Is Child Work Injurious to Health?

Aditi Roy*

Southern Methodist University

June 2009

Abstract

Estimating the causal impact of child work on the contemporaneous health of a child has proven quite challenging given non-random selection into the labor market and the inability to find strong and valid instruments. Our data, the Indonesian Family Life Survey is no different. Recognizing the lack of a credible instrument, we instead pursue a different strategy based on the methodology of Altonji et al. (JPE, 2005). This method assesses the robustness of the impact of child work estimated under the assumption of random selection (i.e., selection into child work on observable attributes only) to varying degrees of non-random selection (i.e., selection into child labor on unobservable attributes). If the estimated effect is found to be extremely sensitive to selection on unobservables, then one should be wary about infering an adverse causal effect of child work. In addition, the nature of the selection process is identified using parametric assumptions. The results are striking, suggesting positive selection of children into work when we consider underweight and high weight status as dependent variables. This indicates that there is both healthy worker selection effect as well as unhealthy worker selection effect. There is however negative selection into work for the children belonging to the intermediate weight range. This heterogeneity in the selection process across the distribution has not been previously identified in the literature. Moreover, we also find evidence suggesting a heterogeneous impact of child work on health once we allow for a modest amount of selection on unobservables. Specifically, we find evidence of a negative causal effect of work on healthier children, but evidence of beneficial impact of work on the least healthy children.

JEL: I12, J13, J22, J28

Keywords: Child work, health, selection on unobservables, Indonesia

*Department of Economics, Southern Methodist University, Dallas, TX 75275-0496. Tel: (214) 768-3981. Fax: (214) 768-1821. E-mail: aroy@smu.edu

1 Introduction

Children's weight distribution is an important criterion for measuring their health. Using data from Indonesian Family and Life Survey this paper investigates the effect of child work on their health by looking at the children's weight distribution.[1] Our two main findings are: (1) There is evidence indicating positive selection of children into work for both the healthiest and the least healthy children. This suggests the existence of both the healthy worker selection effect and the unhealthy worker selection effect. However, we find negative selection into work for the children who belong to the intermediate weight category. To the best of our knowledge such heterogeneity in the selection of children across the weight distribution has not been reported in the literature earlier. (2) We also find evidence of heterogeneous impact of child work on health. There is indication of negative causal impact of child work on the healthiest children, while there is evidence of beneficial effect of child work on the health of children at the lower end of the weight distribution. We also find that child work increases the probability of the children of being in the intermediate weight range.

The impact of child work on the child's health is an issue of interest both on humanitarian grounds as well as from a policy perspective. According to an ILO (2005) report, it is estimated that around 246 million children are working in the world. Though child health is a vital component of human capital much of the empirical research focuses on the impact of child work on schooling. Poor health can affect school attendance and also make children miserable and unhappy thereby reducing their welfare. Child health also affects adult health which has an impact on labor market outcomes both at the microeconomic and macroeconomic levels. This has been documented in a number of studies. For instance, Strauss and Thomas (1998) find that there is a positive impact of height on earnings while Weil (2005) indicate that variation in health explains approximately 20 percent of the cross-country variation in log income per worker which is around the same fraction as explained by the variation in education. Clearly, further research is necessary to understand how work affects a child's health and what policies need to be prescribed to eradicate child work when it negatively affects health.

In literature it is found that child work can have both negative and positive impact on health. For example, wages earned from child work can improve the living standard of poor households. Better food and a better living style can improve the health of the child. Steckel (1995), Appleton and Song (1999), Smith (1999) find a positive impact of child work on the living standards of families and hence on health. On the other hand, child work can also lead to chronic illness or fatal injuries. According to the Statistical

[1] We believe that this is the first paper to examine the relationship between child work and health using the weight distribution. The weight categories which are formally defined later in the paper are Underweight children, children with high weight and those in the intermediate range.

Information and Monitoring Programme on Child Labor (SIMPOC) survey in Ghana, parents report around 29.4 percent of the child workers to be either injured or ill while in Cambodia this figure is 47.1 percent. A 1991 study by the Committee for the Creative Education of Indonesian Children shows that over 70 percent of the children working in the export factories of Tangerang, Indonesia are sick. So it is necessary to unravel this complex question of the direction of effect of child work on health.

However, the empirical child work literature suffers from a well known endogeneity problem. A number of factors like health endowment and family's attitude towards their children may affect child work but are unobservable to an econometrician. These characteristics are however observable to the households. Parents may decide to send their healthiest child to work which can result in a positive correlation between work and health outcomes. This is called the healthy worker selection effect. However, it is also possible that a family may send their healthier child to school and the least healthy child to work. This may result in a negative relationship between child work and health and may be called the unhealthy worker selection effect. So the bias due to this non-random selection of children into work can be either positive or negative. Note that the standard regression method is inadequate for estimating the causal impact of child work on health in presence of this type of selection bias.

In fact a lot of the empirical work has focused only on the correlation between child work and health. Satyanarayana et al. (1986) show that in the rural areas of Hyderabad, India, boys who work for wages suffer growth deficits when compared to boys who attend school. Kassouf et al. (2001) using information from a cross-section of rural and urban adults of southeast and northeast Brazil show that the probability of reporting poor health in adulthood rises as the age of entry into workforce falls. Guiffrida et al. (2001) use a nationally representative cross-section of 18-60 year old Brazilian adults and find that the entry to labor force at or below the age of 9 has a statistically significant and substantial negative effect on health in adulthood.

