ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208269
New Haven, Connecticut 06520-8269

CENTER DISCUSSION PAPER NO.810

EARNINGS EFFECTS OF HOUSEHOLD INVESTMENT IN HEALTH IN COLOMBIA

Rocio Ribero
Yale University

November 1999

Earnings effects of household investment in health in Colombia[1]

Rocio Ribero[2]

Economic Growth Center
Yale University

New Haven, November, 1999

Abstract

This study considers the links between primary indicators of health and individual labor productivity in Colombia and explores how additional public expenditures on health may improve individuals' health. Sample statistics show that illness occurs more frequently for women than for men, for less educated than for more educated, for rural than urban residents, and for older individuals. The well educated are considerably taller than those without schooling (6 cm. for males and 4 cm. for females).

The empirical evidence confirms that health indicators are related to individual earnings in Colombia. A Mincerian log-earnings equation that includes health indicators as a form of human capital in addition to schooling is specified. When the morbidity variable is treated as endogenous and measured with error and the model is estimated by instrumental variables [IV], it becomes significant and has the expected negative sign. Controlling for age, education, sector of employment, gender and geographic location, an increase by 50% on the average number of days an individual was ill and unable to do his ordinary activities in the last month would imply reductions in labor earnings of 11% for urban males, 8% for urban females, 13% for rural males and 7% for rural females. The estimations with height show a positive sign and high significance even without the IV correction, but the coefficients increased with IV methods by an order of magnitude. Having one more centimeter of stature would increase urban female earnings by 4.7% and urban male earnings by 12%. Individual's wealth and living in a community with better health provision indicators are linked with better health outcomes. An analysis of the returns to schooling shows that schooling captures part of the effect of health on productivity when the health indicator is not included in the Mincer equation.

JEL classification: I12, J00, J24
Keywords: Earnings, Height and Health Indicators

---

[1] The paper is based on a project financed by the Inter American Development Bank, which report is in Ribero and Nuñez (1999). I am thankful to T. Paul Schultz, William Savedoff, Robert Evanson, Ann Stevens, Michael Boozer and participants at the workshops in CEDE of Universidad de los Andes and Economic Growth Center at Yale University for useful comments and suggestions. I am pleased to acknowledge the support of the Rockefeller Foundation Post-Doctoral Fellowship at Yale University during which I finished this paper.

[2] Professor Universidad de los Andes - CEDE Researcher - Post Doctoral Fellow Economic Growth Center - Yale University.

1. Introduction

This study considers the links between primary indicators of health and individual labor productivity in Colombia and explores how additional public expenditures on health may improve individuals' health. The effects of good health as well as the determinants of health outcomes are studied at the individual level. Rural-urban[3] and gender aspects are considered separately. Household resource allocation and consumption decisions determine nutritional status and health of children and adults within the household. These decisions influence adult anthropometric measures such as height and the patterns of illness and disability.

The framework set up by Mincer (1974) is enriched to allow for additional forms of human capital besides schooling. Schultz (1997) analyzes how state and family investments influence the formation of reproducible human capital and how these in turn affect labor earnings and growth. He finds that adult height is an important determinant of adult productivity, that it emerges as inversely correlated to chronic health problems among the middle aged and elderly, and that it is inversely related to mortality and directly related to length of productive life. Fogel (1994) also finds that height and body mass index [BMI] are related to male mortality at later ages and to chronic diseases at ages between 20 and 50. This study confirms that height is positively related to individual earnings in Colombia.

According to Schultz and Tansel (1997), most studies that measure morbidity focus on chronic disabilities among the elderly in high income countries. Other studies analyze the productivity effects of nutrition in developing countries (Behrman, 1993, Deolalikar, 1988, Behrman and Deolalikar, 1988, Sahn and Alderman, 1988) and an extensive literature focuses on child morbidity and malnutrition effects for children (Rosenzweig and Wolpin, 1988, Rosenzweig and Schultz, 1982, 1983). Schultz (1984) analyzes the relation between child mortality and public program interventions. Adult health status measures, such as height, reflect cumulative health, early childhood conditions, and nutrition investments undertaken by the parents of the individual (Strauss and Thomas, 1995, Martorell and Habicht, 1986). In addition, changes in height over time may be attributed to changes in reproducible human capital investments or in disease environments (Fogel, 1994). Thomas and Strauss (1997) find that height has a large and significant effect on wages for males and females in Brazil, even after controlling for education. They find that relative to the returns to education without controlling for health, the estimated returns to education with health controls were 45% smaller for literate men and 30% smaller for men with secondary education or more. Schultz (1996) finds that the estimated wage returns to schooling are reduced between 10% and 20% with the addition of three other human capital inputs in the regression: migration, BMI and height. This paper finds that the returns to education do not change much when the health variables are included in the earnings equation as raw variables, but they do change when the endogeneity of the health indicators is taken into account. The returns to schooling drop 9% for males and 5% for females when the instrumental variable for number of days disabled in the last month is included in the model. When the instrumental variable for height is included in the model, the schooling coefficient drops largely, by 44% for urban males and by 12% for urban females.

This study finds significant, large and negative effects of number of days disabled (i.e. unable to work or perform daily tasks) in the last month on wages of men and women in Colombia. It is found that a man who works in urban areas and has been disabled one day in the last month is expected to perceive labor earnings that are 55% smaller than the wages of a man who was healthy in the previous month. The detrimental effect of bad health is not as large for men in rural areas (49%) or for females (36% in urban areas and 23% in rural areas).

The effects of height on wages are found to be significant and positive, a result documented in other studies such as Schultz (1996). In Colombia a taller man receives hourly earnings 12% higher per centimeter and a woman receives hourly earnings 4.7% higher per centimeter. These effects are in line with those found in Ghana, where a one centimeter increase is associated with a 5.7% wage gain for males and 7.5% for females, holding constant for BMI and migration.

The paper is divided in five sections. Section 2 describes the data sources as well as the health indicators. Section 3 includes the descriptive statistics of the data. Section 4 discusses the empirical specification and estimation of the models of hourly earnings, reduced forms for health indicators

---

[3] The patterns of illness in Colombia differ between rural and urban areas. Communicable diseases occur more frequently in rural areas. In 1993 approximately 30% of the population was located in rural areas (UNDP,1998).

and hourly earnings with instrumental variables for health indicators. Section 5 contains the main conclusions from the analysis.

2. The data

The data are taken from two major household surveys conducted by DANE.[4] The first one is "Encuesta de Caracterización Socioeconómica" [CASEN], a national survey collected in 1993, that contains specific modules on health, education and child mortality. It includes 35,250 records of individuals with positive wages or earnings between 18 and 70 years old, of which 64% are male and 74% live in urban areas. The health indicator used from this survey is the actual number of days of work lost because of an illness or health problem in the last month[5] (as reported by the individual).

As reported in other studies (Murrugarra and Valdivia, 1999, Cortez, 1999), the health indicator "number of days disabled in the last month" exhibits some weaknesses. On one hand it may be subjective or it may suffer from recall errors from the part of the respondent, as well as measurement errors from the part of the interviewer. However, the fact that this question is the third one after a sequence of related questions may properly filter these errors. On the other hand, the health indicator may be correlated with the characteristics of the labor contracts and the formality of the job. Persons with a job that provides social security may be more likely to take days disabled than individuals who are self-employed and uncovered by social security, for whom it may be more costly not to attend their jobs. However, the flexibility of informal jobs may allow the person to take more days of disability, work at home or be replaced by some relative or friend in the job. This factor is taken into account in the models of this paper, where a dummy variable for being a salaried worker is included.

[4] Departamento Administrativo Nacional de Estadística collects national statistical information for the Colombian government.

[5] The sequence of related questions in the survey is as follows: 1) During the last month did you have any illness, accident, and dental problem or health problem? (yes or no) 2) During the last month did you not go to work or did not do your ordinary activities because of the illness or health problem mentioned above? (yes or no) 3) For how many days during the last month were you unable to work or did you stay in bed because of the illness or health problem mentioned above? (number of days disabled). The question used here in the third one.

The second major survey used is the urban part of the "Encuesta Nacional de Hogares - Etapa 74" [ENH-91], collected in 1991, which covers Bogotá, Cali, Medellín, Barranquilla, Bucaramanga, Manizales, Pasto, Cúcuta, Pereira, Ibagué, Montería and surrounding metropolitan areas of these cities.[6] No other household survey in Colombia includes the height of persons,[7] which is used here as the adult health outcome. The sample age range is restricted to those between 25 and 55 years old. This restriction is made because after approximately age 55 individuals may shrink, not necessarily due to their childhood nutritional status and before age 25 the individuals may still be growing and may not have reached yet their adult height. From all persons who earn positive wages or earnings, those with unreasonable heights (less than 135 cm.) were excluded,[8] leaving a working sample of 18,908 adults. The productivity measure used in both data sets is hourly labor earnings.[9]

Additional sources were consulted to obtain community characteristics that were merged with the individual household survey data. The Ministry of Health provided the number of hospitals available in each municipality and information on vaccination programs and public health expenditures in each municipality. The Instituto Geográfico Agustín Codazzi provided geographic and weather information of the municipalities.

