Preference for Redistribution and Inequality Perception in China: Evidence from the CGSS 2006

Zhou Xun

Working Papers / Documents de travail

Preference for Redistribution and Inequality Perception in China: Evidence from the CGSS 2006

Zhou Xun

WP 2015 - Nr 18

Preference for redistribution and inequality perception in China: Evidence from the CGSS 2006 *

Zhou XUN †

April 2015

Abstract

In this paper I investigate the conditional correlation between preference for redistribution and the perceived role of "circumstances" and "effort" using the Chinese General Social Survey. I found very significant correlations, thus validating the hypothesis of "sense of justice" for China. The migrant worker group who has dual identity (living/working in an urban area while being registered as a rural individual) is analysed in order to identify a discrimination effect (induced by the Chinese rural-urban segmentation policy) upon attitudes. However, being migrant is an endogenous variable to the attitude variables and the consistent estimate of this effect is much more important than the effect produced by a naïve estimate. The econometric model is a multivariate triangular LDV system with a binary endogenous explanatory variable estimated via a GHK simulator method. To implement the GHK calculations, I propose a parametric constraint to impose the positivity of the 3×3 correlation matrix. A generalisation for higher dimension cases is provided.

*During the writing of this paper, I have benefited from very useful conversations with Michel Lubrano, Habiba Djebbari, Emmanuel Flachaire and Gilles Stupfler. Of course, remaining errors are solely mine. This work has been carried out thanks to the support of the A*MIDEX project (N° ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR).

†Aix-Marseille Univ. (Aix-Marseille School of Economics), CNRS & EHESS, GREQAM, F-13002 Marseille, France

Keywords: Preference for redistribution, inequality perceptions, conditional correlation, Hukou and migrant worker, binary endogenous, GHK simulator.
JEL classification: C36, D19, H23, J18.

Acknowledgement

Data analyzed in this paper were collected by the research project "China General Social Survey(CGSS)" sponsored by the China Social Science Fundation. This research project was carried out by Department of Sociology, Renmin University of China & Social Science Division, Hong Kong Science and Technology University, and directed by Dr. Li Lulu & Dr. Bian Yanjie. The author appreciate the assistance in providing data by the institutes and individuals aforementioned. The views expressed herein are the author's own.

1 Introduction: Inequality in China

After nearly two decades of economic stagnation, China started market-oriented reforms in December 1978. Being the world's fastest growing economy over the last 30 years, growth came together with an inequality increase and the latter became one of the most important issues. The income distribution changed dramatically since the mid-1980s. According to the China Statistical Yearbook, the over-all Gini index has grown from 0.35 in 1990 to over 0.45 in 2006. Using the Chinese household nutrition survey, Chen and Cowell (2013) claim that in the post-millennium area, climbing on the income ladder has become more difficult.

The social reforms and the fast economic growth have affected different groups of people in different ways. The rural/urban gap has increased. The `Hukou` policy can be thought of being the cause of this widening gap. This state policy was adopted to limit mass migrations from the land to the cities, to ensure both economic and political social stability. The state favours urban residents and discriminates against rural residents in resource allocation, such as education, job vacancies, social benefits, health care, etc... (Afridi et al. 2012). Moreover, the redistribution scheme is decentralized in China as it mostly depends on the local economic level. This signifies that the subsistence allowances are much lower in rural areas than in urban areas (Houkai and Xiaoxia 2009). Thus being "rural" or being "urban" entails a huge discrepancy in terms of living standards and mostly in terms of opportunities. On the other hand, production activities in rural areas depend mostly on land farming although the labour-output ratio can vary a lot. Urban poor people have no access to means of production and they relies more on low paid job salaries and subsistence allowances. Rural individuals can decide to move to or to work in cities in order to have a relative higher income, but they will not necessarily be registered as "urban" officially. Although these rural labour forces contributed to a very important part in the economic development and urban modernization, the discrimination entailed by the `Hukou` system prevents them from having access to the benefits of the fruits of development in an equal way (Wong et al. 2007). The differences in rural and urban living and working circumstances and the isolation of rural people might lead to divergences in the perceptions of poverty between rural and urban people.

The urban group could be the beneficiary of growth in urban areas, but we also observe a significant group of the urban population who lives in a state of poverty (Fang et al. 2002). For example, in the early 90s, a great number of urban workers experienced the privatization of state owned enterprizes where they worked. Following this privatization, they became laid-off workers. This event consistently influenced their own life and that of their families. They dropped into the disadvantaged group.

The groups who benefited from most of these social and economic reforms are the government officers, businessmen, and people who have relations to them. They climbed up fast on the social ladder and distanced themselves from the pack by far. The increasing inequality becomes a potential risk for the stability of the society. The inter-regional inequality is also important. The east coastal provinces developed much faster than the interior provinces due to a series of preferential policies (Demurger et al. 2002). This divergence should contribute to the over-all inequality.

In the literature, we have plenty of evidence and discussions about inequality in China, see for instance, Kanbur and Zhang (1999), Khan and Riskin (1998), Khan and Riskin (2001), etc. Although, how people perceive inequality is less discussed in the literature. Some researchers have noticed the ten times increased number of mass protests in China from 1993 to 1995. This leads to the conclusion that Chinese ordinary people are angrier about the rising inequality (Tanner 2006). The Chinese government also notices this risk. During the term of office of the former leadership of Hu Jintao and Wen Jiabao, they put forward a slogan called "harmonious society" as well as series of policies to stabilize the society and to inhibit the dissatisfied. However, in a recent paper of Whyte et al. (2009), the author provides some evidence averse to these worries. By studying a national-wide inequality attitude survey, he concludes that the dissatisfaction for inequality has been overestimated in China. For example, the rural group is not more depressed by the inequality than the urban group as was assumed. According to this author, the rural group is more dissatisfied by procedural injustice rather than by distributive injustice. Moreover, unlike the urban group, the rural group has a limited perception of the real social ladder and people of that group tend to compare themselves to people who live in the same village. These conclusions come from the fact that rural individuals report the inequality within their community which is moderate. On the other hand, because the urban group has seen many upstart examples around them, they tend to be more depressed by inequality.

One question related to the inequality perceptions is the preference for redistribution, which is not a well developed topic in China. It also reflects people's inequality perception and adverts to how individuals perceive themselves as compared to others. Moreover, the topic of preference for redistribution is a natural experiment that bestows the possibility to survey many other topics, such as altruism and risk aversion. Lastly, does inequality perception influence the preference for redistribution? The correlation between preference for redistribution and inequality perception arises as an important issue in this study. Since the forming of the preference for redistribution is rather complex, many factors that determine the preference are unobservable, especially the value orientation and psychological traits. In Xu and Liu (2013), the authors show the importance of both social justice recognition and self-interest variables for explaining preference for redistribution. Unfortunately, this paper neglects the possibility of endogeny caused by the correlation between preference for redistribution and justice recognition.

The effect of `Hukou` policy is also less discussed in the preference for redistribution literature. As an national-wide policy that discriminates against the rural group, what would be the potential effect of such system upon attitudes? Is being a migrant worker an endogenous factor for attitude analysis? Wong et al. (2007) has shown the policy that discriminates towards migrant workers. Thus the decision of migration is not simply economically driven and rural individuals have to take into account many factors, which might be influenced by different perceptions and attitudes.

The aim of this paper is to provide some evidence concerning the forming of preference for redistribution and poverty perceptions in China and the relation between them. The effect of `Hukou` policy upon preference and perceptions is also discussed. This paper is organized as follows. Section 2 reviews the literature about preference for redistribution and some important theories in this domain. Section 3 introduces the data base and discusses the choice of the potential determinants of preference for redistribution. Section 4 discusses one of the key variables in the preference for redistribution study, which is the occupation prestige scale and its readjustment for the Chinese society. A trivariate ordered probit model is introduced and discussed in section 5. Empirical evidence are given in Section 6 with further discussions of "migrant workers" along with a more realistic specification up to a quadrivariate model that copes with the endogeneity problem. Section 7 concludes.

2 Preference for redistribution and perception of poverty: a literature review

The literature about preferences for redistribution started with the static model of Meltzer and Richard (1981), based on the median voter theory of Romer (1975). It assumes that if the median income is lower than the mean income and if the government does nothing more than taxing the richer group (above the mean income level) and redistributing the taxes to the poor, then they will be a majority of population who will vote for a higher tax rate.

This main result of the model of Meltzer and Richard (1981) can break down if we considerer a dynamic framework where the voter introduces his future income in his utility function. If it is so, then people who earn an income lower than the mean level today are not necessarily interested in a redistribution policy if they anticipate that they would climb up over the mean level tomorrow. This idea was formalized with the *prospect of upward mobility* (POUM hypothesis) of Benabou and Ok (2001). Within a dynamic framework, the median voter theory may no longer hold because individuals maximize their inter-temporal utility where their expected future income appears as an argument. They can vote against redistribution if their anticipation function is a concave function of their income.

