北京大学中国经济研究中心
China Center for Economic Research

讨论稿系列
Working Paper Series

No. E2016005 **March 2016**

How does social capital matter to health status of older adults? Evidence from the China Health and Retirement Longitudinal Survey

Gordon G. Liu[a], Xindong Xue[b,c,*], Chenxi Yu[d], Yafeng Wang[e]

[a] *National School of Development, Peking University, China*
[b] *School of Public Administration, Zhongnan University of Economics and Law, China*
[c] *Co-innovative Research Center for Health Insurance, Hubei University of Economics, China*
[d] *Woodrow Wilson School of Public and International Affairs, Princeton University, United States*
[e] *Institute of Social Science Survey, Peking University, China*

Abstract

This paper uses longitudinal data from China to examine the causal relationship between structural social capital and health among Chinese older adults. We control for the potential endogeneity of social capital via instrumental variable approach and account for the possible contextual confounding effects by including community-level social capital. We use three indicators to measure individuals' general, physical, and mental health respectively. Results indicate that social capital has a significant and positive effect on general and physical health. Based on the IV findings, a one standard-deviation increase in social capital leads to a 4.9 standard-deviation decrease in the probability of having bad health and a 2.2 standard-deviation decrease in physical activity limitations. Our results are robust to a series of sensitivity checks. Further analysis suggests some heterogeneous effects by age, but not by gender and area of residence.

Key words: Social capital; Health; Fixed effects; Instrumental variable; Heterogeneity; China

JEL codes: I14; I18; Z1

[*] Corresponding author. Tel.: +86 27 8838 7901; fax: +86 27 8838 6936.
E-mail address: xuexindong@znufe.edu.cn (Xindong Xue).

Highlights

- We examine the causal effect of structural social capital on health among Chinese older adults.
- We use three indicators to measure general, physical, and mental health respectively.
- Fixed effects and instrumental variable estimates show that social capital has a significant and positive effect on general and physical health, but not on mental health.
- Social capital effect on physical health tends to be stronger for the older old adults.

Acknowledgements

This research is supported by the National Social Science Fund of China (Grant No.14BRK013).We would like to thank three anonymous reviewers and editor Jörg Baten for their insightful comments, which have resulted in a much improved manuscript. We are also grateful to Sherman Folland, Peter Eibich, Lorenzo Rocco, and participants at the Health Economics Seminar (Peking University) and the fifth workshop of Global Network on Social capital and Health (University of Toronto) for their helpful comments and suggestions. All errors are our own.

1. Introduction

With increasing life expectancy and decreasing fertility rate, population aging has become a global challenge in the 21st century. The process of population aging is usually accompanied by higher prevalence of chronic diseases and disabilities among the elderly, which may consequently reduce economic performance and increase public health care expenditures (Sirven and Debrand, 2008). Among the various social policies to cope with the health challenges associated with aging, social capital has been proposed as an important policy instrument (Oxley, 2009). The positive relationship between social capital and health is observed in many countries, for different measurements of social capital, and various indicators of health (Folland and Rocco, 2014). However, there is no consensus on whether this relationship is causal and how large it is. This discussion is of particular significance to China as China has the largest aging population and one of the highest aging rates in the world today (Zhao et al., 2013). If social capital does have a causal effect on the health of older adults in China, then providing them with easier access to social interactions and organizations will – in the long run – be a more effective health promotion strategy than merely increasing public health care expenditures.

Three possible mechanisms of why social capital may causally lead to better health have been identified in the literature (Scheffler and Brown, 2008; Folland, 2008). First, social capital helps to buffer individuals' mental stress by forming an environment that facilitates social interaction and builds mutual trust. Second, social capital helps individuals to obtain information on health care and health, thus improving health production efficiency. Third, social capital can generate a sense of shared responsibility within families and communities and consequently reduces the likelihood of risky behaviors such as drinking and smoking.

The association between social capital and health has been examined in a number of countries including the US (Schultz et al., 2008; Scheffler and Brown, 2008; Scheffler et al., 2007; Folland, 2007), Canada (Nakhaie and Arnold, 2010; Veenstra, 2000), the UK (Rocco, 2014b; Borgonovi, 2010; Snelgrove et al., 2009; Petrou and Kupek, 2008; Poortinga, 2006b), Italy (Fiorillo and Sabatini, 2015), the Nordic countries (Lindstrom, 2009; Lindstrom and Mohseni, 2009; Islam et al., 2008; Iverson, 2008), Europe as a whole (Rocco et al., 2014a; Ljunge, 2014; Sirven and Debrand, 2008, 2012; Rostila, 2007; Poortinga, 2006a), Japan (Murayama et al., 2013, 2015; Yamamura, 2011a, 2011b; Ichida et al., 2009; Fujisawa et al., 2009), former Soviet Union countries (Goryakin et al., 2014; Habibov and Afandi, 2011; D'Hombres et al., 2010; Ferlander and Mäkinen, 2009; Rose, 2000) and Latin American countries (Ronconi et al., 2012; Hurtado et al., 2011). Despite many empirical arguments for the positive association between social capital and health, concerns over interpreting correlation as causation remain. First, spurious correlation may be caused by omitted variables. Some unobserved heterogeneity such as time preference and personal traits could drive both social capital and health to move in the same direction. For example, people with low discount rates place high value on future benefits. These people are more likely to increase investments in social capital and health concurrently. Second, there may be reverse causality bias. On the one hand, people with better health are both physically and financially capable of engaging in more social activities, thus having higher social capital. On the other hand, people with relatively worse health may have increased incentive to invest in social capital as their opportunity cost of time is lower. Third, there may be measurement errors related to the subjective nature of self-reported health. Although self-reported health is commonly used in current literature, misreporting bias still remains in the estimation. Fourth, it is difficult to distinguish social capital effects from other local contextual effects that potentially influence health. Social capital may vary between locations depending on social and economic characteristics of the community, which may confound the effect of individual-level social capital on health (D'Hombres et al., 2010).

A number of studies have attempted to disentangle the causal relationship between social capital and health by employing several identification strategies. These strategies include the use of lagged value in longitudinal data to control for reverse causality bias (Sirven and Debrand, 2012; Nakhaie and Arnold, 2010) and the instrumental variable (IV) approach (Fiorillo and Sabatini, 2015; Ljunge, 2014; Rocco et al., 2014a; Goryakin et al., 2014; Yamamura, 2011b; D'Hombres et al., 2010; Schultz et al., 2008). The most frequently used instrument is the community-level social capital. However, this instrument may not satisfy the exclusion restriction condition because there may be possible contextual effects of social capital on health at the community level. Higher community-level social capital usually indicates a better living environment, which could have a positive effect on individuals' health.

Our study aims to contribute to the literature in several ways. First, this will be the first empirical study in the Chinese context to examine the causal relationship between social capital and health using national longitudinal data. Second, we take into account the potential endogeneity resulting from unobserved heterogeneity and reverse causality by using fixed effects (FE) and IV estimators. Third, we account for the possible contextual effects by including community-level social capital. Fourth, we explore the heterogeneous effects of social capital on health by gender, age, and area of residence. Our study shows that structural social capital has a significant and positive effect on general and physical health, but not on mental health. Social capital effect on physical health tends to be stronger for the older old adults.

This paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the data and variables. Section 4 presents the empirical methodology. Section 5 reports results, robustness checks, and heterogeneity. Section 6 concludes.

2. Literature review

2.1 The definition and measurement of social capital

The concept of social capital was first put forward by Bourdieu (1986) and further popularized by Coleman (1990) and Putnam (2000). Despite some disputes on its definition, social capital is commonly understood to encompass social networks, norms, and trust that facilitate cooperation between individuals in a community (Putnam, 1993). Some researchers suggest that social capital has cognitive and structural dimensions (Harpham et al., 2002). Cognitive social capital relates to people's perceptions of interpersonal trust, solidarity and reciprocity, whereas structural social capital refers to the extent and intensity of participation in associations and other forms of social activity (Rostila, 2007). Structural social capital is particularly conducive to generating the beneficial effects of social capital as participation in social groups may facilitate the transmission of knowledge and increase trust between members of society (Alesina and La Ferrara, 2000; Putnam, 1993). Therefore, social participation is one of the most frequently employed measures of social capital in the empirical literature (Sirven and Debrand, 2008, 2012; Borgonovi, 2010; Berry and Welsh, 2010; Ronconi et al., 2010).

2.2 Empirical studies on social capital and health

Although numerous studies have linked social capital to positive health outcomes, the empirical relationship between different forms of social capital and health is mixed at best. For instance, with regard to the relationship between structural social capital and health, some studies show a positive association between social participation and health (Borgonovi, 2010; Berry and Welsh, 2010; Petrou and Kupek, 2008), while others do not find such relationship (Meng and Chen, 2014; Hurtado et al., 2011; D'Hombres et al., 2010; Snelgrove et al., 2009; Yip et al., 2007; Norstrand and Xu, 2012). Furthermore, some studies find that cognitive social capital is positively related to self-reported health and structural social capital is positively related to somatic symptoms (Goryakin et al., 2014; Yamoka, 2008).

