Socio-Economic Determinants of Health in Rural Pakistan: Relative or Absolute Standards?

Khadija Shams

Abstract

This paper addresses key aspects of health inequality. We analyse in par- ticular to what extent income determines household-specific health outcomes in rural Pakistan using our survey data. Controlling for various socio-economic characteristics, we investigate validity of the three income-health hypotheses: the Absolute Income Hypothesis (AIH), the Relative Income Hypothesis (RIH) and the Income Inequality Hypothesis (IIH). Whilst these hypotheses crucially differ in their exact substance, broadly speaking, those refer to the idea that a household's health status might be linked to the existing socio-economic environment. Households with a more favourable income position (either in absolute, relative or distributional terms) might enjoy a better health sta- tus. We employ a general empirical specification that nests different health functions as special cases. This permits testing the income-health hypotheses separately and jointly. We find that in rural Pakistan both the relative in- come (with respect to the relevant community) and absolute income are major determinants of health. This is in contrast to results typically reported for developed countries, where in particular the household's absolute income posi- tion appears to matter. The study provides important insights into the causes of health inequalities. For instance, higher income improves health directly because of higher social support and other psychosocial reasons. However, we failed to confirm IIH on pure statistical grounds.

JEL Classification: I11; I12; I13; I14; I18; P46.

Key Words: Health measurement; health inequality; health services; socio-economic indicators; rural Pakistan.

1. Introduction

It has become increasingly acknowledged in both social and health sciences that a range of socio-economic factors contribute to inequalities in health. If this was the case, individual health outcomes would more and more hinge on the household's wealth and income status rather than purely immaterial determinants such as ge- netics or luck. In this paper we aim to investigate the link between income and health in rural Pakistan. Against the backdrop of a relatively underdeveloped pub- lic health system there, we seek to analyse in particular several key income-health hypotheses that are commonly considered in the health economics literature. We do so using econometric modelling strategies and by comparing our results with other both developed and emerging economies.

The relationship between health and income tends to be framed in terms of three related hypotheses. The Absolute Income Hypothesis (AIH) states that a household's health depends on its own level of income, independent of the financial situation of its peers. In other words, the AIH suggests that the higher a household's income, the lower the risk of being unhealthy. The Relative Income Hypothesis (RIH), on the other hand, claims that at any given level of income, the household's health depends on its income in relation to society. This implies that a higher social status provides increasing psychosocial satisfaction and thus ensures better health. For the RIH the distribution of income matters such that living in an environment characterized by a relatively unequal income distribution is hypothesized to cause psycho-social stress, leading to a worsening of health outcomes. This is captured by the related Income Inequality Hypothesis (IIH). One may argue that living in a place with an unequal income distribution leads to social or ethnic inequalities such that relatively deprived people do not have access to a proper and efficient health care system (which might explain social and ethnic inequalities in health), and are prone to get sick more frequently.

∗I would like to thank Luis Angeles, Sai Ding, Richard Harris, Alexander Kadow, Philippe LeMay-Boucher and seminar participants at the University of Glasgow for helpful comments.

†University of Glasgow and Shaheed Benazhir Bhuto Women University, Department of Economics, Adam Smith Building, Glasgow, G12 8RT, United Kingdom; Tel: +44 (0)141-330 2592; Fax: +44 (0)141-330 4940; k.shams.1@research.gla.ac.uk

Relative income as well as the income distribution are often labelled as "psychosocial determinants" of health, even though they are triggered by material standards such as income. From a policy maker's perspective, it is therefore essential to disentangle the empirical importance of each of these hypotheses. Does relative income dominate over absolute one?

In which form does the distribution of income matter at all? Different policy recommendations would obviously arise depending on the estimation outcomes. We employ ordered probit regression to justify our results. The majority of the studies for developed countries thus far tends to provide evidence in favour of the AIH in comparison to the RIH with only a few exceptions. Results on the validity of the IIH are mixed, however. Studies based on US popu- lation data, overall, tend to favour the AIH, rejecting the RIH and providing little or no support for the IIH (Wagstaff and Doorslaer, 2000). Similarly, Lindley and Lorgelly (2005) used data for the UK provided by the British Household Panel Sur- vey (BHPS) to test for and to distinguish between the AIH vis-à-vis the RIH. The longitudinal nature of their data moreover allowed them to investigate the dynamic evolution of the given hypotheses. They provide strong evidence to support the AIH, and found that the RIH does not seem to hold over time within the UK. Gerdtham and Johannesson (2004) analysed Swedish panel data comprising more that 40,000 adults who were followed over 10-17 years. Whilst their results are consistent with the AIH, they fail to confirm either the RIH or IIH. Marmot et al. (1991), on the other hand, analysing health inequalities among British Civil Servants (Whitehall II study) found that in lower administrative ranks, it is indeed relative rather than absolute income that is important for health such that a low relative income has a detrimental effect on the health status. Lobmayer and Wilkinson (2000) conducted a multi-country study based on the Luxembourg Income Study (LIS).1 Overall, the results cannot confirm the idea of more egalitarian societies reporting better health suggesting that the IIH does not hold.

Amid the inconclusive evidence on the IIH, Mellor and Milyo (2002) suggest investigating this postulated health-income relationship from two angles (i.e. in the strong or weak form). The strong version of the IIH implies that inequality adversely affects all members in a society equally, regardless of their financial status. The weak version states that income inequality is harmful to the health of only the least well off in a society. Mellor and Milyo (2002)'s analysis is based on the US Current Population Survey (CPS) data from 1995-1999 and examines the effect of income inequality on household's health status for both the general households and those households living in poverty. However, they fail to establish consistent evidence of the IIH in either form. On the other hand, Li and Zhu (2006) tested the two versions of the IIH for China. They used China Health and Nutrition Survey (CHNS) data for five years (1989, 1991, 1993, 1997 and 2000). Overall, their results provide evidence supporting the strong version but failed to confirm the IIH in its weak form.

