Inequality of Opportunity in Adult Health in Colombia

Johanna Fajardo-Gonzalez†
Department of Applied Economics
University of Minnesota

April 2016

Working Paper No. 2016-2
https://doi.org/10.18128/MPC2016-2

†Correspondence should be directed to:
Johanna Fajardo-Gonzalez
1994 Buford Avenue, St. Paul, MN 55108
e-mail: fajar016@umn.edu

Abstract

This paper measures inequality of opportunity in adult health in Colombia using the 2010 Living Standards and Social Mobility Survey, a rich dataset that provides unique information about individual childhood circumstances in that country. Dissimilarity and Gini-opportunity indexes are calculated to provide different measures of inequality of opportunity using a self-reported variable for health status. The Shapley-value decomposition is then used to estimate the contribution of circumstances such as parental education and household socioeconomic status in childhood to the dissimilarity index. In addition to a national-level analysis, estimates for residents in urban and rural areas are provided. The findings suggest that 8 percent to 10 percent of the circumstance-driven opportunities distinctively enjoyed by those who are healthier should be redistributed or compensated for among those who are less healthy in order to achieve equality of opportunity. Differences in household socioeconomic status during childhood and parental educational attainment appear to be the most important dimensions of inequality of opportunity in adult health.

Key Words Childhood; Colombia; health; inequality; opportunity

JEL Classifications: D39, D63, I14

1. Introduction

The 2006 World Development Report on Equity and Development points out that health is not only an important dimension of welfare, but that inequality in health often reinforces and reproduces over time inequality in domains such as income, education or labor (World Bank, 2006.) Reducing inequalities in health has been of major interest of policy makers in both developed and developing countries. The traditional focus of policies that aim to reduce health inequity is the reduction of inequality in specific health outcomes, access to health care services and health insurance. Differences in opportunities driven by individual characteristics such as gender, ethnicity or place of origin have not received such consideration, but they seem to play a key role in determining how health inequality reproduces over time and across generations. For that reason, the study of alternative policies to reduce health inequality has led to an increasing interest in the equality of opportunity literature and its empirical application to health equity (Rosa Dias and Jones, 2007; Fleurbaey and Schokkaert, 2009; Rosa Dias, 2009; Jusot, Tubeuf and Trannoy, 2010; Li Donni, Checchi and Pignataro, 2014)

Roemer (1998)'s theoretical approach to equality of opportunity is based on the idea that the sources of an individual's desirable outcome, like good health, can be separated between circumstances and efforts. Circumstances are factors that are beyond an individual's control and inequalities emerging from such circumstances should be compensated for. Conversely, effort is affected by individual choice and inequalities arising from different efforts are morally and normatively acceptable. The most important implication derived from the equality of opportunity approach is that an equal-opportunity policy should aim at providing everyone with the same opportunity to achieve or enjoy an excellent outcome. A social planner, therefore, would seek to equalize opportunities rather than outcomes and would allow individuals to be fully responsible for their own choices and final results.

Inequality of opportunity, from a theoretical stance, rests on two principles: the compensation principle and the reward principle (Ramos and van de Gaer, 2012). The

compensation principle indicates that inequalities due to circumstances must be compensated, whereas the reward principle indicates that individual efforts must be rewarded. The ex-ante approach to compensation suggests that equality of opportunity holds as long as all individuals face the same opportunities, regardless of the circumstances faced by each one. Under this approach, the observation of all possible efforts is not required for empirical analysis as inequality of opportunity can then be studied focusing on the outcome distributions for different sets of circumstances.

Following an ex-ante approach, inequality of opportunity in adult health has been studied mainly in the context of developed countries. For instance, Rosa-Dias (2009) finds that about 21 percent of health inequality in adulthood, for a cohort of British individuals born in 1956, is related to circumstances in childhood such as maternal education, spells of financial difficulties , as well as poor health and obesity in childhood. The empirical analysis developed in this paper is also grounded on Trannoy et al. (2010) and Li Donni, Peragine and Pignataro (2014). Trannoy et. al study inequality of opportunity among French adults and suggest that such inequality might be halved if the effects of individual circumstances were removed. Li Donni, Peragine and Pignataro, in contrast to Rosa-Dias, apply an alternative empirical approach to data from various waves of the British Household Panel Survey and estimate that about 30 percent of adult health inequality is due to circumstances.

For developing countries the literature is very scarce. For instance, Jusot, Mage and Menendez (2014) study inequality of opportunity in adult health in Indonesia. The authors construct a synthetic index of global health status using information on biomarkers and self-reported health. Their most salient finding is that the existence of long-term inequalities in adult health is related mainly to variables that indicate a sense of community such as religion and language spoken.

