The Australian Journal of Agricultural and Resource Economics, 45:3, pp. 361–381

Changes in regional inequality in rural China: decomposing the Gini index by income sources†

Guang Hua Wan*

A new method is proposed to decompose inequality changes as measured by the Gini index[1] into structural effects, real inequality effects and interactive effects. Application of the method to updated Chinese data reveals that structural effects represent the driving force underlying the increasing trend in regional income inequality in rural China. Policy implications are explored. In addition to these contributions, considerable efforts are made to construct the income data used in the article. Pitfalls in measuring income inequality in rural China are discussed.

1. Introduction

Equity is a central issue in economics, sociology and politics. To study equity at the aggregate level, income distribution is usually examined. In particular, changes in income inequality over time and identification of the relevant contributing factors are of great concern to policy-makers at the macro as well as micro levels. It is a well-known fact that income distribution underlies social stability. Further, variations in income inequality are most likely to alter market demand for various commodities even in the absence of changes in income, relative prices and consumer preferences. Such variations may affect national savings and hence aggregate investment and employment.

† This article was drafted in late 1997 when the author took up a World Bank-sponsored research fellowship awarded by the China Center for Economic Research (CCER), Peking University. It was presented at an CCER seminar and at the 10th Annual Conference of the Association for Chinese Economic Studies (Australia), 13–15 July 1998, University of Sydney. I am grateful to CCER and the World Bank for financial assistance, to Professors Justin Lin, Dayuan Hu, Linxiu Zhang, and Weiming Tian, Drs Yanrui Wu and Zhangyue Zhou, two anonymous referees and the editor of this journal and participants at the seminar and conference for their useful comments on earlier drafts of the article.

* Guang Hua Wan, Senior Lecturer, Department of Agricultural Economics, University of Sydney, NSW 2006, Australia and Research Fellow, Center for Economic Research, Peking University, China.

[1] The terms Gini index, Gini ratio and Gini coefficient are equivalent and will be used interchangeably throughout the article.

Studying income inequality in China is particularly important as equity or 'balanced growth across regions' was a major goal of Chinese government policies until recent economic reforms began in the late 1970s. Research by Tsui (1991) demonstrates the success of these policies: regional inequality is found to be quite low and remained more or less stable in the pre-reform period.[2] Although one of the initial objectives of economic reforms was to break down the egalitarian system and let some regions become rich first by providing them with favourable concessions in investment, taxation and trade, in the last few years serious concerns about regional disparity have been raised inside and outside China. See Griffin and Zhao (1993), Rozelle (1994, 1996), Hu *et al.* (1997, p. 27), Wei *et al.* (1997, section 5.6), Tan (1997) and references therein. As another indication of the importance of this issue, the People's Congress listed regional inequality as one of the most pressing problems to be targeted during the ninth five-year plan (1996–2000). At present, the State Council, the State Statistical Bureau, the Chinese Academy of Social Sciences and the State Planning Commission are all engaged in income distribution studies.

Many attempts have been made to analyse income distribution and regional inequality in rural China.[3] See Hu *et al.* (1997), Yao (1997), Ravallion and Chen (1999), Yang (1994) and references therein. However, there are at least three major deficiencies associated with the existing literature. First, most of the published works only provide a snapshot (Knight and Song 1993; Tsui 1993; Hussain *et al.* 1994), a notable exception being Tsui (1991). This is a limitation in the context of rural income distribution because of the influence of random factors such as weather, policy adjustments, and other unpredictable shocks (Wei *et al.* 1997, p. 3). Findings based on data from a particular year may not be conclusive. Second, many previous studies use proxy variables such as agricultural output (Howes and Hussain 1994), regional national income (Tsui 1991), collective income (Griffin and Saith 1982) or even grain output (Lyons 1991) rather than personal income. These proxies may not adequately represent living standards in China because government transfers — substantial in China (Wei *et al.* 1997, Chapter 8) — among regions via taxation, subsidies and the like are ignored. Also, variables such as national income, GDP or GNP may suffer from the well-documented shortcomings in the social accounting system of China (Rozelle 1994, p. 386; Tsui 1993). For example,

[2] This conclusion is not without controversy, see Rozelle (1994, 1996) and Chen and Fleisher (1996).

[3] It is beyond the scope of this article to include a thorough literature review on regional inequality in rural China. Interested readers are encouraged to consult the reference list.

these measures exclude income from personal remittances, sideline production and service provision. It is useful to add that long and consistent time-series data on national income or GDP do not exist in China. While the former is reported only up to 1992, China did not use the concept of GDP until 1978. Third, little effort has been made to decompose rural inequality by factor income other than by Khan and Riskin (1998). These deficiencies and more are recognised by Griffin and Saith (1982), Tsui (1991), Knight and Song (1993, p. 211), Rozelle (1996) and Chen and Fleisher (1996). The latter appealed for the use of income data.

The main contribution of this article lies in the development of a simple framework that can be used to decompose *changes* in the Gini coefficient. This theoretical development leads to a refined or modified version of the well-known Kuznets hypothesis (Kuznets 1955). The original Kuznets hypothesis states that the relationship between economic development and income inequality follows an inverted U pattern. The proposed framework also permits a more detailed analysis on sources of changing income inequality. On the empirical side, this article will provide a time profile of regional income inequality in rural China for the period 1984–96. Such a profile will shed some light on the impact of economic transition on regional income distribution and will help facilitate testing of the Kuznets hypothesis and its modified version, though we will not carry out such tests in this study. The time profile, based on a properly constructed data set, has not appeared in the English literature. Finally, we will undertake inequality decomposition by income sources. The decomposition yields insightful information that could be useful to policy-makers in designing and implementing inequality-reducing policies. It is noted that the study by Khan and Riskin (1998) differs from the present one in that quite different data sets are employed. Furthermore, Khan and Riskin (1998) covered many aspects of income inequality in China, while this article focuses on rural inequality and more importantly on the changes in the Gini coefficient.

To some extent, our study complements that of Tsui (1991) who studied regional inequality in China for the period 1952–85, notwithstanding that he used national income, not household income, data and he took urban and rural China as a whole. In comparison with earlier publications, we use a more appropriate variable, i.e., per capita total disposable income based on rural household surveys. More importantly, we make extra efforts to take differences in regional price levels into consideration and to deflate the income data by regional cost-of-living indices. Few, if any, of the early studies have done both. In passing, it is noted that this article focuses on income distribution in rural China. Urban China is excluded due to serious data deficiencies recently discovered by Xu (1997). See also Lardy (1984). Xu found that the rural household survey data on income are of good quality.