As noted above, several studies use longitudinal data and an IV approach to deal with the potential endogeneity of social capital. Based on the respective longitudinal data in Europe and Canada, Sirven and Debrand (2012) and Nakhaie and Arnold (2010) use the lagged social capital instead of current social capital to control for reverse causality bias. They both find that social capital contributes to better health. Using an IV approach, most studies find larger coefficient estimates on social capital than in the OLS estimation. Folland (2007) uses state-level characteristics such as employment per capita, geographic latitude, and state government contributions to colleges per capita as IVs and finds social capital to be highly correlated with a number of health measures in the US. Schultz et al. (2008), in their study on the US population, use residence years and religiosity variables as IVs. They find individual-level social capital to be a significant predictor of self-rated health. A similar approach is followed by Yamamura (2011b), who finds social capital to have a significant and positive influence on health for people without a job in Japan. Ronconi et al. (2010) use transportation conditions as IVs and find a significant and positive relationship between social capital and health in Argentina. Rocco et al. (2014a) instrument social capital with crime victimization and physician density in the community and find a causal and positive relationship between social capital and health in European countries. These findings are further confirmed by Fiorillo and Sabatini (2015), Goryakin et al. (2014) and D'Hombres et al. (2011), who address the endogeneity issue by employing community-level social capital as IVs. Ljunge (2014) uses ancestral trust as an IV and finds a causal effect of trust on health in Europe.

Only a few studies investigate the association between social capital and health among older adults. Muchenhuber et al. (2013) find social capital to be significantly more important for health of older people than for younger people in Austria. Sirven and Debrand (2008, 2012) and Leon and Hassel (2015) report that social capital improves health of older people in Europe. Murayama et al. (2013) find that social capital has beneficial effects on health of the older Japanese. Koutsogeorgou et al. (2014) arrive at the conclusion that the association between social capital and self-rated health among older adults differs in the Poland, Finland, and Spain.

In the Chinese context, Yip et al. (2007) are the first to examine the association between social capital and health. Using data from three counties in Shandong province, they find that cognitive social capital (trust) is positively associated with health, while structural social capital (organizational membership) is not. Wang et al. (2009) use household survey data from 22 villages in rural China and find that trust is positively associated with general and mental health. Norstrand and Xu (2012) find no significant relationship between social participation and physical and emotional health among older Chinese. Meng and Chen (2014) document similar findings. The above studies on China suffer from three limitations. First, these studies did not account for the potential biases caused by unobserved heterogeneity and reverse causality. Second, self-reported health is commonly used and possible misreporting bias may exist. Third, they fail to examine the heterogeneity underlying the relationship between social capital and health in China.

3. Data, variables and descriptive statistics

3.1 Source of data

We use data from the 2008 and 2010 waves of the *China Health and Retirement Longitudinal Survey* (CHARLS).[1] CHARLS is a comprehensive biennial survey aiming to collect detailed information about demographics, family, health, work, retirement and pension from a nationally representative sample of Chinese residents aged 45 and older. It was conducted by the National School of Development at Peking University. Considering the enormous complexity involved in a national survey, CHARLS team administered a pilot survey in Gansu and Zhejiang province (see Fig.1) in 2008 and 2012. Wave 2008 is the first survey and Wave 2012 is the follow-up survey. Gansu is a poor inland province located in the less developed northwestern region and Zhejiang

[1] More details of the survey can be accessed at: http://charls.ccer.edu.cn/en.

is a rich coastal province located in the developed eastern region. In 2007, Gansu had a population of 26 million and GDP per capita of 10,346 RMB (about 1,642 USD), while Zhejiang had a population of 50 million and GDP per capita of 37,411 RMB (about 5,398 USD) (National Statistical Bureau, 2008). The sharp contrast between these two pilot provinces reflects the unbalanced development and enormous diversity of China. The survey followed a strict, four-level multi-stage sampling process: county, community, household, and respondent. It collected data from 95 communities (villages) in 32 counties (districts), covering 2,685 individuals living in 1,570 households. In our analysis, the sample was restricted to respondents with non-missing values.

Fig.1 Two provinces in CHARLS pilot survey

3.2 Variables

3.2.1 Measuring health

Since health is a multi-dimensional and general concept, we take a holistic approach and examine three health indicators: general, physical and mental health. We first use self-reported health, as it can provide a more complete picture of one's overall wellbeing and significantly predict mortality and disability, even after controlling for many objective measures of health. In CHARLS, respondents are asked to assess their own health on a five-point scale: very good, good, fair, poor, and very poor. We then condense these responses into a binary variable: bad health (=1, poor and very poor) and good health (=0, very good, good and fair).[2]

Self-reported health may suffer from misreporting bias. Differences in individuals' personal traits and threshold points may cause measurement errors. More optimistic individuals may systematically overstate their health, or vice versa. To address the potential misreporting bias, we use two other objective indicators to measure individuals' physical and mental health respectively: the Limitations in Activities of Daily Living (ADL) Index and the Center of Epidemiological Studies Depression (CES-D) Scale.

[2] Unlike previous studies on China (Yip et al., 2007; Meng and Chen, 2014), we code "fair" as good health. In the Chinese context, "fair" is a conservative expression of "just so-so" or "ok". So when a person rates his/her health as "fair", it by no means implies his/her health is bad.

In contrast to self-reported health, ADL indicators are considered to be more objective and less likely to be affected by differences in individual response scales (Bratti and Mendola, 2014). In particular, ADL indicators have the advantage of recording specific facts related to an individual's daily living rather than his/her opinions on his/her physical well-being. In CHARLS, respondents are asked nine questions on ADL.[3] For each question, there are three choices: "*No, I don't have any difficulty*" (=0); "*I have difficulty*" (=1), and "*I cannot do it*" (=2).[4] We sum these nine items and create an ADL index according to the following formula proposed by Gertler and Gruber (2002):

$$ADL = \left(\frac{score - Minscore}{Maxscore - Minscore}\right)$$

The ADL index ranges from 0 to 1, with the higher number indicating more limitations in activities of daily living and worse physical health.

CES-D was originally developed by Radloff (1977) and is widely used as a measure of mental health. Respondents are asked to indicate the frequency of experiencing ten feelings or emotions during the past week.[5] The four responses are: "*rarely or none of the time*" (=0); "*some or a little of the time*" (=1); "*occasionally or a moderate amount of the time*" (=2), and "*most or all of the time*" (=3). The responses to questions about feeling hopeful about the future and being happy are reversely recoded so that higher values of the CES-D scale represent more depressive symptoms and worse mental health. Following Duncan and Rees (2005), we sum the responses to the ten questions to produce a scale between 0 and 30. Moreover, as the CES-D scale is heavily skewed to the right, we will use the log-transformed CES-D scale as the dependent variable in the following regressions.[6]

3.2.2 Measuring social capital

Social capital is measured using the CHARLS social participation question: "*Have you done any of these activities in the last month*?" There are ten activities for the respondents to choose from.[7] We exclude two activities ("stock investment" and "using the Internet") as they are not social participation in nature. For the remaining eight activities, if the respondents answer "yes", it is coded as 1, and 0 otherwise. We then standardize the total score from these eight questions and construct the social capital index.

---

[3] ADL measures difficulties in the following nine activities: running or jogging about 1 km; walking 1 km; walking 100 meters; getting up from a chair after sitting for a long period; climbing several flights of stairs without resting; stooping/kneeling/crouching; reaching or extending your arms above shoulder level; lifting or carrying weights over 10 *jin* (=0.5 kg); picking up a small coin from a table.

[4] It should be noted in wave 2012, there are four choices for the ADL question: "No, I don't have any difficulty"; "I have difficulty but can still do it"; "Yes, I have difficulty and need help" and "I cannot do it". In order to be consistent with wave 2008, we recode "I have difficulty but can still do it" and "Yes, I have difficulty and need help" into one scale: "I have difficulty".

[5] The ten items are: I was bothered by things that don't usually bother me; I had trouble keeping my mind on what I was doing; I felt depressed; I felt everything I did was an effort; I felt hopeful about the future; I felt fearful; My sleep was restless; I was happy; I felt lonely; I could not get "going".

[6] The skewness value of the CES-D scale is 0.934.

[7] The ten activities are: volunteering or philanthropy activities; taking care of the elderly or disabled that you don't live with free of charge; offering help to relatives, friends or neighbors that you don't live with free of charge; going to school or attending training courses; visiting friends; playing *mahjoon*, chess, poker and going to community activity center; participating in *tai chi*, dancing and other group activities; participating in activities organized by community organizations; stock investment and used the Internet.