The ideal solution to control for unobserved characteristics which may affect both child work and health outcomes, when using non-experimental data, is to instrument for child work. However, a strong and valid instrument is difficult to find in child work literature. Rosati and Straub (2006) attempt to overcome this problem by estimating a conditional fixed effect model using data on siblings. Their results indicate a negative impact of child work on adult health in Guatemala. However, this method can only control for unobservables that are fixed across siblings. O'Donnell et al. (2005), Beegle et al. (2004) and Wolff and Maliki (2008) address this selection bias issue by using the instrumental variable estimation technique. We now discuss these three papers in detail.

Beegle et al. (2005) use data from Vietnam to look at the health outcomes of children who were engaged in child work five years ago. They find that child health is not significantly associated with prior work. The instruments used for child work are community disasters, rice prices and the interaction term of community disasters with log per capita household expenditure in 1992-93. Their results identify a pure labor effect on health since they only consider a sample of school going children. In Beegle et al.'s paper while their instruments are jointly significant, the reported F-statistic is 9.07 which is less than the usual rule of thumb of F=10 recommended by Stock and Yogo (2005) for avoiding problems associated with weak instruments.

The study by O'Donnell et al. (2005) is closest to our work. This interesting paper analyzes the contemporaneous effect of child work on the body mass index (BMI) of children in Vietnam and finds a positive significant relationship between child work and BMI. They claim that this positive association reflects the healthy worker selection effect. Instrumenting for the child work variable they do not find a significant impact of work on the BMI of children. The instruments used are the relative price of rice in the community, demand for labor proxied through indicators of work related migration to and from the community, and the year in which the community primary school was built as a proxy for the school quality, along with the interaction of the above instruments with the child work dummy. However, the F-statistic reported for the significance of instruments is 4.96, well below the usual Stock and Yogo (2005) rule of thumb. Moreover, there may be other problems associated with these instruments. According to the authors, an increase in the price of rice reduces child work. But a change in rice prices may directly affect health through an income effect (for rice growers), or by affecting food consumption choices through the usual income and substitution effects. Next, the year in which a school was built may not be a good indicator of school quality. A family's decision to send their children to work or school depends on the cost of schooling as well as on its availability. It is possible that free meals available in schools can directly affect health of the children. Therefore the instrument indicating the year the school was built may not be appropriate.

A recent paper by Wolff and Maliki (2008) looks at the impact of child work on the health of school going children in Indonesia for the years 1993, 1996 and 2000. They use data from Indonesian Socio-Economic Surveys. The first treatment variable is a dummy for whether a child works to produce marketable goods and services, and the second is the number of hours worked. The three health indicators used are whether a child has fever, cough, or some other illness. The fourth health variable is a dummy indicating whether a child suffers from at least one of the illnesses while the fifth health outcome records the number of days their school or work was disrupted due to illness. Assuming child work to be exogenous, they get a negative association between health and work. Instrumenting for the child work variable they still find a negative relationship between work and health. The instruments used are the adult employment rate and the number of primary school buildings. The reported F-statistic is greater than 1000 and significant at the 1 percent level. The Sargan overidentification test however rejects the validity of the instruments in most of the outcome equations in 1993 and 1996 while for 2000 the instruments are insignificant for all the health variables. In fact the authors themselves concede that the validity of instruments is of concern. The adult employment rate may affect family's living standard and thus affect health directly. Availability of meals and doctors in schools can also influence health directly and the instruments may actually belong in the model. Moreover as the sample is restricted to school going children, the number of primary schools should be irrelevant in the child work decision. The authors admit that the lack of a strong and valid instrument renders their results weak.

In this paper we address the problem of unavailability of strong and valid instruments in the child work literature. As our data set also does not provide a strong and valid instrument we apply an estimation strategy suggested by Altonji et al. (2005). This estimation technique is useful when there are concerns about the endogeneity of a treatment variable, but valid instruments are not available. We use the 2000 wave of the Indonesian Family and Life Survey (IFLS) data set. We consider six different dependent variables: BMI, a z-score variable for BMI-for-age, a z-score variable for weight-for-age and three indicators of underweight, high weight and intermediate weight status. We estimate the impact of child work on health using the full sample as well as a subsample where we only consider the school going children.

Regression results for both the samples indicate a positive relationship between child work and the health variables: BMI, BMI-for-age and weight-for-age. Probit estimation shows that the indicator for high weight has a positive association while the dummy for underweight has a negative relationship with child work in both the samples. For the intermediate weight group we however obtain a negative relationship with child work.

Next, following Altonji et al. we estimate the effect of child work on health outcomes under the assumption of selection on only observed characteristics affecting child health. In absence of a valid instrument the actual amount of selection on unobserved characteristics affecting child health is unknown. So we calculate (under certain assumptions) how strong this selection on unobservables needs to be (relative to the amount of selection on observables) in order to explain the entire effect of child work estimated under the assumption of only selection on observables. In addition, for the discrete health outcomes, we estimate bivariate probit models constraining the correlation between unobservables that affect health outcomes and work participation to take different values and analyze the causal effect of child work on child health outcomes. Finally, we estimate unconstrained bivariate probit models to provide suggestive evidence on the magnitude and direction of selection on unobservables, identifying the parameters on the basis of the parametric assumptions.

The results obtained are striking. We find that there is evidence of a differential pattern in terms of the selection on unobserved variables across the weight distribution. Our results indicate the existence of a healthy worker selection effect. This implies that the children who are more likely to have high weight also have higher probability of working. Our findings also suggest the existence of an unhealthy worker selection effect. In other words, the least healthy children, captured by the dummy for underweight, are also more likely to work. However, we find evidence of negative selection into work when we consider the children belonging to the intermediate weight category. These outcomes imply that our OLS/probit estimates are biased.

