

Relative Deprivation and Health: Which Reference Groups Matter?¹

Eiji Mangyo
Graduate School of International Relations
International University of Japan

Albert Park
Department of Economics
University of Oxford

Abstract

We examine the extent to which self-reported health and psycho-social health are affected by relative economic status in China, for the first time examining the importance of reference groups not defined by geographic location or demographic characteristics. We propose a methodology to address potential bias from subjective reporting biases and control for unobserved community characteristics. Analyzing a nationally representative dataset from China, our findings support the relative deprivation hypothesis and suggest that relatives and classmates are salient reference groups for urban residents and neighbors are important for rural residents.

JEL codes: I12, O15, O53

¹ We thank the Editor and two anonymous referees for very helpful comments.

1. Introduction

A growing literature provides evidence that individual health and subjective well-being are influenced by relative economic status (Eibner and Evans 2005, Ferrer-i-Carbonell 2005, Luttmer 2005). According to the relative deprivation hypothesis, feeling less well off than others creates unhappiness and stress which leads to worse health, contributing to a negative relationship between income inequality and health (Wilkinson 1996). The theory assumes that utility is a function not only of own consumption, but also of the consumption level of others in one's social reference groups. The relative deprivation hypothesis can rationalize empirical evidence that the health gradient exists over the whole spectrum of socioeconomic status in rich countries rather than disappearing above an income threshold (Marmot et al 1991, Davey Smith et al 1990, Der et al 1999). It is also consistent with the strong positive correlation between income and subjective well-being found in cross-sectional data but the failure of average happiness to increase as societies become richer (Easterlin, 1995). Health and subjective well-being are likely to be strongly connected given that medical studies find a large impact of stress on the incidence and progression of many illnesses (Lovallo 1997, Sapolsky 1998).

A major challenge in studying the relative deprivation hypothesis is defining appropriate reference groups to which individuals compare themselves (Eibner and Evans, 2005). Because of the limited content of most social surveys, previous empirical studies all have defined reference groups based on geographic location or demographic characteristics such as age, gender, or ethnicity. However, other social reference groups with which individuals have frequent social contact may be much more salient, for example relatives, co-workers, and former classmates. Even for geographic reference

groups, previous research has not systematically studied which level of geographic aggregation has the greatest influence on individual health and sense of well-being. In particular, neighbors living in close proximity may be a particularly important reference group. It is also likely the relative importance of different types of social reference groups may differ across individuals.

In this paper, we empirically investigate for the first time the impact on health of relative deprivation defined with respect to multiple social and geographic reference groups. We analyze data from a unique national survey in China that includes questions about respondents' subjective assessments of how their living standards compare to different geographic and non-geographic reference groups. China is a particularly interesting case because it is not only the world's most populous country but also a nation which has witnessed a very rapid increase in income inequality during its transition from a socialist planned economy to a market-based system. China's gini coefficient increased from 0.309 in 1981 to 0.453 in 2003 (World Bank, 2009).

This paper also makes a methodological contribution by showing how subjective questions about relative economic status can be used to investigate the relative deprivation hypothesis without being undermined by subjective reporting biases. Such biases arise because individual outlooks (i.e., optimism or pessimism) may influence both subjective welfare assessments and self-reports of health status. We address this problem by directly estimating the magnitude of such biases and controlling for them in our econometric analysis, and also by employing instrumental variables. In addition, we control for unobserved regional characteristics affecting health by including regional fixed effects, which is not possible in most other studies in this literature which use regional income measures to test the relative deprivation hypothesis. The lack of

adequate controls for omitted regional characteristics may be one reason why previous studies find mixed results on the effect of relative deprivation on health.

The rest of the paper is organized as follows. The next section reviews previous studies on the effects of relative income on health. Data and descriptive statistics are presented in Section 3, Section 4 describes the research methodology, Section 5 presents the results, and Section 6 concludes.

2. Relative Deprivation and Health

Relative economic status can influence health through multiple mechanisms. First, as described above feelings of relative deprivation may directly influence an individual's sense of well-being or happiness. Studies have found a link between stress caused by economic hardship and health-related behaviors such as smoking, heavy alcohol use, and less healthy diet (Conway et al 1981, Gorman 1988, Horwitz and Davies 1994, Jensen and Richter 2004, Kristenson et al. 1999). Second, other social and political mechanisms also may link relative economic status and individual health. The relatively poor may lack social cohesion with others; the quality of social relationships has been found to be associated with unhealthy behaviors and poorer health (House et al. 1999). Third, the relative poor may have relatively less access to health care or other services if access is rationed or subject to political influence. Although it is difficult to distinguish empirically between the first two explanations, the third is less likely to strongly affect psychosocial health than the first two.

This study extends the work of previous authors who have examined the empirical relationship between relative deprivation and health outcomes and health behaviors. Eibner and Evans (2005) analyze U.S. microdata and find that relative

deprivation with respect to individuals with similar demographic characteristics reduces self-reported health status, increases mortality, and increases risky health behaviors (smoking, obesity, less exercise). Other research based on survey data also has found an empirical link between relative deprivation and both mortality and suicide (Eibner and Evans 2005, Miller and Paxson 2006, Daly, Wilson, and Johnson 2007), but not all studies find a mortality effect (Gerdtham and Johannesson 2004). Even experimental research on primates has found that low social status leads to higher cholesterol, increased atherosclerosis, obesity, and depression (Shively and Clarkson 1999; Shively, Laber-Laird, and Anton 1999; Sapolsky, Alberts, and Altmann 1999). A number of studies also have found a close relationship between relative economic status and subjective well-being (Luttmer, 2005, Ferrer-i-Carbonell 2005, McBride 2001). As noted earlier, all of these studies examine reference groups defined by geographic and demographic characteristics and are unable to fully rule out bias from omitted regional characteristics.

As first motivated by Wilkinson (1996), the relative deprivation hypothesis could explain a negative relationship between income inequality and average health of the population. However, it is important to point out that the inequality-health relationship could also be influenced by factors unrelated to the relative deprivation hypothesis, such as concavity of the income-health relationship and less provision of public goods (e.g., health services) in communities with greater income inequality due to political economy reasons, etc.² In this paper, we restrict attention to testing the impact of relative economic status on individual health, and do not directly address the

² See Bénabou (1996) for reasons why a more equal income distribution within a region could increase expenditures on public goods within the region. Lynch and Kaplan (1999) argue that as the gap between the rich and the poor widens, interests diverge, translating into reduced social spending and leading to reduced access to life opportunities especially among the poor.

recent literature relating income inequality to aggregate health (Deaton and Paxson 2001, Deaton 2002 and 2003, Kawachi et al 1997; Kennedy et al 1998, Mellor and Milyo 2002).

3. Data

Our data come from the China Inequality and Distributive Justice survey project³ conducted in the fall of 2004 which collected data on a nationally representative sample of 3267 Chinese adults between the ages of 18 and 70 living in 23 of China's 31 provinces and in 65 counties and 85 townships.⁴ Respondents were selected using spatial probability sampling methods. Important for our purposes, the survey included questions asking the respondents to rate their living standards in comparison with multiple reference groups. Specifically, the survey asked: "Compared with the average living standard of [your relatives, classmates with the same level of schooling as you, your coworkers, your neighbors, others in the same county or city, others in the same province, others living in China], do you feel your living standard is much better, a little better, about the same, a little worse, or much worse?" These questions are coded from one to five, with five being "much better".

We use two health measures. The first is self-reported health status (1=very poor, 2=poor, 3=average, 4=good, 5=very good). The second is an index measure of psychosocial health based on eight questions: "Below are some descriptions of people's life conditions. In the past week, did you experience these conditions: often, sometimes,

³ We acknowledge primary funding support for the survey from the Smith Richardson Foundation, with supplementary funding provided by Harvard's Weatherhead Center for International Affairs, the University of California at Irvine, and Peking University.

⁴ The provinces are Liaoning, Heilongjiang, Beijing, Hebei, Shanxi, Shaanxi, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Hunan, Guangdong, Hainan, Anhui, Jiangxi, Henan, Hubei, Ningxia, Xinjiang, Guangxi, Yunnan, and Tibet. In cities, a "county" signifies an urban district and "township" signifies an urban sub-district (*jiedao*).

rarely, or never? a. I worry about some small things. b. I have no appetite for food. c. I cannot focus my attention while doing things. d. I feel my life is a failure. e. The quality of my sleep is poor. f. I feel fortunate. g. I feel alone. h. I feel my life is very happy.” The answers to each question are coded from 1 to 4, with 4 being better psychosocial health. To calculate an index of psychosocial health, the answers to each question are normalized to be standard deviations from the mean, and the index is the mean of the normalized scores for the eight questions. Psychosocial health is a particularly appropriate measurement for examining the relative deprivation hypothesis which posits that health is affected largely through dissatisfaction or stress caused by relative economic status.