Research has shown that micro data rather than macro data appears to be more appropriate to discriminate between the competing income-health hypotheses. For instance, Smith (1999) using aggregate data for international comparisons found that a distinction between the effects of (absolute and relative) income and income inequality on health can hardly be drawn. This is due to a potential concave as- sociation between income and individual health. This form of non-linearity in the income-health relationship causes income inequality and health to move in opposite directions on an aggregate level. Wagstaff and Doorslaer (2000), therefore, favour the use of micro data over macro data as well.

Our study is based on micro survey data and comprises 600 representative house- holds of rural Pakistan. We test the three income-health hypotheses (AIH, RIH, IIH), while controlling for individual socio-economic characteristics of the households like gender, age, education, marital status and family size. There is a considerable liter- ature investigating these hypotheses using individual-level data. Those studies are typically based on self-assessed health, infant health or mortality rates as proxies of health outcomes. In particular, Meara (1999) focused on infant health (i.e. low birth weight), Mellor and Milyo (2002) used the self-reported health status and Gerdtham and Johannesson (2004) considered mortality as relevant health measure. We use the household's self-assessed health status (as health variable) made at the house- hold level during the given year. Our study is particularly in line with the paper by Gerdtham and Johannesson (2004), who were the first to explicitly discriminate between the three health-income hypotheses in a coherent setting.

----

[1] Their analysis covered the following 14 OECD countries in alphabetical order: Australia, Bel- gium, Canada, Denmark, Finland, France, (former West) Germany, Italy, the Netherlands, Norway, Spain, Sweden, the UK, the US; excluding Luxembourg given its relatively small population size

Further studies to test the hypotheses jointly are Fiscella and Franks (1997); Daly et al. (1998); Meara (1999) and Mellor and Milyo (2002). However, most of them tend to focus on the IIH without in some cases even reporting results for RIH. As an important methodological contribution to the analysis, we consider it is important to realise the links and associations among the different health-income hypothesis. We therefore strive to set up a framework that is general enough to permit testing all three key hypotheses in a coherent setting.

Despite these unique strengths (micro data, simultaneous testing set-up, robust- ness checks), our analysis and dataset also carry some limitations which need to be mentioned. First, it should be noted that Pakistan is an agrarian economy with 70% of its population living in rural areas that are engaged in subsistence agriculture. Income is as a result more equally distributed because most workers have very low levels of income. In other words, incomes are concentrated at low levels and that concentration dominates the overall distribution of income. This may make the rural side of Pakistan not an ideal laboratory to test the IIH. However, it may be the case that apart from the level of income inequality, the variations in income inequality across geographical regions also matter. In such a case, there are sufficient variations in income inequality across the sample districts. That is, 97% of the total income inequality in rural Pakistan is explained by inequality between the districts. The remaining 3% are explained by inequality within the districts.2 A further limitation is that we have measured relative income and income inequality only at the district level rather than considering the sample as a whole. It might be the case, how- ever, that it is important to test the given hypotheses at the country level as well

– which is not supported by the available data.3 Furthermore, we assume that the association between health and socio-economic status (SES), whether measured by education, gender, or income is largely due to the effects of SES on health, not vice versa following (Doornbos and Kromhout, 1990; Fox et al., 1985; Power et al., 1990; Wilkinson, 1986). This one-way causality is in line with the most part of the health economics literature.

The remainder of the paper is structured as follows. We first provide in Section 2 general background information on the relationship between healthcare systems and their impact on health inequalities with special reference to Pakistan to motivate our health measure. Section 3 summarises the data and describes the empirical specification. Section 4 presents and discusses our empirical results. The paper ends with concluding remarks and policy recommendations in Section 5.

2. Inequalities in Health and the Healthcare System

Health depends on a number of factors, including biological factors, environmental factors, nutrition, and the standard of living. In other words, health can be seen as a function of welfare. Few of the issues which cause ill health are dealt with directly by ‘health services’; they are, rather, issues in the ‘welfare state’ as a whole. When in the 19th century, Chadwick identified poor health as a major cause of pauperism, his response was to improve sanitation, not to introduce more extensive medical care. Most of the world’s diseases are attributable to poor water supply or nutrition.

‘Health services’ are thus better described as medical services.

There are clear differences in the incidence of ill health by social class. All studies cited above show a close link between health and social inequality. People in lower social classes, including children, are more likely to suffer from infective and parasitic diseases, pneumonia, poisonings or violence. Adults in lower social classes are more likely, in addition, to suffer from cancer, heart disease and respiratory disease. Lower class people have more time off work, pay more visits to the doctor and are likely to be chronically ill. As part of 1999 General Household Survey, ONS statisticians looked at 1,200 workless households, containing at least one person of working age. They found that 32% of the members of workless households reported chronic illnesses, compared with 12.5% of those in working household.[4]..

---

2 Source: Survey 2008.

3 Recall that our dataset considers cross-sections and focuses on rural Pakistan only.

4 The General Household Survey (GHS) is a survey conducted on an annual basis by the Office for National Statistics (ONS) and collects data about private households in Great Britain. The aim of this survey is to provide government departments and organisation with information on a range of topics concerning private households for monitoring and policy purposes

There are several possible explanations for these inequalities. For instance, poverty leads to ill health through nutrition, housing and environment. Another explanation may lie in cultural and behavioural aspects. There are, for example, differences in the diet and fitness of different social classes and in certain habits like smoking. There are moreover often major inequalities in access to health care accord- ing to social class. The problem becomes what Tudor Hart once called an 'inverse care law' implying that those individuals in the worst health condition receive the least services. The inverse care law proposed by Julian Tudor Hart in 1971 states that 'the availability of good medical care tends to vary inversely with the need for it in the population served'. The law explains the fact that poor people with chronic illness and diseases actually need and deserve good medical care but they cannot af- ford it and vice-versa. Apart from income, location, race/ethnicity and gender may also explain health inequalities in general.

