école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Preference for Redistribution, Poverty Perception among Chinese Migrants

Zhou Xun
Michel Lubrano

WP 2022- Nr 28

Preference for Redistribution, Poverty Perception among Chinese Migrants *

Zhou XUN † Michel LUBRANO ‡

November 2022

Abstract

We analyse preference for redistribution and the perceived role of "circumstances" and "effort" in China within the framework of the *belief in a just world* hypothesis (BJW) using the 2006 CGSS. As this very rich data base does not include Dalbert questionnaire on GBJW and PBJW, we have completed the CGSS by a survey led during the COVID episode in Shanghai and Nanjing. Thanks to this new survey, we could identify the components of PBJW and GBJW inside the traditional opinion variables about the causes of poverty and the desire for redistribution of the CGSS. Using a tri-variate ordered probit model for explaining opinions, we show how treating the decision to migrate as an endogenous variable modifies the usual results of the literature concerning migrants and the effects of the Hukou status. The correlations found validate the distinction between personal BJW and general BJW, a distinction that has important policy implications for the status of migrants.

Keywords: Preference for redistribution, inequality perceptions, belief in a just world, Hukou and migrant workers, conditional correlations, binary endogenous, GHK simulator, marginal effects.
JEL classification: C36, D19, H23, J18.

*During the writing of this paper, we have benefited from very useful conversations with Habiba Djebbari, Emmanuel Flachaire and Gilles Stupfler. Of course, remaining errors are solely ours. This work has been carried out thanks to the support of the *National Science Foundation of China* (Grant No.71764008). The project leading to this publication has received funding from the French Government under the "France 2030" investment plan managed by the French National Research Agency (reference: ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University - A*MIDEX.

†School of economics, Nanjing University of Finance and Economics. Email decxun@hotmail.com

‡School of economics, Jiangxi University of Finance and Economics and Aix Marseille Univ, CNRS, AMSE, Marseille, France. Email michel.lubrano@univ-amu.fr, corresponding author. Address: AMSE, 5 Bd Maurice Bourdet, 13001 Marseille, France

[0]**Data acknowledgement:** The data analysed in this paper were collected by the research project

1 Introduction

After starting the market-oriented reforms of 1978, China became the world's fastest growing economy over the last 30 years. However, this significant economic growth came together with an inequality increase which soon appeared to become a major issue. The income distribution changed dramatically after 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. And using the Chinese Household Nutrition Survey, Chen and Cowell (2017) found that in the post-millennium area, climbing on the income ladder has become more difficult.

Social reforms together with fast economic growth have adversely affected different groups of people. More precisely, the rural-urban gap has increased (see e.g. Ravallion and Chen 2007 or Piketty et al. 2017). The *Hukou* policy (household registration system) can be thought of being one of the major causes of this widening gap. This state policy was adopted to limit mass migrations from land to cities, to ensure both economic and political stability. However, this policy had the consequence of favouring urban residents and discriminating against rural residents in public resource allocation, such as education, job vacancies, social benefits, health care, etc... (see Song 2014 for a general introduction to the *Hukou* system and Afridi et al. 2015 for its consequences on social identity). Moreover, as the redistribution scheme is decentralised in China and mostly depends on local resources, social benefits are much lower in rural areas than in urban areas (Wei and Wu 2009). Thus being "rural" or "urban" entails a huge discrepancy in terms of living standards and overmuch in terms of opportunities.

Rural individuals can decide to move to work in cities in order to have a relative higher income, but they will not necessarily be registered officially as "urban". Local governments have no incentive to provide social benefits to this unregistered population and the central government either does not provide any compensation or has very little power to enforce central state regulations (see Wong et al. 2007). Although these rural labour forces contributed a lot to economic development and urban modernisation, the discrimination entailed by the *Hukou* system prevents them from having access to the benefits of the fruits of development in an equal way. The differences in rural and urban living and working conditions and the isolation of rural people might lead to divergences in the perceptions of the causes of poverty between rural and urban people.

Perception of the causes of poverty and the desire for redistribution form a complex and interdependent mechanism. Preference for redistribution adverts to how individuals perceive themselves as compared to others and is thus related to distributive justice. Poverty perception relates to the opposition between efforts and circumstances (bad luck). Both mechanisms can be related to the belief in a just

"China General Social Survey(CGSS)" sponsored by the China Social Science Foundation. 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 authors appreciate the assistance for providing data by the institutes and by the individuals aforementioned. The views expressed herein are those of the authors.

world (BJW), a theory initiated by Lerner (1965) and used by Benabou and Tirole (2006) to explain differences in the desire for redistribution, depending on historical culture. Knowing the recent Chinese history, this belief can be rather different among rural, urban and migrants in China.

The *first aim* of the paper is to extend the empirical literature on preferences for redistribution by relating it to the psychological literature of the *belief in a just world* (BJW) in China. The basic source of information is the 2006 version of the Chinese General Social Survey (CGSS). This data set contains opinion variables about the desire for redistribution and about the two causes poverty. However, this wave of the CGSS contain very little information about the Belief in a Just World, and in particular it does not make mention to Dalbert (1999) questionnaire. We have thus led a new survey in 2022 including both this questionnaire and questions about the desire for redistribution in order to provide clues for interpreting the CGSS in term of BJW.

The *second aim* of the paper is to measure the specific impact of being a migrant on the formation of opinions. Due to a small proportion of these individuals in the samples, this impact does not appear as being significant in the literature (see e.g. Han 2012). However, the decision to become a migrant can be also influenced by the opinions about the causes of poverty and by the belief in a just world. So this variable has to be considered as endogenous when trying to estimate an econometric model of opinion formation. These opinions are shaped in the context of the *Hukou* policy. When they arrive in cities, rural migrants are marginalised. As underlined in Wong et al. (2007), *"they take up jobs that the urban residents are unwilling to do and live in very poor housing conditions, and their children do not have access to public school systems."* So the *Hukou* system has strong chances of changing their vision of *justice as fairness*, compared to the group of rural individuals that have not decided to migrate. Once this variable is treated as endogenous, we shall show that their opinions become quite different from those of the rural group.

The *third aim* of this paper is to propose an adequate econometric method for modelling and estimating all these interrelationships in order to appraise the exact impact of the *Hukou* policy in China and the recent decision to relax it. For that, we use a simultaneous multivariate ordered probit model for which we propose a simulation method to compute marginal effects within the estimation procedure.

The paper is organised as follows. Section 2 reviews the general literature about belief in a just world and how it can be related to preference for redistribution in order to situate our contribution within the specificity of the Chinese context. Section 3 introduces our data base extracted from the Chinese General Social Survey (CGSS) for 2006 together with our new survey led in 2022. It discusses the choice of the potential determinants of preference for redistribution and their interpretation in term of BJW using our 2022 survey. Section 4 introduces our multivariate econometric model of opinion formation and explain how to treat identification using both IV and heteroskedasticity of the error term. Section 5 is devoted to empirical results and details the importance of considering being-a-migrant as endogenous for appraising its impact on poverty perception and preference for redistribution. The correlations found justify the distinction between General and Personal BJW made in Wu et al. (2010) while section 6 details the policy implications of this distinction in terms of subjective welfare. Section 7 concludes. Technical details are given in appendices, which detail in particular our simulated maximum likelihood estimator.

2 China and the belief in a just world

The literature about preferences for redistribution is in a way rather large. The specific case of China has to be detailed carefully as this country has experienced both a communist system with strong redistributive experiences in the past while more recently being the object of vast liberal economic reforms.

2.1 Preference for redistribution and belief in a just world

The preference for redistribution literature usually starts with self-interest variables, the median voter theory of Romer (1975) and its implementation in a taxation model by Meltzer and Richard (1981). But as China is not a country where people vote for a tax level and a redistribution scheme, it is better to start from a different point of view and we have adopted the *belief in a just world* (BJW) initiated by Lerner (1965), Lerner and Miller (1978), Lerner (1980) and empirically validated in Rubin and Peplau (1975). Essentially, individuals have a need to believe that they live in a world where people get what they deserve and they deserve what they get. In such a world it is easy to make plans, to deliver to children the message that efforts are rewarded. Consequently, religions are promising a reward in afterlife; individuals are poor because they are lazy and a victim is blamed for having adopted a risky attitude (Andre and Velasquez 1990). In a series of surveys, Rubin and Peplau (1975) found that believers in a just world tend to be more conservative and are less involved in social activities intended to alleviate the fate of social victims. Using the scale devised by Rubin and Peplau (1975), Furnham and Procter (1989) reported that BJW was correlated with political support to conservative policies, meaning for instance that victims of social injustice like the poor owed their fate to personal inadequacy rather than to the failing of social institutions. Building on these ideas, Benabou and Tirole (2006) have developed a model based on different recollection of misfortune experiences which lead to different beliefs about the reward of efforts to explain the the desire for redistribution and differences between Europe and the US. If the world is thought to be a just place where efforts are rewarded, there exist a political equilibrium based on laissez-faire public policy with a low tax rate and low redistribution. If individuals are more sensitive to unjust situations, with beliefs that poor are trapped in situations despite their efforts, then there exists a second political equilibrium with a higher tax rate and higher redistribution.

2.2 Chinese beliefs in a just world and redistribution

BJW can play a double role. On one side, it leads to blame victims and favours a conservative social belief system. On the other side, it functions as a personal resource when an individual is confronted to adversity leading her to think that things will be better in the future. Building on this double role of BJW, Dalbert (1999) introduced the distinction between *General BJW (GBJW)* and *Personal BJW (PBJW)*. PBJW is mainly related to self-esteem while GBJW is related to social organisation. In individualistic cultures, adults report higher PBJW and lower GBJW. In a collectivistic culture like China, Wu et al. (2010) indicate that Chinese adults are consistently reporting higher GBJW than PBJW. The aim of Wu et al. (2010) was to test if GBJW helped Chinese people for resilience when confronted to adverse shocks like the 2008 Sichuan earthquake or poverty during adolescence and if GBJW could explain positively life satisfaction. When confronted to adverse situations, individuals tended to keep a high GBJW even if their PBJW broke down. This means that they do not change their beliefs about general justice, but think nevertheless that their personal trauma was undeserved.

These reported differences between GBJW and PBJW have obvious meanings for interpreting the causes of poverty. In the literature specifically devoted to redistribution, Im (2014) found for China counter-intuitive patterns in which mutually conflicting ideas can coexist: for example, pro-redistribution individuals might adopt the idea that inequality is necessary for development. The Chinese economic reforms reinforced agreement to GBJW.