The survey also collected information about individual demographic characteristics and household income. Table 1 shows descriptive statistics for the main variables. The sample size is 2,891 individuals with complete information for all variables used in the regression analysis. The self-reported health status of the great majority of individuals is average or better (85.8%). With regard to psychosocial health, a significant share of respondents report sometimes or often having negative experiences such as having no appetite (33.2%), being unable to focus attention (31.0%), feeling their life is a failure (23.8%), having poor quality sleep (39.0%), and feeling alone (26.3%). For a full description of answers to all questions used in the index, see Appendix Table 1.

Next we describe respondents’ perceptions of how their own living standards compare to seven reference groups: three non-geographical (relatives, classmates, and co-workers), and four geographical (neighbors, county or city, province, and nation). Full results are presented in Appendix Table 2. For the non-geographical reference

groups, the most common response is that living standards are similar to others in the reference group, but there is substantial variation in these rankings.⁵ For geographic reference groups, an interesting pattern emerges in which the greater the geographic scope of the reference group, the more likely that individuals report being less well off than the reference group. Thus, a majority of respondents feel that they have lower living standards than the typical person in China.

The survey asks respondents to report their household income for the entire year of 2003. Of the total 3,267 sample individuals, 2,907 report income: 2,517 individuals (77%) report a precise value for their household income and 390 individuals (12%) report a range for their household income in which case income is set equal to the range mid-point.⁶ Less than one percent of respondents (25) report zero household income; in order to be able to calculate log per capita income, for respondents with income in the lowest one percent, we set per capita income equal to per capita income at the first percentile.⁷

Next, we show how self-reported health and psychosocial health are correlated with how individuals rate themselves with respect to different reference groups. We calculate the mean health of individuals who report different relative economic status and plot the results (Figure 1). We also plot the relationship between health and household income level. For self-reported health status, there are strong associations between health and the perception of own living standards in comparison with others in each of the seven reference groups. We also see a strong correlation between

⁵ Some questions have high percentages of missing values because the questions are not applicable, for instance the respondent did not attend school or has little or no contact with former classmates, or the person is not working.

⁶ There are 16 ranges of household income separately for both agricultural and non-agricultural households. For individuals who report the top income range (over 200,000 RMB), we set household income equal to 200,000 RMB.

⁷ Fifty three sample individuals report per-capita household income below the cutoff which is 137.5 RMB. Results are not sensitive to the specific cutoff line chosen.

self-reported health and per-capita income quintile. Turning to the psychosocial health index, the associations are less evident in general, but psychosocial health scores are clearly low when individuals consider that their living standards are somewhat or much worse than that of others.

4. Empirical Methodology

According to the relative deprivation hypothesis, controlling for a person's own income, people feel more deprived and thus have poorer health when others in a reference group have higher incomes. A frequently used objective measure of relative income is the mean living standards of others living in the same geographic area (sometimes also broken down by demographic group). In this study, we use both township-level and province-level mean log income per capita as objective relative deprivation measures in addition to the seven subjective relative income measures described above.⁸ Township may denote a rural town (*xiang* or *zhen*) or an urban subdistricts (*jiedao*), which share the same administrative level in China.

The basic estimating equation is as follows:

$$H = \beta_0 + \beta_1 Y + \beta_2 R + \beta_3 X + \beta_4 TOWN + \varepsilon, \quad (1)$$

where H is self-reported health status or psychosocial health, Y is log per-capita household income, R is a measure of relative income, X is a vector of demographic variables including an indicator variable for whether the place of residence is in an urban or rural community, $TOWN$ is a vector of township dummy variables, ε is an error term, and the β_i are coefficients to be estimated.

⁸ Most sample counties (70.8%) contain only one sample township, making it impossible to test the impact of both county- and township-level mean incomes.

OLS with objective relative income measures

First, we estimate equation (1) using OLS, objective relative income measures, and no regional fixed effects, which is similar to the specification adopted in many previous studies. We cluster standard errors at the township level to take into account potential correlations of health across individuals living in the same community. The relative deprivation hypothesis predicts that mean regional income levels should have a negative impact on health. However, without controlling for unobserved regional characteristics, the coefficient on regional log income per capita is likely to be biased upward if richer communities have better infrastructure and public services, including health care, which improve health outcomes.

OLS with subjective relative income measures

Next, we estimate Equation (1) using OLS with subjective measures of relative income, initially excluding regional fixed effects as before. We estimate the impact of each subjective relative income measure one by one (relatives, classmates, coworkers, neighbors, county, and province).

As noted earlier, one advantage of using subjective measures of relative income with respect to non-geographic reference groups is that we can include regional fixed effects since the subjective measures vary at the individual level.⁹ The lowest level of clustered sampling in our dataset is the township level, so inclusion of township fixed effects maximizes our ability to control for unobserved geographic characteristics. Although it is possible that regional differences at lower levels of spatial aggregation (i.e.,

⁹ As a technical matter, given individual variation in subjective assessments it is possible to estimate the impact of subjective measures of relative income with respect to geographic reference groups even when including geographic fixed effects; however, such a specification is not consistent with a causal interpretation of relative income on health.

rural village or urban neighborhood) could still bias our estimates, in China townships are the lowest administrative level of government and are the site of state-run medical clinics, bank branches, and other important local institutions. Therefore, township fixed effects should effectively control for key unobserved differences in policy, quality of medical services, and other geographic and institutional factors. When we include township fixed effects we are unable to include objective measures of relative income status with respect to geographic reference groups aggregated to the township level or above. We cluster standard errors at the township level for the same reason as before.

Addressing unobserved outlook bias

We are concerned about bias due to unobserved individual outlooks, because our measures of health and relative income are both subjective. For example, people who like to complain or are pessimistic may report poorer self-reported health as well as lower relative income status. In this section, we show how one can directly estimate individual outlook bias and explicitly control for it in estimation. The essential idea is to define the difference between one's subjective rating of relative living standards within one's county of residence and one's actual income position within the county as a measure of optimism or pessimism. We focus on county comparisons because the county is the smallest geographic area with sufficient observations for both objective and subjective income measures.

To see more clearly how this is implemented, consider the following health production function:

$$H = H(Y_i, R_{ij}, u_i, o_i, u_j) \quad (2)$$

Here, Y_i is true household income, R_{ij} is subjective relative income of individual i with

respect to some reference group J , u_i is individual and household unobservables affecting health that are independent of attitudinal biases that affect reporting of relative incomes, o_i is the unobserved outlook of individuals that can affect both health and perceptions of relative income, and u_J is unobserved group characteristics that affect health.

We posit that R_{ij} is a function of household income, mean community income, and outlook bias:

$$R_{ij} = R(Y_i, Y_J, o_i) \quad (3)$$

We recognize that Y_i and Y_J may both be measured with error:

$$\tilde{Y}_i = Y_i + e_i \quad \text{and} \quad \tilde{Y}_J = Y_J + e_J. \quad (4)$$

Thus, \tilde{Y}_i and \tilde{Y}_J are noisy measures of true household income and group mean income.

We propose the following approach to estimating equation (2) to reduce likely biases. We start by first estimating the determinants of the relative income measure with respect to a reference group for which we have multiple observations within each group, i.e., counties. Assuming that relative income can be expressed as a linear function of its elements, we can estimate the following equation:

$$R_{ij} = \alpha_0 + \alpha_1 \tilde{Y}_i + \gamma_J + \varepsilon_i^R \quad (5)$$

The county fixed effect γ_J absorbs the effect of true group mean income (Y_J). Given equations (3) and (4), if we estimate (5) using OLS the error term will have two components, $\varepsilon_i^R = -\alpha_1 e_i + o_i$. Since the error term contains e_i , it is clearly correlated with \tilde{Y}_i , leading to biased estimates due to measurement error. However,

using instrumental variables, it is possible to estimate α_1 consistently so that the error term will consist only of omitted outlook bias. In estimating (5), we instrument log income per capita using the two assessments of the household's standard of living relative to others in the community made by the survey enumerator who interviewed the household, which are plausibly independent of income measurement error and outlook bias.¹⁰ We use the residual from this 2SLS estimation as our estimate of outlook bias ($\hat{\phi}_i$).¹¹ This estimate is unbiased but measures actual outlook bias with noise. If we assume that outlook bias affects all self-reported relative income measures similarly, these residuals can be used to control for outlook bias when estimating the effects of subjective relative income with respect to reference groups for which multiple observations are not available.