3.2.3 *Other control variables*

We also control for individual-, household-, and community-level variables. Individual-level variables are: age, gender, area of residence, education, marital status, insurance status, and health behaviors. Household-level variables are: family size, log of per capita income, availability of water, and whether the house is connected to sewage. Community-level variables are: community social capital, hospital density, and activity center density.[8] The definitions of all variables are presented in Table 1.

Table 1
Variable definitions

Variable	Definition
Dependent variables	
Bad health	=1 if self-reported health is poor or very poor, 0 otherwise
ADL index	From 0~1. Higher values indicate worse physical health
CES-D scale	From 0~30. Higher values indicate worse mental health
Independent variables	
SC	From 0~1; Social capital index
Individual characteristics	
Age	Age in complete years
Female	=1 if female, 0 if male
Urban	=1 if having urban household registration, 0 otherwise
Illiterate(reference)	=1 if illiterate, 0 otherwise
Primary	=1 if primary education, 0 otherwise
Secondary	=1 if secondary education, 0 otherwise
University	=1 if university education, 0 otherwise
Married	=1 if married, 0 otherwise
Insured	=1 if having health insurance, 0 otherwise
Smoke	=1 if ever smoke, 0 otherwise
Drink	=1 if ever drink, 0 otherwise
Household Characteristics	
Family size	Total number of household members
Log income	Log of per capita income in the household
Water	=1 if connected to tap water, 0 otherwise
Sewage	=1 if connected to sewage, 0 otherwise
Community characteristics	
CSC	Community social capital
Hospital	Average number of hospitals per 1000 population
Activity center	Average number of activity centers per 1000 population
Bus distance	Distance to the nearest bus station
Road	=1 if having a road passing through, 0 otherwise

3.3 *Descriptive statistics*

Table 2 presents descriptive statistics of all variables. The results show that 28.9% of the total sample report bad health. The average values of the ADL index and the CES-D scale are 0.145 and 8.023 respectively. The mean social capital index is 0.081, suggesting a relatively low level of social capital among Chinese older adults. In the sub-sample statistics, females on average are more likely to report bad health, higher ADL index (worse physical health) and higher CES-D scale (worse mental health) than males. Females tend to have lower social capital than

[8] In line with current literature, we calculate community social capital by averaging individual social capital index at the community level.

males. Compared to younger-old adults (aged 45 to below 60), older-old adults (aged 60 or above) have worse general health, higher ADL index, higher CES-D scale, and lower social capital. Furthermore, urban older adults are more likely to report better health and have higher social capital than their rural counterparts. In terms of other explanatory variables, females tend to have lower levels of education and fewer unhealthy behaviors than males. Older-old adults have lower levels of education and income than younger-old adults. We also find that urban older adults enjoy higher levels of education and higher income than their rural counterparts.

Table 2

Descriptive statistics

		By gender		By age cohort		By area of residence	
	All	female	male	45-60	60+	urban	rural
Bad health	0.289	0.331	0.245	0.238	0.357	0.192	0.321
ADL index	0.145	0.171	0.117	0.093	0.211	0.123	0.152
CES-D scale	8.023	8.719	7.317	7.646	8.546	6.718	8.472
SC	0.081	0.077	0.087	0.096	0.064	0.127	0.068
Age	60.573	59.490	61.735	52.633	69.953	61.029	60.438
Female	0.517	1.000	0.000	0.554	0.475	0.510	0.519
Urban	0.221	0.218	0.225	0.211	0.234	1.000	0.000
Illiterate	0.428	0.591	0.254	0.362	0.506	0.226	0.486
Primary	0.361	0.266	0.461	0.350	0.372	0.360	0.361
Secondary	0.197	0.135	0.263	0.273	0.106	0.366	0.149
University	0.013	0.006	0.021	0.013	0.013	0.048	0.003
Married	0.832	0.804	0.863	0.929	0.718	0.842	0.830
Insured	0.897	0.889	0.905	0.915	0.876	0.853	0.909
Smoke	0.351	0.029	0.696	0.329	0.376	0.337	0.355
Drink	0.397	0.163	0.652	0.399	0.395	0.402	0.396
Family size	3.206	3.201	3.213	3.229	3.179	2.684	3.355
Log income	7.851	7.774	7.935	8.148	7.501	8.433	7.692
Water	0.785	0.789	0.780	0.801	0.765	0.945	0.739
Sewage	0.449	0.458	0.439	0.458	0.438	0.879	0.317
CSC	0.085	0.082	0.081	0.089	0.079	0.121	0.073
Hospital	1.351	1.335	1.369	1.364	1.335	1.034	1.454
Activity center	5.181	5.128	5.238	5.572	4.693	3.781	5.635
Bus distance	2.117	2.088	2.149	2.117	2.118	1.063	2.433
Road	0.943	0.946	0.940	0.948	0.938	0.994	0.928

4. Empirical strategy

We first pool the observations from the 2008 and 2012 waves of the CHARLS and estimate the following standard Pooled Ordinary Least Squares (POLS) model:

$$H_{it} = \alpha + \beta SC_{it} + X_{it}\delta + D_i + D_t + \varepsilon_{it} \quad (1)$$

where H_{it} measures health outcomes of individual i in year t. SC_{it} denotes the social capital index of individual i at year t, X_{it} a vector of individual-, household-, and community-level variables either time-variant or time-invariant, D_i a province dummy, D_t a year dummy, and ε_{it} the error term. β is the coefficient of interest.

The POLS estimate of β is inconsistent if social capital SC_{it} is correlated with the error term ε_{it}. As previously discussed, endogeneity may arise from several sources. Omitted variables, such as time preference and personal traits, may have an impact on both social capital and health. In addition, reverse causality may be a concern as healthier people are more likely to participate in social activities and have higher social capital.

To address bias caused by time-invariant unobservables, we estimate a model with individual fixed effects as follows:

$$H_{it} = \alpha + \beta SC_{it} + K_{it}\delta + D_t + a_i + \mu_{it} \quad (2)$$

where H_{it} denotes health outcomes, SC_{it} the individual social capital index, K_{it} a vector of time-varying individual-, household-, and community variables, D_t a year dummy, a_i an individual fixed effect, and μ_{it} the error term.

Although FE estimation can correct for the bias caused by time-invariant unobservables, the biases coming from time-variant unobservables and reverse causality are still a concern. One common solution to this problem is the IV approach. A proper IV identification strategy requires the instrument to be correlated with the endogenous social capital variable (relevance condition) and at the same time does not have a direct, independent effect on individual health (exclusion restriction condition). The first stage can be formularized as follows:

$$SC_i = \alpha + X_i\gamma + \varphi Z_i + v_i \qquad (3)$$

In the second stage, we use the predicted value of social capital in an equation similar to (1). The credibility of the IV approach depends on the validity of instrumental variable Z_i, which is excluded in the second stage.

Our IV is whether a respondent's surname belongs to the common surnames in the village.[9] In CHARLS, village leaders are asked whether the village has common surnames, and if they answer "yes," they are further asked to list the three most common surnames.[10] We combine the information on the respondent's surname and the three most common surnames in the village to construct the IV. If a respondent's surname belongs to one of the three most common surnames in the village, the IV is denoted as 1, and 0 otherwise. The surname of a female respondent is coded as the same as her husband's because they tend to share the same social and family networks. Since having common surnames is only a phenomenon in rural villages, this question is only asked for rural respondents. In addition, for reasons of data confidentiality, we only have access to the rural samples of wave

[9] In rural China, although households with the same surname belong to the same ancestry, they live independently under different roofs in the village. For households with many sons, when the sons become adults and get married, they usually move out and form a new household within the village.

[10] The criteria for common surname are that the percentage of local persons who have the same surname is above 20%.

2008 for the IV analysis.[11] Therefore, the IV results represent a local average treatment effect and are a complement to FE estimation.

The validity of the relevance condition of our IV rests on the hypothesis that having a common surname is correlated with a higher social capital index. Surnames in China are usually based on the same lineage identity and act as "glue" to hold family members together. A common surname usually embodies a strong trust and solidarity among its members. It could provide more opportunities for rural residents to communicate with each other and participate more in community affairs (Lei et al., 2014). Therefore, if a respondent has a common surname, we can naturally hypothesize that he/she could have a higher social capital than those who do not.

However, if the common surname is endogenous to health, the exclusion restriction condition will fail and estimation bias will ensue. There are three arguments as to why we expect our instrument to satisfy the exclusion restriction condition. First, the surnames in China have been stable and well preserved through the male line for generations due to the prevalence of Confucian culture in which people do not change their surnames (Du et al., 1992). No surname or lineage has an inherent superiority or inferiority (Baker, 1979). Second, composition of surnames in rural villages is relatively stable as the household mobility is restricted by the household registration (*hukou)* system.[12] Rural residents are tied to the land where they were born and unlikely to move between different places (Pan, 2011). Especially for the older adults in the rural village, the cultural tradition of "*fallen leaves return to root (Luo Ye Gui Gen)*" makes them even more reluctant to move out of their home villages. Therefore, the composition of surnames in a village is a pre-determined condition rather than a result of personal choices.[13] Third, there may not exist self-selection effect in marriage markets. For example, if males with common surnames have higher social status, it is more likely for them to get married with the "healthy" females. This will lead to an upward bias of our estimates. We argue that the self-selection effect in marriage markets tends to diminish as the mean marriage age of older adults in our sample is 22.44, while the mean value of their current age is 60.57.