2.3 Desire for redistribution and the *Hukou* system

When investigating GBJW and PBJW, Wu et al. (2010) considered Chinese society as a whole when in fact it is strongly segregated by the *Hukou* system. Chan and Zhang (1999) describe the *Hukou* system as a "larger economic and political system" designed to control life and aspirations. It started in 1951 in cities and was extended to rural areas in 1955. Its main consequence was to block rural-urban migration when the latter started to cause congestion problems in cities after 1960. However, with the economic reforms at the end of the seventies, rural-urban migration developed a lot as additional labour force was needed in the new economic regions of the East coast. That floating population was discriminated against on the labour market and also for accessing social services (see Li 2008 using the *Chinese Household Income Project* data set (CHIP) of 2008 and the *Rural Migration Survey* of 2004).

Whyte (2009) reported that urban residents are more critical to inequalities compared to rural residents and believe less in the role of effort. Rural citizens live in a world which is closer to traditional Chinese values (Confucianism), they lack information about global inequality, so that it is easier for them to adhere to GBJW (see e.g. Wu et al. 2010 for the role of ancient Chinese philosophies). If we go back to the model of Benabou and Tirole (2006), we would expect urban who have experienced much higher incomes to be closer to a low-tax-low-redistribution equilibrium than rural. In fact, Whyte (2009) observed just the contrary. Urban residents have a higher desire for redistribution and are more sensitive to inequality than rural residents. This gap in beliefs between rural and urban is explained in Whyte (2009) by the fact that rural have experienced positive external shocks such as reduction of taxation and more autonomy in cultivation activities. On the other side, urban experienced many exogenous adverse shocks with urban expansion and destruction of houses, market reforms, privatisation, unemployment, adverse shocks that are not considered in the model of Benabou and Tirole (2006).

This urban-rural opposition is made more complex when we consider the in-between floating population, migrants. Migrants are coming from rural areas and so should share rural beliefs. But when joining urban cities, they are confronted both to different beliefs and to the adverse shocks of discrimination and procedural injustice. Wu et al. (2010) pointed out the importance of adverse shocks to explain a low level of PBJW while having a consistent GBJW. Despite they expect to be victims of discriminations that are out of their control and exogenous, migrants still make the choice to move because they know that their efforts are going to be rewarded in term of higher wages and such expectations are endogenous with respect to their belief about the nature of the world. The question is to determine how their beliefs are going to be modified by confrontation with more opportunities and at the same time with discriminations in access to job positions and social services. We present in the next section the two data bases that will help to shed light on this empirical puzzle.

3 Data: The Chinese General Social Survey completed by a new survey

The *Chinese General Social Survey* (CGSS) is a repeated cross-section survey designed to collect individual opinions, social values and judgements about the quality of life in China for individuals over 18, starting in 2003 on an annual or a bi-annual basis. The last released wave was collected in 2017. The CGSS is a sub-project of the *International Social Survey Programme* (ISSP). Following the same structure as the famous US *General Social Survey* (GSS), the CGSS provides multi-dimensional information on both socio-economic characteristics and attitudes and values about social issues. However, only the 2006 wave includes a module containing questions about the perceived causes of poverty which makes it comparable to the *Social Inequality Programme* of the ISSP. In this wave, 28 provinces are included, with Beijing, Shanghai, and some of the other most developed direct-controlled municipalities. This makes a total of 10,151 observations with weights.

However, the CGSS does not contain enough information concerning the belief in a just world. To remedy to this situation, we conducted a complementary survey in 2022 and collected the answers of 324 respondents in order to quantify PBJW and GBJW in China following Dalbert (1999) questionnaire. This survey was issued online on 2nd of April 2022 and finalised on 5th of April 2022, using the survey platform WenJuanXing (www.wjx.cn). This kind of platforms is mainly designed for marketing purposes, focussing on urban residents in order to reach quickly the targeted number of respondents while eliminating abnormal responses. The sample was thus restricted to Shanghai and Jiangsu.[1]

[1]The whole survey was processed during the Shanghai lock-down due to the Covid outbreak (gradual lock-down since middle of March till full lock-down by the end of March). Potential respondents were chosen according to IP address to ensure the location condition. Respondents have been asked to answer

The survey contained 34 questions, in particular including the same questions about the perceived causes of poverty as in the CGSS. But it includes the 13 questions of the Dalbert (1999) questionnaire about PBJW and GBJW. More details are given in Appendix A.

3.1 Two definitions for being a migrant

The *Hukou* system provides to each individual an official status corresponding to her geographical origin and permanent residence. The latter can be either an agricultural or non-agricultural. For simplicity reasons, we use the terms rural and urban. It was quite difficult to change from a rural to an urban *Hukou* status in 2006, somehow easier in 2022 as we shall see with our 2022 survey. People with an official rural status represented 62% of the sample while true urban residents were only 36% in 2006. Among those having an initial rural status, only a very small proportion have managed to obtain an urban status in the previous 10 years. They represent less than 2% of the total CGSS sample. Changing status implies that their property right on their land is lost.

In cities, a question is asked to the individuals if they are local or not to the city where the interview is taking place and 15% were not officially local to the city. They can come either from rural or from other urban areas. Among these 15%, 26% said they were coming from rural areas. Those migrant workers thus represent 3.9% of the whole CGSS sample (using weights) and are 512 in number.[2]

Because there is a lot of discussions in the literature concerning the definition of a migrant (see e.g. Florence 1999), we have to appraise the representativity of the CGSS sample in this domain. According to Li (2008), rural-urban migrants were 130 million in 2006, thus representing 10% of the total population. Akay et al. (2012) report an even larger proportion of rural-to-urban migrant workers in China with a figure of 18 percent of the Chinese population in 2008, representing 234 million. As our CGSS sample contains only 3.9% of rural-urban migrants, we are far from these proportions, indicating that we are probably missing the seasonal migrants and all those staying in town for less than one year. We thus have a much stricter definition of migrants.

Instead of defining rural-urban migrants as those who, while not being local to the city, answer to a subsidiary question about their origin, we can consider the original *Hukou* variable and cross it with the variable not being local to the city where the interview is taking place. With this alternative convention, we would find a much larger sample of 1,142 individuals, representing 13% of the total population.

the total of 34 questions online, using a computer or a smart phone. Those who were able to pass the control check and answered all the 34 questions could receive a cell phone recharge bill coupon valued five Yuans. The control check included IP constraint (Shanghai and Nanjing), device constraint (not in the black list to avoid professional respondents), trap questions (to avoid random behaving respondents), time constraints (to avoid perfunctory manner) and artificial checks. We have received a total sample of 828 and finally got a valid sample of 324.

[2]When their actual *Hukou* status is checked, only 451 have the official rural status out of the 521 who declared coming from a rural area.

We interpret the small group of 512 individuals as long stayers, while the larger group of 1,142 individuals should include seasonal migrants. This difference in length of stay in the city can be confirmed using the variable *have you bought or do you plan to buy an apartment*. 17% of long stay migrants were in this case while this percentage drops to 2% for short term stayers. The accumulated income (*since your arrival to this place how much have you earned in total*) of long term stayers is on average 18,152 CNY and just 8,607 CNY for the other group.[3]

Our new 2022 survey is quite different for the population of migrants as it was collected in an urban area and in a more recent period. First, only 26% of our sample has a *Hukou* rural status, while 63% had a urban status. 11% have managed to change their status from rural to urban. The pure long term migrants represent 17% of our 2022 sample and this percentage goes up to 19% when including short term migrants.

3.2 Income

The income variable regroups all sources of individual income received in the year 2005.[4] There is a strong asymmetry in the income distribution reflected by the large discrepancy between the median income (5,000 CNY) and the much larger mean income (8,235 CNY). The Gini coefficient in the whole sample is very large, 0.540, in accordance with accepted figures (see e.g. Chen et al. 2019). We give in Table 1 a decomposition of the income distribution by subgroups. There is a huge difference between the income distribution of rural and urban people while the income distribution of the long term migrants tends to be close to that of the urban, however with a slightly lower mean and a much higher Gini. Note also that the income of the urban residents includes subsidies and allowances which are not available to migrants. So for reaching their level of income, long term migrants have to provide higher efforts. When adopting a larger definition of the status of migrants (including seasonal migrants), their mean income is closer to the mean income reported in Li (2008) (around 9,000 CNY for migrants) and moreover much closer to the rural mean income.[5]

3.3 Occupation and Social Mobility

Using US data, Day and Fiske (2017) have shown that perceived social mobility had a strong impact on meritocratic and just-world beliefs. In our case, the *Hukou* system can influence a lot both the possibility of occupation mobility and its perception as migration from rural to urban regions is much constrained, limiting thus the mobility opportunities of "rural" people. However, because rural areas were disconnected

[3]Variable *qd14h*, have you bought an apartment: 1 already, 2 planned to, 3 not attempted. Variable *qd14g*, since your arrival to this place how much have you earned in total.

[4]Variable *qd35a* represents annual total personal income (not household income) including wages, bonuses, subsidies, allowances, insurance, interest, rent, business income, profits. There are 805 missing observations. Among the 9,283 valid observations, there are 1,066 observations reported as being zero for 2006, 71% of them being females but only 2% migrants.

[5]We did not collect income data in our 2022 survey.

Table 1: China individual income decomposition in 2006

	Total	Urban	Rural all	Migrant all	Migrant long term	Migrant seasonal
Mean income	8,235	13,072	5,462	8,183	12,133	6,840
Gini	0.540	0.438	0.542	0.520	0.552	0.475

Notes. Yearly income is $qd35a$, zero and NA excluded. Weights used. Urban is defined as *Hukou* status $qa03a \geq 2$. Long term migrants are the non-local coming from a village $qd14c = 1$, rural are those with a rural *Hukou* $qa03a = 1$. Migrant-all are those with a rural *Hukou* and not local to the city $qd14a = 1$. Migrant-seasonal is the complement of migrant long term.

from urban areas due for instance to a lack access to internet, the vision of the urban society and the opportunities that it offers for social mobility are altered and as a consequence their preference for redistribution.[6] Even if it is not a panel, the CGSS provides information on social mobility in China between father's and son's job occupation using the EGP classification.[7] This is a valuable source to measure the impact of social mobility on opinion formation in China. Five categories are proposed which correspond to the items given in Table 2 together with their frequency in the population and their sub-group decomposition. We had to modify the order of the EGP scale, using the stereotype ordered regression of Anderson (1984), in order to adapt it to China. This led to downgrading the Self-Employed category which is usually a despised category in communist countries, compared to manual workers as those societies have a long tradition to inhibit private property. Table 2 shows that urban is the privileged category as it concentrates the upper categories IV and V. Farm labour is of course over-represented among rural. Migrants are mainly skilled-unskilled workers. However, long term migrants have managed to get significant positions in the higher occupation IV, but are far from the urban for the highest occupation V.