We now are ready to estimate the following equation for the determinants of health outcomes:

$$H_i = \beta_0 + \beta_1 \tilde{Y}_i + \beta_2 R_{iJ} + \beta_3 \hat{\phi}_i + \varepsilon_i^H \quad (6)$$

Ignoring measurement error in the income variables for a moment, given the health production function described in equation (2), in equation (6) the error term includes u_i which is independent by assumption, and u_J . It is clear that we will be unable to fully isolate the effect of relative income on health, because group mean income (or unobservables correlated with group mean income) can have independent effects on health. However, as a general rule, most of the likely effects of greater group mean

¹⁰ The two specific questions are the following: 1) From your impression of the respondent's household, please evaluate whether in the local area the household would be considered a low income household, average income household, upper middle income household, or high income household?; and 2) How does the respondent's home compare to the average home in the area: below average, average, or above average?

¹¹ The excluded instruments have decent explanatory power in the first stage regression as evidenced by $F(2, 2833)=132.87$. The excluded instruments pass the over-id test well above the conventional levels of significance (p-value=0.83).

income on health should be positive. For example, having more affluent relatives, classmates, or coworkers could improve one's health through remittances, information about health, or help in accessing better quality health care services. In contrast, the relative deprivation hypothesis predicts that greater group mean welfare (or lower relative income) reduces a person's own health. This suggests that a positive effect of relative income on health should be viewed as strong evidence in favor of the relative deprivation hypothesis. Differences in the effect of relative income with respect to different reference groups could reflect differences in both the impact of relative deprivation on health and other independent effects of group mean welfare on health.

Instrumenting for income

In the OLS regressions, per capita household income could be correlated with the error term, leading to biased estimation. Income could be measured with error, leading to attenuation bias toward zero, or there could be reverse causality running from health to income leading to upward bias in the coefficient on income. To address these issues, we employ instrumental variables. The instruments are a set of seven wealth indicator variables which reflect whether the household owns the following assets: motorcycle, car, refrigerator, color TV, computer, phone, and washing machine. These instruments have the advantage of being highly objective questions unlikely to be prone to measurement error or reporting bias. To the extent that durables reflect permanent rather than transitory income, the instruments should effectively control for the impact of short-term health shocks on income, part of the second source of bias. However, some bias may still exist if current wealth is greater as the result of better long-term health which persists to the present or if wealth, income, and health are all positively related to

other unobserved individual characteristics.

If OLS estimates of the coefficient on per-capita household income are biased, we expect opposite bias in the coefficient on relative income. To see why this is true, consider the case in which relative income is defined as own income minus mean group income, so that the relationship between health and true own income and relative income variables can be written as follows:

$$H_i = \beta_0 + \beta_1 Y_i + \beta_2 (Y_i - Y_J) + \varepsilon_i^H = \beta_0 + (\beta_1 + \beta_2) Y_i - \beta_2 Y_J + \varepsilon_i^H \quad (7)$$

From equation (7), the impact of an increase in own income on health is equal to the sum of β_1 and β_2 and the impact of an increase in group mean income is equal to $-\beta_2$. We see immediately that the main effect of simultaneity caused by the positive effect of health on income and wealth is that $\hat{\beta}_1$ will be biased upward. Bias in $\hat{\beta}_1$ will cause $\hat{\beta}_2$ to also partially reflect the impact of own income as well. For example, if there were no control for own income ($\hat{\beta}_1 = 0$), then $\hat{\beta}_2$ would reflect the average of the impacts of changes in own and mean group income which would lead to upward bias if $\beta_1 > 0$ and downward bias if $\beta_1 < 0$. More generally, $\hat{\beta}_2$ has the opposite direction of bias of $\hat{\beta}_1$. Thus, if there is attenuation bias in $\hat{\beta}_1$, $\hat{\beta}_2$ from OLS would be biased upward, and if there is reverse causality in $\hat{\beta}_1$, $\hat{\beta}_2$ from OLS would be biased downward.

Use of instruments should correct for either direction of bias. However, if wealth is correlated with health independently of income for the reasons described earlier, the IV estimate of the coefficient on own income would be biased upward, which would lead to downward bias on the coefficient of subjective reports of relative income. Thus, the IV estimates of the effect of relative income on health are likely to be lower-bound estimates.

Initially, we do not instrument for the relative income variables. This is an appropriate choice if we believe that such measures have relatively little measurement error and that perceptions of relative income are as important, or even more important, for health than actual relative income. However, just as for own income, estimates of the impact of subjective relative income on health could suffer from downward bias due to measurement error or upward bias due to reverse causality. In addition, to the extent that our estimate of outlook bias is incomplete, then there could remain upward bias caused by the correlation between unobserved individual outlooks and both subjective health measures and subjective measures of relative income. We have a sufficient number of wealth variables to treat both per capita household income and a relative income variable as endogenous in the same regression. Because wealth variables predict actual relative income (through their correlation with own income), unlike the OLS estimate, the IV estimate does not reflect additional health impacts of perceptions of relative income.

If self-reports of health status reflect comparisons individuals make between their own health and others in their social reference group, then measured impacts of relative income on health could influence how respondents report their health rather than their actual health. Use of instrumental variables cannot effectively address this type of bias.¹² Our measure of psychosocial health could be considered more objective and less subject to such bias, given that it is based on questions which ask about the frequency of specific experiences rather than for a general assessment. To address this

¹² This could be also caused by “focusing illusion” (Kahneman et al., 2006) if in the context of the survey itself, questions about relative economic status make comparisons to specific reference groups more salient at the time respondents answer questions about health. In the survey, health questions are asked some time after a set of questions about relative economic status, which should reduce focusing illusion and make it difficult for such bias to explain different health effects of different reference groups.

concern, ideally one would test how relative income affects highly objective measures of health status, such as physical health exam results or incidence of specific diseases or health conditions; unfortunately, such data are not available in the survey.

5. Results

OLS with objective relative income measures

We first report results of OLS estimation using objective relative income measures with respect to geographic reference groups. Columns 1 and 2 of Table 2 present the results for self-reported health status and psychosocial health, respectively. For both measures, higher own income is significantly associated with better health. Based on the point estimates a doubling of income would increase self-reported health status by 0.113 ranks (0.109 standard deviations) and improve psychosocial health by 0.098 standard deviations. Township mean log per-capita household income negatively affects both health measures after controlling for own income, consistent with the relative deprivation hypothesis. However, the coefficient on township mean income is statistically significant at conventional levels only for psychosocial health (column 2). In fact, for psychosocial health, the magnitude of the negative effect of mean township income is even greater than the positive effect of own income.

For other demographic variables, the coefficient estimates are mostly as expected. Males report themselves to be healthier than females. As people age, their health becomes worse at a decreasing rate (the coefficient on the age squared term is positive and statistically significant at the 10% level for psychosocial health). Psychosocial health of unmarried persons is significantly worse than that of married persons, but self-reported health status is similar. Education has a nonlinear relationship

with health. Health increases with educational attainment through middle school, but with more education beyond middle school it fails to increase or even falls. Residing in urban areas is associated with poorer health after controlling for other covariates, but the effect is not statistically significant at conventional levels.¹³

We also investigate whether provincial income per capita predicts health outcomes. Results are presented in columns 3 and 4 of Table 2. Provincial income per capita is negatively associated with psychosocial health but positively associated with self-reported health; however, the associations are not statistically significant at conventional levels. This suggests that others in the township are a more salient comparison group than others in the province, or that unobserved provincial differences associated with mean income have more positive impacts on health than unobserved township differences.

OLS with subjective relative income measures

Next, we report the coefficients when we use subjective assessments of relative income with different reference groups in Table 3.¹⁴ In each of these regressions, we include the same covariates as before, but because they do not differ substantially from the patterns found in Table 2, we do not report their coefficients in this or subsequent tables. As noted earlier, some subjective assessments have a nontrivial number of missing observations; to deal with this, we add a dummy variable for whether the relative income measure is missing and assign a zero to the measure of relative income.¹⁵

¹³ The control variables also include a dummy for bottom-coding of per-capita household income (set equal to the one percent cutoff value).

¹⁴ We do not examine subjective comparisons with China as a whole because this would lead to a fallacy of composition in that since the sample is a national one, for national comparisons relative incomes and absolute incomes are indistinguishable.

¹⁵ Results do not differ appreciably if we simply drop the observations with missing values.

We start with the exact same specification as in Table 2, but using subjective relative income measures rather than objective ones. Results are reported in columns 1 and 2 of Table 3. In general, the coefficient on the household's own income is positive and statistically significant, but somewhat smaller in magnitude than when objective relative income measures are used. This is to be expected, since the relative income measure is the difference between own income and that of others rather than the mean income of others. This means the coefficient on relative income captures part of the effect of own income and that the coefficient on own income captures the extent to which own income is more important to health than group mean income.

