

The Impact of Agricultural and Industrial Development on Income Inequality in Rural China

Yuanyuan Chen *

October 20, 2005

Abstract

This paper assesses the contribution of the agricultural and industrial income for the rise of rural household income inequality in China by counterfactual decomposition techniques. Using the Chinese Household Income Project data in 1988 and 1995, I find that the rise in inequality reflected both the stagnant development of agriculture and the rural industrialization. Most of the rise inequality was accounted for spatial inequality, the increase of which reflected growing unevenly regional development of industrialization.

JEL Classification: J22, J21, C2

*Author's Email Address: ychen@imf.org

I am grateful to Peter Gottschalk, Author Lewbel, Donald Cox and all participants in the seminar at Boston College Economics Department for helpful guidance and comments. All errors are mine.

1 Introduction

Since the late 1970s, the average household income of rural areas, as well as the income inequality in China, has increased substantially. According to the official statistics, the Gini coefficient in rural China was .21 in 1978 and increased by 43 percent to .30 in 1990. It then continued to rise to .34 in 1995.¹

As a large country as China, the diffusion of economic reform over the whole population takes a long time. Targeting on the ideology of egalitarianism in income distribution led to a low efficiency in the economy. Therefore, in the initial stage of the reform, policies were intended to overcome egalitarianism, promote efficiency, increase mean in income distribution and ultimately realize common prosperity based on an enlarged “pie”. However, the increase in inequality was much more rapid than originally expected at the outset of the reform.

The rising in income inequality has been a subject of great interest to both researchers and policy makers. A lot of researches (Hare, 1994; Knight and Song, 2000; Rozelle, 1994 and 1996; Tsui, 1995 and 1998; Zhang, 1992; and Zhu, 1992) have focused on the extent to which the increase in inequality is accounted for changes in the sources of income. In particular, does the increase in inequality reflect the growth of agricultural income under the system of private household farming or the growth of nonagricultural income with rural industrialization?²

While it is commonly accepted that the rise in inequality is a consequence of the expansion

¹Li, et al. (1997).

²Industrial income in this paper is defined as wages, dividends and welfare from township, village, and private enterprise.

of rural industry (Hare 1994; Knight and Song 2000; Rozelle 1994, 1996; Tsui 1995, 1998; Zhang 1992; and Zhu 1992), there is no consensus whether it can also be attributed to the imbalance of growth in agricultural sector. Rozelle (1994) decomposes the Gini coefficient by agricultural income and industrial income in rural Jiangsu province in 1983-1989 and suggests that increasing the importance of agriculture in the economy leads to a reduction in inequality; expanding rural industry gives rise to greater inequality. Knight and Song (2000) uses a national wide dataset and obtains the same conclusion. These findings, however, contrast with the decomposition results of Cheng (1996) showing that inequality in crop income is the major source of inequality based on the analysis of five provinces in grain producing areas. More recently, Tsui (1998) argues that “while nonagricultural additivities have exerted an increasingly important influence on rural income inequality, income from agriculture still contributes significantly to overall inequality. Hence, policy makers should not overlook the inequality reducing potential of policies focused on agriculture”.

While these studies differ in several important aspects (such as spatial coverage of data, aggregate or micro level of data, years considered, definition of income), “contribution” of income components to overall inequality is commonly obtained by using decomposition techniques proposed by Shorrocks (1983) and Lerman and Yitzhaki (1985). The same methodology has been widely used in other studies of China’s income inequality (Coady and Wang, 2000; Khan, et al., 1992 and 1998).

The analysis based on standard decomposition techniques, however, has been questioned by Cancian and Reed (1998 and 1999) in the framework of studying the impact of wives’

earnings on family income inequality. They argue that the impact of an income source can only be measured meaningfully by comparing the observed distribution with a reference distribution. The standard decomposition component has no implicit reference distribution. Therefore, it should not be interpreted as a meaningful measure of the contribution of an income source to overall inequality.

This paper introduces the concept of Cancian and Reed (1998 and 1999)'s counterfactual decomposition into the analysis of the impacts of income sources on overall income inequality and spatial inequality in China. Three measures of the impacts of the income components are described based on counterfactual reference distribution: (1) the impacts of agricultural income and industrial income on the overall rural inequality at a point in time, (2) the impact of the change in each income component on the rise in inequality over time, (3) the impact of each income component on spatial inequality.

It is necessary to point out that the measures of the impact of particular income source on the distribution of household income in this paper are fundamentally accounting concepts. I do not consider the behavioral linkage among each income source. In particular, in the absence of agricultural activities, household might have more time for nonagricultural operation or off-farm employment, so that the true counterfactual distribution would differ from the observed total income minus the agricultural income. Estimating this "indirect" effect requires modelling on behavior relationships, which is beyond the scope of this paper. The limitation should be kept in mind when interpreting the implication of the results on the impacts of income sources.