This paper fits in this line of research. Specifically, I address the following research question: among the set of observed circumstances, which particular early life circumstances have a salient long-term association with observed inequality of opportunity in adult health

Colombia as a whole, and in both rural and urban areas of the country? To the best of my knowledge, this paper is among the first to answer this question using data from a developing country.

Colombia is undergoing rapid demographic changes. The Colombian population predominantly lives is urban areas, is aging (life expectancy at birth has increased from 65 to 75 years in the last 35 years), and has a declining fertility rate (decreasing from 4.0 in 1980 to 2.0 births per woman in 2015.) Not surprisingly, health outcomes appear to be worse in rural areas than in urban areas. The perception of health status varies greatly between rural and urban residents: 32 percent of the rural population reports a poor or fair health status whereas 22 percent of the urban population reports a similar status. It is worth noting that access to health care services has considerably increased in the country. The World Health Organization (2014) reports that the health system achieved 96% coverage of the population in 2013. Yet, some important differences persist between urban and rural areas. Findings from a few studies (Restrepo et. al., 2009; Florez et. al, 2007) suggest that the area of residence is an important determinant of the use of health services in Colombia. Differential health care use between urban and rural residents may reflect both a major difficulty in securing the availability of health care providers in rural areas and a large concentration of private health care providers in urban areas (Vargas, 2009.) Besides important differences in the density of medical care access or income, exposure to different childhood circumstances may still play an important role in adult health outcomes currently observed in urban and rural areas.

I use data from the 2010 Colombian Living Standards and Social Mobility Survey, a rich dataset that provides retrospective information about individual childhood circumstances in the country. In the empirical analysis, I use first-order stochastic dominance analysis to provide a weak test of inequality of opportunity in the conditional distributions of self-assessed health status (following Lefranc, Trannoy and Pistolesi, 2008.) I also compute a dissimilarity index and a Gini-opportunity index as direct measures of inequality of opportunity (Paes de Barros, Molinas and Saavedra, 2008; Paes de Barros et al., 2009; Rosa Dias, 2009.) I then use the

Shapley-value decomposition to calculate the specific contribution of childhood circumstances such as parental education and household socioeconomic status at age 10 to inequality of opportunity.

The findings suggest that 8 percent to 10 percent of the circumstance-driven opportunities distinctively enjoyed by those who are healthier should be redistributed or compensated for among those who are less healthy in order to achieve equality of opportunity. Differences in household socioeconomic status during childhood and parental educational attainment appear to be the most important dimensions of inequality of opportunity in adult health. Household socioeconomic status at age 10 contributes between 15 and 22 percent to the dissimilarity index, whereas parental education between 10 and 13 percent. In contrast with Indonesia (Jusot, Mage, and Menendez, 2015), the influence of parental education on adult health is highly relevant whereas ethnicity and region of birth appear to be less important.

The remaining of the paper is organized as follows. Section 2 describes the 2010 Living Standards and Social Mobility Survey and provides some descriptive statistics. Section 3 explains the empirical methods. Estimation results are presented in Section 4. Section 5 provides a discussion of the limitations of this paper and concluding remarks.

2. Data

The main data source is the 2010 Colombian Living Standards and Social Mobility Survey (LSSM – *Encuesta de Calidad de Vida y Movilidad Social*) carried out by the Colombian Bureau of Statistics (*Departamento Administrativo Nacional de Estadistica* – DANE.) This survey provides current and retrospective measures of socioeconomic characteristics. The LSSM is representative for the entire country, urban and rural areas, and for nine different subnational regions.[1] The LSSM includes recall questions on living conditions when the respondent was 10 years old. This set of questions provides information on parental educational attainment and

[1]The regions are: Atlantic, Eastern, Central, Pacific, Orinoquia-Amazonia, Antioquia, Valle del Cauca, San Andrés and Providencia, and Bogotá. Rural areas in the regions of Orinoquia-Amazonia and San Andrés and Providencia were not surveyed due to prohibitive costs and poor road access.

ownership of durable assets during childhood. The social mobility module in the LSSM only considers heads of household who are between 25 and 65 years old. The sample design ensures that the final sample of 2,253 individuals represents about 9.57 million heads of household in Colombia. See Table 1 for a summary of descriptive statistics for the full sample. [2]

The outcome of interest is health status in adulthood. It is measured by self-assessed health status, which has been demonstrated to be effective in predicting mortality (Idler and Benyamini, 1997; van Doorslaer and Gerdtham, 2003) and health care utilization (De Salvo et al., 2005.) In the survey, individuals rank their health as either poor (1), fair (2), good (3) or excellent (4) when answering the question "*In general, how do you rate your health status?.*" Around 73 percent of the respondents reported a good or an excellent health status whereas 2.2 percent reported a poor health status. By area, 78 percent of urban residents reported at least a good health status whereas 68 percent of rural residents reported a similar status.

