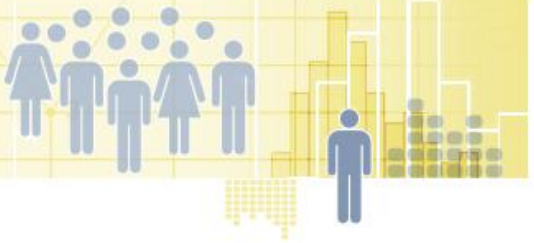
EQUALITAS
Economics of Inequality and Poverty Analysis



EQUALITAS Working Paper No. 32

Rural Poverty and Ethnicity in China

Carlos Gradín

2014

Rural Poverty and Ethnicity in China

Carlos Gradín*

(Universidade de Vigo, EQUALITAS)

September, 2014

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. Results show that rural poverty would be higher among minorities if they had the same regional distribution of Han. On the contrary, the ethnic poverty differential is reduced after equalizing other 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. Additionally, I show that poverty among minorities is not even higher because some of these adverse characteristics have a smaller negative impact on them than on Han. Finally, the ethnic per capita (log)income differential is shown to be higher for higher percentiles, with an increasing role of location as the main driver of these 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* and *Xunta de Galicia*.

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. 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.

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 Ding (2009), after documenting that poverty was higher among ethnic minorities in rural China, focused on a geographical explanation for this. According to them this ethnic differential was mainly due to the higher concentration of minorities in the less-developed western provinces 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 (2010) agree with this view after estimating a model for probability of being poor conditioned on a dummy indicating the minority status plus a set of covariates reflecting household- and community-level characteristics.[1]

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 be larger, not smaller, because minorities are underrepresented in the part of the country (central region) where the ethnic gap is the largest and overrepresented where there is no such a gap (western region). 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

[1] Other studies have focused on specific ethnic groups or provinces (e.g. Hui in Ningxia in Gustafsson and Ding, 2011 and Sato and Ding, 2012).

and less developed villages. These have been used in Hannum and Wang's (2010) 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). 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 region from the National Bureau of Statistics (NBS).[2]

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 provinces 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

[2] *China Statistical Yearbooks Database*, 2002 at http://www.yearbookchina.com.

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.[3] Furthermore, the use of CHIP 202 allows to read the results in the context of the previous literature that has also used this same database.

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 (2009) 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 Chuang, Hui, Uygurs, Yi, Miao, Manchu, and an additionally 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.[4] Nonetheless, for most of the paper all minorities will be 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

[3] 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 interethnic inequality. For that reason, it has not been used in this study.

[4] See Hannum and Wang (2010) for a more detailed discussion of this classification.

coefficients. This regression is estimated separately for Han and ethnic minorities. 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). 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):

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

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):

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

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.[5] 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.[6]

---

[5] 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.

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

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 $IF(y; q_\tau)$[7]:

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

Where $\mathbf{1}()$ is an indicator function that takes value 1 if the specified condition is satisfied and 0 otherwise. 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:

$$q_\tau^0 - q_\tau^1 = \overline{RIF(y; q_\tau)}^0 - \overline{RIF(y; q_\tau)}^1 = \bar{X}^0\hat{\gamma}_\tau^0 - \bar{X}^1\hat{\gamma}_\tau^1 = (\bar{X}^0 - \bar{X}^1)\hat{\gamma}_\tau^1 + \bar{X}^0(\hat{\gamma}_\tau^0 - \hat{\gamma}_\tau^1). \textbf{(8)}$$

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 $W_k^{\Delta\beta} = \bar{x}_k^0(\hat{\gamma}_{\tau k}^1 - \hat{\gamma}_{\tau k}^0)$.[8] 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).[9] 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