In order to further assess whether our instrument is as good as random we tested for differences of characteristics of respondents (Table 3). The comparison of average characteristics between individuals with or without common surnames (CS) does not suggest any consistent differences. An exception is the percentage of

[11] The sample size was further reduced after deleting the observations with missing values. Nonetheless, the mean value of the variables in the IV sample (Table 3) and the rural sample (Table 2) is very similar, indicating that the missing value is random in the IV sample.

[12] It should be noted that the *hukou* system has been partially relaxed in recent years. Many rural surplus young laborers go to cities to find jobs, but the rural older adults are unlikely to do the job in cities because of their old age. Moreover, for the rural young labors working in the city, it is extremely difficult for them to get the urban *hukou* and settle in the city. Therefore, the number of registered people in rural village remains relatively unchanged.

[13] More importantly, villages in rural China were set up administratively and arbitrarily in 1949 out of the needs of collective farming rather than the blood connections among people (Pan, 2011). So, there is no self-selection effect in the surname composition of the village.

secondary education, which is significantly (at the 10% level) lower in individuals with common surnames than in individuals without common surnames. The difference is, however, only 0.039, thus economically not very significant. Moreover, we control for education in our regressions. At the bottom of Table 3, there are two other significantly different community variables: hospital and activity center. However, the hospital and activity center do not explain health effects as the variables are insignificant throughout our main regression results reported below. For other community-level variables, two-sided t-tests do not reveal any significant, systematic differences.

Table 3
Characteristics of respondents by the IV

	All	Without CS	With CS	Difference[a]
SC	0.069	0.076	0.086	-0.010*
Age	58.808	58.487	59.299	-0.812
Female	0.503	0.504	0.501	0.003
Illiterate	0.496	0.493	0.499	-0.006
Primary	0.338	0.327	0.355	-0.028
Secondary	0.163	0.178	0.139	0.039*
University	0.004	0.001	0.007	-0.006
Married	0.844	0.848	0.838	0.010
Insured	0.937	0.946	0.924	0.022
Smoke	0.415	0.418	0.409	0.009
Drink	0.419	0.429	0.404	0.025
Family size	3.370	3.383	3.298	0.085
Log income	7.696	7.733	7.638	0.095
Water	0.635	0.641	0.627	0.014
Sewage	0.189	0.196	0.178	0.018
CSC	0.071	0.070	0.073	-0.003
Hospital	1.741	1.910	1.483	0.427***
Activity center	6.056	5.133	7.464	-2.331***
Bus distance	3.581	3.762	3.304	0.458
Road	0.907	0.912	0.899	0.013
N	1,124	679	445	

[a] Two-sided t-test
*** p<0.01.
** p<0.05.
* p<0.1.

One of our dependent variables (bad health) is expressed as a binary value, either 1 (bad health) or 0 (good health). The use of a two stage least squares (2SLS) model that works with continuous dependent variables will yield inconsistent results. It has been suggested that the two-stage residual inclusion (2SRI) approach is most suited for IV estimation with a non-linear second-stage regression (Blundell and Powell, 2004; Terza et al., 2008). Bhattacharya et al. (2006) show that in a test on simulated data for binary outcome variables, the 2SLS estimator exhibited greater bias than the 2SRI estimator. Based on the Health and Retirement data in the US, Basu and Coe (2015) also find that, compared to the 2SRI estimator, the 2SLS estimator produced substantially different and potentially biased results for binary outcome variables. Therefore, we believe the 2SRI to be a better choice for the analysis of the "bad health" model. For the other two continuous outcome variables (the ADL index and the CES-D scale), we will still use the 2SLS estimator. The main difference between 2SRI and 2SLS is that in 2SRI, the actual endogenous variable (rather than the predicted value of the endogenous variable) and the residuals from the first stage estimation are included as additional regressors in the second stage equation. As residuals are correlated with the unobservable characteristics that affect both the endogenous variable (social capital) and the dependent variable in the second-stage regression (bad health), their inclusion ensures that the social capital coefficient in the second-stage regression only reflects the causal effect of social capital on health.

5. Results, robustness checks, and heterogeneity

5.1 Main results

5.1.1 OLS regressions

Table 4 reports the OLS estimates of the effect of social capital on three health measures. Controlling for all observable variables, year dummy and province dummy, social capital is significantly associated with health. A one standard-deviation increase in social capital is associated with a 0.23 standard-deviation decrease in reporting bad health, a 0.142 standard-deviation decrease in the ADL index and a 0.295 standard-deviation decrease in the CES-D scale.[14] However, we must cautiously interpret this association as OLS is unable to control for the possible unobserved heterogeneity and reverse causality bias. Next, we use FE estimation to address the bias caused by time-invariant unobservables.

Table 4
OLS estimates of effect of social capital on health

	Bad health[a]	ADL index	CES-D scale[b]
SC	-0.230***	-0.142***	-0.295**
	(0.051)	(0.021)	(0.116)
Age	0.005***	0.006***	0.002
	(0.001)	(0.000)	(0.002)
Female	0.071***	0.055***	0.146***
	(0.018)	(0.009)	(0.043)
Urban	-0.001	0.008	0.011
	(0.023)	(0.009)	(0.048)
Primary	-0.026	-0.017**	-0.069**
	(0.017)	(0.007)	(0.031)
Secondary	-0.066***	-0.026***	-0.202***
	(0.022)	(0.008)	(0.055)
University	-0.057	-0.034*	-0.244**
	(0.044)	(0.020)	(0.119)
Married	-0.048*	-0.022**	-0.255***
	(0.024)	(0.009)	(0.040)
Insured	0.032	-0.002	0.028
	(0.024)	(0.011)	(0.050)
Smoke	0.020	-0.003	0.072*
	(0.019)	(0.009)	(0.043)
Drink	-0.048***	-0.004	-0.041
	(0.015)	(0.006)	(0.034)

[14] We also estimate the untransformed ADL and CES-D using Poisson regression (which takes into account that untransformed ADL and CES-D can only be zero or a positive number) as a sensitivity check and find that our conclusions remain unaffected. The results are available upon request.

Family size	0.004	0.006***	0.010
	(0.005)	(0.002)	(0.011)
Log income	-0.012***	-0.003**	-0.018***
	(0.003)	(0.001)	(0.007)
Water	-0.069***	-0.035***	-0.163***
	(0.023)	(0.012)	(0.056)
Sewage	-0.062**	-0.022***	-0.028
	(0.024)	(0.008)	(0.047)
CSC	-0.145	0.079	-0.866**
	(0.223)	(0.086)	(0.419)
Hospital	0.003	-0.001	0.007
	(0.003)	(0.001)	(0.007)
Activity center	0.001	0.000	0.001
	(0.001)	(0.001)	(0.003)
Bus distance	0.002	0.000	0.007**
	(0.001)	(0.001)	(0.003)
Road	-0.113**	-0.023	0.142
	(0.056)	(0.020)	(0.121)
Year 2012	-0.002	0.049***	-0.073**
	(0.014)	(0.007)	(0.032)
Zhejiang	-0.126***	-0.078***	-0.338***
	(0.022)	(0.008)	(0.042)
Constant	0.373***	-0.108***	2.360***
	(0.092)	(0.037)	(0.213)
R^2	0.133	0.294	0.143
N	3,920	3,920	3,215

[a] Bad health is estimated by a linear probability model. The results do not change substantially using a logit model.
[b] CES-D scale is log transformed.
Robust standard errors corrected for clustering on community are in parentheses.
*** p<0.01.
** p<0.05.
* p<0.1.

5.1.2 Individual fixed effects

Table 5 presents the individual FE estimation. We start with a parsimonious model where we only examine the impact of social capital on health. Then, we add a set of control variables. Our results show that social capital has a significant effect on having bad health and the ADL index. A one standard-deviation increase in social capital is associated with a 0.186 standard deviation decrease in reporting bad health and a 0.075 standard deviation decrease in the ADL index. However, there appears to be no effect of social capital on the CES-D scale. There are two possible explanations. First, the relationship between social capital and the CES-D scale may be driven by a common, time-invariant unobserved heterogeneity such as personal traits. The more optimistic and extroverted people are more likely to participate in social activities and at the same time less likely to have depressive feelings. Second, participation in volunteer organizations in China could often be simply a practice where the "young old" take care of the "very old" (Norstrand and Xu, 2012), which is likely to reduce the beneficial effects of social capital on mental health.