Regardless of the reference group (relatives, classmates, coworkers, neighbors, county, or province), higher subjective assessments of relative income are associated with better health, and almost all of the coefficients on the relative income variables are statistically significant at the one or five percent level. In terms of magnitude, for both health measures, relative economic statuses with respect to relatives, classmates, and neighbors have larger coefficient estimates than relative comparisons with the other groups. A one rank increase in relative living standards in comparison to relatives/classmates/neighbors increases self-reported health status by 0.169 to 0.188 ranks (0.163 to 0.182 standard deviations) and improves psychosocial health by 0.096 to 0.128 standard deviations. Further, the importance of relative income comparisons with non-geographic reference groups and neighbors is consistently greater in magnitude than that of comparisons with conventionally defined, larger geographic reference groups.¹⁶ This highlights the importance of considering non-geographic reference groups and neighbors living in close proximity in studies of the relative deprivation

¹⁶ The only exception is that for self-reported health status, the coefficient on relative income in comparison to others in the same county is higher than that on relative income in comparison to co-workers.

hypothesis.

Finally, we estimate a specification in which we include all six subjective measures of relative income together (relatives, classmates, coworkers, neighbors, county, and province). For SRHS, comparisons with relatives, classmates, neighbors, and others in the same county support the relative deprivation hypothesis and are statistically significant at the ten percent level or better in one or more specification. For psychosocial health, the same is true for comparisons with relatives and classmates. Not surprisingly, the individual magnitudes of the coefficients fall, but the relative importance of different comparison groups is the same as in the individual regressions. Relative income with respect to coworkers is no longer statistically significant. The only odd result is that for psychosocial health better living standards in comparison to others in the same province has a negative statistically significant effect on health status. As noted before, this could reflect the importance of province-level unobservables that are positively correlated with both provincial income per capita and individual health (such as health care service quality).

As pointed out earlier, the simple OLS estimates presented thus far are subject to a number of potential sources of bias. To begin addressing these, we first examine how the results change when we include township fixed effects in the regressions examining the importance of relative income comparisons with non-geographic reference groups or neighbors. Inclusion of the township fixed effects controls for potential bias associated with unobservable community level factors that are correlated with both income levels and health. Results are presented in Table 3, columns 3 and 4. We find that the coefficients on the relative income measures become smaller in magnitude with township fixed effects in almost all cases, typically by 5-20 percent, but

remain highly statistically significant. This reduction in impact is not surprising since without township fixed effects, differences in self-reported relative economic status is likely to be positively correlated with community wealth differences and the quality of health services, which promote better health and so lead to upward bias in the coefficient estimate. Township fixed effects reduce the bias from such differences in regional wealth by controlling for the wealth of the respondent's own region. With township fixed effects, the coefficient on household income becomes greater; this likely corresponds with the lower coefficient on relative income, since that coefficient also captures part of the impact of household income on health outcomes.

Next, we add our estimated outlook bias as a control variable, following the methodology outlined earlier. Results are presented in Table 3, columns 5 and 6. The outlook bias term is positive and statistically significant in all of the regressions. Controlling for outlook bias also reduces the estimated impact of all subjective relative income measures on health outcomes, typically by 15-35 percent, but most of them remain statistically significant. The exceptions are comparisons to coworkers which have lost statistical significance for both health measures and comparisons to neighbors which have lost statistical significance for psychosocial health. The magnitude of outlook bias is remarkably consistent across comparison groups, underscoring the importance of dealing with reporting biases in the estimation procedure.

Using the specification with township fixed effects and controlling for outlook bias, we also examine the relative income hypothesis separately for urban and rural residents. Results are presented in Table 4. First, we find that own income matters to health more for rural ~~and less-educated~~ respondents, which makes sense since ~~these~~ this populations tends to be poorer. Second, comparisons with relatives and with

classmates matter more for urban residents, and comparisons with neighbors are more important for rural residents. Finally, the estimated coefficients on estimated optimism are larger and more statistically significant for rural respondents for self-reported health, but the opposite is true for psychosocial health. We also estimated results separately for those with high school education and above and those with middle school education and below, and found similar but less pronounced results. This makes sense given that rural populations are much less well educated than urban populations.

Instrumental Variables

Next, we introduce instrumental variables to address bias created by the endogeneity of per capita household income. Table 5 presents the results of the 2SLS regressions when the seven asset indicator variables are used as instruments for per capita household income. In different estimations, we use different measures of objective and subjective relative income while continuing to control for outlook bias, township fixed effects, and other covariates. For each estimation, we report first-stage F-statistics and over-id test results. The F-statistics for the excluded instruments are all ~~above~~ 17 or above, sometimes well above, suggesting that there is no weak instrument problem. The instruments pass the over-ID tests in all cases using a 10% significance threshold.

A comparison of the IV results presented in Table 5 with the OLS results presented in Table 2 and in the last two columns of Table 3 yields some strong regularities. When testing objective relative income measures, using IVs increases the magnitude of the positive coefficient on per capita household income, which is consistent with measurement error and attenuation bias plaguing the OLS estimates and

dominating any opposite bias caused by reverse causality. However, when using subjective relative income measures, the IV estimates produce larger coefficients on household income per capita when the dependent variable is psychosocial health but smaller coefficients when it is SRHS. This difference could reflect the fact that SRHS relates to physical health which has a greater impact on income than psychosocial health, leading to greater simultaneity bias.

For the objective relative income measures, township mean income and province mean income, the coefficient on the mean income of the reference groups are more negative, providing stronger support for the relative deprivation hypothesis. However, the coefficient estimates are statistically significant (at the 10% level) only for the effect of township income per capita on psychosocial health.

In comparison to OLS the IV estimates of the coefficients on subjective relative income are smaller in magnitude when the dependent variable is psychosocial health but larger in magnitude when the dependent variable is SRHS. This is to be expected, since as explained earlier downward bias in the coefficient on per-capita household income leads to upward bias in the coefficient on subjective relative income and vice versa. The standard errors also tend to be larger with the instrumental variables estimation. Nonetheless, for SRHS the relative comparisons with relatives, classmates, and neighbors remain statistically significant, ~~most~~ at the 5% significance level. Only for coworkers is the impact of relative living standards not statistically significant for either health measure. For psychosocial health, the impact of relative deprivation with respect to classmates and neighbors are not statistically significant. The ordering of the magnitudes of the coefficients on the four subjective relative income measures is similar for both health measures. The most salient reference group affecting self-reported

health status is classmates, followed by neighbors, relatives, and coworkers; for psychosocial health the order is the same except the importance of neighbors and relatives is reversed. These orderings are similar to the ordering we found using OLS. These health effects are driven by differences in actual relative incomes or differences in perceptions of relative income. If the asset instruments are positively associated with health independent of income, then the IV estimates are lower bounds for the actual importance of relative deprivation to health.

We also present the IV results for rural and urban residents separately in Table 6. The differences are quite striking. For urban residents, the salience of classmates is the most important for both health outcomes, followed by relatives. For these two reference groups, all of the impacts are statistically significant and the magnitudes of the effects are notably larger than for the pooled sample (0.291 for SRHS and 0.148 for psychosocial health for classmates compared to 0.153 and 0.074 for the pooled sample, and 0.185 for SRHS and 0.106 for psychosocial health for relatives compared to 0.120 and 0.065 for the pooled sample). For self-reported health status, relative income with respect to co-workers is also statistically significant, but less important in magnitude. For neither health outcome are neighbors a salient reference group. Nor are mean township or province income per capita.

For rural residents, the salient reference groups are completely different than for urban residents. Relative income with respect to classmates, relatives, and coworkers has no impact on health outcomes, but relative income compared to neighbors affects both SRHS (coefficient of 0.157) and psychosocial health (coefficient of 0.072). The latter effect is statistically significant at the 10 percent level even though it was not statistically significant for the pooled sample. Also, unlike for urban residents,

the impact of regional mean income levels, especially township income per capita, has a large and statistically significant negative impact on both health outcomes. These findings suggest that while geographic reference groups are salient for rural residents, classmates and relatives are more salient for urban residents.

What happens if we use the asset variables to instrument for both per capita household income and subjective relative income at the same time? To get a sense of how this affects the estimation results, we present the results for the pooled sample in Appendix Table 3.¹⁷ The coefficients on subjective relative income measures increase substantially in magnitude for comparisons with relatives, classmates, and coworkers, consistent with significant measurement error, but are imprecisely estimated. However, the coefficients on relative living standards with respect to neighbors are negative, though not statistically significant. The only estimates which are statistically significant at the 10% level are the impact of relative income of relatives on self-reported health status and the impact of relative income of classmates on psychosocial health.

