The Role of Social Capital in Accumulation of Human Capital: A Study of Social Network and College Attendance in Rural China

Zhiqing Li (zl3064)

Columbia University in the City of New York

Dec 7th, 2021

Introduction

Social network, defined by Granovetter as "a combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" (Granovetter, 1973), has been playing an indispensable role in China both in the ancient and contemporary context. It is mainly due to Confucian's emphasis on person-to-person relationships and the obligation attached to them. Interpersonal ties contain abundant information, which may increase the possibility of being employed (Knight, 2008), escalate willingness to establish own business because of informal credit (Ma, 2011), and so on.

Many people in rural China regard higher education as "a useless investment", for they have a relatively low perception of returns to universities, especially in labor markets. According to a 2011 report of the Bureau of Education of China, despite the largely expanded college enrollment rates, a declining trend can be observed from the registered *Gaokao* (National College Entrance Examination in China) participants in recent years, which is projected to

persist and widespread in rural areas. In addition, the rapid growth of household income did not promote investments in education for rural families. (Yin, 2016)

Hence, social networks may also be helpful in terms of disseminating more accurate information about the value of schooling from cities to the countryside, thereby increasing rural households' propensity of taking part in college education. Moreover, this question remains untouched until recent years in China, and it has not been answered convincingly.

In order to disentangle the causations, I adopt the measurement of social networks designed in accordance with Chinese society (Bian, 1997), namely, "the web of greetings in Chinese traditional Spring Festivals", and apply it to the 2007 CHIP house survey data. Moreover, I utilize family name structures of a certain village as an instrumental variable to cope with potential endogeneity. A series of robustness checks are included, which excludes provinces that have extremely scarce educational resources or ample supply of universities to rule out the effect of outliers, and also try to alter the assessment of social networks. I find out that greeting one more citizen during Spring Festival increases children's probability of partaking in colleges by 0.2 percentage points in rural China.

Empirical Analysis

(a) Explanation of Variables and Hypothesis

My hypothesis is that social networks with citizens are related to college attendance for Chinese rural residents. I use OLS first.

$$College\ Attendance_i = \alpha + \beta * Social\ Network_i + \delta * X_i + \varepsilon_i$$

The Independent variable is the family household head's or spouse's social network with citizens. ε_i captures the error term. X represents the control variables.

Variable	Definition
Citizens.num	Number of citizens sent greetings to by the surveyed heads of families during 2007 Spring Festival (independent variable)
College	Whether attending higher education (1 = yes, 0 = no); higher education is defined as undergraduate (bachelor's degree) plus polytechnic college.
Gifts. giving	The percentage of household income used as gifts/banquet/money (in 2007), converted to monetary value, that rural families gave to their closest persons who were living mainly in cities
Gifts.receiving	Gifts/banquet/money (in 2007) relative to household income, converted to monetary value received from their closest citizens.
Household characteristics control variables	
Income_net	Household net income in 2007 (yuan in ten thousand)
Fathedu	Years of official education received by father
Mothedu	Years of official education received by mother
Siblings	Number of children in the family
Gender	Gender of children (1 = male, 0 = female)
Dad.ethnicity	Father's ethnicity (1 = majority, 0 = minority)
Mum.ethnicity	Mother's ethnicity (1 = majority, 0 = minority)
Ethnicity	Individual's ethnicity
Village characteristics control variables	
Ave.income	Average income of a certain village
Policy	Whether the village is listed as one of the "state poverty counties" (1 = yes, 0 = no)
Distance1	Time cost to the nearest cities (1 = within 30 minutes, 2 = half-to-one hour, 3 = one-to-two hours, 4 = two-to-four hours, 5 = more than four hours)
Distance2	Time cost to the nearest primary school (1 = within 30 minutes, 2 = half-to-one hour, 3 = one-to-two hours, 4 = two-to-four hours, 5 = more than four hours)
Distance3	Time cost to the nearest middle school (the same as above)

Parents' ethnicity may result in different social behaviors and views on the importance of education. The average income of village may cause different investment to education due to different tax revenue in a certain village; "state poverty counties" are regions that received a series of policy benefits including education and economic development, which may therefore affect the outcome variable.

(b) Data

The data I am going to use is CHIP 2007(Chinese Household Income Project), which is carried as part of a collaborative research project on incomes and inequality in China organized by Chinese and international researchers, with the assistance from National Bureau of Statistics in 2007.