2. Decomposing the Gini coefficient and its change

An observed income distribution can be graphically represented by plotting cumulative shares of income-receiving units against the corresponding cumulative shares of income, where units are arranged in ascending order of their income, or per capita average income if units are grouped (Bhattacharya and Mahalanobis 1967). The plot is similar to the conventional cumulative distribution function (CDF) where the horizontal axis indicates cumulative shares of persons under study, starting from the poorest person or group of persons. The vertical axis indicates the cumulative income shares received by these persons or groups. Connecting the plots results in the discrete Lorenz curve. Therefore, the Lorenz curve would be a straight line with a constant slope of 0.5 if income were evenly distributed. Otherwise, it must be concave with increasing slopes. The degree of concavity signifies the extent of income inequality.

While intuitively appealing, the graphical representation of the Lorenz curve is not convenient nor precise enough for income distribution analysis. Consequently, alternative parametric functional forms are proposed in the literature for approximating the curve. None of them, however, is found to be satisfactory (Wan 1999). As an alternative to the Lorenz curve representation (being graphic or algebraic), various summary measures of income distribution or inequality have been developed. The Gini ratio, the Theil ratio, the Atkinson ratio and the generalised entropy ratio have been widely used in empirical research, as have the coefficient of variation and income range. Each of them possesses its own advantages and disadvantages (Cowell 1975).

A main criticism of the Gini ratio was its lack of decomposibility (Cowell 1975). This, however, is invalid as Kakwani (1977) showed that it is possible to decompose the Gini by factor components. Subsequent works discovered more ways of decomposing the Gini by income sources. While one method cannot be claimed to be superior to others (Shorrocks 1982), the covariance methods developed later require, implicitly or explicitly, special assumptions on population grouping. See Fei *et al.* (1978), Pyatt *et al.* (1980), Lerman and Yitzaki (1985) and Lambert and Aronson (1993). Such assumptions, however, are not needed for deriving Kakwani's results; see equation (4). For empirical applications of the covariance method to China, see Rozelle (1994, 1996).

On the other hand, Dagnum (1990) proves that only the Gini ratio is supported by observed economic units' behaviour as it is based on non-individualistic or interpersonal utility and disutility functions. Other inequality measures generally imply decreases in social welfare following a drop in income of an economic unit independent of what might happen to

the overall income inequality. In contrast, the social welfare function underlying the Gini index is an increasing function of the mean income and decreasing function of the income inequality Gini ratio. Therefore, the Gini index takes into account both the income and the inequality effects. Apart from this distinct and desirable property, the Gini index is a dimensionless measure, takes values in the unit interval and satisfies the aversion to poverty and inequality principles. As pointed out by Dagum (1990, p. 99), the Gini index allows a much more realistic interpretation of both social welfare and social income inequality than the Theil, the generalised entropy and the Atkinson inequality ratios. Moreover, the Gini ratio is the most popular and the oldest inequality measure (Nygard and Sandstrom 1989, p. 82). Accordingly, the Gini index will be used to measure inequality in this article.

By definition, the Gini ratio to be denoted by G is equivalent to one minus twice the area under a Lorenz curve. If a satisfactory Lorenz function can be obtained, the area could be easily computed by numerical or analytical integration over the interval (0, 1). However, given the findings of Wan (1999), this approach as taken by Yang (1992) and others is not recommended. Consequently, the discrete Gini to be denoted by G_d will be used instead. Following Silber (1989), G_d can be obtained via simple matrix manipulation:

$$G_d = PQI, \tag{1}$$

where P is a row vector of the shares of income-receiving units, and I is a column vector of income shares. Both are ranked by increasing values of per capita income of the income-receiving units. Finally, Q is a square matrix with appropriate dimensions whose elements q_{ij} are equal to 0 if $i = j$, to $+1$ if $i < j$ and to -1 if $i > j$.

Suppose total income Y is composed of K sources or components, i.e., $Y = Y_1 + Y_2 + \cdots + Y_K$. When the above equation is applied to a particular source of income Y_k, the corresponding Gini ratio, denoted by G_k, indicates inequality in the k-th factor income and can be computed using the following equation:

$$G_k = P_k Q I_k, \tag{2}$$

where P_k is a row vector of shares of persons who receive Y_k, and I_k is a column vector of income shares based on the k-th income data. Both are ranked by increasing values of per capita Y_k.

A related concept is the concentration index C_k, which can be defined as:

$$C_k = P_k^* Q I_k^*, \tag{3}$$

where P_k^* is a row vector of shares of persons who receive Y_k, and I_k^* is a

column vector of income shares based on the k-th income data. Both are ranked by increasing values of per capita Y. Clearly, the concentration index C_k is similar to G_k in definition except that the relevant income and population shares are ranked differently. When computing C_k, they are ranked by increasing values of per capita total income rather than per capita component income. The latter is used when computing G_k. Thus, $C_k = G_k$ if and only if the ordering of the total income and that of the factor income are identical.

While any Gini ratio G_k or G_d only takes value in the interval (0, 1), C_k lies between $-G_k$ and $+G_k$ (see Kakwani 1977, p. 721). Also, G_d only gives the lower bound of the true Gini index G (Gastwirth 1972). This is partly why Kakwani and Podder (1973) and Basmann *et al*. (1990) proposed alternative functional forms of the Lorenz curve. Their purpose was to seek better approximations to G than G_d. However, their propositions do not lead to improvement in the Gini ratio, at least in the case of income distribution data from rural China (Wan 1999).