Counterfactual analysis show that agricultural income was an equalizing income source whereas rural industrialization was the opposite. Rise in inequality over time reflects both diminishing equalizing impact of agricultural income and increasing disequalizing impact of industrial income. In addition, the spatial analysis show that the equalizing impact of agricultural income was as large on within-regional as on between-regional inequality. The growing regional inequality can be largely explained by the increasingly uneven regional development of industrialization.

Therefore, the rise in inequality was attributed to both the stagnant development of agriculture and development of rural industrialization. Spatial inequality accounted for most of the rise in inequality. And rise in spatial inequality itself reflected growing uneven regional development of industrialization.

The paper is divided into six sections. The following section begins by generally describing China's rural economic reform and its impact on rural income inequality. The third section discusses the decomposition techniques. Data and definition of income sources are presented in Section 4. Section 5 shows the results and the last section concludes the paper.

2 Rural Economic Reform in China

Before proceeding to the analysis, it may be useful to briefly describe the economic reform of rural China and to explore the economic reasons for the rise of income inequality. China's rural economic reform started in 1978, when a household contract responsibility system was first adopted. Before that, the Chinese government practised a fully-planned macro

policy favoring heavy industry. The state imposed a tremendous burden on the peasantry in terms of indirect taxation through mandatory procurement of agricultural products. All farmers worked on state-owned or collective-owned land and got equal low payments. Rigid regulations restricted farmer's non-farm employment opportunities and farmers could not produce according to local market demands because of forced production quotas and state intervention in food and agricultural market. Rural economy and the living standards of farmers stagnated for more than two decades before 1978, with income equally distributed among farmers.

The household contract responsibility system started in 1978 and allowed the state-owned land being contracted to individual households on egalitarian principles. Farmers obtained income by selling their household products to the states and meanwhile the government continued to raise state purchase prices. As a result, agricultural production in 1978-1984 achieved the greatest improvements in the history, by increasing 7.4 percent per annum, according to Huang (1998). The development of rural economy was accompanied by a rise in income inequality. Gini index of rural area increased from .21 in 1978 to .26 in 1984 (see Li Shi et al., 1997).

However, since the mid 1980's, the agricultural economy as an engine of growth has apparently run out of steam. Growth rates in agriculture dropped back to 3.8 percent during the period of 1984-1995. The following are the two most important reasons for this phenomenon. First, the household contract responsibility system itself resulted in sharp decline in farmers' long-term investment in land. It can be explained by the nature of land

contracts. According to the law, farmers do not own the land, and therefore they were reluctant to make any significant long-term investment in land because of the expectation that the land will one day be taken away from them. Instead, they invested in housing and non-agricultural activities. Second, agriculture is low-profit production, due to both the high agricultural tax and the government interference in food prices to protect urban consumers. The net subsidy for agricultural production – by agricultural tax - was actually negative, ranged from -40 to -60 percent of producer value in the mid 1980s, reflecting severe discrimination against agriculture. After many years of successful agricultural reform, farmers lost one-third of the product value due to various distortion policies.

On the other hand, a number of policies were implemented in the mid 1980's to encourage diversification of non-agricultural, non-grain production and rural industrial development, with many township and village enterprises established during that period. Rural economy experienced the structural change with a falling share in agriculture and increasing in light industry.

Rural income inequality during this period continued to rise. The Gini index of rural area rose from .26 in 1984 to .34 in 1995. The remainder of this paper explores the extent to which the increase in inequality reflected agricultural development and industrialization.

3 Decomposition Method

In the spirit of Cancian and Reed (1998), this paper measures the impact of a particular income component on overall inequality by comparing actual income distribution with its

counterfactual reference distribution. The following section compares the decomposition by income components proposed by Shorrocks (1983) with Cancian and Reed (1998), and also describes decomposition technique by regional groups.

3.1 Decomposition by income components

The coefficient of variation (CV) is used as the summary measure of the level of inequality in this paper. CV is defined as the standard deviation of income divided by the arithmetic mean.

In the Shorrocks (1993) decomposition, the coefficient of variation of total income is decomposed as:

$$CV_t \equiv S_a R_a CV_a + S_i R_i CV_i + S_o R_o CV_o \quad (1)$$

where for each source k , $S_k = \mu_k / (\mu_a + \mu_i + \mu_o)$, CV_k is the coefficient of variation of income component k , and R_k is the correlation between income source and total family income. S_k is the share of total family income from component k , and μ_k is the mean of income from component k . The subscription a denotes agricultural income, i industrial income, and o other income. The decomposition term of $S_k R_k CV_k$ is interpreted as the contribution of source k to income inequality.

