Rural Poverty and Ethnicity in China

Carlos Gradín*

(Universidade de Vigo, EQUALITAS)

In: Thesia I. Garner , Kathleen S. Short (ed.), Measurement of Poverty, Deprivation, and Economic Mobility (Research on Economic Inequality, Volume 23) Emerald Group Publishing Limited, pp.221 – 247.

Abstract

In this paper I investigate the nature of the differential in poverty by ethnicity in rural China using data from the Chinese Household Income Project in 2002. For that, I compare observed poverty with that in a counterfactual distribution in which ethnic minorities are given a set of relevant village and household characteristics of the Han majority. In particular, I investigate the importance of the location of minorities in explaining their higher poverty levels. The ethnic poverty differential does not change after equalizing the distribution of the population by geographical region (unless we use a higher poverty line). However, it is reduced after equalizing other locational characteristics of minorities (such as them living in less developed and mountainous areas), their larger number of children, their low education, and their fewer skilled non-agriculture workers. Finally, the ethnic per capita (log)income differential is shown to be higher for higher percentiles, with an increasing role of the geographical region as the main driver of these higher differentials.

Keywords: China, poverty, rural, ethnicity, decomposition.

JEL Classification: D63, I31, I32, J15.

Address: Facultade de CC. Económicas e Empresariais, Universidade de Vigo, Campus Lagoas-Marcosende, 36310 Vigo, Galicia, Spain. E-mail: cgradin@uvigo.es; Phone: +34 986813527; Fax. +34 986812401.

(*) I acknowledge financial support from the Spanish *Ministerio de Ciencia e Innovación* (grant ECO2013-46516-C4-2-R).

1. Introduction

China is a country of extreme contrasts. Despite the extraordinary growth path that followed the economic reforms that began in the late 1970s, rural and urban areas still represent two worlds apart. Sicular *et al.* (2007) estimated that even after correcting for price differences, 26 percent of overall income inequality in the country was associated with this urban-rural gap. About a half of the gap was explained by differences in endowments, of which education turned out to be the most important. Thus, it does not come as a surprise that several studies have shown so far that most poverty in China is rural, and so were most of the recent gains in reducing poverty.[1] For example, Ravallion and Chen (2007) report a reduction in rural poverty from 76 to 12.5 percent between 1980 and 2001 (from 6 to 0.5 percent in urban areas), with the most impressive reductions during the earlier 1980s that they attribute most of it to the agrarian reform, along increasing local and provincial public spending or macroeconomic stability. Xia (2009), analyzing another period with strong reduction in rural poverty (1995-2002) maintains that despite rising inequality, poverty was reduced because household incomes grew thanks to the growing importance of market forces such as rural entrepreneurship and human capital in determining rural household income, with a decreasing relevance of political-related factors (e.g. party membership, government officials, …). The engagement in off-farm occupation was also important in raising household income level but its importance declined. Rural poverty continued to be reduced in the 2002-07 period (Luo and Sicular, 2013). In their analysis of poverty dynamics for the 1989-2009 period, Imai and You (2014) showed that farming and out-migration were more effective strategies for escaping from persistent rural poverty than engagement in local non-agricultural employment.

China is also an ethnically diverse country with multiple ethnicities cohabiting with the majoritarian Han. However, the difference in poverty among these ethnic groups has not yet been investigated in depth. There are a few exceptions. Gustafsson and Li (2003) analyzed the average ethnic minority-majority income gap and its changes over time (1988-1995) in rural China using a Blinder-Oaxaca decomposition. Bhalla and Luo (2013) provided a comparative discussion of the situation of minorities in China and India, with detailed information about poverty incidence, regional concentration, living conditions or social inclusion along a number of dimensions. Gustafsson and Ding (2009a), after documenting that poverty was higher among ethnic minorities in rural China, focused on a geographical explanation for this (in line with Gustafsson and Li, 2003). According to them this ethnic differential was mainly due to the higher concentration of minorities in the less-developed western region of the country. The argument behind this explanation is that there is no such ethnic differential in the west region. In the same line, Hannum and Wang (2012) agree with this view after estimating a model for probability of being poor conditioned on a

[1] For a more complete analysis of recent income distribution trends in China see the contributions edited by Wang (2008a,b) or Li, Sato, and Sicular (2013).

dummy indicating the minority status plus a set of covariates reflecting household- and community-level characteristics.[2]

In our view, the role that geography plays in explaining ethnic differences in rural poverty in China might be a bit more complex. On the one hand, as we will later show, if the regional distribution of minorities were similar to that of Han, the poverty differential would still be large or even higher, depending on the geographical classification used. This is so because minorities are underrepresented in the part of the country (middle or central region) where the ethnic gap is the largest and overrepresented where there is no such a gap or this is smaller. On the other hand, we need to explore the role of other poor endowments of minorities in explaining their higher poverty levels, such as their lower education, their larger number of children, or their concentration in mountainous and less developed villages.[3] These have been used in Hannum and Wang's (2012) estimations but we lack a quantification of their contribution to explain the ethnic poverty gap, of how much of the ethnic gap in poverty remains after controlling for all these factors (i.e. the conditional poverty gap that is the result of those factors having a different impact on poverty in the case of Han and minorities).

The aim of this paper is precisely to shed some new light on the nature of the ethnic poverty gap in rural China using regression-based decompositions of the gap in poverty rates and in income at different quantiles. For that, using the most common public available data used in previous studies we will measure poverty in a counterfactual distribution in which these minorities are given the relevant characteristics of Han. Among these characteristics we include those found to be most highly associated with the higher poverty of minorities such as their geographical location, lower education, less skilled occupations, or higher number of children. The aggregate decomposition based on comparing the actual and counterfactual distributions allows us identifying the global contribution of ethnic divergence in households' attributes to explain the observed poverty differential, as well as the conditional poverty gap that remains unexplained. The detailed decomposition allows us to identify the individual contribution of each of those factors.

The rest of the paper is organized as follows. The next sections subsequently describe data, methodology and results. The final section summarizes the main empirical findings.

2. Data

The data used in this paper come from the rural sample of the Chinese Household Income Project (CHIP) based on questionnaire-based interviews conducted in 2002 by the Institute of Economics, Chinese Academy of Science (CASS) (see Li,

[2] Other studies have focused on specific ethnic groups or provinces (e.g. Hui in Ningxia in Gustafsson and Ding, 2014 and Sato and Ding, 2012).

[3] Gustafsson and Ding (2009b) provided a detailed discussion of the characteristics of villages for the majority and minority groups in China.

2009). This database has been the main source of research on poverty in China during the last years and is hosted at the Inter-University Consortium for Political and Social Research (ICPSR), Institute for Social Research, University of Michigan. CHIP contains very rich information on household income, expenditure and an array of characteristics of individuals and households (including a social network questionnaire) as well as village-level data, which was obtained by interviewing village leaders. The original sample does not provide sampling weights, these were constructed in order to reproduce the provincial distribution of rural population in China according to the estimates of rural population by province from the National Bureau of Statistics (NBS).[4]

CHIP 2002 also has some limitations. The most important are that it is an old database and that it has a restricted geographical coverage: 22 out of 31 provinces (or autonomous regions). Among the excluded areas are some of the most relevant for important ethnic minorities (e.g. Tibet, Ningxia, Inner Mongolia). The small sample size also constraints the scope of the analysis, especially given the wide diversity within Chinese minorities. Despite these facts, it provides the best available sample to analyze rural poverty by ethnicity in China, an important issue that deserves international attention due not only to its ethical implications, but also for being a source of political and social unrest. Furthermore, the use of CHIP 2002 allows to read the results in the context of the previous literature that has also used this same database.[5]

For consistency with previous studies, poverty is measured using household per capita disposable income. Income is defined as cash payments plus a range of additional in-kind components (including agricultural output produced for self-consumption valued at market prices, the value of ration coupons and other direct subsidies, and the imputed value of housing). In analyzing income poverty we define the same poverty line used by Gustafsson and Ding (2009a) of 878 Yuan per person/year based on the NBS low income level, adapted to take into account the bias in average income in CHIP data.