The two models (static and dynamic) were tested on different data set. Even if it is now out of fashion, the static model of Meltzer and Richard (1981) was tested by Karabarbounis (2011) using the OECD SOCX data set over 14 countries. He found that more inequality was related to an increase in the demand for redistribution. However, people have become much more concerned the dynamic model. Using several data sets, economists and sociologists found proofs of the

"POUM" effect in majoritarian democratic societies. For instance, Alesina and La Ferrara (2005) used the GSS (General Social Survey) and the PSID to relate income dynamics to preference for redistribution in the US. They found a POUM effect. Using the BHPS, Clark and d'Angelo (2008) found the upward/downward mobility effect when analyzing intergenerational mobility.

Economically, the "POUM" hypothesis relies on the specification of a particular individual dynamic utility function where only self interest is at work. However, preferences could be impacted by ideology, culture and family traditions and not simply by income levels or by income expectations as detailed in Piketty (1995) or in Benabou and Tirole (2006). See also Neustadt (2011) and Scheve et al. (2006) for the effect of beliefs and religion.

Researchers have paid a lot of attention to the role of political ideology that generates differences between voters, differences based on issues such as equality, fairness, and the role of government, see Alesina and Glaeser (2004), Bean and Papadakis (1998), Feldman and Zaller (1992). Most of the discussions are around the relation between preference for redistribution and ideology as led around stylized facts. For example, Alesina and Glaeser (2004) try to explain why the EU society is more supportive of redistribution while the US society is much less supportive. They found that in US the majority tends to believe that poverty is generated by a lack of effort while the EU society tends to impute poverty to misfortune. The author points out that the link between preference for redistribution and beliefs about the nature of poverty relies on the `sense of justice`: "if you believe that luck (or inherited wealth) determines differences in income, you are more favorable to redistribution. If you believe that individuals' effort and individual's ability determine income, you are less favorable to redistribution" (Alesina and Angeletos 2005). In other words, it is a common sense that people should hold responsibilities for factors which are under their own control (i.e. lack of efforts) while they hold no responsibilities to external factors which are out of their control (i.e. circumstances), see for instance Rawls (1971) and Sen (1980) and Sen (1999).

Of course, differences in perception for the role of effort, luck/misfortune and preference for redistribution go back to long lasting historical and cultural differences between the two sides of the Atlantic. Alesina and Glaeser (2004) point out that the correlations are strong and *provocative*. This indicate that we cannot explain preference for redistribution as a function of poverty perception, but that these variables have to be explained simultaneously. At the individual level, many factors including one's life experiences, family background, psychological traits, social attitudes, ethics and the world outlook are usually unobservable. However, people who believe that luck determines success might still know the importance of effort. Alesina and Angeletos (2005) further point out that recognitions of luck and effort can be influenced by economic and political policies while redistribution policies are the revealed preference for redistribution of the majority in a democratic society. Fong (2001) emphasises that poverty perceptions are correlated with self-interest variables which leads to the question: what are the determinants of poverty perceptions? The causality of preference and perceptions is complex and simultaneous, while it is possible to quantitatively measure the correlations among these factors conditioned on exogenous variables.

In a recent paper of Xu and Liu (2013), authors enter self-interest variables and "key to success" variables jointly in the preference for redistribution equation. However, they assume that "key to sucess" variables are exogenous. Ignoring the simultaneity of the system would entail a serious endogenous problem. Although their results provide some evidence concerning the `sense of justice`, the problem is now how to determine the relation between preference for redistribution and the poverty perception in an efficient way. We shall provide a specific econometric model for that.

3 The Chinese General Social Survey

The Chinese General Social Survey (CGSS) is an annual or biannual repeated cross-section survey designed to collect individual opinions on social trends and the changing relationship between social structure and quality of life in China. CGSS is a sub-project of the International Social Survey Programme (ISSP). Following the structure of the famous general social survey (GSS), the CGSS provides multi-dimensional information on both socio-economic characteristics, attitudes and values on social issues. For the same reason as for the GSS, the respondents of each survey wave are randomly selected so that they cannot be supposed to be followed repeatedly so as to avoid selection bias and so as to ensure that the sample is representative of the whole population in each wave. The first wave was collected in 2003 and the last wave in 2010. The first wave provides very limited information while the social value part of the 2010 wave is not yet published. In this paper, we choose the 2006 wave because it contains the richest information available on social values. In this wave, 28 provinces are included, including Beijing, Shanghai, and some of the other most developed direct-controlled municipalities. This makes a total of 9 517 observations.

The 2006 CGSS is organized in four parts: The individual socio-demographics characteristics, the occupation status, the household components and status, and most importantly for us the subjective attitude variables.

3.1 Exogenous variables

Two types of explanatory variables are considered:

- *Socio-demographic variables:* region, gender, birth cohort, party membership, religion belief, material status, rural/urban status and years of education.
- *Individual socio-economic variables:* income, occupation prestige, occupation mobility with respect to that of the parents, and subjective expectation of household future socio-economic status.

There are 28 provinces included in the data set. We regroup them into three regions:

- E.C. China (41.7%): East coast of China. The most developed provinces and the big cities of China (including Beijing, Shanghai and Shenzhen) and the three northeast provinces. This region has the most developed industries and the most developed third sector.
- C. China (26.5%): Central China. Less developed than the E.C, including the traditional agriculture provinces (Henan, Hunan, Hubei, etc).
- W. China (31.8%): The west of China, the least developed region.

The weighted sample proportions are given in parentheses. Descriptive statistics are given in Table 1.

The income variable includes all sources of individual income received in the year 2005 (currency unit: RMB). The summary table is shown in Table 2. There are 805 income missing observations and 1 003 observations with a zero income.

A subjective measure of the future upward mobility is also considered, the self-reported question: *How do you perceive your future household financial situation in three yeas ahead*, is it *better* (coded as 1) or not (coded as 0, includes *the same* and *worse*).

Table 1: Socio-economic descriptive statistics
using individual weights

Gender	
female	0.516
male	0.484
Birth cohorts	
-1959	0.162
1960-1979	0.462
1980-	0.376
Party membership	
member	0.152
mass	0.846
Religious beliefs	
Believer	0.137
atheist	0.846
Living in a couple	
yes	0.866
no	0.134
"Rural" in 2005	
yes	0.619
no	0.381
(Rural)"Migrant worker" in 2005	
yes	0.060
no	0.940
Being "New urban" (change within 10 years)	
yes	0.046
no	0.954
Years of education	
1st Qu.	6
Median	9
average	9.1
3rd Qu.	12

Table 2: Annual Income distribution in Chinese currency

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Income distribution	0	2 000	5 000	8 824	11 300	250 000

3.2 Dependent variables: Social values and opinions

In the attitude part of the survey, respondents are requested to report their opinions on a four-level scale tracing agreement to a given proposition (1 for totally disagree and 4 for totally agree). We have selected the three following questions:

1. Government should tax the rich more to help the poor.
2. Individuals are poor because society is not well functioning, especially because of misgoverning.
3. Individuals are poor because they are lazy

Descriptive statistics for social values and opinions are given in Table 3. A first glance at this Table gives us the impression that these are

Table 3: Attitude and perceptions

	1	2	3	4
Redis. Pref.	0.030	0.173	0.528	0.275
Poor. misgov	0.025	0.206	0.588	0.181
Poor. lazy	0.201	0.467	0.265	0.078

very similar questions. The distribution of preferences for redistribution is roughly the same as that of the *poor.misgov* variable. The *poor.lazy* variable is distributed just as the complementary distribution of the above two variables. This again confirms the warning of potential endogeneity problems. The ideology variables and preference for redistribution are jointly determined by unobserved factors such as value orientations and character traits.

4 Occupation and Social mobility

Clark and d'Angelo (2008) have shown that the inter-generational prestige mobility between father and son have a significant effect upon the preference for redistribution. In that paper, job prestige is measured by the Hope-Goldthorpe Scale(HGS)[1] which transforms the

[1] The HGS is an occupational index that reflects the job's reputation and classifies jobs according to their social desirability.

occupation norms over a continuous scale. A precise measure of the occupation prestige should reflect the social ranking and the social class of individuals which are correlated to their preferences and perceptions.

In China, the `Hukou` system also contributes to the differences in occupations as well as to the occupation mobility. The migration from rural to urban regions is constrained which would limit the opportunities (discrimination) of "rural" people. Consequently, the `Hukou` system may have far-reaching influence upon preferences and perceptions through their life-cycle experiences and occupations.

In the CGSS, the individual's current job occupation and father's occupation are coded using the usual Erikson, Goldthorpe, and Portocarero (EGP) classification. This corresponds to:

1. Category I (40%): farm labor
2. Category II (26%): skilled/unskilled worker
3. Category III (12%): self employed
4. Category IV (10%): lower sales-service/routine non-manual
5. Category V (11%): higer/lower controllers

The EGP classes are ranked on the basis of two dimensions: Employee monitoring difficulties and human asset specificity (required on the job training), see for instance Edlund (2008). Both the HGS and EGP scales are designed for measurement purpose in western societies. It is not evident that the ordering entailed by the EGP classification is well adapted for China. In the following subsections, we will discuss the properties of inter-generational mobility in China and its consequences. Based on the following discussion, we will show that the EGP occupation prestige scale is not suitable for China. As a consequence, we have to build a proper measure of occupation prestige and the corresponding occupation mobility instead of using directly the EGP classification.

