UNITED NATIONS UNIVERSITY

UNU-WIDER

World Institute for Development Economics Research

Discussion Paper No. 2006/07

Income, Income Inequality, and Health

Evidence from China

Hongbin Li[1] and Yi Zhu[2]

August 2006

Abstract

This paper tests using survey data from China whether individual health is associated with income and community-level income inequality. Although poor health and high inequality are key features of many developing countries, most of the earlier literature has drawn on data from developed countries in studying the association between the two. We find that self-reported health status increases with per capita income, but at a decreasing rate. Controlling for per capita income, we find an inverted-U association between self-reported health status and income inequality, which suggests that high inequality in a community poses threats to health. We also find that high inequality increases the probability of health-compromising behavior such as smoking and alcohol consumption. Most of our findings are robust to different measures of health status and income inequality.

Keywords: health, income, inequality, behaviour, China

JEL classification: D63, I10, O15, O53

[1] Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, People's Republic of China. [2] Department of Economics, Michigan State University, East Lansing, MI, USA.

This study has been prepared within the UNU-WIDER project on Inequality and Poverty in China.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Norway (Royal Ministry of Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

Acknowledgements

We would like to thank the Carolina Population Center for kindly supplying the data and for the World Institute for Development Economics Research for an award. We are very grateful to Loren Brandt, Julan Du, Kai Yuen Tsui, Guanghua Wan, and an anonymous referee for very helpful comments.

UNU-WIDER held an essay competition on Inequality and Poverty in China that attracted over 300 entries and was judged by a selection committee headed by Anthony Shorrocks with members John Bonin, Fang Cai, Shi Li, Justin Lin, and Guanghua Wan. This is one of the prize winning papers.

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Prepared by TechBooks (www.techbooks.com) and by Adam Swallow at UNU-WIDER.
Printed at UNU-WIDER, Helsinki

ISSN 1609-5774
ISBN 92-9190-866-5 (printed publication)
ISBN 92-9190-867-3 (internet publication)

1 Introduction

China has recorded impressive growth over the past 25 years since the introduction of the market economy, and there has been a substantial increase in average living standards. However, in recent years there has been growing concern about the large increase in income inequality during the same period. For example, Bramall (2001) shows that the Gini coefficient for rural China has increased by almost 50 percent from 1980 to 1999. The rising inequality has had and will have important impacts on various aspects of social life, resulting, for example, in frequent social conflicts (Alesina and Perotti, 1996), higher levels of violent crime (Hsieh and Pugh, 1993), and ultimately in a slowing down of economic growth (Aghion *et al.*, 1999). While inequality may affect the society and its economic development in many ways, we focus in this paper on a particular aspect of the socioeconomic effects of inequality, i.e., its impact on health.

The relationship among income, income inequality, and health is an issue that has attracted the attention of a variety of social science disciplines such as economics, sociology, and public health. From an early stage in the debate, it was argued that income has a positive effect on health (Grossman, 1972; Preston, 1975). This is called the absolute income hypothesis. However, some researchers assert that relative income or income inequality plays an equally important role in determining health. According to the relative income hypothesis (or the weak income inequality hypothesis), people who feel more economically disadvantaged than their peers in a reference group are more likely to have poorer health (Marmot *et al.*, 1991; Wilkinson, 1997). Low relative income may cause stress and depression leading to illness (Cohen *et al.*, 1997) or weaken one's power in the allocation of local health-related resources (Deaton, 2003). Some (Wilkinson, 1996) go even further and argue that income inequality may affect the health of both the poor and the well-off in a society (referred to as the strong income inequality hypothesis), possibly through disinvestment in public health and human capital, the erosion of social capital, or stressful social comparisons (Kawachi and Kennedy, 1999).

The relative income or income inequality hypotheses have been empirically tested, but almost exclusively drawing on data from industrialized countries, and the results have been mixed.[1] The tests have been conducted at both the aggregate and individual levels. At the aggregate level, a number of studies have shown a robust association between income inequality and public health (e.g., Waldmann, 1992; Kaplan *et al.*, 1996; Kawachi *et al.*, 1997; Lynch *et al.*, 1998). However, the use of aggregate data may be unconvincing. As noted by Gravelle (1998), income inequality may be spuriously correlated with the aggregate measure of health if individual health is a concave function of income. It is therefore difficult to discriminate between the effects of income and income inequality using aggregate data. To differentiate between the absolute income and income inequality effects, recent studies employ individual data. Among these studies, some support the income inequality hypothesis (e.g., Kennedy *et al.*, 1998a; Soobader and LeClere, 1999; Blakely *et al.*, 2001), while others find no significant effects of inequality (e.g., Meara, 1999; Blakely *et al.*, 2002; Mellor and Milyo, 2002).

1 For a systematic review of previous empirical work, see Deaton (2003) and Lynch *et al.* (2004).

The goal of this paper is to test the above hypotheses and investigate the relationship between income, income inequality, and health in China, using the individual data from the China Health and Nutrition Survey (CHNS). We find evidence supporting the absolute income hypothesis that income has a positive effect on self-reported health status. Consistent with findings by Daly *et al.* (1998), we also find evidence supporting the strong version of the income inequality hypothesis but not the weak version. However, unlike previous findings of a linear relationship, our results show an inverted-U association between self-reported health status and inequality, i.e., the detrimental effect of income inequality on health only appears in communities with high inequality. We also test the effect of relative deprivation and income rank on health but find little effect of relative income on health. This is in contrast with the work of Eibner and Evans (2005), who find relative deprivation important in explaining individual health with the exception of rank. Finally, we also show that rising inequality can significantly increase one's probability of engaging in health-compromising behavior such as smoking and alcohol abuse.

We contribute to the literature studying the relationship between income inequality and health in the following ways. First, this paper is one of the first studies to use individual data from a developing country. Although poor health and high inequality are key features of many developing countries, the earlier literature has studied their association drawing mainly on data from the United States and other industrialized countries.[2] Moreover, as pointed out by Gerdtham and Johannesson (2004), industrial countries like Sweden may not be the best places for studying the effects of income inequality, because these countries are typically more egalitarian and do not have sufficient variation in income inequality across regions. In contrast, China has both rising inequality and a large variation in inequality across localities (Gustafsson and Li, 2002). Second, we extend the previous work by explicitly distinguishing between the relative income hypothesis and the income inequality hypothesis in the same study. Previous studies have tested either the relative income hypothesis (Deaton, 2001; Eibner and Evans, 2005) or the income inequality hypothesis (e.g., Mellor and Milyo, 2002).[3] Finally, we measure the income inequality at the community level, so that our focus is more locally defined than in most previous studies, which focus on the state or county level. Using community-level inequality not only facilitates the empirical test by allowing us to work with a larger variation in inequality, but also permits us to examine the potential impacts of inequality within a society by taking a set of people who are more closely related.

The structure of the paper is as follows. Section 2 presents the hypotheses and literature review. Section 3 describes the data and some measurement issues. Section 4 reports our estimation results. The paper concludes with Section 5.

---

2 For example, Osler *et al*. (2002), Shibuya *et al*. (2002), and Gerdtham and Johannesson (2004) employ data from Denmark, Japan, and Sweden, respectively.

3 Gerdtham and Johannesson (2004) test both hypotheses, but their measure of relative income is a simple one.

2 Hypotheses and previous research

In our study, we attempt to examine whether health outcomes and behavior are correlated with income and income inequality in China. We begin with a discussion of several hypotheses that link income and income distribution to health, followed by a selected review of previous empirical work. We then specify the empirical test for each hypothesis.

Hypothesis 1. Absolute income hypothesis

The absolute income hypothesis argues that people with higher incomes have better health outcomes, but income inequality or relative income has no direct effect on health. A related concept is the poverty hypothesis, which emphasizes that ill health is a consequence of low income or extreme poverty. The idea that health improves with income goes back a long way in the literature. One of the most influential works in this area is by Preston (1975), who finds that the impact of additional income on mortality is greater among the poor than the rich. In other words, there is a concave relationship between income and health.

A large number of empirical studies in a variety of disciplines (such as economics, sociology, and epidemiology) demonstrate a robust association between income and health (no matter how income and health are measured) using individual data, and most of the evidence points to a nonlinear relationship.[4] We follow the literature and test whether per capita income has a positive effect on individual health.[5] However, since the protective effect of absolute income on health is relatively uncontested (compared with the effect of income inequality or relative income), we do not place too much emphasis on this test.