Ethnicity is based on the official classification that distinguishes 55 ethnic minorities or nationalities along with the majority group (Han). The main minorities covered by the survey are Zhuang, Hui, Uygurs, Yi, Miao, Manchu, and an additional category is included for the other ethnicities. This classification might be controversial as many ethnic groups rather use different denominations, and some of these categories represent a wide range of heterogeneous groups.[6] Nonetheless, for most of the paper all minorities will be

[4] See *China Statistical Yearbooks Database*, 2002 at http://www.yearbookchina.com. These weights are similar to those proposed in Song, Sicular, and Yue (2013). Unfortunately, due to the lack of information, it was not possible to estimate more appropriate weights fully accounting for the probability of being surveyed. In particular, this might affect the representativeness of each minority in each province.

[5] Note that the more recent CHIP 2007 has already been released. However, the geographical scope of the rural sample is much narrower, excluding the most important areas with high concentration of minorities, making useless any analysis of inter-ethnic inequality. For that reason, it has not been used in this study.

[6] See Hannum and Wang (2012) for a more detailed discussion of this classification.

considered as one single group to overcome small sample problems. The sample consists of 37,910 complete individuals observations (9,183 households), of which 5,294 individuals (1,136 households) belong to any of the minority groups.

3. Methodology

In order to obtain a decomposition of the gap in poverty rates between Han and minorities in China, we use an extension of the well-known regression-based Blinder (1973) and Oaxaca (1973) decomposition approach based on non-linear probability models. Let us consider that the *i*th person in group g (g =0, Han; g =1, minority) is poor whenever her per capita household income y_i^g falls below poverty line z. We first estimate for each group the statistical association between the probability of being poor and household-level characteristics with a logit probability model, where the likelihood of this person being poor (P_i^g) is given by:

$$P_i^g = \Pr(y_i^g < z) = F(X_i^g \hat{\beta}^g) = \frac{\exp(X_i^g \hat{\beta}^g)}{1+\exp(X_i^g \hat{\beta}^g)} \quad . \qquad \textbf{(1)}$$

F represents the logistic probabilistic cumulative distribution, X_i^g is a vector of characteristics describing i's household, and $\hat{\beta}^g$ is the associated vector of coefficients. This regression is estimated separately for Han and ethnic minorities, thus allowing for a different impact of characteristics on poverty. Given that observations are individuals but all explanatory variables are collected at the household level, we estimated robust standard errors taking into account (perfect) correlation between observations within the same sample cluster (household), while assuming independence across clusters (see Cappellari and Jenkins, 2004). In this context, this is equivalent to running the regressions over households, with the sample weight of each household multiplied by the number of household members of the corresponding ethnicity.

The head-count ratio of poverty in group g, H^g, is equal to the average predicted probability for this group (with population N^g):

$$H^g = \overline{P^g} = \overline{F(X_i^g \hat{\beta}^g)} = \frac{1}{N^g}\sum_{i=1}^{N^g} F(X_i^g \hat{\beta}^g). \qquad \textbf{(2)}$$

Thus, using the counterfactual distribution $F(X^0\hat{\beta}^1)$ in which minorities are given the characteristics of Han while keeping their own estimated coefficients, we can rewrite the differential in poverty rates between minorities and Han as the sum of the *aggregate characteristics effect* (gap explained by shifting characteristics valued at the coefficients of the target group) and the *aggregate coefficients effect* (unexplained or conditional gap due to differences in coefficients given the characteristics of the target group):

$$H^1 - H^0 = \overline{F(X_i^1\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^0)} = \left[\overline{F(X_i^1\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^1)}\right] + \left[\overline{F(X_i^0\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^0)}\right]. \textbf{(3)}$$

The evaluation of the individual contribution of each variable to the total explained difference, the detailed decomposition, is more complicated because of the nonlinearity of F (there is not a unique procedure). We followed the linear

approximation proposed by Even and Macpherson (1990, 1993) for the characteristics effect, later extended by Yun (2004) to the coefficients effect.[7] Thus, $W_{H,k}^{\Delta X} = \frac{(\bar{x}_k^0 - \bar{x}_k^1)\hat{\beta}_k^1}{(\bar{X}^0 - \bar{X}^1)\hat{\beta}^1}\left[\overline{F(X_i^1\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^1)}\right]$ is the individual contribution of characteristic k (k=1,…, K) to the aggregate characteristics effect $W_H^{\Delta X} = \left[\overline{F(X_i^1\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^1)}\right]$, while $W_{H,k}^{\Delta\beta} = \frac{\bar{x}_k^0(\hat{\beta}_k^0 - \hat{\beta}_k^1)}{\bar{X}^0(\hat{\beta}^0 - \hat{\beta}^1)}\left[\overline{F(X_i^0\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^0)}\right]$ is its contribution to the aggregate coefficients effect $W_H^{\Delta\beta} = \left[\overline{F(X_i^0\hat{\beta}^1)} - \overline{F(X_i^0\hat{\beta}^0)}\right]$. To prevent the identification problem associated with the detailed decomposition of the coefficients effect (the results for categorical variables depend on which is the omitted category, Oaxaca and Ransom, 1999), we use the normalization proposed in Yun (2005, 2008). Reported standard errors are based on the Delta method.[8]

Once we have identified the main factors that are associated with the higher poverty of ethnic minorities, it is interesting to ask whether the same pattern of differences in wellbeing can be extended to other parts of the income distribution. For that, we used another regression-based decomposition method that allows to evaluate the impact of changes in the distribution of household attributes on different quantiles of the unconditional (marginal) distribution of household disposable log income (Firpo, Fortin and Lemieux, 2007, 2009). This method consists on applying the conventional Blinder-Oaxaca decomposition to the differential in quantiles. For that, using the same explanatory variables, we run OLS regressions in which the dependent variable is the recentered influence function (RIF) of the unconditional income quantiles.