Self-reported health status has some limitations that have been previously identified in the health literature (Jusot, Mage, and Menendez, 2014) The first limitation is that sub-groups of the population may use different thresholds and reference points when assessing their health status, although their objective health conditions are probably the same, leading to a problem known as reporting bias. The second limitation is the lack of cardinality and continuity of the self-assessed health status variable. This problem proves difficult for the use of standard inequality measures.

The set of *early-life circumstances* includes parental educational level and household socioeconomic status at age 10. Parental educational attainment is a categorical variable that indicates whether a parent completed or not a specific level (primary school, secondary school or higher education). In this sample, approximately 60 percent of the heads of household reported that their parents did not attend school or did not complete primary education. In contrast, less than 9 percent indicated that their parents completed secondary school or a higher education level. In urban areas, 46 percent of fathers and 51 percent of mothers did not

[2] For the urban and rural subsamples, see Tables 10 and 11 in the Appendix.

complete primary education. In rural areas, the percentages for incomplete primary education are even higher: 54 percent for fathers and 62 percent for mothers.

Household socioeconomic status at age 10 is a categorical variable that indicates the quintile in which a household falls into, based on an asset index following the methodology by Vyas and Kumaranayake (2006).[3] For the full sample, about 25 percent[4] of the heads of household are assigned to the first quintile of the socioeconomic index, according to their reports of assets ownership.[5] In urban areas, each of the five quintiles has approximately the same number of individuals. In rural areas, in contrast, 25 percent of individuals belong in the first quintile. Retrospective data are far from ideal and measurement error and recall bias could be problematic, in particular when income or earnings data are asked. It is still possible to argue that the variables for assets ownership that are used in this paper could be remembered with some reasonable accuracy.

I also consider other variables that are likely to affect individual health status. In the set of *demographic controls* I include ethnicity (indigenous, African Colombian, or none), urban or rural location of birth, and region of birth. About 9 percent of household heads reported being a member of an ethnic minority. Indigenous minorities are mostly located in rural areas, in contrast with African Colombian minorities who are uniformly distributed between urban and rural areas.[6] Regarding location of birth, most urban (rural) residents were born in urban (rural) areas.

The LSSM does not provide information on individual or parental health-related behaviors. The only circumstance in the data that is partly affected by individual effort is years of education. Educational attainment is an important variable in the analysis of health inequality,

[3] Variables in the socioeconomic status index include type of floor materials, source of water supply, type of toilet available, availability of electricity, and ownership of appliances like washing machine, vacuum cleaner, refrigerator, gas or electric stove, gas or electric oven, television set, as well as ownership of dwelling, automobile, or motorcycle.
[4] Quintiles of the wealth index do not contain equal numbers of individuals, since many respondents in rural areas have the same or very similar index scores in the lower part of the distribution.
[5] One potential concern that arises from the use of these data is the recall nature of the early-life circumstances. A threat to this analysis comes from the possibility that the information reported is less accurate for longer recall intervals, in particular, for older adults regarding assets ownership in their childhood.
[6] The choice between ethnicity and region is not of particular concern here. The correlation between these variables is low. Predicting ethnicity from region of birth, or vice versa, gives a variance inflation factor of 1, which is well below the rule of thumb of 10.

as it has been shown to have a positive and large association with health status (Lleras-Muney, 2005; Arendt, 2005; Cutler, Lleras-Muney and Vogl, 2008). The average number of years of education of the heads of household in this sample is seven years.

Throughout the analysis, additional controls include gender and age group. In the full sample, about 71 percent of household heads are males. In rural areas this figure is of 79 percent, whereas in urban areas is slightly smaller with only 64 percent male heads.

3. Empirical Methods

3.1 Stochastic-dominance tests

In order to provide an initial assessment of inequality of opportunity, I rely on the comparison of the cumulative conditional distributions of the self-assessed health status variable. Lefranc, Trannoy and Pistolesi (2008) show that under equality of opportunity the probability distribution of health status, given effort, does not depend on how different two sets of circumstances are. The notion of first-order stochastic dominance is then used to construct a weak test of inequality of opportunity. According to the test, there is inequality of opportunity if and only if the conditional distributions of health status can be ordered by first-order stochastic dominance.