Hypothesis 2. Income inequality hypothesis

The income inequality hypothesis presumes that income inequality per se is a threat to the health of individuals within a society, even holding their incomes constant. It focuses on the direct tie between health and income inequality, regardless of a person's particular income level. There are several potential pathways through which income inequality might harm an individual's health directly. For example, high levels of inequality might produce instabilities in the social capital by increasing mistrust and stress, or declining social cohesion, which in turn adversely influence an individual's own health through psychosocial responses like violent crime or self-destructive behavior.[6]

This hypothesis has two versions (Mellor and Milyo, 2002). The strong version states that inequality affects all members in a society equivalently, irrespective of their income levels. The weak version suggests that income inequality may harm the health of only

---

4 See the review by Feinstein (1993), and a more recent discussion by Smith (1999).

5 We also control for income squared to capture the nonlinear relationship between income and health.

6 Kawachi and Kennedy (1999) summarize three plausible mechanisms linking income inequality to health: disinvestment in human capital, the erosion of social capital, and stressful social comparisons.

the least well-off in a society, or that the harmful effect of inequality on health decreases with one's income rank.

Early studies use aggregate data to test the correlation between income inequality and health. Various works by Wilkinson over the past decade (e.g., Wilkinson, 1992, 1996) present evidence of a relationship between income inequality and life expectancy across a number of industrialized countries, both at a point in time and over time. While Wilkinson reports correlation coefficients, a growing body of literature tests this hypothesis using regression frameworks. A link between income inequality and health measures (mortality, morbidity, etc.) has been discerned repeatedly at the level of countries (Waldmann, 1992; Wennemo, 1993), and across states, counties, and cities within nations (Kaplan *et al.*, 1996; Ben-Shlomo *et al.*, 1996; Kennedy *et al.*, 1996; Kawachi *et al.*, 1997; Kawachi and Kennedy, 1997; Lynch *et al*., 1998). In addition, some studies find an association between income distribution across U.S. states and state-level measures of smoking (Kaplan *et al*., 1996), alcohol consumption (Marmot, 1997), and firearm crimes (Kennedy *et al*., 1998b).

Although these studies are informative, they use aggregate data, making it hard to differentiate between the hypotheses for absolute income and income inequality. The aggregate association between income inequality and health may merely reflect the nonlinear relationship between income and health at the individual level. For example, if a transfer of one dollar from the rich to the poor improves the health of the poor more than it diminishes the health of the rich, this income-equalizing transfer will increase the average health of the whole society.[7] If all that matters to individual health is income, then for two communities with identical average income, the community with a more equal income distribution tends to have better average health than the one with greater inequality. Thus, in aggregate studies, it is hard to distinguish this statistical artifact (Gravelle, 1998) from mechanisms in which income inequality has a direct effect on individual health. In order to identify the true effect of inequality, one should employ individual data.

A number of studies using U.S. data find that income inequality does indeed have a negative effect on individual health. For instance, Kennedy *et al*. (1998a, 1998b), Soobader and LeClere (1999), Fiscella and Franks (2000), and Blakely *et al*. (2001) all show a significant association between inequality (at state or county level) and self-rated health status. Daly *et al*. (1998) examine the effects of several measures of state-level income inequality on individual mortality, and find supporting evidence for the income inequality hypothesis in a particular time period. Using county and tract-level inequality data, LeClere and Soobade (2000) find supporting evidence as well, but only for some specific subgroups in high-inequality counties.

In contrast, some studies indicate no association between income inequality and individual health. Measuring inequality by the proportion of income earned by the poorest 50 percent of the population, Fiscella and Franks (1997) find no effects of county-level inequality on mortality. Meara (1999) examines the relationship between state-level inequality and birth outcomes (such as infant mortality and low birth weight), and finds no significant relation. Mellor and Milyo (2002) construct several inequality

---

7 Using a new data set, Deaton (2003) shows a recent version of the Preston curve and suggests that income redistribution from rich to poor countries will in principle increase average health worldwide.

measures both at the level of states and metropolitan areas, and show that their effects on self-rated health status are eliminated once individual income and locality effects are controlled. Using the same data as Mellor and Milyo (2002), Blakely *et al*. (2002) draw a similar conclusion, finding that, after controlling for income, there is little association between income inequality and individual health. A few studies using data outside the United States provide further evidence against the income inequality hypothesis (Osler *et al*., 2002; Shibuya *et al*., 2002; Gerdtham and Johannesson, 2004).

Most of the existing literature focuses on the strong version of the income inequality hypothesis. Only a few studies (Daly *et al*., 1998; Meara, 1999; Mellor and Milyo, 2002; Gerdtham and Johannesson, 2004) implicitly or explicitly test the weak version, but none of their findings support the hypothesis.

In this paper, we test both the strong and weak versions of the inequality hypothesis. The strong version of the income inequality hypothesis is specified as follows,

$$H_{ij} = \beta_0 + \beta_1 Q_j + \beta_2 Q_j^2 + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}, \qquad (1)$$

where i and j are subscripts for individual and community, respectively. H_{ij} denotes a number of health outcomes and behavior (self-reported health status, objective body conditions, smoking, alcohol use, etc.). Q_j stands for the community-level income inequality. I_{ij} is the vector of per capita income and income squared, and X_{ij} is the vector of other individual, household, and community variables. We also include the squared term of inequality to capture the potential nonlinear effect. We hypothesize that health outcomes deteriorate with income inequality ($\beta_1 < 0$), but the relation might not be linear ($\beta_2 \neq 0$).

To test the weak version, we extend equation (1) by introducing the interaction between inequality and a person's rank (in the ascending order of income), denoted by R_{ij}, to allow the effects of income inequality to vary by the relative income level. The model is

$$H_{ij} = \beta_0 + \beta_1 Q_j + \beta_2 Q_j^2 + \delta R_{ij} + \eta Q_j \cdot R_{ij} + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}. \quad (2)$$

We expect a positive coefficient of the interaction term ($\eta > 0$), or that the negative effect of inequality on health outcomes is smaller for people with higher income rankings.

Hypothesis 3. Relative income hypothesis

The relative income hypothesis states that health depends on an individual's income relative to others in his or her group, rather than an individual's absolute income. According to this hypothesis, health declines when one is financially deprived relative to one's peers, and improves when one is prosperous relative to others. A similar hypothesis is the relative position hypothesis, which stresses that one's relative rank in a group is related to health outcomes.[8]

---

[8] The rank extends the concept of relative income as it can be measured by socioeconomic factors other than income, such as occupation and education.

Some psychosocial and material factors may play a role in the mechanisms connecting relative income to health. Perceptions of being relatively deprived compared to their peers may make people stressed and depressed, thus diminishing their health directly through diseases or indirectly via health-compromising behavior.[9] Another possibility is that within a community, relative income (or rank) may be more important in determining an individual's access to material goods or services that are correlated with health.[10]

The relative income hypothesis is consistent with an effect of income inequality, but the two are not totally equivalent. If inequality increases, the poor are made even poorer in relative terms, and the rich become relatively more prosperous. Thus the harmful effect of income inequality is greater among the least well-off. In this sense, the relative income theory parallels the weak version of the income inequality hypothesis. However, the strong version of the income inequality hypothesis goes further than the relative income hypothesis. According to the strong version, even rich people, who are least deprived in terms of relative income, may still suffer the adverse impacts of high income inequality. Thus, the strong version suggests that income inequality might directly influence health through channels independent of relative income.

Studies using different measures of relative income generate mixed results. Some recent research uses the mean (or median) income of a community as a proxy for relative income, but finds no evidence supporting the hypothesis (e.g., Robert, 1998; Gerdtham and Johannesson, 2004). However, the Whitehall study in Britain (Marmot *et al*., 1984; Marmot *et al*., 1991), one of the most widely known studies on relative income (position), finds higher rates of morbidity and mortality among civil servants in the lower administrative ranks. The contributions by Deaton (2001) and Eibner and Evans (2005) are more interesting, since they measure the level of relative income more specifically by the differences between an individual's income and the incomes of the richer members of the group. Using these measures, which are called relative deprivation (RD),[11] they both find a significant relative income effect on individual mortality from U.S. data. Moreover, Eibner and Evans (2005) show that relative deprivation also influences the probability that an individual will engage in health-compromising behavior, such as smoking and not wearing a seatbelt while driving.

Following Eibner and Evans (2005), we test the relative income hypothesis using the following specification:

$$H_{ij} = \beta_0 + \beta_1 RD_{ij} + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}. \qquad (3)$$

9 Some research on monkeys and primates (e.g., Cohen *et al*., 1997; Shively *et al*., 1997) provides biological evidence of how relative status may affect health.

10 Deaton (2003) takes the case of local housing in a town: the richest people are able to get the hilltop plots with fine views while the poorest are left with the plots downward of the smokestacks. This is an example "where it is not money itself that is important, but rank, here determined by money."

11 The definition of relative deprivation is originally proposed by Runciman (1966), who argues that one is deprived if others in the group possess something that one does not have. Yitzhaki (1979) develops the definition by viewing income as personal possessions, and shows the link between relative deprivation and income inequality.

Equation (3) is similar to equation (1), except that we replace Q_j with RD_{ij}, which stands for relative deprivation indices that measure an individual's relative income (see Section 4.2 for details). The difference in subscripts between Q_j and RD_{ij} means that income inequality is an aggregate measure for the whole community, while the relative income measures that we use are individual specific. We hypothesize that higher relative deprivation of income (or lower relative rank) reduces the probability of being healthy, and increases the probability of participating in health-compromising behavior.