For any τ-th quantile of the income distribution (now expressed in logs), q_τ, its recentered influence function $RIF(y; q_\tau)$ is given by adding the quantile to its influence function $IF(y; q_\tau)$[9]:

$$RIF(y; q_\tau) = q_\tau + IF(y; q_\tau) = q_\tau + [\tau - \mathbf{1}(y \leq q_\tau)]/f(q_\tau). \quad \textbf{(4)}$$

Where $\mathbf{1}()$ is an indicator function that takes value 1 if the specified condition is satisfied and 0 otherwise. Note that $E\big(RIF(y; q_\tau)\big) = q_\tau$ because $IF(y; q_\tau)$ has zero expectation. If we label $\hat{\gamma}_\tau^g$ the vector of coefficients estimated by regressing $RIF(y; q_\tau)$ on X in group g, it can be shown that:

$$q_\tau^0 - q_\tau^1 = \overline{RIF(y; q_\tau)}^0 - \overline{RIF(y; q_\tau)}^1 = \bar{X}^0\hat{\gamma}_\tau^0 - \bar{X}^1\hat{\gamma}_\tau^1 = (\bar{X}^0 - \bar{X}^1)\hat{\gamma}_\tau^1 + \bar{X}^0(\hat{\gamma}_\tau^0 - \hat{\gamma}_\tau^1). \quad \textbf{(5)}$$

---

[7] This technique has some advantages over other proposed methods in the literature. First, the weights are quite transparent and simple to compute, because this only requires estimates of the coefficients and sample means for the characteristics. Second, this procedure overrides the problem of path dependency that is common to all sequential approaches to nonlinear models, in which values of characteristics and/or coefficients of one group need to be switched with those of the other group. Third, unlike these sequential approaches, the detailed characteristics effect can be obtained without making any assumptions to match individuals of one group with the characteristics of another. Finally, the original Blinder-Oaxaca approach is shown to be a particular case of this decomposition when F is a linear function.

[8] The results were obtained using the *OAXACA* Stata module (RePEc:boc:bocode:s456936) written by B. Jann.

[9] An influence function is a statistical tool used in robustness analysis that measures the influence of each individual observation on any statistic.

Thus, we obtain the corresponding Blinder-Oaxaca aggregate explained and unexplained effects: $W^{\Delta X} = (\bar{X}^0 - \bar{X}^1)\hat{\gamma}_\tau^1$ and $W^{\Delta \beta} = \bar{X}^0(\hat{\gamma}_\tau^1 - \hat{\gamma}_\tau^0)$. Similarly to the case of poverty rates, the detailed effects are estimated using the specific characteristics and their corresponding coefficients as $W_k^{\Delta X} = (\bar{x}_k^0 - \bar{x}_k^1)\hat{\gamma}_{\tau k}^1$ and $W_k^{\Delta \beta} = \bar{x}_k^0(\hat{\gamma}_{\tau k}^1 - \hat{\gamma}_{\tau k}^0)$.[10] Repeating the procedure for different quantiles we are able to explain the ethnic gap along the entire income distribution.

4. Empirical results

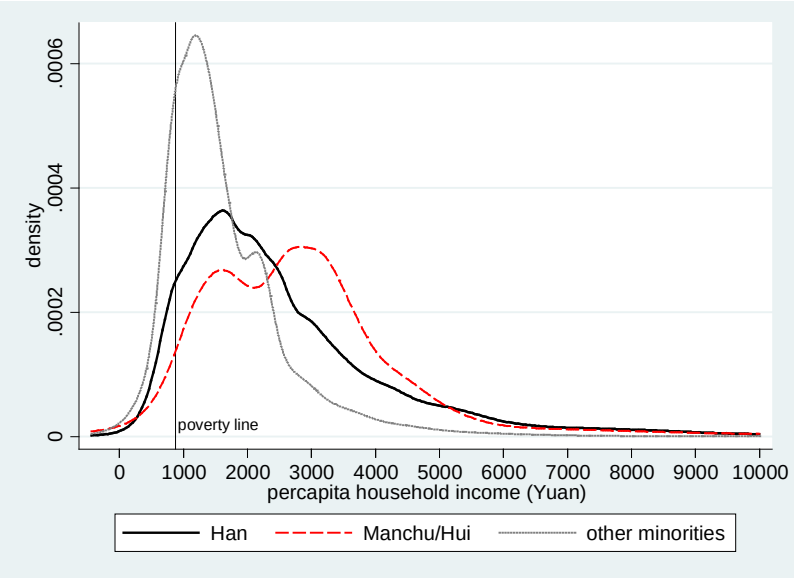
a. Poverty among ethnic groups

A large majority of the CHIP target population in rural China identified themselves as Han (89 percent, see Table 1). The remaining 11 percent of the population in the survey belongs to any of the national minorities officially recognized in the country, with the largest category being the conglomerate of other ethnicities (4 percent).[11] All minorities report lower median per capita household income than that of Han, with the only exception of Manchu, an elite minority that historically ruled the country until the end of the Qing Dynasty, whose income is 32 percent above the majoritarian group. Another peculiar group is the Hui minority, whose median income lies close to that of Han, 91 percent. The other minorities are clearly more disadvantaged, with median incomes ranging from 50 percent for Miao to 73 percent of Uygurs. The median per capita income of these disadvantaged groups considered together is 64 percent of that of Han (68 percent when Manchu and Hui are included). Figure 1 depicts the whole income distribution for the population according to their ethnicity: Han, Manchu and Hui, and disadvantaged minorities. It shows that there is a clear overrepresentation of the latter at the bottom of the income distribution.

[10] The RIF of different unconditional quantiles is obtained using the *RIFREG* Stata code (http://faculty.arts.ubc.ca/nfortin/datahead.html) from Firpo, Fortin and Lemieux (2009), and then the *OAXACA* code is used for the decomposition.

[11] The proportion of the whole Chinese population that belonged to any of the 55 national minorities was about 8.5 percent according to the 2010 Census (http://www.stats.gov.cn).

Figure 1. Income distribution densities by ethnicity in rural China, 2002



Source: Own construction using CHIP, 2002. Based on per capita household disposable income and an individual annual poverty line of 878 Yuan. Kernel density estimations using a Gaussian kernel function and adaptive optimal bandwidth.

Table 1. Ethnicity and poverty in rural China, 2002

Ethnic group	Population	Income (Yuan)		Poverty indices					
	%	Median	Ratio (Han=100)	H	Ethnic gap	HI	Ethnic gap	SPG	Ethnic gap
Han	89.0	2,147	100	8.4		2.1		0.9	
Zhuang	1.8	1,368	64	12.8	4.4	3.3	1.2	1.3	0.4
Hui	0.3	1,959	91	4.3	-4.1	2.7	0.6	2.0	1.1
Uygurs	1.3	1,561	73	14.8	6.4	4.9	2.8	3.0	2.1
Yi	1.1	1,378	64	15.6	7.3	3.7	1.6	1.5	0.6
Miao	1.2	1,063	50	33.9	25.5	8.3	6.2	2.8	1.9
Manchu	1.2	2,825	132	5.4	-3.0	1.6	-0.5	1.0	0.1
Other	4.2	1,398	65	15.5	7.2	3.8	1.7	1.5	0.6
Total	100	2,066	96	9.1	0.8	2.3	0.2	1.0	0.1
All minorities	11.0	1,457	68	15.5	7.2	4.0	1.9	1.8	0.9
Disadvantaged Minorities (Manchu/Hui excluded)	9.5	1,370	64	17.2	8.8	4.4	2.3	1.8	0.9

Source: Own construction using CHIP, 2002. Based on per capita household disposable income and an individual annual poverty line of 878 Yuan.

As a consequence of the general lower income of disadvantaged minorities, these face higher poverty rates (head-count ratio H, Table 1). While 8.4 percent of Han are poor, according to the 878 Yuan poverty line used (also depicted in Figure 1), the percentage of poor rises to 34 percent among Miao and 13-15 percent among the other disadvantaged minorities. Poverty is however lower in the case of Manchu (5.4 percent) and Hui (4.3 percent). Figure 1 allows us to infer that the ethnic gap would be even higher for the disadvantaged minorities had the poverty line be fixed at a more generous level. Also the advantage for Manchu would increase (but the differential for Hui would in fact be reversed as many of

them concentrate just above the poverty line). For example, arbitrarily increasing the poverty line by a half, the poverty rate would be 47 percent for the disadvantaged minorities, in contrast with 12 percent for Manchu, 21 percent for Han, or 25 percent for Hui.