Before proceeding with testing all the three income-health hypotheses in rural Pakistan, it may be helpful to obtain a better understanding of the existing health- care system in Pakistan. In Pakistan, the public health services and hospitals are relatively cheap compared to the private ones, but the standard they provide is not satisfactory. It involves a lot of administrative delays and poor health facilities. The public hospitals are very inefficient compared to the private clinics/hospitals in providing proper treatments to their patients and on proper time. This leads to prolonged and chronic sickness and diseases and the patients have to suffer rather strongly and for long. However, the majority of the people still go to these hospitals because they cannot afford the private doctors. Those who are relatively better-off financially, on the other hand, can visit the private doctors and receive the proper treatment within good time. Based on those features inherent in the health care sys- tem in Pakistan, it thus appears interesting to test the three income related health hypotheses, with special reference to rural Pakistan which represents 70% of the country's population (CIA, 2008)).

3. Data

3.1. Research Design

Our analysis is based on a survey of households in rural Pakistan conducted in 2008. The dataset comprises all four provinces of Pakistan: Punjab, Sind, the North Western Frontier Province (NWFP) and Baluchistan. To ensure representativeness, we decided to sample households in 10 districts (i.e. roughly one tenth of the total number of districts) across the country (stratified sampling). Based on population figures, we allocated the number of districts across the provinces as follows: four districts from Punjab, three from Sind, two from NWFP and one from Baluchistan. The selected districts in Punjab are Attock, Layyah, Rahimyarkhan and Sahiwal; Badin, Mirpurkhas and Thatta in Sind; Dir and Malakand in NWFP and Kalat from Baluchistan. Those districts were selected for various reasons. First, they are geographically in a range that offered easy access without raising security concerns for the interviewers compared to further remote areas. Second, these districts paint a representative picture of the socio-economic environment in rural Pakistan. Due to the geographic scope of the districts, great care has been taken, where necessary, in sampling households from villages which are reasonably far away from major cities such as Lahore in Punjab, Karachi in Sind, Peshawar in NWFP and Quetta in Baluchistan. Two villages were chosen from each district. Within these predefined strata, households have been selected randomly. Our target was to achieve a total of 30 responses per village, that is 60 households per district, yielding an overall sample size of N = 600.

Thus, we have sampled a total of 240 households from Punjab, 180 households from Sind, 120 households from NWFP and 60 households from Baluchistan. More- over, to ensure a good representation of rural Pakistan, we apply weights to each household with respect to the district of origin as shown in Table A-1. All our econometric results are based on that weighting scheme.

Our survey data provides household-level data on health and different socio- economic variables like income, family size, education, marital status, age and sex of the household's head. The household's health is determined by the household's self- assessed health status. The following question has been asked from the household head: "During the last 12 months, how would you assess the health status of your household?". The answers included: excellent, good, fair, poor, very poor. In the following, we denote the health variable as (h). More specifically, households are ranked according to their self-assessed health status. A higher rank (4, 3, 2, 1, 0) corresponds to the higher health status of a household and vice-versa.

3.2. Descriptive Statistics

Table 1 illustrates the concept and provides some descriptive statistics. The health distribution is skewed towards the lower health scales with the mass of the distri- bution being concentrated on that side. For instance, the given distribution shows that 11.33% and 31.50% of the observations are reporting higher health i.e. 4 and 3, respectively. While more than half of the population (nearly 60%) finds itselt at the lower health scales. This suggests a overall fairly unequal distribution of health.

	Health Variable
Mean	-
Standard deviation	-
Health status:	
Excellent	4
Good	3
Fair	2
Poor	1
Very poor	0

Table 1: Descriptive statistics for household's health variables.
Source: Survey 2008.

Inequalities in health can arise for various reasons. Most countries identify dif- ferences in health status by social grouping and economic status. To single out the effects of income on health inequalities, we investigate in particular the AIH, RIH and IIH. We first want to ensure that such a proposed link between income and health exists for our data. We therefore calculate the correlation coefficient between health and the rank of a household within the income distribution (used as proxy of relative income).5 We use relative income to capture the socio-economic status which is said to affect the health status (see e.g. Wagstaff et al. (1989); Humphries and van Doorslaer (2000); Gerdtham and Johanneson (2000); Doorslaer et al. (1997); Doorslaer and Koolman (2000) and Bommier and Stecklov (2002)). This can be computed as follows:

$$C_{h,R} = \frac{\operatorname{cov}(h_i, R_i)}{\sqrt{\operatorname{var}(h)\operatorname{var}(R)}} = 0.224, \quad (i = 1, 2, 3.....600) \qquad (1)$$

where hi is the health variable of a household, Ri is the ith household's fractional rank in the income distribution, cov is the covariance between the two and var measures the variance of the given variables.[6] Equation (1) suggests that health is positively correlated with the income ranking of a household in the given income distribution such that higher income groups are relatively better-off in terms of health and vice versa.

The household's relative income in social comparisons is assessed with respect to the average income of the overall society or community.7 The average income, how- ever, may vary across a certain group, community or region (Wagstaff and Doorslaer, 2000). Economists therefore prefer to measure relative household income at the com- munity level for cross-sectional data. In our case, the community might be considered to correspond to a district.8 Thus, the relative income at community/district level is given by

$$(y_r)_{ij} = \frac{y_i}{\bar{y}_j},$$

where (yr)ij is the relative income of ith household in district j, yi denotes absolute income of the ith household and yj represents the average income of the jth district.

---

[5] See Li and Zhu (2006) for further methodological details.

[6] In fractional ranking, items that compare equal receive the same ranking number, which is the mean of what they would have under ordinal rankings. Furthermore, in ordinal ranking, all items receive distinct ordinal numbers (1, 2, 3, and so on..), including items that compare equal.