Table 5

FE estimates of effect of social capital on health

	Bad health[a]		ADL index		CES-D scale[b]	
SC	-0.197**	-0.186*	-0.127***	-0.075**	-0.017	0.081
	(0.097)	(0.104)	(0.036)	(0.036)	(0.165)	(0.210)
Age		0.004		0.003		0.009
		(0.008)		(0.006)		(0.016)
Female		0.148**		0.222***		-0.242
		(0.068)		(0.035)		(0.170)
Urban		0.055**		0.033**		0.016
		(0.027)		(0.016)		(0.081)
Primary		-0.068*		-0.010		-0.274**
		(0.041)		(0.014)		(0.115)
Secondary		-0.112*		-0.021		-0.303*
		(0.058)		(0.023)		(0.168)
University		-0.136*		0.009		0.032
		(0.076)		(0.037)		(0.403)
Married		-0.068		-0.033		-0.295*
		(0.066)		(0.032)		(0.152)
Insured		0.001		-0.006		0.043
		(0.042)		(0.017)		(0.078)
Smoke		0.017		0.019		-0.017
		(0.063)		(0.021)		(0.103)
Drink		-0.083**		0.007		-0.006
		(0.037)		(0.013)		(0.068)
Family size		-0.004		0.004		0.036*
		(0.011)		(0.004)		(0.019)
Log income		0.001		0.002		0.007
		(0.004)		(0.002)		(0.009)
Water		-0.008		-0.008		-0.073
		(0.039)		(0.013)		(0.093)
Sewage		0.014		-0.009		0.020
		(0.063)		(0.017)		(0.183)
CSC		0.616*		0.130		0.703
		(0.362)		(0.165)		(1.015)
Hospital		-0.004		-0.001		0.006
		(0.007)		(0.003)		(0.019)
Activity center		0.008*		-0.001		-0.001
		(0.004)		(0.002)		(0.013)
Bus station		0.002		-0.001		0.005
		(0.002)		(0.001)		(0.007)
Road		-0.352***		-0.034		-0.075
		(0.057)		(0.047)		(0.305)
Year 2012		0.017		0.060**		-0.084
		(0.037)		(0.023)		(0.084)
Constant	0.308***	0.383	0.157***	-0.131	1.911***	1.781*
	(0.009)	(0.476)	(0.003)	(0.342)	(0.015)	(0.986)
Sargen-Hansen test [*p*-value][c]	18.825 [0.000]	64.961 [0.000]	31.963 [0.000]	170.147 [0.000]	23.048 [0.000]	58.892 [0.000]
N	3,920	3,920	3,920	3,920	3,215	3,215

[a] Bad health is estimated by a linear probability model. The results do no change substantially using a xtlogit model.

[b] CES-D scale is log transformed.

[c] Sargen-Hansen test is a test for fixed vs. random effects with robust standard errors clustered on community. The fixed effects model is consistent under H_0 and *Ha*, the random effect model is inconsistent under *Ha*.

Robust standard errors corrected for clustering on community are in parentheses.

*** $p<0.01$.

** $p<0.05$.

* p<0.1.

5.1.3 Instrumental variable

Table 6 reports the results of IV estimates. As expected, the first stage results show that having a common surname increases respondents' social capital index by 0.013 percentage points, which is statistically significant at 5% level. The under-identification test (Anderson canon. corr. LM statistic) rejects the null hypothesis that the IV is under-identified. Thus, our instrument is a significant predictor of respondents' social capital.

The second stage results indicate social capital has a significant effect on general and physical health, but not on mental health. A one standard-deviation increase in social capital leads to a 4.863 standard-deviation decrease in reporting bad health and a 2.226 standard-deviation decrease in physical activity limitations. It is interesting to note that the coefficients on the social capital are larger in the IV regressions than in the OLS and FE estimates, suggesting OLS and FE estimates may underestimate the effect of social capital on health. Suppose that those individuals who are healthy are also more likely to participate in social activities, this would cause an upward bias (in absolute value) of the estimated coefficients of social capital. An alternative argument is that, as investment in social capital is driven in part by the opportunity cost of time (Glaeser, et al., 2002; Putnam, 2000), individuals with relatively worse health have increased incentives to participate in social activities as their opportunity cost of time is lower. This would imply a downward bias of the coefficients of social capital. The patterns of the coefficients in our IV estimation seem to support the latter interpretation.

One of the most common concerns about the IV analysis is the weak instrument problem. The Cragg-Donald Wald F statistics in our regression is 4.268, much lower than the critical value of 10 proposed by Stock et al. (2002).[15] The weak instrument means the standard test statistics used for inference may be unreliable and point estimates can be imprecise. To dispel this concern, we use the conditional likelihood ratio (CLR) test to ascertain whether our instrument is weak. The CLR test is proposed by Moreira (2003), Mikusheva and Poi (2006) and Moreira (2009), which has a correct coverage probability in confidence regions under weak instruments. If the CLR test has finite confidence intervals, it means our data contain sufficient information on the social capital effect. Otherwise, if the CLR test has an infinite interval, i.e. two intervals $(-\infty, x1)$ and $(x2, +\infty)$, where $x1<x2$ or a single interval $(-\infty, +\infty)$, it means there is a weak instrument problem. The CLR test shows that the effects of social capital on the ADL index and the CES-D scale have finite confidence intervals.[16] Furthermore, as our model has one endogenous variable and one IV, it is a just-identified model. According to Angrist and Pischke (2009, p209), "even with weak instruments, just-identified 2SLS is approximately centered where it should be. We therefore say that just-identified 2SLS is median-unbiased." Thus, we argue that our IV estimates are reliable

[15] It is noteworthy that the low value of F-statistics might be just a small sample problem – i.e. in a bigger sample the F-statistic might look more convincing.

[16] The confidence regions of the two models are [-193.013, -0.696] and [-753.651, 59.501] respectively.

despite the low Wald F statistics.

Table 6

IV estimates of effect of social capital on health

	Bad health(2SRI)		ADL index(2SLS)		CES-D scale(2SLS) [b]	
	1st stage	2nd stage [a]	1st stage	2nd stage	1st stage	2nd stage
SC		-4.863***		-2.226*		-1.014
		(1.65)		(1.219)		(4.464)
Age	-0.002***	-0.003	-0.002***	0.002	-0.002***	0.002
	(0.000)	(0.003)	(0.0004)	(0.002)	(0.0004)	(0.008)
Female	-0.006	0.035	-0.006	0.007	-0.004	0.117
	(0.011)	(0.047)	(0.01)	(0.030)	(0.011)	(0.077)
Primary	-0.006	-0.06*	-0.006	-0.016	-0.007	-0.041
	(0.007)	(0.034)	(0.007)	(0.019)	(0.008)	(0.055)
Secondary	0.031**	0.05	0.03**	0.054	0.033**	-0.198
	(0.014)	(0.082)	(0.014)	(0.045)	(0.014)	(0.164)
University	0.084	0.585***	0.087	0.219	0.085	0.059
	(0.091)	(0.132)	(0.091)	(0.229)	(0.094)	(0.545)
Married	-0.008	-0.130***	-0.009	-0.046**	-0.007	-0.301***
	(0.009)	(0.047)	(0.009)	(0.021)	(0.011)	(0.066)
Insured	-0.005	-0.039	-0.004	-0.007	-0.005	-0.149
	(0.014)	(0.065)	(0.013)	(0.033)	(0.011)	(0.105)
Smoke	0.002	0.0003	0.003	-0.040*	0.005	-0.029
	(0.010)	(0.05)	(0.009)	(0.022)	(0.011)	(0.077)
Drink	0.013*	0.009	0.011	0.033	0.015*	0.022
	(0.007)	(0.041)	(0.007)	(0.023)	(0.008)	(0.082)
Family size	0.001	0.007	0.0007	0.005	0.001	0.004
	(0.002)	(0.009)	(0.002)	(0.005)	(0.002)	(0.021)
Log income	0.003*	-0.002	0.003*	0.003	0.003*	-0.015
	(0.001)	(0.008)	(0.001)	(0.005)	(0.002)	(0.017)
Water	0.012*	0.073*	0.012**	0.010	0.009	-0.131*
	(0.007)	(0.04)	(0.006)	(0.024)	(0.007)	(0.074)
Sewage	0.005	0.012	0.009**	0.015	0.019**	-0.095
	(0.005)	(0.075)	(0.004)	(0.021)	(0.008)	(0.118)
CSC	0.825***	-1.418*	0.824***	-1.991*	0.864***	-1.664
	(0.070)	(0.777)	(0.056)	(1.016)	(0.061)	(3.979)
Hospital	0.002*	0.006	0.001**	0.001	0.001**	0.022*
	(0.001)	(0.007)	(0.0004)	(0.003)	(0.0006)	(0.013)
Activity center	-0.000	0.003	-0.002***	-0.000	-0.003***	0.019
	(0.000)	(0.004)	(0.0007)	(0.004)	(0.0001)	(0.020)
Bus station	-0.000	-0.006*	-0.0001	-0.000	-0.0005	-0.002
	(0.000)	(0.003)	(0.0002)	(0.001)	(0.0004)	(0.006)
Road	0.005	-0.05	0.009**	-0.025	0.009	0.069
	(0.006)	(0.117)	(0.004)	(0.034)	(0.007)	(0.135)
Zhejiang	0.006	-0.165**	0.007	-0.092***	0.004	-0.343***
	(0.007)	(0.075)	(0.006)	(0.017)	(0.008)	(0.079)
IV	0.013**		0.013**		0.013**	
	(0.006)		(0.006)		(0.006)	
Constant	0.071**	1.548**	0.074**	0.090	0.088**	2.674***
	(0.033)	(0.705)	(0.031)	(0.138)	(0.042)	(0.513)
Anderson canon. corr. LM statistic	----	----	4.33[0.037]		4.35[0.037]	
Cragg-Donald Wald F statistics	----	----	4.268		4.268	
Endogeneity test	----	----	11.836[0.0006]		0.221[0.638]	
N	1124		1124		948	

