Intergenerational Mobility Pathways: Evidence From a Long Panel from Rural China

Albert Park and Xianqiang Zou*

November 2017

Abstract

This paper documents the basic facts about the intergenerational mobility in rural China setting. We also account for the intergenerational transmission by de- composing the elasticity to the part explained by cognitive skills, noncognitive skills, and education. Utilizing an unique data set GSCF, this study is the first to provide empirical evidence in developing countries rural setting. We find that the intergenerational education transmission coefficient is 0.391 in our analysis, which is a little bit lower than the estimation from other papers. Our estimated intergenerational income elasticity is 0.099, which indicates the intergenerational mobility is quite high within this rural area. The decomposition results show that the role of cogni- tive skills is important for accounting for intergenerational persistence of education and income, noncognitive skills is relatively less important, and health variables play very small role in intergenerational education transmission, play modest role in intergenerational income transmission. The decomposition results conditional on education reflects cognitive skills affect income partly because of their influence on education.

*Park: Department of Economics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: albertpark@ust.hk); Zou: Department of Economics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: xzou@ust.hk). Data collection for the Gansu Survey of Children and Families was supported by grants from The Spencer Foundation Small and Major Grants Programs (wave 1); NIH (1R01TW005930-01 and 5R01TW005930-02) and the World Bank (wave 2), ESRC/DFID [RES167250250] (wave 3), and the Hong Kong Research Grants Council General Research Fund [692713] (wave 4) which also supported the research for this paper.

1. Introduction

Intergenerational mobility is one measurement for the degree of family's socioeconomic status affect the outcome of children. A strong association between father's socioeconomic status and children's socioeconomic status means low mobility among the society. It is an indicator for equality of opportunity, and relate to the intergenerational efficiency and intergenerational equality.

In this chapter, we utilize a unique long panel data extending 15 years from rural Gansu province to study the intergenerational mobility in rural China con- text systematically. We estimate both the intergenerational education correlation and intergenerational income correlation. For the estimation of intergenerational income correlation, we estimate both the traditional intergenerational income elasticity and the intergenerational income rank-rank slope (Chetty et al., 2014).

To understand the intergenerational mobility more thoroughly, we explore the channels through which family outcome or father's outcome affect children's educational attainment and labor market outcome. Recently, large literature focus on the importance of the cognitive skills and noncognitive skills measurement to explain children's labor market outcome. Utilizing the good measurement on cognitive skills and noncognitive skills over years in our data set, we decompose the intergenerational persistence to the role played by the cognitive skills and noncognitive skills. We decompose the intergeneration income elasticity firstly without conditional on children's educational attainment, we examine the direct role of cognitive and noncognitive skills to account for the intergenerational income elasticity. We then do the decomposition conditional on children's education, and explore the conditional contribution of cognitive skills and noncognitive skills. We also study the conditional intergenerational educational correlation and conditional intergenerational income correlation, the also decompose the conditional correlation to be role played by cognitive and noncognitive skills. The gender difference of the decomposition is also explored.

We find that the intergenerational education transmission coefficient is 0.344, which means the modest relationship between father's educational attainment and children's educational attainment. Our estimation of intergenerational income elasticity is 0.111, and the intergenerational income rank-rank slope is 0.103, which is quite low compared with the estimation of intergenerational income elasticity in the US or in urban China. The transitional matrix between family net income and children's income is consistent with the low intergenerational income correlation.

We decompose the estimated intergenerational education transmission coefficient to the part explained by cognitive skills and noncognitive skills. The decomposition results show that the cognitive skills play an important role to account for the intergenerational education transmission, while the role of noncognitive skills are also important but smaller than cognitive skills. The decomposition of intergenerational income elasticity without conditional on children's education attainment shows the similar role played by cognitive skills, but the noncognitive skills are less important in this analysis. The decomposition conditional on education suggests children's education explains a significant proportion of intergenerational persistence. The part played by cognitive skills drops much more than noncognitive skills, which indicate the stronger correlation between cognitive skills and children's educational attainment. The conditional intergenerational education correlation and conditional intergenerational income correlation is quite similar to the results above, and the decomposition patterns are also similar.

There exists a gender difference in the decomposition of intergenerational income correlation. Using male subsample, we find that education still explain certain proportion of the intergenerational correlation given the cognitive skills and noncognitive skills, cog- nitive and noncognitive skills plays a similar role. But for female subsample, education plays no role in the intergenerational correlation transmission. And cognitive skills is more important than noncognitive skills to explain the intergenerational correlation. Similar patterns exist if we do the rank rank slope decomposition for both gender. Our expla- nation for this is that the girls are selectively to go to the school, cognitive skills are more likely to be highly correlated with the girl's education; while boys get more family resource, the education level has less correlation with the cognitive skills.

Our paper builds upon the literature of intergenerational mobility. This literature starts from the seminal work by Becker and Tomes (1979), Becker and Tomes (1986). In these papers, they emphasize the role of human capital investment on children's outcomes, and highlight the impact of credit constraints on human capital investment. Lot of papers, Mulligan (1999), Mulligan and Grawe (2002), Han and Mulligan (2001) and Solon (2004) extend the above framework to other contexts but mostly are still emphasize on the role of education. Most of the empirical papers on intergenerational mobility focus on documenting the basic facts about intergenerational mobility in different countries, aiming to provide some policy implications to reduce the intergenerational persistence in low mobility countries. Solon (1999), Black and Devereux (2011) and Corak (2013) provide an excellent summary of this literature. Recently, Chetty et al. (2014) uses tax data to analyze intergenerational mobility in the USA across regions, and finds significant variation of intergenerational mobility across states. They also examine what factors are correlated with intergenerational transmission across states.

The literature on intergenerational mobility in China consists mainly of empirical papers trying to estimate the intergenerational income elasticity in China, they focus on urban China due to the data scarcity for rural China. Gong et al. (2012) estimate the intergenerational income elasticity is 0.63 for father and son using UHEES 2004, Deng et al. (2013) estimate the intergenerational income elasticities to be 0.477 using 1995 CHIPS, 0.508 using CHIPS 2002. There is no estimates of intergenerational income elasticities in rural China. Our paper filled this gap using an unique long-panel data set we collected ourselves in Rural China.

Recently, the literature on noncognitive skills is emerging. It's well known before that the cognitive skills plays an important role in the labor market. Heckman et al. (2006) study the impact of cognitive skills and noncognitive skills in the labor market and social behavior. And they found that the noncognitive skills is as importance as cognitive skills in the labor market. Another strand of this literature focus on the formation of cognitive and noncognitive skills. Cunha and Heckman (2008) found that the parental investment are generally more effective in the formation of non-cognitive skills. Most of the paper on this literature using data from developed countries. Few papers study the importance of noncognitive skills in developing countries, especially for the rural area. Using the first three waves of the GSCF, Glewwe et al. (2013) find that early cognitive and noncognitive skills influence educational attainment but do not affect early wages independently of years of education. Due to the important of cognitive and noncognitive skills in the labor market documented in the literature, we introduce the cognitive and noncognitive skills as potential channels through which intergenerational persistent.

The only paper in economics literature which did similar thing is Blanden et al.

(2007), they use the British data to examine the role of cognitive and noncognitive skills in the intergenerational transmission. Our paper is a good complementary to that paper. Using a unique long panel data set in rural China, we also find that cognitive skills play more important role than noncognitive skills for the intergenerational transmission.

Overall, this paper contributes to two strands of literature. The first literature is about the estimation of intergenerational mobility in developing countries. Our paper is the first paper trying to document the intergenerational mobility in rural poor areas systematically. The second literature is about the pathways or mechanisms of the intergenerational mobility. Potentially, the father's socioeconomic status tends to affect the children's socioeconomic status through multiple mechanisms. We have rich data on children's cognitive and noncognitive skills, which enable us to decompose the intergenerational transmission through cognitive skills and noncognitive.

The remainder of our paper is structured as follows. In the next section, we review the literature on intergenerational mobility and on the role of cognitive skills and noncognitive skills to explain the labor market outcome. Section 3 introduces the data set we use in the analysis, and shows the summary statistics. Section 4 presents the empirical results. We conclude at the last section.