Using the corrected EGP scale, we build a dummy variable of intergenerational upward mobility which is coded as 1 if the status of the male respondent (female) is better than that of his father (mother), and 0 otherwise. We get a contrasted picture of social mobility. A majority of urban people has managed to climb the social ladder, even if a significant minority experienced a social downgrading. On the contrary, most rural people did not change. A huge majority of rural who decided to migrate had an upward mobility. However, this result should be taken with care as most rural when they migrate leave the lowest occupation (farmer) for the higher occupation skilled-unskilled workers. The literature on attitudes to redistribution and poverty in China has much insisted on the rural-urban contrast and eventually on some specificities of the migrant group. Table 2 illustrates the complexity of that

[6] The CGSS documents leisure activities. There is no difference concerning television access between urban and rural areas. But 91% of rural households had no access to internet while this figure was only 65% in urban areas at the time of the 2006 survey.

[7] The Erikson, Goldthorpe, and Portocarero classification is a classification initially due to Erikson et al. (1979). The EGP classes are ranked on the basis of two dimensions: Employee monitoring difficulties and human asset specificity (required on the job training).

Table 2: Modified EGP categories and social mobility

EGP	Occupation	Total	Urban	Rural	Mig. all	Mig. LT	Mig. ST
				Frequency			
I	Farm labour	41%	3%	**64%**	12%	13%	11%
II	Self-employed	10%	12%	9%	12%	18%	10%
III	Skilled-unskilled worker	26%	**37%**	20%	**58%**	**43%**	**64%**
IV	Lower sales-service-routine	12%	25%	4%	13%	19%	11%
V	Higher-lower controller	11%	24%	4%	5%	8%	5%
			Intergenerational social mobility				
	Downward	10%	19%	6%	5%	3%	6%
	Same	**49%**	28%	**62%**	17%	17%	16%
	Upward	41%	**53%**	32%	**78%**	**80%**	**78%**

Notes. The frequencies were estimated using sampling weights. Mig. LT means long term migrants, ST short term or seasonal migrants. In this table, the three groups (Urban, Rural, Migrant long term) are mutually exclusive while seasonal migrants covers also some rural people. Migrant all is the sum of seasonal and long term migrants.

group.

3.4 Social values and opinions and their relation to BJW

Three variables are concerned with the desire for redistribution and the perception of poverty in the CGSS. They are phrased as follows (translated from Chinese):

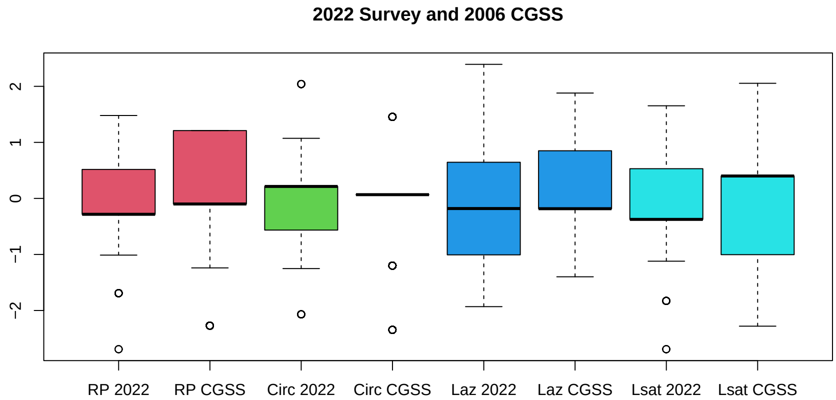
1. *Government should tax more the rich to help the poor.*
2. *Individuals are poor because society is not well functioning, especially because of misgoverning.*
3. *Individuals are poor because they choose to be lazy and this is of their own responsibility.*

These opinions are reported on a four-level Likert scale, tracing the agreement to a given statement (1 for totally disagree to 4 for totally agree). Two variables report life satisfaction:

1. *General life satisfaction up to now*
2. *Generally speaking how satisfied you are with your current life*

In order to be coherent with the Dalbert questionnaire, opinion values in our 2022 survey are reported on a Likert scale between 1 (totally disagree) and 6 (totally agree), with no possibility for a neutral response. The *Poor.misgov* variable was phrased slightly differently, replacing *especially because of misgoverning* by *especially because of reasons out of their control.* The life satisfaction questions were identical to those of the CGSS.

Answers to these questions are visualised and compared with box plots reported in Figure 1. In order to cope with the differences in scales, we have transformed the variables using the POLS transformation of van Praag and Ferrer-i-Carbonell (2004, pages 29-34).



RP is for redistribution preferences in red with in order 2022 survey and CGSS. Poor because of circumstances in green, poor because of laziness in blue.

Figure 1: Box plots for opinions about preferences for redistribution and causes of poverty

There is a strong majority in favour of redistribution in the CGSS sample, slightly less in the 2022 survey. Presumably because of its phrasing, the CGSS question *Poor.misgov* is very concentrated around middle values, while it covers a wider range in the 2022 survey, with values slightly lower than for *Redis.pref*. For *Poor.lazy*, there are more people that agree in the CGSS, probably because it covers a great proportion of rural people. People in the CGSS seem to be more satisfied than in our 2022 survey, probably because some people were in COVID quarantine.

In Table 3, we have computed the polychoric correlation matrix of the three redistribution questions for both data bases (together with life satisfaction). Benabou and Tirole (2006) predict a negative correlation between *Redist.pref* and *Poor.lazy* for a BJW equilibrium. The correlation found is -0.09 in the CGSS and -0.08 in our survey. A second equilibrium can exist where *Redist.pref* and *Poor.misgov* are positively correlated. The correlation found is 0.21 in the CGSS and 0.27 in our survey. So the two data bases present the same correlation structure

Table 3: Polychoric correlation matrix

	2006 CGSS				2022 survey			
	Redist .pref	*Poor* .misgov	*Poor* .lazy	Life .sat	*Redist* .pref	*Poor* .misgov	*Poor* .lazy	Life .sat
Redist.pref	1.00	0.21	-0.09	-0.13	1.00	0.27	-0.08	-0.07
Poor.misgov	0.21	1.00	0.01	-0.07	0.27	1.00	-0.20	-0.15
Poor.lazy	-0.09	0.01	1.00	0.03	-0.08	-0.20	1.00	0.16
Life.Sat	-0.13	-0.07	0.03	1.00	-0.07	-0.15	0.16	1.00

for these two important points. Our tentative to relate these three questions to the belief in a just world by using our more recent survey is thus justified. The sole difference between the two data bases is the zero correlation between *Poor.misgov* and *Poor.lazy* (0.01) in the CGSS, correlation which becomes strongly negative in our 2022 survey (-0.20). Finally, the chosen *Poor.misgov* and *Poor.lazy* variables are coherent with the definition of "bad luck" and "effort" that we discussed in section 2. In both samples, the correlation is negative between $Life.sat$ and $Redist.pref$, $Poor.misgov$ and become positive for $Poor.lazy$.

3.5 Identifying GBJW and PBJW

The correlation matrix given in Table 3 provides a first clue for interpreting *Redist.pref*, *Poor.misgov* and *Poor.lazy* in term of BJW. But we need more clues if we want to introduce the distinction between GBJW and PBJW. This is the reason why we led our 2022 survey. We first note that Dalbert (1999) phrases PBJW as *I deserve more than what I have*, while GBJW is phrased as *in general efforts are rewarded, despite individual efforts might not be*. This might suggest that *Redis.pref* and *Poor.misgov* could be associated to PBJW while *Poor.lazy* mainly concerns GBJW.

The Dalbert (1999) questionnaire is divided into 6 questions for GBJW and 7 questions for PBJW. Our survey provided coherent 324 answers for GBJW and PBJW as the Cronbach (1951)'s Alpha were equal to 0.825 for GBJW, 0.828 for PBJW and 0.887 for the 13 items together. These values are very similar to those found by Wu et al. (2010) for their general sample. Figure 2 shows that mainly people answer in a positive way to most of the questions, but some questions can be viewed in a quite negative way by a minority of people who gave answers below the neutral red dashed line.

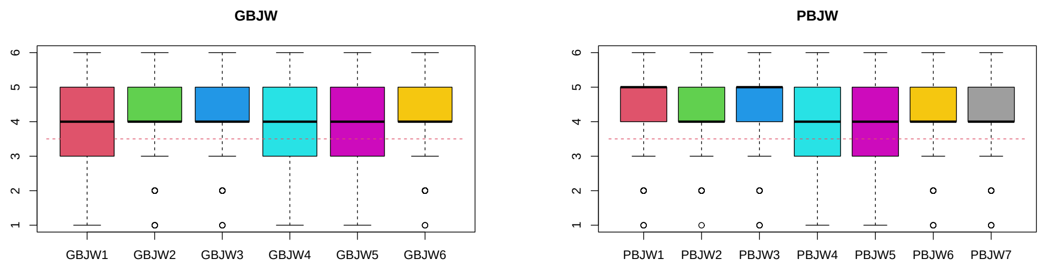


Figure 2: Box plots for GBJW and PBJW questions

Following Wu et al. (2010, page 433, section Method), we have extracted two factors from the 13 GBW questions, a number validated by a BIC criterion. However, with this method, it difficult to identify the meaning of each factor. So we decided to extract one factor from the six GBJW questions calling it GBJW and one factor from the seven PBJW questions, calling it PBJW. The first factor coming from the 13 questions is correlated at 99% with PBJW while the second factor is correlated at 99% with GBJW.

We have now to check that the variable *Poor.lazy* is tightly related to GBJW while *Pref.redist* and *Poor.misgov* are more related to PBJW. However, there is a large difference in term of coverage between our survey and the CGSS. The coverage of the CGSS is the whole China with a large proportion of rural people. In particular, it includes poor migrants with a rural origin who correspond to the victims analysed in Wu et al. (2010). They have bad experiences such as trauma and discrimination. They are the primary concerned about asking for redistribution and for blaming circumstances. This type of migrants is not covered by our survey, because the survey was done by internet and restricted to urban areas. The migrants covered by our survey have the same level of education as the general population and very few have a rural origin. We have to keep in mind those limitations.

Table 4 gives the results of an ordered probit model explaining *Poor.lazy* by the two independent factors PBJW, GBJW and gender. In this model, clearly only

Table 4: Relation between *Poor.lazy* and GBJW

	Estimate	Std. Error	t value
PBJW	0.022	0.036	0.60
GBJW	0.139***	0.040	3.45
Gender	0.338	0.204	1.65

$GBJW$ is significant, justifying the interpretation of *Poor.lazy* in term of GBJW only.