Let μ and μ_k denote the mean of the total income and the mean of the k-th factor income, decomposing the Gini index by income sources can be implemented using the equation below (Kakwani 1977, p. 724):

$$G_d = \Sigma_k(\mu_k/\mu)C_k. \tag{4}$$

Based on equation (4), the Gini index can be interpreted as a weighted average of K concentration indices, with weights being given by the aggregate income shares μ_k/μ.[4]

It is important to point out that in reality, changes in income shares can be independent of changes in the distributions of individual income components. To elaborate, suppose there was a rise in the m-th factor income by x per cent for all the units under consideration or for the whole population. Clearly, the rise will not affect the distribution of any factor income including the m-th income source. However, μ_m is now greater than before and all other shares will alter. If the m-th income is more equally distributed relative to other components, the inequality of total income following such a rise should drop, and vice versa. This change in total income inequality is of course completely due to variations in income shares. It is interesting to discover that the Gini index, expressed in the form of equation (4), accurately portrays this aspect of reality. As long as income shares (μ_k/μ) alter, the Gini index G_d will change even if the distributions of factor income as measured by G_k or C_k remain unchanged.

[4] Note $\mu_k/\mu = \mu_k N/(\mu N)$ = the k-th income share at the aggregate or national level, where N = total number of units or national population.

Let $S_k = \mu_k/\mu$ denote the k-th income share, then G_d is equivalent to:

$$G_d = \Sigma_k S_k C_k. \tag{5}$$

Thus, $(S_k C_k)/G_d \times 100$ can be utilised to represent the percentage contribution of income source k to the overall income inequality. Adams (1994) proposes to define C_k/G_d as the relative concentration coefficient. A factor income is said to be inequality-increasing or inequality-decreasing if its relative concentration coefficient is greater or smaller than unity.

It is often necessary to analyse *changes* in income inequality over time or differences in inequality across countries. The percentage contributions obtained using equation (5) are useful, but a dominant contributor in terms of the percentage contribution may not be a major contributor to the changes or the differences in the overall income inequality. An analogous case is the study of sources of economic *growth* versus description of the composition of an economy. For example, the rural sector may represent a major component of a developing economy, but economic growth could come entirely from nonrural sectors. Similarly, wage income may constitute a large part of the overall income inequality in China, but it might be a small contributor to inequality changes. It is worth noting that analysing inequality changes is perhaps more important than merely focusing on the composition of Gini values.

Using t and $t+1$ to index time or country, we have:

$$G_{d\,t+1} - G_{d\,t} = \Sigma_k (S_{k\,t+1} C_{k\,t+1} - S_{k\,t} C_{k\,t}). \tag{6}$$

Defining $\Delta G = G_{d\,t+1} - G_{d\,t}$, $\Delta S_k = S_{k\,t+1} - S_{k\,t}$ and $\Delta C_k = C_{k\,t+1} - C_{k\,t}$, a change in the Gini coefficient can be decomposed as:

$$\begin{aligned} \Delta G &= \Sigma_k [(\Delta S_k + S_{k\,t})(\Delta C_k + C_{k\,t}) - S_{k\,t} C_{k\,t}] \\ &= \Sigma_k \Delta S_k C_{k\,t} + \Sigma_k \Delta C_k S_{k\,t} + \Sigma_k \Delta C_k \Delta S_k. \end{aligned} \tag{7}$$

Equation (7) indicates that changes in inequality from one period or country to another can be attributed to three sources: (i) changes or differences in income shares, i.e., $\Sigma_k \Delta S_k C_{k\,t}$; (ii) changes or differences in concentration indices, i.e., $\Sigma_k \Delta C_k S_{k\,t}$; and (iii) changes or differences in interactive terms, i.e., $\Sigma_k \Delta C_k \Delta S_k$. Clearly, each of these three sources can be further broken down into K components. It should be stressed that, just like the Gini index itself (G_d), changes in the Gini ratio (ΔG) may be decomposed in other ways. For example, ΔG can be decomposed into factor components according to equation (6), each to be represented by $S_{k\,t+1} C_{k\,t+1} - S_{k\,t} C_{k\,t}$. This decomposition, however, is neither interesting nor useful.

As previously mentioned, changes in income shares are not necessarily related to spatial distributions of individual income sources. In theory, it is possible that all changes in the overall income inequality are attributable to changes in income shares which may reflect some fundamental shifts in an

economy. In this case, $\Delta G = \Sigma_k \Delta S_k C_k$. For instance, an increase in the share of the j-th factor income that accrues mostly to the poor will not necessarily cause a decline in the overall inequality; though C_j is now smaller and thus source (ii) is negative, source (i) or share changes weighted by C_k may be positive and large enough to offset source (ii). This is exactly what happened in rural China, as is discussed later. Of course, these two sources may reinforce each other, in which case the interactive terms will be positive. Otherwise, they will be negative. Given the foregone discussion, it is appropriate to term source (i) structural effects, and source (ii) real inequality effects or pure concentration effects.

The significance of equation (7), developed in this article, is self-evident: policy options to deal with pure concentration effects could be quite different from those to deal with changing income shares or structural effects. In most developing countries where farmers derive income from farming, off-farm activities and other sources, increases in inequality due to changing income shares (decline in the farm income share is commonly observed over time in less developed countries (LDCs)) are inevitable as urbanisation, rural industrialisation and agriculture modernisation take place. In fact, these increases are closely associated with, sometimes even indispensable to, the pace of transformation of the national economy. In other words, LDCs and transitional economies almost certainly have to experience this kind of increase sooner or later. Under this circumstance, any government attempt to curb such increases in inequality seems inappropriate. By the same token, the structural effects will decrease as an economy becomes mature or approaches its completion of structural adjustment. These arguments are applicable to any economy at any development stage as long as it experiences significant structural changes.

The preceding reasoning depicts an inverted U relationship between development and inequality. As structural change is a necessary and sufficient condition for development, inequality will have to rise as an economy enters the phase of radical transformation. Consequently, our decomposition of the Gini index leads to a refined or modified version of the Kuznets hypothesis as follows: it is the structural effects of the changing (income, GDP or GNP) inequality that really underlie the systematic pattern of an inverted U. The structural effects of income inequality will increase initially, reach a peak and then decrease while the other effects may not exhibit any trends as an economy experiences fast growth or significant reforms. These arguments are supported by empirical evidence to be presented later. At some point in time when the economy reaches an equilibrium state, one would expect changes in income shares to be more or less random in the long run. When that happens, the structural effects approach zero and the overall inequality tends to be stable.

In light of the above discussion, some of the previous works on testing the Kuznets hypothesis need re-examination and the existing mixed findings may well be reconcilable. In particular, it is not difficult to conceive that when the structural effects are dominant, the original Kuznets hypothesis is likely to be accepted. Otherwise, it may be refuted even if the structural effects imply an inverted U pattern.