1 Empirical Framework

The basic regression equation to estimate the intergenerational education correlation is

$$\mathrm{edu}_c = \alpha + \beta \mathrm{edu}_f + \mathrm{Controls} + s \quad (1)$$

The basic regression equation us to estimate the intergenerational income correlation is

$$y_c = \alpha + \beta y_f + \text{Controls} + s \tag{2}$$

In equation (1), edu_c is children's year of education, edu_f is father's year of schooling, β is the intergenerational education correlation. In equation (2), y_c is log children's income or children's income rank, y_f is log family net income or family net income rank. β measures the intergenerational correlation, we estimate both traditional intergenerational income elasticity and rank-rank slope (Chetty et al., 2014). Control variables include children's birth year dummies, father's age, father's age square, and gender dummy variable. Besides the estimation of intergenerational mobility, we seek to uncover the channels through which family's socioeconomic status or parents' outcome affect children's education outcome and labor market outcome. The intergenerational transmission happens through various channels. The parental educational investment is commonly mentioned in the literature. Besides the educational attainment, the cognitive and noncognitive skills or personality are also very important for the children's outcome. And this potentially are also transmits through the generations. Blanden et al. (2007) measures the intergenerational mobility and accounts for the intergenerational persistence using the mediating factors including cognitive skills, noncognitive skills, educational attainment, and labor market attachment using British data sets. We adopt a similar approach with this paper. To decompose the intergenerational education correlation into role played by cognitive and noncognitive skills, we firstly examine the impact of father's education on each skill variables,

$$H_i = \alpha_{1i} + \lambda_i \text{edu}_f + u_{1i} \tag{3}$$

And then we examine the impact of each skill variables on children's education

$$\text{edu}_c = \alpha_2 + \sum \rho_i H_i + u_2 \tag{4}$$

The intergenerational education correlation can be decomposed into:

$$\beta = \sum \rho_i \lambda_i + \frac{cov(u_2, edu_f)}{Var(edu_f)} \tag{5}$$

To explore the role of each skill measurement on the intergenerational income correlation, we first examine the impact of family net income or family net income rank on each skill variable.

$$H_i = \alpha_{1i} + \lambda_i y_f + u_{1i} \tag{6}$$

And then we study the relationship between each skill variable and children's income or children's income rank,

$$y_c = \alpha_2 + \sum \rho_i H_i + u_2 \tag{7}$$

The Intergenerational income correlation can be decomposed into the

$$\beta = \sum \rho_i \lambda_i + \frac{cov(u_2, y_f)}{Vary_f} \tag{8}$$

However, it's well known that skill measurement also affects the educational performance and attainment, the above decomposition method shows the unconditional influence of these skill measurements on the intergenerational persistence, to account for the channel non-cognitive skills also affect the educational performance and attainment, we add the children's educational variable as control variable.

We examine the influence of each skill variable on children's outcome.

$$y_c = \alpha_1 + \delta_i H_i + \pi edu_c + u_3 \tag{9}$$

And then explore the relationship between log family income and children's educational attainment.

$$edu_c = \alpha_2 + \gamma y_f + u_4 \tag{10}$$

The intergenerational income correlation can be decomposed to

$$\beta = \sum \delta_i \lambda_i + \pi\gamma + \frac{cov(u_3, lnY_f)}{var(Y_f)} \tag{11}$$

The conditional contribution of each skill variable is $\delta_i\lambda_i$.

Another question we want to ask is whether parents' education play a role when we study the relationship between family income and children's income, and whether con- trolling for family income would affect the relationship between father's education and children's education. If controlling for father's education does not affect the intergenera- tional income correlation, which means

There is a significant partial correlation between family income and father's education. We firstly estimate the intergenerational education correlation conditional on family income and also estimate the intergenerational income correlation conditional on father's education.

$$y_c = \alpha + \beta_1 y_f + \beta_2 \text{edu}_f + \epsilon \tag{12}$$

And then we explore the impact of log family net income and father's education on skill variables

$$H_i = \alpha_{1i} + \lambda_{1i} y_f + \lambda_{2i} \text{edu}_f + u_{1i} \tag{13}$$

The relationship between skill variables and children's education, children's income are shown as above,

$$y_c = \alpha_2 + \sum \delta_i H_i + u_2 \tag{14}$$

Hence, The conditional intergenerational correlation can be decomposed to

$$\beta_1 = \sum \delta_i \lambda_{1i} + \frac{cov(u_2, \hat{y_f})}{var(\hat{y_f})} \tag{15}$$

$$\beta_2 = \sum \delta_i \lambda_{2i} + \frac{cov(u_2, \hat{\text{edu}}_f)}{var(\hat{\text{edu}}_f)} \tag{16}$$

2 Data and Summary Statistics

The data set used in this paper is the Gansu Survey of Children and Families (GSCF), a panel study of rural children conducted in Gansu province, China. Gansu, located in northwest China, is one of the poorest and most rural provinces in China. Figure 1 shows the sample location of GSCF. The sampling county spread widely across the province. Our sampling strategy ensure our samples are representative.

The first wave of the survey, conducted in 2000, surveyed a representative sample of 2000 children aged 9-12 in 20 rural counties, as well as their mothers, household heads, teachers, principals, and village leaders. All but one of these 2000 children have complete information in the first wave. The second wave, implemented in 2004, re-surveyed almost all of these children at age 13-16, and also added a survey of their fathers; 93.6% of the original sample, or 1872 children, were re-interviewed in the second wave, and 1773 completed achievement tests that were administered in their schools.

The third wave, completed in early 2009, re-interviewed the original sample children (who at that time were young adults) during Spring Festival, a peak time for young people to visit their parents' homes in rural China. If the sampled individual was not available, parents were asked questions about their child's education and employment status; however, skill measures could be collected only from the children who had returned to their parents' homes. Of the original 2000 children, 1437 (72% of the original sample) were interviewed directly and completed skill tests in this wave. In addition, information was collected for an additional 426 sample children by surveying their parents.

The fourth wave of the survey, conducted in 2015, age of the sample children should be between 24-27, most of them have entered the labor market. We collected detailed job and income information for each child. Combined with the household questionnaires, and the parents' questionnaires. We could get the income information for the household and also imputing the parent's income. Utilizing the panel data set, we could average the family income over the first three waves, which provide a relatively good measurement for family socioeconomic status.

We focus on the measurement of intergenerational income elasticity and intergenerational education transmission. The information on education is more accurate and less noisy than the income data. For parent's education, we use their education information from the first wave as measurement, and for the sample with missing information on parent's education, we replace the missing value with the parents' education information in the wave 4. We also across tabulate the parent's education measurement in both waves, although there are observations report different education level in different years, the inconsistency case is minor. We believe the measurement error on education is not a big issue for education information. For income measurement, we impute family income, income per capita, wealth per capita for the first three waves. We take the average value of these measurements for the first three waves to measurement family socioeconomic status. We also impute father's income by multiply the family income and labor share of father measured by the father's working time over total working time of the family mem- ber on farming and plus the father's migration income. The family income or father's income is zero for some observations in certain waves in our sample set. And the zero income might reflect both not working or the non-response. We use the fixed effect model to predict the family income or father's income if it is zero in certain years but positive in other years. Children's income information is from wave 4, in which we ask about the children's income from the current job, or previous jobs if there are not working currently. The summary statistics of each variable are shown in Table 1. We have 848 observations with complete information, both on income, education, health and various ability measurement. The observations we suspicious not the same kids across all four waves are dropped. Income measurements all take the log to make the distribution more likely to be normal. From Table 1, log children's income is higher than log family income or log father's income, this is not surprising since the log of family income and log of father's income measurements are the average of income in the first three waves, which were implemented in 2000, 2004, 2007, respectively. The wage and income grow significantly over these years in China. The average education year for children is also much higher than the average years of schooling of fathers and mothers. The potential experience equals to age minus education year minus 7. We also provide summary statistics for some other characteristics of the sample kids, such as hukou, marriage status, health measurement, height and weight, etc.

