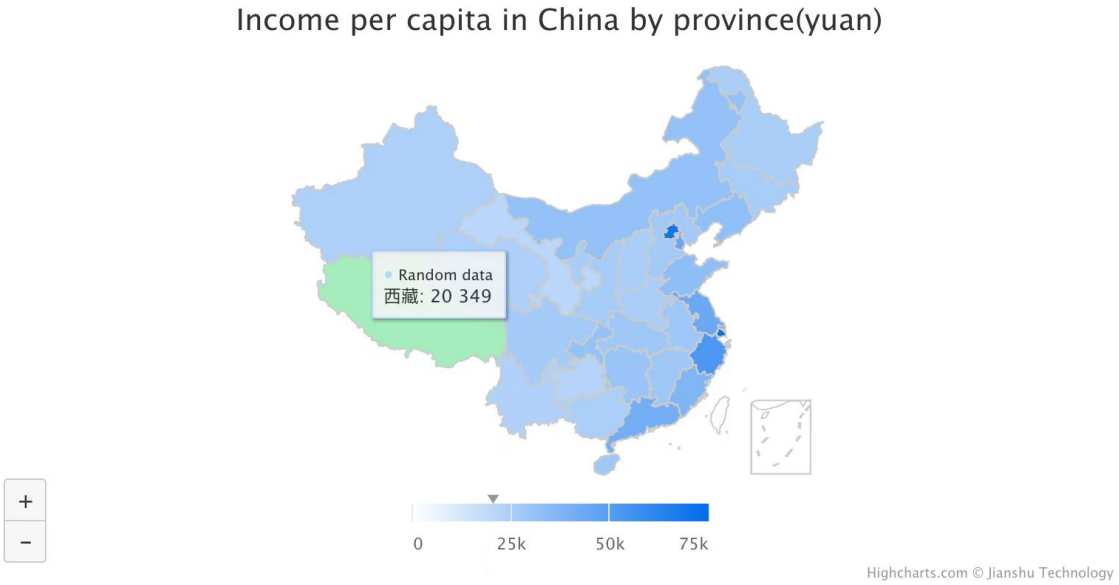


Figure 3: Income per capita in China by province(2020)

the ratio of per capita income between East, Central, and Western regions. Per capita income in each region is a weighted sum of per capita urban disposable income and rural net income with the population of urban and rural as weights. The following graph illustrates the evolution of income gaps among the East, Central, and Western regions in China from 1978 to 2016.

From the above Figure 4, we can tell that per capita household income in the East was constantly higher than those in the Central and Western regions, with widening gaps over time. The gap between the East and Central regions grew relatively faster in the 1980s and 1990s, whereas the gap between the East and Western regions widened more quickly after the 2000s. Throughout the economic transformation, reform policies favoring the coastal provinces were widely criticized for increasing the regional gaps. One main consequence of regional inequality is the rising opportunity inequality, which in turn enhances the regional income disparity. In the years after 2009, the gap among the East, Central, and Western regions has become stable and has slightly decreased due to the Western Development Strategy initiated in 1999 to rebalance the development among the three regions in the new century. While regional inequality between provinces fluctuates more strongly, interregional inequality between the Eastern, Western, and Central regions keeps increasing. However, both interregional and interprovincial inequalities have declined substantially since the global economic crisis in 2008 (Liao and Wei 2016).

This regional inequality situation is closely related to the policy and history of the national economic development in China. The fact is that during the 18 years after the establishment of the People's Republic of China (PRC), the GDP of China has been less than 200 billion (Hanrahan 2022). The residence poverty problem had been one of the biggest challenges in China at that time. Within the implementation of the open-door policy in 1978, one critical adjustment of the development policy in China has been the emphasis on the importance of regional comparative advantages. Some regions are allowed and even encouraged with the mentality "get rich faster" and others to follow. This encouragement proved efficient to stimulate economic development in China. Eleven years after the implementation of the open-door policy, the GDP of China has reached 456 billion (Hanrahan 2022), more than 2 times the GDP before the implementation of the policy (Hanrahan 2022). However, it greatly favors coastal provinces because of their geographical advantages, thereby widening the income disparities among East, Central, and Western China (Association 2021).