In most cases where the IV coefficients for relative income have increased toward the positive direction in comparison with OLS coefficients, the coefficients on per capita household income have lost statistical significance at the conventional levels and become smaller and in one case negative. In cases where the IV coefficients for relative income have become negative, the positive coefficients on own income have increased in the IV compared to the OLS. This again is consistent with opposite directions of bias in the coefficients for per capita household income and for subjective relative income. Also, unlike for the OLS regressions, estimated optimism/pessimism is

¹⁷ If our goal is to focus on whether true relative income in comparison with reference groups affects health rather than perceptions of relative income, then instrumenting both own and relative incomes may be more appropriate.

no longer statistically significant. Overall, although most of the IV coefficient estimates for subjective relative income are not very precisely estimated, they suggest the importance of relatives and classmates as reference groups to health.— Throughout the different empirical specifications, we consistently find that relatives and classmates are the most salient comparison groups in impacting health outcomes, especially for urban and relatively educated individuals. Neighbors may be an important reference group for rural and relatively less educated individuals. Comparison with coworkers does not have a consistent or statistically significant impact on health.¹⁸

6. Conclusion

In this paper, we examine for the first time the importance of social reference groups other than those defined on the basis of geographic or demographic characteristics. Our methodology advances the previous literature by controlling for unobserved regional omitted factors and demonstrating how subjective relative income measurements can be used to test the importance of social reference groups. We propose a method for controlling for unobserved reporting biases affecting both subjective relative income measures and self-reported health, and also use instrumental variables to deal with the endogeneity of income measurements.

Our results indicate that for urban residents, relative income with respect to former classmates and relatives, and to a lesser extent coworkers are important social reference groups in China.— This demonstrates the importance of examining non-geographical reference groups in testing the relative deprivation hypothesis. We

¹⁸ The standard errors for the coefficients on relative living standards are too large in the preferred IV specifications to prove that the differences in magnitudes for different reference groups are statistically different than zero.

also find that the relative importance of different social reference groups may differ across individuals. In contrast to urban residents, geographic reference groups do appear to be salient for rural residents. These findings suggest that future research on the importance of more salient social reference groups to health may hold great promise for improving understanding of how relative deprivation affects individual health outcomes. Put another way, if previous studies focused on reference groups that were less salient to individuals, they may underestimate the actual importance of relative deprivation. We recommend that future work on this topic extend the current study in two directions. The first is to examine impacts on physical health measurements or more objective self-reported health outcomes, such as health expenditures, activities of daily life, or diagnosis of major diseases, which can reduce the likelihood of reporting biases affecting the results. The second is to study in greater depth the factors which explain variation in subjective relative income measurements, in order to control better for possible reporting biases.

References

Benabou, Roland. 1996. "Equity and Efficiency in Human Capital Investment: the Local Connection." *Review of Economic Studies* 63 (2): 237-64.

Conway, Terry, Ross Vickers, Harold Ward, Richard Rahe. 1981. "Occupational Stress and Variation in Cigarette, Coffee, and Alcohol Consumption." *Journal of Health and Social Behavior* 22 (2): 155-165.

Daly, Mary C., Daniel J. Wilson, & Norman J. Johnson. 2007. "Relative status and well-being: evidence from U.S. suicide deaths." Working Paper Series 2007-12, Federal Reserve Bank of San Francisco.

Davey Smith, G., M. Shipley, and G. Rose. 1990. "The Magnitude and Causes of Socio-Economic Differentials in Mortality: Further Evidence from the Whitehall Study." *Journal of Epidemiology and Community Health* 44 (4): 265-270.

Deaton, Angus and Christina Paxson. 2001. "Mortality, Education, Income, and Inequality among American Cohorts," in D. Wise (ed.), *Themes in the Economics of Aging*, Chicago: University of Chicago Press.

Deaton, Angus. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature* 41(1): 113-158.

Deaton, Angus. 2002. "Policy Implications of the Gradient of Health and Wealth." *Health Affairs* 21 (2): 13-30.

Der, D., S. Macintyre, G. Ford, K. Hunt, and P. West. 1999. "The Relationship of Household Income to a Range of Health Measures in Three Age Cohorts from the West Scotland," *European Journal of Public Health*. 9 (4): 271-277

Easterlin, Richard. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27: 35-48.

Eibner, Christine E., and William N. Evans. 2005. "Relative Deprivation, Poor Health Habits and Mortality." *Journal of Human Resources* 40(3):591-620.

Ferrer-i-Carbonell, Ada. "Income and well-being: an empirical analysis of the comparison income effect." *Journal of Public Economics*, June 2005, 89(5-6), pp. 997-1019.

Gerdtham, Ulf-G., and Magnus Johannesson. 2004. "Absolute Income, Relative Income, Income Inequality, and Mortality." *Journal of Human Resources*, 39 (1): 228-47.

Gorman, Dennis. 1988. "Employment, Stressful Life Events and the Development of Alcohol Dependence." *Drug and Alcohol Dependence* 22 (1-2): 151-159

Horwitz, Allan, and Lorraine Davies. 1994. "Are Emotional Distress and Alcohol Problems Differential Outcomes to Stress? An Exploratory Test," *Social Science Quarterly*

75 (3): 607-21.

House, James, Karl Landis, and Debra Umberson. 1999. "Social Relationships and Health." In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Jensen, Robert, and Kaspar Richter. 2004. "The health implications of social security failure: evidence from the Russian pension crisis." *Journal of Public Economics* 88(1-2): 209-236.

Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur A. Stone. 2006. "Would You Be Happier If You Were Richer? A Focusing Illusion," *Science* 312 (5782): 1908 – 1910.

Kawachi, Ichiro, and Bruce Kennedy. 1997. "The Relationship of Income Inequality to Mortality: Does the Choice of Indicator Matter?" *Social Science and Medicine* 45 (7): 1121-27.

Kennedy, Bruce, I. Kawachi, R. Glass, and D. Prothrow-Stith. 1998. "Income Distribution, Socioeconomic Status, and Self Rated Health in the United States: Multilevel Analysis." *British Medical Journal* 317 (7163): 917-21.

Kristenson, Margareta, Kristina Orth-Gomer, Zita Kucinskiene, Bjorn Bergdahl, Henrikas Calkauskas, Irena Balinkyniene, and Anders Olsson. 1999. "Attenuated

Cortisol Response to a Standardized Stress Test in Lithuanian versus Swedish Men.” In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Lovallo, William. 1997. *Stress and Health*. Sage Publications: London.

Luttmer, Erzo F.P. “Neighbors as Negatives: Relative Earnings and Well-Being.” *Quarterly Journal of Economics*, 2005, 102(3), pp. 963-1002.

Lynch, John, and George Kaplan. 1999. “Understanding How Inequality in the Distribution of Income Affects Health.” In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Marmot, Michael. George Davey Smith, Stephen Stansfeld, Chadra Patel, Fiona North, J. Head, Ian White, Eric Brunner, and Amanda Feeny. 1991. “Health Inequalities among British Civil Servants: the Whitehall II Study.” *Lancet* 337 (8754): 1387-93.

McBride, Michael, “Relative-Income Effects on Subjective Well-Being in the Cross-Section,” *Journal of Economic Behavior and Organization*, XLV (2001), 251–278.

Mellor, Jennifer, and Jeffrey Milyo. 2002. “Income Inequality and Health Status in the United States: Evidence from the Current Population Survey.” *Journal of Human Resources* 37 (3): 510-39.

Miller, Douglas, and Christina Paxson. "Relative Income, Race, and Mortality." *Journal of Health Economics*, 2006, 25(5), pp. 979-1003.

Shively, Carol, and Thomas Clarkson. 1999. "Social Status and Coronary Artery Atherosclerosis in Female Monkeys." In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Shively, Carol, Kathy Laber-Laird, and Raymond Anton. 1999. "Behavior and Physiology of Social Stress and Depression in Female Cynomolgus Monkeys." In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Sapolsky, Robert. 1998. *Why Zebras Don't Get Ulcers*. W.H. Freeman and Company: New York.

Sapolsky, Robert, Susan Alberts, and Jeanne Altmann. 1999. "Hypercortisolism Associated with Social Subordination or Social Isolation among Wild Baboons." In *The Society and Population Health Reader: Income Inequality and Health*, ed. Kawachi, Ichiro, Bruce Kennedy, and Richard Wilkinson. The New Press: New York.