4.1 Markov Inter-generational mobility

Like any other occupation variable, the EGP classification provides not only the occupation categories but also their corresponding prestige ranking. If the assumed prestige ranking does not adapt properly to China, it provides a misleading information for analyzing the inter-generational mobility and thus the preference for redistribution. The occupation mobility is one way to verify the validity of the assumed prestige ranking because the mobility monotonicity property holds if and only if the prestige ranking is monotone increasing.

As we have this classification both for the respondent and his/her father (current occupation or before retirement), we can model inter-generational mobility. Using these five ordered categories, we estimate a weighted Markov transition matrix to model an inter-generational transition matrix[2] which is reported in Table 4. We see that mobility

Table 4: Inter-generational mobility, Prais Index = 0.774

	EGP:1	EGP:2	EGP:3	EGP:4	EGP:5
EGP:1	0.54	0.23	0.09	0.07	0.07
EGP:2	0.09	0.44	0.08	0.23	0.15
EGP:3	0.17	0.19	0.31	0.25	0.08
EGP:4	0.06	0.28	0.14	0.32	0.19
EGP:5	0.11	0.25	0.13	0.21	0.29

is lower in lower rows and that the first row is the most sticky one. The first row corresponds to the mobility probabilities of individuals having a father in the farm labor category. Implicitly, this refers to the mobility of individuals who originally come from rural areas which indicates a significant policy barrier effect brought by the `Hukou` system. When his father was working in a farm, an individual has a probability of 0.54 to also work in a farm. For all the other categories, the probability for an individual to occupy the same type of job as his father is much lower.

4.2 Occupation mobility

However, the commonly applied EGP prestige scale might not fit the Chinese social ladder as it is designed for Western societies. In this subsection, I use the stereotype ordered regression (SOR) model[3] to revise the ordering of the EGP scale. We want to explain the social class segmentation determined by the EGP scale by observable control variables in order to estimate the implicit ranking of the social occupations that individuals have in mind. The model is given as

[2]The transition matrix is estimated as:

$$P_{jk} = n*_{jk}/n*_{j.},$$

where P_{jk} refers to the probability of moving from origin category j to destination category k. $n*_{jk}$ is the weighted frequency of observations that move from j to k and $n*_{j.}$ refers to the weighted frequency of all the observations that origin from j.

[3]The "stereotype ordered regression" (SOR) is reported for instance in Anderson (1984). See Hendrickx (2000) for an implementation in `Stata`.

follows:

$$s_{ki}^* = \alpha_k + \phi_k * X_i\beta + \epsilon_{ki}. \tag{1}$$

The s_{ki}^* corresponds to the latent score (propensity) of category k for individual i while the α_k refers to the category specific intercepts. X_i is a set of observed variables that controls for the human capital (years of education), basic demographic variables (birth cohort, gender and rural/urban). Similarly to the unconditional Markov chain analysis of previous subsection, we also include the occupation category of the father in order to introduce inter-generational mobility. ϵ_{ki} is the error term which follows an extreme value distribution. The probability is then delivered through a logit type link function.

Unlike the linear part of the standard multinomial logit model, a SOR model constrains the category specific linear parameters β_k to be the same over all the categories. So $\beta_k = \beta$ whatever the value of k. However at the same time, a new multiplicative variable ϕ_k is introduced which relaxes in a way that restriction. It serves to measure the ordinal scale of the destination category ladder while this is not considered in the standard model. For identification reasons, we have to impose $\phi_1 = 0$ and $\phi_5 = 1$. To understand the new scaling metric parameter ϕ_k, we can write the log odds ratio of the two event probabilities $P(y_i = k)$ versus $P(y_i = k')$ as:

$$\log\left[\frac{P(y_i = k)}{P(y_i = k')}\right] = \alpha_k - \alpha_{k'} + (\phi_k - \phi_{k'})X_i\beta. \tag{2}$$

We shall see that the explanatory variable effects are measured by multiplying the category constant parameter β by an estimated category scaling metric ϕ_k. The higher the distance between ϕ_k and $\phi_{k'}$, the higher the magnitude of the effect given by X. In order to estimate the scaling parameter ϕ_k and the linear parameter β, an iterative method is used.[4] Table 5 reports the estimate of the intergenerational mobility (father's occupation versus current occupation of respondent). We shall see that the scaling metric ϕ_k is not monotone increasing. The scale of category III (skilled/unskilled worker) is higher than that of category II (self-employed). This result is comparable to the finding of Wu (2007). In western societies, category III is ranked higher than category II, see for instance Ganzeboom et al. (1989). One possible explanation given by Wu (2007) is that Chinese society (or more

[4]The estimation procedure is as follow: first take the ϕ_k scaling metric as given and estimate β, then take the estimated β as fixed and estimate ϕ_k. The standard errors of ϕ_k are not identified while the standard errors of β are conditional on the given scaling metric ϕ_k. For more details see Hendrickx (2000). The estimation is achieved by the "mclgen" and "mclest" commands in the software `Stata`.

generally all communist societies) has a long tradition to inhibit private property. And eventually, becoming self-employed is easier than finding a stable job in the administration. Self-employment is less preferred in China and only concerns farm labor. The highest gap occurs between the first place (category I) to the second (category III) on the ladder, after that the differences are much smaller. This result shows an extreme low prestige of farm labor. Clearly, this occupation is found only in the "rural" group. The "SOR effect" reported in Table 5 are the estimates of β. We see that higher human capital is associated with higher probability of upward mobility. Being rural reduces the upward mobility probability while it could be weaken by the birth cohort effect (being born after 1980). Individuals are more likely to have a decent occupation if their fathers' occupation prestige is higher. We should also notice the stickiness of farm labor category. Lower categories have important negative effects upon the mobility for the next generation, which might be due to the policy barrier made by the `Hukou` system. From this analysis we see that the ordering of the EGP scale is not suitable for China. Henceforward, we avoid inserting directly the EGP category information. Instead, we use the transformed scaling metric parameter ϕ_k because it corrects the order of the occupation category with information of the scaling metric of each category. A dummy variable of inter-generational upward mobility is then coded as 1 if the following mobility event $\mathbb{1}\{s_{father} = j, s_{son} = k\}$ between two generations satisfies the condition that $\phi_k > \phi_j$.

5 A theoretical econometric model for attitude variables

We have three opinion variables that correspond to discrete observations which might be correlated. To each of these m opinion variables corresponds a level of unobserved utility z_m. This level, for every $m \in \{1, 2, 3\}$ is explained by a linear combination of exogenous variables X so that:

$$z_m = X'\beta_m + \epsilon_m.$$

The observation rule, relating the unobserved utility level z_m to the response variable Y_m is

$$Y_m = k \text{ if } \tau_{m,k-1} < z_m < \tau_{m,k}.$$

Y_m is the observed category vector which is reported in the survey, taking the ordered values from 1 to 4 in our case. $\tau_{m,k}$ is the threshold parameter that locates the boundaries for the discrete responses over

Table 5: Inter-generational occupation mobility the SOR model

Scaling metrics	
EGP:1	0.0000
EGP:2	0.6201
EGP:3	0.5865
EGP:4	0.8481
EGP:5	1.0000
intercepts	
EGP:2	0.454 (0.137)
EGP:3	−0.490 (0.133)
EGP:4	−0.608 (0.184)
EGP:5	−1.129 (0.214)
SOR effect	
Father EGP:2:	0.321 (0.143)
Father EGP:3:	−0.506 (0.256)
Father EGP:4:	0.821 (0.246)
Father EGP:5:	1.109 (0.196)
Cohort 60-70	−0.328 (0.126)
Cohort post. 80	0.456 (0.134)
female	−0.357 (0.080)
yeduc	0.361 (0.015)
Rural	−4.153 (0.129)
$Pseudo - R^2$	0.234
N	8007

the support of the continuous latent utility variable z_m.The error term ϵ_m is supposed to be normal of zero mean. For identification reasons, the three variances are set equal to 1. The three ordered probit models can be estimated separately, if the three error terms are uncorrelated. In this case the probability of a basic event is equal to:

$$\Pr(Y_m = k) = \Phi(\tau_{m,k} - X'\beta_m) - \Phi(\tau_{m,k-1} - X'\beta_m).$$

If the m error terms are correlated, the basic event is much more complex and the three ordered probit models have to be estimated jointly. This model is related to the multivariate probit model (see e.g. Cappellari and Jenkins 2003). But here of course the dependent variables are ordered and not just binary.

5.1 A trivariate ordered probit model

As a starting point, let us consider the distribution of the error term:

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \sim N(0, \Sigma) \tag{3}$$

where we have 1's on the diagonal of the symmetric covariance matrix Σ. If the off-diagonal elements (ρ_{mn}) are all 0, then the model reduces to three independent ordered probit models. In order that Σ be positive definite symmetric, the elements of ρ in

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix} \tag{4}$$

must verify some constraints, see next subsection.