Poverty among the disadvantaged minorities is also twice as big as it is among Han using other indices of the Foster-Greer-Thorbecke family such as the poverty gap ratio (HI) that computes the average normalized poverty gap and apart from incidence takes into account its intensity (the average income gap) among the poor, or the average squared normalized poverty gap (SPG) that incorporates also sensitivity to inequality of income gaps among the poor (see Table 1).

b. A regional compositional effect?

As previous research in other countries suggests (e.g. Gradín 2009, 2012, 2013) the existence of important ethnic differentials in wellbeing might be largely driven by ethnic minorities having poorer endowments. For historical reasons minorities are indeed highly concentrated in specific areas in China where they are native or where they settled long time ago. Given the extraordinary extension and diversity of rural China, the location of ethnic minorities is one important candidate to explain their higher poverty rates. In order to explore this possibility, Tables 2 and 3 display the distribution of the rural population by ethnicity and location following different classifications: provinces in Table 2, geographical regions in Table 3. These tables also report the average per capita income, as well as the poverty rates (with their corresponding robust standard errors, that take into account clustering within households).

The distribution of the population by province displayed in Table 2 highlights that about 96 percent of the minority rural population in the sample live in one of six selected provinces (in contrast with less than 16 percent of Han).[12] The largest minority groups reside in three of the poorest provinces in the sample: Yunnan (27 percent), Guangxi (17.5 percent), and Guizhou (16 percent). About 10-12 percent live in each of the other three provinces (Xinjiang, Liaoning and Hunan), with per capita income closer to the sample average. Ethnic minorities made up the largest share of the rural sample in Xinjiang (83 percent) and Yunnan (66 percent), near a half in Liaoning and Guizhou (47 and 45 percent respectively), with smaller shares in Guangxi (15 percent) and Hunan (11 percent). The concentration of disadvantaged minorities in these provinces was even larger, except in Liaoning (with a high concentration of Manchu). These provinces, jointly with Gansu and Shannxi, outstand for displaying the largest rural poverty rates in the sample (except Guangxi, with a poverty rate close to the average).

The use of geographical aggregates might help to better understand the specific geographical pattern followed by the distribution of minorities across rural

[12] Given that we do not know for sure if all minorities in each province had the same probability of being selected in the sample, we cannot claim this distribution is a good representation of the population.

China. Table 3 offers two alternative classifications. For the sake of comparability, the first one is the same classification used in Gustafsson and Ding (2009a), in which the country is divided into three regions: eastern, middle, and western.[13] According to this classification, a total of 75 percent of minorities in the CHIP rural sample (85 percent in the case of disadvantaged minorities) live in the western region, while Han are more evenly spread all over the country.[14] Given that the western region displays the highest poverty rate (near 15 percent), this fact lead Gustafsson and Ding (2009a) to suggest that higher poverty among minorities were the result of a geographical compositional effect.

The second panel of Table 3 shows an alternative classification of four regions based on the original classification of six regions in CHIP data, but after pooling together the three regions with the smallest share of minorities. This classification includes central, northwestern, southwestern, and northern/eastern (including the northeastern, north, and eastern regions).[15] According to this classification, near a half of the minority rural population lives in the poor northwestern, and another 28 percent in the central region (with larger shares for disadvantaged minorities).

In order to explore the hypothesis of the geographical compositional effect, in a first stage we undertake a simple shift-share analysis, which gives minorities the same geographical representation by province/region of Han. The results in the first panel of Table 3 show that had minorities been more evenly distributed across the country into eastern, western and middle regions, the ethnic poverty gap would be higher (at least 2.7 percentage points), not lower, because the poverty rate would be 18 percent among minorities (22 percent among the disadvantaged), compared with 15 (17) percent that was actually observed. While the overall poverty rate is globally higher in the western region (15 percent) compared with middle and eastern regions (8 and 5 percent respectively), it is in the middle region where the highest ethnic differential can be found due to the larger (and statistically significant) poverty incidence among minorities there (above 30 percent). Thus if ethnic minorities had a higher share of the rural population in this region, their poverty rate would be higher.

---

[13] See Gustafsson and Ding (2009a: note 8, page 593). The eastern region in the sample comprises Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shandong and Guangdong; the middle region (we use this term instead of "central" to prevent confusion with the other classification) includes Shanxi, Jilin, Anhui, Jiangxi, Henan, Hubei and Hunan. Finally, the western region comprises Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu and Xinjiang. This same classification was used, for example, in Jiang (2004: pages 79-80), although might be controversial.

[14] More precisely, in the western region we find all Uygur (Xinjiang) and Yi (Yunnan, Sichuan and Guizhou) populations, almost all Zhuang (Guangxi), around three quarters of Miao (Guizhou) and of those in the "other minorities" category (Yunnan and Guizhou). Most Manchu (93 percent), however, live in the east (Liaoning); Hui are split between 58 percent in the west (Yunnan and Xinjiang) and 42 percent in the east (Hebei and Liaoning). There are also significant groups (around 23 percent) of Miao, and "other minorities" in the middle region (Hunan).

[15] Central: Henai, Hubei, Hunan, Guangdong, and Guangxi; northwest: Chongqing, Sichuan, Guizhou, and Yunnan; southwest: Shannxi, Gansu, and Xinjiang; northern/eastern: Beijing, Hebei, Shanxi, Liaoning, Jilin, Jiangsu, Zhejiang, Anhui, Jiangxi, Shandong.

If we use the distribution across the six selected provinces in which most minorities live, instead of the region, the gap also increases in five percentage points (to 22 and 23 percent, Table 2) in the counterfactual distribution. This changes if we include other provinces in the same analysis, but there is still no reduction in poverty among minorities (17 and 18 percent).

However, if the distribution of minorities were the same as that of Han across the four regions shown in the second panel in Table 3, one third of the gap would be gone for all minorities (almost nothing for disadvantaged minorities). This reduction in the ethnic poverty gap by one third in the counterfactual situation is mostly driven by the higher concentration of minorities in the northwestern region. Based on this classification, there is a large ethnic gap in the central region (16 percentage points), a smaller one in the northwestern region, and virtually no ethnic gap in poverty rates in the northern/eastern region. Paradoxically, poverty rates are higher among Han in the southwestern region.

The facts that the location across provinces/regions alone does not seem to help much to explain the ethnic poverty gap, and that there is a lot of heterogeneity in the situation of minorities within areas, call for more in-depth analysis of the factors associated with rural poverty differentials by ethnicity in China, which is undertaken in the next subsection.