3 Data

In this paper, we use the China Health and Nutrition Survey (CHNS) data, which were collected by the Carolina Population Center (CPC) at the University of North Carolina at Chapel Hill, the Institute of Nutrition and Food Hygiene, and the Chinese Academy of Preventive Medicine.[12] The CHNS was a longitudinal survey with five waves in 1989, 1991, 1993, 1997, and 2000. The sample households were randomly drawn from eight provinces including Liaoning, Shandong, Jiangsu, Henan, Hubei, Hunan, Guangxi, and Guizhou.[13] Two cities and four counties were sampled in each province. Four neighborhoods in each city, and one county-town neighborhood, and three villages in each county, were then randomly selected. We define a neighborhood or village as a community unit.[14] Approximately 20 households were sampled per community.

The CHNS data contain detailed information on household and individual characteristics as well as health-related information such as self-reported health status, physical functions, activities of daily living, and health behavior. We use the wave of 1993 for our basic cross section analysis because the 1993 CHNS has the richest set of health variables. We restrict our sample to men and women who were at least 20 years old in 1993 and had a complete set of data on health and demographic variables (age, sex, marital status, education, etc.). As we need to construct income inequality and relative deprivation indices, we also exclude those with nonpositive household income. In total, we have 7286 observations in the 1993 sample.

We also conduct some panel analysis using four waves of 1991 to 2000, though the panel analysis is limited by the data. Although the CHNS data are longitudinal, some health measures are not consistently reported across all the rounds. For example, the 1989 survey did not report many health outcomes such as the self-reported health status. The 1991 survey did not have questions on activities of daily living, while the 1997 and 2000 waves changed physical functions to rarer diseases. The only health variables consistently available through the later four waves are self-reported health status, blood pressure, and health behavior.

---

[12] A detailed description of the data and quality control procedures can be obtained from http://www.cpc.unc.edu/projects/china/.

[13] Liaoning was replaced by Heilongjiang in the round of 1997 and returned to the survey in 2000.

[14] As a community is defined to be a subunit of an urban city or a rural county, constructing the income inequality and relative income at the community level avoids the case that rural and urban households are pooled in the same reference group. Thus we highlight the within-community inequality rather than the substantial urban–rural gap in our analysis.

Table 1. Definitions of key variables.

Variables	Definition
Self-reported health status (SRHS)	1 if health is excellent or good, 0 if fair or poor
Physical functions (PF)	
Heart	1 if normal in condition of heart, lungs, and stomach, 0 if otherwise
Blood	1 if normal blood pressure, 0 if high blood pressure
Upper	1 if normal in upper extremities, 0 if otherwise
Lower	1 if normal in lower extremities, 0 if otherwise
Urine	1 if normal in urine and bowel control, 0 if otherwise
Activities of daily living (ADL) (for 50+ years old)	
Walk	1 if able to walk for a kilometer, 0 if with limitation
Lift	1 if able to lift a 5 kg bag, 0 if with limitation
Climb	1 if able to climb a staircase, 0 if with limitation
Shower	1 if able to take the shower alone, 0 if needs help
Eat	1 if able to eat alone, 0 if needs help
Health behaviors	
Current smoker	1 if smokes at the survey time, 0 if not
Cigarettes per day	Average number of cigarettes smoked per day
Current drinker	1 if drinks alcoholic beverage in the previous year of the survey, 0 if not
Drinking frequency	0 if does not drink, 1 if no more than once a month, 2 if once or twice a month, 3 if once or twice a week, 4 if 3–4 times a week, 5 if daily or almost everyday
Inequality and relative deprivation	
Gini	Gini coefficient of income within the community
Rank	Centile rank (in the ascending order of income) within the community
RDA	Yitzhaki's relative deprivation index: $RDA_i = \sum(y_j - y_i)/N$, for all $y_j > y_i$, where y_i is the income of person i and N is the size of the community
RDL	Substituting log(y) for y in RDA
RDI	RDA/y, i.e., dividing RDA by one's own income
Other variables	
Income	Deflated per capita household income
Education	Years of formal schooling
Family size	Number of household members
Tap water	1 if pipe or tap water inside house or courtyard, 0 if otherwise
Distance	Average distance (km) from the community to frequently used facilities
Rural	1 if the community is a village unit, 0 if an urban unit

Table 1 summarizes the definitions of key variables in our sample. We now discuss a variety of measurement issues that need to be clarified before we present the estimation results.

3.1 Health indicators

The CHNS data offer several potential health measures, as shown in the top panel of Table 1. Self-reported health status (SRHS) is the main health measure we use. Although SRHS is a subjective measure of individual health,[15] previous studies show that SRHS is highly correlated with subsequent mortality, even when controlling for more objective health measures (Idler and Benyamini, 1997; Deaton and Paxson, 1998). We construct a binary variable, SRHS, which equals 1 if excellent or good health is reported and equals 0 if fair or poor health is reported.[16]

We also use several objective health measures such as physical functions (PF) and activities of daily living (ADL), which are recorded in the physical examination section of the survey. PF provides information on the status of various body functions associated with heart, hearing, eyesight, arms, legs, etc. We construct five indicators as PF measures: heart, lungs, and stomach condition (henceforth heart), blood pressure (blood), upper extremities condition (upper), lower extremities condition (lower), and urine and bowel control (urine). As with SRHS, we define them as binary variables that equal 1 if the function is normal and 0 otherwise. ADL measures whether or not the individual is physically restricted or unable to perform daily activities, such as walking for a certain distance (walk), lifting a certain weight (lift), climbing a staircase (climb), taking a bath alone (shower), and eating and drinking alone (eat). Again, we create binary variables that are equal to 1 if respondents were able to perform the activities, and equal to 0 if respondents reported any difficulty in these activities. However, ADL measures are unavailable for individuals under 50; thus we can only use this measure for a smaller sample of the elderly.

In addition to these direct measures, the CHNS data also contain information on some health-compromising behavior such as smoking and alcohol consumption. Regarding smoking behavior, we have knowledge of whether or not an individual smoked at the time of the survey, and the number of cigarettes smoked per day. Regarding drinking behavior, we know whether or not an individual had drunk any alcoholic beverage in the year prior to the survey, and the frequency of drinking. In total, we have four variables to measure health behavior, i.e., current smoker, cigarettes per day, current drinker, and drinking frequency, as illustrated in Table 1. As most of the smokers and drinkers were men in our sample, we limit the analysis of health behavior to men who had nonmissing behavior variables.

---

15 In the survey, the interviewees were asked the question: "right now, how would you describe your health compared to that of other people of your age?"

16 In the survey, SRHS is a categorical variable coded on a scale of 1 (excellent) to 4 (poor).

Table 2. Descriptive statistics of health, inequality, and other variables in China.

	Mean and (standard deviation)			
		Healthy versus unhealthy		
	Full sample	SRHS = 1	SRHS = 0	*t*-statistics
Variables	(1)	(2)	(3)	(4)
SRHS all	0.730 (0.444)	-	-	-
SRHS men	0.757 (0.429)	-	-	-
SRHS women	0.703 (0.457)	-	-	-
SRHS 50+ years old	0.556 (0.497)	-	-	-
Heart	0.928 (0.259)	-	-	-
Blood	0.948 (0.221)	-	-	-
Upper	0.936 (0.245)	-	-	-
Lower	0.936 (0.245)	-	-	-
Urine	0.995 (0.070)	-	-	-
Walk [a]	0.758 (0.428)	-	-	-
Lift [a]	0.726 (0.446)	-	-	-
Climb [a]	0.661 (0.473)	-	-	-
Shower [a]	0.938 (0.242)	-	-	-
Eat [a]	0.989 (0.103)	-	-	-
Current smoker [b]	0.688 (0.463)	-	-	-
Cigarettes per day [b]	10.82 (10.07)	-	-	-
Current drinker [b]	0.629 (0.483)	-	-	-
Drinking frequency [b]	2.275 (2.008)	-	-	-
Gini	0.323 (0.099)	0.323 (0.098)	0.322 (0.100)	0.61
Rank	0.498 (0.303)	0.508 (0.304)	0.471 (0.300)	4.65***
RDA (/1000)	0.429 (0.408)	0.423 (0.414)	0.443 (0.392)	1.84*
RDL	0.377 (0.512)	0.367 (0.514)	0.405 (0.505)	2.86***
RDI	1.224 (5.066)	1.177 (4.666)	1.353 (6.015)	1.32
Income (1000 yuan)	1.373 (1.246)	1.411 (1.271)	1.273 (1.171)	4.20***
Age	43.47 (14.85)	40.89 (13.69)	50.44 (15.62)	25.43***
Male	0.498 (0.500)	0.516 (0.500)	0.448 (0.497)	5.20***
Married	0.834 (0.372)	0.845 (0.362)	0.803 (0.398)	4.30***
Education	6.052 (4.381)	6.536 (4.219)	4.742 (4.541)	15.78***
Family size	4.414 (1.590)	4.463 (1.541)	4.283 (1.706)	4.29***
Tap water	0.629 (0.483)	0.631 (0.482)	0.621 (0.485)	0.81
Distance (km)	1.495 (2.767)	1.438 (2.592)	1.650 (3.187)	2.90***
Rural	0.676 (0.468)	0.686 (0.464)	0.651 (0.477)	2.81***
Households per community	18.37 (2.04)	-	-	-
Individuals per community	75.78 (16.09)	-	-	-
Sample size	7286	5320	1966	-

Source: 1993 CHNS, adults 20+ years old.