3. Data issues and changing patterns of income inequality in rural China

Original data on rural household disposable income are compiled from various issues of *Rural Household Survey Statistics*, a publication of the State Statistical Bureau of China (SSB various years). They are aggregated to the provincial or regional level and are available on a per capita basis. Although the SSB survey data suffer from a number of shortcomings such as undervaluation of self-consumed products (Khan and Riskin 1998; Ravallion and Chen 1999), they represent the only data source covering a reasonably long period of time. Moreover, according to Yang (1999), the data shortcomings are likely to be time-invariant. Thus, they may not introduce significant biases into studies on changes in income inequality over time.

The SSB publication goes back at least to the 1970s. However, data prior to 1981 are not accessible. This omission is of little consequence as we focus on the post-reform period. Total disposable income consists of three factor components: wage income, family operation income that contains a portion imputed from self-consumed products, and residual income which includes remittances from relatives and family migrants. Since family operation mainly consists of farming, the second component will be referred to as farming income hereafter. While total disposable income data are available from 1982 onwards, consistent data on income components by regions cannot be found for 1982 and 1983. As a consequence, only the 1984–96 observations will be analysed. The needed rural population data are sourced from *China Agricultural Yearbook* (SSB 1984–96a) and *China Rural Yearbook* (SSB 1984–96b). Rural population instead of agricultural population is selected as the rural survey covers non-urban households, which corresponds to the rural rather than agricultural population. It should be pointed out that the computation results are robust to the choice of population data.

Given the area of China, the inadequate transportation and communication infrastructure in rural areas and the relatively underdeveloped rural markets, general price levels can vary substantially across regions, as can inflation rates. It is thus necessary to deflate the income data by regional inflation rates as well as by regional standard-of-living indices. There exist many different price indices, of which the Cost-of-Living Index or Consumer Price Index (CPI) for Farmers seems most appropriate for present purposes.

This index is published in various issues of *China Statistical Yearbook* and the name has been changed to Rural CPI as from 1994. The published regional CPIs take the preceding year as the base and they are all converted so to make 1981 the common base. Regarding different price levels across regions, Howes and Lanjouw (1991) constructed regional standard-of-living indices and the deflated income data are further deflated by them so that comparable real income data across regions are obtained. Although the regional standard-of-living indices were constructed for urban areas, they are used here in the absence of better alternatives on the ground that local price levels in rural and urban areas are expected to be highly correlated.

Since the CPIs for Xizhang (Tibet) do not exist and Hainan was part of Guangdong province prior to 1988, Xizhang is excluded from the following analysis while Hainan is merged with Guangdong throughout. The exclusion of Xizhang may slightly distort the results but is considered to be acceptable given the small population of the region. Further, the income level of Xizhang rural residents is closer to the national average than many of the other regions (see Wei *et al.* 1997, p. 92). In short, for each of the remaining 28 regions in China, 13 yearly observations on per capita disposable income and its components, an index of price level, rural population and rural CPI are collected and will be used in this study.

Table 1 reports national average disposable income (μ) of rural residents, factor income (μ_k) and their shares at the 1981 prices. It is obvious that, taking rural China as a whole, farming still represents the dominant source

Table 1 Per capita disposable income in rural China at 1981 prices

	Level of disposable income (RMB)				Composition (%)		
Year	Total	Wage	Farming	Residual	Wage	Farming	Residual
1984	335.60	57.58	250.81	27.21	17.16	74.73	8.11
1985	337.97	60.41	255.69	21.87	17.87	75.66	6.47
1986	350.75	66.50	266.55	17.70	18.96	75.99	5.05
1987	357.12	72.91	267.72	16.49	20.42	74.97	4.62
1988	352.45	76.29	260.47	15.68	21.65	73.90	4.45
1989	322.31	73.21	234.59	14.51	22.71	72.78	4.50
1990	356.16	72.13	269.02	15.01	20.25	75.53	4.21
1991	357.37	75.83	264.79	16.75	21.22	74.09	4.69
1992	374.63	88.77	267.36	18.51	23.69	71.37	4.94
1993	383.86	82.04	281.56	20.26	21.37	73.35	5.28
1994	416.16	89.86	301.09	25.21	21.59	72.35	6.06
1995	461.89	103.61	330.83	27.46	22.43	71.62	5.95
1996	524.15	123.63	370.68	29.85	23.59	70.72	5.69
Growth Rate (%)	3.79	6.57	3.31	0.77			

of income, although non-farming income may constitute more than 60 per cent of total disposable income in the developed south of China. From 1984 to 1996, real total disposable income of rural residents grew just under 4 per cent annually. This rate is almost identical to the figure presented in Zhao and Li (1997) who used an unspecified deflator and aggregated national rather than regional data in finding the real income. The growth was mainly derived from the wage component which rose at nearly 6 per cent per annum. Increases in farming income averaged 3.3 per cent with almost no change in the residual income. The differing growth rates signify the rising importance of rural industries and off-farm employment opportunities in China. As a consequence, farming income dropped from almost 75 per cent of the total in 1984 to just over 70 per cent in 1996 while the wage component gained more than 6 percentage points. It is undoubted that such a trend will continue and is essentially driven by industrialisation and urbanisation in rural China. It is useful to recall that structural changes will lead to a worsening of income distribution in the early stage of economic development even if the concentration or Gini indices of factor income remain constant or decline.