I rely on a non-parametric test[7] proposed by Yalonetzky (2013), which is extended to the univariate case by Anand, Roope and Gray (2013.) The test is well suited for categorical variables, as the more familiar statistical tests for stochastic dominance such as the Kolmogorov-Smirnov or the Davidson-Duclos cannot be directly applied to outcomes that lack any cardinal meaning. The Yalonetzky test is a pairwise test that specifically compares the cumulative distributions of two specific types: e.g., the health distribution of individuals whose mothers have incomplete primary education against the health distribution of individuals whose mothers have incomplete secondary education. The null hypothesis that the distribution for a certain

[7] The Online Supplementary Material provides more details on the non-parametric test for stochastic dominance.

type *does not* first-order-stochastic dominate the distribution for another type is tested using a z_k^l statistic. This statistic uses the probabilities or proportions that a person of a particular type reports a particular health status. Another feature of the test is that no assumptions about the particular health distributions need to be made.

A major disadvantage of the stochastic dominance approach is that controlling for demographic characteristics entails a loss of precision in the statistical tests of inequality of opportunity since this type of analysis usually requires splitting the sample into many different groups. Moreover, a test where multiple circumstances are analyzed simultaneously is difficult to implement. Nonetheless, the dominance analysis has the advantage of allowing a direct test on the differences between distributions, compared to a regression analysis which is more restrictive and focuses on the mean differences.

I also use a parametric approach to test for inequality of opportunity following Paes de Barros et al. (2009.) I obtain direct estimates of inequality of opportunity, controlling for age and gender, using a non-linear model for health status. The predicted probability of reporting at least a good health status is used to calculate a dissimilarity index. The index is then decomposed using the Shapley-value. The decomposition measures the contribution of each circumstance to the observed inequality of opportunity in adult health. To provide an alternative measure of inequality of opportunity, I also calculate a Gini-Opportunity Index.

3.2 The Dissimilarity Index of Inequality of Opportunity

In the LSSM sample, 2.2 percent of the respondents report a poor health status (category 1) whereas 7.1 percent report an excellent health status (category 4.) For the subsequent analysis, I group the two lower categories (1 and 2) and the two upper categories (3 and 4) to define a dichotomous variable that equals 0 if the respondent reports a poor or fair health status, and equals 1 if the respondent reports a good or excellent health status.

I measure inequality of opportunity using the dissimilarity index, which has been used in inequality analysis using binary outcomes (Paes de Barros, Molinas and Saavedra, 2008; Paes de

Barros et al., 2009.) The dissimilarity index is a measure proportional to the absolute distance between the distribution of circumstances among those with high outcomes (i.e., excellent health) and the distribution among those with low outcomes (i.e., poor health.)

Paes de Barros, Molinas and Saavedra (2008) show that the dissimilarity index is also a measure of the absolute distance between the distribution of circumstances among those with high outcomes and the overall distribution of circumstances. Mathematically, Paes de Barros, Molinas and Saavedra (2008) show that the index can be expressed as

$$D = \frac{1}{2}\sum_{k=1}^{M} |f_1(x_k) - f(x_k)| \qquad [1]$$

where $f_1(x_k)$ denotes the distribution of circumstances among those who enjoy an excellent health and $f(x_k)$ the overall distribution of circumstances. M denotes the total number of possible circumstances groups that can be formed with J circumstance variables c_j, with j=1,…,J. If each circumstance takes a value among g_j categories, then $\mathrm{M} = \prod_{\mathrm{j}=1}^{\mathrm{J}} \mathrm{g_j}$. The set of all possible values of the set of circumstances is $\{x_1, \dots, x_M\}$, with x_k a member of the set.

Paes de Barros, Molinas and Saavedra show that a consistent estimator for the dissimilarity index for binary outcomes is given by

$$\widehat{D} = \frac{1}{2\bar{p}}\sum_{i=1}^{n} w_i |\widehat{p_i} - \bar{p}| \qquad [2]$$

where $\widehat{p_i}$ is the predicted probability of achieving a good or excellent health status for individual *i=1,…,n*. The estimated conditional probability is $\bar{p} = \sum_{i=1}^{n} w_i \widehat{p_i}$, where $\mathrm{w_i}$ denote sampling weights.

The dissimilarity index of inequality of opportunity can be interpreted as the minimum fraction of the number of healthier persons that need to be redistributed across circumstance groups in order to achieve equal opportunity, that is, an equal proportion of less healthy persons

in all circumstance groups (Paes de Barros, Molinas and Saavedra, 2008)[8] The index ranges between 0 and 1, with 0 indicating a situation with equality of opportunity.