Cancian and Reed (1998) argues that the terminology “contribution” is misleading.³

³They argue against the Gini coefficient decomposition derived by Lerman and Yitzhaki (1985): $G_t = \sum_k S_k R_k G_k$, where the notation is similar as (3.1), except the inequality index is the Gini coefficient. And the Gini correlation of source k is defined as the covariance between income from source k and the rank of total income, divided by covariance between income from source k and the rank of source k . I choose to use CV

The word *contribution* suggests that if the term corresponding to income source k is positive ($S_k R_k CV_k > 0$), then income source k increases inequality. But compared to what? The source of the problem is that the form of decomposition has no implicit reference distribution. For example, if $S_a R_a CV_a$ is excluded from equation (1), the remaining terms ($S_i R_i CV_i + S_o R_o CV_o$) do not add up to $CV_t - S_a R_a CV_a$, since R_i and R_o are the correlation between the corresponding income source and *total* income, the latter of which would change if one of its income component changed.⁴

To allow for a counterfactual analysis, Cancian and Reed (1998) proposed the following decomposition technique:

$$\begin{aligned}
 CV^2(y_t) \equiv & S_a^2 CV_a^2 + S_i^2 CV_i^2 + S_o^2 CV_o^2 + 2\rho_{ai} S_a S_i CV_a CV_i \\
 & + 2\rho_{ao} S_a S_o CV_a CV_o + 2\rho_{io} S_i CV_o CV_i
 \end{aligned} \tag{2}$$

where S_k, CV_k are the same as the definition for equation (1), ρ is the correlation *between a pair of income components*. Here, total inequality is fully decomposed by source of income, translating as that it avoids using the correlation of income component k with the total income, the part of which is through k . The counterfactual analysis can now be conducted by comparing the observed distribution with counterfactual reference distributions which for a better comparison of the results between the two methods.

⁴The share of mean of each component would also change, but it can be easily obtained by recalculating the summary statistics. A calculation in counterfactual 3 would be a more complicated but defensible example, since it is impossible to obtain counterfactual correlation between agricultural earnings and total earnings had agricultural earnings not changed since the base year.

can be precisely obtained through equation (2).

To demonstrate that counterfactual analysis is a better measure of the effect of income source k on total inequality than the decomposition component $S_k R_k CV_k$, Cancian and Reed (1998) suggests a hypothetical situation in which an income source, such as agricultural income, is equal across all households. In the absence of agricultural income, the distribution of household income would become less equal. In this sense, the agricultural income are equalizing. However, in the standard decomposition, the contribution of agricultural income in the form of (1) is zero ($CV_a = S_a R_a CV_a = 0$), indicating that agricultural income have no impact on household income inequality.

3.2 Decomposition by regional groups

Decomposition by income sources and decomposition by subgroups are two distinct concepts. The former examines the contribution of an income source to total inequality, the latter investigates the extent to which within group and between group inequality accounts for total inequality. To assess the impact of an income source on regional inequality requires both. In this paper, the square of the coefficient of variation is decomposed by T subgroups as follows:⁵

$$CV_t^2 \equiv \sum_{i=1}^T P_i (cv_i^2) \lambda_i^2 + \sum_{i=1}^T P_i (\lambda_i - 1)^2 \quad (3)$$

⁵Stephen P. Jenkins (1995) proposes the decomposition of half of the squared coefficient of variation, $CV^2/2 = \sigma^2/2\mu^2 = \sum_{i=1}^N P_i (cv_i) \lambda_i^2 + \sum_{i=1}^N P_i (\lambda_i^2 - 1)$, I choose to use CV^2 to be consistent through the paper.

where cv_i is coefficient of variation of subgroup i , P_i is the population share of group i , λ_i is subgroup i 's mean relative to the population mean. The first term is within group inequality, which is determined by the sum of inequality within each group weighted by their population proportion and share of means. The second term is between group inequality, determined by the subgroup mean relative to the population mean, and the population proportion in each group. The derivation of (3) can be found in the appendix.