The discrete Gini values are tabulated in table 2, together with concentration indices and relative concentration coefficients. The Gini values are smaller than those reported by researchers in China (e.g., Zhang 1992). After accounting for the pitfalls (see section 4), the results are comparable. Clearly, there is an increasing trend in the Gini ratio of total disposable income in rural China. This trend is also present when undeflated data are

Table 2 Calculated Gini ratio and related inequality measures

		Concentration index			Relative concentration coefficient		
Year	Gini Ratio	Wage	Farming	Residual	Wage	Farming	Residual
1984	0.106	0.249	0.082	0.034	2.338	0.767	0.319
1985	0.091	0.250	0.054	0.081	2.750	0.596	0.893
1986	0.107	0.241	0.080	0.017	2.246	0.745	0.156
1987	0.109	0.269	0.070	0.036	2.460	0.644	0.332
1988	0.113	0.288	0.065	0.076	2.542	0.569	0.671
1989	0.118	0.284	0.068	0.080	2.417	0.578	0.681
1990	0.108	0.271	0.067	0.050	2.513	0.625	0.461
1991	0.122	0.295	0.073	0.101	2.425	0.603	0.832
1992	0.125	0.290	0.071	0.118	2.314	0.568	0.941
1993	0.139	0.324	0.087	0.116	2.326	0.626	0.831
1994	0.148	0.327	0.101	0.078	2.205	0.680	0.527
1995	0.162	0.330	0.114	0.112	2.035	0.701	0.690
1996	0.160	0.329	0.108	0.110	2.056	0.673	0.689

used. Ignoring differences in regional population or regional price levels does not alter the trend either.[5] In contrast, spatial inequality of factor income as indicated by the relevant concentration indices does not exhibit any clear time trend except that for the wage component. Examination of the relative concentration coefficients reveals that only the wage income is inequality-increasing, in the sense of Adams (1994), as its relative concentration coefficients are greater than unity over the entire period under study. Conversely, both the farming income and residual income are inequality-decreasing.

Against the fairly consistent increasing trend in the Gini ratio, there occurred two noticeable reductions in the regional inequality, from 1984 to 1985 and from 1989 to 1990, as well as a negligible drop from 1995 to 1996. These reductions coincide with the tightening of monetary policy in China. Because non-farming sectors are much more capital-intensive than the farming sector, credit squeezes and increases in interest rates hurt the non-farming sectors more than farming. As a consequence, the farming sector became more important in income generation. Recalling that farming income is inequality-decreasing in China, its increased importance naturally led to reductions in overall inequality. Also, capital is more important in developed areas than in poor regions, as a factor of farming or non-farming production. Thus, income gaps between regions are expected to narrow following the tightening of monetary policy. Similar arguments can be found in Wei *et al.* (1997, Chapter 12) and Yuan (1996). It seems that the year 1990 deserves special attention. In 1990, an austerity program was implemented which caused drops in the overall inequality, in wage dispersion as well as in the wage income share from 1989 to 1990.

The percentage composition of the reported Gini ratios can be found in table 3. Table 3 shows that although wage income only accounts for some 17–24 per cent of the total disposable income, its contribution to the overall income inequality ranges from 40 per cent to 55 per cent. This is consistent with the findings of Rozelle (1994) and Wei *et al.* (1997). However, contrary to some speculations, the contribution percentage of the wage component does not display any upward trend. Further, the trend existing in the concentration index of the wage income does not always match the trend in its contribution percentage. Referring to equation (7), it can be concluded that variations in income shares must have played a role. It is thus important to undertake decomposition of the *changes* in the Gini ratio, to which we now turn.

[5] The numerical results are available from the author upon request.

Table 3 Percentage contribution to the Gini ratio by factor income

		Contribution by		
Year	Total	Wage	Farming	Residual
1984	100	40.12	57.30	2.58
1985	100	49.14	45.07	5.78
1986	100	42.58	56.63	0.79
1987	100	50.22	48.24	1.53
1988	100	55.02	41.99	2.99
1989	100	54.89	42.05	3.06
1990	100	50.90	47.16	1.94
1991	100	51.47	44.63	3.90
1992	100	54.83	40.52	4.65
1993	100	49.72	45.89	4.39
1994	100	47.61	49.20	3.19
1995	100	45.66	50.24	4.11
1996	100	48.49	47.58	3.92

In table 4, we present the structural effects due to variations in income shares, the real inequality effects due to changes in income concentration indices and the interactive effects. Focusing on the columns under the heading 'Total', the structural effects are mostly positive or disequalising except from 1989 to 1990 and from 1992 to 1993. On the other hand, the real inequality effects were equalising five out of twelve times although the magnitudes of the equalising effects are small relative to those of their disequalising effects. The interactive effects were disequalising earlier and tended to be equalising later. Note also that the interactive terms are all small in comparison with the other two terms. More research is needed to explain these patterns.

Further examination of table 4 reveals that the total structural effects are dominated by changes in the wage income share in terms of both magnitudes and signs. Conversely, the real inequality effects are dominated by changes in the farming income concentration in terms of magnitudes and signs. As speculated earlier, there exist no trends in the real inequality effects or their components even if positive values and negative values are examined separately. These provide some justification for our refinement to the original Kuznets hypothesis. In other words, the empirical evidence presented here points to the acceptance of the *refined* but not necessarily the *original* Kuznets hypothesis.

Three major findings can be summarised from table 4: (a) only the structural effects are strongly and positively correlated with the overall income inequality trend. In particular, the only negative structural effects that occurred from 1989 to 1990 and from 1992 to 1993 correspond well to

Table 4 Sources of changes in regional income inequality: rural China

Year	Change in Gini ratio	Structural effects				Real inequality effects				Interactive effects			
		Wage	Farming	Residual	Total	Wage	Farming	Residual	Total	Wage	Farming	Residual	Total
1984–85	−0.0153	0.0018	0.0008	−0.0006	0.0020	0.0003	−0.0204	0.0039	−0.0163	0.0000	−0.0003	−0.0008	−0.0010
1985–86	0.0160	0.0027	0.0002	−0.0012	0.0017	−0.0017	0.0194	−0.0042	0.0134	−0.0001	0.0001	0.0009	0.0009
1986–87	0.0022	0.0035	−0.0008	−0.0001	0.0026	0.0054	−0.0072	0.0010	−0.0008	0.0004	0.0001	−0.0001	0.0004
1987–88	0.0041	0.0033	−0.0007	−0.0001	0.0025	0.0040	−0.0044	0.0018	0.0014	0.0002	0.0001	−0.0001	0.0002
1988–89	0.0041	0.0031	−0.0007	0.0000	0.0024	−0.0009	0.0026	0.0002	0.0018	0.0000	0.0000	0.0000	−0.0001
1989–90	−0.0097	−0.0070	0.0019	−0.0002	−0.0054	−0.0029	−0.0004	−0.0014	−0.0047	0.0003	0.0000	0.0001	0.0004
1990–91	0.0139	0.0026	−0.0010	0.0002	0.0019	0.0049	0.0045	0.0022	0.0116	0.0002	−0.0001	0.0002	0.0004
1991–92	0.0037	0.0073	−0.0020	0.0003	0.0056	−0.0011	−0.0016	0.0008	−0.0018	−0.0001	0.0001	0.0000	0.0000
1992–93	0.0137	−0.0067	0.0014	0.0004	−0.0049	0.0079	0.0113	−0.0001	0.0191	−0.0008	0.0003	0.0000	−0.0005
1993–94	0.0093	0.0007	−0.0009	0.0009	0.0007	0.0008	0.0102	−0.0020	0.0089	0.0000	−0.0001	−0.0003	−0.0004
1994–95	0.0136	0.0027	−0.0007	−0.0001	0.0019	0.0006	0.0092	0.0020	0.0118	0.0000	−0.0001	0.0000	−0.0001
1995–96	−0.0018	0.0038	−0.0010	−0.0003	0.0025	−0.0001	−0.0042	−0.0001	−0.0043	0.0000	0.0001	0.0000	0.0001