Across all four waves, we have rich information of cognitive skills and noncognitive skills about sample kids. In Table 2, we show all the measurement in each wave for cognitive and noncognitive skills (Glewwe et al., 2013). The definition of cognition is all forms of knowing and awareness, such as perceiving, conceiving, remembering, reasoning, judging, imagining and problem solving. The measurement of cognitive skills in the Gansu survey contains a general cognitive skills test in wave 1, a Chinese and math achievement tests in waves 1 and 2, A literacy (life skills) test which includes in waves 2 and 3. Noncognitive skills can be defined as patterns of thoughts, feelings and behavior that affect social interactions with others. In the Gansu survey, the measurement of noncognitive skills includes the internalizing behavior, externalizing behavior and educational aspiration in both wave 1 and wave 2. Internalizing behavior problems are intrapersonal in nature, such as anxiety, depression and withdrawal. Externalizing problems are interpersonal in nature and characterized by destructive behavior, impulsivity, aggression and hyper-activity (Achenbach and Edelbrock, 1978). Child psychology research suggests that environments that destabilize a child's sense of self control over his or her life can in- crease internalizing problems (Chorpita and Barlow, 1998; Dearing et al.,2006), while environments that impede a child's self-regulatory efforts, or the presence of anti-social role models, can increase externalizing problems (Evans, 2004). The questions we asked to measure the internalizing and externalizing behavior is listed in Table 3. The educational aspiration is measured by the children's desire to go to college. The self-esteem (Rosenberg Self-Esteem Scale) and depressive symptoms (Center for Epidemiological Studies Depression Scale, CES-D) in wave 3 and wave 4. And we also compute the corresponding self-esteem scale and depressive symptoms scale for wave 1. We also have the big five personality test in wave 4. The five factors are defined as openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. We adopt the widely used Big five inventory which contains 44-item inventory that measures an individual on the Big Five Factors (dimensions) of personality.

In the analysis, the cognitive skills measurements are based on examination test. We measure the cognitive skills by summing up all the scores on each question in the test. For the noncognitive skills measurements, we use the item response theory from psychometrics literature to fitting the model and predict the latent ability measurements. The Item response theory is widely used to score subjects on their abilities, attitudes, or other latent traits. This method assumes the probability of an keyed response to an item is a function of latent trait or item parameters, and then estimate the latent trait by fitting the model to data. To make the estimation results more comparable, we standardize all the cognitive skills and noncognitive skills measurements. Table 3 shows the correlation between all the ability measurement in each year. For cognitive skills in each wave, they are positively correlated with each other as expected. For the correlation between cognitive and noncognitive ability, all of the correlation has the right sign. Since for each year, we have multiple dimensions on cognitive skills and noncognitive skills, we also use the principal component analysis to reduce the dimension. Specifically, principal component factor method is used to analyze the correlation matrix, and the communalities are assumed to be 1 in the analysis. This principal component analysis can be achieved in the Stata directly. We have one factor variable combining the key information of ability and noncognitive skills for each wave, except for the noncognitive skills measurement in wave 4, we use two measurements to capture it since we have seven variables in this wave. Table 4 shows the correlation over time for skill measurement. Both cognitive and noncognitive skill measurement have high correlation over time. The internalizing scale is still positively correlated but not significant.

2. Empirical Results

In this part, we first show the basic results for the estimation of intergenerational correlation, for both intergenerational education correlation and intergenerational income correlation. For income measurement, we show both the intergenerational income elasticity and rank rank slope. And the intergenerational income elasticity is estimated using the sample set drop the top and bottom 1% sample. The second step is to decompose the intergenerational education correlation , intergenerational income elasticity and income rank-rank slope. We also show various robustness check for the decomposition pattern. Moreover, we show the intergenerational educational correlation conditional on log family income and intergenerational income correlation conditional on father's education. At last, we show the gender difference in the decomposition results.

4.1 Intergenerational Mobility: Basic facts

Table 5 shows the intergenerational education correlation estimation results, for the whole sample, the intergenerational education correlation is 0.391. This estimation results are comparable to other estimation results in the literature. The estimation correlation is 0.351 and 0.422 for female and male, respectively.

Table 6 shows the intergenerational income correlation estimation results. For the first three columns, we drop the top and bottom 1% log family net income observations, and the sample size drop to 832. For the traditional intergenerational income elasticity estimation, we control for children's birth year dummies, father's age, father's age square, and gender dummy variable. The estimation results show the intergenerational income mobility is quite high in this area, the intergenerational income elasticity is as low as 0.099, 0.046, and 0.140 for whole sample, female subsample and male subsample, respectively. If we use the rank-rank specification as in Chetty et al. (2014), the rank-0rank slope is 0.091, 0.052 and 0.122 for whole sample, female subsample, male subsample, respectively, which is quite similar to the intergenerational income elasticity.

We also use the transitional matrix between family net income and children's income to show the income mobility pattern to see whether it is consistent with the estimation results. In Table 7, we can tell that the numbers in this table are all around 25, which indicate the mobility is quite high in this area.

4.1 Decomposition of Intergenerational Education Correlation

Table 8-11 shows the impact of father's education on cognitive, noncognitive skills and health. All cognitive and noncognitive skill measurement are standardized in the table. Table 8 shows that father's years of schooling have a positive and statistically significant effect on children's cognitive skill measurements in each year. Table 9 shows the impact of father's year of schooling on the noncognitive skill measurement before 2009. The skill measurements are all standardized to make the estimation results comparable. The father's education has a negative impact on children's internalizing and externalizing behavior in 2000. For measurement in 2004, the results indicate that the father's education has positive impact on internalizing and externalizing behavior, although it is not significant. The impact of father's education on educational aspiration are both positive in Wave 2000 and Wave 2004, which means the children with better educated fathers have more desire to go to the college in the future. And the impact on Rosenberg self-esteem scale measurement and depression measurement in 2009 are in expect, though the impact on depression measurement is not statistically significant. Table 10 shows the impact on noncognitive skill measurements in last wave. All the coefficients have expected sign, though the impact on extraversion measurement and openness to experience is not statistically significant. Overall, we find that the father's education does have significant impact on children's noncognitive skills.

Figure 2 shows the results of the children's education determinants regression, the estimation coefficients are shown as dots in the figure. The coefficients are all positive, and statistically significant except the coefficient for math achievement score. The literacy score in 2009 is dropped from the regression since at 2009, some students have already finished the education. And the students who are still in the schools make them more likely to achieve better score in the literacy test in 2009, which makes the estimation have a sample selection bias. In general, we could only keep the sample who are still in the school by 2009, however, that would make the sample size shrink a lot. Hence, we drop the 2009 literacy measurement in the education regression. Consider the noncognitive skills, the motivation going to college variable in 2004 have strong impact on children's educational attainment, and also Rosenberg self-esteem scale measurement. The noncognitive skills in 2015 are not included in the regression because at 2015, almost all of the sample observations have already finished education. For the health variables, the coefficients are mostly around 0, except for the weight variables in wave 2009.

Table 12 shows the decomposition of intergenerational education correlation results. Column (1) shows the coefficients explained by the corresponding variables in the left-hand side. The number in the column are obtained by multiplying the impact of father's education on each skill variables coefficients in Table 8-11 and the coefficients in the corresponding children's education regression in Figure 1. Column

(2) show the percent- age of intergenerational education transmission explained by the corresponding variables. Overall, the cognitive ability variables explain 28.2% of the intergenerational education transmission, Noncognitive ability explains 14.0%, and health variables explain 1.9% of the total transmission coefficients.

4.2. Decomposition of Intergenerational Income Correlation

The impact of log family net income on the skill variables are presented in Table 13-16. Table 13 shows that log family net income has a positive and statistically significant effect on children's cognitive skill measurements in each year and of course on children's educational attainment measured by children's years of schooling. Table 14 shows the impact of log family net income on the noncognitive skill measurement. The skill measurements are all standardized to make the estimation results comparable. The log family net income has a negative impact on children's internalizing and externalizing behavior in 2000. For measurement in 2004, the results seem to be weird because the father's education has positive impact on internalizing and externalizing behavior, the impact of log family net income on educational aspiration are both positive but not significant in Wave 2000 and Wave 2004. And the impact on Rosenberg self-esteem scale measurement and depression measurement in 2009 are in expected sign. Table 15 shows the impact on noncognitive skill measurements in Wave 2015. All the coefficients have expected sign, though the impact is not significant in general. Overall, we find that log family net income has a positive and significant impact on children's cognitive skills, but we do not find a modest impact of log family net income on children's noncognitive skills. Table 16 shows the impact of family net income on health variables, the impact on birthweight, height and weight are statistically significant, which means children from relatively rich family enjoy better health condition, they have higher birthweight, they grow up taller, and gain more weight than children from relatively poor family.