Let us now consider the probability of the trivariate event ($Y_1 = j, Y_2 = k, Y_3 = l$). The evaluation of this probability requires the evaluation of a trivariate Gaussian CDF:

$$\Pr[Y_1 = j, Y_2 = k, Y_3 = l] = \int_{\tau_{1,j-1}-\hat{z}_1}^{\tau_{1,j}-\hat{z}_1} \int_{\tau_{2,k-1}-\hat{z}_2}^{\tau_{2,k}-\hat{z}_2} \int_{\tau_{3,l-1}-\hat{z}_3}^{\tau_{3,l}-\hat{z}_3} \phi_3(\epsilon_1, \epsilon_2, \epsilon_3, \rho) \mathrm{d}\epsilon_1 \mathrm{d}\epsilon_2 \mathrm{d}\epsilon_3,$$

where $\hat{z}_1, \hat{z}_2$ and $\hat{z}_3$ are the linear predictors $X'\hat{\beta}_m$ ($m = 1, 2, 3$), ϕ_3 is the PDF of a trivariate normal distribution and ρ represents the vector of all correlation parameters. There are good numerical methods for evaluating a bivariate normal CDF that are included in standard packages. But for higher dimensions, simulation methods are usually preferred. In our case, because of the truncation problem, the GHK (Geweke-Hajivassiliou-Keane) simulator seems to be a good candidate because the truncations could be directly simulated.

In order to apply the GHK simulator, let us rewrite the previous event probability as a product of conditional and marginal probabilities:

$$\begin{aligned}
&\Pr(Y_1 = j, Y_2 = k, Y_3 = l) = \\
&\Pr(Y_1 = j) \times \Pr(Y_2 = k | Y_1 = j) \times \Pr(Y_3 = l | Y_1 = j, Y_2 = k) = \\
&\qquad\qquad \Pr(\tau_{1,j-1} < \hat{z}_1 + \epsilon_1 < \tau_{1,j}) \times \\
&\qquad\qquad \Pr(\tau_{2,k-1} < \hat{z}_2 + \epsilon_2 < \tau_{2,k} | \tau_{1,j-1} < \hat{z}_1 + \epsilon_1 < \tau_{1,j}) \times \\
&\Pr(\tau_{3,l-1} < \hat{z}_3 + \epsilon_3 < \tau_{3,l} | \tau_{1,j-1} < \hat{z}_1 + \epsilon_1 < \tau_{1,j}, \tau_{2,k-1} < \hat{z}_2 + \epsilon_2 < \tau_{2,k}).
\end{aligned} \tag{5}$$

The difficulty comes from the fact that the ϵ_m are correlated. Let A be the lower triangular Cholesky decomposition of Σ such that $AA' = \Sigma$. Let us introduce three *iid* standard normal random variables η_m so that we can express the ϵ_m as a linear combination of the three independent η_m:

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \tag{6}$$

or in an expanded notation:

$$\begin{aligned}
\epsilon_1 &= a_{11}\eta_1, \\
\epsilon_2 &= a_{21}\eta_1 + a_{22}\eta_2, \\
\epsilon_3 &= a_{31}\eta_1 + a_{32}\eta_2 + a_{33}\eta_3.
\end{aligned}$$

Following this triangular system, we can decompose the joint probability (5) into the product of three conditional independent Gaussian probabilities. The first marginal probability is defined as:

$$\Pr(\tau_{1,j-1} < \hat{z}_1 + a_{11}\eta_1 < \tau_{1,j}) = \Phi[\frac{\tau_{1,j} - \hat{z}_1}{a_{11}}] - \Phi[\frac{\tau_{1,j-1} - \hat{z}_1}{a_{11}}], \tag{7}$$

and can be evaluated directly because $\Phi(.)$ is the CDF of η_1. The second conditional probability is:

$$\begin{aligned}
&\Pr(\tau_{2,k-1} < \hat{z}_2 + a_{21}\eta_1 + a_{22}\eta_2 < \tau_{2,k} | \tau_{1,j-1} < \hat{z}_1 + a_{11}\eta_1 < \tau_{1,j}) \\
&\qquad = \Phi[\frac{\tau_{2,k} - \hat{z}_2 - a_{21}\eta_1}{a_{22}}] - \Phi[\frac{\tau_{2,k-1} - \hat{z}_2 - a_{21}\eta_1}{a_{22}}],
\end{aligned}$$

where $\Phi(.)$ is the CDF of η_2. The third conditional probability is:

$$\begin{aligned}
&\Pr(\tau_{3,l-1} < \hat{z}_3 + a_{31}\eta_1 + a_{32}\eta_2 + a_{33}\eta_3 < \tau_{3,l} \\
&\quad | \tau_{2,k-1} < \hat{z}_2 + a_{21}\eta_1 + a_{22}\eta_2 < \tau_{2,k}; \tau_{1,j-1} < \hat{z}_1 + a_{11}\eta_1 < \tau_{1,j}) \\
&= \Phi[\frac{\tau_{3,l} - \hat{z}_3 - a_{31}\eta_1 - a_{32}\eta_2}{a_{33}}] - \Phi[\frac{\tau_{3,l-1} - \hat{z}_3 - a_{31}\eta_1 - a_{32}\eta_2}{a_{33}}],
\end{aligned} \tag{8}$$

where $\Phi(.)$ is the CDF of η_3. The first marginal probability can be evaluated directly, using a standard numerical routine for Gaussian CDFs. The second probability is conditional on the distribution of η_1, which is unobserved. The idea of the GHK algorithm is to replace η_1 by a random draw from a truncated Gaussian distribution in order to evaluate the probability of a basic event and write the likelihood function. Of course, several draws have to be made as we shall detail below. Let us call η_1^{*r} the r^{th} draw of η_1 so that we have now:

$$\Phi[\frac{\tau_{2,k} - \hat{z}_2 - a_{21}\eta_1^{*r}}{a_{22}}] - \Phi[\frac{\tau_{2,k-1} - \hat{z}_2 - a_{21}\eta_1^{*r}}{a_{22}}], \tag{9}$$

where η_1^{*r} comes from a truncated standard normal density with lower and upper truncation points equal to $(\tau_{1,j-1} - \hat{z}_1)/a_{11}$ and $(\tau_{1,j} - \hat{z}_1)/a_{11}$ respectively. The third conditional probability includes two Gaussian random variables, the same η_1 as before and η_2. We use the same η_1^{*r} as before and draw η_2 from a truncated Gaussian so as to have:

$$\Phi[\frac{\tau_{3,l} - \hat{z}_3 - a_{31}\eta_1^{*r} - a_{32}\eta_2^{*r}}{a_{33}}] - \Phi[\frac{\tau_{3,l-1} - \hat{z}_3 - a_{31}\eta_1^{*r} - a_{32}\eta_2^{*r}}{a_{33}}]. \tag{10}$$

This time, η_2^{*r} is drawn from a standard normal density with lower and upper truncation points $(\tau_{2,k-1} - \hat{z}_2 - a_{21}\eta_1^{*r})/a_{22}$ and $(\tau_{2,k} - \hat{z}_2 - a_{21}\eta_1^{*r})/a_{22}$. We explain in Appendix C how to draw truncated random numbers using GHK algorithm.

Since the computation of Equation (7) is straightforward, we shall initialize the algorithm by computing it first and then recursively evaluating Equation (9) and (10). Now if we have R draws of η_1^* and η_2^*, the simulated probability is then the arithmetic mean of each probability given the r^{th} random draw of ξ^r (see Appendix C):

$$\overline{\Pr}(Y_1 = j, Y_2 = k, Y_3 = l)_{GHK} = \frac{1}{R}\sum_{r=1}^{R}[\Pr_1 \times \Pr_2^r \times \Pr_3^r]$$

where $\Pr_2^r$, $\Pr_3^r$ refer to Equations (9) and (10) respectively given r^{th} draw of ξ. Finally, the simulated likelihood function is given by:

$$L_{GHK} = \prod_{i=1}^{N} \overline{\Pr}(y_{i,m} = k)_{GHK}$$

for $m = \{1, 2, 3\}$ and $k = \{1, 2, 3, 4\}$ and $Y_m = \{y_{1m}, \cdots, y_{im}, \cdots, y_{Nm}\}$. The weighted likelihood function is:

$$WL_{GHK} = \prod_{i=1}^{N} \overline{\Pr}(y_{i,m} = k)_{GHK}^{w_i}, \tag{11}$$

where w_i is the weight value assigned to individual i as our data set is a weighted sample.

5.2 Monte Carlo simulation

Now let's consider a Monte Carlo simulation example in order to verify that our method is working correctly. We have selected a sample size of 1 000 and a number of replications equal to 1 000. We first draw the three independent explanatory variables X_1, X_2 and X_3 from a standard normal distribution with mean zero and standard deviation 1.5. Once we have the X, we select values for the βs so as to generate the latent utilities. We have selected the following structure:

$$\begin{aligned} z_1 &= 0.3 * X_1 - 0.6 * X_2 + 0.9 * X_3 + \epsilon_1, \\ z_2 &= 0.2 * X_1 - 0.3 * X_2 + 0.6 * X_3 + \epsilon_2, \\ z_3 &= -0.2 * X_1 + 0.9 * X_2 + 1.5 * X_3 + \epsilon_3. \end{aligned}$$

The coefficients of X in each equation are chosen arbitrarily. The error terms ϵ_1, ϵ_2, ϵ_3 are simulated from a trivariate normal distribution with zero mean and covariance matrix:

$$\text{Cov}\begin{pmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{pmatrix} = \begin{pmatrix} 1 & 0.25 & -0.4 \\ 0.25 & 1 & 0.6 \\ -0.4 & 0.6 & 1 \end{pmatrix}. \tag{12}$$

The threshold parameters are chosen so as to correspond to the (0.25, 0.50, 0.75) quantiles of the $\hat{z}_m$. The ordinal responses $Y_{\{1,2,3\}}$ are then generated accordingly.