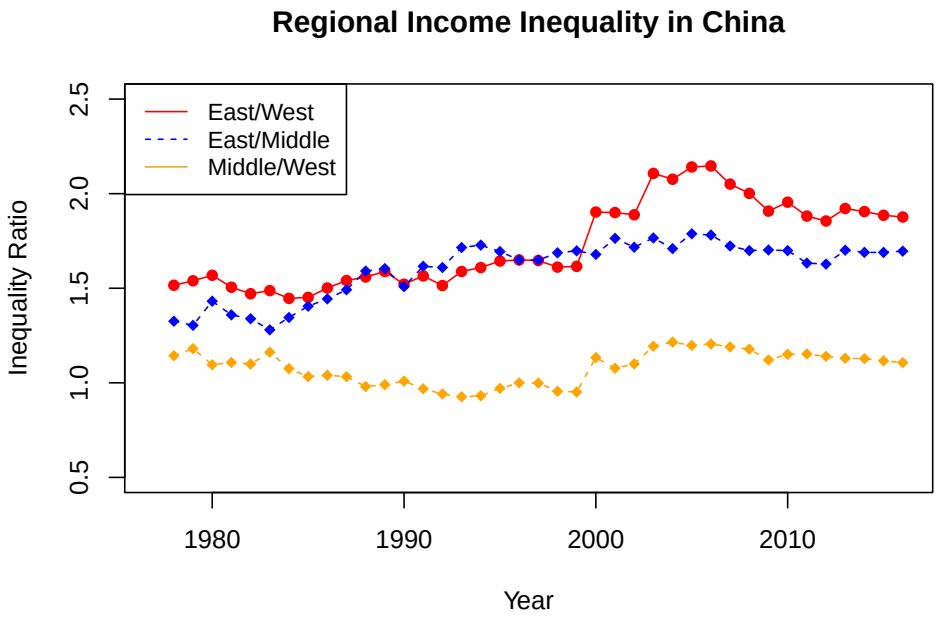


Figure 4: Regional income inequality in China among East/West, East/Middle and Middle/West

This analysis further unfolds that regional inequality is influenced by multiple factors, and in particular by China's transitions' decentralization, marketization, and globalization. Overall, the uneven regional development in China is embedded in its specific political contexts and rooted in the transitional nature of the reform. It has been increasingly subject to global economic integration and external shocks (Liao and Wei 2016). We still need to pay attention to the multi-scalar regional inequality and the investigations of the dynamics and the various mechanisms that affect regional inequality, helping to release the problem of regional disparity in China.

3 Discussion

Since 1978, both Chinese urban and rural districts had a dramatic increase in income. The startling income growth mainly attributes to the implementation of reform and opening-up policy introduced by Chinese leader Xiaoping Deng. The "internal reform" in this policy implements a mixed economic model and the household contract responsibility system in agriculture. It overturned the planned economy system before 1978 that the Chinese government controlled and managed social resources. The reform introduced the principal of market economy in two stages. The first stage from the late 1970s to the early 1980s involved abolishing the agricultural collective system and opening the local market to foreign capital, allowing local entrepreneurs to start a business. The second stage was carried out from the late 1980s to the 1990s. Its policy includes the reform of the land system and the privatization of state-owned enterprises. The government no longer controls prices and abolishes some protectionist policies. Nevertheless, the state still controls key industries such as the banking and petroleum industry. Additionally, the "opening to the outside world" in the reform has enabled many overseas enterprises to enter the Chinese market, and a considerable number of Chinese enterprises have opened up overseas markets. In the context of economic globalization, since the reform and opening up, for many reasons, the United States, Japan, France, Germany, and Britain have provided various forms of long-term assistance to China, including free assistance. Hongkong, Macao, and Taiwan also provide a lot of capital and technology to the Chinese mainland. By 2009, China had received nearly $6.7 billion in free aid from developed countries (Ballawar 2021).