Wilkinson, Richard. 1996. *Unhealthy Societies: the Afflictions of Inequality*. London,

Routledge.

World Bank. 2009. *From poor areas to poor people: China's evolving poverty reduction agenda*.

Washington, D.C.: The World Bank.

Figure 1: Self-Reported Health and Psychosocial Health, by Relative Income Status with Respect to Different Reference Groups

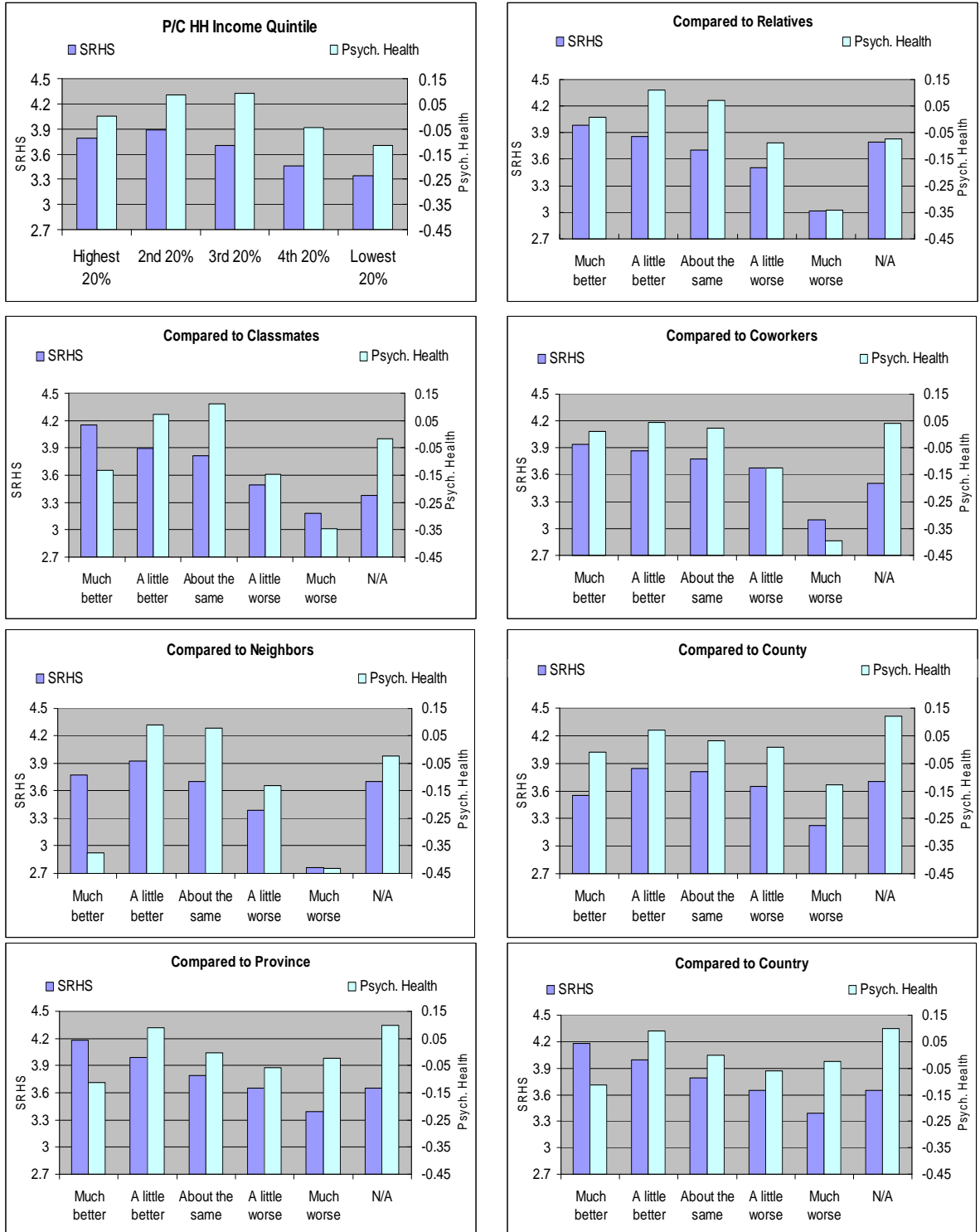


Table 1. Descriptive statistics (N=2891)

	Mean	Std. Dev.	Min	Max
Self reported health status (v. good=5, v.poor=1)	3.631	1.034	1.0	5.0
Self-reported health (shares): Very good	0.228			
Good	0.342			
OK	0.288			
Poor	0.118			
Very poor	0.024			
Psychosocial health (N=2895)	0.0028	0.6324	-2.1199	1.1949
Per-capita hh income (RMB)	4,061.07	10,153.9	137.5	375,000.0
Mean per-capita hh inc within TOWNSHIP (RMB)	4,692.9	4,754.7	328.5	30,381.6
Mean per-capita hh inc within COUNTY (RMB)	4,530.5	4,294.6	382.9	30,381.6
Mean per-capita hh inc within PROVINCE (RMB)	4,091.6	2,464.4	623.0	11,097.1
Male (dummy)	0.517	0.500	0.000	1.000
Age	38.6	13.2	18.0	70.0
Marital status (shares): Married	0.798			
Never married	0.161			
Other marital status	0.041			
Education level (shares): No schooling	0.251			
Primary school	0.155			
Middle school	0.338			
High school	0.160			
College and above	0.096			
Urban (dummy)	0.398	0.490	0.000	1.000
Agricultural household registration (dummy)	0.703	0.457	0.000	1.000
Major income from agriculture (dummy)	0.590	0.492	0.000	1.000
Interviewer assessment of respondents economic status in village (shares): Low income	0.192			
Average income	0.577			
Upper-middle income	0.204			
High income	0.027			
Interviewer assessment of respondents housing (shares): Poor	0.121			
Middle	0.697			
High	0.182			
Own assets (dummies): Motor cycle	0.352	0.478	0.000	1.000
Car	0.046	0.209	0.000	1.000
Refrigerator	0.343	0.475	0.000	1.000
Color TV	0.779	0.415	0.000	1.000
Computer	0.112	0.316	0.000	1.000
Telephone/cellphone	0.637	0.481	0.000	1.000
Washing machine	0.458	0.498	0.000	1.000
Household size	4.1	1.5	1.0	14.0

Table 2: OLS with objective measures of relative income

Dependent variable	(1) SRHS	(2) Psych.	(3) SRHS	(4) Psych.
Log per-capita hh income	0.113** (0.043)	0.098** (0.039)	0.095** (0.039)	0.084** (0.034)
Mean log per-capita hh inc within TOWNSHIP	-0.041 (0.071)	-0.102** (0.049)		
Mean log per-capita hh inc within PROVINCE			0.022 (0.085)	-0.100 (0.062)
Male (dummy)	0.156*** (0.057)	0.158*** (0.050)	0.163*** (0.054)	0.164*** (0.048)
Age/10	-0.284* (0.156)	-0.191* (0.104)	-0.298* (0.160)	-0.202* (0.106)
(Age/10)^2	0.012 (0.017)	0.018* (0.011)	0.013 (0.018)	0.019* (0.011)
Married==Never married (dummy)	-0.076 (0.101)	-0.164*** (0.050)	-0.087 (0.099)	-0.173*** (0.049)
Married==Other marital status (dummy)	0.066 (0.127)	-0.153 (0.110)	0.051 (0.130)	-0.152 (0.111)
Married==No response (dummy)	0.105 (0.498)	0.320 (0.300)	0.115 (0.494)	0.383 (0.301)
Degree==Primary (dummy)	0.239** (0.094)	0.135** (0.055)	0.226** (0.091)	0.138** (0.056)
Degree==Middle school (dummy)	0.383*** (0.087)	0.147*** (0.055)	0.363*** (0.087)	0.144** (0.058)
Degree==High school (dummy)	0.355*** (0.129)	0.048 (0.068)	0.330** (0.128)	0.035 (0.067)
Degree==Tertiary (dummy)	0.309 (0.204)	0.031 (0.103)	0.280 (0.194)	-0.000 (0.095)
Degree==No response (dummy)	-0.339 (0.305)	0.014 (0.091)	-0.372 (0.308)	0.012 (0.102)
Urban resident (dummy)	-0.109 (0.096)	-0.093 (0.075)	-0.136 (0.087)	-0.123* (0.069)
Bottom coding p/c hh income (dummy)	-0.102 (0.289)	-0.055 (0.144)	-0.112 (0.289)	-0.066 (0.148)
Constant	3.696*** (0.617)	0.373 (0.385)	3.419*** (0.705)	0.507 (0.476)
Observations	2891	2895	2891	2895
Adj R Square	0.152	0.064	0.151	0.064

Notes:

1) Omitted or reference categories are married for marital status and less than primary schooling for education.