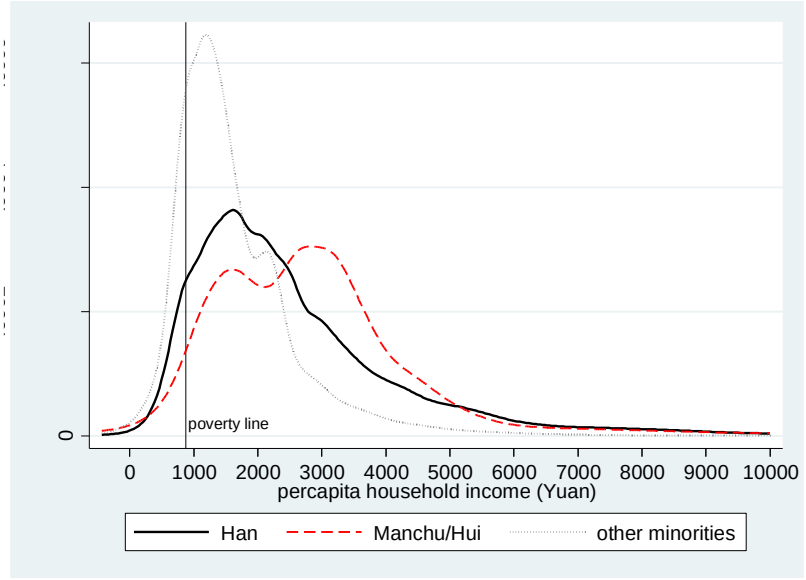
[7] An influence function is a statistical tool that measures the influence of each individual observation on any statistic.

[8] 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.

[9] 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).

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.

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	
Chuang	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 these ethnic differentials in wellbeing might be largely driven by ethnic minorities having poorer endowments. For example, as previous literature highlights (Gustafsson and Ding, 2009), given the extraordinary extension and diversity of rural China, the region of residence is one important candidate. That is, it could be the case that higher poverty among minorities were the result of them being overrepresented in the poorest region of the country (western) thus producing a mere geographical compositional effect. For historical reasons minorities are indeed highly concentrated in specific areas in China where they are native or where they settled long time ago. As Table 2 reports, 75 percent of minorities in the CHIP sample (85 percent in the case of disadvantaged minorities) live in the western region, while Han are more evenly spread all over the country.[10] But this geographical distribution has something that seems to be peculiar. While the overall poverty rate is globally higher in the western region (15 percent) compared with central and eastern regions (8 and 5 percent respectively), it is in the central region where the highest ethnic differential can be found due to the larger poverty incidence among minorities there.[11]

[10] More precisely, in the western region we find all Uygur (Xinjiang) and Yi (Yunnan, Sichuan and Guizhou) populations, almost all Chuang (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 central region (Hunan).

[11] The eastern region in the sample comprises Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shandong and Guangdong; the central region includes Shanxi, Jilin, Anhui, Jiangxi, Henan,

A simple shift-share analysis, shown in Table 2, that gives all minorities the same geographical representation of Han, shows that had minorities been more evenly distributed across the country, 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. This calls 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 by region, 2002

Population	Eastern	Central	Western	Total	
Han	34.6	38.9	26.6	100	
All minorities	13.3	11.9	74.8	100	
Disadvantaged Minorities (Manchu/Hui excluded)	2.6	13.0	84.4	100	
All	32.2	35.9	31.9	100	
Poverty rate (H)	**Eastern**	**Central**	**Western**	**Total**	**Shift-share**
Han	5.0	6.9	14.9	8.4	8.4
All minorities	6.5	31.1	14.7	15.5	18.2
Disadvantaged Minorities (Manchu/Hui excluded)	16.8	32.3	14.9	17.2	22.3
All	5.0	7.8	14.8	9.2	11.6

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

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. 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 (east/central or west) and 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

Hubei and Hunan. Finally, the western region comprises Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu and Xinjiang.

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 3. Results for the country as a whole show that about 62 percent of the observed ethnic gap in rural poverty is associated with characteristics varying across ethnic groups. Another 38 percent remains unexplained.