the downturns in China's economic reforms. This is not surprising given our refinement to the Kuznets hypothesis; (b) if the Kuznets hypothesis is applicable to rural China, it is obvious that the underlying increasing trend in regional inequality is driven by the structural effects as we propose in this article; and furthermore, (c) it is the increases or decreases in the wage income share rather than its concentration which ultimately determine the variations in the overall income inequality across regions in rural China.

Based on equation (7), it can be seen that the signs of the individual structural effects are the same as those of changes in the relevant income shares. Therefore, the structural effects must contain negative as well as positive values in each year unless there is no change in income shares at all. When income shares remain unchanged, the total structural effects would be zero. The same, however, cannot be said of the real inequality effects. Although the signs of the individual concentration effects must be identical to the signs of the changes in the concentration indices, it is possible for all the indices to go up (e.g., from 1992 to 1993 to 1994 to 1995) or down (e.g., from 1991 to 1992 and from 1995 to 1996) at the same time.

An interesting observation is that wage income concentration actually decreased five times over the sample period. In other words, average wage income levels across regions in rural China became more equal in almost half the years studied. In light of this finding, earlier conclusions regarding the role of town and village enterprises (TVEs) in enhancing regional inequality are not precise enough, to say the least. Our analysis suggests that the development of rural industries leads to higher regional inequality due to their structural effects, not due to the imbalanced growth of TVEs. This finding has important policy implications: tackling the spatial imbalance in TVEs growth may be an ineffective or even mistaken strategy in solving the increasing inequality problem in China. As long as rural industries grow faster than other sectors and the structural effects are large enough, overall income inequality will rise, regardless of what might happen to wage income disparities across regions. Of course, more balanced growth of TVEs across regions will, to a certain extent, help offset the structural effects as it did from 1985 to 1986, from 1988 to 1989 to 1990, from 1991 to 1992 and from 1995 to 1996.

To find the full contribution to the changes in the overall inequality made by factor income, we can add the structural, real inequality and interactive effects together for every income component. The results are shown in table 5. A positive value implies that the factor is inequality-enhancing and vice versa. For wage income, the full contribution is positive all the time except for 1989–90. Likewise, farming income and residual income are inequality-enhancing most of the time. While the structural effects dominate the wage and residual income contributions, the real inequality effects

Table 5 Contribution to changes in regional income inequality, by factor income

	Wage income				Farming income				Residual income			
Year	Structural effect	Real inequality effect	Interactive effect	Full contribution	Structural effect	Real inequality effect	Interactive effect	Full contribution	Structural effect	Real inequality effect	Interactive effect	Full contribution
1984–85	0.0018	0.0003	0.0000	0.0021	0.0008	−0.0204	−0.0003	−0.0199	−0.0006	0.0039	−0.0008	0.0025
1985–86	0.0027	−0.0017	−0.0001	0.0009	0.0002	0.0194	0.0001	0.0197	−0.0012	−0.0042	0.0009	−0.0044
1986–87	0.0035	0.0054	0.0004	0.0093	−0.0008	−0.0072	0.0001	−0.0079	−0.0001	0.0010	−0.0001	0.0008
1987–88	0.0033	0.0040	0.0002	0.0075	−0.0007	−0.0044	0.0001	−0.0051	−0.0001	0.0018	−0.0001	0.0017
1988–89	0.0031	−0.0009	0.0000	0.0021	−0.0007	0.0026	0.0000	0.0018	0.0000	0.0002	0.0000	0.0002
1989–90	−0.0070	−0.0029	0.0003	−0.0096	0.0019	−0.0004	0.0000	0.0015	−0.0002	−0.0014	0.0001	−0.0015
1990–91	0.0026	0.0049	0.0002	0.0078	−0.0010	0.0045	−0.0001	0.0035	0.0002	0.0022	0.0002	0.0026
1991–92	0.0073	−0.0011	−0.0001	0.0061	−0.0020	−0.0016	0.0001	−0.0035	0.0003	0.0008	0.0000	0.0011
1992–93	−0.0067	0.0079	−0.0008	0.0004	0.0014	0.0113	0.0003	0.0130	0.0004	−0.0001	0.0000	0.0003
1993–94	0.0007	0.0008	0.0000	0.0015	−0.0009	0.0102	−0.0001	0.0092	0.0009	−0.0020	−0.0003	−0.0014
1994–95	0.0027	0.0006	0.0000	0.0033	−0.0007	0.0092	−0.0001	0.0084	−0.0001	0.0020	0.0000	0.0019
1995–96	0.0038	−0.0001	0.0000	0.0037	−0.0010	−0.0042	0.0001	−0.0051	−0.0003	−0.0001	0.0000	−0.0004

dominate the farming income contribution. We also note that the positive values under wage income always more than offset the negative values under farming income. While it has been said that no government intervention is justified for tackling the structural effects of the increasing regional inequality in China, efforts can be made to promote farming in the less developed regions. This strategy will slow down the declines in the farming income share, thus helping to mitigate the inevitable structural effects. It may also cause the negative values under 'Farming income' in table 5 to rise in absolute terms so that the overall regional inequality can be further reduced. In passing, it is noted that the full contribution of wage income went against the trend in the overall Gini ratio three times.