We report in Table 6 the mean bias when Σ is treated as an identity matrix, and the mean bias when Σ is estimated using the GHK simulator. Table 7 reports the same results for the MSE. It is well

Table 6: Mean bias for comparing three independent ordered probit with a trivariate ordered probit

	Independent ordered probits			Trivariate ordered probit		
	y_1	y_2	y_3	y_1	y_2	y_3
X_1	-0.219	-0.138	0.154	-0.000	0.000	-0.000
X_2	0.443	0.210	-0.697	0.002	0.001	0.003
X_3	-0.665	-0.419	-1.167	-0.004	-0.001	0.003
ρ_{21}		0.250			0.000	
ρ_{31}		-0.400			-0.005	
ρ_{32}		0.600			0.006	

apparent from these tables, first that when there is correlation, a serious error is committed if the model is treated as three independent

Table 7: MSE for comparing three independent ordered probit with a trivariate ordered probit

	Independent ordered probits			Trivariate ordered probit		
	y_1	y_2	y_3	y_1	y_2	y_3
X_1	0.050	0.021	0.026	0.0007	0.0006	0.0007
X_2	0.199	0.046	0.492	0.0010	0.0007	0.0023
X_3	0.447	0.178	1.374	0.0014	0.0009	0.0050
ρ_{21}		0.063			0.0016	
ρ_{31}		0.160			0.0018	
ρ_{32}		0.360			0.0015	

ordered probit models. And second, which is in a way the most important result, that our method using the GHK simulator managed to give quite accurate results as well as the access to the correlations.

5.3 Evaluation strategy

The trivariate ordered probit model was programmed using the software **R**. The basic idea is to maximize the simulated likelihood function, this is done via the package "maxLik" using the "BHHH" algorithm in **R**. In order to initialize the evaluation, a reasonable set of starting values has to be provided. In this study, the staring values are chosen from the coefficient estimated from the independent ordered probit models while the starting values for the correlations are set equal to zero. As we have discussed already, several constraints have to be imposed to ensure that the model is identifiable, e.g. the variance of error terms have been normalized to 1 in this paper.

Another thing which is important in the evaluation of the model is how to ensure the positive definite property of the variance covariance matrix Σ. If ρ_{21} and ρ_{31} are freely chosen between 0 and 1, then the third term ρ_{32} must verify the constraint:

$$\rho_{21}\rho_{31} - \sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{31}^2} \leq \rho_{32} \leq \rho_{21}\rho_{31} + \sqrt{1-\rho_{21}^2}\sqrt{1-\rho_{31}^2}. \quad (13)$$

A proof of this result is given in Appendix B. This condition is essential to improve the efficiency of the evaluation along with the iterations of the MLE process, or try-check (accept/reject) sampling strategy could be used which is quite slow especially when the correlation dimension is more than 2. This correlation condition does not seem to be commonly applied in most of the statistical softwares and packages (e.g. the try-check strategy has been used in command "triprobit" in software **STATA** without penalty in the likelihood function).

However this strategy also has limitations when the dimension is higher than three thus we need to find another way of constructing Σ. We shall discuss this later.

6 Empirical results for preference for redistribution in China

Table 8 reports estimation results for the trivariate ordered probit model (11) which was discussed in the previous section and which is meant to explain the answers given to the three ordinal variables depicting preference for redistribution and poverty perceptions. We first present the unrestricted version of the model. If we set all the individually insignificant coefficients to zero (lower than a 90% significance level), the log-likelihood value drops from -17 834 to -17 838.92, with DF of 20, so the overall restrictions are not rejected. If we now try to restrict to zero the structural variances (equivalent to considering three independent ordered probit models), the likelihood value drops to -17929.04. The difference is 95.04 with DF of 3, so that this null restriction is rejected. It is necessary to consider a joint model.

6.1 Structural correlation

Our model provides an efficient way of estimating a correlation matrix among ordinal variables, conditionally on exogenous variables. Non zero correlation implies that the three variables are mutually endogenous. Table 8 shows that these correlations are not important in magnitude, but are highly significant. More precisely, correlation is very significant between the *preference for redistribution* variable and the two other variables. However, it is not significant between the *poor-misgovernment* and the *poor-lazy* variables. This means that these two variables provide independent information on *preference for redistribution*. Remember that Alesina and Glaeser (2004), when comparing subjective poverty perception in Europe and in the US, were considering two exclusive justifications: lack of effort in the US and absence of luck in Europe. In China the two types of explanation can play a complementary role, and that at the same time in the same mind.

The correlation between *preference for redistribution* and *poor-misgovernment* is positive. This means that the variables explaining the opinion about mis-governance will have also an indirect influence upon the preference for redistribution with presumably the same sign. This means also that individuals thinking that mis-governance is a cause

Table 8: Preference for redistribution
and poverty perception, a trivariate estimation

	Redis. Pref.	Poor-misgov	Poor-lazy
C. China	0.019 (0.033)	-0.077^{*} (0.033)	0.080^{*} (0.032)
W. China	0.006 (0.029)	-0.108^{***} (0.029)	0.146^{***} (0.029)
birth 60-79	$0.081^{\cdot}$ (0.043)	0.118^{**} (0.042)	-0.090^{*} (0.043)
birth post 80	0.111^{*} (0.046)	0.115^{*} (0.045)	-0.119^{**} (0.045)
female	$0.065^{\cdot}$ (0.033)	-0.043 (0.035)	-0.096^{**} (0.033)
party	-0.036 (0.036)	-0.086^{*} (0.034)	-0.053 (0.035)
coupled	-0.001 (0.039)	-0.037 (0.040)	0.026 (0.038)
yeduc	-0.005 (0.005)	-0.017^{***} (0.004)	0.004 (0.004)
ln income	0.104^{***} (0.020)	0.075^{***} (0.019)	-0.008 (0.019)
ln income squared	-0.009^{***} (0.002)	-0.007^{***} (0.002)	0.000 (0.002)
Occup. prestige	0.069 (0.069)	-0.023 (0.069)	0.031 (0.066)
Upward (fath./son)	-0.131^{**} (0.047)	0.014 (0.048)	-0.118^{**} (0.045)
Upward (moth./dau.)	-0.132^{*} (0.052)	$0.099^{\cdot}$ (0.053)	-0.024 (0.050)
better finance	0.003 (0.026)	0.028 (0.026)	0.164^{***} (0.026)
Rural	-0.105^{*} (0.041)	-0.086^{*} (0.040)	0.134^{***} (0.039)
migrant worker	0.187^{**} (0.067)	-0.046 (0.076)	$-0.112^{\cdot}$ (0.060)
1—2	-1.787^{***} (0.093)	-2.123^{***} (0.091)	-0.689^{***} (0.089)
(2—3)-(1—2)	1.075^{***} (0.029)	1.261^{***} (0.032)	1.265^{***} (0.017)
(3—4)-(2—3)	1.491^{***} (0.018)	1.687^{***} (0.019)	1.001^{***} (0.019)
$\rho_{R.P,misgov}$		0.202^{***} (0.016)	
$\rho_{R.P,lazy}$		-0.086^{***} (0.016)	
$\rho_{misgov,lazy}$		0.013 (0.016)	
N		5138	
Loglik		-17 834	
R		25	

The two Upward dummy variables are 1 if the prestige of a son is higher than his father's, the same for the variable measuring the upward mobility of a daughter compared to her mother. They are built using the revised order of occupation categories provided by the SOR model.
p-value codes: "***" for 0.001, "**" for 0.01, "*" for 0.05 and "." for 0.1.

for poverty would tend to believe that poverty has to be compensated by redistribution.

The correlation between *preference for redistribution* and *poor-lazy* is negative and significant. However, this correlation is much lower. So individuals thinking that the main cause of poverty is laziness are also less in favour of redistribution. However, because of the smaller correlation, that effect is less important than the previous one. Consequently the `sense of justice` detailed and explained in Alesina and Glaeser (2004) is justified in China as the impact of *poor-misgovernment* is more important than that of the *poor-lazy* for explaining the preference for redistribution.

The final consequence of the presence of significant correlations is that even if a variable does not appear in the equation explaining preference for redistribution, it can have an indirect effect provided it appears significantly in one of the other equations.

6.2 Poverty perceptions

The estimates of the two equations corresponding to beliefs in the causes of poverty are reported in column 2 and 3 of Table 8. For these two equations, both regions and birth cohorts are significant. The estimated sign of these two sets of dummy variables in these two equations shows an exclusive pattern. Generally speaking, individuals living in the Central or in the Western part of China (compared to the Eastern Coast region) support less the idea that poverty is generated by misgovernment while they tend to support the idea that laziness is the main cause of poverty. New generations support more the idea that poverty is caused by misgovernment rather than by laziness. The gender dummy variable enters as a significant factor only in the *poor-lazy* equation. Females tend to impute poverty to laziness more than males. The negative sign of the party membership dummy variable is expected in the *poor-misgov* equation. Party members tend to be more confident in the ability of the Party to fight against poverty. While being a member of the Party makes no significant differences in answering the *poor-lazy* question. The strong effect of party membership in the *poor-misgov* equation is then an evidence that a self-interest variable can influence poverty perception.