Wu et al. (2010) argue that PBJW is shattered by trauma while in traditional Chinese philosophy GBJW would not be affected. Let us consider the variable *In the past few years I have experienced discrimination or trauma* of our survey and create a dummy variable *Itr* with value 1 when the POLS transformation is positive and 0 otherwise. We then explain *Poor.misgov* by $PBJW$, the interaction of $PBJW$ with *Itr*, $GBJW$, gender and the POLS transform of *Pref.redist*. In this model, all

Table 5: Relation between *Poor.misgov* and PBJW

	Estimate	Std. Error	t value
$PBJW$	-0.060**	0.021	-2.76
$Itr \times PBJW$	0.015*	0.006	2.29
$GBJW$	-0.036	0.023	-1.55
$Pref.redist$	0.232***	0.063	3.65
Gender	-0.269*	0.119	-2.26

the variables are significant, except $GBJW$. So *Poor.misgov* is related to $PBJW$ and not to $GBJW$. *Pref.redist* plays also a significant role, confirming the high positive correlation that we found both in the CGSS and our 2022 survey between *Poor.misgov* and *Pref.redist*. So a linear combination of *Pref.redist* and *Poor.misgov* could be a proxy for $PBJW$.

4 A Multivariate Ordered Probit Model with Endogeneity

We want to explain jointly three opinion variables (*Pref.redist*, *Poor.misgov* and *Poor.lazy*) by mean of a multivariate ordered probit model using explanatory variables. Among the explanatory variables, the decision of *being-a-migrant* is likely related to the underlying psychological traits explaining the previous three opinion variables. We have thus an econometric model with four equations. Identification of the last equation (decision to migrate) is achieved by introducing an instrumental variable. The robustness of the results is checked by using an alternative identification strategy based on heteroskedasticity.

4.1 A Structural Ordered Probit Model

A first group of 3 equations explains the utility y^*_{im} of individual i when answering in a joint manner each of the three opinion questions with $m = 1, 2, 3$. A fourth equation explains the utility d^*_i of individual i when she decides to migrate if rural. So we have the following structural system:

$$\begin{aligned} y^*_{im} &= X'_i\beta_m + d_i\kappa_m + \epsilon_{im}, \qquad m = 1, 2, 3, \quad i = 1, n, & (1)\\ d^*_i &= X'_i\alpha + Z'_i\gamma + \nu_i, & (2) \end{aligned}$$

where X_i is a matrix of explanatory variables, Z_i a set of instruments not contained in X_i and d_i the observed decision of migrating when rural. The error term is composed of ϵ_{im} and ν_i, with joint distribution a 4×4 multivariate normal density with zero mean and correlation matrix Σ:

$$(\epsilon_{im}, \nu_i)' \sim N_4(0, \Sigma), \qquad \Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}. \qquad (3)$$

The upper left 3×3 part of the correlation matrix involves the whole sample while the last column and last row of this correlation matrix concern only the rural group. The endogeneity problem exists whenever the correlation between ϵ_{im} and ν_i is non-zero. Because the utility level has no scale, the diagonal of this matrix is set to 1. This system is completed by a set of observation rules. Corresponding to the utility level y^*_{im}, we observe K levels y_{im} of a Likert scale, while for the utility level d^*_i we observe the zero-one variable d_i. This leads to the system:

$$\begin{aligned} y_{im} &= k \times \mathbb{1}(\tau_{m,k-1} < X'_i\beta_m + d_i\kappa_m + \epsilon_{im} < \tau_{m,k}), \quad k = 1, \ldots, K, & (4)\\ d_i &= \mathbb{1}(X'_i\alpha + Z'_i\gamma + \nu_i > 0) & (5) \end{aligned}$$

where $\mathbb{1}(\cdot)$ is the indicator function equal to 1 when the condition is true and zero otherwise. This observation rule introduces $(K - 1) \times M$ parameters $\tau_{m,k}$ which are unobserved bounds common to all the individuals. This writing is quite general if we suppose that $\tau_{m,0} = -\infty$ and $\tau_{m,K} = +\infty$. For our data set, $K = 4$, the number of levels in each opinion question of the CGSS.

Because the variance-covariance matrix Σ is not diagonal, we have to consider the joint probability of four events, e.g., $(y_{i1} = j, y_{i2} = k, y_{i3} = l, d_i = 1)$ for each individual i for writing down the likelihood function. As computing the probability of a basic event requires the evaluation of a four-dimensional integral, we have to rely on simulated maximum likelihood. Following Geweke et al. (1994), the GHK simulator seems to be the best choice for this class of models. The computation of the joint probabilities and the implementation of the algorithm is explained in Appendix D as well the computation of marginal effects as a by-product of inference in Appendix F. As the algorithm may fail if positivity constraints are not imposed on the variance-covariance matrix when the dimension of Σ is larger than 3, we detail in Appendix E how to impose directly those constraints.

4.2 Identification strategies

The model is identified if the dimension of Z_i is at least one and if $(X_i, Z_i) \perp (\epsilon_{im}, \nu_i)$, where $\perp$ denotes statistical independence. The usual solution to identify the κ_m parameters in the above triangular system is to consider an instrument Z_i which enters the "migration equation" and not the other equations. However, valid instruments are hard to find for explaining the decision to migrate while still being orthogonal to all three opinion variables.[8] Consequently, when Z_i is not available (and thus the exclusion restriction cannot be met), we should identify the endogenous treatment effect by another strategy. One way of obtaining identification is to control for the heteroscedasticity of the error terms and to develop a feasible control, different for each equation (see e.g. Farré et al. 2013 for a linear model). For our ordered probit model, we choose a fairly general form of heteroscedasticity with:

$$\sigma_{im} = \exp(W_i \delta_m),$$

where W_i is a set of observed variables explaining residual dispersion and δ stands for a vector of unknown parameters. Thus the marginal probability function of event k becomes:

$$\Pr(Y_{im} = k) = \Phi\left(\frac{\tau_{m,k} - X_i'\beta_m}{\exp(W_i\delta_m)}\right) - \Phi\left(\frac{\tau_{m,k-1} - X_i'\beta_m}{\exp(W_i\delta_m)}\right).$$

Another advantage of using error heteroscedasticity identification is to improve efficiency for probit and ordered probit models in presence of heteroscedasticity (Litchfield et al. 2012).

[8]There is controversy in the literature concerning the exclusion restrictions in a limited dependent variable model with discrete endogenous variables. On one side, Wilde (2000) argue that the exclusion restrictions are not required because the model is automatically identified due to its nonlinearity. On the other side, Chesher and Smolinski (2012), Meango and Mourifie (2014) among others, pointed out that the exclusion restrictions are indeed essential to identify properly the model.

5 Migrants and BJW in China

We estimate the system of four equations (4)-(5) to measure opinion interactions and the induced support to BJW. We compute marginal effects and then see what would have happened if we had neglected the endogeneity bias for the *being-a-migrant* variable. We discover that this has a tremendous impact on the coefficient of this variable in the three attitude equations, explaining some of the paradoxes contained in the literature.

5.1 Opinion Interactions and the Support for BJW

We consider long term migrants for the variable *Being-a-migrant* because they are those who are the most different from rural people. Treating it as an endogenous variable requires some care. We have selected the traditional instrument *Being-eldest-among-the-siblings*. It means that the eldest among siblings takes responsibility of sharing the burden of raising the younger siblings.[9] We have also excluded from this equation a number of variables which are present in the other equations in order to avoid other endogeneity issues. In the decision to migrate, it is obvious that income and social mobility variables are endogenous. *Education* is also endogenous as argued in Xu et al. (2019). Parents decide to educate their kids with the desire for them to climb the social ladder, which is possible only by migrating to a big town.[10] In Wong et al. (2007) the majority of migrants have a junior secondary school level (technical training school, like cooking, construction, vehicle maintenance), a type of education which is not very useful in rural areas. In Cao and Song (2017), a quarter of the migrant believers have been converted after migration, to compensate their social difficulties in cities, so *Believer* is also endogenous to the decision to migrate. Under all these assumptions, we get inference results as given in Table 6.

In the case of a BJW, the model of Benabou and Tirole (2006) predicts a first equilibrium where people are responsible of their fate and choose low redistribution. In coherence with this prediction, we found a negative correlation (-0.073***) between *Redis.pref* and *Poor.lazy*. Individuals thinking that the main cause of poverty is laziness are also less in favour of redistribution. But in the same model of Benabou and Tirole (2006), a second equilibrium is predicted when some individuals are more sensitive to unjust situations. As we found a positive (0.218***) and stronger correlation (in absolute value) between *Redis.pref* and *Poor.misgov*, poverty has to be compensated by redistribution.[11] The existence of those two equilibria is made possible thanks to the heterogeneity of opinions in the population. Migrants provide an example of the distinction made by Wu et al. (2010) between GBJW and PBJW. When migrants are confronted to difficult situations or trauma, their PBJW is decreased and thus they are more in favour of redistribution (our 0.710 coefficient in Table 6). But at

[9]The one-child-policy was not very strictly applied in rural areas. A single child is not considered as being eldest among siblings.

[10]Remember that in Table 2 a large majority of rural people did not experience any upward social mobility.

[11]Note that the correlation found between the two causes of poverty is not significant.

Table 6:Preference for redistribution and poverty perception

	Redis. Pref.	Poor-misgov	Poor-lazy	Migrant
Constant				-1.884^{***}
				(0.108)
Birth 50-60	0.045	$0.071^{\cdot}$	-0.115^{**}	0.318^{**}
	(0.039)	(0.040)	(0.041)	(0.114)
Birth 70-80	0.042	0.040	-0.154^{***}	0.611^{***}
	(0.043)	(0.043)	(0.044)	(0.116)
Female	0.065^{*}	0.007	-0.075^{*}	-0.156^{**}
	(0.032)	(0.033)	(0.031)	(0.060)
Party	-0.029	-0.098^{**}	$-0.062^{\cdot}$	$0.154^{\cdot}$
	(0.035)	(0.034)	(0.039)	(0.078)
Eldest sibling				0.145^{*}
				(0.063)
Believer	$-0.077^{\cdot}$	-0.118^{**}	$-0.068^{\cdot}$	
	(0.040)	(0.038)	(0.039)	
Years educ	-0.006	-0.011^{*}	0.001	
	(0.005)	(0.004)	(0.004)	
Ln income	0.088^{***}	0.067^{***}	-0.012	
	(0.019)	(0.017)	(0.018)	
Ln income squared	-0.008^{***}	-0.007^{***}	0.001	
	(0.002)	(0.001)	(0.001)	
Upward (fath./son)	-0.113^{***}	0.013	-0.109^{**}	
	(0.033)	(0.034)	(0.033)	
Upward (moth./dau.)	-0.089^{*}	0.047	-0.045	
	(0.038)	(0.039)	(0.038)	
Better finance	0.005	$0.043^{\cdot}$	0.181^{***}	
	(0.025)	(0.025)	(0.025)	
Rural	-0.164^{***}	-0.133^{***}	0.052	
	(0.039)	(0.038)	(0.036)	
Being-a-migrant	$\mathbf{0.710}^{***}$	0.831^{***}	$\mathbf{0.647}^{**}$	
	(0.204)	(0.205)	(0.211)	
(1|2)	-1.917^{***}	-1.986^{***}	-0.837^{***}	
	(0.080)	(0.078)	(0.074)	
(2|3) − (1|2)	1.058^{***}	1.220^{***}	1.270^{***}	
	(0.028)	(0.032)	(0.019)	
(3|4) − (2|3)	1.470^{***}	1.661^{***}	0.992^{***}	
	(0.019)	(0.022)	(0.020)	
		conditional correlations		
$\rho_{R.P,misgov}$	0.218^{***}	$\rho_{migrant,R.P}$	-0.244^{***}	
	(0.009)		(0.015)	
$\rho_{R.P,lazy}$	-0.073^{***}	$\rho_{migrant,misgov}$	-0.422^{***}	
	(0.016)		(0.020)	
$\rho_{misgov,lazy}$	0.035	$\rho_{migrant,lazy}$	-0.352^{***}	
	(0.017)		(0.094)	
N valid		5,402		
Loglik		-19,424		
R		100		