I only choose families whose heads and spouses had never worked in urban regions until 2007 and both have rural *hukou* (a system of household registration in mainland China and Taiwan, in order to identify a person as a resident of an area;户口), in order to limit the possibility of other mechanisms. For instance, since cities have more well-equipped schools and more job opportunities, villagers may want to migrate to urban areas with the aim of providing their children with higher-quality education. Rural dwellers who have living experiences in cities also tend to be acquainted with more urban residents. Thus, it is hard to say that such inclination is caused by social networks or subjective motivations. Also, all the children in my dataset are all adult biological descendants, since having a stepparent is negatively correlated with educational attainment when suffering from budget constraints.

My filtered dataset is individual-level and contains 2233 observations. As shown in table 1, we can see that most families have income that is below the average level of entire China (< 50,000 RMB per year); most of them are from the major ethnic group; the majority of the adult children do not have a college degree. (See figure 1 also)

Descriptive Statistics

Statistic	college	gender	income_net	fathedu	mothedu	siblings	citizens.num	dad.ethnicity	mum.ethnicity	ave.income	policy	Distance1	Distance2	Distance3
N	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233	2,233
Mean	0.2	0.4	1.8	6.5	4.1	2.6	6.9	1.0	1.0	10.7	1.8	2.2	1.6	2.1
St. Dev.	0.4	0.5	1.4	2.9	3.2	1.3	11.6	0.1	0.1	3.5	0.4	0.8	1.0	0.9
Min	0	0	-3.6	0	0	0	0	0	0	0	0	0	0	0
Pctl(25)	0	0	1.0	5	0	2	0	1	1	9	2	2	1	1
Pctl(75)	0	1	2.2	9	6	3	9	1	1	12	2	3	2	3
Max	1	1	26.1	19	16	11	120	1	1	18	2	5	5	5

Table 1

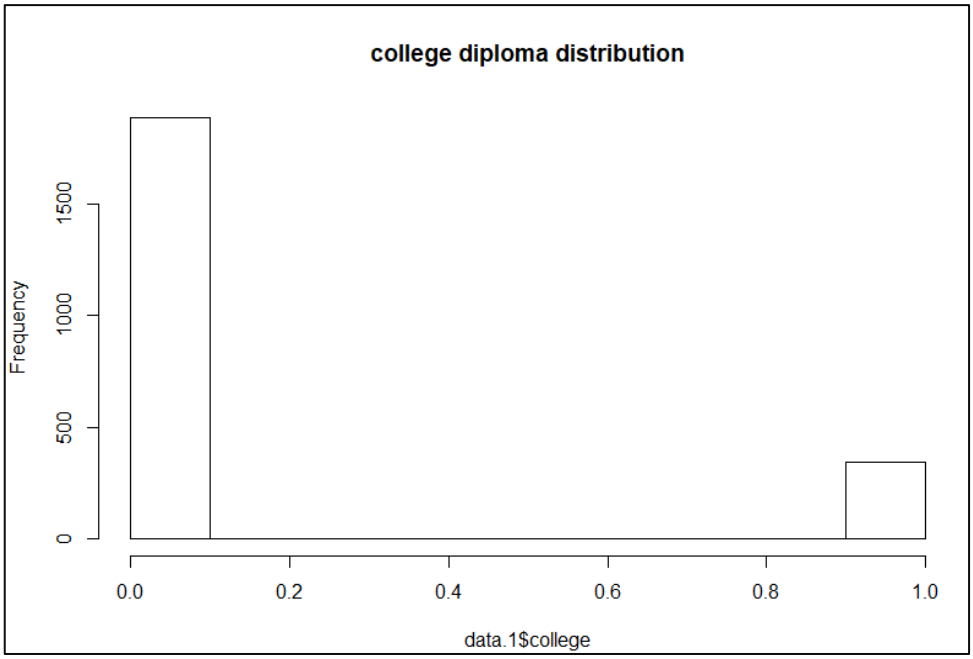

Figure 1

(c) Methodology and Identification

The social network is hard to quantify due to its vague definitions and broad interpretations in various societies. In this paper, I adopt the measurement designed by Bian Yanjie (2012), who counts the number of people greeted by the respondents during the Chinese traditional

Spring Festival. The greeting behavior is called *bainian* in Chinese, and it is regarded as an indispensable obligation for Chinese citizens, especially the majority ethnicity group, namely, *Han* people, to strengthen the ties of kinship and friendship. In this context, I use the number of citizens greeted by the surveyed heads of families during the 2007 Spring Festival to evaluate the density and strength of networks with people living in the urban area.