The surveys used here do not provide sufficient information to describe or explain migration.[10] It must be assumed, therefore, that the place of residence of the individual is exogenous, even though people may have migrated after making health investments and moved to a specific area or community due to the local conditions. This may introduce bias in our estimates (Rosenzweig and Wolpin, 1988).

---

[6] These are the eleven major cities of the country, representing close to 40% of the entire population of the country and about 70% of the urban population. The smallest of these cities at the time of the survey had at least two hundred thousand people.

[7] This survey also includes a rural area, but for the rural area the height of individuals is not collected. If possible, the height is measured with a tape measure. When the person is not present, the height estimate is based on the respondent's height and the height of some of the persons who are present. The survey does not indicate whether a person's height was actually measured or it was estimated.

[8] The number of observations dropped at this stage was approximately 7% of the total sample.

[9] Hourly earnings are constructed from adding wages (for wage earners) and labor earnings (for self-employed) from main activity, and divide the sum by the number of hours worked per month. The number of hours worked per month are the number of hours normally worked per week multiplied by four.

[10] In the ENH-91 the municipio of birth is known only for those born in the same municipio of the survey. CASEN-93 gives the municipality of birth, but the set of health related environmental characteristics is not available for all the municipalities where the individuals may be born.

Domestic servants in general receive food and shelter for their work and earn low cash wages. Since both effects cannot be disentangled with the surveys, domestic servants are excluded from both samples.[11] Given that the goal of the paper is to estimate labor productivity effects of health, only individuals with positive wages or earnings are used for the figures and descriptive statistics of Section 3 and for the estimation of the empirical models of Section 4.[12]

3. Descriptive statistics of the data

For the estimation of the models two samples were consulted. The urban sample obtained from ENH-91 has 62% males and 38% females,[13] 15% of the individuals have less than complete primary school and 16% have more than complete high school. The national sample obtained from CASEN is 67% male and 33% female, only 8% of the individuals have more than high school, 49% have partial or complete primary schooling and 9% have zero years of schooling. In this case, 82% of the sample are located in urban areas. The differences in human capital accumulation between rural and urban areas are large. In urban areas the mean schooling is 7.3 years whereas in rural areas the average schooling is 3.6 years. Similarly, the average number of days disabled in the last month in rural areas is 0.45 and in urban areas is 0.33. The main characteristics of the samples and health indicators are reported in Tables 1 and 2.

Table 1 shows the patterns of illness for the sample of workers between 18 and 70 years old. It is clear that that illness, measured by the average number of days disabled (i.e. unable to work or perform daily tasks) in the last month,[14] is more frequent for women than for men at all ages and all education levels. The mean number of days disabled in the last month increases monotonically with age from 0.25 for individuals between 18 and 24 years old to 0.72 for individuals between 60 and

[11] In CASEN-93, domestic servants are 4.6% of the occupied labor force. In ENH-91, domestic servants are 5.3% of the occupied labor force.

[12] The paper takes account of the sample selection bias that arises when only the sample of workers is considered. The details of this procedure are described later, but since this correction did not generate significant differences on the health variables estimates, the results are not reported.

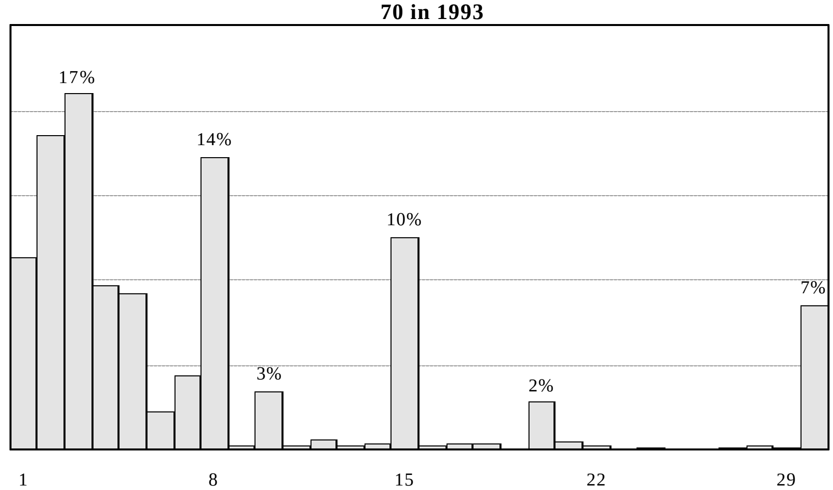
[13] The disparity between the numbers of males and females in the samples is due to the fact that only workers are considered, and the female labor force participation rate in Colombia is approximately 65% of the male labor force participation rate.

[14] All individuals who have not been disabled in the last month have a value of zero in this variable.

70 years old. The largest jump occurs at the oldest ages (between 45-59 and 60-70 years old). Illness occurs more frequently among rural than urban residents at all levels of education (except higher education).[15] The number of days disabled diminishes monotonically with education.

Figure 1 shows the histogram of number of days disabled in the last month for the population with positive days disabled. The bulk of this sample (78%) has fewer than ten days of disability, 10% have 15 days disabled,[16] and 7% of the sample have been disabled for the entire past month. The shape of this histogram is very similar to the one reported in Perú (Cortez, 1999).

Figure 1. Histogram of number of days disabled in last month-Workers aged 18-70 in 1993



Source: CASEN

In the survey, the reference period for the number of days disabled is the last month, therefore the variable is truncated at 30 days disabled. This implies that all individuals with disability periods that last 30 days or more are grouped in this category. The descriptive statistics shown in Table A1 in the Appendix confirm that the group of individuals with 30 days disabled are very different from the rest of the population. This sample contains individuals that are in general older, less educated and more informal than the samples of workers with less than 30 days of disability or with zero

[15] Some cells in Table 1 have very few observations such as the cells of rural residents with higher education, and the cell of individuals 60-70 years old with higher education.

[16] The peaks at 8 and 15 days disabled may be due to the fact that the answers of one or two weeks are rounded and in Colombia are commonly associated with 8 and 15 days, respectively.

days of disability. On average these individuals have more than twice the non labor incomes and report higher hourly wages or earnings than the rest of the population.

Given that approximately 11% of the Colombian population suffers from chronic illnesses (according to Encuesta de Condiciones de Vida, a survey conducted in 1997), the individuals that report 30 days disabled in CASEN most likely suffer from chronic illnesses or are recuperating from major accidents.[17] Since the nature of chronic disabilities or accidents may be different than the nature of "sporadic" disability, the underlying reasons for the two types of disability might also be different. Consequently, the 7% of the population that has been disabled for 30 days has been excluded from the sample. This selection does not alter substantially the results of this paper, as shown in Table A4 in the Appendix.[18]

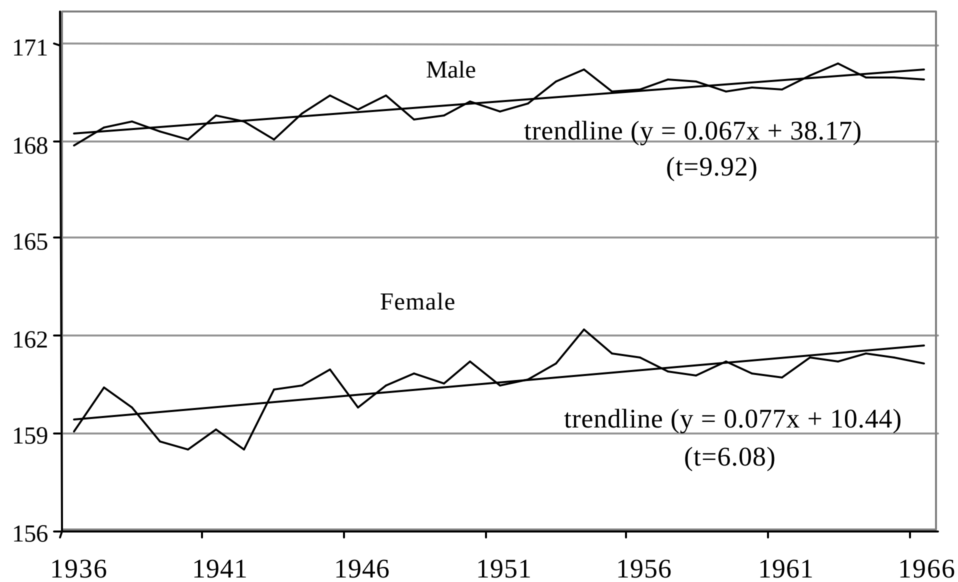
The patterns of height summarized in Table 2 refer to the urban workers between 25 and 55 years old. They indicate that between young (25-34) and old (45-55) age groups, women have gained 2.41 cm. and men have gained 1.98 cm. The best educated males have a 5.81 cm. advantage over the zero year educated, although this result mixes age and class. For females the advantage is of 3.97 cm. Within the education groups and across ages (between youngest and oldest age groups) there is less than 1 cm. gain, except for the group with zero years of schooling, where there was a 1.7 cm. gain between the cohorts of 35-44 and 45-55. Between zero and more than twelve years of schooling and within age groups, the gap has declined from 4.85 cm. for the oldest to 2.79 cm. for the youngest among females. For males the gap across education levels and within age cohorts has declined less but the size of the gap is larger (from 5.92 cm. for the oldest to 5.29 cm. for the youngest). This means that for males the differences in height between education groups tend to disappear in a slower fashion than for females, a fact that may be related to recent increases in overall female education.