4. Pitfalls in measuring income inequality in rural China

As pointed out by Hu (1997), there exist some pitfalls in previous studies on income inequality in rural China. They are partly attributable to the mixed, sometimes even conflicting, findings reported in the literature.[6] See Rozelle (1994). The purpose of this section is to briefly assess distortions caused by three main pitfalls: (a) use of nominal values; (b) ignoring regional difference in price levels and inflation rates; and (c) ignoring regional difference in total population. The assessment will be accomplished by computing Gini ratios using our data with and without those pitfalls.

Comparing the results obtained with the pitfalls[7] against those presented earlier in the article, it is clear that ignoring differences in the regional price levels leads to upward biases in the Gini ratio for all the years. The biases can be as large as 0.1 in absolute terms, increasing the Gini ratio more than 60 per cent. Ignoring differences in the regional inflation rates leads to further upward bias except in 1984 and 1996. The combined over-estimations add some 0.15 or more to the Gini ratio, which would result in a Gini value twice the properly estimated counterpart.

Identification of these biases is important due to the fact that prominent economists in China seem to take a Gini value of 0.2 as the critical point (State Council Research Office 1997). A value smaller than 0.2 signals reasonable equity and a larger value is used in appeals for policy action. Meanwhile, the practice of ignoring differences in inflation and regional price levels results in a Gini value greater than 0.2. Recall that the properly

[6] As the focus of this article is on decomposing *changes* in (not the level of) income inequality as measured by the Gini index, our major findings are not directly comparable with those from earlier studies.

[7] These results were obtained by the author, but are not reported for reasons of brevity.

estimated Gini never exceeded 0.16. The double over-estimation mostly carries over to the inequality components associated with farming and residual income. This implies that more developed areas not only faced higher prices, but also experienced faster increases in price levels over the years.

5. Conclusion

In this article, we developed a framework within which sources of *changing* income inequality as measured by the Gini ratio can be identified and quantified. It is shown that a change in the Gini ratio is composed of structural effects, real inequality effects and interactive effects. This development leads to a refinement of the well-known Kuznets hypothesis: it is the structural effects that may underlie the popular inverted U pattern of income distribution. As transitional or developing economies are usually characterised by structural changes, increases in income inequality are inevitable initially but will disappear later. In some sense, our derivation provides an analytical foundation to substantiate the Kuznets hypothesis.

When the framework is applied to properly constructed income data from rural China, several major conclusions can be drawn. First, regional income inequality in rural China has been increasing since the onset of the economic reforms. While the inequality level is still low by international standards, the rising trend may justify some concerns, particularly in the context of social stability and unity with minorities who are largely scattered in less developed regions. Second, although the contribution of wage income amounts to almost half of the Gini values, the increases or changes in regional inequality are caused by both wage and non-wage income sources. There is no strong evidence to suggest that the former is dominating. On the contrary, regional concentration in wage income did help reduce Gini values in some years. Whether this can be attributed to policy initiatives requires further analysis. Thus, contrary to earlier conclusions, it is incorrect to single out rural industry development as being the major contributor to the growing regional inequality. Third, structural change is the most significant force driving the increases in the Gini ratio. As long as China embarks on structural adjustment and economic transformation, further increases in regional inequality are inevitable. Decision-makers must be aware of this fact and perceive this as a normal phenomenon. Fourth, the dominant role played by wage income in raising regional inequality is not sustainable in the long run as its contribution is mainly coming from income share increases. These increases are a direct consequence of the development process of urbanisation and industrialisation.

Thus, there appears to be no basis for government intervention because

rural industrialisation is inevitable and, most importantly, the imbalanced growth of town and village enterprises across regions is not the dominant source of increased income inequality in rural China. In other words, attempting to bridge the gap in industrial development across regions, while not without justifications, may not be effective in reducing regional inequality. Moreover, the current level of the properly-computed Gini ratio in rural China is still within the low range. These findings can be useful to those who are over-concerned about rising regional inequality in rural China.

In recent years, the base of agriculture in China has been shifting from the relatively rich south and east to the north and the west. Thus, farming income is becoming more equally distributed across regions in rural China. Note that when there is a decline in the farming concentration index there is a decline in the total real inequality effect. Therefore, assistance to non-industrial activities in less developed areas can be quite effective in tackling increasing regional inequality in rural China. One way of doing this, as suggested by Rozelle and Huang (2000), is for the Chinese government to provide price support to cropping. The most recent campaign of 'developing the west', initiated by the central government in early 2000, may have to emphasise support to the farming sector if the large-scale campaign is going to lead to declines in regional income inequality in China.

It must be pointed out that this article suffers from a number of limitations. First, the short time span under study prevents a proper statistical test of the Kuznets hypothesis. Second, distortions to the results arising from the unknown method of imputation of self-consumed products remain to be addressed. Finally, decomposition of changes in inequality as measured by other indices is yet to be attempted.

References

Adams, R. 1994, 'Non-farm income and inequality in rural Pakistan: a decomposition analysis', *Journal of Development Studies*, vol. 31, pp. 110–33.

Basmann, R.L, Hayes, K.H., Slottje, D.J. and Johnson, J.D. 1990, 'A general functional form for approximating the Lorenz curve', *Journal of Econometrics*, vol. 43, pp. 77–90.

Bhattacharya, N. and Mahalanobis, B. 1967, 'Regional disparity in household consumption in India', *Journal of the American Statistical Association*, vol. 62, pp. 143–62.

Chen, J. and Fleisher, B.M. 1996, 'Regional income inequality and economic growth in China', *Journal of Comparative Economics*, vol. 22, pp. 141–64.

Cowell, F.A. 1975, *Measuring Inequality*, LSE Handbooks in Economics Series, Prentice-Hall, London.

Dagum, C. 1990, 'On the relationship between income inequality measures and social welfare functions', *Journal of Econometrics*, vol. 43, pp. 91–102.

Fei, J., Ranis, G. and Kuo, S. 1978, 'Growth and the family distribution of income by factor components', *Quarterly Journal of Economics*, vol. 92, pp. 17–53.