Table 3. Explaining the ethnic gap in poverty rates in rural China, 2002

	China			East/central			West	
	Estimate	%	Std. Err.	Estimate	%	Std. Err.	Estimate	Std. Err.
Overall								
Minority	15.5		1.2	18.1		2.6	14.7	1.4
Han	8.4		0.3	6.0		0.4	14.9	0.8
Poverty gap	7.2	100	1.3	12.1	100	2.6	-0.2	1.6
explained	4.4	62.0	1.5	7.2	59.3	3.2	4.6	1.0
unexplained	2.7	38.0	1.6	4.9	40.7	3.0	-4.8	1.4
Explained								
Region	-4.4	-61.4	1.9					
Other geographical variables(*)	4.5	62.5	1.2	6.0	49.5	3.1	1.9	0.7
N. of children	1.7	24.0	0.5	0.7	5.9	0.7	0.9	0.4
Household median education	1.4	19.7	0.6	-0.0	-0.3	0.3	0.9	0.4
Skilled non-farmer household head	1.2	17.1	0.4	0.5	4.1	0.5	0.8	0.4
Unexplained (only shown those significant at 10%)								
Region	3.1	42.8	1.1					
Other geographical: electricity before 1990	-2.3	-31.2	1.3				-3.1	1.2

Source: Own construction using CHIP, 2002. Individual annual poverty line of 878 Yuan.

(*) Mountainous area, village population size, electricity before 1990

After having controlled for an array of village and household characteristics, there remains a large compositional effect driven by the region that is even larger than the one obtained using the shift/share analysis. The different regional distribution of Han and minorities is associated with -4.4 percentage points of the differential. The negative sign indicates that the gap would be higher if minorities shared the same geographical distribution of Han (between both east/central and west). The difference is that now we show that this is robust to controlling for

other possible factors that could be driving the result. Thus, the observed 7.2 percentage-points differential would increase up to 11.6. 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, 2013) showed that African descents in Brazil and South Africa tend to concentrate in the poorest regions or provinces compared with whites and this largely contributed to explain their racial poverty gaps (although the focus there was overall poverty, not only rural).[12] However, the role of the 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 disappears.[13] 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, mountainous areas of the country (that are expected to make assets less productive and reduce their labor market opportunities). This explains about 4.5 percentage points (62.5 percent) of the observed gap, indicating that this is a very important factor to understand why rural poverty is higher among minorities in China.[14] Note that the village's effect cancels out with the regional effect, such that globally speaking the geography does not explain anything of the observed ethnic gap in rural poverty in China. But one should understand that while the regional distribution of minorities is actually hiding a higher ethnic poverty, minorities living in less developed areas largely explains their higher poverty.

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 1.7 percentage points (24 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

[12] 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 (2010) already suggested.

[13] In fact, if we increased the poverty line by a half, equalizing the regional distribution of minorities with that of Han would have no effect on the ethnic differential in rural poverty in China, unlike the other factors mentioned later. In this case there is a differential in poverty rates in west China between minorities and Han using this higher poverty line (44 versus 34 percent), although the gap is still much higher in the case of central China (63 versus 20 percent) and lower in the east (15 versus 12 percent). However, after controlling for other characteristics, the equalization of the distribution by region among Han and minorities does not increase the ethic poverty gap.

[14] 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).

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 1.4 percentage points (20 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 1.2 percentage points of the ethnic poverty gap (17 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 would be of about 2.7 percentage points. This is what remains after equalizing the distribution across all the discussed dimensions. An inspection of the detailed decomposition of the unexplained differential shows that again location is somehow related with this remaining gap. In fact only two detailed effects are significant, region and other geographical factors (more specifically, having electricity before 1990). And again both with opposite signs. The positive sign of region is the consequence of the fact that, other things being equal, minorities tend to be poorer if they live in the east/central region compared to minorities in the west, while the opposite is true for Han. The negative sign of other geographical factors is the results that living in villages that had electricity before 1990 reduces more poverty risk among Han than among minorities. That is, although minorities live in areas that developed later, this seems to be less relevant to explain their poverty than it is for Han.

The strong importance of the regional distribution and the fact that region differentially affects minorities and Han suggest that the same decomposition exercise could be reproduced separately in each region to see how factors explaining the poverty gap vary in the east/central and western regions. This is done in Table 3 too.