Notes: standard deviations are shown in parentheses; *, **, and *** represent significance levels of 10, 5, and 1 percent. [a] Reported only by individuals aged 50+. [b] Women are excluded.

Table 2 (column 1) provides descriptive statistics concerning these health measures. SRHS and PF measures are available for the whole sample, but ADL and health behavior variables are only available for smaller samples. Among all individuals, 73 percent reported being in good health. Examining the data in two sex groups, we find that men are more healthy than women, with 76 percent of men but only 70 percent of women reporting themselves in good health. The proportion declines with age, as only 56 percent of those over 50 report themselves to be in good health. By contrast, higher normal rates are reported for the PF measures, all exceeding 90 percent for the whole sample. The proportion of people with no limitations in ADL is closer to that for SRHS, although it should be remembered that the sample is much smaller. Finally, 69 percent of men were smoking at the time of the survey, and 63 percent reported that they drank during the year prior to the survey.

3.2 Income inequality and relative income measures

In this paper, we use the Gini coefficient to measure the community-level income inequality.[17] For every community, we calculate the Gini based on household income weighted by the family size.[18] In total there are about 180 communities in our sample. The Gini ranges from 0.1 to 0.6, with the average value around 0.32 (Table 2).

Following Eibner and Evans (2005), we construct several relative deprivation indices as the proxy for relative income, i.e., relative deprivation of absolute income (RDA), relative deprivation of log income (RDL), relative deprivation over individual income (RDI), and individual rank. Based on the theory developed by Yitzhaki (1979), RDA is defined as

$$\mathrm{RDA}_i = \frac{1}{N}\sum_j (y_j - y_i) \qquad \forall \quad y_\mathrm{j} > y_i. \tag{4}$$

It measures the relative deprivation of person i with income y_i in a reference group of N people by the normalized total incomes of other group members who earn more than i does. RDL is the same as RDA except that it uses log(y) rather than y in (4). RDI equals RDA_i/y_i, namely the ratio of RDA relative to person i's own income. The final index we use is the individual's centile rank within the reference group (where income is sorted in the ascending order). In contrast to the first three measures, the rank ignores the magnitude of the income difference between individuals. While larger values in RDA, RDL, and RDI indicate higher levels of relative deprivation, a higher centile rank means a lower level.

---

17 Kawachi and Kennedy (1997) show that the six inequality measures (including the Gini coefficient and the Theil index) used in their study are highly correlated with each other, and the choice of inequality indicators does not change the relationship between income inequality and mortality. We also use another inequality index, the Theil index, to test the robustness of our results, and the different measure of inequality does not change our results qualitatively.

18 The household income used here has taken account of production costs. We also calculate the gross income by not subtracting production costs, and generate similar estimates (not shown but available upon request).

As the Gini coefficient depicts the overall income distribution of a society, relative deprivation reflects a person's position or rank relative to the incomes of others within a reference group. In order to be consistent with the Gini coefficient, we use households in the same community as the reference group to generate these RD measures.[19] The summary statistics of our relative deprivation measures are reported in Table 2. Unlike the Gini, which is bounded between 0 and 1, relative deprivation measures (RDA, RDL, and RDI) are not limited in value and therefore have larger variations in the sample.

3.3 Other explanatory variables

In the individual-level analysis, we control for variables including per capita income and income squared, age and age squared, education, indicators for sex and marital status, family size, source of drinking water (tap water or not), the distance from the community to nearby medical facilities,[20] and rural and provincial indicators. We show the descriptive statistics for these variables in the bottom panel of Table 2. Individuals in our sample have an average income of 1373 yuan,[21] an average age of 43, and averagely 6 years of schooling.

In Table 2 we also divide the sample into two subsamples: good health and poor health (columns 2 and 3). The differences in personal characteristics between the two subsamples are what we would intuitively expect. Specifically, we find that on average healthy people have higher per capita income and education level, and are much younger than unhealthy ones. Those in good health also live in larger families and closer to medical facilities. The role of income inequality is less explicit, as the average Gini coefficients for the two groups are very close. On the other hand, the poor health group on average is slightly more deprived, as indicated by its smaller mean of individual ranks and larger mean of the other three indices. The *t*-ratios in column 4 show that most of the means are significantly different between the two subsamples, except for some inequality and relative deprivation variables.

4 Estimation results

In this section, we use the 1993 CHNS sample to systematically test various hypotheses discussed in Section 2. The main purpose of our study is to examine the correlation between individual health and income inequality or relative income. We also perform panel data analysis using four rounds of the CHNS data from 1991 to 2000.

---

19 On average, each community has 18 households and 75 individuals in the sample (Table 2).

20 The distance to medical facilities is obtained from the CHNS community survey and measures the availability of public health services to the community. We use the average distance if more than one facility is frequently used.

21 We use the consumer price index provided in the CHNS data to adjust per capita income to prices in urban areas in the Liaoning province.

4.1 Income, income inequality, and health

We first employ probit model to test the income inequality hypothesis (Hypothesis 2), in both the strong and weak versions. We apply models (1) and (2) to various health measures such as SRHS, PF, ADL, and health behavior, using individual-level data. Our specifications also allow for a test of the absolute income hypothesis (Hypothesis 1), even though it is not our focus.

Self-reported health status

Table 3 presents the results of probit regressions using SRHS as the dependent variable. The results exhibit an inverted-U, i.e., a quadratic relationship between SRHS and income inequality. We report d*F*/d*x*, or the marginal change of probability of reporting excellent or good health when the independent variable increases.[22] In the first column, we have the Gini as the only independent variable. The coefficient of the Gini is positive but not significant. When we add the squared term in the second column, the correlation is still insignificant. However, in column 3, the coefficients of the Gini and Gini squared both become significant at the 5 percent level, after we include other control variables such as per capita income, and personal and household characteristics. The positive coefficient of the Gini and negative coefficient of Gini squared mean that SRHS increases with inequality when Gini is less than 0.42 (75 percentile in the sample) and decreases with inequality for larger Gini. The results suggest that the strong version of the income inequality hypothesis (Hypothesis 2) is only supported for communities with large inequality.[23]

We also find evidence supporting the absolute income hypothesis (Hypothesis 1). Column 3 shows that there is a concave relationship between individual health and per capita income. The positive coefficient of income and negative coefficient of income squared are both significant at the 1 percent level. The critical point of the health–income quadratic curve is about 7667 yuan, but 99 percent of the values for income in our sample are below this figure. This means that for most of our sample health increases with absolute income, but at a decreasing rate.

Other control variables also have the expected signs in column 3. The probability of being in good health decreases with age at a rate of 1 percentage point per year (not shown). One more year of schooling increases the probability of being in good health by 0.4 percentage point. Men have a 3.7 percentage points higher probability of being in good health than women, and married people have a 4.3 percentage points higher probability than single people. A one standard deviation increase in family size (1.6) raises the probability by 1.8 percentage points. Having access to tap water increases the probability of reporting good health by 3 percentage points. Finally, the distance to medical facilities has a negative sign but is statistically insignificant.

---

22 The statistics that are reported here, as in all of the regressions in this paper, allow for the correlation of errors within the household.

23 This is consistent with the findings of LeClere and Soobade (2000) who use U.S. data.

Table 3. Probit regressions measuring the effects of income inequality on self-reported health status.

	Dependent variable: self-reported health status (1 = excellent or good, 0 = fail or poor)			
	(1)	(2)	(3)	(4)
Gini	0.032	0.313	1.029**	1.147**
	(0.46)	(0.80)	(2.33)	(2.51)
Gini squared		–0.416	–1.236**	–1.176*
		(–0.73)	(–1.96)	(–1.87)
Rank				0.140*
				(1.78)
Gini * rank				–0.352
				(–1.54)
Income			0.046***	0.038**
			(3.81)	(2.34)
Income squared			–0.003***	–0.003*
			(–2.59)	(–1.83)
Education			0.004**	0.004**
			(2.00)	(1.98)
Male			0.037***	0.037***
			(4.26)	(4.25)
Married			0.043**	0.043**
			(2.28)	(2.32)
Family size			0.011**	0.011**
			(2.12)	(2.20)
Tap water			0.030*	0.032*
			(1.72)	(1.80)
Distance			–0.002	–0.002
			(–0.71)	(–0.76)
Rural			0.013	0.012
			(0.70)	(0.64)
Provincial indicators	No	No	Yes	Yes
Number of observations	7286	7286	7286	7286
Pseudo *R*-squared	0.00	0.00	0.09	0.09

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust t-statistics, which allow for correlation of errors within household, are shown in parentheses; regressions 3 and 4 include age and age squared.