Notes. The optimisation algorithm converged in 15 iterations, with starting values obtained using the univariate ordered probit models. Standard deviations are given between parenthesis. The correlation between the attitude equations are jointly significant. If we try to restrict to zero these structural correlations (but not the migrant correlations), the likelihood value would drop from -19 424 to -19 547, thus rejecting this restriction at any reasonable significance level.
p-value codes given for t ratios: "***" for 0.001 ($t = 3.29$), "**" for 0.01 ($t = 2.58$), "*" for 0.05 ($t = 1.96$) and "." for 0.1. ($t = 1.64$). The likelihood value -19,424 was computed retaining the four equations of the likelihood function. For the GHK simulator, $R = 100$.

the same time they are also more in favour of GBJW (more compared to the other groups) in order to pursuit a long run resilience improvement (our 0.647 coefficient in Table 6) as in the long term efforts are supposed to be rewarded. We shall discuss this aspect in Section 6 in relation to the duration of migration.

Table 7: Marginal effects of being rural or migrant on opinion variables

	Redistribution	Circumstances	Laziness
Rural	-0.090^{***}	-0.058^{*}	0.027
	(-4.10)	(-2.22)	(1.07)
Migrant	0.225^{**}	0.295^{***}	0.569^{***}
	(3.28)	(4.08)	(4.01)

Notes. Marginal effects are obtained by adding the positive probability changes of opinions (3 and 4) and subtracting the negative probability changes (1 and 2). Standard errors are obtained by simulating the parameters, assuming normality with mean and variance at their estimated values. Urban is the reference. $t-$ratios are reported between parentheses.

Let us give the detailed marginal effects for the two variables *Rural* and *Being-a-migrant* for each opinion equation in Table 7. The method is explained in Appendix F. **??** of Appendix F.

Compared to urban, *rural individuals* tend to impute poverty less to misgovernment by 6% while they impute it more to laziness by 3%, even if the latter effect is not very significant. This combination of opinions leads them to be less in favour of redistribution by 9% which is the dominant effect. *Long term migrants*, while being a subgroup of the rural group (they all have the same *Hukou* status), have much clearer opinions. They strongly think that laziness is the cause of poverty by 57% when this opinion was much weaker among the rural people (3%). However, contrary to the latter, they also strongly think that circumstances are important by 30%, even if this effect is roughly twice weaker. Finally, and contrary to the rural people, they turn up to be in favour of redistribution by 23%, a strength comparable to their opinion concerning circumstances. Note that these figures are slightly different from the estimated coefficients of the model, showing the interest in computing marginal effects.

5.2 Explaining a Literature Puzzle

In their study on attitudes toward government responsibility for social services, Han (2012) did not manage to find differences in opinion between rural and migrant categories because of large confidence intervals. Similarly, Whyte and Maocan (2009) and Whyte (2010a) found that most of the time migrants had the same opinion than rural people who were against redistribution and thought that poverty was explained by laziness and not by circumstances: "*Migrants join other rural respondents in having less critical attitudes on Harmful Inequality and Unfair Inequality*".

In Table 6, we have found that long term migrants were quite different from rural people, at odds with the above quoted literature. This is because we have treated migrants as an endogenous variable. A naïve model with only three equations would treat the migrant variable as an exogenous variable. Inference results reported in

Table 8 below shows that in this case the *Being-a-migrant* variable has no specific effect, meaning that long term migrants are not different from the rural population. We can conclude that the literature puzzle comes from an inadequate treatment of endogeneity.[12]

Table 8: The endogeneity bias for measuring preferences for redistribution among the migrants

	Redis. Pref.	Poor-misgov	Poor-lazy
		Migrant exogenous	
		Likelihood: -18,652	
Rural	-0.131^{***}	-0.075^{*}	0.100^{**}
	(0.036)	(0.035)	(0.033)
Being-a-migrant	0.085	-0.104	0.017
	(0.070)	(0.082)	(0.066)
		Migrant endogenous	
		Likelihood: -18,511	
Rural	-0.164^{***}	-0.133^{***}	0.052
	(0.039)	(0.038)	(0.036)
Being-a-migrant	0.710^{***}	0.831^{***}	0.647^{**}
	(0.204)	(0.205)	(0.211)

Notes. We report estimation results only for two variables. The other variables are the same as those reported in Table 6. Standard deviations between parenthesis. The likelihood value -18,511 was computed retaining only the three opinion equations of the likelihood function for ease of comparison with the three equations of the naïve model. Parameter for GHK $R = 100$.

Long term migrants, despite the fact that they have a rural origin, have different opinions. They become in favour of redistribution, strongly think that poverty is explained by circumstances (deficiencies in social organisation) as they are facing discrimination, but also think that poverty is explained by laziness as for earning more they have made a lot of efforts. So migration and its conditions have changed the mind of the former rural persons they were. By using an adequate model with four equations, we have managed to identify long term migrants as a specific population.

5.3 Robustness to Identification

The traditional instrument *eldest-among-siblings* can be seen as unreliable. A way of checking the robustness of our result is to identify our model using heteroskasticity variations. The latter can be explained by variables such as the number of siblings (not eldest among siblings) and regional dummies (non-costal regions, and west regions). The Latent Instrumental Variables (LIV) method of Ebbes et al. (2005) is another possibility, using regional dummies (see Ebbes et al. 2009 for a survey). With Table 9, we find that the identification method has little consequences for infer-

[12] A likelihood ratio test shows that the naïve model is rejected by the data. The three correlation coefficients between the migrant error term and the other error terms, ρ_{ν,ϵ_m} cannot be set to zero. The statistics $2(18,652 - 18,511) = 282$ is far away from the 5% $\chi^2(3)$ critical value of 7.8.

Table 9:The impact of alternative identification strategies

	Redis. Pref.	Poor-misgov	Poor-lazy
	Identification with an instrument likelihood -19,424		
Rural	-0.164^{***}	-0.133^{***}	0.052
	(0.039)	(0.038)	(0.036)
Being-a-migrant	0.710^{***}	0.831^{***}	0.647^{**}
	(0.204)	(0.205)	(0.211)
	Identification through heteroskedasticity likelihood -19,375		
Rural	-0.163^{***}	-0.124^{***}	$0.061^{\cdot}$
	(0.040)	(0.037)	(0.036)
Being-a-migrant	0.554^{**}	0.665^{***}	0.674^{**}
	(0.198)	(0.171)	(0.195)
	Identification through LIV regional dummies Likelihood: -19,420		
Rural	-0.164^{***}	-0.125^{***}	0.049
	(0.039)	(0.038)	(0.035)
Being-a-migrant	0.554^{**}	0.623^{**}	0.780^{***}
	(0.189)	(0.194)	(0.188)

ence on the parameters of the rural and migrant variables. Looking at the reported log-likelihood values, the identification strategy through heteroskedasticty seems to be the most efficient one, but it is also the one which uses the greatest number of variables.

6 GBJW, PBJW and policy implications

The main variable for economic policy adjustment is the *Hukou* status. Rural people when they stay in rural areas do not experience the constraints imposed by the *Hukou* regulations, while migrants become confronted to it as soon as they arrive in towns. Any changes in these regulations will have strong consequences on migrant welfare. At this point, it is important to consider the difference between short term or seasonal migrants and longer term migrants. This difference in length of stay of rural people in an urban environment should influence their perception of a just world, the longer they stay, the longer they are exposed to discriminations.[13] The welfare question is which targeted changes in the *Hukou* system could really improve the situation for both types of migrants.

6.1 Decomposition of SWB

In order to identify the points on which a different *Hukou* policy could improve welfare, we estimate an ordered probit model explaining in a series of equation estimated on the rural sub-sample a decomposition of subjective well-being, considering *Cur-*

[13]At the time of the CGSS, the *Hukou* restrictions were pretty strong. They have been slightly softened since that date.

rent life satisfaction, *family Financial situation* (which should be improved by the decision to migrate), *Family situation* (migrants are most of the time separated from their children that stay in the country side with their parents), *Job satisfaction*, *Housing situation*, *Health status* and two indicators related to *social interactions*. We are going to explain these sub-domains of life satisfaction by $\widehat{Redist.pref}$, $\widehat{Poor.misgov}$ and $\widehat{Poor.lazy}$, which are the predicted values of our general model that can be interpreted in term of PBJW and GBJW, and by two dummy variables to make the difference in satisfaction level for seasonal and for long term migrants. The results presented in Table 10 allow us to compare the situation of both types of migrants, first between them and second with respect to the whole population of rural people.