Paes de Barros et al. (2009) and Yalonetzky (2012) show that the dissimilarity index for binary outcomes satisfies some important properties of inequality indexes. First, the index equals 0 if the conditional distributions of health given circumstances are identical (that is, perfect between-type equality in access to opportunities), and equals 1 when one individual always attains an excellent health status while others do not. Second, the dissimilarity index is scale-invariant, so that rescaling the outcome by some scalar does not alter the index. Third, the index exhibits anonymity as it does not vary when individuals switch between two dichotomous states of health status. Fourth, the index is invariant to population replication. Fifth, the dissimilarity index is insensitive to balanced increases in opportunities, which suggests that the index does not change when the predicted probability of achieving a better health status increases for each type in such a way that the original distribution is preserved. That is, the index is insensitive to transfers of opportunities between circumstance groups that are above or below the average population achievement because the balanced increases do not alter the proportion of the population in each type or the proportion of the population enjoying an excellent health status.

Gignoux and Ersado (2012) also show that the index can only increase when new circumstances are added. Elaborating on the last property, Gignoux and Ferreira (2011) show that the measure of inequality of opportunity obtained with a set of observed circumstances is a lower bound on the true inequality of opportunity that would be captured if the full vector of circumstances was observed.

Empirically, the calculation of the dissimilarity index first requires the estimation of a logistic regression model to obtain the predicted probability of achieving a good or excellent

[8] An alternative interpretation: the index indicates the percentage of available opportunities for enjoying a better health status that need to be reallocated from the adults who are healthier to the adults who are less healthy, in order to achieve equality of opportunity.

health status ($\widehat{p_i}$). In the following sub-section, I provide further details of the model to be estimated.

3.2.1 Parametric Model of the Relationship between Health Status and Early Life Circumstances

The predicted probability of achieving a good or excellent health status is obtained after the estimation of a logit model in which the dependent variable is the dichotomous health status indicator previously defined. Thereafter, I use the predicted probability to calculate the dissimilarity index. This procedure is performed for the entire sample, and for the subsamples of urban and rural residents.

First consider a health production function,

$$\mathbf{H} = f(C, D, e, u) \quad [3]$$

where C is a vector of individual circumstances, D a vector of demographic controls and e a vector of effort. The residual term u captures luck and other random factors that are not measured by the other variables in the health production function. Notice that effort can also be affected by individual circumstances. Previous studies suggest that an individual choice variable like educational attainment is associated with a circumstance like parental socioeconomic background since more educated parents provide more inputs into the production of education of their children (Hanushek, 1986.) This association between own education and circumstances implies that,

$$\mathbf{H} = f(C, D, e(D, C, v), u) \quad [4]$$

This relationship can be empirically approximated using a linear specification. Let $\mathbf{H_i}$ be the health outcome for individual *i*, e_i the vector of individual effort variables, and u_i and v_i be error terms that capture luck and other random factors. Equation [4] may be estimated using the following linear system of equations

$$\mathbf{H_i} = C_i\alpha + e_i\beta + D_i\vartheta + u_i \quad [5]$$

$$e_i = C_i\gamma + D_i\theta + v_i \quad [6]$$

For the purpose of this study, instead of estimating the full system of equations, I estimate a reduced-form derived from equations [5] and [6]. The reduced-form model for health status is given by

$$\mathbf{H_i} = C_i\alpha + [C_i\gamma + D_i\theta + v_i]\beta + D_i\vartheta + u_i$$

$$= C_i[\alpha + \beta\gamma] + D_i[\theta\beta + \vartheta\,] + \; v_i\beta + u_i$$

$$= C_i\varphi + D_i\psi + \varepsilon_i \quad [7]$$

where $\varphi = \alpha + \beta\gamma, \psi = \theta\beta + \vartheta$, and $\varepsilon_i = v_i\beta + u_i$

The parameter φ measures both the direct effect and the indirect effect of circumstances through individual effort choice. The estimation of equation [7], using the LSSM data, only provides evidence of the correlation between early life circumstances and health status, and cannot be given any causal interpretation.

As mentioned in section 2, the following circumstances are observed in the 2010 LSSM data: ethnicity (*E*), father's highest educational level (*FE*), mother's highest educational level (*ME*), quintile of household socioeconomic status index during childhood (*WS*), urban or rural area of birth (*LB*), and region of birth (*RB*). The only circumstance partly affected by individual choice that is observed in the dataset is years of education (*ED*). Demographic controls include gender (*M*) and age group (*AG*)

Roemer's definition of equality of opportunity assumes that efforts are orthogonal to circumstances. This assumption suggests that any other determinant of health status that is correlated with circumstances is also understood as a circumstance. For instance, an effort variable such as educational attainment is included in C_i as it is partly correlated with circumstances.