Next in column 4, we test the weak version of the income inequality hypothesis, i.e., whether the effects of inequality differ by relative income. As in the previous regression, the Gini has a quadratic form of effect on health. Although the individual rank has a positive effect (significant at the 10 percent level), the interaction between the Gini and the rank is negative and not statistically different from zero. The result implies that the effect of income inequality does not vary with relative income, and the negative sign on the interaction seems to contradict what is predicted by the weak version of the income inequality hypothesis that income inequality harms the health of the poor more than the rich.

In short, the results in Table 3 show that the community-level income inequality influences the individual health status in a nonlinear way. According to the estimated coefficients, income inequality tends to have a detrimental impact on health when a community has large inequality (the Gini above 0.40, in column 3). The higher individual rank is beneficial to one's health, but income inequality has the same effect for each community member, regardless of his or her rank.

Physical functions

Table 4 reports estimations using two PF variables as dependent variables: the condition of heart, lungs, and stomach, and the condition of blood pressure.[24] We find a nonlinear relationship between the Gini and heart function (columns 1 to 3), but no correlation between the Gini and blood pressure (columns 4 to 6). The effects are not altered by one's relative income position, as the coefficients of the interaction term are both insignificant (columns 3 and 6). Compared to previous estimates for SRHS, fewer control variables remain significant. These results are likely due to the lack of variation for the PF measures. For example, the proportion of people reporting normal heart condition amounts to 93 percent, and the proportion reporting normal blood pressure is 95 percent.

Activities of daily living

As another check, we estimate the influence of income inequality on ADL measures in a restricted subsample of elderly people in Table 5. The two dependent variables we use here are indicators of whether one is able to walk for 1 km and lift a 5 kg bag without difficulty. We follow the estimation specifications that were previously applied to PF indicators.

The regression results in Table 5 further confirm our finding that income inequality has an impact on individual health. The community Gini has a negative effect on both walking and lifting abilities (columns 1 and 4). Moreover, inequality has a nonlinear effect on the lifting ability (columns 5 and 6). The estimates imply that the probability of being able to lift the bag decreases with income inequality when the Gini is greater than 0.29 (about 38 percentile in the subsample). The impacts of income inequality on ADL limitations are independent of the individual rank, since the interaction of Gini and rank is not significant in columns 3 and 6.

---

[24] To save space, we only report two PF measures here and two ADL measures in Table 5. The estimates for other PF and ADL variables are compiled into Table 6.

Table 4. Probit regressions measuring the effects of income inequality on physical functions.

	Dependent variable: heart (1 = normal in heart, lung, and stomach, 0 = otherwise)			Dependent variable: blood (1 = normal blood pressure, 0 = high blood pressure)		
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	0.051	0.423**	0.441**	0.005	-0.069	-0.050
	(1.42)	(2.51)	(2.53)	(0.33)	(-0.81)	(-0.57)
Gini squared		–0.543**	–0.548**		0.112	0.106
		(–2.20)	(–2.23)		(0.90)	(0.86)
Rank			0.003			-0.001
			(0.08)			(-0.09)
Gini * rank			–0.023			-0.020
			(–0.23)			(-0.43)
Income	0.004	0.005	0.006	–0.004*	-0.004*	-0.002
	(0.84)	(0.91)	(0.93)	(–1.83)	(-1.83)	(-0.61)
Income squared	–0.000	–0.000	–0.000	0.000	0.000	0.000
	(–0.52)	(–0.53)	(–0.65)	(0.64)	(0.59)	(0.00)
Education	0.000	0.000	0.000	–0.000	–0.001	–0.001
	(0.07)	(0.23)	(0.21)	(–1.21)	(–1.26)	(–1.34)
Male	0.013**	0.012**	0.012**	0.004	0.004	0.004
	(2.05)	(1.98)	(1.98)	(1.37)	(1.37)	(1.45)
Married	0.006	0.005	0.005	–0.007*	–0.007*	–0.007*
	(0.66)	(0.59)	(0.59)	(–1.68)	(–1.66)	(–1.67)
Family size	0.000	0.000	0.000	0.000	0.000	0.000
	(0.17)	(0.10)	(0.08)	(0.11)	(0.16)	(0.12)
Tap water	0.016**	0.017**	0.016**	–0.007*	–0.007*	–0.007**
	(2.03)	(2.05)	(1.99)	(–1.86)	(–1.87)	(–2.02)
Distance	–0.002**	–0.002**	–0.002**	0.000	0.000	0.000
	(–2.24)	(–2.34)	(–2.31)	(0.13)	(0.20)	(0.25)
Rural	0.017**	0.014	0.014*	0.011***	0.011***	0.011***
	(2.00)	(1.64)	(1.66)	(3.00)	(3.08)	(3.15)
Number of observations	6349	6349	6349	6033	6033	6033
Pseudo *R*-squared	0.08	0.08	0.08	0.21	0.21	0.21

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust *t*-statistics, which allow for correlation of errors within household, are shown in parentheses; all regressions include age and age squared, and provincial indicators.

Table 5. Probit regressions measuring the effects of income inequality on activities of daily living.

	Dependent variable: walk (1 = able to walk for 1 km, 0 = having limitation)			Dependent variable: lift (1 = able to lift a 5 kg bag, 0 = having limitation)		
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	–0.484***	0.145	0.208	–0.281**	1.267*	1.316*
	(–4.43)	(0.24)	(0.34)	(–2.28)	(1.88)	(1.87)
Gini squared		–0.897	–0.850		–2.218**	–2.191**
		(–1.08)	(–1.03)		(–2.38)	(–2.34)
Rank			0.122			0.083
			(1.13)			(0.66)
Gini * rank			–0.237			–0.166
			(–0.76)			(–0.46)
Income	0.005	0.006	–0.009	0.028	0.031	0.021
	(0.32)	(0.35)	(–0.37)	(1.45)	(1.59)	(0.79)
Income squared	0.001	0.001	0.002	–0.000	–0.000	0.000
	(0.45)	(0.45)	(0.92)	(–0.20)	(–0.19)	(0.17)
Education	0.003	0.003	0.003	–0.003	–0.002	–0.002
	(0.98)	(1.05)	(1.05)	(–0.87)	(–0.67)	(–0.70)
Male	0.127***	0.127***	0.127***	0.160***	0.159***	0.159***
	(6.17)	(6.14)	(6.13)	(7.49)	(7.42)	(7.45)
Married	–0.009	–0.010	–0.009	0.023	0.022	0.022
	(–0.35)	(–0.39)	(–0.37)	(0.83)	(0.79)	(0.80)
Family size	0.007	0.007	0.007	0.005	0.005	0.005
	(1.20)	(1.16)	(1.22)	(0.87)	(0.78)	(0.82)
Tap water	–0.055**	–0.055**	–0.051**	0.039	0.038	0.041
	(–2.15)	(–2.16)	(–1.96)	(1.31)	(1.28)	(1.37)
Distance	0.001	0.001	0.000	–0.002	–0.003	–0.003
	(0.24)	(0.14)	(0.13)	(–0.47)	(–0.63)	(–0.64)
Rural	–0.003	–0.007	–0.009	0.037	0.027	0.026
	(–0.11)	(–0.26)	(–0.33)	(1.32)	(0.95)	(0.91)
Number of observations	2007	2007	2007	1988	1988	1988
Pseudo *R*-squared	0.17	0.17	0.17	0.19	0.19	0.19

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust *t*-statistics, which allow for correlation of errors within household, are shown in parentheses; all regressions include age and age squared, and provincial indicators.

Table 6. Probit regressions measuring the effects of income inequality on other objective health measures.