Table 10: Components of Subjective Well-being

	Lifesat	Finance	Family	Social
Long Mig.	-0.564^{**} (−3.102)	0.371^{*} (2.141)	0.121 (0.640)	−0.056 (−0.288)
Seas. Mig.	-0.195^{***} (−4.030)	-0.148^{**} (−3.202)	0.042 (0.847)	0.059 (1.141)
Redist.pref	-0.905^{***} (−3.641)	-1.166^{***} (−4.923)	-0.682^{**} (−2.637)	-0.964^{***} (−3.637)
Poor.misgov	−0.087 (−0.273)	-1.232^{***} (−4.019)	−0.327 (−0.980)	0.338 (0.991)
Poor.lazy	1.687^{***} (10.082)	1.957^{***} (12.256)	0.764^{***} (4.409)	0.630^{***} (3.586)
	Health	**Housing**	**Community**	**Job**
Long Mig.	0.550^{**} (3.107)	-0.663^{***} (−3.851)	-0.645^{***} (−3.568)	0.061 (0.335)
Seas. Mig.	0.247^{***} (5.184)	-0.160^{***} (−3.483)	-0.203^{***} (−4.164)	-0.204^{***} (−4.206)
Redist.pref	−0.266 (−1.103)	-0.476^{*} (−2.029)	-1.387^{***} (−5.553)	-1.404^{***} (−5.626)
Poor.misgov	−0.520 (−1.667)	−0.397 (−1.308)	0.890^{**} (2.747)	0.081 (0.249)
Poor.lazy	0.210 (1.306)	1.397^{***} (8.855)	1.178^{***} (7.104)	1.351^{***} (8.089)

Each column represents a different ordered probit model with regressors indicated in rows. Weights used with full sample. Current life satisfaction is *qe488*, Family financial condition *qe481*, Family relationships *qe482*, Social relationships *qe483*, Personal health condition *qe484*, Housing situation *qe485*, Community *qe486* (are you satisfied with the community you are living in), Job satisfaction *qe487*. t statistics are given between parenthesis.

There are strong differences between long term and seasonal migrants. On the whole, long term migrants are less satisfied of their current life, of their housing condition, of the community they are living in than seasonal migrants, but also than other rural people. Conversely, they are more satisfied than rural people of their financial situation and of their health condition. Seasonal migrants are less satisfied of their financial situation, housing condition, community relationships, job situation than other rural people, but more satisfied of their health situation. We have seen in Figure 1 that the raw variable *Poor.misgov* had a low variance because of its formulation. Consequently it rarely appears as significant in Table 10. So PBJW is mainly represented by *Redist.pref* which has the strongest negative sign for job satisfaction.

Overall, those who hold strong preference for redistribution are less satisfied in both general and domain satisfactions which reflects the reaction of weakened Personal Belief in a Just World. On the other side, those who believe in $\widehat{Poor.lazy}$, i.e. stronger GBJW, are more satisfied in all respects which shows the role of resilience of General Belief in a Just World.

6.2 Policy measures

Although the Hukou system has been soften recently so that the rural population has more channels to work or to settle down in urban areas, it remains an open question to know how rural migrants could merge to urban life, what would be the challenges and the corresponding integrating policies. This paper may shed some lights to these points. Clearly policy measures could be targeted where subjective well-being is the lowest. And we have found that the population which was the most affected by their rural Hukou status were the long term migrants. The housing situation is not satisfactory for both categories, but it is the most crucial question for long stayers. Both types of migrants were in favour of more redistribution as shown in Table 10 and housing is an item where redistribution could apply, for instance in the form of specific lodging subsidies. Both type of migrant also felt that poverty is due a lot to circumstances, so an adequate housing policy could be an example. Concerning job satisfaction, long stayers are not different than other rural Hukou holders. They believe that efforts are rewarded. But seasonal migrants are dissatisfied with their jobs. Table 2 has indicated that they occupy more often unskilled labour positions and less often sales-services positions (see also Zhang and Wu 2017). We also saw that long stayers identify poverty to laziness while for seasonal migrants opinions concerning the role of laziness was much less pronounced, if not insignificant. We are recognising here the discriminating role of the Hukou barrier for jobs. There is certainly something to do in the direction of the labour market, both around the Hukou regulation and in term of vocational education. Long stayers are more dissatisfied with both their family relationships and their family financial situation. They had to leave their children behind with the grandparents because they mostly cannot access public schools in towns. They could go barely only to private schools which are too expensive or less qualified Chan and Zhang 1999). Seasonal migrants are less sensitive to that question because they are returning often to their home region. We have here another example where the Hukou could be released for long stayer migrants in term to access to public services.

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, situating these preferences in the more general theoretical framework of *Beliefs in a Just World*. The presence of common unobserved factors means that preference for redistribution and the subjective perception of the origins of poverty are mutually endogenous. The correlations found provide a proof for the existence in China, of a *sense of fairness* along the words of Alesina and Glaeser (2004). Meanwhile, laziness and misgovernment are not seen as two negatively correlated causes for poverty, at least in the perception of the Chinese people. These correlations also imply the possibility of indirect effects of explanatory variables in this simultaneous system, implying the necessity to compute global marginal effects.

Migrants have a different perception of poverty and redistribution than the rural people. These differences can be measured only when the migrant variable is treated as endogenous. And this result is robust to the way the model is identified. The explanation of these differences could come from the rural-urban policy barrier. Migrants experience the constraints of the *Hukou* system while rural people who have not yet decided to migrate have not experienced it. And the longer migrants stay in town, the stronger they experience the constraint. We could interpret these differences at the light of the distinction made in Wu et al. (2010) between PBJW and GBJW. This distinction has policy implications in term of targeting.

In the data set they examine, Whyte and Maocan (2009) found that rural people are happier than urban residents because they compare themselves to groups within the same village. Between villages inequality can be important, but within inequality is much lower. And there were economic reforms concerning agricultural prices that had a strong impact for diminishing poverty as documented in Ravallion and Chen (2007) while there is less stress in rural areas than in urban areas (Whyte and Maocan 2009). This can explain the negative desire for redistribution among rural people that we found. However two facts come against the main conclusions of Whyte and Maocan (2009). First, using CGSS data, rural people are less happy than urban people in term of *up-to-now life satisfaction*. And second, long term migrants are a specific population with a strong desire for redistribution, suffering from their *Hukou* status. Not changing the rigidity of this status could amplify the social volcano that Whyte (2010b) tends to minimise. Knowing that the migrant population is tremendously increasing since recent years, ignoring their dissatisfaction could ignite the social volcano.

References

Afridi, F., Li, S., and Ren, Y. (2015). Social identity and inequality: The impact of China's Hukou system. *Journal of Public Economics*, 123:17–29.

Akay, A., Bargain, O., and Zimmermann, K. F. (2012). Relative concerns of rural-to-urban migrants in China. *Journal of Economic Behavior and Organization*, 81(2):421–441.

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

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

Andre, C. and Velasquez, M. (1990). The just world theory. *Issues in Ethics*, 3(2).

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

Cao, Y. and Song, Y. (2017). How does difficulties in social assimilation affect region conversion of rural migrants in China. *Population and development*, 06:70–78.

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

Chan, K. W. and Zhang, L. (1999). The Hukou system and rural-urban migration in China: Processes and changes. *The China Quarterly*, 160:818–855.

Chen, J., Pu, M., and Hou, W. (2019). The trend of the Gini coefficient of China (1978-2010). *Journal of Chinese Economic and Business Studies*, 17(3):261–285.

Chen, Y. and Cowell, F. A. (2017). Mobility in China. *Review of Income and Wealth*, 63(2):203–218.

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

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334.

Dalbert, C. (1999). The world is more just for me than generally: About the personal belief in a just world scale's validity. *Social Justice Research*, 12:79–98.

Day, M. V. and Fiske, S. T. (2017). Movin' on up? how perceptions of social mobility affect our willingness to defend the system. *Social Psychological and Personality Science*, 8(3):267–274.

Ebbes, P., Wedel, M., and Backenholt, U. (2009). Frugal IV alternatives to identify the parameter for an endogenous regressor. *Journal of Applied Econometrics*, 24(3):446–468.

Ebbes, P., Wedel, M., Backenholt, U., and Steerneman, T. (2005). Solving and testing for regressor-error (in)dependence when no instrumental variables are available: With new evidence for the effect of education on income. *Quantitative Marketing and Economics*, 3(4):365–392.

Erikson, R., Goldthorpe, J. H., and Portocarrero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30:415–441.

Farré, L., Klein, R., and Vella, F. (2013). A parametric control function approach to estimating the returns to schooling in the absense of exclusion restrictions: an application to NLSY. *Empircal Economics*, 44:111–133.

Florence, E. (1999). Les migrations en Chine et leurs implications sur le plan urbain. *Hommes et Migrations*, pages 73–95.

Furnham, A. and Procter, E. (1989). Belief in a just world: Review and critique of the individual difference literature. *British Journal of Social Psychology*, 28:365–384.

Geweke, J., Keane, M., and Runkle, D. (1994). Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics*, 76(4):609–632.

Greene, W. H. (1996). Marginal effects in the bivariate probit model. Working paper 96-11, Department of Economics, Stern School of Business, New-York.

Greene, W. H. and Hensher, D. A. (2010). *Modeling Ordered Choices: A Primer*. Cambridge University Press, Cambridge.

Han, C. (2012). Attitudes toward government responsibility for social services: Comparing urban and rural China. *International Journal of Public Opinion Research*, 24(4):472–494.

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

Im, D.-K. (2014). *Attitudes and Beliefs about Distributive Justice in China*. PhD thesis, Harvard University.

Lerner, M. J. (1965). Evaluation of performance as a function of performer's reward and attractiveness. *Journal of Personality and Social Psychology*, 1:355–360.

Lerner, M. J. (1980). *The Belief in a Just World: A Fundamental Delusion*. Plenum Press, New York.

Lerner, M. J. and Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, 85(5):1030–1051.

Li, S. (2008). Rural migrant workers in China: Scenario, challenges and public policy. Working Paper 89, Policy Integration and Statistics Department, International Labour Office, Geneva.

Litchfield, J., Reilly, B., and Veneziani, M. (2012). An analysis of life satisfaction in Albania: An heteroscedastic ordered probit model approach. *Journal of Economic Behavior and Organization*, 81(3):731–741.

Meango, R. and Mourifie, I. (2014). A note on the identification in two equations: Probit model with dummy endogenous regressor. *Economic Letters*, 125(3):360–363.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927.

Mullahy, J. (2017). Marginal effects in multivariate probit models. *Empirical Economics*, 52:447–461.

Piketty, T., Yang, L., and Zucman, G. (2017). Capital accumulation, private property and rising inequality in China, 1978-2015. Working Paper 23368, NBER, Cambridge, MA.

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

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

Rubin, Z. and Peplau, L. A. (1975). Who believes in a just world? *Journal of Social Issues*, 31(3):65–89.

Song, Y. (2014). What should economists know about the current Chinese Hukou system? *China Economic Review*, 29:200–212.

van Praag, B. and Ferrer-i-Carbonell, A. (2004). *Happiness Quantified: A Satisfaction Calculus Approach*. Oxford University Press, Oxford.

Wei, H. and Wu, X. (2009). The review and prospect of China's anti-poverty policies. *The Journal of Shanghai Administration Institute*, 2:56–69.

Whyte, M. K. (2009). How do Chinese citizens view current inequalities? *Sociological Studies*, 1:96–120.