Figure 2 shows the average heights for the workers born between 1936 and 1966 (25-55 years old in 1991) in relation to their dates of birth. The increasing trend in height is similar in shape and size to the one observed in Brazil by Thomas and Strauss (1997). The slope in the trend-line is steeper

[17] Unfortunately chronic illness or chronic disability is not reported in CASEN.
[18] This Table should be consulted after reading the paper.

for females than for males (0.77 cm. per decade for females and 0.67 cm. per decade for males). The height gains per decade in Colombia over the entire population aged 25-55 in 1991 are 0.71 cm. for urban women and 0.65 cm. for urban men. The greater slope for workers than for the entire population may indicate that the urban labor market is selecting younger individuals into employment who have higher childhood nutritional levels.

Figure 2: Mean height by year of birth - Workers aged 25-55 in 1991



Trendlines calculated with OLS. T-statistics of coefficient of year in brackets. Curves are not smoothed.

Source: ENH-91

Strauss and Thomas (1998) show that in the United States the mean male stature increased 1.25 cm. per decade between 1910 and 1950. The relative figure in Vietnam was 1.05 cm. per decade and in Brazil 0.77 cm. per decade. Fogel (1994) reports that in Sweden, between the third quarter of the last century and the third quarter of this century, mean male height increased 0.81 cm. per decade and in France 0.64 cm. per decade. For the same period, the increase per decade in male height in Norway was 0.57 cm. and 1.07 cm. in Denmark. Schultz (1996) reports that the gain of height per decade has been almost 1.33 cm. for men and 1 cm. for women in Cote d'Ivoire, and 0.66 cm. for men and 0.33 cm. for women in Ghana. Although the figures of height gains across countries found in the literature are not strictly comparable because they have been taken in different periods of time, they illustrate that height changes in Colombia are similar to that occurring in other countries.

4. Estimation of productivity of health investments

4.1 Hourly earnings equations

An earnings function is estimated of the type:

$$\log (w_i) = a_1 + S\ b_{1j} X_{ji} + S\ c_{1k} C_{ki} + S\ d_{1h} H_{hi} + e_{1i} \qquad (1)$$

where w_i is the individual's hourly wage, X_{ji} contains only exogenous endowments that are not modified by the individual or family, C_{ki} are reproducible forms of human capital, and H_{hi} are the health status indicators. In this section, the health status indicators are assumed to be exogenous to the hourly earnings function and not correlated with the error term e_{1i} in equation (1). The parameters a_1, b_1, c_1 and d_1 are estimated, the error term e_{1i} is assumed to be zero mean independently distributed, i refers to individuals, and j, k and h refer to the specific variables in the sets denoted by X, C and H respectively. The sample includes wage earners and also the non-wage workers with positive earnings and hours worked.

Among the exogenous endowments X, age and age squared are included. The variable in C is the number of years of schooling. Although a dummy variable for recent migration was initially included in set C, this did not affect substantially the coefficients of health or education in any manner and therefore these results are not reported.[19] In the set of health status indicators H, two variables are considered in separate regressions:

1. Number of days that the person was sick and unable to work or perform regular tasks in the last month.[20]
2. Height of individual (measured in centimeters).

[19] The variable is equal to one if the person lives in a different place from where she did five years before the survey, and zero otherwise. This variable is available only from CASEN-93.
[20] The persons who were not prevented from engaging in their regular activity in the last month had a value of zero for this variable.

The “number of days disabled in the last month” uses the threshold of inability to work to make the sickness less subjective, and adds the information on how long the individual is incapacitated, up to 30 days. On the other hand, height for adults is used as an indicator of child nutritional status, exposures to diseases and variation in other environmental health factors (Schultz, 1997).

In the models with height, a dummy variable equal to one if the person was born in rural areas is included as an additional exogenous endowment. Unfortunately, this variable is not available in the data set containing the number of days disabled in the last month. The model was estimated separately for wage earners and self-employed, without uncovering substantial differences. These two groups are distinguished by a dummy variable equal to one for wage earners and zero otherwise, under the assumption that it is exogenously determined. Although this assumption is not ideal, since the decision to belong to a specific sector of the economy might be endogenous, it does not affect the results emphasized in this paper. Similarly, although the education human capital variables may be correlated with the error, the standard assumption is adopted that education is exogenous. The earnings function was estimated separately for men and women, taking into account that some of the health status and control variables may differ by sex (height in particular). The earnings function was also estimated separately for rural and urban areas, though the estimation may be affected by selective migration. Descriptive statistics of the variables used in the empirical models are included in Tables A2 and A3 in the Appendix.

The models were estimated with ordinary least squares [OLS]. The hourly earnings regressions are shown in Tables 3 and 4. Height is highly significant and has the expected sign (positive). Height benefits men’s earnings more than women’s earnings. On the contrary, a very weak correlation is observed between wages and the number of days disabled in the last month. This health indicator is not significant and does not show the expected sign (negative) for the male estimations. The model is also estimated excluding the health variables from the right hand side of equation (1). This exercise shows that the coefficients of age and the dummy for “wage-earner” do not change much when the health variables are included in the earnings equation.

The returns to education in general are not altered by the inclusion of the health variables in the regressions. The estimates of the return to education decrease slightly only for height. In most cases there is a significant positive coefficient for the dummy of being a salaried worker, meaning that a salaried worker has higher wages. The age variables are significant and have the expected signs.

In additional regressions it was found that height and education are positively correlated in Colombia, so that when controlling for education, the coefficient of height drops markedly.[21] Estimations of the model for the whole sample with a gender dummy, indicated that being a female is negatively related to wages, a result reported in other studies (Ribero and Meza, 1997, Ribero and García, 1996). Similarly, the inclusion of a dummy for urban-rural areas suggests that workers in rural areas have lower productivity, a result also documented by Leibovich et al. (1997).

When an earnings equation is estimated based only on a sample of individuals who are working, it is possible to introduce bias in the estimated parameters (Heckman, 1979). This selection bias may be more serious in estimations of female earnings because relatively fewer women make the decision to enter the labor force.[22] To correct this problem some variables must be specified that determine the decision to work but not the observed market wage offers. Ribero and Nuñez (1998) corrected the estimates of equation (1) for possible selection bias generated by using a sample of only individuals who are working. In addition to age and education, participation in the working force is assumed to depend on non labor income,[23] a dummy for living in a house or apartment,[24] a dummy variable for having adequate floors[25] in the house, and a dummy for being the owner of the house where the individual lives.[26] These variables associated with the individual's wealth diminish

[21] The coefficients of height without education were 0.023 and 0.014 for men and women respectively. When education is included they drop to 0.009 and 0.005 for men and women respectively. These coefficients are significant. These regressions are not included.

[22] As reported in Ribero and García (1996), in 1993 the urban female rate of labor force participation was 49.6% and the male rate was 76.4%.

[23] Non–labor income is defined as the sum of four variables in the survey. The actual question of the survey is: did you receive money in the last month from any of the following: a) interest (yes, no, amount), b) rent (yes, no, amount) c) pensions or retirement benefits (yes, no, amount) and d) monetary assistance (yes, no, amount). Since non-labor income was not a very powerful instrument for explanation of participation in the labor force, other housing variables were used as proxies for wealth. Non-labor income is measured at the level of the individual and not for the family.

[24] The survey question for "type of housing" has four options: 1. House 2. Condominium or apartment 3. Room or rooms and 4. Others: trailer, natural refuge, tent, etc.. The dummy built here takes the value 1 when the answer was 1 or 2 and zero when the answer was 3 or 4.