2) Standard errors are clustered by township.

3) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: OLS with subjective measures of relative income

Dependent variable	SRHS	Psych.	SRHS	Psych.	SRHS	Psych.
Township fixed effects?	No	No	Yes	Yes	Yes	Yes
(1) Log p/c hh income	0.062* (0.037)	0.039 (0.031)	0.077* (0.043)	0.079** (0.038)	0.113** (0.047)	0.098** (0.041)
RELATIVES	0.169*** (0.035)	0.114*** (0.027)	0.140*** (0.034)	0.102*** (0.022)	0.112*** (0.033)	0.087*** (0.023)
Est. optimism					0.107*** (0.040)	0.057** (0.023)
(2) Log p/c hh income	0.077** (0.038)	0.049 (0.032)	0.086* (0.044)	0.086** (0.041)	0.118** (0.050)	0.105** (0.046)
CLASSMATES	0.187*** (0.052)	0.128** (0.051)	0.177*** (0.050)	0.110** (0.042)	0.149** (0.057)	0.093* (0.048)
Est. optimism					0.104** (0.043)	0.061** (0.028)
(3) Log p/c hh income	0.085** (0.040)	0.063* (0.032)	0.095** (0.044)	0.089** (0.037)	0.139*** (0.048)	0.113*** (0.037)
COWORKERS	0.097** (0.037)	0.082*** (0.030)	0.079** (0.034)	0.076** (0.034)	0.034 (0.033)	0.051 (0.034)
Est. optimism					0.135*** (0.042)	0.074*** (0.024)
(4) Log p/c hh income	0.058* (0.035)	0.043 (0.028)	0.069* (0.039)	0.078** (0.033)	0.106*** (0.039)	0.100*** (0.031)
NEIGHBORS	0.188*** (0.055)	0.096** (0.040)	0.150*** (0.056)	0.087** (0.041)	0.118** (0.054)	0.067 (0.044)
Est. optimism					0.103*** (0.036)	0.061** (0.025)
(5) Log p/c hh income	0.070* (0.036)	0.051* (0.030)				
COUNTY	0.137*** (0.038)	0.060** (0.025)				
Est. optimism						
(6) Log p/c hh income	0.086** (0.034)	0.068** (0.027)				
PROVINCE	0.062 (0.051)	-0.015 (0.039)				
Est. optimism						

Table 3: (continued)

Dependent variable	SRHS	Psycho.	SRHS	Psycho.	SRHS	Psycho.
Township fixed effects?	No	No	Yes	Yes	Yes	Yes
(7) Log p/c hh income	0.038 (0.036)	0.045 (0.027)	0.053 (0.040)	0.066* (0.033)	0.083* (0.042)	0.082** (0.034)
RELATIVES	0.084** (0.037)	0.071** (0.028)	0.073* (0.037)	0.064*** (0.024)	0.064* (0.036)	0.059** (0.024)
CLASSMATES	0.130* (0.067)	0.104* (0.062)	0.134** (0.066)	0.073 (0.061)	0.123* (0.069)	0.068 (0.062)
COWORKERS	-0.068 (0.048)	-0.005 (0.037)	-0.057 (0.043)	-0.004 (0.044)	-0.067 (0.042)	-0.009 (0.044)
NEIGHBORS	0.127** (0.055)	0.051 (0.042)	0.108* (0.060)	0.047 (0.044)	0.091 (0.057)	0.038 (0.045)
COUNTY	0.092** (0.043)	0.063 (0.040)				
PROVINCE	-0.061 (0.061)	-0.101* (0.052)				
Est. optimism					0.078** (0.037)	0.040 (0.026)

1) The sample sizes are 2,891 when the dependent variable is SRHS and 2,895 when the dependent variable is psychosocial health.

2) All regressions include as covariates the same set of non-income variables reported in Table 2.

3) Standard errors are clustered by township.

4) * significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: OLS with subjective relative income measures, by sub-groups

Dependent variable	SRHS	SRHS	Psych.	Psych.
Sub-group	Rural	Urban	Rural	Urban
(1) Log per-capita hh income	0.169*** (0.045)	0.091 (0.066)	0.169*** (0.033)	0.044 (0.042)
RELATIVES	0.088* (0.044)	0.129*** (0.045)	0.058* (0.031)	0.121*** (0.029)
Estimated optimism	0.155** (0.060)	0.046 (0.044)	0.056 (0.036)	0.063*** (0.021)
(2) Log per-capita hh income	0.191*** (0.048)	0.067 (0.058)	0.181*** (0.034)	0.035 (0.051)
CLASSMATES	0.051 (0.047)	0.243*** (0.064)	0.028 (0.033)	0.166** (0.068)
Estimated optimism	0.173*** (0.064)	0.010 (0.039)	0.071* (0.038)	0.051* (0.030)
(3) Log per-capita hh income	0.189*** (0.048)	0.114 (0.070)	0.177*** (0.032)	0.062 (0.041)
COWORKERS	0.051 (0.058)	0.057 (0.049)	0.057 (0.045)	0.063 (0.040)
Estimated optimism	0.181*** (0.065)	0.068 (0.043)	0.067* (0.035)	0.083*** (0.023)
(4) Log per-capita hh income	0.127*** (0.043)	0.135** (0.058)	0.146*** (0.032)	0.079** (0.033)
NEIGHBORS	0.208*** (0.051)	-0.028 (0.083)	0.119*** (0.027)	-0.008 (0.079)
Estimated optimism	0.117** (0.057)	0.091** (0.034)	0.034 (0.035)	0.102*** (0.028)

1) The sample size for rural areas is 1,377 when the dependent variable is SRHS and 1,379 when the dependent variable is psychosocial health. The sample size for urban areas is 1,514 when the dependent variable is SRHS and 1,516 when the dependent variable is psychosocial health.

2) All regressions include as covariates the same set of non-income variables reported in Table 2. All regressions control for township fixed effects.

3) Standard errors are clustered by township.

4) * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: 2 SLS results instrumenting for per-capita log household income

Dependent variable	SRHS	Psycho.
(1) Log per-capita household income	0.175 (0.129)	0.162** (0.074)
Mean log per-capita hh income within TOWNSHIP	-0.085 (0.135)	-0.150* (0.080)
Estimated optimism	0.140** (0.054)	0.084** (0.034)
1 st -stage F-statistic	20.19	19.74
Over-id test on exogeneity of instruments (p-value)	0.25	0.28
(2) Log per-capita hh income	0.135 (0.087)	0.129** (0.053)
Mean log per-capita hh income within PROVINCE	-0.003 (0.113)	-0.127 (0.078)
Estimated optimism	0.129*** (0.049)	0.078** (0.033)
F stat on own income in the 1 st stage	37.68	36.75
Over-id test on exogeneity of instruments (p-value)	0.27	0.25
(3) Log per-capita hh income	0.085 (0.136)	0.180** (0.085)
Living standard compared to RELATIVES (1: much worse to 5: much better)	0.120** (0.053)	0.065* (0.034)
Estimated optimism	0.097 (0.063)	0.087** (0.036)
F stat on own income in the 1 st stage	18.31	17.76
Over-id test on exogeneity of instruments (p-value)	0.19	0.19
(4) Log per-capita hh income	0.101 (0.135)	0.189** (0.090)
Living standard compared to CLASSMATES	0.153** (0.072)	0.074 (0.059)
Estimated optimism	0.098 (0.064)	0.090** (0.040)
F stat on own income in the 1 st stage	17.28	16.98
Over-id test on exogeneity of instruments (p-value)	0.17	0.17

Table 5: (continued)

Dependent variable	SRHS	Psycho.
(5) Log per-capita hh income	0.139 (0.131)	0.207*** (0.075)
Living standard compared to COWORKERS	0.034 (0.044)	0.020 (0.037)
Estimated optimism	0.135** (0.060)	0.107*** (0.033)
F stat on own income in the 1 st stage	23.59	22.97
Over-id test on exogeneity of instruments (p-value)	0.16	0.18
(6) Log per-capita hh income	0.080 (0.125)	0.205*** (0.069)
Living standard compared to NEIGHBORS	0.127** (0.057)	0.028 (0.052)
Estimated optimism	0.093* (0.055)	0.102*** (0.038)
F stat on own income in the 1 st stage	23.60	23.43
Over-id test on exogeneity of instruments (p-value)	0.21	0.19

Notes:

1) The sample sizes are 2,891 when the dependent variable is SRHS and 2,895 when the dependent variable is psychosocial health.