Living in a couple has no effect in both cases. Having more years of education weaken the recognition that misgovernment causes poverty while it has no effect in the *poor-lazy* equation. The effect of log-income has an inverted U-shape. But as the first 25% quantile level of log-income (7.601) locates on the right hand of the curve peak, the effect of log-income has then only a monotone decreasing trend with an increasing speed. Occupation prestige has no effect in both equations. Having an upward mobility experience (compared to the father) drifts negatively the recognition that idleness causes poverty. Lastly, people who anticipate an upward household financial situation agree more that laziness causes poverty.

Rural individuals tend to impute poverty less to misgovernment while they impute it more to laziness. This result might not be coherent with the evidence found in section 4.2 about occupation mobility. Farm labour (in the rural group) is the most static category with the lowest prestige. Implicitly, being rural reduces upward mobility opportunities a lot, compared to the urban group. According to the literature, people having a low upward mobility prospect should be more in favour of redistribution. However, it is just the reverse here. This result is comparable to that of the Whyte et al. (2009). We may guess that it is because the return of physical efforts in farm labour is more clearly perceived than the return of effort in the urban group. Rural people tend to believe that poverty is a direct indicator of a lack of effort. Moreover, as argued in Whyte et al. (2009), compared to the urban group, the rural group is a relatively more closed society with much fewer upstart examples so that people living there do not perceive an important level of within-group inequality. The different redistribution schemes applied in rural and urban areas also entail differences in poverty perception. This might be an evidence of perception distortion due to `Hukou` system. We shall discuss this distortion effect in subsection 6.4.

6.3 Preference for redistribution

Now let's look at the first column of Table 8 which corresponds to the estimate of the preference for redistribution equation. The region dummy effects are not significant contrary to the birth cohort effects which are significant and have a positive effect. This means that people who were born later are more supportive of redistribution (Chinese society is changing). Females are more supportive of redistribution (the positive effect of gender upon preference could be emphasized via an indirect channel through *poor-lazy* equation since the two variables are negatively correlated). This effect has also been found in the literature for many different countries. The party membership has no significant direct effect in the preference equation (but a significant negative indirect effect via *poor-misgov* channel). Living in couple has no significant effect.

The number of years of education has no significant effect in the preference for redistribution equation. Clark and d'Angelo (2008) found that more educated people are less in favour of redistribution, using the BHPS. However, the effect of education upon the preference for redistribution (after controlling for income) can be ambiguous as this has been pointed out in Alesina and Giuliano (2009). Higher educated people could be more altruistic while they could also take into account the potential loss of their education premium entailed by redistributive policies. Although we find no direct effect of education upon preference for redistribution, the significant indirect effect of education (through the *miss-gov* channel) cannot not be neglected.

The income effect is monotone and negative as discussed for the *poor-misgov* equation. The occupation prestige measured by the scaling metric ϕ_k has no significant effect while the comparisons made between the scaling metric of different generations are very significant. This means that people who experienced an upward mobility compared to their parents are less in favour of redistribution. This result is then coherent with that of Clark (2003). Remember that the sign effect for upward mobility between son and father had to be negative in the *poor-lazy* equation while it is again negative in the preference equation and the correlation between these two equations are found to be negative. This could be due to two reasons: i) the correlation parameter is much smaller than the one between *preference for redistribution* and *poor-misgovernment* equations and ii) one who has experienced an upward mobility should not agree that poverty is caused by laziness because that would be equivalent to say that his father was lazy.

The variable *better finance* which captures the subjective measure of the expected household financial situation has no significant direct effect. It only enters with a significant negative impact via the *poor-lazy* channel.

Being rural reduces the support for redistribution but this negative effect. We also found a very significant effect of the "rural" variable in the other two equations. So, the "rural" variable plays a very important role in poverty perceptions and thus influences indirectly preference for redistribution. However, the effect of *Hukou system* is still unclear. If the rural-urban barrier (Hukou system) which prevents rural area migrants from reaching the urban areas did exist, then society would be divided into two isolated parts. Thus the Hukou barrier could influence in a diverging way preference for redistribution as well as poverty perception of the two sub-populations (rural and urban). Nevertheless, we cannot ignore the fact that some individuals, while being registered as rural are in fact migrant workers, living and working in urban areas. Most of them are occupying low paid physical jobs and they receive much less social benefits compared to the native urban residents, see for instance Wong et al. (2007). On the other hand, what they have experienced and seen in urban areas is totally different from what they have seen in their hometown (discrimination and between group inequality). Thus their preferences and perceptions are drifted compared to those who have remained in rural areas. The change in attitude of individuals who have a dual identity (rural identification and migrant worker status) should modify the interpretation we have of the rural-urban dichotomy, because the group identified as being rural is heterogenous. Moreover, the dual identity of migrants provides a natural experimental subject to understand what would be the effect of the barrier. I insert a dummy variable "being a migrant worker" which corresponds to those who are working in an urban area while still being registered as rural in the year 2006. We see from the last line of Table 8 that migrant workers believe less that poverty is caused by laziness but this effect is only significant at 10% level. Meanwhile, migrant workers are much more in favour of redistribution (than those who are simply "rural" or "urban") even if "rural" individuals have lower preference for redistribution compared to urban individuals and thus the migrant group desire most the redistribution among the population. Clearly, the discriminations entailed by the *Hukou policy* have significant effects upon inequality perceptions and preference for redistribution.

6.4 Being migrant as an endogenous variable

However, we might ask the following question: What are the characteristics of these "migrant workers", and what are their motivations to migrate given the fact that they might face discriminations, have a weak social capital, that they have left behind their children and so on? Despite the prospect of a better payment, the decision to migrate is not an easy one, as described in Wong et al. (2007). Thus, it is reasonable to think that "being migrant" is an endogenous variable in our attitude equations. There are two potential sources of endogeneity:

- Omitted variables that determinate both attitudes and the migration decision, e.g. experiences documented by other people.
- Reverse causality, i.e. although attitudes may be affected by migration experience, people holding different attitudes about poverty and redistribution may also take different decisions whether to migrate or not.

Ignoring the endogeneity bias may lead to estimation and interpretation problems. A standard consistent estimate of the treatment effect in a linear framework is IV estimation. However, as we are within a high non-linear framework, especially when the endogenous variable is also discrete ("being migrant" is a binary variable), the solution is less evident.

Instead of a 2SLS or Heckman two-step estimation as for linear models,[5] I consider a simultaneous triangular system:

$$\begin{aligned} Y_m &= k \times \mathbb{1}[\tau_{m,k-1} \leq X'\beta_m + \kappa_m D + \epsilon_m \leq \tau_{m,k}], \\ D &= \mathbb{1}[\tilde{X}'\alpha + Z'\gamma + \nu > 0]. \end{aligned}$$

The first equation represents our previous ordered probit model when the second equation is a simple probit model that determines what are the characteristics of the "migrant workers" (here the zero-one variable D) within the rural group, given that $X, \tilde{X}$ may or may not have identical elements, that $\tilde{X} \subseteq X$ and $(X, Z) \perp (\epsilon_m, \nu)$, where $\perp$ denotes statistical independence. The error terms of these two groups of equations are ϵ_m and ν. The joint distribution of the error terms is a multivariate normal with

$$(\epsilon_m, \nu)' \sim N_{m+1}(F_{\epsilon_m}(\epsilon_m), F_\nu(\nu), \rho).$$

The endogeneity problem exist whenever $\rho_{\epsilon_m,\nu}$ is non-zero. The decision of migration concerns only the rural group, i.e. individuals having a "rural" status, which is a subgroup of our full sample. As we consider an enlarged model, we have to define an enlarged correlation matrix Σ as follows:

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}. \tag{14}$$

The upper left 3×3 part of this matrix involves the whole sample while the last column and last row of this correlation matrix concern only the rural group.

Empirically, we are facing three difficulties:

i) by adding a new probit equation, the dimension of our system increases to 4, thus it is not evident how to construct a PDS correlation matrix Σ for the GHK algorithm[6]

[5] 2SLS does not always provide an average causal effect, but it provides a local average causal effect, as it has been documented in the literature by Angrist and Pischke (2008). Another reason for not using 2SLS is that it would ignore the non-linear structure. Heckman two step estimation (with a correction term) is neither a good candidate as indicated in Freedman and Sekhon (2010).

[6] The correlation constraints found in case of 3×3 do not hold in the 4×4 case because for example, by given value of ρ_{21}, ρ_{31}, ρ_{41}, the interval of ρ_{32} is jointly determined by the function of (ρ_{21}, ρ_{31}) and (ρ_{42}, ρ_{43}) while the later pair is also unknown.

ii) how to identify the κ_m in this triangular system?

iii) what would be a proper instrumental variable?