	Dependent variables: physical functions (1 = normal, 0 = otherwise)					
	Upper		Lower		Urine	
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	–0.029	–0.207	0.021	0.020	–0.004	–0.049**
	(–1.02)	(–1.27)	(0.72)	(0.12)	(–0.97)	(–2.37)
Gini squared		0.259		0.002		0.068**
		(1.11)		(0.01)		(2.22)
Income	–0.003	–0.003	–0.001	–0.001	–0.000	–0.000
	(–0.75)	(–0.75)	(–0.13)	(–0.13)	(–0.42)	(–0.51)
Income squared	–0.000	–0.000	–0.000	–0.000	0.000	0.000
	(–0.15)	(–0.18)	(–0.45)	(–0.45)	(0.55)	(0.51)
Number of observations	6447	6447	6443	6443	6444	6444
Pseudo *R*-squared	0.13	0.13	0.14	0.14	0.15	0.16
	Dependent variables: activities of daily living (1 = no limitation, 0 = otherwise)					
	Climb		Shower		Eat	
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	–0.040	0.109	0.010	0.194	0.023**	0.011
	(–0.29)	(0.14)	(0.24)	(0.91)	(2.05)	(0.23)
Gini squared		–0.214		–0.268		0.019
		(–0.20)		(–0.90)		(0.28)
Income	0.032	0.032	–0.006	–0.006	0.001	0.001
	(1.56)	(1.57)	(–0.87)	(–0.86)	(0.96)	(0.96)
Income squared	–0.001	–0.001	0.000	0.000	–0.0002*	–0.0002*
	(–0.50)	(–0.51)	(0.36)	(0.42)	(–1.77)	(–1.82)
Number of observations	2001	2001	1971	1971	1764	1764
Pseudo *R*-squared	0.13	0.13	0.23	0.23	0.22	0.22

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust t-statistics, which allow for correlation of errors within household, are shown in parentheses; all regressions include age and age squared, education, gender, marital status, family size, tap water, distance to the medical facility, rural dummy, and provincial indicators.

Health behavior

Previous results show that income inequality is strongly correlated with health outcomes. We now explore one of the potential mechanisms of their correlation by examining whether an increase in income inequality raises the probability that an individual engages in health-compromising behavior, i.e., smoking and alcohol consumption. The estimation results using different dependent variables are reported in Table 7.[25]

Table 7. Estimations of the effects of income inequality on health behavior

	Dependent variable: smoking		Dependent variable: drinking	
	Current smoker	Cigarettes per day	Current drinker	Drinking frequency
	Probit	Tobit	Probit	OLS
	(1)	(2)	(3)	(4)
Gini	0.206**	8.212***	0.177*	0.645
	(2.11)	(2.84)	(1.73)	(1.61)
Income	0.011	0.423*	0.030***	0.135***
	(1.50)	(1.91)	(3.68)	(4.41)
Education	–0.010***	–0.315***	0.005**	0.003
	(–3.66)	(–4.00)	(2.01)	(0.25)
Married	0.123***	4.514***	0.060**	0.410***
	(4.20)	(5.06)	(1.96)	(3.55)
Family size	0.001	0.051	0.001	0.016
	(0.20)	(0.30)	(0.17)	(1.61)
Tap water	0.009	0.814	0.032	0.293***
	(0.44)	(1.28)	(1.46)	(3.32)
Distance	–0.010***	–0.452***	0.001	–0.013
	(–2.97)	(–4.45)	(0.42)	(–1.04)
Rural	0.039*	1.789***	0.011	0.153
	(1.76)	(2.69)	(0.45)	(1.61)
Number of observations	3004	2899	3092	3083
(Pseudo) *R*-squared	0.04	0.01	0.04	0.04

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust *t*-statistics, which allow for correlation of errors within household, are shown in parentheses; all regressions include age and age squared, and provincial indicators.

25 An alternative test is to control for health behavior in health outcome regressions to see whether these controls attenuate the effect of income inequality. We present the estimates for SRHS (male sample only) in Table 8, which shows that adding health behavior does not change the estimates of income inequality significantly. These results suggest that there may exist some other mechanisms through which income inequality influences health, but exploration of these mechanisms is beyond the scope of this study.

Table 8. Probit regressions measuring the effects of income inequality on SRHS with control of behavior.

	Dependent variable: self-reported health status (1 = excellent or good, 0 = fail or poor)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	0.959*	1.068*	1.047*			
	(1.95)	(1.97)	(1.90)			
Gini squared	–1.186*	–1.290*	–1.271			
	(–1.67)	(–1.65)	(–1.61)			
Income	0.042***	0.037***	0.032**			
	(3.28)	(2.57)	(2.10)			
Income squared	–0.003**	–0.003*	–0.002			
	(–2.30)	(–1.76)	(–1.10)			
Smoker		0.018	0.017			
		(0.67)	(0.63)			
Cigarettes per day		0.002	0.001			
		(1.29)	(0.95)			
Drinker			–0.046			
			(–1.38)			
Drinking frequency			0.023***			
			(2.80)			
Lagged Gini				1.748***	1.600***	1.673***
				(4.30)	(3.62)	(3.70)
Lagged Gini squared				–2.537*	–2.371***	–2.459***
				(–4.11)	(–3.49)	(–3.53)
Lagged income				0.030	0.022	0.025
				(1.64)	(0.99)	(1.12)
Lagged income squared				0.001	0.001	0.000
				(0.19)	(0.27)	(0.10)
Lagged smoker					0.026	0.010
					(1.06)	(0.41)
Lagged cigarettes per day					0.001	0.001
					(1.07)	(1.13)
Lagged drinker						0.044
						(1.45)
Lagged drinking frequency						0.000
						(0.06)
Number of observations	3626	2899	2875	3626	2987	2928
Pseudo R-squared	0.09	0.09	0.10	0.10	0.10	0.10

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust *t*-statistics, which allow for correlation of errors within household, are shown in parentheses; lagged variables are obtained from 1991 CHNS. All regressions include age and age squared, education, gender, marital status, family size, tap water, distance to the medical facility, rural dummy, and provincial indicators.

Table 7 suggests a strong correlation between inequality and smoking habits (columns 1 and 2). In the first column we have the current smoker indicator as the dependent variable. The coefficient of the Gini is positive and significant at the 5 percent level. It predicts that a one standard deviation increase in community Gini (0.10) will increase the probability of smoking by 2.1 percentage points. We then use the Tobit model to estimate the effects on cigarettes consumed per day in the second column. As with the estimation on current smoker, the Gini has a strong positive effect.

However, columns 3 and 4 show that the association between inequality and drinking behavior is not as strong. The effect of income inequality on the probability of being a current drinker is positive and significant at the 10 percent level, but the effect on drinking frequency appears marginally insignificant, albeit the same sign. In particular, the coefficient of the Gini indicates that a rise in the Gini by one standard deviation (0.10) will increase the probability of drinking alcohol by 1.8 percent.

4.2 Relative income and health

We now test the relative income theory (Hypothesis 3) by replacing the independent variables of inequality with relative deprivation measures: RDA, RDL, RDI, and individual rank. The model to be estimated is equation (3). Because these measures are highly correlated with each other, their effects are estimated separately.

The estimation results with SRHS as the dependent variable (Table 9) show that the relative income hypothesis is not supported for any relative deprivation measure examined.[26] Across all the columns, none of the coefficients of RDA, RDL, RDI, or rank is statistically different from zero at the 10 percent level. We conduct the same estimations taking PF/ADL and health behavior measures as dependent variables, and again do not find any significant correlations with the relative deprivation indices (hence not reported). Our results differ from those of Eibner and Evans (2005), who find that the relative deprivation has a strong impact on health when it reflects income differences between individuals (measured in RDA, RDL, and RDI), although their results are imprecise in many cases when they measure relative deprivation using rank.

---

26 This is in line with our previous finding that the weak version of the income inequality hypothesis cannot be confirmed by the sample (Table 3). As discussed in Section 2, the relative income hypothesis is a similar argument to the weak version of the income inequality hypothesis.

Table 9. Probit regressions measuring the effects of relative deprivation on self-reported health status.

	Dependent variable: self-reported health status (1 = excellent or good, 0 = fail or poor)			
	(1)	(2)	(3)	(4)
RDA (/1000)	0.006			
	(0.30)			
RDL		–0.001		
		(–0.05)		
RDI			0.000	
			(0.78)	
Rank				0.048
				(1.49)
Income	0.040***	0.038***	0.038***	0.023
	(3.23)	(2.77)	(3.32)	(1.50)
Income squared	–0.003**	–0.003*	–0.003**	–0.002
	(–2.10)	(–1.87)	(–2.08)	(–1.08)
Education	0.003	0.003	0.003	0.003*
	(1.58)	(1.60)	(1.59)	(1.71)
Male	0.039***	0.039***	0.039***	0.038***
	(4.50)	(4.49)	(4.50)	(4.42)
Married	0.044**	0.044**	0.044**	0.044**
	(2.34)	(2.35)	(2.34)	(2.37)
Family size	0.011**	0.011**	0.011**	0.012**
	(2.17)	(2.18)	(2.18)	(2.23)
Tap water	0.025	0.026	0.025	0.031*
	(1.41)	(1.50)	(1.50)	(1.75)
Distance	–0.002	–0.002	–0.002	–0.002
	(–0.71)	(–0.74)	(–0.74)	(–0.84)
Rural	0.025	0.024	0.024	0.021
	(1.40)	(1.38)	(1.38)	(1.20)
Number of observations	7286	7286	7286	7286
Pseudo *R*-squared	0.09	0.09	0.09	0.09

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust *t*-statistics, which allow for correlation of errors within household, are shown in parentheses; all regressions include age and age squared, and provincial indicators.