In the case of the east/central region, the observed gap is larger, 12 percentage points. In this case near 60 percent of it can be explained by the different characteristics of Han and minorities. The only single factor that is statistically significant and explains half the entire gap (6 percentage points) is minorities living in less developed and mountainous villages. In the west of the country there is virtually no poverty gap between Han and minorities. However, it is still interesting to check whether this is the consequence of higher similarity in characteristics or rather the result of strong counterbalancing forces. The large and significant characteristics effect of 4.6 percentage points indicates the poorer endowments of minorities in this region across all dimensions, geographically

(explaining a 1.9 percentage-points differential), having more children, less attained education and skilled non agriculture workers (explaining near 1 percentage-point each). Why, then, do not we find any differential in poverty rates among minorities in the west region? This is due to the lower incidence of poverty associated with living in villages with later development among minorities there. Note that this coefficient effect is only found in the west, thus explaining the overall effect found for the whole sample.

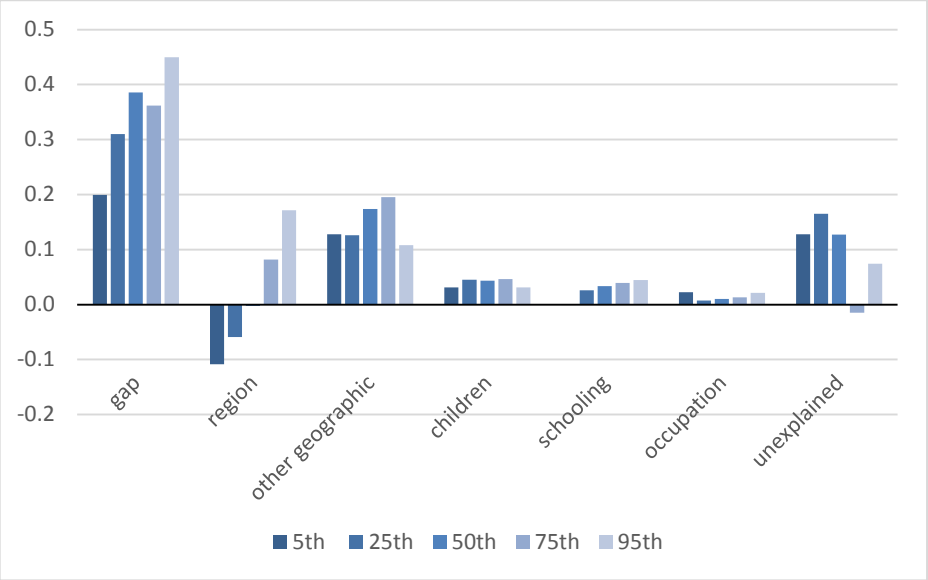
d. Ethnic income gap decomposition across the distribution

In this section we go a step further and analyze in a more consistent way whether the results shown in the previous section 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 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles using the same set of explanatory variables used in the previous epigraph.

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 negative values (consistent with the previous decomposition) to play no role at the median, and helping 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, the opposite to what was found at the bottom. The role of other geographical factors is however more similar along the distribution, although with the largest absolute effect at the 75^{th} percentile. 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)



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 higher in the central and western regions, and lower in the east. 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 in fact be higher if ethnic groups were not concentrated on the western areas of the country because the ethnic gap is the largest in central China (even if overall poverty is lower). However this regional effect vanishes with higher poverty thresholds because the differential in poverty by ethnicity increases all over the country.

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. Especially in the western region, 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 location being the main driving force. While a more equal regional distribution would increase the ethnic gap for incomes below the median, the opposite is true above the median. The importance of other locational factors is also increased as we move up in the income scale.

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.

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, (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.

Gustafsson, B. and S. Ding (2006), "Villages where China's Ethnic Minorities Live", IZA DP No. 2418.

Gustafsson B. and S. Ding (2009), "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 (2012), "Assessing Ethnic Disparities in Income and Poverty in China – The Case of Han and Hui in Ningxia", Paper prepared for the 4th World Conference on Remedies to Racial and Ethnic Economic, Inequality, Minnesota.

Hannum, E.C. and M. Wang (2010), "Ethnicity, Socioeconomic Status, and Social Welfare in China", World Bank Report, G. Hall and H.A. Patrinos (eds.), Washington DC: World Bank.