The Altonji et al. methodology provides evidence consistent with a beneficial causal impact of child work on health of the children at the low end of the weight distribution. However, there is substantiation of a negative causal effect of child work on the healthiest children. We also find child work increases the probability of children to be in the intermediate weight category. Whether this effect is beneficial or detrimental is unknown. Finally, our results also show that a small amount of selection on unobserved characteristics affecting child health will be able to explain the total effect of child work on health. Similar results are obtained for our subsample where we focus only on the school going children.

The rest of the paper is structured as follows: Section 2 will describe the econometric methodology while the Indonesian Family Life Survey data will be discussed in Section 3. Section 4 will give the results while Section 5 will conclude.

2 Empirical Methodology

To identify the impact of child work on contemporaneous health we begin with the following empirical model of health determination:

$$y_i = x_i\beta + \tau D_i + \varepsilon_i \tag{1}$$

where y denotes the health status and D denotes the work status of a child, x is a vector of observable attributes of child, while τ is the constant treatment effect. ε captures the effect of unobservable attributes on a child's health. To get a consistent estimate of τ using OLS estimation technique we need child work to be independent, conditional on x, of unobservables that affect health of a child.

The OLS estimator is susceptible to bias from selection on unobservables. So to analyze the sensitivity of our results to any selection on unobservables we follow the procedure developed by Altonji et al. (2005). We need to mention that the methodology suggested by Altonji et al. (2005) does not provide us with point estimates. This method assesses the robustness of the impact of child work estimated under the assumption of random selection (i.e., selection into child work on observables attributes only) to varying degrees of non-random selection (i.e., selection into child labor on unobservable attributes). The nature of the selection process is identified using parametric assumptions.

First, for the three discrete outcomes we use the bivariate probit model to analyze the relationship between child work and binary health outcomes. The model is given by

$$y_i = I(x_i\beta + \tau D_i + \varepsilon_i > 0) \tag{2}$$

$$D_i = I(x_i\lambda + v_i > 0) \tag{3}$$

where I($\cdot$) is an indicator function, y is the binary indicator of child health (underweight, intermediate weight and high weight status), D is the dummy representing child work and $\varepsilon, v \sim N_2(0, 0, 1, 1, \rho)$. ρ indicates the correlation between the unobservables that affect the health outcome and the child work variable. $\rho > 0$ represents a positive selection on unobservables while $\rho < 0$ represents a negative selection on unobservables. The bivariate normality assumption implies that the model is technically identified although there is no exclusion restriction.

Altonji et al. (2005) suggest another method to assess the role of selection bias and this technique can be applied to continuous and discrete outcomes. This methodology assesses how large the selection on unobservables needs to be relative to the selection on observables to fully explain the relationship between child work and health outcomes under the null hypothesis of no average treatment effect.

The normalized amount of selection on unobservables is given by the ratio

$$\frac{E[\varepsilon|D=1] - E[\varepsilon|D=0]}{Var(\varepsilon)} \tag{4}$$

where D is the child work variable and ε captures the unobservables in the outcome equation. The normalized amount of selection on observables is given by the ratio

$$\frac{E[x\beta|D=1] - E[x\beta|D=0]}{Var(x\beta)} \tag{5}$$

where x represents the set of the observable variables in the outcome equation and β represents the corresponding parameter vector.

Let the actual child work variable be defined as

$$D_i = x_i\lambda + \nu_i \tag{6}$$

Substituting (6) in (1) we get

$$y_i = x_i(\beta + \tau\lambda) + (\tau\upsilon_i + \varepsilon_i) \tag{7}$$

The probability limit of the OLS estimator of τ in (7) is given by

$$\begin{aligned} p\lim\tau &= \tau + \frac{Cov(\nu,\varepsilon)}{Var(\nu)} \\ &= \tau + \frac{Var(D)}{Var(\nu)}\{E[\varepsilon|D=1] - E[\varepsilon|D=0]\} \end{aligned} \tag{8}$$

Assuming that the degree of selection on unobservables given by (4) is equal to the degree of selection on observables given by (5) the bias term in (8) can be written as

$$\frac{Cov(\nu,\varepsilon)}{Var(\nu)} = \frac{Var(D)}{Var(\nu)}\{\frac{E[x\beta|D=1] - E[x\beta|D=0]}{Var(x\beta)}Var(\varepsilon)\} \tag{9}$$

β can be consistently estimated from equation (7) using OLS restricting τ to zero. The estimate of the asymptotic bias is obtained using the estimates of β and $Var(\varepsilon)$ and the sample values of $Var(D)$ and $Var(\nu)$ under the assumption of equality of selection on observables and selection on unobservables. Dividing the unconstrained estimate of τ from (7) by the estimate of bias given by equation (9) we get an implied ratio. This implied ratio shows how large the selection on unobservables needs to be relative to the selection on observables to explain the entire treatment effect. A small implied ratio indicates that if selection on unobservables exists, then the treatment effect is highly sensitive to selection on unobservables. According to Altonji et al. (2005) if the set of variables in x is drawn randomly from all the factors affecting child health outcomes and no factor, both observed and unobserved, plays too large a role in affecting child health, then an implied ratio of value less than 1 implies that the estimate is not robust.

3 Data

The data are obtained from the 2000 wave of Indonesian Family Life Survey (IFLS). The IFLS data set which is longitudinal in nature, represents about 83% of the Indonesian population living in 13 of the nation's 26 provinces in 1993. Our analysis is restricted to rural areas only since there is insufficient information for urban areas.