4 Data

This paper uses cross-sectional data from a survey of rural incomes in 1988 and 1995.⁶ The sample includes 10,258 rural households with 51,352 individual members in 28 provinces in 1988 and 7,998 rural household with 34,739 individuals in 19 provinces in 1995. The collection of sample was organized by an international group of economists working with colleagues in the Economics Institute of the Chinese Academy of Social Sciences. It is now possible, for the first time, to estimate income according to its standard international definition (Riskin, et al, 1998). Moreover, with its widely spatial coverage, it provides the best data currently available for addressing the income distribution picture in China as a whole.⁷

⁶The data set is obtained from Riskin, Carl, Zhao Renwei, and Li Shi. Chinese Household Income Project, 1988 and 1995 [Computer file]. ICPSR version. Amherst, MA: University of Massachusetts, Political Economy Research Institute [producer], 2000. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2000.

⁷For more detail information on the datasets, please refer to Khan et al. (1992, 1998, 2000). Compared to the statistics from Chinese State Statistical Bureau (SSB), per capita rural income in our data is 760

In this paper, rural disposable household income is divided by four income sources: (1) agricultural income from household production (household net cash income from the sale of the products of household farming production and market value of self-consumption of farm products, minus the cost of inputs). (2) industrial income which includes wages and other compensation received from township, village, collective and state and other enterprises. (3) nonagricultural household self-employed income (household net cash income from the sale of the products of industrial and subsidiary activities, and income from family small business, minus the cost of inputs). (4) other income includes: (a) household income from property: income from interest, rent or leasing out land and other property, and market-rental value of housing, (b) net transfer from and to the state, local government and the “collective”, and (c) private transfer payments, the major part of which are transfers from household member working in urban area.

The proportion of each source to the total income are shows in Table 1. All income amounts are in 1988 yuan (RMB), using the rural consumer price index (CPI) of the SSB (Chinese State Statistics Bureau), which is 220.1 in 1995 (1988=100).⁸ Household production accounts for a very large proportion of income, and most of this income source comes from agricultural operation. With the deepening of rural economic reform, the proportion of yuan, about 39 percent higher than the SSB estimate of 545 yuan in 1988. The difference is almost entirely because of the difference in definition rather than the difference in measurement. Using the same definition of income as SSB definition yield per capita income of 548 yuan.

⁸Although it is said to understate the true increase in cost of living. I choose to use it to be consistent with other studies using this dataset (Khan et.al, 1998).

agricultural income source decreased, from 74 percent in 1988 to 56 percent in 1995, and it was replaced by the increasing share of wages and compensation from off-farm employment.

5 Results

This section begins by showing the result obtained by standard decomposition of Shorrocks (1983) as a benchmark, and then focuses on counterfactual analysis. The impacts of agricultural income and industrial income on overall household income inequality are assessed separately.⁹

5.1 Standard decomposition by income components

As a benchmark, Table 2 provides the results by standard decomposition method. The first three columns show the coefficient of variation, the share of total income and the correlation between income source and total income. The inequality of agricultural income, measured by CV , was the lowest among all income sources. In 1988, the cv of agricultural income was .82, compared with 3.17 for industrial income. The development of agricultural activity were stagnant over the period between 1988 to 1995. The amount of agricultural income decreased from 2439 to 2116 yuan in real terms, and its share to total income fell from 64 percent to 46 percent, indicating that more and more rural labors flowed from agriculture

⁹In this paper, I present the results of decomposing total household income. The similar results are obtained by decomposing income per capita, as CV is invariant to scale. The results are available from the author upon request.

to rural industrial sectors.

The changes of the inequality within each income source show that while the *cv* of agricultural income slightly increased from .82 to .92 in 1995; the *cv* of industrial income, decreased from 3.17 to 2.88 in 1995, which was a 10 percent decline. The reason why industrial income became more equalized was that most of the rural enterprises were labor intensive and offered relatively equal wages. The inequality within the income source from wage earnings were due to the differences in employment opportunities. As the developing of rural industrialization, more and more families had opportunities to work in firms and, therefore, the inequality within the income source from wage earnings decreased.

Column (4) and (5) show the contribution of each income component in numeric term and in percentage. By using the standard decomposition, the contribution is calculated as the product of the share, the inequality of an income source and its correlation with total income. In 1988, agricultural income contributed to more than half of the inequality, even though agricultural income had the lowest *CV*. Most of the contribution was through its high share of total income of 64 percent. On the other hand, the industrial income contributed only 24 percent to the total inequality, ignoring the fact that industrial income had the highest *CV*. The story in 1995 was different. Agricultural income only contributed 17 percent to total inequality, as its share of total income decreased; the contribution of industrial income increased to 64 percent.

In sum, the results using standard decomposition show that both income sources contributed to the overall income inequality, with a decreasing impact of agricultural income

and an increasing impact of industrial income. But this reflects only the decline/rise in the proportion of an income source, not a change in the inequality of this source.