However, the disparity between urban and rural income was still critical after nearly 38 years of development. From 1998 to 2016, the Gini coefficient in China is above the warning line of 0.4, indicating that the income gap between urban and rural is large and also serious. Besides, the urban area seems to benefit more from the new policies with a more rapid increase in the economic income. In order to analyze the current situation, this paper makes a further discussion of several reasons for this income gap in China.

3.1 Rural Development Obstacles

After the reform and opening up, China's residents' income has generally been greatly improved, but the rural development is not optimistic. First of all, the rural infrastructure is backward and the transportation facilities are not perfect. It is difficult to use scientific progress to drive the development of productive forces. Secondly, farmers are mainly engaged in traditional industries related to the cheap labor force, and the productivity is low. Compared with this, the urban population is mostly engaged in mental labor, which widely exists in modern industries such as the information industry, manufacturing industry, and service industry. Although the government has taken a series of measures such as financial transfer payments in recent years to try to reduce this expansion trend, on the whole, the income gap between urban and rural areas is still widening. The gap between urban and rural areas is mainly reflected in the income gap between urban and rural residents. Analyzing and comparing the per capita income and income source can clearly reflect the gap between the rich and the poor between urban and rural residents.

3.1.1 Unreasonable Urban and Rural Structure

China's long-term urban-rural dual economic structure system is the fundamental reason for the slow growth of farmers' income and the large income gap between urban and rural areas (Zhang 2018). The city's registered residence system is mainly reflected in two aspects, such as the two-element structure, which is the long-term existence of the urban and rural household registration system. The rural population's free movement to the city is restricted by registered residence, which leads to a large surplus labor force in the countryside. Even though the registered residence reform in China has allowed the transfer of agricultural workers to the city, the city population is obviously better than the farmers in terms of medical treatment, social security, and welfare. As far as the living environment is concerned, the rapidly rising house prices are beyond the reach of many migrant workers who have no foundation. In terms of their wage income, the per capita wage income of urban residents was 20665 yuan in 2016 and that of rural residents was 5021 yuan, with a difference of about four times (Hanrahan 2022). On the whole, urban economic growth and urban living standards are higher than those in rural areas. The existence of this phenomenon will lead to the richer urban rich and the poorer rural poor. On the other hand, rural areas are mainly engaged in agricultural labor dominated by primary productivity, and urban residents are mainly engaged in labor dominated by secondary and tertiary industries. Therefore, at present, China's rural labor productivity is far lower than that of urban labor productivity, resulting in the widening income gap between urban and rural residents. After the reform and opening up, the production efficiency of various industries in China's cities has been greatly improved. However, due to the existence of urban-rural dual economic structure, many rural areas cannot share advanced information technology, the improvement of agricultural production mode is limited, the reasonable flow of production factors between urban and rural areas is blocked, and rural areas are backward and scattered compared with urban areas (Zhang 2018).

3.1.2 Inclined Policies

The government's macro-control policy has a great correlation with the problem of the large gap. On the one hand, since the reform and opening up, China has put forward the policy of getting rich first and then driving the later people to get rich (Zhang 2018). This policy has indeed brought significant effects to China's economic development. Especially after the reform and opening up, the active fiscal policy inclined to cities has played a great role in promoting urban residents. However, after more than 30 years of development, the driving effect of the first rich group on the later rich group is not as obvious as expected, and the problem

of increasing farmers' income has not been well solved, which has exacerbated the income gap between urban and rural areas. On the other hand, according to the distribution principle of "giving priority to efficiency and giving consideration to fairness" put forward by China, the economic policies formulated first pay attention to the improvement of efficiency, and pay little attention to fairness. Although the state has increased the scope and intensity of agricultural tax relief in recent years, China's distribution principle will inevitably lead to the widening gap between the rich and the poor in practice.