In analysing the relationship between income and health, we consider in par- ticular two key hypotheses: the AIH and the RIH. The AIH seeks to investigate a relationship between the household's income and health status, whereas accord- ing to the RIH, health is rather affected by relative income differentials. The RIH builds upon the claim that low relative income increases psychosocial stress which may lead to physical illness (Cohen et al. (1991); Cohen et al. (1997)).9 Similarly, several studies suggest that it is in fact an individual's relative income instead of absolute income that matters (Marmot et al. (1991); Wilkinson (1997); Wilkinson (1998)). If this were the case, a doubling of everyone's income would have no ef- fect on health. Such arguments explain income effects psycho-socially, rather than in materialistic terms and have lead to a model of health in which social coherence plays an important role. Additionally, the individual's health is also attributed to the distribution of income within a society. For instance, living in a region with an unequal distribution of income by itself afflicts health (Wilkinson, 1996), which is related to psychosocial mechanisms rather than material deprivation.

Table. 2 Provides summary statistics of material and psychosocial determinants of health for our sample.10 Income inequality is measured here in terms of the Gini coefficient. This is arguably the most commonly used measure of income inequality in testing the IIH in an attempt to establish a relationship between income inequality and health (see, amongst others, Kennedy et al. (1998); Mellor and Milyo (2001);
Soobader and LeClere (1999)).

Variable	Obs	Mean	Std.Dev	Min	Max
Absolute income y_i	600	3515.58	1207.88	1596	6938
Relative income$(y_r)_{ij}$	600	1.00	0.08	0.79	1.24
Income inequality of district G_j%	10	4.42	2.19	1.6	8.1

Table 2: *Material and psychosocial determinants of health*
Source: Survey 2008.
Note: Al l income figures are related to the households' monthly incomes in Pakistani currency (PKR).

In our survey, individuals are grouped into 10 districts. We estimated the in- equality index Gij at the household level such that households living within the same district have been assigned the same index. Inequality hence differs across districts but not across households within any given district.11

	h_i	y_i	$(y_r)_{ij}$
h_i	1		
y_i	0.1865***	1	
$(y_r)_{ij}$	0.7045***	0.2148***	1

Table 3: Correlation matrix between health and income variables
N = 600 households. *, **, *** indicates significance level of 10%, 5% and 1%, respectively.

----

[7] This implies that relative incomes with positive or negative values indicate the household's income to be greater than or less than the average income of society, respectively.
[8] We are using the words "community" and "district" interchangeably in the remainder.
[9] Low social status/prestige and lack of control and awareness are often labelled as psychosocial determinants of health, even though they may be triggered by material factors such as lack of income or bad housing (Kawachi et al., 2002).
[10] See Kawachi et al. (2002) for further methodological details.
[11] Gij is a contextual variable that varies across districts but has the same value for all the households within a district. A similar idea has been presented by (Blalock, 1984) and (Lindley and Lorgelly, 2005) in order to explain individual-level variables by using group-level variables

Table 3 reports the correlations between health and the income variables. Ac- cording to Table 3, health is positively correlated to income in general. But the correlation is much stronger in case of relative income ($(yr)_{ij}$) compared to the ab- solute income (y_i) of the household.

3.3. Empirical Specification

The measure of health that we use as endogenous variable throughout the paper is the health h of ith household, h_i. Health is measured on an ordinal scale ranging from 0 to 4, derived from the answer to the question: "During the last 12 months, how would you assess the health status of your household?". As discussed before, answers included: excellent, good, fair, poor, very poor. A higher number (4, 3, 2,1, 0) corresponds to the better health of a household and vice-versa.

Our regressions include a list of control variables relating to the households whose effects on health have been shown to be important in the literature. These control variables are in particular: age, sex, income (both absolute and relative), district- specific income inequality, marital status, children, family size and education. More specifically, to capture age we use the household's head age. We create two dummies for sex of the household's head such that if f emale = 1 otherwise 0. The household's absolute income and relative income (with respect to the district) are expressed in logarithmic terms. Similarly, the variable measuring the degree of income inequality in the districts will be considered in percentage terms in our regressions. For marital status and children we create dummies. If it is a couple = 1 (dual parent family) or otherwise 0. Similarly, the case of the household having children (aged less than 16 years and living with their parents) is assigned 1 or 0 otherwise. The total number of household members determines the f amilysize. The education variable is measured by the number of years of education of the household's head. Table B-1 presents the correlations between the given controls that we use as socio- economic determinants of health in our regressions. Our discussion focuses on the two relationships that are found to be statistically significant. According to Table B-1, age is negatively correlated with being a female ($r = -0.1264$), the significance of the estimated coefficient suggests that female heads of households are younger than male heads of household. Which in a development setting where male headed households are the norm, suggests that when males are die, their wives usually take over and there is normally an age gap between husband and wife. The correlation between family size and children is positive and statistically significant ($r = 0.1592$). This positive relationship confirms our initial conjecture because having children at home obviously adds to the family size.

Following much of the literature, the baseline empirical specification that we use for studying the determinants of health is as follows

$$h_i = \alpha + \beta_1 \log(y_i) + \beta_2 \log(yr)_{ij} + \gamma(G)_{ij} + B(X)_i + e_i, \qquad (2)$$

where the subscripts i and j in (2) refer to households and districts, respectively. The household's absolute income is denoted by y_i; $(yr)_{ij}$ and G_{ij} represent district- specific measures of relative incomes and income inequality. X_i denotes the remaining controls mentioned.

Model (2) can be considered as the empirical counterpart of a health function of the general form $h = f(y, yr, G, X)$ which we estimate using ordered probit regression. This specification will serve in the remainder as our benchmark model to investigate the three income hypotheses.

4. Empirical Results

4.1. Ordered Probit Estimation

4.1.1. Absolute And Relative Income Hypotheses: Separate And Joint Tests

Assuming for now that it is absolute rather than relative levels of income that matter, we would consider a health function of the form $h(y, X)$, where y measures absolute income and X represents the given set of control variables. We estimate this health function with the following empirical specification in order to test the AIH:

$$h_i = \alpha + \beta \log(y_i) + B(X)_i + e_i. \qquad (3)$$

Alternatively, we may also consider that it is not absolute but rather relative levels of income that affect health.