5.2 Counterfactual analysis on static and dynamic income inequality

In counterfactual analysis, the impacts of agricultural income are assessed by comparing the observed level of inequality of total household income, measured by the coefficient of variation, with three alternative reference levels of inequality that would have prevailed if,

1. there was no agricultural income ($\mu_a = 0, CV^2(y_t) = S_i^2 CV_i^2 + S_o^2 CV_o^2 + 2\rho_{io} S_i CV_o CV_i$)
2. the mean (μ_a) and dispersion (CV_a) of agricultural income had not changed
3. the mean (μ_a), dispersion (CV_a) and correlation of agricultural income with other sources (ρ_{ai} & ρ_{ao}) had not changed.

In each case, if observed inequality is less than the level of counterfactual distribution, agricultural income is considered as an equalizing income source. The first counterfactual distribution measures the impact of agricultural income on static overall income inequality and spatial inequality, the next two measure the impact of changes of agricultural income on the rising inequality. A similar analysis is applied to industrial income.

Table 3 shows the result of counterfactual (1). If there was no agricultural income and all other components of the CV of household income remained unchanged, the CV of total

income would have been 1.46 in 1988, compared with the observed .77, suggesting an equalizing impact of agricultural income of 47 percent. As agricultural income lost its share in total income, its equalizing impact decreased from 47 percent in 1988 to 41 percent in 1995. Under the same measure, industrial income increased overall inequality by 4 percent in 1988 and dramatically by 29 percent in 1995, since it became a more important income source for the farmers.

In addition in examining the impacts of income components on household income inequality in a given year, I now examine the extent to which the changes in these income components accounted for the rise in household income inequality. I use counterfactual (2) and (3) to calculate the level of inequality that would have occurred if a particular income source had not changed. First, the observed CV of total income is compared with the CV of total income that would have occurred had the mean and CV of the particular income component in 1995 not changed, based on year 1988.

The contribution of agricultural income and industrial income are shown in panel a and b of Table 4 respectively. The coefficient of variation of household income was .77 in 1988 and increased to 1.08 in 1995, suggesting an increase of 40 percent in income inequality. Had the mean and CV of agricultural income of 1995 remained at the same level as in 1988 while all other income inequality components changed to 1995 levels, the CV of total income would have been 1.0 in 1995. In this context, the changes of agricultural income increased the CV of total income by 9.5 percent, as it would have been lower if the distribution of agricultural income had not changed. The changes of industrial income, under the same

measure, increased the inequality by 27.4 percent.

Second, I compare the observed level with what would have occurred if the mean and *CV* of a particular income source and its correlation with other income sources had not changed. For agricultural income, had its correlations with industrial income and other income sources, as well as the mean and *CV*, remained at their 1988 levels, the *CV* of total income would have been 1.01 by 1995 and the rise in inequality would have been 8.2 percent less, indicating that the changes of the correlation between agricultural income and other income sources had only a negligible impact on the changes of total income inequality of 1.3 percent (8.2-9.5). This is also true for the changes in the correlation between industrial income and other income sources. This indicates that changes in industrial income and agriculture income were not highly correlated. The household opportunities to obtain industrial employment did not depend on their agricultural income or the total household income, and vice versa.

To sum up, agricultural income had an equalizing impact on total income inequality, whereas industrial income disequalized income inequality. The rise in inequality, however, was attributed to both the diminishing equalizing impact of agricultural income, because of the stagnant development of agriculture after the mid 1980s, and increasing disequalizing impact of industrial income, as industrial wages became a more important household income source. As a result, it was the structural changes (share of income sources) rather than increase in inequality of individual income sources that caused the rise in inequality.

5.3 Counterfactual analysis on spatial inequality

None of previous studies have linked the changes in a particular source of income to the changes in spatial inequality, as decomposition by income components and decomposition by subgroups are two independent concepts. The method developed in this paper makes it possible to link the two and to examine how each income source contributed to the spatial inequality in rural China.

Table 5 illustrates the impact of agricultural income and industrial income on spatial inequality. Column (1) shows the observed inequality in each region.¹⁰ The rise of inequality over the period of 1988-1995 was common phenomenon throughout the country. Within-regional inequality was the major source of overall inequality in rural China, as shown in other studies (Coady and Wang, 2000; Tsui, 1993, 1995 and 1996; and Lee, 2000). However, increasing within-regional inequality was much slower than between group inequality. Within group inequality increased from .54 in 1988 to .99 in 1995, indicating a rise of 83 percent, while between group inequality increased more than 300 percent during the same period, from .05 to .18. This shows that most of the increase in overall inequality during 1988 to 1995 was accounted for the rise of spatial inequality.