[25] Adequate floors are defined as those made of tile, brick, carpet, marble or hard wood. The alternatives were cement or dirt.

[26] This variable, called "owner-occupied housing", is a dummy variable equal to 1 if individual lives in a house that is owned by him or his family, and 0 if he lives in a rented place or other.

the probability of participation for the rural and urban samples, except for adequate floors in the rural sample. The model, estimated separately for males and females, indicated no sample selection bias for males and weak effects for urban females. In the absence of uniform and significant selection bias, the uncorrected estimates are more efficient as well as consistent (Heckman, 1979). For rural women, even if the correction of selection bias was significant, it did not affect in size or significance the estimated coefficients for the health indicators which are the focus of this paper.

Because the health variables included in ***H*** may themselves be simultaneously determined with wages and measured with error, the next step is to instrument for health status.

4.2 Individual reduced forms for health indicators

This section explores the determinants of the observed health outcomes. Using information on individual's endowments, education and wealth and the community health infrastructure prices and policies, the model tries to account for the individual indicators of health status. The estimated equation is:

$$H_i = a_2 + S\ b_{2j} X_{ji} + S\ c_{2k} O_{ki} + S\ d_{2h} P_h + e_{2i} \qquad (2)$$

where H_i is the individual health status indicator, X_{ji} are exogenous endowments to the individual, O_{ki} represent private opportunities and P_h public policies that may have an impact on health outcomes. The error term e_{2i} is assumed to be zero mean independently distributed, and a_2, b_2, c_2 and d_2 are the estimable parameters. The letters ***j, k*** and ***h*** index the sets ***X***, ***O*** and ***P*** respectively, and ***i*** indexes the individual.

The model is estimated with OLS for both health indicators. Since the number of days disabled in the last month is truncated at zero and 29, Tobit models were also estimated, but the derivatives at sample means do not differ substantially from the OLS (results are not reported). The variables in the set ***X*** are age, schooling and being a salaried worker. When the health indicator is height, the model controls for being born in rural areas. The variables in the set ***O*** are included to take into account that wealth might shift health outcomes positively. These variables are (i) non labor income, (ii) a dummy for having electricity in the house and (iii) a dummy for having telephone service in the house. Given that each health variable comes from different surveys, the available instruments to explain each variable differ slightly. When the health variable is "number of days disabled in the last month," a type of housing variable is included (dummy to indicate the type of housing equal to one if the person lives in a house, condominium or apartment and zero otherwise). When the health indicator is height, the model controls for owner-occupied housing (dummy equal to one if the individual lives in a house that is owned by him or his family, and zero if he lives in a rented place or other).

The variables in the set ***P*** are obtained using data from other sources.[27] They describe the community-specific environment and are expected to be related to the health outcomes. These variables are health provision indicators (coverage of public vaccination programs and per capita expenditure in health in each municipality), education access indicators (mean time to schools in region and number of primary schools per capita in region), labor market indicators (unemployment rates and access to public credit in region) and climate indicators (temperature and yearly average rainfall in each municipality). These characteristics are given at the time of the surveys since it was not possible to match the community-specific environment characteristics at the time each individual was born or in infancy.[28]

Given that the childhood nutritional and environmental factors determine height, it would have been desirable to consider the community-specific characteristics of the environment in which the person was born and raised in the first years of life instead of those of the place of current residence. However, the survey does not clearly account for the place of birth of the individual, making this an impossible task.

The estimation results are in Tables 5 and 6. Age is an important factor that explains both health indicators, with older individuals tending to have worse health.[29] Age is more significant for height

---

[27] The Ministerio de Salud and the Instituto Geográfico Agustin Codazzi.

[28] The task to match the community-specific environment characteristics at the time each individual was born would require collecting time series data for seven decades on the aforementioned variables, data which for the most part does not exist.

[29] Estimation of these models only with a linear term for age showed that an additional year increases the average number of days disabled by 0.005 and 0.008 for urban males and females respectively, and by 0.01 and 0.02 for

than for the number of days disabled. Height increases with age until around 33 years old for females and then begins to decrease. The shape of a graph with height in the Y-axis and age in the X-axis with the coefficients from Table 6, is that of an inverse U indicating that for older women an additional year implies a larger decrease in height than for younger women. For men also height decreases with age, but the pattern is flatter along the relevant range, indicating that the decreases in height with age are more or less constant, tending to decrease when reaching age 52.[30] This outcome was expected from the analysis of the descriptive statistics included in Section 3[31] and Figure 2. The negative effects of age on days disabled are larger in rural than in urban areas, and for females than for males. These effects are monotonic for all the samples during most of the age range. As expected, schooling has a positive and very significant effect on height (higher for males), but on days disabled it is significant and negative only in urban samples.

Non labor income is not significant, but the wealth proxy "living in a house or apartment" is negatively related to the number of days disabled and is significant in the urban samples. Even though owner-occupied housing is not significantly related to height, the dummy for having a telephone in the house is positively associated with height. These results suggest that wealthier individuals may have invested in better health, controlling for the individual and community characteristics listed in the tables.

Some of the municipality level indicators exhibit different effects on the two health indicators. The coverage of vaccinations is positively and significantly associated with height, though it has little impact on the number of days disabled. The per capita expenditure on health has the expected impact on the number of days disabled in the last month, however, it has a negative effect on height. Other municipality level indicators exhibit different effects depending on the sample. The

---

rural males and females respectively. Similarly, one more year of age in the model implies decreases in height of 0.034 cm. for urban men and 0.023 cm. for urban women. These regressions are not reported.

[30] Further regressions of height against age (not reported) indicate that one year older cohorts have 0.06 cm. less of height, the same for males and females. Holding only age constant, a woman with one more year of schooling is expected to be 0.3 cm taller and a man 0.4 cm.. When other individual variables e.g. non-labor income, owner-occupied housing, and community characteristics are taken into account, the partial association between schooling and height decreases.

[31] The height gaps between the cohorts 45-55 and 35-44 are 1.45 cm. for females and 0.75 cm. for males and the gaps between the cohorts 35-44 and 25-34 are 0.04 cm. for females and 0.54 cm. for males.

number of primary schools per capita in the region is negatively associated with days disabled for rural females, but shows the opposite sign for urban males. [32, 33]

In order to see the impact as a related group of explanatory variables, the F-tests of joint significance are reported at the bottom of Tables 5 and 6. These tests imply that the identifying variables in general are jointly significant at the 5% level. The models of height show higher significance than the models of number of days disabled in the last month.

4.3 Hourly earnings equations with instrumental health variables

The observed human capital health stocks may be correlated with the earnings error and be measured with error. To deal with these problems, equation (1) is estimated using instrumental variable [IV] methods, separately for rural and urban areas and by gender. With the IV methods both health variables are more statistically significant and they affect wages in the expected directions. The results are reported in Tables 7 and 8.

The effect of number of days disabled on earnings is negative and significant for urban and rural samples, and the size of the coefficients indicates that one more day of disability reduces male earnings by a larger percentage than female earnings. Simulations of the model indicate that a rise of 50% in the average number of days disabled in each sample would imply an earnings reduction of 11% for urban males, 8% for urban females, 13% for rural males and 7% for rural females.[34]

---

[32] Colombia's proximity to the Equator implies a strong negative correlation between temperature and altitude. Rosenzweig and Schultz (1998) find that altitude is significant and negative in the determination of child mortality. However, Table 6 shows that temperature is significant and positive for height, implying that in places with higher temperature health tends to be better. The discrepancy might be due to the fact that this sample covers only the eleven major cities of the country. In the rural samples the variable is significant only at the 15% level for the male number of days disabled in the last month. Estimations with altitude and altitude squared (these regressions are not reported) suggested that non-linearities in the effect of altitude on the health indicators are not strong.

[33] The mean time to schools in the region, an access restriction indicator, was expected to be negatively associated with height, but it shows a positive sign for females.

[34] These results imply that the elasticities of log (hourly earnings) with respect to days disabled in the last month are -0.03817 for urban males, -0.02803 for urban females, -0.05067 for rural males and -0.02504 for rural females.