The first column in Table 2 reports the OLS results of the regression, according to which we are 99% confident that greeting one more citizen increases the children's probability of attending college education by 0.2 percentage points.

However, there may be endogenous problems in the OLS estimates. For instance, there may be reverse causality: having a child with a college diploma may be correlated with a denser social web with urban civilians. Moreover, omitted variables may also affect both social networks and college attendance. It is possible that parents with a broader social web may have systematic differences compared with others who do not, which may also have impacts on children's chance of going to college simultaneously. Besides, the methods adopted for measuring social networks may be inappropriate. There are 55 minority groups in China, some of which even do not have the tradition of the Spring Festival.

Therefore, I use "whether more than 50% of the villagers have the same family names" as an instrumental variable, with 1 denoting "yes" and 0 otherwise. Traditionally, branches of families are tightly related with each other; therefore, rural residents living in the village comprising only their relatives may be less willing to migrate or tend to migrate together. The latter one cannot be observed in my sample, since I only select rural civilians with no experience in urban districts on a long-term basis. Owing to the reasons above, I expect to see a negative correlation between the independent variable and the instrumental variable.

Next, I will explain why it may be a proper instrument. First, this variable has significant predictive power of the variable of interest, for the coefficient is -2.627 and it is statistically significant at a 99% confidence level (see the last column in table 2). Second, feedback from college enrollment to the structures of village members is unlikely to occur; hence, it seems to be exogenous in the context. Third, components of neighboring members seem to only correlate with college participation through altering the size and density of social networks intuitively.

Social Network and College Attendance

	Dependent variable:			
	college			citizens.num
	OLS		*instrumental variable*	*OLS*
	(1)	(2)	(3)	(4)
citizens.num	0.002***	0.002***	0.003	
	(0.001)	(0.001)	(0.006)	
IV				-2.627***
				(0.500)
gender	0.145***	0.139***	0.138***	0.267
	(0.015)	(0.015)	(0.015)	(0.501)
fathedu	0.005*	0.006**	0.006**	-0.043
	(0.003)	(0.003)	(0.003)	(0.092)
mothedu	0.014***	0.014***	0.013***	0.015
	(0.003)	(0.002)	(0.003)	(0.084)
siblings	-0.029***	-0.026***	-0.026***	-0.030
	(0.006)	(0.006)	(0.006)	(0.201)
income_net	0.014**	0.006	0.006	0.235
	(0.006)	(0.006)	(0.006)	(0.192)
dad.ethnicity	0.141	0.131	0.128	3.337
	(0.124)	(0.124)	(0.125)	(4.171)
mum.ethnicity	-0.066	-0.083	-0.084	1.436
	(0.121)	(0.120)	(0.120)	(4.028)
ave.income		0.010***	0.010***	0.300***
		(0.002)	(0.003)	(0.080)
policy		-0.086***	-0.084***	-2.093***
		(0.021)	(0.025)	(0.710)
Distance1		-0.043***	-0.042***	-0.492
		(0.010)	(0.010)	(0.333)
Distance2		-0.016*	-0.015*	-0.601**
		(0.008)	(0.009)	(0.280)
Distance3		-0.007	-0.008	1.257***
		(0.010)	(0.012)	(0.340)
Constant	-0.035	0.188**	0.185*	3.489
	(0.085)	(0.095)	(0.097)	(3.204)
Observations	2,233	2,233	2,233	2,233
R^2	0.084	0.113	0.112	0.030
Adjusted R^2	0.080	0.108	0.107	0.024
Residual Std. Error	0.347 (df = 2224)	0.341 (df = 2219)	0.342 (df = 2219)	11.497 (df = 2219)
F Statistic	25.333*** (df = 8; 2224)	21.833*** (df = 13; 2219)		5.310*** (df = 13; 2219)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2

After analyzing the results of two-stage least square regressions, we can see that the correlation between the two variables is weaker in terms of statistical significant level; nevertheless, the coefficient increases from 0.002 to 0.003, which means greeting one more citizen by the head of the family increases the probability of having a child with a college education by 0.3 percentage point, but the standard error is relatively large. (see table 2)

There appears to be no substantial difference between the OLS estimate and IV results. The biggest problem that lies behind the instrument is that people sharing identical last names may not be members of the same family; instead, it is plausible that their family names are simply too common. It may render the instrument invalid. Hence, I accept the OLS estimate this time. Also, because inhabitants in China can choose their own ethnicities based on either one of their parents' ethnicity, the majority of people prefer themselves to be a member of a minority ethnic group due to policy benefits. For example, the beneficiaries may be able to enter universities with much lower college entrance examination scores compared with the majority *Han* people (汉族), which suggests the minorities may be overrepresented.