Column (2) shows the CV that would have occurred if there was no agricultural income or industrial income and all other income sources remained as observed. Column (3) calcu-

¹⁰ The regions are divided as follows. East region includes Guangdong, Fujian, Jiangsu, Zhejiang, and Shandong. The middle includes Hunan, Hubei, Sichuan, Anhui, Jiangxi, Guangxi, Henan, Hebei, Shanxi, Shangxi, Liaoning, Jilin, and Heilongjiang. And Xinjiang, Xizang, Qinghai, Ningxia, Gansu, Neimenggu, Yunan and Guizhou are included in Western region.

lates the difference between Columns (1) and (2) and indicates the impacts of the particular income source on both within and between regional inequality. The equalizing impact of agricultural income was the smallest in east coastal area, as it was a smaller share of income than in central and western area. This equalizing impact decreased for both east coastal and western area from 1988 to 1995, but increased in central region. The equalizing impact of agricultural income on within-regional inequality was as large as between-regional inequality, implying that agricultural development was fairly similar among different regions. In 1988, within-group inequality would have been 1.94 in the absence of agricultural income, suggesting the equalizing impact of agricultural income on within-group inequality of 72 percent. Within this context, agricultural income had a same magnitude of impact on between group inequality of 74 percent. In 1995, the story was similar except that the equalizing impact of agricultural income decreased, its contribution decreased from 72 to 65 percent for within-group inequality and 74 to 64 percent for between group inequality. Therefore, the imbalance of agricultural development was as severe within regions as between regions. On the other hand, the impact of industrial income showed a different pattern. Industrial income increased within-regional inequality by only 4 percent in 1988 in that within-regional inequality would have been less in its absence; its impact on between-regional inequality, very remarkably, were 96 percent under the same context. In 1995 when a lot of town and village owned enterprises were established and most of them were concentrated in the east region, the impact of industrial income increased to 59 percent on within-regional inequality and as large as 128 percent on between-regional inequality.

Therefore, the most of the growing overall inequality was accounted for the rise in spatial inequality. The growing regional inequality can largely be explained by the increasingly uneven regional development of industrialization. In particular, it was the rise in inequality of employment opportunities that was the culprit. In developed regions the high level of township industry development offered many employment opportunities, so that household income from industrial sources were much higher than from farming. In contrast, in underdeveloped regions such employment opportunities were more limited, so that farmers experienced difficulty in obtaining any income from industrial work. According to Riskin, et al. (1998), the interprovincial Gini index among rural industrial workers was .14 in 1995, whereas the Gini index of the proportion of industrial workers in rural labor force was as high as .63. This also indicates that wage inequalities were very small within the industrial sector, it was the inequality of opportunity to gain industrial income that caused the rise in inequality.

6 Conclusion

This paper assesses the contribution of income sources to the rise in rural income inequality in China. The results indicate that the rise in inequality was attributed to both the stagnant development of agriculture and development of rural industrialization. In particular, agricultural income was an equalizing income source but its share of total income decreased, whereas industrial income was disequalizing and its share rose as rural industrialization developed. Most of the rise inequality was accounted for spatial inequality, the increase of

which reflected growing uneven regional development of industrialization. Nevertheless, it is important to point out that income from rural industrialization itself did not contribute to the rise in inequality in that wages from rural industry, mostly labor intensive sectors, were fairly equal. As the rural industry developed and more and more rural laborers got the chance to work in firms, the inequality of wage income itself actually declined.

7 Appendix

7.1 Decomposition by subgroups:

Define:

T : T subgroups in total

N : population

cv_t : coefficient of variation of total population

cv_i : coefficient of variation of sub_group i

$P_i \equiv N_i/N$: population share of group i

$\lambda_i \equiv \mu_i/\mu$: group i 's mean income relative to the population mean

To Prove:

$$CV_t^2 \equiv \sum_{i=1}^N P_i (cv_i^2) \lambda_i^2 + \sum_{i=1}^N P_i (\lambda_i - 1)^2$$

Prove:

$$\begin{aligned}\sigma^2 &= \frac{1}{N} \sum_{k=1}^N (x_k - \mu)^2 \\ &= \frac{1}{N} \left[\sum_{k=1}^{N_1} (x_k - \mu_1 + \mu_1 - \mu)^2 + \dots + \sum_{k=N_{T-1}+1}^N (x_k - \mu_2 + \mu_2 - \mu)^2 \right]\end{aligned}$$