Whyte, M. K. (2010a). Fair versus unfair: How do Chinese citizens view current inequalities? In Oi, J. C., Rozelle, S., and Zhou, X., editors, *Growing Pains: Tensions and Opportunity in China Transformation*, pages 305–332. Shorenstein Center, Stanford.

Whyte, M. K. (2010b). *Myth of the Social Volcano: Perceptions of Inequality and Distributive Injustice in Contemporary China*. Stanford University Press, Stanford.

Whyte, M. K. and Maocan, G. (2009). How angry are Chinese citizens about current inequalities? Evidence from a national survey. In Kwok-bun, C., Ku, A. S., and Yin-wah, C., editors, *Social Stratification in Chinese Societies*, pages 17–54. Brill, The Netherlands.

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, M. S., Yan, X., Zhou, C., Chen, Y., Zhu, J. L. Z., Shen, X., and Han, B. (2010). General belief in a just world and resilience: Evidence from a collectivistic culture. *European Journal of Personality*, 25(6):431–442.

Xu, C., Wei, T., and Yang, Z. (2019). Education and social inclusion of migrants: Empirical analysis of social identity. *Journal of Shanghai University of Finance and Economics*, 21(02):81–93.

Zhang, Z. and Wu, X. (2017). Occupational segregation and earnings inequality: Rural migrants and local workers in urban China. *Social Science Research*, 61:57–74.

Appendix

A Organisation of the 2022 survey

The questionnaire of the new survey was build around seven different sections with a total of 34 questions:

1. Socio economic questions about gender, age, occupation, education, religion belief, member of the PCC, expected financial situation, Hukou status, location, local to the city, from which we could define migrants (Questions 1 to 11, similar to the CGSS)
2. Three questions about desire of redistribution and causes of poverty (Questions 12 to 14, similar to the CGSS)
3. Dalbert questionnaire for GBJW (Questions 15 to 20)
4. Subjective well-being (Questions 21 to 23, similar to CGSS)
5. Dalbert questionnaire for PBJW (Questions 24 to 30)
6. Resilience questions from Wu et al. (2010) (Questions 31 and 32)
7. Trauma and Covid quarantine (Questions 33 and 34).

With this survey, we should be able to build a relation between the CGSS questions about the causes of poverty and the concepts of Personal and General Belief in a Just World. All the opinion related questions were on a scale between 1 and 6 in conformity with the Dalbert questionnaire, with no possibility for a neutral response. 1 is strongly disagree and 6 strongly agree.

B Inference in the SOR model

The stereotype ordered regression (SOR) model of Anderson (1984) is designed to make inference on endogenous ordering in a multinomial model. It explains the propensity score s^*_{ki} of individual i for EGP category k as:

$$s^*_{ki} = \alpha_k + \phi_k(X_i\beta + \sum_{j=1}^{K} \phi_j \mathbb{1}(O_i = j)\gamma_j) + \epsilon_{ki}. \tag{6}$$

α_k refers to the category specific intercepts, X_i is a set of observed variables that controls for human capital (years of education), basic demographic variables (birth cohort, gender and rural/urban status). The influence of initial status of the father is measured by γ_j and O_i is the observed occupation category of the father. The error term ϵ_{ki} is assumed to follow an extreme value distribution. The link function with the status of the son is of the logit type. The SOR model introduces a multiplicative parameter ϕ_k varying with k. It serves to measure the ordinal scale of the destination category. The identification problem is solved by imposing $\phi_1 = 0$ and $\phi_K = 1$. The scaling parameter ϕ_k and the linear parameter β are estimated using an iterative method (Hendrickx 2000). The log odds ratio of the two event probabilities $P(y_i = k)$ versus $P(y_i = k')$ is:

$$\log\left[\frac{P(y_i = k)}{P(y_i = k')}\right] = \alpha_k - \alpha_{k'} + (\phi_k - \phi_{k'})(X_i\beta + \sum_{j=1}^{K}\phi_j \mathbb{1}(O_i = j)\gamma_j). \tag{7}$$

Table 11: Inter-generational occupation mobility using the Stereotype Ordered Regression model

	ϕ_k	α_k	γ_j
EGP1 Farm labour	0.000	0.000	0.000
		(−)	(−)
EGP2 Skilled-unskilled worker	0.620	0.454^{***}	0.321^{*}
		(0.137)	(0.143)
EGP3 Self-employed	0.587	-0.490^{***}	-0.506^{*}
		(0.133)	(0.256)
EGP4 Lower sales service routine	0.848	-0.608^{***}	0.821^{***}
		(0.184)	(0.246)
EGP5 Higher lower controller	1.000	-1.129^{***}	1.109^{***}
		(0.214)	(0.196)
	β		
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 effects of explanatory variables are found by multiplying the constant category parameter β by the 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. The estimation is done using the "mclgen" and "mclest" commands in `Stata`. Table 11 reports the estimates of the intergenerational mobility (father's occupation versus current occupation of the respondent). The "SOR effect" corresponding to the estimates of β is reported in the bottom panel of Table 11. The influence of the occupation of the father is measured by γ_j.

C A multivariate ordered probit

In a multivariate ordered probit model, an individual i has an unobserved utility level y^*_{im} which depends linearly on a set of exogenous variables X_{im} plus an error term:

$$y^*_{im} = X'_{im}\beta_m + \epsilon_{im}, \qquad m = 1, \ldots, M. \tag{8}$$

The error terms ϵ_{im} are distributed according to a normal density of zero mean and variance-covariance matrix Σ with its diagonal elements set to 1 for identification reasons. An observation rule relates the item responses to the utility levels by means of:

$$y_{im} = k \times \mathbb{1}(\tau_{m,k-1} < y^*_{im} < \tau_{m,k}), \qquad k = 1, \ldots, K, \tag{9}$$

where $\mathbb{1}(\cdot)$ is the indicator function equal to 1 when the condition is true and zero otherwise. The $(K-1) \times M$ unobserved bounds $\tau_{m,k}$ are parameters common to all the individuals that have to be estimated. This writing is quite general if we suppose that $\tau_{m,0} = -\infty$ and $\tau_{m,K} = +\infty$. For our data set, $K = 4$.

Let us now introduce a fourth equation so as to consider the binary explanatory variable d_i, *being-a-migrant* as an endogeneous variable. The previous M reduced form equations:

$$y_{im} = k \times \mathbb{1}(\tau_{m,k-1} \leq X'_{im}\beta_m + d_i\kappa_m + \epsilon_{im} \leq \tau_{m,k}), \tag{10}$$

are completed by a structural equation, a probit model:

$$d^*_i = \mathbb{1}(X'_i\alpha + Z'_i\gamma + \nu_i > 0). \tag{11}$$

The error term is composed of ϵ_{im} and ν_i, with joint distribution a 4×4 multivariate normal density with zero mean and correlation matrix Σ:

$$(\epsilon_{im}, \nu_i)' \sim N_4(0, \Sigma), \qquad \Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix}. \tag{12}$$

We have to consider the joint probability of four events, e.g. $(y_{i1} = j, y_{i2} = k, y_{i3} = l, d_i = 1)$ for each individual i, a probability defined by a four dimensional integral:

$$\Pr[y_{i1} = j, y_{i2} = k, y_{i3} = l, d_i = 1] = \int_{\tau_{1,j-1}-\hat{y}^*_1}^{\tau_{1,j}-\hat{y}^*_1} \int_{\tau_{2,k-1}-\hat{y}^*_2}^{\tau_{2,k}-\hat{y}^*_2} \int_{\tau_{3,l-1}-\hat{y}^*_3}^{\tau_{3,l}-\hat{y}^*_3} \int_{-\infty}^{0-d^*_i} \phi_4(\epsilon_1, \epsilon_2, \epsilon_3, \rho)\, \mathrm{d}\epsilon_{i1}\mathrm{d}\epsilon_{i2}\mathrm{d}\epsilon_{i3}\mathrm{d}d_i, \tag{13}$$

where $\hat{y}^*_1, \hat{y}^*_2$ and $\hat{y}^*_3$ are the linear predictors $X'_{im}\hat{\beta}_m$ $(m = 1, 2, 3)$ and d^*_i refers to the linear predictor for the fourth probit equation, ϕ_4 the PDF of a quadri-variate normal distribution, ρ representing the vector of all correlation parameters. Under an IID assumption, the log-likelihood of the entire sample is:

$$\log L = \sum_{i=1}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} \sum_{l=1}^{K} \sum_{d} \times \log \Pr[y_{i1} = j, y_{i2} = k, y_{i3} = l, d_i]. \tag{14}$$

As computing the probability of a basic event requires the evaluation of a four-dimensional integral, we have to rely on simulated maximum likelihood. Following Geweke et al. (1994), the GHK simulator seems to be the best choice for this class of models.[14]

D Implementing the GHK Simulator

The GHK simulator exploits the fact that a multivariate distribution can be decomposed into the product of sequential conditional univariate distributions, which can be easily simulated on a truncated range. The simulator is used to approximate the joint event probability (10) and (11) R times. The average of these R evaluations is then introduced as proxy of log-likelihood function which is then maximized using a standard algorithm like `BHHH` in the package `maxLik` of `R`.

To compute the joint probabilities, the GHK simulator has to generate R draws of the ϵ's and ν_i's. Let A be the lower triangular Cholesky decomposition of the variance-covariance Σ such that $AA' = \Sigma$ with:

$$A = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

We consider four *iid* standard normal random variables η_m, so that we can express the ϵ_m as a linear combination of the four independent η_m, $\epsilon = A\eta$ in a matrix notation, or in an expanded notation:

$$\begin{aligned} \epsilon_1 &= a_{11}\eta_1, \\ \epsilon_2 &= a_{21}\eta_1 + a_{22}\eta_2, \\ \epsilon_3 &= a_{31}\eta_1 + a_{32}\eta_2 + a_{33}\eta_3, \\ \epsilon_4 &= a_{41}\eta_1 + a_{42}\eta_2 + a_{43}\eta_3 + a_{44}\eta_4. \end{aligned}$$

[14] Cappellari and Jenkins (2003) have proposed an implementation of the GHK simulator for evaluating the likelihood function of a multivariate probit model. Here we generalize their approach to the case of ordered probit models, but we also treat specifically the question of the positivity of the variance-covariance matrix of the error terms in appendix E, as positivity can be a serious problem for larger models. The algorithm may fail if positivity constraints are not imposed on the variance-covariance matrix.