Table 2. Ethnicity and poverty in rural China in selected provinces, 2002

Province	Average income (Yuan)	Population (%)					Poverty rate (%)			
		All	Han	All Minorities	% in area	Disad. Minorities	All	Han	All Minorities	Disad. Minorities
All sample	2,539	100	100	100	11.0	100	9.1	8.4	15.5	17.2
							(0.3)	(0.4)	(1.2)	(1.4)
Liaoning	2,497	2.9	1.7	12.1	46.6	2.3	11.9	16.3	6.9	18.2
							(1.7)	(2.6)	(1.9)	(7.0)
Hunan	2,253	6.3	5.8	10.8	18.8	12.5	11.8	6.8	33.4	33.4
							(1.6)	(1.4)	(5.5)	(5.5)
Guangxi	1,721	5.2	3.6	17.5	37.2	20.1	9.3	7.2	12.9	13.0
							(1.6)	(1.8)	(2.2)	(2.9)
Guizhou	1,363	3.9	2.4	15.9	45.2	18.4	29.9	32.1	27.3	27.2
							(2.4)	(3.3)	(3.4)	(3.4)
Yunnan	1,618	4.5	1.7	27.4	66.3	30.7	20.4	44.4	8.2	8.2
							(2.6)	(5.2)	(2.2)	(2.2)
Xinjiang	2,112	1.6	0.3	12.4	82.9	13.3	12.6	5.3	14.0	14.8
							(1.8)	(2.6)	(2.1)	(2.2)
All 6 provinces	1,900	24.4	15.6	96.1	43.3	97.3	15.8	15.9	15.6	17.1
							(0.9)	(1.1)	(1.3)	(1.4)
Other provinces	2,746	75.6	84.4	3.9	0.6	2.7	7.0	7.0	13.2	18.4
							(0.4)	(0.4)	(7.0)	(10.9)
Shift-share (6 provinces)									21.6	22.8
Shift-share (all sample)									16.6	18.0

Source: Own construction using CHIP, 2002. Based on per capita household disposable income and an individual annual poverty line of 878 Yuan. Robust standard errors in parentheses. Shift-share analysis computes the average poverty rate using the distribution of Han by province.

Table 3. Ethnicity and poverty in rural China by region, 2002

Region	Average Income (Yuan)	Population (%)					Poverty rate (%)			
		All	Han	All Minorities	% in area	Disad. Minorities	All	Han	All Minorities	Disad. Minorities
All sample	2,539	100	100	100	11.0	9.7	9.1	8.4	15.5	17.2
							(0.3)	(0.4)	(1.2)	(1.4)
Eastern	3,574	32.3	34.6	13.3	4.5	2.6	5.0	5.0	6.5	16.8
							(0.5)	(0.5)	(1.7)	(6.2)
Middle	2,236	35.9	38.9	11.9	3.7	13.0	7.8	6.9	31.1	32.2
							(0.5)	(0.5)	(5.1)	(5.3)
Western	1,832	31.9	26.6	74.8	25.8	84.4	14.8	14.9	14.7	14.9
							(0.7)	(0.8)	(1.4)	(1.4)
Shift-share (3 regions)							11.6	8.4	18.2	22.3
Central	2,523	31.9	32.4	28.4	9.8	32.7	6.4	4.9	20.7	20.8
							(0.6)	(0.5)	(2.8)	(2.9)
Northwestern	1,914	18.8	16.5	44.9	26.3	51.0	13.8	13.2	15.6	15.7
							(1.0)	(1.1)	(1.9)	(2.0)
Southwestern	1,710	7.9	7.4	12.4	17.2	13.3	20.8	22.2	14.0	14.7
							(1.4)	(1.6)	(2.1)	(2.2)
Northern/eastern	2,994	41.4	44.7	14.3	3.8	3.0	6.9	6.9	6.5	14.1
							(0.5)	(0.5)	(1.6)	(5.6)
Shift-share (4 regions)							8.8	8.4	13.1	16.6

Source: Own construction using CHIP, 2002. Based on per capita household disposable income and an individual annual poverty line of 878 Yuan. Robust standard errors in parentheses. Shift-share analysis computes the average poverty rate using the distribution of Han by region. Check in section 4.b the list of provinces in each region.

c. A regression-based decomposition analysis

The previous shift-share analysis did account only for the geographical variation of ethnic groups across regions. There are other factors that could influence the larger poverty levels of some groups with respect to the others, and some of them could be captured by the region of residence when only this is controlled for. When we modified the distribution of ethnic minorities by region or province in the re-weighted distribution, we also changed their distribution by these other factors.

In the next exercise we undertake a more complete counterfactual analysis in which the counterfactual distribution of minorities reproduces on average not only the distribution by region but also other relevant factors that were found to be significant in explaining the ethnic gap in poverty rates. We explored the role of a large number of potential factors that could influence either the needs of the household or the opportunities of their members to get income. In the final regressions and decompositions only those that were found robustly significant in most specifications (and at most at 10 percent in the final one) were included to prevent potential problems induced by the small sample size. We controlled for region using two alternative specifications following our previous discussion. In the first one, we introduced a dummy indicating if the region is eastern/middle or west (omitted). In the second one, we included dummies for

the 4-region classification (omitting central). We have also included other geographical factors such as village's total population (in thousands), the timing of village's development (proxied here by whether it had electricity before 1990) and its accessibility (mountainous area or not). Among demographic factors we controlled for the number of children below 15 years old, while householder age, sex or marital status were not found to be significant in any specification. Education is taken into account using the median years of schooling among adults (at least 16 years old) in the household (that were found to be more important than householder's years of schooling). Labor characteristics of the household were proxied only by whether the householder was or not a skilled non agriculture worker because others such as the number of workers or workers in other occupations were not found significant. Other dimensions such as social capital (trust, mutual help …), cultural aspects such as attitudes towards money or competition, the possession of productive assets (amount of dry and irrigated land and the value of other productive assets), social status (whether the householder was a cadre, or she was member of the Communist Party), etc. turned out not to play any significant role in explaining the ethnic gap in income poverty rates. See the appendix for sample means of the explanatory variables and the auxiliary regressions used in the analysis.

The results of applying the decomposition methodology to explain the gap in rural poverty rates between Han and minorities are shown in Table 4. The second specification (based on the four-region classification) turned out to be much more explicative of the ethnic gap in poverty rates. About 82 percent of the observed ethnic gap in rural poverty was associated with characteristics varying across ethnic groups, while another 18 percent remains unexplained. In the first specification, more consistent with Gustafsson and Ding's (2009a) regional classification, the proportion explained was 62 percent. This first specification was analyzed in more detail in a previous version of this paper (Gradín, 2014) and raises similar qualitative results.[16] In what follows we will focus the analysis on the second specification.

[16] The main difference is that using the first classification the contribution of the region of residence is large and negative, with also a larger positive contribution of other locational factors. The contribution of the other factors is just slightly smaller.

Table 4. Explaining the ethnic gap in poverty rates in rural China, 2002

	China					
	(1)			(2)		
	Estimate	%	s.e.	Estimate	%	s.e.
Overall						
Minority	15.5		1.2	15.5		1.2
Han	8.4		0.3	8.4		0.3
Poverty gap	7.2	100	1.3	7.2	100	1.3
Explained	4.4	62.0	1.5	5.9	82.4	1.6
Unexplained	2.7	38.0	1.6	1.3	17.6	1.7
Explained						
Region(*)	-4.4	-61.4	1.9	-0.3	-3.7	1.3
Other geographical variables(**)	4.5	62.5	1.2	2.4	33.9	1.8
N. of children	1.7	24.0	0.5	1.5	20.2	0.5
Household median education	1.4	19.7	0.6	1.2	16.8	0.5
Skilled non-farmer household head	1.2	17.1	0.4	1.1	15.3	0.4
Unexplained (Only shown those significant at 10%)						
Region	3.1	42.8	1.1	2.4	34.1	1.4
Other geographical: electricity before 1990	-2.3	-31.2	1.3			

Source: Own construction using CHIP, 2002. Individual annual poverty line of 878 Yuan.
(*) Region: (1) eastern/middle, west; (2) central, northwestern, southwestern, northern/eastern.
(**) Mountainous area, village population size, electricity before 1990.