This would suggest a health function of the form h(yr , X), where yr represents the relative income of the household with respect to the given district and X represents the set of control variables. For testing the RIH, we estimate the health function as follows:

$$hi = \alpha + \beta \log(yr)ij + B(X)i + ei. \quad (4)$$

We may moreover assume that people care about both absolute and relative levels of income. This implies a combined health function of the form h(y, yr , X), with y, yr and X defined as above. This can be translated in econometric terms as follows to test the income hypotheses jointly in absolute and relative terms:

$$hi = \alpha + \beta 1\log(yi) + \beta 2\log(yr)ij + B(X)i + ei. \quad (5)$$

According to the RIH, however, the household's health is additionally affected by the distribution of income (Gij) within a society so that living in a place with an unequal income distribution is anticipated to lead to a worsening of the health experience. This would suggest a health function of the form h(y, yr , G, X), which could be expressed in econometric terms as follows:

$$hi = \alpha + \beta 1\log(yi) + \beta 2\log(yr)ij + \gamma 1(G)ij + B(X)i + ei. \quad (6)$$

Since health is usually assumed to be curvilinear in income inequality, we in- troduce a squared inequality term (G2) to allow for such potential non-linearities in health outcomes. This would imply accordingly a health function of the form h(y, yr , G, G2, X), which can be specified as:

$$hi = \alpha + \beta 1\log(yi) + \beta 2\log(yr)ij + \gamma 1(G)ij + \gamma 2(G)2 + B(X)i + ei. \quad (7)$$

The estimation strategy is as follows. First, we use (3) and (4) to investigate the AIH as well as the RIH in turn. Having estimated both models separately, we use in a next step (5) to test both hypotheses jointly. Model (6) will then be used to test all three hypotheses (i.e. including the IIH) simultaneously. Finally, (7) is estimated to capture any potential non-linearity in health with respect to the distribution of income within a district.

Table B-2 catalogues the coefficient estimates of five different specifications of the ordered probit regression in order test for the three income-health hypotheses, separately as well as jointly. Ensuring well-specified models throughout, we also con- ducted the linktest. The idea of this test is that if the model is properly specified, one should not be able to find any additional predictors that are statistically significant except by chance. The variable _hat should thus be a statistically significant pre- dictor since it is the predicted value from the model. This will be the case unless the model is misspecified. On the other hand, if our model is properly specified, variable

_hatsq should not have much predictive power except by chance. Therefore, the linktest is significant for a significant _hatsq. Such a test outcome would usually suggest that either we have an omitted-variable bias or it might be the case that the link function is not correctly specified. According to Table B-2, the variable _hatsq appears to be statistically insignificant for all the given specifications of the ordered probit model for health. The main aspect to consider in this context is the signifi- cance of _hat. This basically checks whether we need more variables in our model by running a new regression with the observed outcome variable. In our case, the _hat is significant only for the joint test of the AIH and the RIH as shown in Table B-2, column (3), which means that our ordered probit model for health is correctly specified and we therefore do not require any additional variables that significantly determine the health outcome. The corresponding marginal effects for the joint test of the AIH and the RIH are reported in Table B-3.

In the ordered probit regression, a positive coefficient indicates an increased chance that a subject with a higher score on the independent variable will be ob- served in a higher health category. Similarly, a negative coefficient implies a relatively greater chance that a subject with a higher score on the independent variable will be observed in a lower health category. For instance, as shown in Table B-2, higher income increases the chances of being in a higher health category and the result obtained is statistically significant. However, since the protective effect of absolute income on health is relatively uncontested (compared with the effect of relative in- come and income inequality), we do not place very much emphasis on it. Similarly, the results in Table B-2 show a positive association between relative income and the household's health, which means higher relative income brightens the chances of being in a higher health category.

Results suggest evidence in favour of both the AIH as well as the RIH. Unlike Li and Zhu (2006) who failed to establish a significant impact of relative income on health, we find that both measures are statistically significant and of positive signs. Ferrer-i-Carbonell (2005) use German data to empirically analyse the importance of relative income for household's well- being. He finds that absolute income has a very small and not significant coefficient when included alongside relative income. It is also interesting to note that once we control for both absolute and relative income, the effects of the relative and absolute incomes remain unaffected and statistically significant as well.

We test additionally for the effects of income inequality in linear (column 4) and quadratic (column 5) terms. The positive coefficient of "income-inequality linear" and the negative coefficient of "income-inequality squared" suggest an inverted U- shaped pattern between health and income inequality. The maximum corresponds to an income inequality of 6.57%. This suggests that any increase in income inequality is beneficial for health till reaching the 6.57% threshold after which higher income inequality poses threats to health. We may therefore conclude that the IIH is only supported by districts with high income inequalities – with "high" being pinned down to about 7% in this context. However, on pure statistical grounds this relationship does not exist at all. A similar story applies to China as for instance suggested by Li and Zhu (2006) who found an inverted U-shaped (and statistically siginificant) association between household's health and community level income inequality.

In addition, we obtain a significant and negative relationship between family size and health, which indicates that households with larger families are most likely to be observed in a lower health category and vice versa. Health appears to be U-shaped in age – the estimated turning point corresponds to the age of 53 years. Such a pattern might explain the mid-age crises as young parents with children struggle relatively strongly for their career and financial management. As they turn old and their children grow up, however, their lives normally become more stable which positively affects the overall health status of a household. The remaining controls included in the model are found to be statistically insignificant.