Health status is redefined as a dichotomous variable H^*. Therefore, I estimate a logistic regression model for self-assessed health status controlling for circumstances $C_i \equiv \{E_i, FE_i, ME_i, WS_i, LB_i, RB_i, ED_i\}$ and demographic controls $D_i \equiv \{M_i, AG_i\}$

The logistic regression model takes the form

$$\Pr[H^* = 1|C_i, D_i] = \frac{\exp\{d + C_i a + D_i b\}}{1 + \exp\{d + C_i a + D_i b\}}$$

$$\Pr\left[\mathrm{H}^* = 0\middle|\mathrm{C_i}, \mathrm{D_i}\right] = 1\text{-}\Pr[\mathrm{H}^* = 1|\mathrm{C_i}, \mathrm{D_i}] \qquad [8]$$

In order to estimate the global effect of observed circumstances on health status, I also clean years of education of any influence coming from the other observed circumstances. In a related study, Trannoy et al. (2010) proposed a two-step procedure to estimate the correlation of circumstances and health status in a non-linear model. The first step involves the estimation of the residuals from an auxiliary regression of each of the circumstance variables affected by individual effort on the full set of observed circumstances. In the second step, these residuals are included in the estimable health status equation along with the same vector of observed circumstances. Trannoy et al. emphasize that the residuals from step one represent effort, luck and unobserved circumstances that allow an individual to reach a higher education level, for a given vector of observed circumstances. In this paper, I adopt Trannoy et al (2010)'s empirical strategy.

The logistic regression model now takes the following form:

$$\Pr[H^* = 1|C'_i, \hat{v}_i^e, D_i] = \frac{\exp\{d + C'_i a_1 + \hat{v}_i^e a_2 + D_i b\}}{1 + \exp\{d + C'_i a_1 + \hat{v}_i^e a_2 + D_i b\}}$$

$$\Pr[H^* = 0|C'_i, \hat{v}_i^e, D_i] = 1 - \Pr[H^* = 1|C'_i, \hat{v}_i^e, D_i] \qquad [9]$$

where $C'_i \equiv \{E_i, FE_i, ME_i, WS_i, LB_i, RB_i\}$. Vector C_i includes years of education, whereas vector C'_i does not.

Note that the logistic regression model now contains the term $\hat{v}_i^e$, which corresponds to the residuals obtained from the OLS estimation of the following model:

$$ED_i = k + C_i'g + D_i w + v_i \quad [10]$$

where v_i is a disturbance assumed to be normally distributed.

By construction, the residuals $\hat{v}_i^e$ are orthogonal to circumstances in the equation for health status and represent the share of individual educational attainment explained by individual responsibility, luck and unobserved characteristics and circumstances, for the given vector of observed circumstances, as shown by Trannoy et al. (2010.)

The interest in this paper is to gauge what circumstances are more correlated with health status reported by residents in rural areas and respondents living in urban areas. Therefore, I estimate logistic regression models for the subsample of individuals residing in rural areas and the subsample of individuals residing in urban areas using similar specifications to those presented in equations [8], [9] and [10].[9] Note that I do not perform this analysis for the full sample controlling for a dichotomous variable that indicates current urban or rural residence status, because current residence is considered an effort variable in Roemer's framework that may not be controlled for in the ex-ante approach followed in this paper.

A contribution of my paper comes from the joint estimation of equations [9] and [10]. I provide suggestive evidence regarding the possible transmission channels of health inequalities by defining whether the effect is direct or indirect. For instance, if the estimated coefficient on a particular circumstance is only statistically significant in the estimation of the education equation but not so in the estimation of the health status equation, then it can be argued that the circumstance has an indirect effect. That is, the circumstance only has an effect on self-reported health through its effect on education. Alternatively, if the coefficient on a circumstance is significant in the health status equation only, then it can be argued that the effect is direct. Note that a circumstance may also have both direct and indirect effects. In my view, this type of

[9] I retain both significant and insignificant coefficients in the estimation of the dissimilarity index, following Paes de Barros, Molinas and Saavedra (2008)

analysis is consistent with the transmission channels proposed by Trannoy et. al. (2010.) More specifically, the authors suggest that human capital investments during childhood and the transmission of parental socioeconomic status have an indirect influence on health status in adulthood, whereas a specific risk that takes place during childhood has a direct influence on adult health following a latency period.

3.3 Gini-Opportunity Index

In order to provide a measure of inequality of opportunity that is sensitive to transfers of opportunities between circumstances (Lefranc, Trannoy and Pistolesi, 2008), I calculate a Gini-opportunity index. This index computes the weighted sum of all the differences among areas of opportunity sets and then divides that sum by the mean outcome of the entire population.