In Table 8 the size of the IV coefficients on height is much larger than the corresponding OLS estimates from Table 4. The male coefficient is twelve times larger and the female coefficient is nine times larger. This indicates that when the endogeneity and measurement problems of height are taken into account, the effect of a variation in (predictable) height on productivity is increased substantially. Simulations of this model suggest that a one centimeter increase in the mean height of males would lead to an earnings increase of 12% for males and 4.7% for women. [35]

The education returns are smaller than those obtained when the health variables are not included in the estimation of equation (1). When the instrumental variable for number of days disabled is included, the returns to schooling drop from 0.87 to 0.79 for urban males, from 0.106 to 0.101 for urban females, from 0.078 to 0.071 for rural males and from 0.103 to 0.098 for rural females. When the instrumental variable for height is included, educational returns drop more substantially from 0.095 to 0.053 for urban males and from 0.094 to 0.083 for urban females. Similarly, when the endogeneity of health is taken into account, the estimated effects of being a salaried worker are reduced (except in the models with height for males). This indicates that when the health indicator is not in the model the variables of schooling and being a salaried worker are capturing part of the effects of health on productivity.

5. Conclusions

The purpose of this study was to understand how public and private investments in health in Colombia might be related to future earnings of individuals. The study identified the magnitude of the returns to having good health status through the direct effect of health variables on earnings of individuals.

Descriptive statistics showed that illness is more frequent for women than for men, for less educated than for more educated, for rural than urban residents, and for older individuals. Corresponding patterns were found with height, although this sample was only urban. The well educated males are almost 6 cm. taller than those with no years of education and for females the

[35] These results imply that the elasticities of log (earnings) with respect to height are 3.173225 for urban males and 1.253418 for urban females.

difference is 4 cm. The secular increases in height imply that in Colombia the height gains for workers per decade are approximately 0.67 cm. for urban men and 0.77 cm. for urban women.

The initial estimations of a Mincerian log-earnings equation that included number of days disabled as a form of human capital parallel to schooling exhibited a weak partial correlation between the health status variable and earnings. When the morbidity variable is treated as endogenous and measured with error and the model is estimated by instrumental variables [IV], it becomes significant and has the expected negative sign. The regressions with height showed the expected positive sign and high significance even without the IV correction for health, but the coefficients increased with IV methods by an order of magnitude. Correcting for the selection bias introduced when only individuals who are earning positive wages are analyzed, made little difference in the hourly earnings equation estimates.

Controlling for age, education and other observable characteristics of individuals, it was found that increasing by 50% the average number of days disabled would imply a reduction in labor earnings of 11% for urban males, 8% for urban females, 13% for rural males and 7% for rural females. On the other hand, having one more centimeter of stature would increase urban female earnings by 4.7% and urban male earnings by 12%. The sizes of the coefficient of height in Colombia are in line with the returns found in Ghana (Schultz, 1996), where a one centimeter increase is associated with a 5.7% wage gain for males and 7.5 % for females, holding constant for BMI and migration. As reported in other studies (Thomas and Strauss, 1997, Schultz, 1996), the returns to education are reduced with the introduction of the health indicators in the earnings equations. They drop from 9.5% without height to 5.3% with height for urban men and from 9.4% without height to 8.3% with height for urban women. When the number of days disabled is included in the IV estimates of the earnings function, the returns to schooling of urban males drop from 8.7% to 7.9%, for rural males from 7.8% to 7.1%, for urban females from 10.6% to 10.1% and for rural females from 10.3% to 9.8%. This result implies that when the health indicators are omitted from the earnings function, the education coefficient is capturing part of the effects of health on productivity.

A general result that does not depend on the measure of health status used is that wealthier individuals (those who have higher non labor incomes, own the house where they live or have better access to public services), controlling for age, community characteristics and geographic location, tend to have better health. Community health provision indicators such as public vaccination programs and the per capita expenditures in health as a set are linked with better health for individuals. Geographic location is associated with health suggesting a better health status for individuals who live in areas with a higher temperature.

The health indicators used in this study are not ideal. The answer to the question "how many days were you disabled in the last month?" is to some degree subjective and may exhibit recall as well as measurement errors. On the other hand, the measurement errors in height could be partially corrected if the measure was taken only from individuals present at the time of the survey, instead of including an approximation for those individuals who were absent.

Finally, the lack of good information on timely public health interventions is a limitation of the study. Despite the effort of collecting data at the "departamento" and municipality levels to describe the individuals' health environment and merging it with the household surveys data for the analysis, most of these indicators could not account for much of the variation in individual health indicators. Although several patterns are suggestive, variables that were expected to be correlated with health outcomes, such as the supply of hospitals in the region or the number of hospital beds in each region, were not significant in explaining the available health indicators. This fact may reveal the poor quality of the information collected from sources other than the surveys, and the need for better indicators of the local quality and prices of the health services. It may reveal also that the measured health services may not be particularly effective in improving the adult health indicators used in the study.

Future research should extend this analysis of height in combination with household survey measures of acute and chronic illnesses and BMI, which could be jointly explained by local policy and environmental factors (Fogel, 1994). With these data, a firmer case may be made for investments in particular health programs and policies that would be expected to raise labor productivity by improving the Colombian population's long run and current health status.

6. References

Behrman, 1993. The economic rationale for investing in nutrition in developing countries. World Development, Vol. 21, No. 11.

Behrman and Deolalikar, 1988. "Health and Nutrition" in H.B. Chenery and T.N. Srinivasan Eds. Handbook of Economic Development, Vol. 1, Amsterdam North Holland Publishing Co.

Cortez, Rafael. Salud y productividad en el Peru un análisis empírico por género y región. R-363, 1999. Final Report presented to the Interamerican Bank of Development. Available in the Web page: http://www.iadb.org/oce/44i.cfm

Deolalikar, Anil B.,1988. "Nutrition and labor productivity in agriculture: estimates for rural South India". Review of Economics and Statistics, Vol. 70, No. 3.

Encuesta CASEN - Material de soporte - DNP - Bogotá, Diciembre de 1993

Fogel, R.W., 1994. "Economic growth, population theory and physiology: the bearing of long term processes on the making of economic policy", American Economic Review 84(3): 369-395.

Heckman, J., 1979. "Sample selection bias as a specification error", Econometrica 47(1): 153-161.

Instituto Geográfico Agustín Codazzi, 1996. Diccionario Geográfico de Colombia. Bogotá, D.C. , Colombia, 1996.

Leibovich, J. et al. 1997. El empleo en el sector rural colombiano: qué ha pasado en los últimos años? qué se puede prever? Documento CEDE # 97-08. Universidad de los Andes. Bogotá.

Martorell, R., and J.P. Habicht, 1986. "Growth in early childhood in developing countries". In Human Growth, Vol. 3, rds. F. Falkner and J.M. Tanner. New York: Plenum Press.

Mincer, J., 1974. Schooling, experience and earnings. New York: Columbia University Press.

Murrugarra, E. and Valdivia, M. The Returns to Health for Peruvian Urban Adults: Differentials Across Genders, The Life-Cycle and The Wage Distribution. R-352, 1999. Final Report presented to the Interamerican Bank of Development. Available in the Web page: http://www.iadb.org/oce/44i.cfm

Ribero, R. and García, C., 1996. "Estadísticas descriptivas del mercado laboral masculino y femenino en Colombia: 1976-1995". Archivos de macroeconomía. Departamento Nacional de Planeación.

Ribero, R. and Meza, C., 1997. "Earnings of men and women in Colombia: 1976 - 1995". Archivos de macroeconomía. Departamento Nacional de Planeación.

Ribero, R. and Nuñez, J., 1998. “Productivity of household investment in health – The case of Colombia” Final Report presented to the Interamerican Bank of Development. Available in the Web page: http://www.iadb.org/oce/44i.cfm

Rosenzweig, M.R., and T.P. Schultz, 1982. “Market opportunities, genetic endowments and intrafamily resource distribution: child survival in rural India” American Economic Review 72 (4).

_______. 1983. “Estimating a household production function: heterogeneity, the demand for health inputs and their effects on birthweight” Journal of Political Economy 92 (5).

_______. 1998. “Child mortality and fertility in Colombia”, Health Policy and Education. Amsterdam : Elsevier Science Pub. Co.

Rosenzewig M.R. and K. Wolpin, 1988. Heterogeneity, intrafamily distribution and child health. The Journal of Human Resources. Vol 23, No. 4.

Sahn and Alderman, 1988. “The effect of human capital on wages, on the determinants of labor supply in a developing country”. Journal of Development Economics, Vol. 29 No. 2.

Schultz, T.P., 1984. “Studying the impact of household economic and community variables on child mortality” Population and Development Review 10 (Supplement.):215-235.

_______. 1996. “Wage rentals for reproducible human capital: evidence from two West African countries”. Working Draft, Yale University.

_______. 1997. “Assessing the productive benefits of nutrition and health: an integrated human capital approach”. Journal of Econometrics 77, NH Elsevier.