Let us now replace the ϵ's by their corresponding values in term of the independent η's. We have a marginal Gaussian probability and two conditional probabilities which are independent by construction. The first marginal probability is defined as:

$$\mathrm{Pr}_1 = \mathrm{Pr}(\tau_{1,j-1} < \hat{y}_1^* + a_{11}\eta_1 < \tau_{1,j}) = \Phi\left(\frac{\tau_{1,j} - \hat{y}_1^*}{a_{11}}\right) - \Phi\left(\frac{\tau_{1,j-1} - \hat{y}_1^*}{a_{11}}\right), \quad (15)$$

where $\Phi(.)$ is the cumulative distribution of $a_{11}\eta_1$. Conditionally on the value of η_1, the second probability is given by:

$$\begin{aligned}&\mathrm{Pr}(\tau_{2,k-1} < \hat{y}_2^* + a_{21}\eta_1 + a_{22}\eta_2 < \tau_{2,k} | \tau_{1,j-1} < \hat{y}_1^* + a_{11}\eta_1 < \tau_{1,j}) \\ &\qquad = \Phi\left(\frac{\tau_{2,k} - \hat{y}_2^* - a_{21}\eta_1}{a_{22}}\right) - \Phi\left(\frac{\tau_{2,k-1} - \hat{y}_2^* - a_{21}\eta_1}{a_{22}}\right). \quad (16)\end{aligned}$$

where $\Phi(.)$ this time is the cumulative distribution of the random variable $a_{22}\eta_2$, η_1 being considered as a fixed quantity. Following the same principle, the evaluation of the third probability writes as:

$$\begin{aligned}&\mathrm{Pr}(\tau_{3,l-1} < \hat{y}_3^* + a_{31}\eta_1 + a_{32}\eta_2 + a_{33}\eta_3 < \tau_{3,l} | \\ &\qquad \tau_{2,k-1} < \hat{y}_2^* + a_{21}\eta_1 + a_{22}\eta_2 < \tau_{2,k}; \tau_{1,j-1} < \hat{y}_1^* + a_{11}\eta_1 < \tau_{1,j}) \\ &\qquad\qquad = \Phi\left(\frac{\tau_{3,l} - \hat{y}_3^* - a_{31}\eta_1 - a_{32}\eta_2}{a_{33}}\right) \\ &\qquad\qquad\qquad - \Phi\left(\frac{\tau_{3,l-1} - \hat{y}_3^* - a_{31}\eta_1 - a_{32}\eta_2}{a_{33}}\right). \quad (17)\end{aligned}$$

Thee same logic applies for the fourth probability.

The first marginal probability is evaluated directly, using a standard numerical routine for Gaussian CDFs. The second probability is conditional on 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 simulate the consequence of the first basic event and then describe the conditional event accordingly. Let us call η_1^* a draw of η_1 coming from a truncated standard normal density with lower and upper truncation points respectively $(\tau_{1,j-1} - \hat{y}_1^*)/a_{11}$ and $(\tau_{1,j} - \hat{y}_1^*)/a_{11}$. The second conditional probability is given by:

$$\mathrm{Pr}_2^r = \Phi\left(\frac{\tau_{2,k} - \hat{y}_2^* - a_{21}\eta_1^*}{a_{22}}\right) - \Phi\left(\frac{\tau_{2,k-1} - \hat{y}_2^* - a_{21}\eta_1^*}{a_{22}}\right). \quad (18)$$

The third conditional probability includes two Gaussian random variables, η_1 and η_2. We use the same η_1^* as before and draw η_2^* from a truncated Gaussian with lower and upper truncation points $(\tau_{2,k-1} - \hat{y}_2^* - a_{21}\eta_1^*)/a_{22}$ and $(\tau_{2,k} - \hat{y}_2^* - a_{21}\eta_1^*)/a_{22}$ so as to have:

$$\mathrm{Pr}_3^r = \Phi\left(\frac{\tau_{3,l} - \hat{y}_3^* - a_{31}\eta_1^* - a_{32}\eta_2^*}{a_{33}}\right) - \Phi\left(\frac{\tau_{3,l-1} - \hat{y}_3^* - a_{31}\eta_1^* - a_{32}\eta_2^*}{a_{33}}\right), \quad (19)$$

the same for the last conditional probability Pr_4^r. Since the computation of (15) is straightforward, we initialize the algorithm by computing it first and then recursively evaluate (18) and (19) and the last probability. Now, if we have R draws of η_1^*, η_2^*, $\eta_3^*$and η_e^*, the simulated joint probability can be approximated by the arithmetic mean of each probability given the r^{th} random draw of η^r:

$$\overline{\Pr}_i(y_1 = j, y_2 = k, y_3 = l, d = 1)_{GHK} = \frac{1}{R}\sum_{r=1}^{R}\left[\Pr_1{\times}\Pr_2^r{\times}\Pr_3^r{\times}\Pr_4^r\right],$$

where $\Pr^r$ refers to the simulated probability given the r^{th} draw of η and $\Pr_1$ is simply (15). Finally, the simulated likelihood function is given by:

$$L_{GHK} = \prod_{i=1}^{N}\overline{\Pr}_i(y_1 = j, y_2 = k, y_3 = l, d)_{GHK}^{w_i}, \tag{20}$$

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

E Imposing Positivity Constraints

We treat the positivity constraints at the level of the Cholesky decomposition of Σ in the GHK algorithm.[15] If $\Sigma = AA'$, A has to be built according to:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a_{21} & \sqrt{1-a_{21}^2} & 0 & 0 \\ a_{31} & a_{32} & \sqrt{1-a_{31}^2-a_{32}^2} & 0 \\ a_{41} & a_{42} & a_{43} & \sqrt{1-a_{41}^2-a_{42}^2-a_{43}^2} \end{pmatrix}. \tag{21}$$

This matrix exists if for every line i starting at line two, the following condition is met:

$$\Sigma_{j=1}^{i-1} a_{ij}^2 < 1 \qquad \forall i > 1.$$

The resulting matrix $\Sigma = AA'$ is automatically positive definite symmetric if this condition is met. One way of imposing this condition is obtained for a matrix of dimension n by the spherical coordinate system defined for the $n-$dimensional Euclidean space 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 $[0, 2\pi]$. The different lower diagonal elements of A, for row i, are given by:

$$\begin{aligned} a_{i1} &= r_i \cos(\omega_1) \\ a_{i2} &= r_i \sin(\omega_1)\cos(\omega_2) \\ &\vdots \\ a_{i,i-1} &= r_i \sin(\omega_1)\cdots\sin(\omega_{i-3})\sin(\omega_{i-2}). \end{aligned} \tag{22}$$

[15]Cappellari and Jenkins (2003, page 290) use a similar decomposition in their `mvprobit` routine. However, they do not impose positivity. If positivity fails, they take the previous value obtained in the optimization, leading presumably to a local optimum.

We first impose the restrictions $r_i = 1$ in order to have unit variances. It remains $n-1$ parameters ω to be estimated freely. At the end of the optimization process, the original parameters have to be reconstructed and the Delta method applied for finding standard deviations.

F Computing Marginal Effects in Ordered Probit Models

Computing marginal effects is a partially unsettled topic in a multivariate setting. The bivariate case was explored in Greene (1996) for probit models. The multivariate probit and ordered probit are treated in Greene and Hensher (2010, Chap 10) and Mullahy (2017). We propose an alternative way of defining marginal effects, adapted to our simulation context.

F.1 Defining Marginal Effects

When Σ is not a diagonal matrix, we have to take into account the interaction between the three equations. So, instead of computing simple derivatives, we would have to evaluate:

$$\frac{\partial \Pr(y_{i1} = j, y_{i2} = k, y_{i3} = l | X_i, \widehat{\Theta})}{\partial x_i}.$$

This is not a convenient task, first because it involves a multidimensional integral and second because there would be $4^3 = 64$ different values to compute, depending on the configuration of events. We prefer to consider the derivative of a conditional probability:

$$\frac{\partial \Pr(y_{i,m} = j | y_{i,-m}, X_i, \widehat{\Theta})}{\partial x_i}.$$

This means that we are considering the probability of event $y_{i,m} = j$ as described by equation m, conditionally on the realization of other events as they appear in the observed sample and as they are explained by the remaining equations indexed here schematically by $-m$. $\widehat{\Theta}$ represents all the parameters at their estimated value and X_i the vector of exogenous variables for individual i. x_i is the particular variable for which we want to compute the derivative. To compute an average marginal effect (AME), we average the n obtained values over the sample, integrating over the empirical distribution of the observed sample:

$$\begin{aligned} AME &= \mathrm{E}_{x_i, y_{i,-m}} \frac{\partial \Pr(y_{i,m} = j | y_{i,-m}, X_i, \widehat{\Theta})}{\partial x_i}, \\ &\simeq \frac{1}{n} \sum_i \frac{\partial \Pr(y_{i,m} = j | y_{i,-m}, X_i, \widehat{\Theta})}{\partial x_i}. \end{aligned}$$

This average will give us, for each equation, an average marginal effect of the influence of a change in a variable for each of the four categories of opinion. The question is now to adapt the GHK simulator for evaluating the average marginal effect, when taking the parameters at their MLE values.

F.2 Evaluation of Marginal Effects Using Simulation

When $m = 3$, the conditional probability we are looking for is directly estimated when using the GHK simulator. It can be deduced from (17), giving:

$$\begin{aligned}
\Pr(y_{i3} = l | y_{i1} = k, y_{i2} = j, X_i + \Delta x_i, \widehat{\Theta}) = \\
\int\int \Bigg(\Phi[\frac{\widehat{\tau_{3,l}} - \hat{y}_3^*(X_i + \Delta x_i) - \widehat{a_{31}}\eta_1 - \widehat{a_{32}}\eta_2}{\widehat{a_{33}}}] \\
- \Phi[\frac{\widehat{\tau_{3,l-1}} - \hat{y}_3^*(X_i + \Delta x_i) - \widehat{a_{31}}\eta_1 - \widehat{a_{32}}\eta_2}{\widehat{a_{33}}}] \Bigg) d\eta_1 d\eta_2. \quad (23)
\end{aligned}$$

So that the derivative we are looking can be approximated by:

$$\begin{aligned}
\frac{\Pr(y_{i3} = l | y_{i1} = k, y_{i2} = j, X_i + \Delta x_i, \widehat{\Theta})}{\Delta x_i} \\
- \frac{\Pr(y_{i3} = l | y_{i1} = k, y_{i2} = j, X_i, \widehat{\Theta})}{\Delta x_i}. \quad (24)
\end{aligned}$$

In (23), the double integral will be evaluated by the GHK simulator. This means that we must draw R values for η_1 and η_2 according to a truncated normal density. The bounds are determined by the observed sample values $y_{i1} = k$ and $y_{i2} = j$. When computing the average marginal effect, we are going to average over all the sample values of these bounds.

Of course the conditional probability (23) relies on the way we have decomposed the joint probability during the estimation process. In particular, it is specific to the particular Choleski decomposition. For each conditional probability (here we have three conditional probabilities because we have three equations), we must compute a new Choleski decomposition.