In our study a child is defined as an individual between ages 6-14 years. To measure child health, we use the age (in years), gender, weight and height of each child and construct six measures of child health. These measures are body mass index (BMI), two z-score variables, BMI-for-age and weight-for-age and three indicators for underweight, intermediate weight and high weight status. [2] We define the underweight variable as BMI-for-age below the 5th percentile based on the growth charts from the Center for Disease Control (CDC) while high weight status is defined as BMI-for-age equal or above the 50th percentile. We define the high weight status to indicate children who are at the higher end of the BMI distribution. Given that our focus is not on obesity we chose not to use the overweight category defined by CDC. Instead we use the median as the cutoff point for high weight category since in the context of a developing country these represent the healthier children. Note also that increasing the cutoff point to something beyond the median will only make our result stronger. Another important consideration behind the cutoff points for the different weight categories is to ensure that each group has enough observations. Finally, the range in between the high weight and underweight is termed as intermediate weight. The treatment variable is a dummy equal to 1 if the child has worked in the month preceding the date of interview and zero otherwise. A child may be engaged in income-generating work in a household business and/or paid work outside the household.

We control for a child's age, gender and an interaction term of age and gender as well as their higher orders. To control for the household factors, the following covariates are included in the x vector: BMI of parents, earnings of father and mother, age of mother, logarithm of household food expenditure, dummy for whether the household receives any assistance from the regional or central government, whether the house is self-owned, whether it has electricity, an indicator for not drinking safe water, dummies representing whether the household owns a toilet, has a health card, participates in a health fund and receives any assistance from community. The other household variables are whether the family owns an agricultural land and/or a non-agricultural land. We also include variables indicating whether the floor, the outerwalls and the roof of the house are made of solid materials.

The community variables in the x vector are the number of health centers and integrated health posts in the village, whether the health centers have doctors, nurses, midwives, a pharmacy and check-up rooms. Other community variables included are presence of solid roads, formal and informal credit opportunities, whether there were natural disasters or epidemics in the last year, and whether there was any rice assistance program in the village.

Children with missing age and health variables are dropped from our sample. Missing values for the remaining control variables are imputed and the imputation dummies are included in the control set. The final sample includes 3033 children of which 312 are working. Tables 1A, 1B and 1C provide the summary

[2] BMI is defined as weight(kg)/square of height(metre square). BMI-for-age and weight-for-age are calculated using the "zanthro" command in STATA. The "zanthro" command helps to generate z-scores for anthropometric measures in children according to US or UK reference growth charts. Percentile BMI can be obtained from the z-score variable, BMI-for-age.

statistics. The mean BMI of the working children is 16.71 while that of the non-working children is 15.53. The average age of working children is around 12. Among child workers about 20.5 percent children belong to the high weight category while for the non-working children this number is 15 percent. The percentage of children belonging to the intermediate weight group is higher for non-working children. The mean BMI of the fathers is lower for the child workers. The opposite is true when we consider mothers. When compared with the non-working children, the mean earnings of the father of the child who works is lower. But the household consumption expenditure on food is higher for the families of working children. The households of the child workers are more likely to own agricultural land. Around 94 percentage of the children in the full sample are in school. Hence the descriptive statistics of the sample consisting of school going children is not much different from that of the full sample.

4 Results

4.1 Full Sample

The OLS and probit results for the full sample are presented in Table 2. The OLS regression results show that there exists a positive relationship between child work and the three continuous health outcomes: BMI, BMI-for-age and weight-for-age. However the positive association is significant only for BMI and weight-for-age. Probit results indicate that both underweight and intermediate weight variables have a negative relationship with child work. The association between work and high weight status turns out to be positive. The RESET test results for the continuous outcomes (BMI-for-age, BMI, weight-for-age) and the binary outcomes (underweight and high weight status) reject the null of no omitted variables. The failure of this test implies that non-linear combinations of the explanatory variables have power in explaining the exogenous variable and therefore the model is misspecified. Hence to obtain the causal impact of work on health we move on to Altonji et al. (2005) approach.

We first estimate the bivariate probit model without constraining the value of ρ. We find a negative but significant effect of work on the underweight status while for the intermediate weight status the positive effect of work is significant only at the 10 percent level. There is a negative but insignificant effect of work on high weight status. The estimated values of ρ are 0.810, -0.491 and 0.105 for the underweight, intermediate weight and high weight status respectively.

These estimates of ρ indicate a differential pattern in terms of selection on unobservables across the weight distribution. There is indication of positive selection into work for the healthiest child workers suggesting the healthy worker selection effect. There is also evidence of the unhealthy worker selection effect as the least healthy children are positively selected into work. This selection into work for the children in the two extreme ends of the weight distribution is consistent with the RESET tests for the two extremes. The negative selection into work indicates that children belonging to the intermediate weight group are less likely to work.

We then constrain ρ to take different values and estimate the effect of child work on the health outcome variables without making any formal distributional assumptions. For the high weight and underweight categories we constrain ρ to taking values in the interval [0, 0.5] in increments of 0.1. This will indicate increasingly strong levels of positive selection on unobservables. Next, for the intermediate weight status we constrain ρ to take the values 0, -0.1, -0.2...., -0.5 to indicate increasingly strong levels of negative selection on unobservables. The results given in Table 3 show that if we consider the high weight status as the dependent variable, the positive effect of work disappears when we assume $\rho = 0.1$. It becomes negative and significant at 5 percent level when ρ is increased to 0.2. The negative effect of child work on the intermediate weight status turns positive at $\rho = -0.1$ and the positive effect becomes significant at the 1 percent level when $\rho = -0.2$. The negative effect of work on the underweight status becomes significant at $\rho = 0.1$.

The bivariate probit results indicate that the relationship between child work and health obtained from probit estimation are extremely sensitive to selection on unobservables. For the high weight and intermediate weight status, a modest amount of selection on unobservables, completely eliminates or reverses the results obtained from the probit estimation where we do not consider any selection on unobservables. Even for the underweight case the negative effect becomes significant when we allow a small amount of selection on unobservables.