As far as the ancillary parameters (or cut points) are concerned, Cut1 is the estimated cut point of the latent variable $\tilde{h}$ (which is a continuous and unobservable response variable) used to differentiate very poor health from higher health cate- gories (i.e. poor, fair, good and excellent). For example, households that had a value of -1.7035 or less on the underlying latent variable that gave rise to our health cate- gory variable would be classified as very poor. Cut2 is the estimated cut point used to differentiate very poor and poor health categories from higher health categories (fair, good and excellent). This means that households that had a value of -0.2750 or greater on the latent variable would be classified in the higher health categories. Cut3 differentiates very-poor, poor and fair health categories versus higher health categories like good and excellent such that households that had a value of 0.5775 or greater on the given latent variable would be classified in those higher health categories. Cut4 distinguishes very poor, poor, fair and good health categories from the highest category; excellent. Cut4 which is equal to 6.262 indicates that house- holds that had a value 1.6451 or greater on health variable would fall in the excellent category for health.

The corresponding marginal effects are presented in Table B-3. According to these marginal calculations, health appears to be inverted U-shaped in age for the lower health category (1), the estimated turning point is around 56 years. In contrast, health is U-shaped in age for the higher health categories (3) and (4) with estimated tipping points of 54 to 56 years, respectively. We may conclude that before 54-56 years (middle-age) with each year increase in the age of the household, the probability to be in the lower health category (1) increases by 2.59% and at the same time the likelihood to be observed in the higher health categories (3) and (4) decreases by 1.96% and 1.80%, respectively. Conversely, after crossing the middle-age, the probability of being in the lower health category (1) decreases by 0.023% and at the same time the chance to be in the higher health categories (3) and (4) increases by 0.018% and 0.016%, respectively. Overall, we may conclude that age has a non-linear (U-shaped) effect on household's health status. Before the age of around mid-fifties households are more likely to be observed in lower health categories but after crossing that age they are most probably to lie in the higher health categories. Furthermore, the likelihood of being in the lower health categories (0) and (1) increases by 1.6% and 17.87% respectively, if the households are having kids in their homes. These findings are in line with initial conjectures as households with kids potentially have more health problems related with kids compared to the households without children.

Similarly, a larger family size is inversely related to the health of a household. For instance, a unit increase in the family size of a household increases its chance to lie in the lower health categories (0), (1) and (2) at the rate of 0.35%, 2.27% and 0.68%, respectively. At the same time, it lowers its chance to be observed in the higher health categories (3) and (4) by the amount 1.72% and 1.58%, respectively. As far as household's absolute and relative incomes are concerned, those have a positive influence on household's health. If, for example, the household income increases by 1%, its probabilty to be observed in the lower health categories (0), (1) and (2) decreases, while at the same time it is more likely to be observed in the higher health categories such as (3) and (4).

Overall, we find evidence supporting both the AIH and the RIH. However, the effect of the hypothesis in relative terms appears more pronounced than in absolute form. We may conclude that being in better health involves more psychosocial fac- tors rather than absolute material standards. This means that wealthier households' health is positively influenced by their higher incomes relative to their reference group. On the other hand, lower relative income weakens one's power in the alloca- tion of efficient local health-related resources and thus leads to a poor health status, stress and potential depression.12 However, we failed to confirm the hypothesis that more egalitarian societies are characterised by better health. Hence, we cannot pro- vide evidence for the IIH. Similarly, Lobmayer and Wilkinson (2000), using data on 14 OECD countries, failed to accept the IIH either suggesting that income inequality is rather beneficial for health.

5. Concluding Remarks

In this paper we have sought to investigate the validity of three key income-health hypotheses: the AIH, the RIH and the IIH using ordered probit regression. The analysis is based on household survey data (Survey 2008) for rural Pakistan. We specify a health function that is general enough to permit a separate as well as simultaneous investigation. We find evidence in favour of the RIH and AIH and no support for the IIH. Relative income appears to have a significantly positive effect on health outcomes in rural Pakistan. This finding is in contrast to the general view established for developed countries according to which incomes in absolute terms seem to be the main driver of a household's health status.

Understanding the relationship between income and health is of obvious relevance to policymakers. In a "first-best" world, health differentials should not hinge on material factors after all. The relationship between the socio-economic environment and health outcomes may be particularly relevant for regions with weak public health provision such as rural Pakistan. The Pakistani government should therefore take note of this issue in areas, where relatively poor health care system exists. Thus, policymakers concerned about health would be well advised to improve the quality as well as the number of public health care units and hospitals, especially in areas which are considerably far away from the major cities like Islambad in Punjab, Karachi in Sind, Peshawar in NWFP and Quetta in Baluchistan.

A Weighting matrix

District	No. of households	Total Population	Rural Population	Sample Population	pweights	pweights-normalised
	Census 1998			Survey 2008		
					$\{(RP)^0/(SP)_j\}$	$\{(Pw)_j/\Sigma(Pw)_j\}$
(j)	$(hh)^0_j$	$(TP)^0_j$	$(RP)^0$	$(SP)_j$	$(Pw)_j$	
Attock	206,678	1,274,935	1,003,843	266	3,773.85	0.10
Layyah	152,050	1,120,951	976,748	289	3,379.75	0.09
RahimyarKhan	3,141,053	17,743,645	2,524,471	246	10,262.08	0.27
Sahiwal	n.a	1,843,194	1,541,204	269	5,729.38	0.15
Badin	211,354	1,136,044	949,556	267	3,556.39	0.09
Mirpurkhas	148,470	905,935	605,760	251	2,413.39	0.06
Thatta	220,068	1,113,194	988,455	259	3,816.43	0.10
Lower Dir	76,531	717,649	673,314	241	2,793.83	0.07
Malakand	49,330	452,291	409,112	234	1,748.34	0.05
Kalat	34,410	237,834	204,040	215	949.02	0.02
Total	-	11,943,080	9,876,503	2,537	38,422.46	1.00

Table B-1: *Correlation matrix between health determinants.*N = 600 *households. *, **, *** indicates significance level of 10%, 5% and 1%, respectively.*

Table B-2: AIH and RIH: separate and joint tests using ordered probit regression.Note: regression coefficients are in bold and standard errors appear below them. *,**,*** denote statistical significance at 10%, 5% and 1% levels. Figure in brackets[] are p values.