Gastwirth, J.L. 1972, 'The estimation of the Lorenz curve and Gini index', *Review of Economics and Statistics*, vol. 54, pp. 306–16.
Griffin, K. and Saith, A. 1982, 'The pattern of income inequality in rural China', *Oxford Economic Papers*, vol. 34, pp. 172–206.
Griffin, K. and Zhao, R.W. 1993, *The Distribution of Income in China*, New York: St Martin Press.
Howes, S. and Hussain, A. 1994, *Regional Growth and Inequality in Rural China*, Working Paper EF/11, London School of Economics, London.
Howes, S. and Lanjouw, P. 1991, *Regional Variations in Living Standards in Urban China*, Working Paper CP/17, London School of Economics, London.
Hu, A.G., Wang, S.G. and Kang, X.G. 1997, *Regional Disparities in China*, Liaoning People's Press, Shengyang, China (in Chinese).
Hu, D.Y. 1997, *Economic Transition and Regional Inequality*, Bulletin of the China Center for Economic Research 20, Peking University, Peking, (in Chinese).
Hussain, A., Lanjouw, P. and Stern, N. 1994, 'Income inequalities in China, evidence from household survey data', *World Development*, vol. 22, pp. 1947–57.
Kakwani, N.C. 1977, 'Applications of Lorenz curves in economic analysis', *Econometrica*, vol. 45, pp. 719–27.
Kakwani, N.C. and Podder, N. 1973, 'On the estimation of Lorenz curves from grouped observations', *International Economic Review*, vol. 14, pp. 278–92.
Khan, A. and Riskin, C. 1998, 'Income and inequality in China: composition, distribution and growth of household income, 1988 to 1995', *China Quarterly*, vol. 154, pp. 221–53.
Knight, J. and Song, L. 1993, 'The spatial contribution to income inequality in rural China', *Cambridge Journal of Economics*, vol. 17, pp. 195–213.
Kuznets, S. 1955, 'Economic growth and income inequality', *American Economic Review*, vol. 45, pp. 1–28.
Lambert, P.L. and Aronson, J.R. 1993, 'Inequality decomposition analysis and the Gini coefficient revisited', *The Economic Journal*, vol. 103, pp. 1221–7.
Lardy, N. 1984, 'Consumption and living standard in China: 1978–83', *China Quarterly*, vol. 100, pp. 849–65.
Lerman, R.I. and Yitzaki, S. 1985, 'Income inequality effects by income source: a new approach and application to the United States', *Review of Economics and Statistics*, vol. 67, pp. 151–6.
Lyons, T. 1991, 'Interprovincial disparities in China: output and consumption, 1952–1987', *Economic Development and Cultural Change*, vol. 39, pp. 471–506.
Nygard, F. and Sandstrom, A. 1989, 'Income inequality measures based on sample surveys', *Journal of Econometrics*, vol. 42, pp. 81–95.
Pyatt, G., Chen, C-N and Fei, J. 1980, 'The distribution of income by factor components', *Quarterly Journal of Economics*, vol. 94, pp. 451–73.
Ravallion, M. and Chen, S. 1999, 'When economic reform is faster than statistical reform: measuring and explaining income inequality in rural China', *Oxford Bulletin of Economics and Statistics*, vol. 61, pp. 33–56.
Rozelle, S. 1994, 'Rural industrialisation and increasing inequality: emerging patterns in China's reforming economy', *Journal of Comparative Economics*, vol. 19, 362–91.
Rozelle, S. 1996, 'Stagnation without equity: patterns of growth and inequality in China's rural economy', *The China Journal*, vol. 35, pp. 63–92.
Rozelle, S. and Huang, J.K. 2000, 'Transition, development and supply of wheat in China', *The Australian Journal of Agricultural and Resource Economics*, vol. 44, pp. 543–71.
Shorrocks, A.F. 1982, 'Inequality decomposition by factor components', *Econometrica*, vol. 50, pp. 193–211.

Silber, J. 1989, 'Factor components, population subgroups and the computation of the Gini index of inequality', *The Review of Economics and Statistics*, vol. 71, pp. 107–15.

SSB (State Statistical Bureau) 1984–96a, *China Agricultural Yearbook*, China Statistical Publication House, Beijing.

SSB (State Statistical Bureau) 1984–96b, *China Rural Yearbook*, China Statistical Publication House, Beijing.

SSB (State Statistical Bureau) various years, *Rural Household Survey Statistics*, internal publication, State Statistical Bureau, Beijing.

State Council Research Office 1997, 'Analysis of personal income inequality in China and policy recommendations', *The Economic Research Journal*, vol. 8, pp. 3–10 (in Chinese).

Tan, C.L. 1997, *A Study on Regional Economic Inequality in China*, China Economics Publishing House, Beijing (in Chinese).

Tsui, K.Y. 1991, 'China's regional inequality, 1952–1985', *Journal of Comparative Economics*, vol. 15, pp. 1–21.

Tsui, K.Y. 1993, 'Decomposition of China's regional inequality', *Journal of Comparative Economics*, vol. 17, pp. 600–27.

Wan, G.H. 1999, 'An empirical assessment on alternative functional forms of the Lorenz curve', *Applied Economics Letters*, vol. 6, pp. 597–99.

Wei, H.K., Liu, K., Zhou, M.L., Yang, D.L. and Hu, W.X. 1997, *Regional Development in China: Economic Growth, System Transition and Regional Inequality*, Business Management Publishing House, Beijing (in Chinese).

Xu, L.Z. 1997, 'Differences and changes in monetary income and spending of urban and rural residents in China', *Economic References Newspaper*, 6 December, p. 3 (in Chinese).

Yang, D.T. 1999, 'Urban-biased policies and rising income inequality in China', *American Economic Review*, vol. 89, pp. 306–10.

Yang, K.Z. 1994, 'Study on changing economic inequality across regions in China', *The Economic Research Journal*, vol. 12, pp. 20–8 (in Chinese).

Yang, W.M. 1992, 'Changes in regional income disparities: an empirical analysis', *The Economic Research Journal*, vol. 1, pp. 70–4 (in Chinese).

Yao, S. 1997, 'Industrialisation and spatial income inequality in rural China, 1986–92', *Economics of Transition*, vol. 5, pp. 97–112.

Yuan, G.M. 1996, 'Regional economic disparities and macro economic fluctuations', *The Economic Research Journal*, vol. 10, pp. 49–56 (in Chinese).

Zhang, P. 1992, 'Regional income distribution in rural China', *The Economic Research Journal*, vol. 2, pp. 62–9 (in Chinese).

Zhao, R.W. and Li, S. 1997, 'Increase in income disparities in China and its causes', *The Economic Research Journal*, vol. 9, pp. 19–28 (in Chinese).