Figure 3 shows the income regression results. Except for the ability measurement, children's birth year dummy variables, father's age, father's age square in the regression, the gender dummy are also added when we run the regression with whole sample. In general, we find that most variables have coefficient around zero in this regression, this might be because of there are too much variables related to each other in one regression, after conditional on other variable, one variable can hardly play significant role to explain the final income variable. We run the income regression with and without conditional on children's educational attainment variable, since this skill measurement also have a impact on children's educational attainment, one channel for these variables to affect the children's income variable is through children's education attainment. Hence, it would be interesting to see the results with and without conditional on children's educational attainment.

Table 17 shows the decomposition results for intergenerational income elasticity. The number in all odd columns are obtained by multiplying the coefficients of the impact of log family net income on each ability measurement and the coefficients in the corresponding children's income regression. The number in the even columns are the percentage of intergenerational income elasticity explained by the corresponding variables in the corresponding specifications. Columns (1)-(2) don't add children's educational attainment into the specification, without conditional on children's education, cognitive skill accounts for 16.8% of intergenerational income elasticity while noncognitive skills accounts for 0.2% of intergenerational income elasticity, and health variables explain 13.7% of intergenerational income elasticity, they have in total explained 30.7% of the intergenerational income elasticity. Columns (3)-(4) are the decomposition results conditional on children's education. The education variable explains a modest intergenerational income elasticity, around 6.4%. Both the role of cognitive ability and noncognitive ability drops when we add children's educational attainment variables, which indicate the skill variable affect the children's income through the education channel, although the impact is minor. In total, we explained 33.7% of intergenerational income elasticity if we add every mediating variables into the regression.

Table 18 shows the decomposition of rank-rank slope. The general pattern is quite similar with the decomposition of traditional intergenerational income elasticity as shown in Table 17. The cognitive skills explain a larger proportion of intergenerational income correlation compared with noncognitive skills, health variable plays modest role in intergenerational income persistence. One potential concern for our results is that when we add all the skill variables into the specification, the overall proportion explained by noncognitive skill are negative, which is mainly caused by the negative role played by externalizing behavior in Wave 2000 and Openness in Wave 2015. These two measurements might have a severe measurement error issue, and cause the odd estimation and decomposition results.

4.2. Robustness Check

Since lots of ability measurement are not significant in the children's education or income regression, and to explore the differentiate impact of cognitive and noncognitive skills measurement over time, we reduce the dimension of ability measurement using the princi- pal component analysis. We create a variable synthesizing cognitive skills, noncognitive skills and health information in each wave. The procedure of principal component analysis is that we firstly reverse the ability measurement which are indicate bad behavior, such internalizing, externalizing, depression and Neuroticism in each wave. And then use the principal component factor analysis to get the factor variables, each factor variable can be seen as a linear combination of the corresponding noncognitive skills measurement. For the noncognitive skills measurement in 2015, since there are seven measurements in total, hence, we use two factor variables combining the information from the seven variables. The two factor variables are orthogonal to each other. For health variables, for education decomposition, since the number of variables are relatively small, we use the one factor variable to combine all the information, for income decomposition, since we include more health measurement variables, we use two factor variables to combine the information, which is shown as Hlhinc1 Hlhinc2.

Table 19 shows the decomposition of intergenerational education correlation using the reduced dimension of skill variables. For the intergenerational education decomposition, the factor variables combine the cognitive skill measurement we have in each wave explained 24.2% of the total intergenerational education persistence, compared with 3.2% of factor variables indicating the noncognitive skill measurement we have in each wave health variables explain 1.8% of the intergenerational education transmission. In total, they explained about 29.2% of the total intergenerational education transmission coefficient. This is smaller than the decomposition results using the whole sets of cognitive skills variables, noncognitive skills variables and health variables because this factor variables only capture the major information from those variables, some secondary information lost due to the decomposition, hence, the explained role of these factor variables tend to be lower. However, this decomposition results show the similar pattern as above intergenerational education transmission decomposition. The cognitive skills explained more intergenerational education transmission compared with the noncognitive skills. and health variable play minor role in intergenerational education transmission.

The decomposition of intergenerational income elasticity in Table 20 also show the similar but slightly different pattern as the decomposition results from Table 17. Without conditional on children's education, the cognitive skill factor variables explain a 4.2% of the intergenerational income elasticity, the noncognitive skill explains 2.9%, health variables explain 11.4%, and in total, they explain 18.5% of the income persistence across generations. The role of cognitive ability drops a lot in results, this might because of we drop some key information from cognitive ability during the principal component analysis procedure. The decomposition results conditional on education also show similar pat- tern. Cognitive skill captures certain part of variation explained by children's educational attainment, they provide the same source of variation to explain the children's labor market outcome. However, the role of education becomes negative in this case.

4.2 Gender Difference of Decomposition of Intergenerational In- come Elasticity

Table 21 shows the gender difference of decomposition of intergenerational income elasticity, we find that for male subsample, Cognitive and noncognitive skills plays similar role, and health plays minor role for the intergenerational income persistence. For female subsample, cognitive skills is much more important than noncognitive skills to explain the intergenerational income correlation. And health variables play a very big role for intergenerational income persistence for female subsample, it explains 47.8% of intergenerational income elasticity. Similar patterns exist if we do the rank-rank slope decomposition for both gender. We propose a potential explanation for the findings. The girls are selectively to go to the school, cognitive skills are more likely to be highly correlated with girls' education; while boys get more family resource, the education level has less correlation with the cognitive skills.

5. Conclusion

This paper documents the basic facts about intergenerational mobility in rural China, using a unique long-panel data from Gansu Province, China. We estimate both intergenerational education transmission and intergenerational income elasticity systematically. We find that the intergenerational education transmission is around 0.391, this result is quite similar with the estimation of intergenerational education transmission in other papers in the literature. However, our estimation of intergenerational income elasticity is 0.099, which is quite low compared to the other estimation. We try to average the family income in three waves extending over 7 years when parents of most sample kid are in the age between 30-50 to eliminate the impact of measurement error. And the low elasticity might be explained by the facts that most children migrate out to work, which makes their income less likely to depend on the family income. One point worth to point out is that we estimate of intergenerational mobility only restrict to the rural sample in this certain area, since most families in our sample is poor family in our sample, within this poor families, we know little about whether children from relatively better family gain any advantage than the children from relatively poor family. And our results show that they don't. Children's income doesn't depend on their family income within these poor families. Of course, this result does not necessary hold true if we expand our sample to include more rich families from urban area. We also show the transitional matrix using the family income quartile and children's income quartile, and the results also suggest a high intergenerational income mobility pattern.

The second aim of this paper is to understand the role of cognitive skills, noncognitive skills and health variables in intergenerational transmission. We account for 41.1% of intergenerational education transmission by considering all the cognitive skills, noncognitive skills and health variable measurement we have. And the cognitive skills play much a more important role in the intergenerational education transmission. The cognitive skills accounts for 28.2% of the intergenerational education transmission, while noncognitive skills accounts for 14.0%, health variables play very small role here. We account for intergenerational income elasticity using two frameworks, with and without conditional on children's educational attainment. Without controlling for children's education, the cognitive skills explained 16.8% of total intergenerational income elasticity, noncognitive skills explained 0.2% of the elasticity, health variable explains 13.7% of total intergenerational income elasticity, in total, they explained 30.7% of the elasticity. After controlling for children's educational attainment, we find that children's educational attainment accounts for 6.4% of intergenerational income elasticity, and cognitive skill variables, noncognitive skill variables and health variables account for 14.7%, -0.8%, 13.4%, respectively in the decomposition framework.

Overall, the cognitive skills and noncognitive skills play an important role in account- ing for the intergenerational persistence. Cognitive ability accounts for intergenerational persistence in greater proportion, both in the decomposition of intergenerational education transmission and intergenerational income elasticity. Health variable play very small role in intergenerational education transmission, but play much larger role in intergenerational income persistence.

And we also find a systematical gender difference between the decomposition of inter- generational income elasticity. Our proposed explanation for the finding is that the girls are selectively to go to the school, cognitive skills are more likely to be correlated with girls' education; while boys get more family resource, the education level has less correlation with the cognitive skills. This point needs more systematic study.