Schultz T.P. and Tansel A., 1997. “Wage and labor supply effects of illness in Cote d'Ivoire and Ghana: instrumental variable estimates for days disabled” Journal of Development Economics Vol. 480

Strauss, J., and D. Thomas, 1995. “Human resources: empirical modeling of household and family decisions.” In Handbook in Development Economics, Vol. 3A, eds. J. Behrman and T.N. Srinivasan. Amsterdam : Elsevier Science Pub. Co.

________, 1998. “Health, nutrition and economic development.” Journal of Economic Literature, Vol. XXXVI, No. 2.

Thomas, D. and J. Strauss, 1997. “Health and wages: evidence on men and women in urban Brazil.” Journal of Econometrics, 77.

UNDP, 1998. “Human development report.” Oxford University Press.

Table 1

Mean number of days disabled in last month by sex, age education and area[1,2,3]

	SEX		
AGE	**Male**	**Female**	**TOTAL**
18-24	0.25 (1.57)	0.26 (1.45)	0.25 (1.53)
25-34	0.27 (1.67)	0.35 (1.89)	0.30 (1.75)
35-44	0.32 (1.81)	0.32 (1.61)	0.32 (1.74)
45-59	0.37 (1.99)	0.58 (2.45)	0.44 (2.15)
60-70	0.67 (2.76)	0.89 (3.27)	0.72 (2.89)
TOTAL	0.33 (1.85)	0.40 (1.97)	0.35 (1.89)

	AREA		
EDUCATION	**Rural**	**Urban**	**TOTAL**
0 years	0.57 (2.35)	0.50 (2.44)	0.53 (2.4)
Primary	0.45 (2.14)	0.37 (1.96)	0.40 (2.02)
Secondary	0.34 (1.76)	0.29 (1.7)	0.30 (1.71)
Higher	0.14 (0.74)	0.22 (1.52)	0.22 (1.51)
TOTAL	0.45 (2.12)	0.33 (1.88)	0.35 (1.89)

	EDUCATION				
AGE	**0 years**	**Primary**	**Secondary**	**Higher**	**TOTAL**
18-24	0.29 (1.67)	0.27 (1.63)	0.24 (1.47)	0.22 (1.47)	0.25 (1.53)
25-34	0.37 (1.87)	0.34 (1.83)	0.29 (1.72)	0.23 (1.63)	0.30 (1.75)
35-44	0.37 (1.95)	0.36 (1.86)	0.31 (1.74)	0.17 (1.07)	0.32 (1.74)
45-59	0.48 (2.24)	0.46 (2.21)	0.39 (1.95)	0.31 (2.05)	0.44 (2.15)
60-70	0.94 (3.35)	0.67 (2.74)	0.51 (2.5)	0.06 (0.44)	0.72 (2.89)
TOTAL	0.53 (2.4)	0.40 (2.02)	0.30 (1.71)	0.22 (1.51)	0.35 (1.89)

	SEX		
EDUCATION	**Male**	**Female**	**TOTAL**
0 years	0.46 (2.2)	0.70 (2.88)	0.53 (2.4)
Primary	0.36 (1.96)	0.48 (2.15)	0.40 (2.02)
Secondary	0.28 (1.7)	0.32 (1.72)	0.30 (1.71)
Higher	0.18 (1.43)	0.26 (1.6)	0.22 (1.51)
TOTAL	0.33 (1.85)	0.40 (1.97)	0.35 (1.89)

Source: CASEN

1) Standard deviation in brackets.
2) Sample including all workers between 18 and 70 years old in 1993.
3) Sample excluding individuals with 30 days of disability and domestic servants.

Table 2

Mean height in centimeters by sex, age and education[1,2,3]

FEMALE			**EDUCATION**		
AGE	**0 years**	**Primary**	**Secondary**	**Higher**	**TOTAL**
25-34	159.68 (7.58)	159.30 (6.4)	161.23 (6.43)	162.47 (6.64)	161.16 (6.57)
35-44	158.98 (6.92)	159.68 (7.03)	161.78 (6.47)	162.18 (6.43)	161.12 (6.74)
45-55	157.27 (6.95)	158.52 (7.05)	161.02 (6.42)	162.12 (6.65)	159.67 (6.95)
TOTAL	158.37 (7.11)	159.24 (6.84)	161.38 (6.45)	162.34 (6.57)	160.91 (6.71)

MALE			**EDUCATION**		
AGE	**0 years**	**Primary**	**Secondary**	**Higher**	**TOTAL**
25-34	167.46 (6.35)	167.68 (6.46)	169.80 (6.09)	172.75 (6.52)	169.72 (6.46)
35-44	166.33 (8.24)	167.57 (6.62)	169.79 (6.46)	171.24 (6.16)	169.18 (6.66)
45-55	165.48 (6.62)	167.34 (6.7)	169.63 (5.96)	171.40 (6.85)	168.43 (6.66)
TOTAL	166.18 (7.13)	167.53 (6.59)	169.77 (6.18)	171.99 (6.47)	169.28 (6.59)

Source: ENH-91

1) Standard deviation in brackets.
2) Sample including all workers aged 25 - 55 in 1991 taller than 135 cm.
3) Sample excludes domestic servants.

Table 3

Initial wage equations with and without days disabled in the last month as a regressor - dependent variable log (hourly labor income)

	Male											
	Urban						**Rural**					
Variables	Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic	
Health indicator												
Number of days disabled/1000	--	--		5.388	(1.66)		--	--		0.923	(0.16)	
Individual variables												
Age	0.069	(24.54)	*	0.069	(24.55)	*	0.035	(6.77)	*	0.035	(6.77)	*
Age squared /1000	-0.674	(19.74)	*	-0.675	(19.75)	*	-0.345	(5.55)	*	-0.345	(5.55)	*
Years of schooling	0.087	(61.17)	*	0.087	(61.20)	*	0.078	(18.62)	*	0.078	(18.61)	*
Dummy wage earner	0.003	(0.23)		0.003	(0.23)		0.207	(8.24)	*	0.207	(8.24)	*
Intercept	4.159			4.157			4.514			4.514		
Number of observations	18,593			18,593			4,940			4,940		
F-test	1193.06			955.09			111.46			89.16		
Adj R-squared	0.2041			0.2042			0.0821			0.0819		

	Female											
	Urban						**Rural**					
Variables	Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic	
Health indicator												
Number of days disabled/1000	--	--		-1.665	(0.37)		--	--		-4.646	(0.36)	
Individual variables												
Age	0.074	(16.52)	*	0.074	(16.52)	*	0.048	(3.30)	*	0.047	(3.26)	*
Age squared /1000	-0.709	(12.64)	*	-0.709	(12.64)	*	-0.422	(2.50)	*	-0.417	(2.46)	*
Years of schooling	0.106	(47.37)	*	0.106	(47.34)	*	0.103	(10.75)	*	0.103	(10.73)	*
Dummy wage earner	0.212	(10.81)	*	0.212	(10.81)	*	0.278	(3.90)	*	0.277	(3.89)	*
Intercept	3.479			3.479			3.884			3.892		
Number of observations	10,424			10,424			1,293			1,293		
F-test	852.18			681.72			46.37			37.1		
Adj R-squared	0.2462			0.2462			0.1232			0.1226		

Source: CASEN

Samples exclude domestic service. Ages 18-70. Sample exclude days disabled =30.

* = significant at the 5% level.

Table 4

Initial wage equations with and without height as a regressor - dependent variable log (hourly labor income)

Variables	**Male**				**Female**			
	1		2		3		4	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Health indicator								
Height in cm./100	--	--	0.929	(10.04) *	--	--	0.525	(4.56) *
Individual variables								
Age	0.049	(7.28) *	0.050	(7.51) *	0.055	(6.10) *	0.054	(6.02) *
Age squared /1000	-0.408	(4.70) *	-0.422	(4.88) *	-0.537	(4.57) *	-0.526	(4.49) *
Years of schooling	0.095	(66.08) *	0.092	(62.15) *	0.094	(49.78) *	0.093	(48.61) *
Dummy wage earner	0.088	(6.66) *	0.090	(6.82) *	0.220	(12.42) *	0.217	(12.24) *
Dummy born in rural area	-0.060	(3.12) *	-0.051	(2.67) *	-0.041	(1.66)	-0.034	(1.40)
Intercept	3.994		2.418		3.668		2.848	
Number of observations	11,772		11,772		7,136		7,136	
F-test	1021.3		875.1		684.2		575.2	
Adj R-squared	0.3024		0.3082		0.3238		0.3256	

Source: ENH-91

Samples exclude domestic service, height smaller than 135 cm. and incomplete information on wages and earnings. Age is restricted to 25-55 years old.

* = significant at the 5% level.