[a] The coefficient of the second stage is the marginal effect.
[b] CES-D scale is log transformed.
---- means STATA cannot report such statistics in 2SRI.
Robust standard errors corrected for clustering on community are in parentheses.
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$.

5.2 *Robustness checks*

We also carry out a series of robustness checks on the FE estimation.[17] One of the concerns about the social capital index is the potential measurement error. If individuals misreport their social activities in a systematic way, our estimations would be biased. For example, an individual with cognitive impairment and bad health may have difficulties in recalling past events, thus may underreport social activities. Our estimations would then be biased downward. To test for measurement error, we recode the social capital index into a dummy variable. If an individual participated in one of the eight activities in the past month, the social capital index is recoded as 1, and 0 otherwise. By this way, we can significantly reduce the likelihood of misreporting errors and check the robustness of the FE estimates. Table 7 presents the results with a dichotomous indicator of social capital. Compared to Table 5, the effects of social capital on bad health and the ADL index are slightly reduced but still statistically significant. Therefore, our main findings that social capital has a positive effect on general and physical health remain intact.

Table 7
FE estimates by using alternate definition of social capital

	Bad health[a]	ADL index	CES-D scale[b]
SC dummy	-0.060**	-0.027***	0.019
	(0.027)	(0.008)	(0.046)
Age	0.003	0.003	0.009
	(0.008)	(0.006)	(0.016)
Female	0.120*	0.211***	-0.231
	(0.068)	(0.034)	(0.157)
Urban	0.056**	0.033**	0.016
	(0.026)	(0.016)	(0.081)
Primary	-0.067	-0.010	-0.273**
	(0.041)	(0.015)	(0.115)
Secondary	-0.112*	-0.021	-0.302*
	(0.058)	(0.024)	(0.168)
University	-0.135*	0.009	0.031
	(0.081)	(0.040)	(0.401)
Married	-0.062	-0.031	-0.297*
	(0.066)	(0.032)	(0.152)
Insured	0.001	-0.005	0.043
	(0.041)	(0.017)	(0.078)
Smoke	0.016	0.018	-0.017
	(0.062)	(0.021)	(0.103)
Drink	-0.082**	0.007	-0.006
	(0.037)	(0.013)	(0.068)
Family size	-0.005	0.003	0.036*

[17] We use the fixed effects model to carry out the robustness checks for two reasons. First, compared to the IV model, the fixed effects model represents lower bound estimates. Second, the sample size of IV analysis is small and may suffer from explanatory power.

	(0.011)	(0.004)	(0.019)
Log income	0.001	0.002	0.007
	(0.004)	(0.002)	(0.009)
Water	-0.008	-0.008	-0.073
	(0.040)	(0.013)	(0.093)
Sewage	0.012	-0.010	0.020
	(0.064)	(0.017)	(0.183)
CSC	0.671*	0.161	0.711
	(0.353)	(0.165)	(1.009)
Hospital	-0.004	-0.001	0.006
	(0.007)	(0.003)	(0.019)
Activity center	0.007	-0.001	-0.001
	(0.004)	(0.002)	(0.013)
Bus station	0.002	-0.001	0.005
	(0.002)	(0.001)	(0.007)
Road	-0.356***	-0.036	-0.072
	(0.057)	(0.047)	(0.305)
Year 2012	0.015	0.059**	-0.084
	(0.037)	(0.023)	(0.084)
Constant	0.417	-0.117	1.764*
	(0.475)	(0.342)	(0.992)
Sargen-Hansen test [*p*-value][c]	64.321 [0.000]	147.696 [0.000]	70.055 [0.000]
N	3,920	3,920	3,215

[a] Bad health is estimated by a linear probability model. The results do no change substantially using a xtlogit model.
[b] CES-D scale is log transformed.
[c] Sargen-Hansen test is a test for fixed vs. random effects with robust standard errors clustered on community. The fixed effects model is consistent under H_0 and *Ha*, the random effect model is inconsistent under *Ha*.
Robust standard errors corrected for clustering on community are in parentheses.
*** $p<0.01$.
** $p<0.05$.
* $p<0.1$.

Additionally, as noted in previous sections, one source of endogeneity unaddressed by the FE model might come from the unobservable, time-varying shocks that may simultaneously affect both social capital and health. These shocks can occur at individual, family, and community levels. For individuals, the shocks might be serious accidents that cause a reduction of social capital and deterioration of health. For families, they might be serious accidents to other family members, which may crowd out respondents' time for social participation and incur a negative impact on their own health. For communities, they might be natural disasters that may affect respondents' social participation and health simultaneously.[18]

To avoid these confounding cases, we checked the sensitivity of our estimates by adding the individual, family, and community-level shocks into our empirical specifications. If time-varying unobservable shocks represent the main cause for the correlation between social capital and health, we would expect the social capital coefficients to decline significantly after including those exogenous shocks. We use the detailed information on individual and

[18] Regarding the relationship between social capital and natural disasters, Toya and Skidmore (2014) and Yamamura (2015) find natural disasters help social capital formation. However, according to Aida et al. (2013), natural disasters destroy not only physical environments but social networks and relationships in communities. Thus, they have a negative effect on social capital. Our data shows natural disasters and social participation are negatively correlated (-0.097, p=0.000), suggesting they tend to have a negative effect on social capital among older adults in China.

community-level shocks in CHARLS and carry out separate estimations on shocks at each level.[19] Table 8 presents the separate estimations with the inclusion of these time-varying shocks. The magnitude of the coefficients of social capital remains relatively unchanged and statistically significant, suggesting time-varying, unobserved heterogeneity should not pose considerable threat to our FE results.

Table 8

Robustness checks by including individual, family and community-level shocks

	Bad health[a]	ADL index	CES-D scale[b]
Panel A: Individual-level injury			
SC dummy	-0.059**	-0.018**	0.019
	(0.026)	(0.008)	(0.046)
Individual injury	0.074*	0.053**	0.088
	(0.037)	(0.022)	(0.084)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	66.203 [0.000]	146.13 [0.000]	72.149 [0.000]
Panel B: Family-level injury			
SC dummy	-0.061**	-0.027***	0.016
	(0.027)	(0.008)	(0.047)
family member's injury	0.081*	0.042**	0.281***
	(0.047)	(0.02)	(0.090)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	62.79 [0.000]	153.747 [0.000]	81.324 [0.000]
Panel C: Community-level disaster			
SC dummy	-0.067**	-0.025***	0.024
	(0.028)	(0.008)	(0.049)
Community disaster	-0.043	-0.002	0.045
	(0.028)	(0.017)	(0.061)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	64.321 [0.000]	139.892 [0.000]	75.052 [0.000]
N	3,920	3,920	3,215

[a] Bad health is estimated by a linear probability model. The results do no change substantially using a xtlogit model.
[b] CES-D scale is log transformed.
[c] Sargen-Hansen test is a test for fixed vs. random effects with robust standard errors clustered on community.
The fixed effects model is consistent under H_0 and *Ha*, the random effect model is inconsistent under *Ha*.
All control variables in Table 5 are included.
Robust standard errors corrected for clustering on community are in parentheses.
*** p<0.01.
** p<0.05.
* p<0.1.

5.3 Heterogeneity by gender, age, and area of residence

To explore if social capital has differential effects by gender, age, and area of residence, we include the interaction terms into the FE model. Table 9 shows that none of gender differences were statistically significant when tested in a fully interacted model. This is in contrast to the literature on gender differences in other countries, which finds that women tend to derive fewer health benefits from social capital compared to men (e.g., Berry and

[19] CHARS collects individual and family-level shock information by asking "Have you ever been in a traffic accident or any other kind of major accidental injury and received medical treatment?" If respondents answer "yes", we recode it as 1, otherwise 0. As for community-level shock information, it asks "Has your village/community experienced serious natural disasters in the past five years?" If village heads answer "yes", we recode it as 1, otherwise 0.