After having controlled for an array of village and household characteristics, there is no compositional effect driven by the region (as opposed to results found in the shift/share analysis). The different distribution of Han and minorities between geographical regions in China does not explain the ethnic differential in poverty rates.[17] This result indicates that the spatial distribution of minorities in China plays a different role in explaining the ethnic poverty gap than in other economic contexts. For example, Gradín (2009) showed that African descents in Brazil tend to concentrate in the poorest northwestern region compared with whites and this largely contributed to explain their racial poverty gaps (although the focus there was overall poverty, not only rural). However, the lack of relevance of this regional distribution in explaining rural poverty in China seems to be linked to the analysis of the bottom of the distribution using the official low income line, while using a significantly higher poverty line this role would increase.[18] We will turn back to this in the next section.

Regarding the other characteristics, all of them have positive sign indicating the higher prevalence among minorities of characteristics associated with higher poverty. Among them, the most important are the other geographical factors capturing the higher concentration of some minorities in less developed,

[17] Note, however, that in a more general analysis of poverty over the whole country, the overrepresentation of minorities in the rural areas, given the large rural-urban income gap, is expected to play the most crucial factor, as Hannum and Wang (2012) already suggested.

[18] In fact, if we increased the poverty line by a half, equalizing the regional distribution of minorities with that of Han would have a reduction in the ethnic differential in rural poverty in China of about 4 percentage points out of 21 (about 18 percent of the gap).

mountainous areas of the country (that are expected to make assets less productive and reduce their labor market opportunities). This explains about on third of the observed poverty gap (2.4 percentage points), indicating that this is a very important factor to understand why rural poverty is higher among minorities in China.[19] It seems that is this factor the main determinant of the higher (severe) poverty of minorities rather than the geographical region of residence.

Another important factor driving higher rural poverty among minorities is the larger number of children they have (1.3 versus 0.9 aged 15 or less) that increases household's needs and thus the poverty differential by about 1.5 percentage points (20 percent of the observed gap). This factor was also found to be important in other cases, such as African descents in the US, Brazil, or South Africa. However, the nature of this is different in the case of China, where the larger number of children is likely to be the effect of the more flexible application of the one-child policy among minorities, their couples are often granted a second-child or a third-child exemption (e.g. Baochang et al, 2007) rather than cultural factors or a more limited access to family planning that could be more important in other countries. The lower attained education among minorities also explains a significant differential of about 1.2 percentage points (17 percent of the observed gap). Despite the existence of affirmative action policies for matriculation of people claiming minority status into colleges and universities, and subsidies for minority students, the inadequacy of educational resources in many rural and particularly national minority communities implies an important educational barrier for minorities (e.g. Ross, 2006). Indeed, the average median years of schooling among adults in the household is 6.5 for minorities and 7.3 for Han. The lower number of skilled non-agriculture workers among minorities' householders explains an additional 1.1 percentage points of the ethnic poverty gap (15 percent). The proportion of population where the household head is a skilled worker in the non-agriculture sector is 6.4 and 12.4 percent respectively.

As a consequence of the role played by the different distribution of characteristics among minorities and Han, the conditional poverty gap - what remains after equalizing the distribution across all the discussed dimensions- would be of about 1.3 percentage points and is not statically significant. We still inspect the detailed unexplained effect as this could be the result of counterbalancing effects. Again location is somehow related with the gap. Only the coefficients associated with region are significant. The positive sign of this unexplained effect (significant at 10-percent level) is the consequence of the fact that, minorities tend to be poorer than Han when they live in the same region, even if they have other similar characteristics.

[19] The average population size of the village was also smaller among minorities (1.9 versus 2.1 thousands). Additionally, 59 percent of people reporting to belong to any of the ethnic minorities live in mountainous areas and 19 percent in villages without electricity before 1990 (compared with 16 and 9 percent of Han, respectively). This is expected to make the larger assets of minorities less productive (3.3 and 4.9 Mu of irrigated and dry land versus 2.9 and 2.5 Mu; an average value of 5,351 Yuan of other productive assets versus 4,640 Yuan).

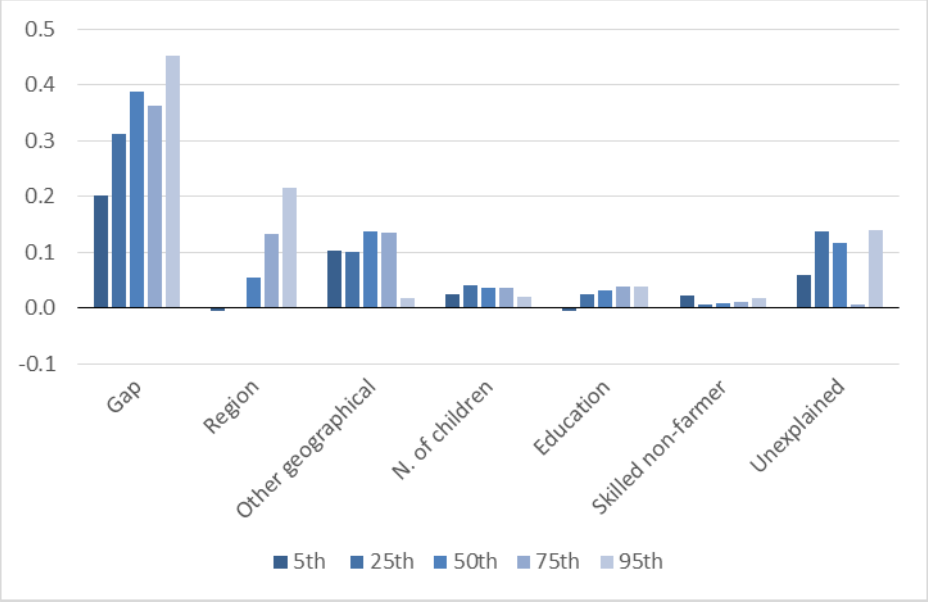
d. Ethnic income gap decomposition across the distribution

In this section we go a step further and analyze in a more consistent way whether the results shown in the previous subsection are specific for the bottom of the income distribution or, on the contrary, can be extended elsewhere. For that, we undertake regression-based decomposition at different points of the distribution, based on RIF functions of log-income quantiles. More specifically, we decompose the gap at 5th, 25th, 50th, 75th, and 95th percentiles using the same set of explanatory variables used in the previous epigraph in (specification 2, the same analysis with specification 1, producing similar results, was done in Gradín, 2014).

The results are reported in Figure 2. The log-income gap is generally higher for higher quantiles. But it turns out that the role of the ethnic regional distribution also changes dramatically as we move up to higher quantiles from virtually zero values (consistent with the previous decomposition) to help to explain a substantial and significant share of the gap at the top of the distribution. This indicates that with a more even regional distribution the gap would be much lower at high quantiles, despite the lack of effect found at the bottom. The role of other geographical factors is however more similar along the distribution, with the largest absolute effect at the 75th percentile, although at the very top we found no effect. The role of other factors such as the number of children, schooling and occupation is more modest, but generally significant and increasing as we move up to higher quantiles.

Figure 2. Decomposition of the ethnic income gap in rural China, 2002

RIF decomposition at various percentiles of household disposable income (in logs)



Region: central, northwestern, southwestern, northern/eastern.
Other geographical variables: Mountainous area, village population size, electricity before 1990.
For the numbers see Table A3 in the Appendix.
Source: Own construction using CHIP, 2002.