The Gini-opportunity index has been applied to the study in health inequalities by Rosa Dias (2009.) The index was first proposed by Lefranc, Trannoy and Pistolesi (2008) to quantify the Gini index for each type G_c, so that the opportunity set for each type is denoted by $\bar{h}_c(1 - G_c)$, where $\bar{h}_c$ represents the average health outcome for type *c*. Rosa Dias (2009) then defines the Gini-Opportunity index in health for k types as:

$$G_{opp} = \frac{1}{\bar{h}} \sum_{i=1}^{k} \sum_{i<j} p_i p_j [\bar{h}_j (1 - G_j) - \bar{h}_i (1 - G_i)] \qquad [11]$$

where $\bar{h}$ denotes the mean of the health distribution, *p* the population share, *G* the Gini coefficient, and *i* the set of circumstances.

Lefranc, Trannoy and Pistolesi (2008) show that the index is bounded between 0 and 1, and that it satisfies almost all of the required properties of inequality indexes. The index, in particular, is not invariant to the scale in which the health outcome is measured. The most salient limitation is that the index, as currently applied, does not account for the ordinal nature of the health status measure. Moreover, the Gini opportunity index is shown to be highly sensitive to the number of types considered by the researcher (Rosa Dias, 2014.)

3.4 Decomposition of the Dissimilarity Index through the Shapley-Value

The Shapley-value decomposition allows estimating what circumstances correlate the most with the observed inequality of opportunity. The Shapley-value is a central solution concept in cooperative game theory and has been extended to inequality analysis by Shorrocks (2012.) I follow the methodology of Hoyos and Narayan (2012) and Gignoux and Ersado (2012) to perform the decomposition. These authors explain that the change in inequality that arises when a new circumstance is added to a set of circumstances depends on the sequence of inclusion of the different circumstance variables. The contribution of each circumstance is measured by the average change in inequality over all possible inclusion sequences. Formally, the change in the dissimilarity index when circumstance c is added to a subset M of circumstances is given by

$$\Delta \mathrm{D}_c = \sum_{M \subset C \setminus \{c\}} \frac{|m|!(\kappa - |m| - 1)!}{\kappa!} [D(M \cup \{c\}) - D(M)] \quad [12]$$

where C denotes the entire set of κ circumstances, and M is a subset of C that includes m circumstance variables except c. $D(M)$ is the dissimilarity index for the subset M and $D(M \cup \{c\})$ is the index obtained after adding circumstance c to subset M.

Let $D(\kappa)$ be the dissimilarity index for the set of κ circumstances. Therefore, the contribution of circumstance κ to $D(\kappa)$ is defined by

$$S_c = \frac{\Delta \mathrm{D}_c}{D(\kappa)} \text{ where } \sum_{i \in C} S_i = 1 \quad [13]$$

As a result, I have an additive decomposition of the dissimilarity index that measures the contribution (in terms of correlation, not causation) of each circumstance to observed health inequality.

4. Results

In this section, I first present a brief summary of the results obtained using non-parametric statistic tests for stochastic dominance. Lefranc, Trannoy and Pistolesi (2008) propose a criterion to assess inequality of opportunity using stochastic dominance, and show

that inequality of opportunity is satisfied if and only if the distributions of health status conditional on different sets of circumstances can be ordered by first-order stochastic dominance (Please see the Online Supplementary Material for further details of the test.) A non-parametric test suitable for categorical variables was introduced by Yalonetzky (2013), and I provide here an extension to assess inequality of opportunity in adult health.

I then examine the estimation results of the logistic regression model for the correlates of self-assessed health status, as well as the calculation and decomposition of the dissimilarity index of inequality of opportunity. I also provide an estimation of the Gini opportunity index, a measure that is sensitive to transfers of opportunities between circumstances, in contrast to the dissimilarity index.

4.1 Stochastic Dominance Tests

In the LSSM data, health status is an ordinal variable which takes on values h=1, 2, 3, 4. Responses to the health status question concentrate in categories 2 (fair) and 3 (good). Thus, for the stochastic dominance analysis, I group the lower two categories together (1 and 2) to define a new categorical variable which equals 1 if the respondent reports a poor or a fair health status, and equals 2 and 3 if the respondent reports a good and an excellent health status, respectively.

In order to compare the conditional distributions of health status, I rely on a non-parametric test proposed by Yalonetzky (2013.) This test is implemented for every pair of categories within a variable of interest. In this subsection, the variables of interest are parental and maternal educational attainment and socioeconomic status at age 10.