The causal estimates obtained after constraining ρ thus suggest that child work modestly reduces the probability of being high weight. This implies that the healthiest children who go to work have negative causal impact on their health due to work. There is however evidence of beneficial effect of child work on children who are at the lowest end of the weight distribution as child work reduces the probability of being underweight. Finally, the estimates of the bivariate probit indicates that child work leads to rise in the probability of being in the intermediate weight category. This may or may not be a beneficial impact.

Our findings indicate heterogeneity in the selection of children into work as well as in the suggested impact of work on health. The results suggest that families favor sending children with very healthy and unhealthy unobservables into the labor market. It is easy to see why the healthiest children are sent to work. As their productivity is higher, the healthiest children are possibly sent to work outside for wages or are involved in more demanding jobs at home. For the less healthy children, one possibility is that the families may keep their least healthy children at home and allocate them to less strenuous work on the household farm or business. Thus even while they are working, these weak children are under the scrutiny of the parents, and at the same time contribute positively to household income.

We find that the impact of such labor is beneficial for the children with poor unobservables, but detrimental for children with healthy unobservables. Because child work can affect health via two routes − a positive income effect and a negative direct effect from work. It is worth noting that the latter effect can vary with the type of job the child is engaged in. The results indicate that the income effect dominates (is dominated) for children with poor (healthy) unobservables. One possible explanation for the result can be that light work may not affect the health of the least healthy children to a great degree, but the money saved from not having to hire outside labor results in a higher net income for the parents. This results in the positive income effect outweighing the negative health impact. For the healthiest children however, it is possible that their work is quite demanding and negatively affects their health outweighing the positive income effect.

The results for the second method which account for both the continuous and binary health variables are given in Table 4. The implied ratio is 0.311 when we consider weight-for age as the health indicator. From this it follows that if there is selection on unobservables and if the (normalized) selection on unobservables is 0.311 or 31.1 percent as much as the amount of (normalized) selection on observables, then that would be sufficient to completely explain the positive effect of child work on weight-for-age. Similarly, when the dependent variable is BMI, the implied ratio of 0.11 indicates that if the selection on unobservables is 11 percent as much as the selection on observables then the entire effect of child work can be attributed to selection bias. In this sample the implied ratio is never greater than 0.4. Hence if there is selection on unobservables and if the (normalized) selection on unobservables is at least 0.4 or 40 percent as much as the amount of (normalized) selection on observables, then this would be sufficient to completely explain the entire effect of child work on health. The above results corroborate the findings of the bivariate model indicating that a small amount of selection on unobservables can explain the relationship between child work and health. For the binary outcomes, and the BMI-for-age variable, the estimate of child work is insignificant under the selection on observables. Hence the second approach is irrelevant for these health variables.

4.2 School going children

For the sake of comparison with the existing literature, we also consider the subsample of school going children. Majority of the results from the full sample continue to hold. The OLS and probit results are presented in Table 2. However, here the positive association between work and the continuous health outcomes is significant only for the weight-for-age variable. Like the full sample case, the probit results for the subsample indicate a positive relationship between work and health when we consider the high weight status as the dependent variable. Underweight and intermediate weight variables have negative association with work. The RESET test shows that the null of no omitted variables is rejected in almost all the cases implying misspecification in the model.

Following Altonji et al. (2005), we find that the unconstrained bivariate probit indicates a negative but significant estimate for the underweight status, while for the intermediate weight status the positive effect is significant only at the 10 percent level. There is a negative but insignificant effect of work on the high weight status. The estimated values of ρ indicate evidence of negative selection into work when we consider the intermediate weight status as the dependent variable, and positive selection into work for the other two binary health outcomes. Hence we find a heterogeneous selection process in the subsample as well.

The constrained results are given in Table 3. Much like the full sample case, the bivariate probit results for the subsample suggest that the relationship between child work and binary health variables are extremely sensitive to selection on unobservables. Our results indicate that the healthiest child workers have negative causal impact of work on their health, while there is evidence of beneficial effect of work on the children who are at the lowest end of the weight distribution. There is also indication that child work leads to a rise in the probability of being in the intermediate weight category. So like the full sample we find a heterogeneous impact of child work on health.

The results for the second approach are given in Table 4. We find that if there is selection on unobservables and if the selection on unobservables is 0.268 or 26.8 percent as much as the amount of selection on observables, then that would be sufficient to completely explain the entire positive effect of child work on weight-for-age. In the full sample case this percentage is 31.1. The coefficient on child work variable is insignificant under the selection on observables for the other health outcomes. So this method is irrelevant for those health indicators. Our descriptive statistics have shown that 94 percent of the children in the full sample are in school. This explains why the results obtained for the subsample are very similar to the full sample case. It also assures us that the problem which may arise due to the trade-off between schooling and work will not be a significant issue for our full sample.

To sum up, the results indicate heterogeneity in the selection of children into work as well as in the suggested impact of work on health. The results attest that families select their healthiest and least healthy children to work. There is also indication of negative causal effect of labor on the children with healthy unobservables while there exists beneficial causal impact on children with poor unobservables.