Oprobit regression	Dependent Variable: Health (h)				
	Separate tests			Joint tests	
Specification:	1	2	3	4	5
Age	**-0.0913***	**-0.0943***	**-0.0953***	**-0.0943***	**-0.0863***
	0.0502	0.0522	0.0511	0.0510	0.0514
Age Squared	**0.0008***	**0.0009***	**0.0009***	**0.0008***	**0.0008**
	0.0005	0.0005	0.0005	0.0005	0.0005
Female	**-0.2338**	**-0.2736**	**-0.2401**	**-0.2485**	**-0.2432**
	0.1881	0.1841	0.1871	0.1866	0.1873

Table B-2: Continued

Oprobit regression	Dependent Variable: Health (h)				
	Separate tests		Joint tests		
Specification:	1	2	3	4	5
\cut1	**-1.4137**	**-5.7580**	**-1.6970**	**-1.1352**	**-0.5236**
	1.9256	1.6332	1.9473	2.0520	2.1524
\cut2	**-0.0074**	**-4.3379**	**-0.2677**	**0.2976**	**0.9074**
	1.9381	1.6281	1.9605	2.0645	2.1695
\cut3	**0.8404**	**-3.5019**	**0.5849**	**1.1518**	**1.7625**
	1.9418	1.6272	1.9642	2.0685	2.1744
\cut4	**1.9055**	**-2.4557**	**1.6522**	**2.2194**	**2.8327**
	1.9483	1.6273	1.9709	2.0763	2.1827
Log pseudolikelihood	-821.229	-826.863	-817.727	-816.992	-816.367
Number of obs	600	600	600	600	600
Wald $\chi^2(k)$	Wald $\chi^2(11)$	Wald $\chi^2(11)$	Wald $\chi^2(12)$	Wald $\chi^2(13)$	Wald $\chi^2(14)$
	= 28.49	= 22.02	= 34.67	= 35.30	= 35.86
Prob $> \chi^2$	0.0027	0.0242	0.0005	0.0008	0.0011
Pseudo R^2	0.0217	0.0150	0.0258	0.0267	0.0275

Table-B: Continued

Oprobit regression	Dependent Variable: Health (h)				
	Separate tests		Joint tests		
Specification: .linktest	1	2	3	4	5
_hat	**0.6224**[0.329] 0.6262	**0.7526**[0.842] 3.7695	**0.9568****[0.029] 0.4	**0.9298**[0.319] 0.9329	**1.2930**[0.378] 1.4672
_hatsq	**0.2457**[0.529] 0.3904	**-0.0346**[0.947] 0.5245	**0.0425**[0.912]	**0.0324**[0.937] 0.4130	**-0.0859**[0.839] 0.4214
\cut1	**-1.5376** 0.2460	**-5.3180** 6.7612	**-1.7035** 0.1517	**-1.1696** 0.4918	**-0.2836** 1.2288
\cut2	**-0.1342** 0.2397	**-3.8977** 6.7715	**-0.2750** 0.1119	**0.2627** 0.4992	**1.1490** 1.2463
\cut3	**0.7132** 0.2435	**-3.0616** 6.7719	**0.5775** 0.1165	**1.1168** 0.5049	**2.0044** 1.2518
\cut4	**1.7800** 0.2505	**-2.0155** 6.7662	**1.6451** 0.1294	**2.1846** 0.5095	**3.0740** 1.2547

Table-B: Continued

Oprobit regression	Dependent Variable: Health (h)				
	Separate tests			Joint tests	
Specification: .linktest	1	2	3	4	5
Log pseudolikelihood	-821.028	-826.8599	-817.7186	-816.9880	-816.3364
Number of obs	600	600	600	600	60 0
Wald $\chi^2(2)$	27.18	21.46	33.43	34.26	33.85
Prob > χ^2	0.0000	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.0219	0.0150	0.0259	0.0267	0.0275

Table B-3: Marginal effects after oprobit using AIH and RIH joint test.

*Note: 1. marginal effects are in bold and standard errors appear below them. 2. *,**,*** denote statistical significance at 10%, 5% and 1% levels. 3. (*) dy/dx is for discrete change of dummy variable from 0 to 1.*

Marginal effects after oprobit

y = Pr(Health category==0) (predict, p outcome(0))

= 0.01680492

y = Pr(Health category==1) (predict, p outcome(1))

= 0.22656655

y = Pr(Health category==2) (predict, p outcome(2))

= 0.31904463

y = Pr(Health category==3) (predict, p outcome(3))

= 0.32718251

y = Pr(Health category==4) (predict, p outcome(4))

= 0.1104014

Outcome	(0)	(1)	(2)	(3)	(4)
Variable	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Age	0.0040*	0.0259*	0.0077*	-0.0196*	-0.0180*
	0.0024	0.0139	0.0045	0.0106	0.0098
Age Squared	-0.00004	-0.00023*	-0.00007	0.00018*	0.00016*
	0.00002	0.00013	0.00004	0.00010	0.00010

Table-3 Continued

Outcome	(0)	(1)	(2)	(3)	(4)
Variable	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Female* Absolute	0.0125	0.0677	0.0122**	-0.0526	-0.0398
	0.0124	0.0545	0.0051	0.0427	0.0275
income Relative	-0.0214***	-0.1394***	-0.0415***	0.1055***	0.0968***
	0.0075	0.0414	0.0149	0.0324	0.0277
income Couple*	-0.0562*	-0.3654**	-0.1088**	0.2766**	0.2537**
	0.0315	0.1548	0.0486	0.1205	0.1084
Children*	-0.0066	-0.0398	-0.0099	0.0305	0.0258
	0.0058	0.0310	0.0066	0.0240	0.0188
Family size	0.0160***	0.1787**	0.1546	-0.0686	-0.2807
	0.0062	0.0794	0.1475	0.0721	0.2990
Years of Education	0.0035*	0.0227*	0.0068*	-0.0172*	-0.0158*
	0.0021	0.0131	0.0041	0.0098	0.0091
	0.0003	0.0020	0.0006	-0.0015	-0.0014
	0.0006	0.0039	0.0012	0.0029	0.0027