It would be interesting to understand the difference in magnitude of intergenerational education transmission and intergenerational income elasticity. The children with well- educated fathers tend to enjoy more education compared with the children with less- educated fathers. The labor market return to education for children's generation is about 5.3% in our sample, and the return to education for father's generation is about 4.2% if we use the log father's income as dependent variable and 3.2% if we use log family net income as dependent variable in the return to education regression. We propose that the small intergenerational income elasticity might due to the high migration rate for the people in these areas. Most of the migration workers work at the manufacturing industry in remote cities. The labor market outcome tends not to be affected by family's income at home that much. It might be interesting to understand the role of migration at the intergenerational mobility systematically and see whether it does help us to explain the small intergenerational income elasticity we found in our date.

One point worth mentioning is that, though the intergenerational income elasticity in our sample is low, the intergenerational income elasticity for the whole country might still be high. In China, the rural urban income gap has always been an issue. And this gap is also reflected in the children's equality of opportunities. We've already show the importance of village fixed effects to explain the variation of children's labor market outcome. Compared with the difference between different villages, the differences of environment between children from rural area and from urban area are more than huge. It is quite difficulty for children from rural area to get high quality education and find a good job in the labor market compared with the children from urban area. It is important for us to understand the role of rural urban gap in explaining the intergenerational mobility in whole China systematically, we leave this to our future research.

References

Aaronson, D. and B. Mazumder (2008). Intergenerational economic mobility in the united states, 1940 to 2000. *Journal of Human Resources 43* (1), 139–172.

Achenbach, T. M. and C. S. Edelbrock (1978). The classification of child psychopathology: a review and analysis of empirical efforts. *Psychological bulletin 85* (6), 1275.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and. *The Journal of Political Economy* , 1153–1189.

Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of labor economics*, S1–S39.

Björklund, A. and M. Jäntti (1997). Intergenerational income mobility in sweden compared to the united states. *The American Economic Review* , 1009–1018.

Black, S. E. and P. J. Devereux (2011). Recent developments in intergenerational mobility. *Handbook of labor economics 4*, 1487–1541.

Blanden, J., P. Gregg, and L. Macmillan (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal 117* (519), C43–C60.

Bowles, S. and H. Gintis (2002). The inheritance of inequality. *The Journal of Economic Perspec- tives 16* (3), 3–30.

Brown, P. H. and A. Park (2002). Education and poverty in rural china. *Economics of Education Review 21* (6), 523–541.

Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal 112* (482), 705–734.

Caucutt, E. M. and L. Lochner (2012). Early and late human capital investments, borrowing constraints, and the family. Technical report, National Bureau of Economic Research.

Chadwick, L. and G. Solon (2002). Intergenerational income mobility among daughters. *American Economic Review* , 335–344.

Chen, N., P. Conconi, and C. Perroni (2013). Multi-trait matching and gender differentials in intergen- erational mobility. *Economics Letters 120* (2), 292–296.

Chen, Y., S. Naidu, T. Yu, and N. Yuchtman (2015). Intergenerational mobility and institutional change in 20th century china. *Explorations in Economic History 58*, 44–73.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics 129*(4), 1553– 1623.

Chorpita, B. F. and D. H. Barlow (1998). The development of anxiety: the role of control in the early environment. *Psychological bulletin 124* (1), 3.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *The Journal of Economic Perspectives*, 79–102.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43* (4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78* (3), 883–931.

Dahl, M. W. and T. DeLeire (2008). *The association between children's earnings and fathers' lifetime earnings: estimates using administrative data*. University of Wisconsin-Madison, Institute for Research on Poverty.

Davies, J. B., J. Zhang, and J. Zeng (2005). Intergenerational mobility under private vs. public education*.
The Scandinavian Journal of Economics 107 (3), 399–417.

Dearing, E., K. McCartney, and B. A. Taylor (2006). Within-child associations between family income and externalizing and internalizing problems. *Developmental psychology 42* (2), 237.

Deng, Q., B. Gustafsson, and S. Li (2013). Intergenerational income persistence in urban china. *Review of Income and Wealth 59* (3), 416–436.

Emran, M. S. and Y. Sun (2014). Are the Children of Uneducated Farmers Doubly Doomed? Farm, Nonfarm and Intergenerational Educational Mobility in Rural China. MPRA Paper 59230, University Library of Munich, Germany.

Emran, S. and Y. Sun (2015, June). Magical transition ? intergenerational educational and occupational mobility in rural China : 1988-2002. Policy Research Working Paper Series 7316, The World Bank.

Ermisch, J., M. Francesconi, and T. Siedler (2006). Intergenerational mobility and marital sorting*. *The Economic Journal 116* (513), 659–679.

Gaviria, A. (2002). Intergenerational mobility, sibling inequality and borrowing constraints. *Economics of Education Review 21* (4), 331–340.

Glewwe, P., Q. Huang, and A. Park (2013). Cognitive skills, non-cognitive skills, and the employment and wages of young adults in rural china.

Goldberger, A. S. (1989). Economic and mechanical models of intergenerational transmission. *The American Economic Review* , 504–513.

Gong, H., A. Leigh, and X. Meng (2012). Intergenerational income mobility in urban china. *Review of Income and Wealth 58* (3), 481–503.

Grönqvist, E., B. Ockert, and J. Vlachos (2010). The intergenerational transmission of cognitive and non-cognitive abilities.

Groves, M. O. (2005). Personality and the intergenerational transmission of economic status. *Unequal chances: Family background and economic success* , 208–231.

Gell, M., J. V. Rodrguez Mora, and C. I. Telmer (2015). The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. *The Review of Economic Studies 82* (2), 693–735.

Han, S. and C. B. Mulligan (2001). Human capital, heterogeneity and estimated degrees of intergenera- tional mobility. *The Economic Journal 111* (470), 207–243.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*

(3), 411–482.

Hsin, A. and X. Yu (2012). Hard skills, soft skills: the relative role of cognitive and noncognitive skills in intergenerational social mobility. *Report 12-775, Feb. 2012, Population Studies Center*.

Jantti, M., B. Bratsberg, K. Roed, O. Raaum, R. Naylor, E. Osterbacka, A. Bjorklund, and T. Eriksson (2006). American exceptionalism in a new light: A comparison of intergenerational earnings mobility in the nordic countries, the united kingdom and the united states.

Lee, C.-I. and G. Solon (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics 91* (4), 766–772.

Lee, S. Y. T. and A. Seshadri (2014). On the intergenerational transmission of economic status. *Unpub- lished manuscript, University of Wisconsin–Madison, Department of Economics*.

Leight, J., P. Glewwe, and A. Park (2015). The impact of early childhood rainfall shocks on the evolution of cognitive and non-cognitive skills.

Lindqvist, E., R. Vestman, et al. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics 3*(1), 101–28.

Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in great britain and the united states since 1850. *The American Economic Review 103*(4), 1109–1137.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *Review of Economics and Statistics 87*(2), 235–255.

Mueller, G. and E. Plug (2006). Estimating the effect of personality on male and female earnings. *Industrial & Labor Relations Review 60* (1), 3–22.

Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of political Economy 107* (S6), S184–S224.

Mulligan, C. B. and N. D. Grawe (2002). Economic interpretations of intergenerational correlations. *Journal of economic perspectives 16* (3), 45–58.

Olivetti, C. and M. D. Paserman (forthcoming). In the name of the son (and the daughter): intergener- ational mobility in the united states, 1850-1930. *The American Economic Review*.

Pitt, M. M., M. R. Rosenzweig, and N. Hassan (2012). Human capital investment and the gender division of labor in a brawn-based economy. *The American economic review 102*(7), 3531.

Restuccia, D. and C. Urrutia (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review* , 1354–1378.

Schwenkenberg, J. M. (2014). Occupations and the evolution of gender differences in intergenerational socioeconomic mobility. *Economics Letters 124* (3), 348–352.

Solon, G. (1999). Intergenerational mobility in the labor market. *Handbook of labor economics 3*, 1761– 1800.

Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. *The Journal of Eco- nomic Perspectives 16* (3), 59–66.

Solon, G. (2004). A model of intergenerational mobility variation over time and place. *Generational income mobility in North America and Europe*, 38–47.