Furthermore, it is noticeable that the mean number of citizens greeted is only 6.9, which indicates social networks can only lead to an approximately 1.4 percentage point increase of college attending for rural students on average. We may conclude that social network with citizens bears no significant causal relationship with higher education participation in rural China.

The reasons behind this may be multifaceted. First, the observations are mainly the poorest families, since their annual net household income is predominantly less than 20,000 RMB (equivalent to around 3730 USD, 2019). It means their friends and relatives may not be able

to afford to live in metropolitans where the labor market was relatively mature and determined mostly by supply and demand. Instead, they may gather in small towns or underdeveloped cities, where the labor markets were not market-oriented in 2007 and thereby biasing the coefficient downward. Second, the surveyed area is simply too small, for the majority of the respondents come from the central part of China who may not be able to represent people elsewhere. Third, it is highly possible that rural residents are not able to afford or even approach the high quality of fundamental education such as enrolling in good middle schools due to regional inequality and corruption; thus, even though they have perceptions of universities that are closer to the reality, there is little they can do to change the status quo.

Finally, some details are worth mentioning here. Firstly, the selection bias may prevent me from understanding the real mechanism. Rural-to-urban migrants may have a stronger desire for a better quality of education in the first place; therefore, a higher preference for college education may result from contacting more education-ambitious persons rather than citizens with more accurate information. Moreover, many parents in the sample were born in the 1960s, suggesting that the Reform and Opening Up policy was introduced exactly when they were about to enter the workforce. Hence, hundreds of thousands of rural workers that are peers to my selected observations chose to leave their hometown and settle down in cities that were distant and unfamiliar. At that time, cheering for a decision made by the central government and following the instructions of the ‘great leader’ were commonly seen in China, regardless of the consequences. In addition, rural areas tend to be more conservative and obey the traditional Chinese values, especially Confucianism, and Confucius specifically asked people not to travel long distances when their parents are still alive because they have

to attend to the elderly.[1] Otherwise, they will be considered to be "unfilial" and disrespectful, which is extremely serious since they can be despised and criticized by neighbors. The factors mentioned above as well as the personal risk preference and aptitude are all influential in terms of the choices of migrating or not, making it more difficult to find out the true channel.

Secondly, endogeneity is not well tackled by this paper, and valid instrumental variables or panel data are required. Thirdly, since some children in my dataset were already adults, they may also contribute to the overall household income and potentially result in a biased estimate.

Conclusions

I use data collected by Beijing Normal University during 2007 with the assistance of the Bureau of Statistics and academics overseas to identify the relationship between social networks with people living in cities and higher education attendance in rural China. Perception of returns to college diplomas is generally more accurate in cities due to the freer flow of information; by contrast, rural inhabitants often regard attending universities as a "useless" investment because they cannot observe benefits brought by higher education in the labor markets. Therefore, contacting more citizens may change villagers' views on higher education.

After using an instrumental variable, the accepted OLS estimates suggest that there is no significant causal relationship between the mentioned two variables. For rural households, associated with one more citizen can only lift the possibility of children's attending colleges

[1] According to The Analects of Confucius. 父母在，不远游。It can be translated into "While his parents are alive, the son may not go to distant places."

by 0.2 percentage point. Owing to the limitation of cross-sectional data and relatively small sample size, as well as unsolved endogeneity, this conclusion may not be fully convincing and reliable; nonetheless, I shed light on such an interesting and important problem.

Reference

Bian, Y. (1997, March). Guanxi Networks and Job Mobility in China and Singapore. *Social Forces*, pp. 981-1005.

Granovetter, M. (1973, May volume 78(6)). The Strength of Weak Ties. *American Journal of Sociology*, pp. 60-80.

Knight, J. (2008, Volume 16(3)). The Role of Social Capital in the Labour Market in China. *Economics of Transitions*, pp. 189-414.

Ma, G. (2011, Volume 3). Social Networks, Informal Finance and Entrepreneurship. *Economic Research Journal*, pp. 83-93.

Yin, Y. (2016, October). Impact Factors in Decreasing Numbers of Applicants for National College Entrance Examination after Educational Reform: Based on Provincial Panel Data. *Lanzhou Academic Journal*, pp. 166-175.