Welsh, 2010; Ferlander and Mäkinen, 2009). One of the arguments for the gender differences is that women usually play a supportive and subordinate role in social interactions, which may decrease their sense of control and enhance feelings of being burdened and emotional stress. However, some studies find that with increased marketization and economic development in China, the role of women and men in the society is becoming equally divided (Yu and Xie, 2012; Matthews and Nee, 2000). This may explain why we find no differential effect by gender on health.

Table 9

FE estimates of heterogeneous effects by gender, age and area of residence

	Bad health[a]	ADL index	CES-D scale[b]
Panel A: Female-male difference			
SC	-0.189	-0.104**	-0.098
	(0.127)	(0.045)	(1.512)
SC×female	0.007	0.067	0.413
	(0.19)	(0.052)	(0.335)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	61.671 [0.000]	57.606 [0.000]	55.174 [0.000]
Panel B: Age difference			
SC	-0.102	0.003	0.176
	(0.122)	(0.038)	(0.255)
SC×age60	-0.250	-0.232***	-0.306
	(0.155)	(0.058)	(0.401)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	58.691 [0.000]	60.01 [0.000]	68.712 [0.000]
Panel C: Urban-rural difference			
SC	-0.197	-0.04	-0.089
	(0.132)	(0.042)	(0.226)
SC×urban	0.035	-0.098	0.474
	(0.168)	(0.069)	(0.408)
Control variables	YES	YES	YES
Sargen-Hansen test [*p*-value][c]	61.682 [0.000]	59.751 [0.000]	59.172 [0.000]
N	3,920	3,920	3,215

[a] Bad health is estimated by a linear probability model. The results do no change substantially using a xtlogit model.
[b] CES-D scale is log transformed.
[c] Sargen-Hansen test is a test for fixed vs. random effects with robust standard errors clustered on community. The fixed effects model is consistent under H_0 and *Ha*, the random effect model is inconsistent under *Ha*.
All control variables in Table 5 are included.
Robust standard errors corrected for clustering on community are in parentheses.
*** p<0.01.
** p<0.05.
* p<0.1.

Table 9 also reveals that the effect of social capital on the ADL index is significantly stronger for older-old adults. The negative sign of the interaction term indicates that social capital leads to a greater decrease in the ADL index (thus better physical health) for older-old adults than for younger-old adults. This finding is consistent with Muckenhuber et al. (2013), who also find that social capital affects health of older people more strongly than that of younger people. The possible explanation is that, compared to younger-old adults, older-old adults lack better coping and compensatory opportunities to make up for their lower social capital. Thus, social capital plays a greater role in the health production for older-old adults.

Our estimates further show that the effect of social capital on health is not heterogeneous across area of residence. This finding stands in contrast to an earlier study on China (Norstrand and Xu, 2012), which finds that urban people benefit more from social capital than rural people do. They argue that limited resources in rural areas make it difficult for rural people to access certain material resources that are essential to their health. Therefore, rural people might not be able to benefit as much from their family members, friends, and people around them, as they are also resource constrained. However, our results are not surprising as China has initiated a new round of health care reform in recent years, which has invested heavily in health care infrastructure in rural areas and hence narrowed the resources distribution gap between urban and rural areas. This narrowing gap could have ameliorated the contributing role of social capital in the urban-rural health divide.

6. Conclusion

Employing detailed longitudinal data from the China Health and Retirement Survey, this paper examines the causal effect of social capital on health among Chinese older adults. Although many studies have shown the positive relationship between social capital and health, this is the first longitudinal study using FE and IV estimations to examine their potential causal relationship in a nationally representative Chinese population setting. Additionally, our study contributes to the literature by exploring the heterogeneous effects of social capital on health by gender, age, and area of residence.

We find evidence that social capital has a significant and positive effect on general and physical health, but not on mental health. Our estimates pass a series of robustness checks. We also find some evidence on the heterogeneous effects of social capital on health. Our results indicate that the social capital effect on physical health tends to be stronger for the older old adults. However, we find no differential effects by gender and area of residence.

A central message from our study is that, confronted with health challenges caused by population aging, social capital does serve as a key instrument for health promotion among older adults in China. The implication of our study is that government can improve health of the older population by investing in social capital. This can be achieved through the following approaches. First, directly provide funding or subsidies to promote individuals' social capital formation, for example, organizing community events to facilitate social interaction and community participation. Second, create favorable policies for the establishment of a variety of voluntary organizations and community groups to stimulate community participation and develop individuals' social capital. Third, pay attention to the heterogeneous effects of social capital on health. More efforts need to be exerted in strengthening social capital formation for females and younger-older adults.

Our study is subject to several limitations. First, we focus only on the structural component of social capital. Social capital is multi-dimensional, so our measurement is not comprehensive and cannot reflect the complex relationship between social capital and health. Future research can explore the complex relationship between different forms of social capital and different health outcomes. Second, the subjects in our dataset are limited to urban and rural residents aged 45 or higher and our IV analysis is restricted to the rural samples only. Thus, our conclusion cannot be generalized to the whole population. Third, our study only empirically examines the causal relationship between social capital and health. The mechanisms underlying this relationship were not explored. We leave them for future research.

References

Aida, J., Kawachi, I., Subbramanian, S.V., Kondo, K., 2013. Disaster, Social Capital, and Health. In: Kawachi, I., Takao, S., Subramanian, S.V. (Eds.), Global perspectives on social capital and health. Springer.

Alesina, A., La Ferrara, E., 2000. Participation in heterogeneous communities. Q. J. Econ. 115, 847-904.

Angrist, J., Pischeke, J., 2009. Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Princeton, NJ.

Baker, H., 1979. Chinese Family and Kinship. Columbia University Press, New York, NY.

Basu, A., Coe, N., 2015. 2SLS vs. 2SRI: Appropriate methods for rare outcome and/or rare exposure, Available at https://bfi.uchicago.edu/sites/default/files/research/RareOutcomes%202SLS%20vs%202SRIv3BFI.pdf.

Bhattacharya, J., Goldman, D., McCaffrey, D., 2006. Estimating probit models with self-selected treatments. Stat. Med. 25, 389–413.

Blundell, R.W., Powell, J.L., 2004. Endogeneity in semiparametric binary response models. Rev. Econ. Stud. 71, 655-679.

Borgonovi, F., 2010. A lifecycle approach to the analysis of the relationship between social capital and health in Britain. Soc. Sci. Med. 71, 1927–1934.

Bourdieu, P., 1986. The forms of capital. In: Richardson, J. G. (Eds.), Handbook of theory and research for the sociology of education, Greenwood Press, Westport, CT.

Bratti, M., Mendola, M., 2014. Parental health and Child schooling. J. Health Econ. 35, 94-108.

Coleman, J., 1990. Foundations of Social Theory. Harvard University Press, Cambridge, MA.

D'Hombres, B., Rocco, L., Suhrcke, M., McKee, M., 2010. Does social capital determine health? Evidence from eight transition countries. Health Econ. 19, 56-74.

Du, R., Yuan, Y., Juliana, H., Joanna, M., Cavalli-Sforza, L.L., 1992. Chinese surnames and the genetic differences between North and South China. J. Chin. Ling. Monogr. Ser. 5, 1-66, 68-93.

Duncan, B., Rees, D.I., 2005. The effect of smoking on depressive symptomatology: a reexamination of the data from the National Longitudinal Study of Adolescent Health. Am. J. Epidemiol. 162, 461–470.

Ferlander, S., Mäkinen, I.H., 2009. Social capital, gender and self-rated health: Evidence from the Moscow Health Survey 2004. Soc. Sci. Med. 69, 1323-1332.

Fiorillo, D., Sabatini, F., 2011. Structural social capital and health in Italy. Econ. Hum. Biol. 17, 129-142.

Folland, S., 2007. Does "community social capital" contribute to population health? Soc. Sci. Med. 64, 2342-2354.

Folland, S., 2008. An economic model of social capital and health. Health Econ. Policy Law. 3, 333-348.

Folland, S., Rocco, L., 2014. The economics of social capital and health. World Scientific, Singapore.

Fujisawa, Y., Hamano, T., Takegawa, S., 2009. Social capital and perceived health in Japan: An ecological and multilevel analysis. Soc. Sci. Med. 69, 500-505.

Gertler, P., Gruber, J., 2002. Insuring consumption against illness. Am. Econ. Rev. 92, 51–76.

Glaeser, E.L., Laibson, D., Sacerdote, B., 2002. An economic approach to social capital. Econ. J. 112, 437-458.

Goryakin, Y., Suhrcke, M., Rocco, L., 2014. Social capital and self-reported general and mental health in nine Former Soviet Union countries. Health Econ. Policy Law. 9, 1-24.