Table-3 Continued

Outcome	(0)	(1)	(2)	(3)	(4)
Variable	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
Region:					
Punjab*	-0.0084	-0.0535	-0.0152	0.0406	0.0366
	0.0068	0.0426	0.0124	0.0321	0.0292
NWFP*	0.0016	0.0099	0.0028	-0.0075	-0.0067
	0.0073	0.0450	0.0120	0.0344	0.0298
Sind*	0.0003	0.0018	0.0005	-0.0013	-0.0012
	0.0066	0.0425	0.0125	0.0322	0.0294

References

Blalock, H. M. (1984). Contextual-effects models: theoretical and methodological issues, Department of Sociology, University of Washington, Seattle, Washington 98195.

Bommier, A. and Stecklov, G. (2002). Defining health inequality: why Rawls suc- ceeds where social welfare theory fails, Journal of Health Economics 21: 497–513.

CIA (2008). The World Factbook.

Cohen, S., David, T. and Andrew, S. (1991). Psychological stress and susceptibility to the common cold, New England Journal of Medicine 325(9): 606–12.

Cohen, S., Scott, L., Stephen, M., Bruce, R., Eugene, H. and Jay, K. (1997). Chronic social stress, social status and susceptibility to upper respiratory infections in nonhuman primates, Psychosomatic Medicine 59(3): 213–21.

Daly, M., Duncan, G., Kaplan, G. and Lynch, J. (1998). Macro-to-micro links in the relation between income inequality and mortality, Milbank Quarterly.

Deaton, A. (2003). Health inequality and economic development, Journal of Eco- nomic Literature 41: 113–158.

Doornbos, G. and Kromhout, D. (1990). Education level and mortality in a 32-year follow-up study of 18-year-old men in the netherlands, International Journal of Epidemiology 19: 374–79.

Doorslaer, E. V. and Koolman, X. (2000). Income-related inequalities in health in Europe: evidence from the European Community Household Panel., Ecuity II Project, Working Paper 1, Erasmus University, Rotterdam.

Doorslaer, E. V., Wagstaff, A. and Bleichrodt, H. (1997). Income-related inequalities in health: some international comparisons, Journal of Health Economics 16: 93–112.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect, Journal of Public Economics 89: 997–1019.

Fiscella, K. and Franks, P. (1997). Poverty or income inequality as predictor of mortality: longitudinal cohort study, British Medical Journal 314(7096): 1724–27.

Fox, A., Goldblatt, P. and Jones, D. (1985). Social class mortality differentials: artifact, selection or life circumstances?, Journal of Epidemiology and Community Health 39: 1–8.

Gerdtham, U.-G. and Johanneson, M. (2000). Income-related inequality in life-years and quality-adjusted life-years., Journal of Health Economics 19(6): 1007–1026.

Gerdtham, U.-G. and Johannesson, M. (2004). Absolute income, relative in- come, income inequality and mortality, The Journal of Human Resources (JHR) XXXIX(1): 228–247.

Humphries, K. and van Doorslaer, E. (2000). Income-related inequalities in health in Canada, Social Science and Medicine 50: 663–671.

Kawachi, I., Subramanian, S. V. and Almeida-Filho, N. (2002). A glossary for health inequalities, J Epidemiol Community Health 56: 647–652.

Kennedy, B., Kawachi, I., Glass, R. and Prothrow-Stith, D. (1998). Income distribu- tion, socioeconomic status and self-rated health in the United States: Muti-level analysis, British Medical Journal 317(7163): 917–921.

Li, H. and Zhu, Y. (2006). Income, income inequality and health: evidence from
China, Journal of Comparative Economics 34: 668–693.

Lindley, J. and Lorgelly, P. (2005). The relative income hypothesis: does it exist over time?, Sheffield Economic Research Paper Series.

Lobmayer, P. and Wilkinson, R. (2000). Income, inequality and mortality in 14 developed countries, Sociology of Health and Il lness 22(4): 401–414.

Marmot, M., George, S., Fiona, N. and Amanda, F. (1991). Health inequalities among British civil servants: the Whitehall II study, Lancet 337(8754): 1387–93.

Meara, E. (1999). Inequality and infant health, Harvard Medical School. Manuscript.
Boston, Massachusetts:.

Mellor, J. and Milyo, J. (2001). Reexamining the evidence of an ecological association between income inequality and health, Journal of Health Politics, Policy and Law 26(3): 487–522.

Mellor, J. and Milyo, J. (2002). Income inequality and health status in the United states: evidence from the current population survey, Journal of Human Resources

Power, C., Manor, A. F. and Fogelman, K. (1990). Health in childhood and social inequalities in health in young adults, Journal of the Royal Statistical Society 153: 17–28.

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

Soobader, M. and LeClere, F. (1999). Aggregation and the measurement of income inequality: effects of morbidity, Social Science and Medicine 48(6): 733–744.

Wagstaff, A. and Doorslaer, E. V. (2000). Income inequality and health: what does the literature tell us?, Public Health 21: 543–67.

Wagstaff, A., van Doorslaer, E. and Paci, P. (1989). Equity in the finance and delivery of health care: some tentative cross-country comparisons, Oxford Review of Economic Policy 5(1): 89–112.

Wilkinson, R. (1996). Unhealthy societies: the afflictions of inequality, London Routledge.

Wilkinson, R. (1997). Health inequalities: relative or absolute material standards?,
British Medical Journal 314(7080): 591–5.

Wilkinson, R. (1998). Low relative income affects mortality, British Medical Journal
316(7144): 1611.

Wilkinson, R. G. (1986). Class and health: research and longitudinal data, London, England: Tavistock.