To answer the first question, instead of treating with the correlation constrains, I start by building a lower triangular matrix A such that $\Sigma = AA'$ which would correspond to the Cholesky decomposition of Σ. This matrix A is built in such a way that:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ s_{21} & \sqrt{1-s_{21}^2} & 0 & 0 \\ s_{31} & s_{32} & \sqrt{1-s_{31}^2-s_{32}^2} & 0 \\ s_{41} & s_{42} & s_{43} & \sqrt{1-s_{41}^2-s_{42}^2-s_{43}^2} \\ \end{pmatrix} \quad (15)$$

with

$$\Sigma_j^{i-1} s_{ij}^2 < 1 \qquad \forall i > 1.$$

This last condition has to be imposed line by line and is obtained by the spherical coordinate system defined for n-dimensional Euclidean space[7] with a radical coordinate variable $r \in [0,1]$ and $n-1$ angular coordinates $\omega_1, \omega_2, \cdots, \omega_{n-1}$ where $\omega_{n-1} \in [0, 2\pi[$ and other angles range over $[0, 2\pi]$. We then have:

$$\begin{aligned} s_{i1} &= r_i \cos(\omega_1) \\ s_{i2} &= r_i \sin(\omega_1)\cos(\omega_2) \\ &\vdots \\ s_{i,i-1} &= r_i \sin(\omega_1)\cdots\sin(\omega_{i-3})\sin(\omega_{i-2}) \end{aligned} \quad (16)$$

Consequently, the resulting $\Sigma = AA'$ will automatically fulfill the positive definite condition and will have unit elements along its diagonal. This method can be generalized to higher dimension problems straightforwardly as shown in Equations (16). The method we designed for the 3×3 case is more efficient as it is purely analytical, but cannot be easily generalized for larger cases, if possbile.

Secondly, there is much controversy in the literature concerning the exclusion restrictions in a LDV model with discrete endogenous variables. For example, Wilde (2000) argue that the exclusion restrictions are statistically not required. However, as pointed out by Chesher and Smolinski (2012) and Meango and Mourifie (2013) among others, the exclusion restrictions are indeed essential to identify the model. In our model, the instrumental variable Z only appears in the probit equation and is excluded from the attitude equations.

[7] n stands for the matrix dimension

Thirdly, we need an instrument that is exogenous and brings effect upon attitudes through the channel of the "migration worker" dummy variable. In this paper I have chosen the dummy variable: being the eldest among respondent's siblings. The birth order is random for a given person, is exogenous and has no direct effect upon attitudes for inequality and redistribution. However, being the eldest person among siblings would have a higher incentive to be migrant worker in order to support the family or even to raise the younger siblings.

The GHK simulation estimation for trivariate ordered probit model can be extended to a trivariate ordered probit + a probit model once we are able to construct a positive definite correlation matrix making use of (15).

In Table 9, I give the estimation results for this new specification and I contrast them to the results of Table 8. First of all, what are the characteristics of migrant workers? They are more likely to be from the East-coast region, to belong to the younger generation, to be the eldest among siblings. In the last probit equation, I do not insert other variables that are present in the attitude equations because they are potentially not exogenous variables for my dummy dependent variable "being a migrant". Then we shall notice that by correcting for the endogeneity bias of being a "migrant worker", the treatment effect of the dummy "migrant worker" has dramatically increased at very significant level in all of the three attitude equations while we found no significant changes for the other coefficients corresponding to exogenous variables. Except for the coefficient of the variable "being rural" which has also changed as now "rural" individuals do not identify themselves with the *poor-lazy* argument while only the migrant group does so. Among the full sample, migrant workers are more in favour of redistribution, more likely to identify with the arguments that poverty is caused by misgovernment and laziness. Meanwhile, the correlation coefficient between the migrant equation and the other equations, ρ_{ν,ϵ_m} are all significant with a negative sign. Thus the endogeneity bias correction is justified. These spurious correlations ρ_{ν,ϵ_m} actually explain the downward endogenous bias of the naïve model as the last line in Table 8 reports much lower estimates for the migrant variable. The correlation between *poor-misgov* equation and *poor-lazy* equation becomes now significant with a positive sign, although this correlation is very close to zero. This might not be very surprising because migrant workers (a subgroup of the "rural" group) hold the same positive attitude towards *poor-misgov* and *poor-lazy* arguments.

Table 9: Preference for redistribution
and poverty perception: Being migrant is treated as endogeneous

	Redis. Pref.	Poor-misgov	Poor-lazy	Migrant
Constant				-1.687^{***}
				(0.161)
C. China	0.024	-0.073^{*}	0.089^{**}	-0.155^{*}
	(0.034)	(0.034)	(0.032)	(0.076)
W. China	-0.005	-0.111^{***}	0.142^{***}	-0.042
	(0.030)	(0.030)	(0.030)	(0.073)
birth 60-79	$0.080^{\cdot}$	0.110^{*}	-0.100^{*}	0.342^{**}
	(0.044)	(0.044)	(0.045)	(0.126)
birth post 80	0.094^{*}	$0.089^{\cdot}$	-0.148^{**}	0.576^{***}
	(0.048)	(0.047)	(0.048)	(0.131)
female	0.075^{*}	-0.028	-0.081^{*}	$-0.117^{\cdot}$
	(0.034)	(0.035)	(0.034)	(0.065)
party	-0.023	-0.089^{*}	$-0.067^{\cdot}$	$0.148^{\cdot}$
	(0.037)	(0.035)	(0.036)	(0.081)
coupled	0.019	0.013	0.031	$-0.160^{\cdot}$
	(0.042)	(0.042)	(0.040)	(0.089)
eldest among siblings				0.163^{*}
				(0.064)
yeduc	-0.006	-0.016^{**}	0.004	
	(0.005)	(0.005)	(0.005)	
ln income	0.101^{***}	0.079^{***}	-0.012	
	(0.020)	(0.019)	(0.019)	
ln income squared	-0.009^{***}	-0.008^{***}	0.001	
	(0.002)	(0.002)	(0.002)	
Occup. prestige	0.062	-0.051	0.049	
	(0.071)	(0.072)	(0.068)	
Upward (fath./son)	-0.143^{**}	0.038	-0.148^{**}	
	(0.049)	(0.050)	(0.046)	
Upward (moth./dau.)	-0.131^{*}	$0.108^{\cdot}$	-0.062	
	(0.053)	(0.055)	(0.051)	
better finance	-0.000	0.035	0.167^{***}	
	(0.027)	(0.026)	(0.026)	
Rural	-0.155^{***}	-0.160^{***}	0.068	
	(0.046)	(0.044)	(0.043)	
migrant worker	0.693^{***}	0.673^{**}	0.812^{***}	
	(0.206)	(0.217)	(0.197)	
1—2	-1.792^{***}	-2.103^{***}	-0.693^{***}	
	(0.096)	(0.094)	(0.091)	
(2—3)-(1—2)	1.057^{***}	1.259^{***}	1.251^{***}	
	(0.030)	(0.034)	(0.020)	
(3—4)-(2—3)	1.475^{***}	1.666^{***}	0.991^{***}	
	(0.019)	(0.022)	(0.021)	

correlation	$\rho_{R.P,misgov}$	0.214^{***}	$\rho_{\nu,R.P}$	-0.236^{***}
		(0.019)		(0.094)
parameters	$\rho_{R.P,lazy}$	-0.067^{***}	$\rho_{\nu,misgov}$	-0.340^{**}
		(0.006)		(0.115)
(spurious)	$\rho_{misgov,lazy}$	0.040^{*}	$\rho_{\nu,lazy}$	-0.446^{***}
		(0.020)		(0.132)
N		5138		2931
Loglik		-17 829		
R		25		

I have used the same random draws for η_m as for Table 8 and a new set of draws for ν.

7 Conclusion and discussion

In this paper we have discussed the determinants of preference for redistribution along with the subjective perception of the origins of poverty. It is obvious that the self-interest variables are not the sole factors at action here, there are also some unobserved factors such as ideology and psychological traits. The preference for redistribution and the poverty perceptions are correlated topics and the causalities are complex. If they are correlated in unobserved ways (mutually endogenous), the standard independent estimation may leads to an inefficiency problem and less information could be provided. A multivariate ordered probit model is designed to capture the conditional correlations of correlated ordinal variables. An extended triangular specification has been applied in order to correct the endogenous treatment effect bias of a binary variable. The estimation via GHK simulator algorithm allows many modelling flexibilities and the way of constructing a positive definite matrix discussed in this paper is essential when the dimension of the correlation matrix is high (3×3 or more).

Several evidence has been found in this paper, using the proposed model. First of all, the correlations among preference for redistribution and the poverty perceptions are important. These results provide a proof of the existence of `sense of justice`. Meanwhile, laziness and misgovernment are not two negatively correlated causes for poverty, at least in the perception of the Chinese people. The correlations also allow us to investigate the direct and indirect effects of explanatory variables in this simultaneous system, e.g. the effect of party membership and being rural have no direct effect upon preference for redistribution but they could have some influences through indirect channels of the poverty perceptions. From the estimates of *poor-misgov* and *poor-lazy* equation, we see that circumstances drift perceptions too. The differences in perceptions between rural and urban group is mainly due to the rural-urban policy barrier. I found that being migrant is endogenous to inequality attitudes and a consistent estimation of "being migrant" shows very important barrier discrimination effect upon attitudes. Our results also provide proofs of the most discussed economic theory in the redistribution preference topic, i.e. the intergenerational mobility effect.