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.

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.

Li, S. (2009), "Chinese Household Income Project, 2002", ICPSR21741-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

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.

Xia, Q. (2009), "The Pattern and Evolution of Poverty in Rural Poverty", School of Economics, Peking University, mimeo.

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.

APPENDIX

Table A1. Sample means of explanatory variables

China	**Han**	**Minorities**
Eastern/Central (%)	73.4	25.2
Mountainous area (%)	15.9	59.3
Village population (thousands)	1.9	1.9
Electricity before 1990 (%)	91.3	80.7
N of children (<16)	0.9	1.3
Median schooling	7.3	6.5
Skilled worker (hh head) (%)	12.4	6.4
East Central		
Mountainous area (%)	10.5	66.7
Village population (thousands)	1.8	1.4
Electricity before 1990 (%)	94.7	94.4
N of children (<16)	0.9	0.8
Median schooling	7.4	7.6
Skilled worker (hh head) (%)	13.1	10.4
Western		
Mountainous area (%)	30.8	56.8
Village population (thousands)	2.1	2.0
Electricity before 1990 (%)	81.9	76.1
N of children (<16)	1.1	1.5
Median schooling	6.9	6.2
Skilled worker (hh head) (%)	10.4	5.0

Source: Own construction using CHIP, 2002.

Table A2. Probability of being poor of Han and minorities (Logit)

	All rural				Eastern/Central				Western			
	Han		Minorities		Han		Minorities		Han		Minorities	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Eastern/Central	0.732	0.028	-0.680	0.010								
Mountainous area	0.672	0.025	0.941	0.011	0.449	0.060	0.293	0.018	0.605	0.028	1.467	0.015
Village population	-0.086	0.010	-0.103	0.004	-0.380	0.038	-0.161	0.007	-0.062	0.010	-0.082	0.006
Electricity before 1990	-0.803	0.025	-0.209	0.014	0.405	0.088	-0.266	0.024	-0.985	0.026	-0.093	0.018
N of children (<16)	0.411	0.011	0.285	0.005	0.866	0.025	0.274	0.007	0.259	0.012	0.287	0.008
Median schooling	-0.140	0.005	-0.070	0.002	-0.219	0.013	-0.079	0.003	-0.132	0.005	-0.055	0.003
Skilled worker (hh head)	-1.644	0.084	-1.011	0.022	-1.407	0.137	-0.835	0.026	-1.436	0.106	-1.339	0.039
Intercept	-1.254	0.046	-1.534	0.022	-1.243	0.143	-1.878	0.034	-0.931	0.050	-2.021	0.031
N observations	5,302		32,616		1,224		24,055		4,078		8,561	
LR $\chi^2(27)$	7,653		30,714		3,591		5,337		4,973		16,962	
Prob > χ^2	0		0		0		0		0		0	
Pseudo R^2	0.114		0.085		0.195		0.026		0.103		0.121	

Source: Own construction using CHIP, 2002.