The first term in the bracket can be written as:

$$\begin{aligned}& \sum_{k=1}^{N_1} (x_k - \mu_1 + \mu_1 - \mu)^2 \\ &= \sum_{k=1}^{N_1} (x_k - \mu_1)^2 + 2 \sum_{k=1}^{N_1} (x_k - \mu_1)(\mu_1 - \mu) + \sum_{k=1}^{N_1} (\mu_1 - \mu)^2 \\ &= N_1 \sigma_1^2 + 2(N_1 \mu_1 - N_1 \mu_1)(\mu_1 - \mu) + N_1 (\mu_1 - \mu)^2 \\ &= N_1 \sigma_1^2 + N_1 (\mu_1 - \mu)^2\end{aligned}$$

Therefore

$$\sigma^2 = \sum_{i=1}^T [P_i \sigma_i^2 + P_i (\mu_i - \mu)^2]$$

Deriving for cv_t ,

$$\begin{aligned}
cv_t^2 &= \frac{\sigma^2}{\mu^2} \\
&= \sum_{i=1}^T \frac{P_i \sigma_i^2 + P_i (\mu_i - \mu)^2}{\mu_i^2 / \lambda_i^2} \\
&= \sum_{i=1}^T \left[\frac{P_i \sigma_i^2}{\mu_i^2 / \lambda_i^2} + \frac{P_i (\mu_i - \mu)^2}{\mu_i^2 / \lambda_i^2} \right] \\
&= \sum_{i=1}^T \left[P_i (cv_i^2) \lambda_i^2 + P_i \lambda_i^2 \left(\frac{\mu_i - \mu}{\mu_i} \right)^2 \right] \\
&= \sum_{i=1}^T \left[P_i (cv_i^2) \lambda_i^2 + P_i \lambda_i^2 \left(1 - \frac{1}{\lambda_i} \right)^2 \right] \\
&= \sum_{i=1}^T [P_i (cv_i^2) \lambda_i^2 + P_i (\lambda_i - 1)^2]
\end{aligned}$$

References

- [1] Cancian, M. and Reed, D. (1998). Assessing the effects of wives' earnings on family income inequality. *Review of Economics and Statistics*, 80(1): 73-79.
- [2] – (1999). The Impact of Wives' Earnings on income inequality: issues and estimates. *Demography*, 36(3): 173-84.
- [3] Cheng, Y. (1996). A decomposition analysis of income inequality of Chinese rural households. *China Economic Review*, 7(2): 155-67.
- [4] Coady and Wang, (2000). Equity, efficiency, and labor-market reforms in urban China: the impact of bonus wages on the distribution of earnings. *China Economic Review*, 11(2): 213-31.
- [5] Huang, Y. (1998). *Agricultural reform in China*. Cambridge University Press.
- [6] Jenkins, S. P. (1995). Accounting for inequality trends: decomposition analyses for the UK, 1971-86. *Economica*, 62(245):29-63.
- [7] – (1995). How much inequality can we explain? A methodology and an application to the United States. *Economic Journal*, 105(429): 421-30.
- [8] Khan, A. R., Griffin, K., Riskin, C. and Zhao R. (1992). Household income and its distribution in China. *The China Quarterly*, 132: 1029-61.
- [9] – (1993). Sources of income inequality in post-reform China. *China Economic Review*, 4(1): 19-35.
- [10] Khan, A. R., Griffin, K., Riskin, C. (1998). Income and inequality in China. *China Quarterly*, 154: 221-53
- [11] Knight, J. and Song, L. (2000). Workers in China's rural industries. in Keith Griffin and Zhao Renwei (eds), *The distribution of Income in China*, London: Macmillan.

- [12] Li, S., Riskin, C. and Zhao R. (2000). *China's retreat from equality*. Asia and the Pacific.
- [13] Lerman, R. I. and Yitzhaki, S. (1985). Income inequality effects by income source: a new approach and applications to the United States. *Review of Economics and Statistics*, 67(1): 151-56.
- [14] Li, S., Zhao, R. and Zhang, P. (1997). The changes of China's income distribution during the transition. Dissertation for the international workshop in Beijing.
- [15] Rozelle, S. (1994): Rural industrialization and increasing inequality: Emerging patterns in China's reforming economy. *Journal of Comparative Economics*, 19: 362-91.
- [16] Shorrocks, A. F. (1982). Inequality decomposition by factor components. *Econometrica*, 50(1): 193-211.
- [17] – (1983). The impact of income components on the distribution of family incomes. *Quarterly Journal of Economics*, 98: 311-26.
- [18] – (1984). Inequality decomposition by population subgroups. *Econometrica*, 52: 1369-85.
- [19] Tsui, K.-Y. (1991). China's regional inequality, 1952-1985. *Journal of Comparative Economics*, 15: 1-21.
- [20] – (1993). Decomposition of China's regional inequalities. *Journal of Comparative Economics*, 17: 600-27.
- [21] – (1996). Economic reform and interprovincial inequalities in China. *Journal Development Economics*, 50: 353-68.