5. Conclusions

Poverty and geography are strongly related in China. Extreme poverty is mostly rural and shows a clear regional pattern, with poverty being highest in the northwestern and southwestern regions. There is also a connection between rural poverty and ethnicity because its incidence is higher among most ethnic minorities than among Han. We have explored here the role of location, along other socioeconomic factors, in explaining a substantial part of the inter-ethnic gap in poverty levels in rural China in 2002.

We showed that, unlike what was previously suggested in the literature, the ethnic gap in poverty would not be lower if ethnic minorities were distributed like Han across geographical regions. The incidence of poverty is larger among minorities because they tend to live in the least developed and mountainous areas that are being more slowly benefiting from the strong country's economic growth. This points out to the crucial role of the local development of areas predominantly populated by minorities as the key policy to close this gap. Poverty is higher among minorities also because of their less economic opportunities, given their lower education and engagement in off-farm activities, in a scenario where market forces increasingly determine rural incomes in China. Poverty is also higher among minorities because they generally have more children, the natural consequence of ethnic exceptions introduced in the one-child policy.

We have also identified a distributional pattern in the ethnic inequality in incomes. The inter-ethnic differential in rural incomes in China tends to be proportionally higher for higher incomes, with the geographical region of residence being the main driving force. While a more equal regional distribution would not decrease the ethnic gap for incomes below the median, the opposite is true above the median, becoming the main explanatory factor of the ethnic inequality for top incomes. Clearly, the different economic opportunities that are associated with living in a certain geographical region become more important to explain ethnic inequality for higher incomes. This is consistent with the finding that the region of residence becomes also more explicative when measuring the inter-ethnic poverty gap using a higher poverty threshold.

The limitation of the CHIP 2002 database leaves room for more in-depth investigation of the nature of the inter-ethnic inequalities in China as soon as new adequate data are released in the future.

References

Baochang, G., W. Feng, G. Zhigang, and Z. Erli (2007), "China's Local and National Fertility Policies at the End of the Twentieth Century", *Population and Development Review*, 33: 129–148.

Bhalla, A.S. and Luo, D. (2013), *Poverty and Exclusion of Minorities in China and India*, The Palgrave Macmillan: Hampshire, UK.

Blinder, A.S. (1973), ''Wage discrimination: reduced form and structural estimates'', *Journal of Human Resources*, 8(4): 436–55.

Cappellari, L. and S.P. Jenkins (2004), "Modeling low income transitions", *Journal of Applied Econometrics*, 19: 593–610.
Even, W.E. and D.A. Macpherson (1990), "Plant size and the decline of unionism", *Economics Letters*, 32(4): 393-398.
Even, W. E. and D.A. Macpherson (1993), "The decline of private-sector unionism and the gender wage gap", *Journal of Human Resources*, 28(2): 279-296.
Firpo, S., N. M. Fortin, and T. Lemieux (2007), "Decomposing Wage Distributions Using Recentered Influence Function Regressions," Unpublished Manuscript, University of British Columbia.
Firpo, S., N.M. Fortin, and T. Lemieux, (2009), "Unconditional Quantile Regressions", *Econometrica*, 77: 953–973.
Gradín, C. (2009), "Why is poverty so high among Afro-Brazilians? A decomposition analysis of the racial poverty gap", *Journal of Development Studies*, 45(9): 1-38.
Gradín, C. (2012), "Poverty among minorities in the United States: Explaining the racial poverty gap for Blacks and Latinos", *Applied Economics*, 44(29): 3793-3804.
Gradín, C. (2013), "Race, poverty, and deprivation in South Africa", *Journal of African Economies*, 22(2): 187-238.
Gradín, C. (2014), "Rural Poverty and Ethnicity in China", EQUALITAS Working Paper No 32.
Gustafsson B. and S. Ding (2009a), "Temporary and persistent poverty among ethnic minorities and the majority in rural China", *Review of Income and Wealth*, 55(1): 588-606.
Gustafsson, B. and S. Ding (2009b), "Villages where China's Ethnic Minorities Live", *China Economic Review*, 20: 193–207.
Gustafsson, B. and S. Ding (2014), Why Is There No Income Gap between the Hui Muslim Minority and the Han Majority in Rural Ningxia, China?, *The China Quarterly*, 220: 968-987.
Gustafsson, B. and Li, S. (2003), "The Ethnic Minority - Majority Income Gap in Rural China during Transition", *Economic Development and Cultural Change*, 51(4):805-822.
Hannum, E.C. and M. Wang (2012), "China: A Case Study in Rapid Poverty Reduction", in Hall, G and Patrinos, H (Eds) *Indigenous People, Poverty, and Development*, Chapter 5: 149 – 204, Cambridge University Press: Cambridge.
Imai, K.S. and You, J. "Poverty (2014), "Dynamics of Households in Rural China", *Oxford Bulletin of Economics and Statistics*, 76(6): 898–923.
Jiang, X. (2004), *FDI in China: Contributions to Growth, Restructuring, and Competitiveness*, Nova Science Publishers: New York.
Li, S. (2009), "Chinese Household Income Project, 2002", ICPSR21741-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
Li, S., H. Sato, and T. Sicular (2013), *Rising Inequality in China: Challenges to a Harmonious Society*, Cambridge University Press: New York.

Luo, C. and T. Sicular (2013), "Inequality and Poverty in Rural China", in S. Li, H. Sato, and T. Sicular (Eds.) *Rising Inequality in China: Challenges to a Harmonious Society*, Cambridge University Press: New York, 197–229.

Oaxaca, R.L. (1973), ''Male-female wage differentials in urban labor markets'', *International Economic Review*, 14(3): 693–709.

Oaxaca, R.L. and M.R. Ransom (1999), "Identification in detailed wage decompositions", *Review of Economics and Statistics*, 81: 154-157.

Ravallion, M. and S. Chen (2007), "China's (uneven) progress against poverty", *Journal of Development Economics*, 82: 1– 42.

Ross, H. (2006), "Where and Who are the World's Illiterates: China" UNESCO, Global Monitoring Report China Country Study.

Sato, H. and S. Ding (2012), "Ethnicity and Income in China: The Case of Ningxia", CCES Discussion Paper Series, No.46, Center for Research on Contemporary Economic Systems, Graduate School of Economics, Hitotsubashi University.

Sicular, T., X. Yue, B. Gustafsson and S. Li (2007), "The urban–rural income gap and inequality in China", *Review of Income and Wealth*, 53(1): 93-126.

Song, J, T. Sicular, and X. Yue (2013), "Appendix II. The 2002 and 2007 CHIP surveys: sampling weights and combining the urban, rural and migrant samples", in S. Li, H. Sato, and T. Sicular (Eds.) *Rising Inequality in China: Challenges to a Harmonious Society*, Cambridge University Press: New York, 465–485.

Xia, Q. (2009), "The Pattern and Evolution of Poverty in Rural Poverty", School of Economics, China Institute for Employment Research Working Paper, China Institute for Employment Research, Beijing.

Yun, M-S. (2004), "Decomposing Differences in the First moment", Economics Letters, 82(2), pp. 275-280.

Yun, M.-S. (2005), "A simple solution to the identification problem in detailed wage decompositions", *Economic Inquiry*, 43: 766–772.

Yun, M-S. (2008), "Identification problem and detailed Oaxaca decomposition: a general solution and inference", *Journal of Economic and Social Measurement*, 33(1), pp. 27-38.