Table A3. Regressions of RIF of (log) per capita income by ethnic group

	5th		25th		50th		75th		95th	
Minorities	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Eastern/Central	-0.212	0.068	-0.108	0.026	0.009	0.024	0.186	0.032	0.373	0.052
Mountainous area	-0.183	0.044	-0.244	0.019	-0.336	0.020	-0.378	0.026	-0.248	0.036
Village population	-0.004	0.010	0.006	0.005	-0.034	0.007	-0.063	0.007	-0.032	0.009
Electricity before 1990	0.479	0.077	0.231	0.030	0.236	0.024	0.215	0.024	-0.026	0.029
N of children (<16)	-0.090	0.026	-0.131	0.011	-0.127	0.010	-0.135	0.011	-0.090	0.014
Median schooling	-0.002	0.009	0.030	0.005	0.040	0.005	0.047	0.006	0.053	0.007
Skilled worker (hh head)	0.369	0.035	0.121	0.034	0.172	0.039	0.218	0.053	0.352	0.092
Intercept	6.299	0.106	6.889	0.046	7.252	0.042	7.688	0.048	8.132	0.061
N observations	5,298		5,298		5,298		5,298		5,298	
F(7, 5290)	27		88		156		140		35	
Prob > F	0		0		0		0		0	
Pseudo R^2	0.029		0.128		0.166		0.155		0.073	
	5th		**25th**		**50th**		**75th**		**95th**	
Han	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Eastern/Central	0.254	0.020	0.257	0.012	0.287	0.010	0.313	0.010	0.333	0.015
Mountainous area	-0.545	0.030	-0.340	0.014	-0.127	0.011	-0.010	0.011	-0.025	0.019
Village population	0.009	0.007	0.036	0.004	0.030	0.003	0.034	0.004	-0.020	0.007
Electricity before 1990	0.086	0.035	0.141	0.019	0.139	0.014	0.128	0.013	0.072	0.022
N of children (<16)	-0.097	0.010	-0.106	0.006	-0.126	0.005	-0.145	0.005	-0.163	0.009
Median schooling	0.031	0.004	0.031	0.002	0.030	0.002	0.037	0.002	0.043	0.005
Skilled worker (hh head)	0.229	0.016	0.204	0.013	0.223	0.012	0.243	0.016	0.437	0.038
Intercept	6.252	0.049	6.783	0.027	7.166	0.022	7.491	0.022	8.208	0.038
N observations	32,598		32,598		32,598		32,598		32,598	
F(7, 5290)	118		477		673		563		159	
Prob > F	0		0		0		0		0	
Pseudo R^2	0.045		0.097		0.106		0.094		0.036	

Source: Own construction using CHIP, 2002.

Table A4. RIF decomposition of the ethnic gap in (log) income at different percentiles

	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	7.26		0.01	7.67		0.01	8.06		0.01	8.69		0.02
Minority	6.41		0.05	6.95		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		0.03	0.36		0.03	0.45		0.04
Explained	0.07	35.9	0.07	0.15	46.8	0.03	0.26	67.1	0.03	0.38	104.1	0.04	0.38	83.5	0.06
Unexplained	0.13	64.1	0.08	0.16	53.2	0.04	0.13	32.9	0.03	-0.01	-4.1	0.05	0.07	16.5	0.08
Explained															
Region	-0.11	-54.4	0.07	-0.06	-19.0	0.03	0.00	-0.7	0.03	0.08	22.6	0.03	0.17	38.1	0.05
Other geographical	0.13	64.0	0.05	0.13	40.7	0.02	0.17	45.1	0.02	0.20	54.1	0.03	0.11	23.9	0.03
N. of children	0.03	15.6	0.02	0.04	14.5	0.01	0.04	11.3	0.01	0.05	12.9	0.01	0.03	6.9	0.01
Education	0.00	-0.4	0.02	0.03	8.3	0.01	0.03	8.7	0.01	0.04	10.9	0.01	0.04	9.9	0.01
Skilled worker (hh head)	0.02	11.1	0.01	0.01	2.4	0.00	0.01	2.7	0.01	0.01	3.6	0.01	0.02	4.7	0.01
Unexplained															
Region	0.11	56.1	0.03	0.09	28.5	0.02	0.07	17.7	0.01	0.03	9.2	0.02	-0.01	-1.2	0.02
Other geographical	-0.01	-2.9	0.08	0.06	19.7	0.04	0.02	4.6	0.04	0.03	8.1	0.04	0.00	-0.1	0.04
N. of children	-0.01	-2.6	0.06	0.02	7.7	0.03	0.00	0.4	0.02	-0.01	-2.1	0.02	-0.07	-14.7	0.03
Education	0.23	114.3	0.16	-0.01	-3.7	0.08	-0.08	-21.9	0.08	-0.09	-24.8	0.10	-0.09	-20.3	0.12
Skilled worker (hh head)	0.05	26.0	0.03	-0.03	-10.4	0.03	-0.02	-5.3	0.03	-0.01	-2.9	0.04	-0.03	-7.5	0.07
Intercept	-0.25	-126.8	0.20	0.03	11.3	0.11	0.14	37.5	0.10	0.03	8.4	0.12	0.27	60.3	0.15

Source: Own construction using CHIP, 2002.