4.3 Lagged inequality measures

Although the above results show a significant correlation between community-level inequality and individual health, it may not have shown a causal effect. It could be that individual health affects income and thus income inequality, in which case there is a reverse causality. There could also be some unobserved variables that have effects on both income inequality and individual health. Generally, it is very difficult to solve these problems given the limitations of data. Nonetheless, we attempt to partially address these concerns using the panel structure of the CHNS data. Specifically, we replace the income and income inequality measures with lagged values.

In Table 10, we report the same regressions as in Tables 3–5, 7, except that we substitute the lagged value of income and income inequality (from the 1991 CHNS) for the current value. Using lagged variables can help us to identify the causal effect from inequality to health because current health status should not affect past income levels or income inequality. Due to space limitation, we only report the coefficients of inequality and income variables, and suppress the coefficients of other control variables. As shown by Table 10, the lagged income inequality has a highly significant effect on SRHS (column 2).[27] Similar to the estimates in Table 3, the effect takes a quadratic form with the critical value of the Gini at 0.35. Lagged inequality also has a significant effect on the blood pressure, but not on other dependent variables.[28]

5 Conclusion

In this paper, we employ micro data from China to test several hypotheses linking income and income inequality to individual health status. We find some evidence supporting these hypotheses. First, our results show a concave relationship between self-reported health status and per capita income (the absolute income hypothesis). Additional income brings about greater improvement in the health of the poor than of the rich. Second, we find a significant association between self-reported health status and community-level income inequality (the income inequality hypothesis). In fact, the relationship we find appears as an inverted-U shape. That is to say, rising inequality tends to improve health when inequality is low, and to harm health when inequality is above a certain level. We also find evidence that income inequality increases the likelihood and frequency of health-compromising behavior such as smoking and alcohol consumption. However, our findings do not support the relative income hypothesis, or that the effect of inequality varies with income rank.

---

27 The effect remains significant after we control for lagged health behavior in Table 8 (columns 4 to 6).

28 We also tried some fixed effects estimations to control for time-invariant unobservable characteristics using four waves of the CHNS data in 1991, 1993, 1997, and 2000. Due to the changes in survey questions, the only available health measures across all rounds are SRHS, blood pressure, and health behavior. As shown in Table 11, the lagged inequality continues to have a significant effect on SRHS. However, neither current inequality nor lagged inequality has a significant effect on other dependent variables. We may not give too much weight to these fixed effects' estimates because the health measures do not have much variation over time. For example, less than 15 percent of the individuals changed SRHS across waves, even fewer for objective measures.

Table 10. Probit regressions measuring the effects of lagged income inequality on health.

	Dependent variables					
	SRHS		Heart		Blood	
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Gini	0.108	2.130***	–0.022	–0.003	0.036*	0.080
	(1.12)	(5.61)	(–0.54)	(–0.02)	(1.84)	(1.09)
Lagged Gini squared		–3.082***		–0.029		–0.071
		(–5.31)		(–0.13)		(–0.62)
Lagged income	0.020	0.022	0.011*	0.011*	–0.004	–0.004
	(1.15)	(1.24)	(1.76)	(1.76)	(–1.45)	(–1.42)
Lagged income squared	0.002	0.002	–0.001*	–0.001*	0.000	0.000
	(0.60)	(0.67)	(–1.93)	(–1.92)	(0.87)	(0.86)
Number of observations	7286	7286	6349	6349	6033	6033
Pseudo R-squared	0.09	0.09	0.08	0.08	0.21	0.21
	Walk		Lift		Smoker	Drinker
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Gini	–0.165	–0.499	0.036	–0.112	0.086	0.076
	(–1.24)	(–0.86)	(0.25)	(–0.19)	(0.76)	(0.66)
Lagged Gini squared		0.510		0.231		
		(0.58)		(0.27)		
Lagged income	0.024	0.024	0.038	0.038	–0.001	0.000
	(1.14)	(1.12)	(1.55)	(1.54)	(–0.12)	(0.03)
Lagged income squared	–0.000	–0.000	–0.000	–0.000		
	(–0.08)	(–0.08)	(–0.00)	(–0.00)		
Number of observations	2007	2007	1988	1988	3004	3092
Pseudo R-squared	0.16	0.16	0.19	0.19	0.04	0.03

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust t-statistics, which allow for correlation of errors within household, are shown in parentheses; lagged income and income inequality are obtained from 1991 CHNS. All regressions include age and age squared, education, gender, marital status, family size, tap water, distance to the medical facility, rural dummy, and provincial indicators.

Table 11. Fixed effects estimations of the effects of income inequality on health.

	Dependent variables					
	SRHS		Blood		Smoker	Drinker
	(1)	(2)	(3)	(4)	(5)	(6)
Gini	0.087	0.428	0.014	–0.017	–0.012	0.044
	(1.51)	(1.60)	(0.72)	(–0.17)	(–0.22)	(0.68)
Gini squared		–0.506		0.047		
		(–1.28)		(0.31)		
Income	0.024***	0.024***	0.001	0.001	–0.007	0.008
	(2.79)	(2.80)	(0.29)	(0.29)	(–0.89)	(0.84)
Income squared	–0.001	–0.001	–0.000	–0.000	0.001	–0.001
	(–1.28)	(–1.26)	(–0.27)	(–0.28)	(1.24)	(–0.58)
Number of observations	23,597	23,597	21,067	21,067	10,429	10,620
R-squared	0.03	0.03	0.01	0.01	0.02	0.01
	SRHS		Blood		Smoker	Drinker
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Gini	0.011	0.599**	0.027	–0.058	–0.002	–0.045
	(0.20)	(2.40)	(1.59)	(–0.67)	(–0.04)	(–0.70)
Lagged Gini squared		–0.811**		0.118		
		(–2.38)		(1.02)		
Lagged income	–0.001	–0.002	–0.005*	–0.005*	–0.014	–0.006
	(–0.15)	(–0.17)	(–1.85)	(–1.86)	(–1.63)	(–0.63)
Lagged income squared	0.000	0.000	0.000	0.000	0.002*	0.000
	(0.10)	(0.16)	(1.16)	(1.15)	(1.65)	(0.02)
Number of observations	23,066	23,066	20,530	20,530	10,174	10,365
Pseudo R-squared	0.03	0.03	0.01	0.01	0.02	0.01

Note: *, **, and *** represent significance levels of 10, 5, and 1 percent; robust t-statistics, which allow for correlation of errors within household, are shown in parentheses; the sample includes 1991, 1993, 1997, and 2000 CHNS; all regressions include age and age squared, marital status, family size, and tap water dummy.

While this study has its own limitations, it is among the first to provide evidence from a developing country on the negative association between inequality and health, both of which are important issues for the field of development. Although the sample size is relatively small compared with the data in many U.S. studies, the set of CHNS data we have used is so far one of the best data sets used in studying inequality and health in the context of developing economies, and is probably the best Chinese data set. Another limitation is that we only focus on one dimension of inequality, i.e., community-level inequality. We do not claim that community-level inequality is necessarily more important than inequality at county or provincial level; rather, our purpose is to examine the socioeconomic impacts of inequality in a local setting, where we can see the people interacting with each other more closely. Focusing on the community level can also facilitate the empirical tests by allowing a larger variation of inequality in the sample. Finally, strictly speaking, our empirical tests are tests of correlations between community-level inequality and individual health. The causal link may not be established until more evidence becomes available regarding the intermediate mechanisms through which inequality affects health. However, the significant impact of lagged inequality on current self-reported health suggests that the causality is likely to go from inequality to health because it would be difficult to argue that individual health affects the past inequality.

China began its economic reform by abandoning the principle of absolute equality, "eating from the same kitchen system," in agriculture (Lin, 1992), in industry (Li, 1997), and even in government (Qian and Weingast, 1997). The reforms have improved incentives in most workplaces, which in turn has led to historic levels of growth in the past 25 years. However, the ever-increasing inequality that accompanies growth will ultimately slow it down. A recent study by Benjamin *et al*. (2006) finds that village-level inequality is negatively associated with village economic growth in the long run. While there are many channels through which inequality could affect growth, our paper shows a particular one, poor health, which is itself a direct indicator of underdevelopment.

The Chinese government has apparently taken note of the serious issue of inequality. Wen Jiabao, the new premier, has repeatedly told the public that the goal of this government is to achieve equitable growth. The government has recently been shifting its focus from the more developed coastal areas to the poor inland areas, introducing a series of preferential policies in favor of the latter, such as a wider range of fiscal subsidies, lower tax rates, and cheaper loans. The government is also shifting its focus from the fast developing industries to the sluggish agricultural sector, which employs most of China's poor. Recently, it has started to remove all agricultural taxes nationwide. While it remains to be seen how well these policies are implemented and how effective they are, the government is moving in the right direction in fighting inequality. As suggested by our results, income redistribution will improve the health of the population, especially in regions where large inequality prevails.