5 Conclusion

Our paper examines the causal impact of child work on the contemporaneous health of children and adds considerably to the previous literature. Given non-random selection into work and lack of strong and valid instruments in our data set, we follow the approach suggested by Altonji et al. (2005). This methodology is particularly useful when there is unavailability of valid instruments to control for the unobserved characteristics, affecting both the treatment and outcome variables. However, one has to keep in mind that using this approach we do not arrive at the point estimates of the effects of child work. The three main results of this paper are: (1) There exists a positive and significant association between child work and two of our health measures, BMI and weight-for-age. (2) Our results provide evidence of positive selection into work for children belonging to the high weight category supporting the theory of healthy worker selection effect. At the same time, our results indicate evidence of the unhealthy worker selection effect for the least healthy children. However, there is indication of negative selection when we consider the intermediate weight status as the dependent variable. (3) Finally, allowing for a small amount of selection on unobservables renders the initial results obtained from the OLS and probit estimation invalid. There is evidence of negative and significant causal impact of work on high weight status while there is indication of beneficial impact on the children at the lower end of the weight distribution. The results also indicate that child work increases the probability of being in the intermediate weight range. For the children in this category, it can not be ascertained whether this positive effect is helpful or harmful. The conclusions obtained in this paper thus give us an indication of heterogeneity in the selection of children into work, and also in the suggested impact of child work on the health of the children. Future work should take into account the type of work done by children with different attributes and whether the health effects vary by the type of work.

References

[1] Altonji, J.G., Elder,T.E. and Taber C.R. (2005), Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, *Journal of Political Economy*, 113, 151-184.

[2] Appleton, S. and Song, L. (1999), Income and human development at the household level: evidence from six countries, *Mimeo,* University of Oxford.

[3] Beegle, Kathleen, Dehejia, Rajeev H and Gatti, Roberta (2005), Why Should We Care About Child Labor? The Education, Labor Market and Health Consequences of Child Labor, *The World Bank Policy Research Working Paper Series,* Number 3479.

[4] Guiffrida, A., Iunes, R.F. and Savedoff, W.D. (2001), Health and poverty in Brazil: Estimation by structural equation model with latent variables, *Mimeo,* Washington D.C., Inter-American Development Bank.

[5] ILO (2005), Facts on Child labor.

[6] Kassouf, A.L., McKee, M. and Mossialos, E. (2001), Early entrance to the job market and its effects on adult health: Evidence from Brazil, *Health Policy and Planning.* 16(1): 21-28.

[7] Maliki, M. and Wolff, F.C. (2008), Evidence on the impact of child labor on child health in Indonesia, 1993-2000, *Economics and Human Biology,* 6, 143-169.

[8] Millimet, D., Tchernis, R. and Husain, M. (2008), School Nutrition Programs and the Incidence of Childhood Obesity, Forthcoming, *Journal of Human Resources.*

[9] O'Donnell, O., Rosati, F.C. and Van Doorslaer, E. (2002), Child labour and health: Evidence and Research issues, *Understanding Children's Work Discussion Pape*r, Florence, Innocenti Research Centre.

[10] O'Donnell O., Van Doorsaler, E. and Rosati, F. (2005), Health Effects of Children's Work: Evidence from Vietnam, *Journal of Population Economics*, 18, 437-467.

[11] Rosati, Furio C. and Straub, R. (2006), Does work during childhood affect adult's health? An analysis for Guatemala, *UCW Working Paper* 10, Understanding Children's Work (UCW Project).

[12] Satyanarayana, K., Krishna, T.P. and Rao, B.S. (1986), The effect of early childhood undernutrition and child labour on the growth and adult nutritional status of rural Indian boys around Hyderabad, *Human Nutrition and Clinical Nutrition* 40C: 131-9.

[13] Smith, J. (1999), Healthy bodies and thick wallets: the dual relation between health and economic status, *Journal of Economic Perspectives*, 13(2), 145-166.

[14] Steckel, R. (1995), Stature and the standard of living, *Journal of Economic Literature*, 33: 1903-40.

[15] Stock, J.H. and Yogo, M. (2005), Testing for Weak Instruments in Linear IV Regression, in Andrews, D. W. K. and Stock, J. H. (eds.), *Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg,* Cambridge University Press, New York, NY (Chapter 5).

[16] Strauss, J and Thomas, D. (1998), Health, Nutrition and Economic Development, *Journal of Economic Literature,* 36, 766-817.

[17] Weil, D. (2005), Accounting for the Effect of Health on Economic Growth, *NBER Working Paper* No. 11455.

Table 1A. Summary Statistics

Variable	Full Sample		School going children	
	Mean	SD	Mean	SD
Child Work(1=yes)	0.103	0.304	0.090	0.286
Underweight(1=yes)	0.227	0.419	0.228	0.420
High Weight(1=yes)	0.155	0.362	0.151	0.358
Intermediate Weight(1=yes)	0.618	0.486	0.621	0.485
Body Mass Index	15.651	2.243	15.552	2.155
Age(years)	10.283	2.486	10.140	2.441
Gender(1=male)	0.505	0.500	0.506	0.500
Log(household expenditure on food)	9.763	0.608	9.764	0.611
Time taken to go to heath center(min)	24.857	73.905	24.921	75.740
Time taken to go to intergrated health posts(min)	12.036	69.218	12.090	70.989
Household receiving assistance from govt/non-govt(1=yes)	0.029	0.168	0.028	0.164
House self-owned(1=yes)	0.932	0.252	0.932	0.253
House uses electricity(1=yes)	0.828	0.378	0.838	0.368
Household does not drink safe water(1=yes)	0.088	0.284	0.086	0.280
House has a sanitory toilet(1=yes)	0.510	0.500	0.520	0.500
Household has a health card(1=yes)	0.230	0.421	0.232	0.422
Household participated in health fund(1=yes)	0.073	0.260	0.074	0.262
Household received assistance from community(1=yes)	0.024	0.153	0.024	0.153
Floors of house not solid(1=yes)	0.225	0.417	0.219	0.414
Outer walls of house not solid(1=yes)	0.146	0.353	0.142	0.349
Roof of house not solid(1=yes)	0.051	0.220	0.048	0.214
Household owns land for farming(1=yes)	0.583	0.493	0.588	0.492
Household owns non-agricultural land(1=yes)	0.171	0.377	0.172	0.378
BMI of father	21.119	2.418	21.142	2.434
BMI of mother	22.546	3.650	22.589	3.690
Earnings of father	3532809	5183954	3576159	5298973
Earnings of mother	1044012	2255141	1059384	2290546
Age of mother	36.479	6.649	36.376	6.600
Presence of financial institution in the village (1=yes)	0.182	0.990	0.189	1.015
Village connected by a land road(1=yes)	0.978	0.148	0.980	0.141
Occurance of natural disaster in 1999	0.050	0.210	0.050	0.219
Occurance of epidemic in 1999	0.011	0.102	0.010	0.098
Rice Assistance in 1999	0.945	0.228	0.944	0.231
Supplementary food program in 1998	0.935	0.246	0.933	0.250
Toilet Facilities in health center(1=yes)	0.989	0.104	0.989	0.105
Doctors Present in health center(1=yes)	0.922	0.268	0.922	0.267
Midwives present(1=yes)	0.956	0.204	0.957	0.203
Village midwivespresent(1=yes)	0.843	0.363	0.842	0.365
# health centers	1.959	0.198	1.959	0.199
# integrated health posts	2.960	0.193	2.962	0.190