2) All regressions include as covariates the same set of non-income variables reported in Table 2. In Panels (3) through (6), all regressions control for township fixed effects.

3) Standard errors are clustered by township.

4) Instruments are 7 indicators of asset ownership: motorcycle, car, refrigerator, color TV, computer, telephone/cell phone, and washing machine.

5)* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: 2 SLS results instrumenting for per-capita log household income, by sub-groups

Dependent variable	SRHS	SRHS	Psych.	Psych.
Sub-group	Rural	Urban	Rural	Urban
(1) Log per-capita hh income	0.457*** (0.121)	0.012 (0.122)	0.261*** (0.080)	0.099 (0.087)
Mean log per-capita hh income within TOWNSHIP	-0.290* (0.150)	0.052 (0.129)	-0.244*** (0.090)	-0.058 (0.099)
Estimated optimism	0.246*** (0.072)	0.046 (0.049)	0.074* (0.045)	0.109** (0.043)
1 st -stage F-statistic	10.97	17.31	10.69	16.92
Over-id test on exogeneity of instruments (p-value)	0.61	0.33	0.45	0.34
(2) Log per-capita hh income	0.328*** (0.087)	0.042 (0.101)	0.187*** (0.062)	0.093 (0.073)
Mean log per-capita hh income within PROVINCE	-0.101 (0.110)	0.081 (0.170)	-0.178** (0.089)	-0.053 (0.108)
Estimated optimism	0.214*** (0.071)	0.056 (0.038)	0.060 (0.042)	0.108** (0.044)
F stat on own income in the 1 st stage	14.63	48.28	14.19	46.87
Over-id test on exogeneity of instruments (p-value)	0.70	0.34	0.42	0.37
(3) Log per-capita hh income	0.383** (0.151)	-0.094 (0.117)	0.343*** (0.089)	0.094 (0.091)
RELATIVES	0.035 (0.063)	0.185*** (0.065)	0.015 (0.043)	0.106** (0.050)
Estimated optimism	0.229*** (0.077)	-0.027 (0.066)	0.116*** (0.042)	0.083* (0.048)
1 st -stage F-statistic	11.39	13.08	11.01	12.71
Over-id test on exogeneity of instruments (p-value)	0.30	0.32	0.47	0.24
(4) Log per-capita hh income	0.430*** (0.134)	-0.090 (0.111)	0.354*** (0.077)	0.094 (0.090)
CLASSMATES	0.003 (0.053)	0.291*** (0.075)	-0.007 (0.038)	0.148* (0.081)
Estimated optimism	0.250*** (0.074)	-0.053 (0.056)	0.127*** (0.040)	0.074 (0.053)
1 st -stage F-statistic	11.29	18.19	10.98	18.03
Over-id test on exogeneity of instruments (p-value)	0.28	0.38	0.45	0.24
(5) Log per-capita hh income	0.410*** (0.129)	-0.050 (0.123)	0.345*** (0.077)	0.126 (0.080)
COWORKERS	-0.023 (0.069)	0.111* (0.057)	0.000 (0.052)	0.042 (0.051)
Estimated optimism	0.253***	0.003	0.122***	0.109**

	(0.071)	(0.060)	(0.037)	(0.049)
1 st -stage F-statistic	14.33	17.71	14.01	17.14
Over-id test on exogeneity of instruments (p-value)	0.28	0.33	0.46	0.22
(6) Log per-capita hh income	0.287*	0.002	0.289***	0.173**
	(0.154)	(0.118)	(0.087)	(0.084)
NEIGHBORS	0.157**	0.030	0.072*	-0.049
	(0.076)	(0.084)	(0.043)	(0.097)
Estimated optimism	0.175**	0.037	0.087**	0.141**
	(0.075)	(0.050)	(0.042)	(0.060)
1 st -stage F-statistic	10.42	23.69	10.24	23.30
Over-id test on exogeneity of instruments (p-value)	0.31	0.33	0.56	0.17

1) The sample size for rural areas is 1,377 when the dependent variable is SRHS and 1,379 when the dependent variable is psychosocial health. The sample size for urban areas is 1,514 when the dependent variable is SRHS and 1,516 when the dependent variable is psychosocial health.

2) All regressions include as covariates the same set of non-income variables reported in Table 2. In Panels (3) through (6), all regressions control for township fixed effects.

3) Standard errors are clustered by township.

4) Instruments are 7 indicators of asset ownership: motorcycle, car, refrigerator, color TV, computer, telephone/cell phone, and washing machine.

5) * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix Table 1: Psychosocial health index questions (N=2,895)

	Never	Rarely	Sometimes	Often	No response
Worry about small things	826 28.5%	763 26.3%	908 31.4%	374 12.9%	25 0.9%
No appetite for food	1018 35.2%	904 31.2%	738 25.5%	223 7.7%	13 0.4%
Cannot focus attention	908 31.4%	1021 35.3%	727 25.1%	171 5.9%	68 2.4%
My life is a failure	1339 46.3%	678 23.4%	507 17.5%	183 6.3%	189 6.5%
Quality of sleep is poor	961 33.2%	794 27.4%	792 27.4%	336 11.6%	12 0.4%
I feel fortunate	216 7.5%	447 15.4%	1048 36.2%	1119 38.7%	65 2.2%
I feel alone	1172 40.5%	863 29.8%	590 20.4%	171 5.9%	100 3.4%
My life is very happy	169 5.8%	368 12.7%	1013 35.0%	1282 44.3%	63 2.2%

Appendix Table 2: Subjective comparisons of own living standards

Comparison group	Much better	A little better	About the same	A little worse	Much worse	N
Relatives	45 1.57%	380 13.29%	1,502 52.52%	707 24.72%	226 7.90%	2,860
Classmates	37 1.65%	265 11.83%	1,255 56.00%	511 22.80%	173 7.72%	2,241
Coworkers	26 1.62%	197 12.26%	970 60.36%	330 20.54%	84 5.23%	1,607
Neighbors	50 1.80%	437 15.77%	1,568 56.59%	576 20.79%	140 5.05%	2,771
County	8 0.30%	208 7.78%	846 31.66%	1,059 39.63%	551 20.62%	2,672
Province	19 0.76%	212 8.47%	556 22.20%	831 33.19%	886 35.38%	2,504
Country	26 1.11%	182 7.75%	643 27.40%	627 26.71%	869 37.03%	2,347

Appendix Table 3: 2SLS results instrumenting for both log per-capita household income and subjective relative income

Dependent variable	SRHS	Psych.
(1) Log per-capita hh income	-0.182 (0.184)	0.084 (0.104)
Living standard compared to RELATIVES	0.630** (0.308)	0.246 (0.197)
Estimated optimism	-0.094 (0.134)	0.019 (0.079)
F stat on household income in the 1 st stage	23.60	23.11
F stat on relative income in the 1 st stage	7.33	7.34
Over-id test on instruments (p-value)	0.30	0.18
(2) Log per-capita hh income	0.009 (0.182)	0.044 (0.108)
Living standard compared to CLASSMATES	0.416 (0.452)	0.485* (0.290)
Estimated optimism	0.023 (0.133)	-0.027 (0.081)
F stat on household income in the 1 st stage	23.26	22.77
F stat on relative income in the 1 st stage	13.98	13.62
Over-id test on instruments (p-value)	0.18	0.42
(3) Log per-capita hh income	0.015 (0.271)	0.040 (0.221)
Living standard compared to COWORKERS	0.389 (0.831)	0.491 (0.740)
Estimated optimism	0.050 (0.194)	-0.007 (0.176)
F stat on household income in the 1 st stage	23.25	22.73
F stat on relative income in the 1 st stage	6.46	6.31
Over-id test on instruments (p-value)	0.18	0.40
(4) Log per-capita hh income	0.205 (0.209)	0.325* (0.166)
Living standard compared to NEIGHBORS	-0.074 (0.335)	-0.166 (0.322)
Estimated optimism	0.172 (0.133)	0.180 (0.128)
F stat on household income in the 1 st stage	23.57	23.05
F stat on relative income in the 1 st stage	11.13	10.93
Over-id test on instruments (p-value)	0.09*	0.08*

Notes:

- 1) The sample sizes are 2,891 when the dependent variable is SRHS and 2,895 when the dependent variable is psychosocial health.
- 2) All regressions include as covariates the same set of non-income variables reported in Table 2. All regressions control for township fixed effects.
- 3) Standard errors are clustered by township.
- 4) Instruments are 7 indicators of asset ownership: motorcycle, car, refrigerator, color TV, computer, telephone/cell phone, and washing machine.
- 5)* significant at 10%; ** significant at 5%; *** significant at 1%.