References

Aghion, Philippe, Caroli, Eve, Garcia-Penalosa, Cecilia, 1999. Inequality and economic growth: the perspective of the new growth theories. Journal of Economic Literature 37, 1615–1660.

Alesina, Alberto, Perotti, Roberto, 1996. Income distribution, political instability, and investment. European Economic Review 40, 1203–1028.

Ben-Shlomo, Yoav, White, Ian R., Marmot, Michael, 1996. Does the variation in socio-economic characteristics of an area affect mortality? British Medical Journal 312(20), 1013-1014.

Benjamin, Dwayne, Brandt, Loren, Giles, John, 2006. Inequality and growth in rural China: Does higher inequality impede growth? Mimeo. Michigan State University.

Blakely, Tony A., Kennedy, Bruce P., Kawachi, Ichiro, 2001. Socioeconomic inequality in voting participation and self-rated health. American Journal of Public Health 91(1), 99–104.

Blakely, Tony A., Lochner, Kimberley, Kawachi, Ichiro, 2002. Metropolitan area income inequality and self-rated health—-a multilevel study. Social Science and Medicine 54, 65–77.

Bramall, Chris, 2001. The quality of China's household income surveys. China Quarterly 167, 689–705.

Cohen, Sheldon, Line, Scott, Manuck, Stephen, Rabin, Bruce, Heise, Eugene, Kaplan, Jay, 1997. Chronic social stress, social status, and susceptibility to upper respiratory infections in nonhuman primates. Psychosomatic Medicine 59(3), 213–221.

Daly, Mary C., Duncan, Greg J., Kaplan George A., Lynch, John W., 1998. Macro-to-micro links in the relation between income inequality and mortality. The Milbank Quarterly 76(3), 315–339.

Deaton, Angus, 2001. Relative deprivation, inequality, and mortality. Working Paper No. 8099. NBER.

Deaton, Angus, 2003. Health, inequality, and economic development. Journal of Economic Literature 41, 113–158.

Deaton, Angus, Paxson, Christina, 1998. Aging and inequality in income and health. American Economic Review 88(2), 248–253.

Eibner, Christina E., Evans, William N., 2005. Relative deprivation, poor health habits, and mortality. Journal of Human Resources 40(3), 591–620.

Feinstein, Jonathan S., 1993. The relationship between socioeconomic status and health: a review of the literature. The Milbank Quarterly 71, 279–322.

Fiscella, Kevin, Franks, Peter, 1997. Poverty or income inequality as predictor of mortality: longitudinal cohort study. British Medical Journal 314, 1724–1728.

Fiscella, Kevin, Franks, Peter, 2000. Individual income, income inequality, health and mortality: What are the relationships? Health Services Research 35(1), 307–318.

Gerdtham, Ulf-G., Johannesson, Magnus, 2004. Absolute income, relative income, income inequality, and mortality. Journal of Human Resources 39(1), 228–248.

Gravelle, Hugh, 1998. How much of the relation between population mortality and unequal distribution of income is a statistical artefact? British Medical Journal 316, 382–385.

Grossman, Michael, 1972. On the concept of health capital and the demand for health. Journal of Political Economy, 80(2), 223–255.

Gustafsson, Bjorn, Li, Shi, 2002. Income inequality within and across counties in rural China 1988 and 1995. Journal of Development Economics 69, 179–204.

Hsieh, Ching-Chi, Pugh, Meredith D., 1993. Poverty, inequality, and violent crime: a meta-analysis of recent aggregate data studies. Criminal Justice Review 18, 182–202.

Idler, Ellen L., Benyamini, Yael, 1997. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior 38, 21–37.

Kaplan, George A., Pamuk, Elsie R., Lynch, John W., Cohen, Richard D., Balfour, Jennifer L., 1996. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. British Medical Journal 312, 999–1003.

Kawachi, Ichiro, Kennedy, Bruce P., 1997. The relationship of income inequality to mortality: does the choice of indicator matter? Social Science and Medicine 45(7), 1121–1127.

Kawachi, Ichiro, Kennedy, Bruce P., 1999. Income inequality and health: pathways and mechanisms. Health Services Research 34(1), 215–227.

Kawachi, Ichiro, Kennedy, Bruce P., Prothrow-Stith, Deborah, 1997. Social capital, income inequality and mortality. American Journal of Public Health 87(9), 1491–1498.

Kennedy, Bruce P., Kawachi, Ichiro, Glass, Robert, Prothrow-Stith, Deborah, 1998a. Income distribution, socioeconomic status, and self-rated health in the United States. British Medical Journal 317, 917–921.

Kennedy, Bruce P., Kawachi, Ichiro, Prothrow-Stith, Deborah, 1996. Income distribution and mortality: cross-sectional ecological study of the Robin Hood Index in the United States. British Medical Journal 312, 1004–1007.

Kennedy, Bruce P., Kawachi, Ichiro, Prothrow-Stith, Deborah, Lochner, Kimberly, Gupta, Vanita, 1998b. Social capital, income inequality, and firearm violent crimes. Social Science and Medicine 47(1), 7–17.

LeClere, Felicia B., Soobade, Mah-Jabeen, 2000. The effect of income inequality on the health of selected US demographic groups. American Journal of Public Health 90, 1892–1897.

Li, Wei, 1997. The impact of economic reform on the performance of Chinese state enterprises, 1980–1989. Journal of Political Economy 105(5), 1080–1106.

Lin, Justin, 1992. Rural reforms and agricultural growth in China. American Economic Review 82(1), 34–51.

Lynch, John, Davey Smith, George, Harper, Sam, Hillemeier, Marianne, Ross, Nancy, Kaplan, George A., Wolfson, Michael, 2004. Is income inequality a determinant of population health? Part 1. A systematic review. The Milbank Quarterly 82(1), 5–99.

Lynch, John W., Kaplan, George A., Pamuk, Elsie R., Cohen, Richard D., Heck, Katherine E., Balfour, Jennifer L., Yen, Irene H., 1998. Income inequality and mortality in metropolitan areas of the United States. American Journal of Public Health 88(7), 1074–1080.

Marmot, Michael G., 1997. Inequality, deprivation and alcohol use. Addiction 92(Suppl. 1), s13.

Marmot, Michael G., Davey-Smith, George, *et al.*, 1991. Health inequalities among British civil servants: the Whitehall II study. Lancet 337(8754), 1387–1393.

Marmot, Michael G., Shipley, Martin, Rose, Geoffrey, 1984. Inequalities in death-specific explanations of a general pattern. Lancet 323(8384), 1003–1006.

Meara, Ellen, 1999. Inequality and infant health. Manuscript. Harvard Medical School.

Mellor, Jennifer M., Milyo, Jeffery, 2002. Income inequality and health status in the United States: evidence from the current population survey. The Journal of Human Resources 37, 510–539.

Osler, Merete, Prescott, Eva, Gronbaek, Morten, Christensen, Ulla, Due, Pernille, Engholm, Gerda, 2002. Income Inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies. British Medical Journal 324, 13–16.

Preston, Samuel H., 1975. The changing relation between mortality and level of development. Population Studies 29, 231–248.

Qian, Yingyi, Weingast, Barry, 1997. Federalism as a commitment to reserving market incentives. Journal of Economic Perspectives 11(4), 83–92.

Robert, Stephanie A., 1998. Community-level socioeconomic status effects on adult health. Journal of Health and Social Behavior 39, 18–37.

Runciman, Walter G., 1966. Relative Deprivation and Social Justice. Routledge and Kegan Paul, London.

Shibuya, Kenji, Hashimoto, Hideki, Yano, Eiji, 2002. Individual income, income distribution, and self-rated health in Japan: cross-sectional analysis of nationally representative sample. British Medical Journal 324, 16–19.

Shively, Carol A., Laber-Laird, Kathy, Anton, Raymond F., 1997. Behavior and physiology of social stress and depression in female cynomolgus monkeys. Biological Psychiatry 41, 871–882.

Smith, James P., 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. Journal of Economic Perspectives 13, 145–166.

Soobader, Mah-Jabeen, LeClere, Felicia, 1999. Aggregation and the measurement of income inequality: effects on morbidity. Social Science and Medicine 48(6), 733–744.

Waldmann, Robert J., 1992. Income distribution and infant mortality. Quarterly Journal of Economics 107(4), 1283–1302.

Wennemo, Irene, 1993. Infant mortality, public policy, and inequality—a comparison of 18 industrialized countries, 1950–1985. Sociology of Health and Illness 15(4), 429–446.

Wilkinson, Richard G., 1992. Income distribution and life-expectancy. British Medical Journal 304, 165–168.

Wilkinson, Richard G., 1996. Unhealthy Societies: the Affliction of Inequality. Routledge, London.

Wilkinson, Richard G., 1997. Health inequalities: relative or absolute material standards? British Medical Journal 314, 591–595.

Yitzhaki, Shlomo, 1979. Relative deprivation and the Gini Coefficient. Quarterly Journal of Economics 93(2), 321–324.