In this paper we also discussed the occupation mobility and prestige in China. Our evidence suggests that the widely used EGP occupation category might not reflect the correct social ladder in China, thus in order to use it, one shall be cautious and some readjustments are necessary.

A Properties of the Transition Matrix

A.1 Mobility indices

A transition matrix P has K independent rows. Each row indicates the probability to change from status j to status k the next period and sums to 1. Overall mobility can be summarised using a Prais (1955) index,

$$M_P(P) = (K - \mathrm{Tr}(P))/(K-1).$$

$M_P(P) = 0$ is perfect immobility while $M_P(P) = 1$ is perfect mobility.

B Proof

If two vectors a and b with zero mean and variance of 1 who are in an inner product space, according to Cauchy-Schwarz inequality we shall have:

$$|\langle a, b\rangle| \leq \sqrt{\langle a, a\rangle\langle b, b\rangle} \tag{17}$$

so that:

$$-1 \leq \frac{\langle a, b\rangle}{\sqrt{\langle a, a\rangle\langle b, b\rangle}} = \rho_{ab} \leq 1 \tag{18}$$

Given the fact that the correlations ρ_{ab} and ρ_{bc}[8] are within 0 and 1 while the correlation ρ_{ac} is unknown, the problem can be solved by using the orthgonal decomposition. Since both vector a and c are correlated to vector b and their correlations are known, we can rewrite a and c as:

$$\begin{aligned} a &= \langle a, b\rangle b + O_b^a \\ c &= \langle c, b\rangle b + O_b^c \end{aligned}$$

where O_b^a is the orthogonal projection of vector a onto b. Then the correlaion between a and c can be written as:

$$\rho_{ac} = \langle a, c\rangle = \langle \rho_{ab} b + O_b^a, \rho_{bc} b + O_b^c\rangle = \rho_{ab}\rho_{bc} + \langle O_b^a, O_b^c\rangle \tag{19}$$

and because that:

$$-1 \leq \frac{\langle O_b^a, O_b^c\rangle}{\sqrt{\langle O_b^a, O_b^a\rangle\langle O_b^c, O_b^c\rangle}} \leq 1$$

[8]Vector c has also mean of zero and variane of 1

and:

$$\begin{aligned}\langle a, a\rangle &= \langle \rho_{ab} b + O_b^a, \rho_{ab} b + O_b^a\rangle = \rho_{ab}^2 + \langle O_b^a, O_b^a\rangle \\ &\Rightarrow \langle O_b^a, O_b^a\rangle = 1 - \rho_{ab}^2\end{aligned}$$

because variance of a is 1, we then have the following condition that:

$$-\sqrt{(1-\rho_{ab}^2)(1-\rho_{bc}^2)} \leq \langle O_b^a, O_b^c\rangle \leq \sqrt{(1-\rho_{ab}^2)(1-\rho_{bc}^2)} \tag{20}$$

Finally, by replacing Inequation (20) into Equation 19 we have:

$$\rho_{ab}\rho_{bc} - \sqrt{(1-\rho_{ab}^2)(1-\rho_{bc}^2)} \leq \rho_{ac} \leq \rho_{ab}\rho_{bc} + \sqrt{(1-\rho_{ab}^2)(1-\rho_{bc}^2)}$$

C RNG

Draw a random number π from a truncated standard normal distribution, for example from $f(\pi | a < \pi < b)$, I apply an inverse sampling approach:

1. First draw r^{th} random number ξ^r from uniform(0,1) distribution.
2. Define $\bar{\xi}^r = (1 - \xi^r)\Phi(a) + \xi^r\Phi(b)$
3. Obtain $\pi = \Phi^{-1}(\bar{\xi}^r)$ which relies between a and b.

Notice that the random numbers are drawn once and kept (McFadden 1989) when parameters vary during the MLE process.

References

Afridi, F., Li, S., and Ren, Y. (2012). Social identity and inequality: The impact of china's hukou system. Discussion Paper 6417.

Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, pages 960–980.

Alesina, A. and Glaeser, E. L. (2004). *Fighting poverty in the US and Europe: A world of difference*. Oxford Univ Press.

Alesina, A. and La Ferrara, E. (2005). Preferences for Redistribution in the Land of Opportunities. *Journal of Public Economics*, 89(5):897–931.

Alesina, A. F. and Giuliano, P. (2009). Preferences for redistribution. Technical report, National Bureau of Economic Research.

Anderson, J. A. (1984). Regression and ordered categorical variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(1):1–30.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Bean, C. and Papadakis, E. (1998). A comparison of mass attitudes towards the welfare state in different institutional regimes, 1985–1990. *International Journal of Public Opinion Research*, 10(3):211–236.

Benabou, R. and Ok, E. A. (2001). Social mobility and the demand for redistribution: The POUM hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487.

Benabou, R. and Tirole, J. (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics*, 121(2):699–746.

Cappellari, L. and Jenkins, S. P. (2003). Multivariate probit regression using simulated maximum likelihood. *The Stata Journal*, 3(3):278–294.

Chen, Y. and Cowell, F. A. (2013). Mobility in china. Technical report, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Chesher, A. and Smolinski, K. (2012). Iv models of ordered choice. *Journal of Econometrics*, 166(1):33–48.

Clark, A. and d'Angelo, E. (2008). Upward social mobility, wellbeing and political preferences: Evidence from the bhps. Technical report, working paper, Paris School of Economics, 17 October.

Clark, A. E. (2003). Inequality-aversion and income mobility: a direct test. Technical report, Delta.

Demurger, S., Sachs, J. D., Woo, W. T., Bao, S., and Chang, G. (2002). The relative contributions of location and preferential policies in china's regional development: being in the right place and having the right incentives. *China Economic Review*, 13(4):444–465.

Edlund, J. (2008). Nominal category class classifications. CESSDA PPP workpackage 9 meeting, Paris, France.

Fang, C., Zhang, X., and Fan, S. (2002). Emergence of urban poverty and inequality in china: evidence from household survey. *China Economic Review*, 13(4):430–443.

Feldman, S. and Zaller, J. (1992). The political culture of ambivalence: Ideological responses to the welfare state. *American Journal of Political Science*, pages 268–307.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82(2):225–246.

Freedman, D. A. and Sekhon, J. S. (2010). Endogeneity in probit response models. *Political Analysis*, 18(2):138–150.

Ganzeboom, H. B., Luijkx, R., and Treiman, D. J. (1989). Intergenerahonal class mobility in comparative perspective. *Research in social Stratification and Mobility*, 8:3–84.

Hendrickx, J. (2000). Special restrictions in multinomial logistic regression. *Stata Technical Bulletin*, 56:18–26.

Houkai, W. and Xiaoxia, W. (2009). The review and prospect of china's anti-poverty policies [j]. *The Journal of Shanghai Administration Institute*, 2:008.

Kanbur, R. and Zhang, X. (1999). Which regional inequality? the evolution of rural–urban and inland–coastal inequality in china from 1983 to 1995. *Journal of comparative economics*, 27(4):686–701.

Karabarbounis, L. (2011). One dollar, one vote. *The Economic Journal*, 121(553):621–651.

Khan, A. R. and Riskin, C. (1998). Income and inequality in china: composition, distribution and growth of household income, 1988 to 1995. *The China Quarterly*, 154:221–253.

Khan, A. R. and Riskin, C. (2001). Inequality and poverty in china in the age of globalization. *OUP Catalogue*.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *econometrica*, 57(5):995–1026.

Meango, R. and Mourifie, I. (2013). A note on the identification in two equations: Probit model with dummy endogenous regressor. *Available at SSRN 2340187*.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government.

Neustadt, I. (2011). Do religious beliefs explain preferences for income redistribution? experimental evidence. *CESifo Economic Studies*, 57(4):623–652.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3):551–584.

Prais, S. (1955). Measuring social mobility. *Journal of the Royal Statistical Society. Series A (General)*, pages 56–66.

Rawls, J. (1971). *A Theory of Justice*. Oxford University Press, London.

Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics*, 4(2):163–185.

Scheve, K., Stasavage, D., et al. (2006). Religion and preferences for social insurance. *Quarterly Journal of Political Science*, 1(3):255–286.

Sen, A. (1980). *Equality of what?*, volume 1. na.

Sen, A. (1999). Commodities and capabilities. *OUP Catalogue*.

Tanner, M. S. (2006). We the people (of china)... *Wall Street Journal*.

Whyte, M. K., Maocan, G., Kwok-bun, C., Ku, A. S., and Yin-wah, C. (2009). How angry are chinese citizens about current inequalities? Evidence from a National Survey. *Social Stratification in Chinese Societies*, page 17.

Wilde, J. (2000). Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters*, 69(3):309–312.

Wong, K., Fu, D., Li, C. Y., and Song, H. X. (2007). Rural migrant workers in urban china: living a marginalised life. *International Journal of Social Welfare*, 16(1):32–40.

Wu, X. (2007). The chinese household registration and intergenerational social mobility. *Sociological Studies*, 6:38–65.

Xu, J. and Liu, H. (2013). Social justice recognition, fluidity expectation and the preference of residence redistribution-an empirical study based on cgss data. *Journal of Yunnan University of Finance and Economics*, 2.