Table 5

First stage regressions for number of days disabled in the last month - dependent variable number of days disabled in the last month

Variable	Male						Female					
	Urban			Rural			Urban			Rural		
	Coefficient	t-stat.		Coefficient	t-stat.		Coefficient	t-stat.		Coefficient	t-stat.	
Individual variables												
Age	-0.005	(0.75)		-0.021	(1.65)		-0.002	(0.16)		-0.053	(1.71)	*
Age squared/1000	0.121	(1.57)		0.382	(2.51)	*	0.119	(0.97)		0.860	(2.38)	*
Years of schooling/100	-0.800	(2.31)	*	-0.990	(0.92)		-0.889	(1.67)		0.922	(0.42)	
Dummy wage earner	-0.010	(0.34)		-0.121	(1.91)	*	-0.025	(0.57)		-0.084	(0.54)	
Own wealth indicators												
Dummy live in house or apartment	-0.091	(1.61)		-0.097	(0.53)		-0.313	(3.47)	*	0.122	(0.35)	
Non-labor income/10^6	0.063	(0.52)		-0.561	(0.59)		0.392	(1.26)		0.211	(0.09)	
Dummy electricity in house	-0.025	(0.25)		-0.028	(0.41)		-0.143	(0.90)		-0.395	(2.37)	*
Dummy telephone in house	-0.095	(2.76)	*	-0.207	(0.85)		-0.018	(0.40)		0.016	(0.03)	
Health provision indicators												
Triple vaccine/1000	-0.493	(0.95)		0.640	(0.44)		-0.343	(0.47)		0.837	(0.28)	
Tetanus vaccine/1000	0.104	(0.49)		-4.228	(1.94)	*	-0.448	(1.46)		0.280	(0.06)	
Per capita expenditure in health	-1.270	(0.46)		48.661	(1.69)		-5.559	(1.48)		-104.728	(1.63)	
Education access indicators												
Average time to school in region/10	0.141	(1.39)		-0.162	(1.25)		-0.099	(0.69)		0.145	(0.52)	
Primary schools per capita in region	1.101	(2.20)	*	-0.355	(0.55)		-0.342	(0.49)		-3.549	(2.55)	*
Labor markets indicators												
Unemployment rate in region	0.188	(0.21)		6.449	(2.74)	*	0.179	(0.14)		-5.192	(1.01)	
Public credit in region	-0.071	(0.96)		-0.124	(0.66)		-0.113	(1.08)		0.841	(2.12)	*
Climate indicators												
Annual average rain in cubic mm./10^6	-0.424	(0.04)		20.600	(0.52)		25.200	(1.73)	*	6.890	(0.09)	
Temperature in Celsius /100	-0.100	(0.31)		-1.162	(1.69)	*	-0.213	(0.48)		0.753	(0.52)	
Intercept	-0.689			1.368			1.446			4.085		
F-tests of joint significance												
Age and age squared	10.94			13.31			9.19			8.82		
Own wealth indicators	2.8			0.4			3.7			1.42		
Health provision indicators	0.41			1.82			1.42			0.98		
Education access indicators	4.18			0.78			0.43			5.08		
Labor markets indicators	0.46			3.89			0.58			2.53		
Climate indicators	0.05			1.49			1.6			0.15		
Number of observations	18,593			4,940			10,424			1,293		
F-test	3.68			3.32			4.01			3.76		
Adj R-squared	0.0024			0.0079			0.0049			0.0351		

Source: CASEN
Samples exclude domestic service. Ages 18-70. Sample exclude individuals with days disabled =30.
* = significant at the 5% level.

Table 6

First stage regressions for height - dependent variable height in centimeters

Variable	**Male**			**Female**		
	Coefficient	t-statistic		Coefficient	t-statistic	
Individual variables						
Age	-0.144	(2.16)	*	0.139	(1.55)	
Age squared/1.000	1.414	(1.65)		-2.105	(1.78)	
Years of schooling	0.346	(22.77)	*	0.222	(10.92)	*
Dummy wage earner	-0.109	(0.83)		0.714	(3.95)	*
Dummy born in rural area	-0.674	(3.48)	*	-0.689	(2.72)	*
Own wealth indicators						
Dummy owner occupied housing	-0.090	(0.73)		-0.030	(0.18)	
Non-labor income/100.000	0.026	(0.45)		0.309	(1.19)	
Dummy electricity in house	1.354	(1.36)		-0.823	(0.70)	
Dummy telephone in house	0.793	(5.76)	*	0.640	(3.57)	*
Health provision indicators						
Triple vaccine	0.013	(2.72)	*	0.019	(3.00)	*
Tetanus vaccine	0.014	(1.38)		0.038	(2.86)	*
Per capita expenditure in health	-221.6	(2.15)	*	-370.7	(2.67)	*
Education access indicators						
Average time to school in region	0.079	(1.38)		0.414	(5.31)	*
Primary schools per capita in region	0.418	(0.09)		-11.082	(1.80)	
Labor markets indicators						
Unemployment rate in region	0.904	(0.11)		2.836	(0.26)	
Public credit in region	0.716	(0.74)		5.496	(4.37)	*
Climate indicators						
Annual average rain in cubic mm/1.000	-0.103	(0.44)		0.017	(0.05)	
Temperature in Celsius	0.120	(6.22)	*	0.193	(7.70)	*
Intercept	163.08			151.13		
F-tests of joint significance						
Age and age squared	12.86			3.47		
Own wealth indicators	8.97			3.72		
Health provision indicators	4.94			9.74		
Education access indicators	1			14.47		
Labor markets indicators	0.3			10.97		
Climate indicators	19.76			29.78		
Number of observations	11,772			7,136		
F-test	58.61			38.75		
Adj R-squared	0.0824			0.0893		

Source: ENH-91

Samples exclude domestic service, height smaller than 135 cm. and incomplete information on wages and earnings. Age is restricted to 25-55 years old.

* = significant at the 5% level.

Table 7

Wage equations - dependent variable log (hourly labor income) - Instrumental variables for number of days disabled in the last month

Variable	**Male**						**Female**					
	Urban			**Rural**			**Urban**			**Rural**		
	Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic	
Health Indicator												
Number of days disabled	-0.800	(3.72)	*	-0.673	(2.84)	*	-0.453	(3.04)	*	-0.259	(2.49)	*
Individual variables												
Age	0.065	(10.20)	*	0.021	(1.65)		0.073	(10.39)	*	0.029	(1.55)	
Age squared/1000	-0.581	(6.86)	*	-0.087	(0.49)		-0.651	(6.78)	*	-0.147	(0.63)	
Years of schooling	0.079	(20.78)	*	0.071	(10.27)	*	0.101	(27.66)	*	0.098	(10.55)	*
Dummy wage earner	-0.001	(0.04)		0.143	(2.68)	*	0.201	(6.81)	*	0.239	(3.02)	*
Intercept	4.467			4.987			3.658			4.320		
Number of observations	18,593			4,940			10,424			1,293		
F-test	257.78			33.93			342.51			37.48		
R-squared	--			--			--			--		

Source: CASEN

Samples exclude domestic service. Ages 18-70. Sample exclude individuals with days disabled =30.

* = significant at the 5% level.

Robust standard errors

Table 8

Wage equations - dependent variable log (hourly labor income) - Instrumental variables for height

Variable	**Male**			**Female**		
	Coefficient	t-statistic		Coefficient	t-statistic	
Health indicator						
Height in cm./100	11.342	(10.44)	*	4.592	(7.84)	*
Individual variables						
Age	0.065	(6.64)	*	0.048	(4.92)	*
Age squared/1000	-0.576	(4.51)	*	-0.443	(3.39)	*
Years of schooling	0.053	(11.4)	*	0.083	(32.6)	*
Dummy wage earner	0.109	(5.44)	*	0.191	(8.60)	*
Dummy born in rural area	0.046	(1.60)		0.015	(0.54)	
Intercept	-15.248			-3.507		
Number of observations	11,772			7,136		
F-test	393.6			455.9		
R-squared	--			--		

Source: ENH-91

Samples exclude domestic service, height smaller than 135 cm. and incomplete information on wages and earnings. Age is restricted to 25-55 years old.

* = significant at the 5% level.

Robust standard errors

Table A1.

Sample characteristics of labor force participants divided by ranges of days of disability

Variable	Days disabled = 0 Mean	Std. Dev.	0< Days disabled <30 Mean	Std. Dev.	Days disabled =30 Mean	Std. Dev.
Years of schooling	6.74	(4.39)	6.03	(4.26)	5.52	(4.56)
Age	36.48	(12.39)	38.93	(13.25)	44.99	(13.60)
Log (hourly labor income)	6.124	(0.96)	6.058	(0.98)	6.176	(1.02)
Dummy wage earner	0.60		0.55		0.44	
Non-labor income/10.000	0.881	(8.87)	1.033	(4.96)	2.376	(10.40)
Gender (0=male, 1=female)	0.33		0.40		0.32	
Number of observations	32,274		1,976		145	

Source: CASEN

Samples exclude domestic service. Ages 18-70.