Notes: N=3033 (full sample); N=2877 (attend school). .

Table 1B. Summary Statistics

Variable	Full Sample Working		Full Sample Not Working	
	Mean	SD	Mean	SD
Child Work(1=yes)	1.000	0.000	0.000	0.000
Underweight(1=yes)	0.224	0.418	0.227	0.419
High Weight(1=yes)	0.205	0.404	0.150	0.357
Intermediate Weight(1=yes)	0.571	0.496	0.624	0.484
Body Mass Index	16.714	2.715	15.529	2.149
Age(years)	12.199	1.731	10.063	2.465
Gender(1=male)	0.497	0.501	0.506	0.500
Log(household expenditure on food)	9.808	0.636	9.758	0.605
Time taken to go to heath center(min)	26.294	55.638	24.692	75.725
Time taken to go to intergrated health posts(min)	16.139	101.964	11.566	64.419
Household receiving assistance from govt/non-govt(1=yes)	0.035	0.185	0.028	0.166
House self-owned(1=yes)	0.965	0.185	0.928	0.258
House uses electricity(1=yes)	0.769	0.422	0.835	0.372
Household does not drink safe water(1=yes)	0.051	0.221	0.093	0.290
House has a sanitory toilet(1=yes)	0.532	0.500	0.507	0.500
Household has a health card(1=yes)	0.285	0.453	0.224	0.417
Household participated in health fund(1=yes)	0.083	0.277	0.072	0.258
Household received assistance from community(1=yes)	0.038	0.193	0.022	0.148
Floors of house not solid(1=yes)	0.266	0.443	0.220	0.414
Outer walls of house not solid(1=yes)	0.167	0.373	0.143	0.350
Roof of house not solid(1=yes)	0.071	0.256	0.049	0.216
Household owns land for farming(1=yes)	0.679	0.467	0.572	0.495
Household owns non-agricultural land(1=yes)	0.192	0.395	0.169	0.375
BMI of father	20.862	2.073	21.149	2.454
BMI of mother	22.571	4.901	22.543	3.479
Earnings of father	2933991	3204542	3601471	5360589
Earnings of mother	1298253	2910594	1014860	2166132
Age of mother	38.604	6.405	36.235	6.633
Presence of financial institution in the village (1=yes)	0.292	1.352	0.169	0.939
Village connected by a land road(1=yes)	0.978	0.148	0.978	0.148
Occurance of natural disaster in 1999	0.058	0.233	0.049	0.216
Occurance of epidemic in 1999	0.019	0.138	0.010	0.097
Rice Assistance in 1999	0.942	0.234	0.945	0.228
Supplementary food program in 1998	0.910	0.286	0.938	0.241
Toilet Facilities in health center(1=yes)	0.994	0.080	0.989	0.106
Doctors Present in health center(1=yes)	0.926	0.262	0.922	0.269
Midwives present(1=yes)	0.968	0.176	0.955	0.207
Village midwivespresent(1=yes)	0.843	0.364	0.843	0.363
# health centers	1.932	0.251	1.962	0.191
# integrated health posts	2.968	0.176	2.960	0.191

Notes: N=312 (work); N=2721 (no work).