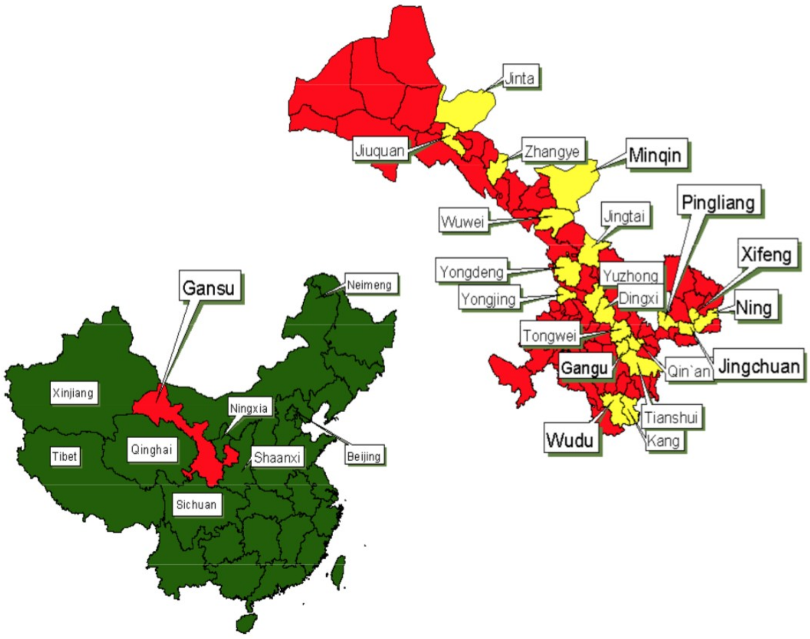

Figure 1: Sample Location of GSCF

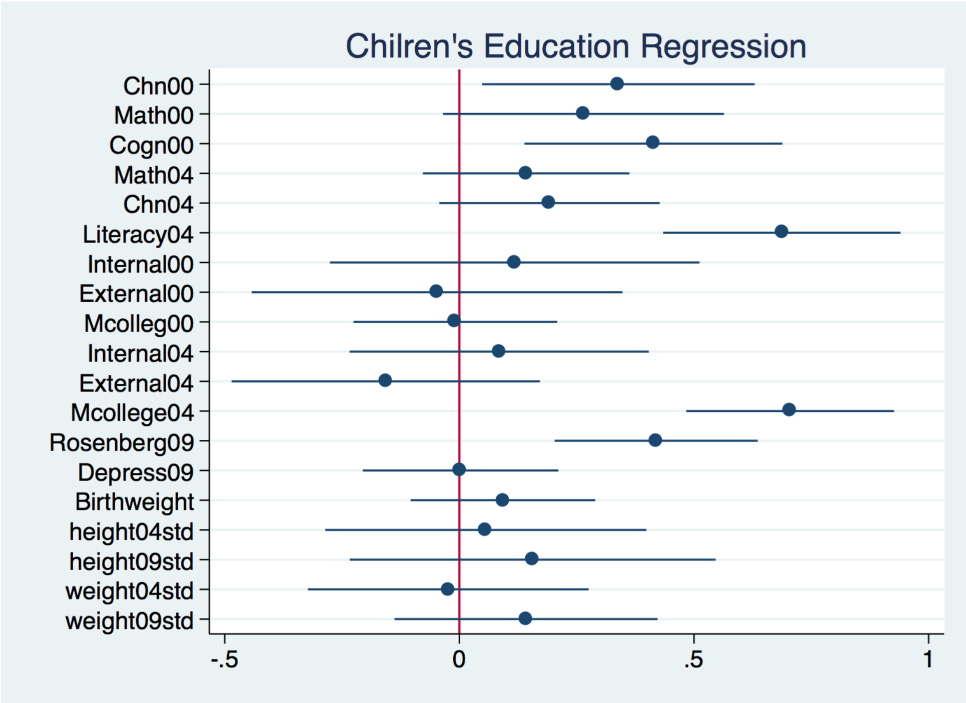

Figure 2: Children’s Education Determinant

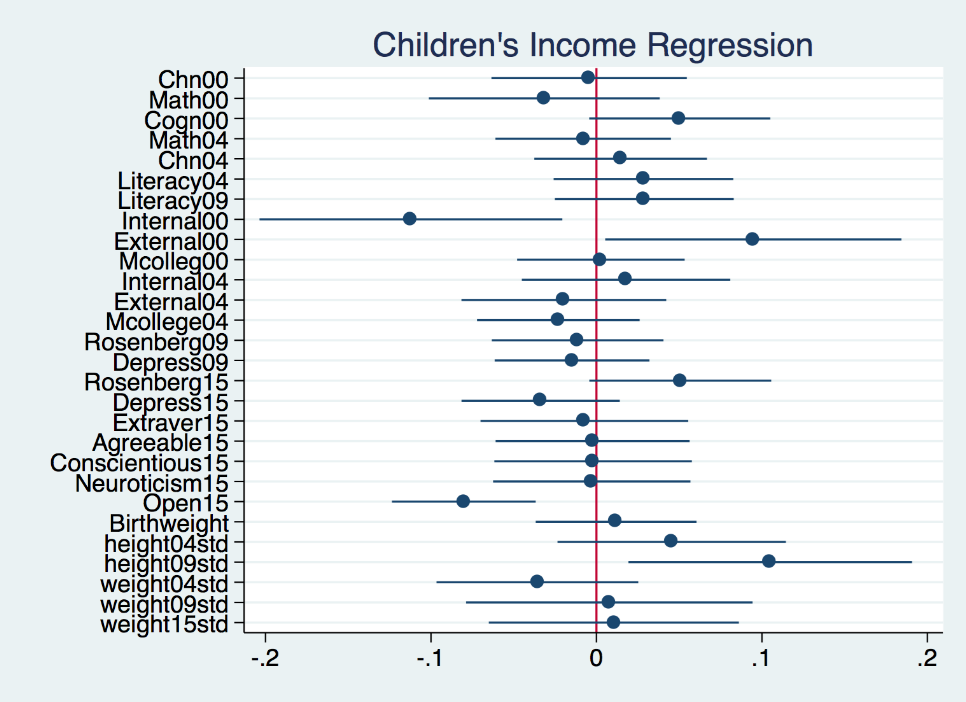

Figure 3: Children’s Income Determinant

Table 1: Summary statistics

Variable	Mean	Std. Dev.	N
Children's income in 2015	39665.129	31849.55	
Family net income average	13324.084	17511.605	
Father's income, three year average	7220.879	8748.751	
Mother's income, three year average	3787.412	5607.909	
Parental income per parent	5504.146	6565.138	
Income per capita, three year average	3324.054	4027.253	
Children gender(male=1)	0.578	0.494	
Children's education year	12.147	3.362	
Age	25.84	1.3	
Potential Exper	6.695	3.456	
Father's age	37.508	5.081	
Years of schooling of Father	6.54	3.109	
Years of schooling of Mother	4.061	3.222	
health00	3.815	1.015	
Health04	3.657	0.86	
Health09	3.813	0.886	
Health15	3.846	0.825	
Birthweight00	5.95	1.011	
Height04	156.046	9.512	
Height09	167.624	7.904	
Height15	168.429	7.402	
Weight04	44.327	8.434	
Weight09	56.008	8.334	
Weight15	60.898	10.695	
N		848	

Note: Data Source: GSCF, Children's income is children's income we interviewed in the last wave, Family net income is the average value for the first three waves. The other children's characteristics, such as gender, Hukou, marriage status, health, height, weight, education level, education year, age and working experience are all taken from wave 4. Years of schooling of father and mother are taken from the first wave in 2000.

Table 2: The Measurement of Cognitive and Noncognitive Skills

Year	Cognitive Skills	NonCognitive Skills
2000 (wave1)	1. Chinese test 2. Math test 3. Cognitive Skills Test	1. Internalizing behavior 2. Externalizing behavior 3. Educational aspiration
2004 (wave 2)	1. Chinese test 2. Math test 3. Literacy Test	1. Internalizing behavior 2. Externalizing behavior 3. Educational aspiration
2009 (wave 3)	1. Literacy test	1. Rosenberg self-esteem 2. Depressive symptoms
2015 (wave 4)		1. Rosenberg self-esteem 2. Depressive symptoms 3. Big Five Personality

Note: this table shows the cognitive and noncognitive skill we have in each wave.