Table A2

Descriptive statistics of variables used in models with days disabled in the last month

	Male			
	Urban		**Rural**	
Variable[1]	Mean	Std. Dev.	Mean	Std. Dev.
Productivity indicator				
Log (hourly labor income)	6.284	(0.89)	5.681	(0.88)
Health indicator				
Number of days disabled	0.300	(1.79)	0.428	(2.06)
Individual variables				
Age	36.59	(12.53)	38.00	(13.89)
Age squared/1000	1495.8	(1031)	1637.3	(1166)
Years of schooling	7.041	(4.26)	3.457	(3.01)
Dummy wage earner	0.633	(0.48)	0.555	(0.50)
Own wealth indicators				
Dummy house or apartment	0.942	(0.23)	0.973	(0.16)
Non-labor income/10.000 [2]	0.974	(10.92)	0.389	(3.12)
Dummy electricity in house	0.982	(0.13)	0.689	(0.46)
Dummy telephone in house	0.232	(0.42)	0.015	(0.12)
Health provision indicators				
Triple vaccine [3]	84.519	(27.34)	82.685	(28.18)
Tetanus vaccine	59.416	(66.50)	38.342	(20.47)
Per capita expenditure in health	0.009	(0.01)	0.007	(0.00)
Education access indicators				
Mean time to school in region	18.146	(1.59)	21.244	(3.34)
Primary schools per capita in region	0.923	(0.03)	0.817	(0.08)
Labor markets indicators				
Unemployment rate in region [4]	0.062	(0.02)	0.021	(0.02)
Public credit in region [5]	0.522	(0.21)	0.540	(0.20)
Climate indicators				
Annual average rain in cubic mm.	1772.99	(1293)	1722.38	(837)
Temperature in Celsius	24.1	(4.65)	22.7	(5.46)
Number of observations	18,593		4,940	

Source: CASEN

Notes:
1) Samples aged 18-70. Exclude domestic service. Exclude individuals with days disabled =30.
2) Monetary figures are in pesos of 1993.
3) Vaccinations are given in % of Health Ministry goals for 1993.
4) Unemployment rate defined as % of unemployed individuals in labor force.
5) Public credit in region defined as % of individuals with acces to public credit.

table continues in next page

Table A2 continuation

Descriptive statistics of variables used in models with days disabled in the last month

	Female			
	Urban		**Rural**	
Variable[1]	Mean	Std. Dev.	Mean	Std. Dev.
Productivity indicator				
Log (hourly labor income)	6.105	(1.00)	5.569	(1.15)
Health indicator				
Number of days disabled	0.378	(1.92)	0.538	(2.32)
Individual variables				
Age	35.73	(11.33)	38.90	(13.23)
Age squared/1000	1405.2	(910)	1688.0	(1123)
Years of schooling	7.939	(4.39)	4.275	(3.72)
Dummy wage earner	0.594	(0.49)	0.382	(0.49)
Own wealth indicators				
Dummy house or apartment	0.954	(0.21)	0.963	(0.19)
Non-labor income/10.000 [2]	1.026	(6.08)	0.496	(2.82)
Dummy electricity in house	0.985	(0.12)	0.756	(0.43)
Dummy telephone in house	0.301	(0.46)	0.022	(0.15)
Health provision indicators				
Triple vaccine [3]	84.946	(27.87)	83.631	(27.78)
Tetanus vaccine	58.133	(65.66)	35.860	(19.02)
Per capita expenditure in health	0.009	(0.01)	0.007	(0.00)
Education access indicators				
Mean time to school in region	18.217	(1.58)	21.374	(3.39)
Primary schools per capita in region	0.924	(0.03)	0.816	(0.08)
Labor markets indicators				
Unemployment rate in region [4]	0.061	(0.02)	0.022	(0.02)
Public credit in region [5]	0.526	(0.22)	0.537	(0.21)
Climate indicators				
Annual average rain in cubic mm.	1851.47	(1465)	1692.58	(954)
Temperature in Celsius	23.8	(4.81)	21.9	(5.88)
Number of observations	10,424		1,293	

Source: CASEN

Notes:
1) Samples aged 18-70. Exclude domestic service. Exclude individuals with days disabled =30.
2) Monetary figures are in pesos of 1993.
3) Vaccinations are given in % of Health Ministry goals for 1993.
4) Unemployment rate defined as % of unemployed individuals in labor force.
5) Public credit in region defined as % of individuals with acces to public credit.

Table A3

Descriptive statistics of variables used in models with height

Variable [1]	**Male** Mean	**Male** Std. Dev.	**Female** Mean	**Female** Std. Dev.
Productivity indicator				
Log (hourly labor income)	6.050	0.765	5.896	0.774
Individual variables				
Age	36.52	8.38	35.46	7.78
Age squared	1404	650	1318	592
Years of schooling	8.269	4.298	9.012	4.374
Dummy wage earner	0.707	0.455	0.712	0.453
Dummy born in rural area	0.110	0.313	0.111	0.315
Health indicator				
Height in cm.	169.27	6.57	160.92	6.69
Own wealth indicators				
Dummy owner occupied housing	0.618	0.486	0.659	0.474
Non-labor income/10.000 [2]	0.630	10.328	0.46	2.93
Dummy electricity in house	0.997	0.589	0.996	0.647
Dummy telephone in house	0.454	0.498	0.552	0.497
Health provision indicators				
Triple vaccine [3]	88.345	18.196	87.920	18.251
Tetanus vaccine	52.904	18.409	54.046	19.219
Per capita expenditure in health	0.007	0.002	0.007	0.001
Education access indicators				
Average time to school in region	20.115	1.974	20.206	2.047
Primary schools per capita in region	0.919	0.051	0.918	0.052
Labor markets indicators				
Unemployment rate in region [4]	0.066	0.023	0.064	0.023
Public credit in region [5]	0.406	0.192	0.391	0.193
Climate indicators				
Annual average rain in cubic mm.	1283.9	446.7	1276.0	428.3
Temperature in Celsius	21.10	5.20	20.53	5.29
Number of observations	11,772		7,136	

Source: ENH-91

Notes:
1) Samples exclude domestic service, height smaller than 135 cm. and incomplete information on wages and earnings. Age is restricted to 25-55 years old.
2) Monetary figures are in pesos of 1991. A 1991 peso is equivalent to 1.53 pesos of 1993.
3) Vaccinations are given in % of Health Ministry goals for 1993.
4) Unemployment rate defined as % of unemployed individuals in labor force.
5) Public credit in region defined as % of individuals with acces to public credit.

Table A4. Main results of paper including individuals with 30 days of disability in the sample.

Initial wage equations with and without days disabled as a regressor - dependent variable log (hourly labor income)

Variables	**Male**				**Female**			
	Urban		**Rural**		**Urban**		**Rural**	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Health indicator								
Number of days disabled/1000	2.000	(0.89)	3.998	(0.99)	2.974	(0.92)	0.039	(0.004)
Number of observations	18,666		4,966		10,464		1,299	
F-test	956.48		90.25		682.33		37.02	
Adj R-squared	0.2038		0.0825		0.2456		0.1219	

These estimates correspond to the model shown in Table 3.
Only the coefficient for number of days disabled in the last month is shown.

Wage equations - dependent variable log (hourly labor income) - Instrumental variables for number of days disabled

Variable	**Male**						**Female**					
	Urban			**Rural**			**Urban**			**Rural**		
	Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic		Coefficient	t-statistic	
Health Indicator												
Number of days disabled	-0.668	(4.83)	*	-0.430	(2.59)	*	-0.230	(2.35)	*	-0.181	(2.14)	*
Number of observations	18,666			4,966			10,464			1,299		
F-test	210.88			39.52			427.97			40.54		

These estimates correspond to the model shown in Table 7.
Only the coefficient for number of days disabled in the last month is shown.

Descriptive statistics of variables used in estimations of models with days disabled

Variable	**Male**				**Female**			
	Urban		**Rural**		**Urban**		**Rural**	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Productivity indicator								
Log (hourly labor income)	6.283	(0.89)	5.682	(0.88)	6.107	(1.00)	5.569	(1.15)
Health indicator								
Number of days disabled	0.416	(2.58)	0.583	(2.96)	0.491	(2.65)	0.674	(3.06)

These statistics correspond to the ones shown in Table A2.
Only the statistics for number of days disabled in the last month and labor income are shown.

Source: CASEN
Samples exclude domestic service. Ages 18-70.
* = significant at the 5% level.