Table 1C. Summary Statistics

Variable	Children are in School & Work		Children are in School & No Work	
	Mean	SD	Mean	SD
Child Work(1=yes)	1.000	0.000	0.000	0.000
Underweight(1=yes)	0.232	0.423	0.228	0.420
High Weight(1=yes)	0.197	0.398	0.146	0.353
Intermediate Weight(1=yes)	0.571	0.496	0.626	0.484
Body Mass Index	16.446	2.505	15.464	2.097
Age(years)	11.981	1.762	9.958	2.424
Gender(1=male)	0.506	0.501	0.506	0.500
Log(household expenditure on food)	9.822	0.652	9.759	0.606
Time taken to go to heath center(min)	26.737	60.577	24.741	77.085
Time taken to go to intergrated health posts(min)	16.717	111.684	11.632	65.621
Household receiving assistance from govt/non-govt(1=yes)	0.027	0.162	0.028	0.165
House self-owned(1=yes)	0.965	0.183	0.928	0.258
House uses electricity(1=yes)	0.807	0.395	0.841	0.365
Household does not drink safe water(1=yes)	0.050	0.219	0.089	0.285
House has a sanitory toilet(1=yes)	0.568	0.496	0.515	0.500
Household has a health card(1=yes)	0.301	0.460	0.225	0.418
Household participated in health fund(1=yes)	0.093	0.291	0.072	0.259
Household received assistance from community(1=yes)	0.039	0.193	0.223	0.148
Floors of house not solid(1=yes)	0.255	0.437	0.215	0.411
Outer walls of house not solid(1=yes)	0.162	0.369	0.140	0.347
Roof of house not solid(1=yes)	0.062	0.241	0.047	0.212
Household owns land for farming(1=yes)	0.703	0.458	0.576	0.494
Household owns non-agricultural land(1=yes)	0.193	0.395	0.170	0.376
BMI of father	20.933	2.111	21.163	2.463
BMI of mother	22.764	5.223	22.572	3.503
Earnings of father	3022296	3411367	3630953	5447697
Earnings of mother	1399634	3148811	1025723	2185355
Age of mother	38.545	6.310	36.161	6.591
Presence of financial institution in the village (1=yes)	0.336	1.476	0.174	0.956
Village connected by a land road(1=yes)	0.981	0.138	0.980	0.141
Occurance of natural disaster in 1999	0.062	0.241	0.049	0.216
Occurance of epidemic in 1999	0.015	0.124	0.009	0.095
Rice Assistance in 1999	0.934	0.248	0.945	0.229
Supplementary food program in 1998	0.900	0.310	0.937	0.244
Toilet Facilities in health center(1=yes)	0.996	0.062	0.988	0.108
Doctors Present in health center(1=yes)	0.931	0.255	0.922	0.269
Midwives present(1=yes)	0.973	0.162	0.955	0.207
Village midwivespresent(1=yes)	0.838	0.369	0.842	0.365
# health centers	1.926	0.261	1.962	0.192
# integrated health posts	2.981	0.138	2.961	0.194

Notes: N=259 (work & school); N=2618 (no work but attend school).

Table 2. OLS/Probit Results

	BMI	BMI-for-age	Weight-for-age	Probability of being underweight	Probability of being intermediate weight	Probability of being high weight
Full Sample						
Child Work	0.279†	0.098	0.242*	-0.056	-0.064	0.149
	(0.144)	(0.074)	(0.076)	(0.091)	(0.081)	(0.095)
School going children						
Child Work	0.140	0.028	0.173†	-0.028	-0.080	0.141
	(0.141)	(0.080)	(0.081)	(0.098)	(0.087)	(0.104)

NOTES: ‡ p<0.10, † p<0.05, * p<0.01. Standard errors in parentheses.

Table 3. Sensitivity Analysis:Bivariate Probit Results with Different Assumptions Regarding Correlation Among the Disturbances

	Correlation of the Disturbances											
	Full Sample						School going children					
	ρ = 0	ρ = 0.1	ρ = 0.2	ρ = 0.3	ρ = 0.4	ρ = 0.5	ρ = 0	ρ = 0.1	ρ = 0.2	ρ = 0.3	ρ = 0.4	ρ = 0.5
A. Probability of being Underweight												
Child work	-0.056	-0.236*	-0.413*	-0.587*	-0.757*	-0.922*	-0.028	-0.211†	-0.391*	-0.568*	-0.742*	-0.911*
	(0.091)	(0.091)	(0.090)	(0.088)	(0.086)	(0.083)	(0.099)	(0.098)	(0.097)	(0.096)	(0.093)	(0.090)
B. Probability of being High Weight												
Child Work	0.149	-0.031	-0.209†	-0.384*	-0.557*	-0.725*	0.141	-0.042	-0.222†	-0.399*	-0.572*	-0.742*
	(0.095)	(0.095)	(0.094)	(0.092)	(0.090)	(0.086)	(0.104)	(0.104)	(0.103)	(0.101)	(0.098)	(0.095)

	Correlation of Disturbances											
	Full Sample						School going children					
	ρ = 0	ρ = -0.1	ρ = -0.2	ρ = -0.3	ρ = -0.4	ρ = -0.5	ρ =0	ρ = -0.1	ρ =- 0.2	ρ = -0.3	ρ =-0.4	ρ =- 0.5
C. Probability of being Intermediate Weight												
Child Work	-0.064	0.118	0.298*	0.478*	0.656*	0.832*	-0.080	0.104	0.287*	0.470*	0.651*	0.830*
	(0.081)	(0.081)	(0.080)	(0.078)	(0.076)	(0.074)	(0.088)	(0.087)	(0.087)	(0.085)	(0.083)	(0.080)

NOTES: ‡ p<0.10, † p<0.05, * p<0.01. Standard errors in parentheses.

Table 4. Sensitivity Analysis: Amount of Selection on Unobservables Relative to Selection on Observables Required to Attribute the Entire Child Work Effect to Selection Bias

	Full sample			School going children		
	Cov(ε,ν)÷ Var(ν)	τ	Implied Ratio	Cov(ε,ν)÷ Var(ν)	τ	Implied Ratio
BMI	2.529	0.279	0.110	2.343	0.140	0.060
		(0.144)			(0.141)	
BMI-for-age	0.352	0.098	0.277	0.302	0.028	0.093
		(0.074)			(0.080)	
Weight-for-age	0.779	0.242	0.311	0.644	0.173	0.268
		(0.076)			(0.081)	
Underweight	0.127	-0.012	-0.090	0.090	-0.003	-0.030
		(0.026)			(0.028)	
Intermediate	-0.889	-0.026	0.029	-0.747	-0.032	0.043
		(0.031)			(0.033)	
High	0.317	0.038	0.118	0.275	0.035	0.127
		(0.024)			(0.026)	

NOTES: Standard errors in parentheses. Cov(ε,ν)/Var(ν) refers to the asymptotic bias of the unconstrained estimate under the assumption of equal (normalized) selection on observables and unobservables.
τ refers to the unconstrained estimate of child work. The implied ratio is the latter divided by the former.
See Table 2 and text for details.