Table 3: Correlation between Cognitive Ability and Noncognitive Ability

2000	Internal00	External00	Mcollege00	Chn00	Math00	Cogn00
Internal00	1					
External00	0.850***	1				
Mcollege00	-0.167***	-0.204***	1			
Chn00	-0.133***	-0.143***	0.0955**	1		
Math00	-0.0836*	-0.0984**	0.123***	-0.000461	1	
Cogn00	-0.258***	-0.265***	0.193***	0.297***	0.201***	1

2004	Internal04	External04	Mcollege04	Math04	Chn04	Literacy04
Internal04	1					
External04	0.753***	1				
Mcollege04	-0.0813*	-0.141***	1			
Math04	-0.0539	-0.0666	0.139***	1		
Chn04	-0.0472	-0.0791*	0.163***	0.490***	1	
Literacy04	0.0204	-0.0516	0.226***	0.251***	0.253***	1

2009	Rosenberg09	Depress09	Literacy09
Rosenberg09	1		
Depress09	-0.308***	1	
Literacy09	0.215***	-0.0467	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2015	Rosenberg15	Depress15	Extraver15	Agreeable15	Conscientious15	Neuroticism15	Ope
Rosenberg15	1						
Depress15	-0.334***	1					
Extraver15	0.414***	-0.217***	1				
Agreeable15	0.405***	-0.216***	0.358***	1			
Conscientious15	0.443***	-0.208***	0.417***	0.547***	1		
Neuroticism15	-0.403***	0.432***	-0.342***	-0.301***	-0.361***	1	
Open15	0.103**	-0.0310	0.138***	0.124***	0.144***	-0.0646	1

Note: this table shows the correlation between cognitive and noncognitive skills for each year separately. The Correlation between chn00 and math00 is none because we randomly choose about half students take math exam, and the other half take Chinese in 2000. All the variables are standardized to make the estimation results more comparable. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Correlation Over time for Skill Measurement

	(00, 04)	(04, 09)	(00, 09)	(09, 15)
Chinese test score	0.141***			
Math test score	0.095**			
Cognitive /Literacy test	0.372***	0.331***	0.438***	
Internalizing scale	0.030			
Externalizing scale	0.094***			
Motivation	0.155***			
Self-esteem Scale				0.276***
Depressive symptoms				0.291***

Note: this table shows the correlation between cognitive and noncognitive skills Over time. All the variables are standardized. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Intergenerational Education Transmission: Baseline Results

	(1)	(2)	(3)
	Whole	Female	Male
Years of schooling of Father	0.391***	0.351***	0.422***
	(0.035)	(0.054)	(0.046)
Observations	848	358	490

Data Source: GSCF, all four waves. Dependent Variables are children's year of schooling. The control variables children's birth year dummies, father's age, father's age square and gender dummy are included but the estimation coefficients before these controls are omitted in this table. Standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Intergenerational Income Correlation: Baseline Results

Dependent Variable:	log children's income			log children's income Rank		
	(1)	(2)	(3)	(4)	(5)	(6)
	Whole	Female	Male	Whole	Female	Male
Family Income Measure	0.099^{***}	0.046	0.140^{***}	0.091^{***}	0.052	0.122^{***}
	(0.037)	(0.052)	(0.050)	(0.032)	(0.046)	(0.044)
Observations	832	353	479	848	358	490

Note: columns (1)-(3) are the estimation of traditional intergenerational income elasticity , for whole sample, female subsample and male subsample, respectively. Columns (4)-(6) is the estimation results for income rank rank specification. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Transitional Matrix between Family Income and Children's Income

Children's Income Group

Family Income Group	1	2	3	4	Total
1	28.77	26.42	23.58	21.23	100.00
2	26.89	23.11	26.42	23.58	100.00
3	28.77	22.17	24.06	25.00	100.00
4	15.57	28.30	25.94	30.19	100.00
Total	25.00	25.00	25.00	25.00	100.00

Note: Transitional matrix between family income group and children's income group. Each group is defined by the quartile of family income or children's income, 1 is the smallest income group and 4 stands for the highest income group.

Table 8: Impact of Father's Education on Cognitive ability

	(1)	(2)	(3)	(4)	(5)	(6)
	Chn00	Math00	Cogn00	Math04	Chn04	Literacy04
Father eduyear	0.061***	0.043***	0.060***	0.014	0.024**	0.068***
	(0.016)	(0.015)	(0.011)	(0.012)	(0.011)	(0.011)
Observations	430	418	848	848	848	848

Standard errors in parentheses Data Source: GSCF, all four

waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Impact of Father's Education on NonCognitive Ability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Internal00	External00	Mcolleg00	Internal04	External04	Mcollege04	Rosenberg09	Depress09
Father eduyear	-0.026**	-0.026**	0.028**	0.017	0.009	0.061***	0.033***	-0.020
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Observations	848	848	848	848	848	848	848	848

Standard errors in parentheses

Data Source: GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Impact of Father's Education on NonCognitive Ability in 2015

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Rosenberg15	Depress15	Extraver15	Agreeable15	Conscientious15	Neuroticism15	Open15
Father eduyear	0.035***	-0.034***	0.011	0.020*	0.027**	-0.028***	0.008
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)
Observations	848	848	848	848	848	848	848

Standard errors in parentheses Data

Source: GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Impact of Father's Education on Health

	(1)	(2)	(3)	(4)	(5)
	Birthweight	height04std	height09std	weight04std	weight09std
Father eduyear	0.035***	0.032***	0.009	0.013	0.009
	(0.012)	(0.011)	(0.008)	(0.011)	(0.009)
Observations	848	848	848	848	848

Standard errors in parentheses Data Source: GSCF, all

four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Decomposition of Intergenerational Education Transmission

	(1)	(2)
chinese00	.020679	5.3
math00	.011352	2.9
cogn00	.02478	6.3
math04	.001988	0.5
chinese04	.004608	1.2
literacy04	.046716	11.9
internal00	-.003068	-0.8
external04	.001222	0.3
Mcollege00	-.000252	-0.1
internal04	.001445	0.4
external04	-.001413	-0.4
Mcollege04	.043005	11.0
rosenberg09	.013827	3.5
depress09	-.00004	-0.0
birthweight	.003255	0.8
height04	.001792	0.5
height09	.001404	0.4
weight04	-.000312	-0.1
weight09	.001278	0.3
Cog	.110123	28.2
Ncog	.054726	14.0
Health	.007417	1.9
Explained	.172266	44.1
Unexplained	.827734	55.9

Notes: Data Source: GSCF. Column (1) show the coefficients explained by the corresponding variables in the left hand side. The number in the column are obtained by multiplying the impact of father's education on each skill variables coefficients and the coefficients in the corresponding children's education regression. Column (2) show the percentage of intergenerational education transmission explained by the corresponding variables under different specifications.

Table 13: Impact of Family Income on Cognitive Ability

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Chn00	Math00	Cogn00	Math04	Chn04	Literacy04	Literacy09	childedu
Log family income	0.267***	0.172**	0.221***	0.096	0.054	0.257***	0.168***	1.263***
	(0.084)	(0.079)	(0.053)	(0.062)	(0.056)	(0.054)	(0.060)	(0.192)
Observations	421	411	832	832	832	832	832	832

Standard errors in parentheses Data Source:

GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Impact of Family Income on Noncognitive Ability00-09

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Internal00	External00	Mcolleg00	Internal04	External04	Mcollege04	Rosenberg09	Depress09
Log family income	-0.049	-0.051	0.058	0.111*	0.103*	0.091	0.126**	-0.125**
	(0.058)	(0.058)	(0.057)	(0.061)	(0.062)	(0.057)	(0.060)	(0.059)
Observations	832	832	832	832	832	832	832	832

Standard errors in parentheses

Data Source: GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Impact of Family Income on Noncognitive Ability 15

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Rosenberg15	Depress15	Extraver15	Agreeable15	Conscientious15	Neuroticism15	Open15
Log family income	0.091	-0.077	0.100*	0.090	0.122**	-0.046	0.065
	(0.057)	(0.057)	(0.056)	(0.058)	(0.054)	(0.052)	(0.060)
Observations	832	832	832	832	832	832	832

Standard errors in parentheses Data

Source: GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Impact of Family Income on Health

	(1) Birthweight	(2) height04std	(3) height09std	(4) weight04std	(5) weight09std	(6) weight15std
Log family income	0.234***	0.226***	0.035	0.147**	0.145***	0.098**
	(0.060)	(0.050)	(0.040)	(0.063)	(0.056)	(0.050)
Observations	832	832	832	832	832	832

Standard errors in parentheses Data Source:

GSCF, all four waves

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Decomposition of Log Log Intergenerational Income Persistence

	(1)	(2)	(3)	(4)
Chn00	-0.001	-1.1	-0.002	-1.6
Math00	-0.006	-5.6	-0.006	-5.7
Cogn00	0.011	11.2	0.011	10.9
Math04	-0.001	-0.8	-0.001	-0.9
Chn04	0.001	0.8	0.001	0.8
Literacy04	0.007	7.3	0.007	7.0
Literacy09	0.005	4.9	0.004	4.2
Internal00	0.005	5.5	0.006	5.6
External00	-0.005	-4.9	-0.005	-4.9
Mcolleg00	0.000	0.2	0.000	0.2
Internal04	0.002	2.0	0.002	2.0
External04	-0.002	-2.1	-0.002	-2.1
Mcollege04	-0.002	-2.1	-0.002	-2.4
Rosenberg09	-0.001	-1.4	-0.002	-1.5
Depress09	0.002	1.9	0.002	1.9
Rosenberg15	0.005	4.7	0.004	4.5
Depress15	0.003	2.6	0.003	2.6
Extraver15	-0.001	-0.7	-0.001	-0.7
Agreeable15	-0.000	-0.2	-0.000	-0.3
Conscientious15	-0.000	-0.2	-0.000	-0.4
Neuroticism15	0.000	0.1	0.000	0.1
Open15	-0.005	-5.3	-0.005	-5.4
Birthweight	0.003	2.8	0.003	2.6
height04std	0.010	10.3	0.010	10.3
height09std	0.004	3.7	0.004	3.7
weight04std	-0.005	-5.3	-0.005	-5.3
weight09std	0.001	1.2	0.001	1.0
weight15std	0.001	1.1	0.001	1.1
childedu			0.006	6.4
Cog	0.017	16.8	0.015	14.7
Ncog	0.000	0.2	-0.001	-0.8
Health	0.014	13.7	0.013	13.4
Explained	0.030	30.7	0.033	33.7
Unexplained	0.069	69.3	0.066	66.3

Notes: Data Source: GSCF. The number in all odd columns are obtained by multiplying the coefficients of the impact of log family net income on each ability measurement and the coefficients in the corresponding children's income regression. The number in the even columns are the percentage of log income intergenerational persistence explained by the corresponding

Table 18: Decomposition of Rank Rank Slope

	(1)	(2)	(3)	(4)
Chn00	0.005	5.1	0.004	4.5
Math00	-0.006	-7.1	-0.006	-7.1
Cogn00	0.006	6.4	0.005	5.8
Math04	-0.001	-1.1	-0.001	-1.2
Chn04	0.000	0.4	0.000	0.4
Literacy04	0.008	8.6	0.007	7.9
Literacy09	0.005	5.1	0.004	4.6
Internal00	0.008	9.0	0.008	9.0
External00	-0.010	-11.2	-0.010	-11.1
Mcolleg00	-0.000	-0.2	-0.000	-0.2
Internal04	0.002	2.0	0.002	1.7
External04	-0.003	-3.0	-0.002	-2.7
Mcollege04	-0.001	-0.7	-0.001	-0.8
Rosenberg09	0.001	1.5	0.001	1.4
Depress09	-0.000	-0.5	-0.000	-0.5
Rosenberg15	0.003	3.3	0.003	3.1
Depress15	0.000	0.0	0.000	0.0
Extraver15	-0.001	-1.2	-0.001	-1.2
Agreeable15	0.000	0.2	0.000	0.2
Conscientious15	0.001	1.0	0.001	0.9
Neuroticism15	0.000	0.0	0.000	0.0
Open15	-0.005	-6.0	-0.006	-6.2
Birthweight	0.002	1.7	0.002	1.9
height04std	0.004	4.0	0.004	4.5
height09std	0.000	0.0	0.000	0.0
weight04std	-0.002	-2.3	-0.002	-2.4
weight09std	0.000	0.0	-0.000	-0.1
weight15std	0.002	2.3	0.002	2.3
childedu			0.004	4.2
Cog	0.016	17.3	0.014	14.9
Ncog	-0.005	-6.0	-0.006	-6.4
Health	0.005	5.7	0.006	6.2
Explained	0.015	17.0	0.017	18.9
Unexplained	0.076	83.0	0.082	81.1

Notes: Data Source: GSCF. The number in all odd columns are obtained by multiplying the coefficients of the impact of log family net income rank on each ability measurement and the coefficients in the corresponding children's income rank regression. The number in the even columns are the percentage of rank rank slope explained by the corresponding variables in the

Table 19: Decomposition of Intergenerational Education Persistence

	(1)	(2)
Fcog00	0.057	14.6
Fcog04	0.037	9.6
FNonCog00	0.001	0.3
FNonCog04	-0.001	-0.3
FNonCog09	0.012	3.1
Health edu	0.007	1.8
Cog	0.094	24.2
Ncog	0.013	3.2
Health	0.007	1.8
Explained	0.114	29.2
Unexplained	0.277	70.8

Notes: Data Source: GSCF. We replace the intermediate variable with factor variables.

Table 20: Decomposition of Intergenerational Income Persistence: PCA

	(1)	(2)	(3)	(4)
	Unconditional on Education		Conditional on Education	
Fcog00	0.001	1.4	0.008	7.9
Fcog04	0.002	1.6	0.005	4.9
Fcog09	0.001	1.2	0.005	5.2
FNonCog00	0.000	0.0	-0.000	-0.1
FNonCog04	0.000	0.0	-0.000	-0.3
FNonCog09	0.000	0.0	-0.001	-1.1
FNonCog15a	0.002	2.3	0.002	1.7
FNonCog15	0.001	0.5	0.003	2.9
Hlhinc1	0.010	9.8	0.005	5.4
Hlhinc2	0.002	1.6	0.005	4.8
Childedu			-0.004	-4.4
Cog	0.004	4.2	0.018	18.0
Ncog	0.003	2.9	0.003	3.2
Health	0.011	11.4	0.010	10.2
Explained	0.018	18.5	0.027	27.0
Unexplained	0.081	81.5	0.072	73.0

Notes: Data Source: GSCF. We use factor variable for ability measurement in the income decom- position.

Table 21: Decomposition of Intergenerational persistence: Gender Difference

	(1)	(2)	(3)	(4)
	Male		Female	
Chn00	0.009	6.1	-0.011	-24.6
Math00	-0.011	-8.0	0.003	7.5
Cogn00	0.011	8.0	0.005	10.1
Math04	-0.002	-1.4	0.009	19.9
Chn04	0.000	0.3	-0.008	-16.7
Literacy04	0.011	7.6	0.005	10.3
Literacy09	-0.001	-0.5	0.009	20.1
Internal00	0.003	1.8	0.010	22.7
External00	-0.001	-0.6	-0.003	-6.0
Mcolleg00	0.001	0.6	-0.000	-0.2
Internal04	-0.006	-4.2	0.003	5.8
External04	0.007	5.0	0.001	2.2
Mcollege04	0.000	0.2	-0.002	-5.4
Rosenberg09	-0.000	-0.3	-0.001	-1.2
Depress09	0.002	1.3	0.004	7.8
Rosenberg15	0.007	4.9	0.001	1.8
Depress15	0.002	1.8	0.001	2.9
Extraver15	-0.003	-2.2	0.005	10.4
Agreeable15	0.000	0.1	-0.004	-8.6
Conscientious15	0.002	1.3	0.000	0.8
Neuroticism15	-0.001	-0.9	-0.001	-3.2
Open15	-0.000	-0.2	-0.017	-37.1
Birthweight	0.004	3.0	0.003	6.3
height04std	-0.001	-0.5	0.005	11.3
height09std	0.001	0.9	0.012	26.4
weight04std	0.002	1.5	-0.006	-12.6
weight09std	-0.006	-4.1	0.011	24.4
weight15std	0.006	4.0	-0.004	-8.0
Cog	0.017	12.2	0.012	26.6
Ncog	0.012	8.5	-0.003	-7.2
Health	0.007	4.8	0.022	47.8
Explained	0.036	25.5	0.031	67.3
Unexplained	0.104	74.5	0.015	32.7

Notes: Data Source: GSCF. The first two columns are for the income persistence decomposition of Male subsample, and column (3)-(4) are for income persistence decomposition of female subsample.