Wang, G. (2008a), *Understanding Inequality and Poverty in China: Methods and Applications*, Palgrave-Macmillan: New York.

Wang, G. (2008b), *Inequality and growth in modern China*, UN-WIDER Studies in Development Economics, Oxford University Press: New York.

APPENDIX

Table A1. Sample means of explanatory variables (other than region)

China	Han	Minorities
Mountainous area (%)	15.9	59.3
Village population (thousands)	1.9	1.9
Electricity before 1990 (%)	91.3	80.7
Number of children (<16)	0.9	1.3
Median schooling	7.3	6.5
Skilled worker (household head) (%)	12.4	6.4

Source: Own construction using CHIP, 2002.

Table A2. Auxiliary regressions by ethnic group: Logit of the probability of being poor and RIF regressions of (log) per capita income

	Logit		RIF regressions									
			5th		25th		50th		75th		95th	
Minorities	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Northwestern	-0.615	0.025	0.176	0.077	0.166	0.029	0.207	0.025	0.285	0.028	0.103	0.031
Southwestern	-0.762	0.051	0.168	0.087	0.186	0.035	0.204	0.034	0.288	0.044	0.299	0.061
Northern/eastern	-0.641	0.047	0.180	0.078	0.195	0.032	0.412	0.030	0.762	0.042	0.857	0.080
Mountainous area	0.461	0.038	-0.157	0.055	-0.205	0.025	-0.276	0.028	-0.266	0.040	-0.048	0.055
Village population	-0.210	0.011	0.030	0.015	0.033	0.007	-0.007	0.007	-0.032	0.008	-0.007	0.011
Electricity before 1990	-0.530	0.024	0.387	0.085	0.166	0.031	0.161	0.024	0.125	0.022	-0.043	0.029
N of children (<16)	0.383	0.011	-0.077	0.025	-0.123	0.011	-0.111	0.010	-0.108	0.011	-0.063	0.014
Median schooling	-0.132	0.005	-0.006	0.010	0.029	0.005	0.039	0.005	0.048	0.006	0.048	0.007
Skilled worker (hh head)	-1.618	0.084	0.363	0.035	0.119	0.035	0.161	0.038	0.194	0.049	0.309	0.087
Intercept	-0.469	0.060	6.119	0.124	6.708	0.056	7.032	0.051	7.370	0.058	7.865	0.075
N observations	5,298		5,298		5,298		5,298		5,298		5,298	
LR χ^2(27) / F(7, 5290)	7,676		22		78		182		174		34	
Prob. > χ^2 / Prob. > F	0		0		0		0		0		0	
Pseudo R^2 / R^2	0.115		0.029		0.136		0.197		0.212		0.102	
			5th		25th		50th		75th		95th	
Han	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Northwestern	0.869	0.015	-0.283	0.028	-0.193	0.017	-0.181	0.014	-0.196	0.014	-0.247	0.021
Southwestern	1.367	0.016	-0.678	0.044	-0.515	0.019	-0.363	0.014	-0.269	0.013	-0.214	0.020
Northern/eastern	0.517	0.013	-0.150	0.017	-0.029	0.011	0.052	0.010	0.103	0.011	0.174	0.021
Mountainous area	0.879	0.011	-0.500	0.030	-0.307	0.014	-0.110	0.011	0.000	0.011	-0.006	0.019
Village population	-0.039	0.005	-0.011	0.007	0.021	0.004	0.015	0.004	0.019	0.004	-0.035	0.007
Electricity before 1990	-0.299	0.014	0.122	0.035	0.177	0.019	0.177	0.014	0.167	0.013	0.103	0.022
N of children (<16)	0.319	0.005	-0.105	0.010	-0.108	0.006	-0.127	0.005	-0.145	0.005	-0.162	0.009
Median schooling	-0.072	0.002	0.033	0.004	0.034	0.002	0.033	0.002	0.040	0.002	0.044	0.005
Skilled worker (hh head)	-1.031	0.022	0.231	0.016	0.201	0.013	0.219	0.012	0.238	0.016	0.431	0.037
Intercept	-2.565	0.026	6.591	0.050	7.025	0.028	7.378	0.023	7.698	0.023	8.416	0.040
N observations	32,598		32,598		32,598		32,598		32,598		32,598	
LR χ^2(27) / F(7, 5290)	34,093		108		422		559		429		118	
Prob. > χ^2 / Prob. > F	0		0		0		0		0		0	
Pseudo R^2 / R^2	0.115		0.053		0.105		0.104		0.090		0.037	

Source: Own construction using CHIP, 2002.

Table A3. RIF decomposition of the ethnic gap in (log) income at different percentiles (displayed in Figure 2)

	5th			25th			50th			75th			95th		
	estimate	%	St. Error	estimate	%	St. Error	estimate	%	St. Error	estimate	%	St. Error	estimate	%	St. Error
Overall															
Han	6.61		0.02	7.27		0.01	7.67		0.01	8.07		0.01	8.70		0.02
Minority	6.41		0.05	6.96		0.02	7.28		0.02	7.70		0.03	8.24		0.03
Gap	0.20		0.05	0.31		0.03	0.39		0.02	0.36		0.03	0.45		0.03
Explained	0.14	70.0	0.06	0.17	55.9	0.03	0.27	70.1	0.03	0.36	97.9	0.04	0.31	69.2	0.06
Unexplained	0.06	30.0	0.07	0.14	44.1	0.03	0.12	29.9	0.03	0.01	2.1	0.05	0.14	30.8	0.07
Explained															
Region	-0.01	-2.6	0.04	0.00	0.4	0.02	0.05	14.1	0.02	0.13	36.9	0.03	0.22	47.9	0.05
Other geographical	0.10	51.5	0.05	0.10	32.2	0.03	0.14	35.7	0.03	0.14	37.2	0.04	0.02	3.9	0.05
N. of children	0.03	12.8	0.02	0.04	13.3	0.01	0.04	9.6	0.01	0.04	10.0	0.01	0.02	4.7	0.01
Education	-0.01	-2.5	0.02	0.02	7.7	0.01	0.03	8.3	0.01	0.04	10.6	0.01	0.04	8.6	0.01
Skilled worker (hh head)	0.02	10.8	0.01	0.01	2.3	0.00	0.01	2.5	0.01	0.01	3.2	0.01	0.02	4.1	0.01
Unexplained															
Region	0.13	63.5	0.03	0.11	36.4	0.02	0.07	16.9	0.01	0.01	3.8	0.02	-0.01	-2.5	0.03
Other geographical	-0.07	-34.2	0.09	0.02	5.7	0.04	-0.01	-1.9	0.04	0.02	5.9	0.03	-0.01	-1.4	0.04
N. of children	-0.03	-13.0	0.06	0.01	4.4	0.03	-0.01	-3.8	0.02	-0.03	-9.3	0.02	-0.09	-20.2	0.03
Education	0.28	141.8	0.16	0.03	10.2	0.08	-0.04	-11.3	0.08	-0.06	-15.7	0.10	-0.03	-5.9	0.12
Skilled worker (hh head)	0.05	24.7	0.03	-0.03	-9.9	0.03	-0.02	-5.7	0.03	-0.02	-4.6	0.04	-0.05	-10.2	0.07
Intercept	-0.31	-152.8	0.20	-0.01	-2.6	0.11	0.14	35.6	0.10	0.08	22.0	0.11	0.32	71.0	0.14

Source: Own construction using CHIP, 2002.