3.1.3 Uneven Distribution of Educational Resources

Education level has a significant impact on income. A person's income level is positively correlated with his education level. The higher the degree of industrialization and marketization, the closer the correlation is. People's education level has greatly affected the polarization between the rich and the poor. Although China has been actively responding to social development in deepening the reform of the education system and increasing investment in education and has achieved positive and remarkable results. However, there are still a series of drawbacks in China's education system. The distribution of educational resources is uneven, and the gap of educational resources between urban and rural areas is particularly obvious. Cities enjoy very superior educational equipment and teachers, while the gap between educational equipment and teachers in rural areas is obvious compared with that in urban areas. The gap between the rich and the poor in education will inevitably lead to the gap between the rich and the poor in the economy (Zhang 2018).

3.2 Future study and Weakness

In this report, our data is only available until 2016 and in recent years, China also implemented a lot of changes regarding income inequality. This analysis can not reflect the latest development of China and further studies should be conducted.

The income inequality in china will still exist in the following years. The Chinese government is making efforts to solidify the achievements in poverty alleviation and common prosperity. They require to solve the rural problems with new strategies. They should try to break the dual structure of urban and rural areas and promote urban-rural integration. Also, the leaders need to pay equal attention to fairness and efficiency. More importantly, the authority should be aware of the need of deepening education system reform and promoting equity. Future studies can focus on the steps or strategies taken by the Chinese government and analyze with the newest data of urban and rural income to see whether the situation has an improvement or not.

References

Association, American Economic. 2021. *A Survey on Income Inequality in China.* https://www.aeaweb.org/articles?id=10.1257/jel.20201495&ArticleSearch%5Bwithin%5D%5Barticletitle%5D=1&ArticleSearch%5Bwithin%5D%5Barticleabstract%5D=1&ArticleSearch%5Bwithin%5D%5Bauthorlast%5D=1&ArticleSearch%5Bq%5D=&JelClass%5Bvalue%5D=0&journal=2&from=j.

Ballawar, Nichole. 2021. *China - a Brief History of Development Assistance.* https://diplomatist.com/2021/03/17/china-a-brief-history-of-development-assistance/.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Hanrahan, Laura. 2022. *Construction of Rental Apartments in the GTA Hits a 30-Year High.* https://dailyhive.com/toronto/construction-rental-apartments-gta-30-year-high.

Kassambara, Alboukadel. 2020. *Ggpubr: 'Ggplot2' Based Publication Ready Plots.* https://CRAN.R-project.org/package=ggpubr.

Liao, Felix Haifeng, and Yehua Dennis Wei. 2016. *Regional Inequality in China: Trends, Scales and Mechanisms.* https://www.rimisp.org/wp-content/files_mf/1473279216202_FH_Liao_YD_Wei.pdf.

Piketty, Thomas, and Li Yang. 2019. "Income Inequality Is Growing Fast in China and Making It Look More Like the US." https://blogs.lse.ac.uk/businessreview/2019/04/01/income-inequality-is-growing-fast-in-china-and-making-it-look-more-like-the-us/.

Poonpatpibul, Chaipat. 2019. *From Shanghai to Gansu: Growing Regional Disparities in China - Causes and Remedies.* https://www.amro-asia.org/from-shanghai-to-gansu-growing-regional-disparities-in-china-causes-and-remedies/.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ravallion, Martin, and Shaohua Chen. 2005. *China's Progress Against Poverty.* https://issuu.com/gkillcity/docs/ravallion____chen___2007_/1.

Stoffer, David. 2021. *Astsa: Applied Statistical Time Series Analysis.* https://CRAN.R-project.org/package=astsa.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2019. *Readxl: Read Excel Files.* https://CRAN.R-project.org/package=readxl.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Zhang, Rui. 2018. *A Study on the Gap Between Urban and Rural Wealth in China.* https://www.hanspub.org/journal/PaperInformation.aspx?PaperID=25951&#abstract.