Table 1 Rural Household Income by Income Sources

Income Sources <i>(in 1988 yuan)</i>	Amount		Percent		Coefficient of Variation	
	1988	1995	1988	1995	1988	1995
Total income	3806	4556	100.0	100.0	0.77	1.08
Agricultural income	2439	2116	64.1	46.4	0.82	0.95
Wage and other compensation from working unit	425	1296	11.2	28.4	3.17	2.88
Non-agricultural income	384	442	10.1	9.7	2.25	3.06
Others						
Property income	375	548	9.8	12.0	1.09	1.04
Government transfer	-73	-22	-1.9	-0.5	-2.32	-15.61
Private transfer	256	175	6.7	3.8	2.76	2.72

Table 2 Income Decomposition by Component

	CV(k)	Mean (k)/ Mean (T)	Correlation with total income	Contribution	Contribution in %
	(1)	(2)	(3)	(4)	(5)
1988					
Income Component					
household agricultural production	.819	.641	.732	.384	50.1
household non-agricultural production	2.254	.101	.376	.086	11.2
wage and compensation from enterprises	3.167	.112	.523	.185	24.2
other income	1.567	.147	.485	.111	14.5
Total Family Income	.766				100.0
1995					
Income Component					
household agricultural production	.951	.464	.421	.186	17.3
household non-agricultural production	3.063	.097	.360	.107	9.9
wage and compensation from enterprises	2.881	.284	.835	.684	63.6
other income	1.297	.154	.492	.098	9.1
Total Family Income	1.075				100.0

Table 3 Contribution of Income Components to Rural Inequality

Panel a: If there were no agriculture income			
	Actual CV (I)	Counterfactual CV (II)	Contribution of Agriculture Income (II-I)/II
1988	.766	1.457	-47.4%
1995	1.075	1.822	-41.0%
Panel b: If there were no industrial income			
	Actual CV (I)	Counterfactual CV (II)	Contribution of Industrial Income (II-I)/II
1988	.766	.737	4.0%
1995	1.075	.834	29.0%

Table 4
Contribution of Changes in Income Component to the Rise in Inequality

Panel a: Agricultural Income

	Coefficient of Variation			Changes in percent
	1988	1995	Changes	
Changes in Means and Dispersion of agriculture income				
Actual CV	.766	1.075	.309	40.4%
No Changes in agriculture Income	.766	1.003	.237	30.9%
Contribution to inequality			.072	9.5%
Changes in Means and Dispersion of Agriculture Income, and its Correlation with Other Income				
Actual CV	.766	1.075	.309	40.4%
No Changes in agriculture Income	.766	1.013	.247	32.2%
Contribution to inequality			.062	8.2%
Contribution of the changes in the correlation of Agricultural income with other income sources				
			-.010	-1.3%

Panel b: Industrial Income

	Coefficient of Variation			Changes in percent
	1988	1995	Changes	
Changes in Means and Dispersion of wages				
Actual CV	.766	1.075	.309	40.4%
No Changes in industrial Income	.766	.865	.099	13.0%
Contribution to inequality			.210	27.4%
Changes in Means and Dispersion of wage, and its Correlation with Other Income				
Actual CV	.766	1.075	.309	40.4%
No Changes in industrial Income	.766	.853	.087	11.3%
Contribution to inequality			.223	29.0%
Contribution of the changes in the correlation of industrial income with other income sources				
			.012	1.6%

Table 5 Contribution of Income Components to the Regional Inequality

Panel a: If there were no agricultural income

Region	Actual		Counterfactual		Contribution of income components	
	(I)		(II)		(I)-(II)/(II)	
	1988	1995	1988	1995	1988	1995
East	.796	.964	1.256	1.313	-37%	-27%
Middle	.643	.836	1.314	1.775	-51%	-53%
West	.633	.831	1.493	1.636	-58%	-49%
Within Group	.540	.990	1.944	2.859	-72%	-65%
Between Group	.047	.167	.179	.460	-74%	-64%

Panel b: If there were no industrial income

Region	Actual		Counterfactual		Contribution of income components	
	(I)		(II)		(I)-(II)/(II)	
	1988	1995	1988	1995	1988	1995
East	.796	.964	.818	.860	-3%	12%
Middle	.643	.836	.631	.635	2%	32%
West	.633	.831	.607	.619	4%	34%
Within Group	.540	.990	.519	.622	4%	59%
Between Group	.047	.167	.024	.073	96%	128%