Habibov, N.N., Afandi, E.N., 2011. Self-rated health and social capital in transitional countries: multilevel analysis of comparative surveys in Armenia, Azerbaijan, and Georgia. Soc. Sci. Med. 72, 1193–1204.

Harpham, T., Grant, E., Thomas, E., 2002. Measuring social capital within health surveys: Key issues. Health Policy Plann. 17, 106–111.

Hurtado, D, Kawachi, I., Sudarsky, J., 2011. Social capital and self-rated health in Colombia: The good, the bad and the ugly. Soc. Sci. Med. 72, 584-590.

Ichida, Y., Kondo, K., Hirai, H., Hanibuchi, T., Yoshikawa, G., Murata, C., 2009. Social capital, income inequality and self-rated health in Chita peninsula, Japan: a multilevel analysis of older people in 25 communities. Soc. Sci. Med. 69, 489-499.

Islam, M.K., Gerdtham, U., Gullberg, B., Lindstrom, M., Merlo, J., 2008. Social capital externalities and mortality in Sweden. Econ. Hum. Biol. 6, 19-42.

Iversen, T., 2008. An exploratory study of associations between social capital and self-assessed health in Norway. Health Econ. Policy Law. 3, 349-364.

Lei, X., Smith, J.P., Sun, X., Zhao, Y., 2014. Gender differences in cognition in China and reasons for change over time: Evidence from CHARLS. J. Econ. Aging. 4, 46-55.

Leone, T., Hessel, P., 2015. The effect of social participation on the subjective and objective health status of the over-fifties: evidence from SHARE. Aging and Society. DOI: http://dx.doi.org/10.1017/S0144686X15000148.

Lindstrom, M., 2009. Marital status, social capital, material conditions and self-rated health: A population-based study. Health Policy. 93, 172-179.

Lindstrom, M., Mohseni, M., 2009. Social capital, political trust and self-reported psychological health: A population-based study. Soc. Sci. Med. 68, 436-443.

Ljunge, M., 2014. Social capital and health: Evidence that ancestral trust promotes health among children of immigrants. Econ. Hum. Biol. 15, 165-186.

Koutsogeorgou, E., Nyqvist, F., Nygard, M., Cerniauskaite, M., Quintas, R., Raggi, A., Leonardi, M., 2014. Social capital and self-rated health among older adults: a comparative analysis of Finland, Poland and Spain. Aging Soc. DOI: http://dx.doi.org/10.1017/S0144686X13001025.

Matthews, R., Nee, V., 2000. Gender inequality and economic growth in rural China. Soc. Sci. Res. 29, 606-632.

Meng, T., Chen, H., 2014. A multilevel analysis of social capital and self-rated health: evidence from China. Health Place 27, 38-44.

Mikusheva, A., Poi, B.P., 2006. Tests and confidence sets with correct size when instruments are potentially weak. Stata J. 6, 335–347.

Moreira, M.J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71, 1027–1048.

Moreira, M.J., 2009. Tests with correct size when instruments can be arbitrarily weak. J. Econometrics 152, 131–140.

Muckenhuber, J., Stronegger, W.J., Freidl, W., 2013. Social capital affects the health of older people more strongly than that of younger people. Aging Soc. 33, 853-870.

Murayama, H., Nishi, M., Matsuo, E., Nofuji, Y., Shimizu, Y., Taniguchi, Y., Fujiwara, Y., Shinkai, S., 2013. Do bonding and bridging social capital affect self-rated health, depressive mood and cognitive decline in older Japanese? A prospective cohort study. Soc. Sci. Med. 98, 247-252.

Murayama, H., Nofuji, Y., Matsuo, E., Nishi, M., Taniguchi, Y., Fujiwara, Y., Shinkai, S., 2015. Are neighborhood bonding and bridging social capital protective against depressive mood in old age? A multilevel analysis in Japan. Soc. Sci. Med. 124, 171-179.

Nakhaie, R., Arnold, R., 2010. A four year (1996-2000) analysis of social capital and health status of Canadians: The difference that love makes. Soc. Sci. Med. 71, 1033–1044.

National Statistical Bureau of China, 2008. China Statistics Yearbook 2008. China Statistics Publishing House.

Norstrand, J., Xu, Q., 2012. Social capital and health outcomes among older adults in China: The urban-rural dimension. Gerontol. 52, 325-334.

Oxley, H., 2009. Policies for Healthy Aging: An Overview. In: OECD Health Working Papers No. 42.

Pan, Y., 2011. Born with the Right Surname? Lineage Networks and Political and Economic Opportunities in Rural China, Available at http://bus.lsu.edu/McMillin/Working_Papers/pap11_15.pdf.

Petrou, S., Kupek, E., 2008. Social capital and its relationship with measures of health status: Evidence from the Health Survey for England 2003. Health Econ. 17, 127–143.

Poortinga, W., 2006a. Social capital: An individual or collective resource for health? Soc. Sci. Med. 62, 292–302.

Poortinga, W., 2006b. Social relations or social capital? Individual and community health effects of bonding social capital. Soc. Sci. Med. 63, 255–270.

Putnam, R., 1993. Making Democracy Work-Civic Traditions in Modern Italy. Princeton University Press, Princeton, NJ.

Putnam, R., 2000. Bowling alone: The collapse and revival of American community. Simon and Schuster, New York, NY.

Radloff, L., 1977. The CES-D Scale: A Self-Report Depression Scale for research in the general population. Appl. Psycho. Meas. 1, 386-401.

Rocco, L., Fumagalli, E., Suhrcke, M., 2014a. From social capital to health--and back. Health Econ. 23, 586-605.

Rocco, L., 2014b. Trust me, you will be in better health. Health Policy. 116, 123-132.

Ronconi, L., Brown, T., Scheffler, R.M., 2012. Social capital and self-rated health in Argentina. Health Econ. 21, 201-208.

Rose, R., 2000. How much does social capital add to individual health? A survey study of Russians. Soc. Sci. Med. 51, 1421–1435.

Rostila, M., 2007. Social capital and health in European welfare regimes: a multilevel approach. J. Eur. Soc. Policy.17, 223-239.

Scheffler, R., Brown, T., Rice, J., 2007. The role of social capital in reducing non-specific mental distress: the importance of controlling for omitted variable bias. Soc. Sci. Med. 65, 842-854.

Scheffler, R., Brown, T., 2008. Social capital, economics, and health: new evidence. Health Econ. Policy Law. 3, 321-331.

Schultz, J., O'Brein, A.M., Tadesse, B., 2008. Social capital and self-rated health: Results from the US 2006 social capital survey of one community. Soc. Sci. Med. 67, 606–617.

Sirven, N., Debrand, T., 2008. Social participation and healthy aging: An international comparison using SHARE data. Soc. Sci. Med. 67, 2017–2026.

Sirven, N., Debrand, T., 2012. Social capital and health of older Europeans: Causal pathways and health inequalities. Soc. Sci. Med. 75, 1288–1295.

Snelgrove, J.W., Pikhart, H., Stafford, M., 2009. A multilevel analysis of social capital and self-rated health: Evidence from the British Household Panel Survey. Soc. Sci. Med. 68, 1993–2001.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. J. Bus. Econ. Stat. 20, 518–529.

Terza, J., Basu, A., Rathouz, P., 2008. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. J. Health Econ. 27, 531–43.

Toya, H., Skidmore, M., 2014. Do Natural Disasters Enhance Societal Trust? Kyklos 67, 255-279.

Veenstra, G., 2000. Social capital, SES and health: an individual-level analysis. Soc. Sci. Med. 50, 619-629.

Wang, H., Schlesinger, M., Wang, H., William, C., 2009. The flip-side of social capital: The distinctive influences of trust and mistrust on health in rural China. Soc. Sci. Med. 68, 133-142.

Yamamura, E., 2011a. Different effects of social capital on health status among residents: evidence from modern Japan. J. Socio-Econ. 40, 475-479.

Yamamura, E., 2011b. Differences in the effect of social capital on health status between workers and non-workers. Int. Rev. Econ. 58, 385–400.

Yamamura, E., 2015. Natural disasters and social capital formation: The impact of the Great Hanshin-Awaji earthquake. Pap. Reg. Sci. DOI: 10.1111/pirs.12121.

Yamaoka, K., 2008. Social capital and health and well-being in East Asia: A population-based study. Soc. Sci. Med. 66, 885-899.

Yip, W., Subramanian, S.V., Mitchell, A., Lee, D., Wang, J., Kawachi, I., 2007. Does social capital enhance health and well-being? Evidence from rural China. Soc. Sci. Med. 64, 35–49.

Yu, J., Xie, Y., 2012. The varying display of "gender display". Chinese Sociol. Rev. 44, 5-30.

Zhao, Y., Strauss, J., Park, A., Shen, Y., Sun, Y., 2013. China Health and Retirement Longitudinal Study-Pilot Users' Guide, Available at http://charls.ccer.edu.cn/en/page/documentation/2008_pilot.