The test results, summarized in Table 2, firstly show that the health distribution for the fifth quintile of socioeconomic status at age 10 dominates the health distribution for all but the first quintile (comparing the fifth and first quintile, the z_k^l statistics are all larger than -1.96, for a confidence level of 95%.) The results also show that the fourth quintile dominates the distribution for the first and second socioeconomic status quintiles (the z_k^l statistics are smaller than -1.96, for a confidence level of 95%.) These dominance relationships are statistically

significant at the 5 percent level. In urban areas, I find that the health distribution for the fifth quintile dominates each of the distributions for the four remaining quintiles. In contrast with the urban sample, the statistical tests results for rural areas suggest that the only statistically significant dominance relationship is that of the health distribution for quintile 5 relative to the first and second quintiles.

Concerning parental education, Table 2 (panel b and panels c) shows that the higher the level of paternal and maternal education the better health opportunities are, in particular, in urban areas. Focusing on urban areas, the distribution of health status of individuals whose fathers have some degree of education dominates the health distribution of individuals whose fathers have no education at all. These results also suggest that there is inequality of opportunity in adult health after comparing the health distribution of individuals whose mothers attained more than secondary education relative to individuals whose mothers attained no more than some primary education.

4.2 Estimation Results from the Logistic Regression Model for Health Status

The calculation of the dissimilarity index first requires the estimation of a logistic regression model since health status is defined as a binary outcome. In this subsection, I briefly describe the estimation results in order to suggest the potential direction of the association between reporting at least a good health status and the observed early life circumstances.

I first examine the results obtained from the estimation of Equation [10], where the variable for individual years of education is cleaned from the effect of circumstances. Note that the coefficients reported in Table 3 on household socioeconomic status at age 10 and parental education are all statistically significant at the 5% level. In particular, the coefficient on socioeconomic status is positive, increasing with quintile. This result suggests how relevant is the capacity of richer households to make more investments in the education of their children. A similar relationship is found for higher education levels attained by both parents. These two results hold for the urban and rural subsamples also.

Considering the remaining individual characteristics in the estimation of the correlates of years of education, being born in an urban area is only statistically significant and positive in the full sample, whereas being male and born in the Central region is a positive significant feature in the urban subsample. In rural areas, being male and older negatively affects educational attainment.

4.2.1 Correlates of health status in the full sample

The first two columns in Table 4 display the estimation results of the logistic regression model for the full sample. In column 1, the results correspond to the estimation of the model controlling for years of education as an additional circumstance (as given in Equation [8]). In this sample, on average, males are more likely to report a good health status than females. The estimated correlation between an individual's educational attainment, measured in years of education, and reporting a good adult health status is positive and highly significant. The coefficient on the age-group variables is negative, statistically significant, and increasing with age. The effect of parental education is positive but not significant, with or without the inclusion of own years of education. Regional differences are slightly important. Being born in the Pacific or Bogota has a negative effect on perceived health status, with the Atlantic and San Andres islands being the reference region. No significant difference is observed by area of birth.

Column 2 in Table 4 presents the results for the binary logistic regression model controlling for years of education purged from the effect of the other observed circumstances (as given in Equation [9].) Note that the variable for years of education purged from circumstances has the same point estimate and standard error as years of education, by construction. Controlling for the correlation between years of education and the circumstance variables, does not change the direction of the basic relationships described in the previous paragraph, except for socioeconomic status during childhood, which becomes highly significant and increasing with the quintile of household wealth at age 10. Cleaning years of education from the influence

of the observed circumstances allows obtaining significant and positive coefficient estimates for almost all quintiles of the socioeconomic status variables.

4.2.2 Correlates of health status in the rural and urban subsamples

Table 4 also presents the estimation results for urban and rural areas. Regarding the results for the urban subsample (columns 3 and 4), I find that early life circumstances like household socioeconomic status and parental education have a significant effect on the likelihood of reporting at least a good health status, although the relationship is not very strong. In particular, when I purge years of education from the influence of observed circumstances, I find a positive relationship between reporting a good health status and coming from the fifth quintile of the socioeconomic status variable.

Regarding the effect of parental education, individuals whose fathers attained no more than some years of secondary education are also more likely to report a good health status, relative to those individuals whose fathers did not complete primary education. In the case of maternal education, the only significant and positive association with better health status is that of mothers having completed secondary education or more, relative to mothers with no education or some years of primary education. Unfortunately, these relationships are barely significant.

Using the sample for rural residents, I only find a positive and significant relationship between reporting a good health status and high socioeconomic status during childhood, only in the comparison of quintiles 3, 4 and 5 against quintile 1, which is the excluded category (columns 5 and 6.) Considering the region of birth, being born in the Eastern, the Pacific, or Antioquia has a negative effect on self-assessed health status, relative to those born in the Atlantic and San Andres islands.

I now turn to the discussion on the potential transmission channels of health inequalities in adulthood. In what follows, I refer to the results presented in Tables 3 and 4. Parental socioeconomic status and parental education attainment have both direct and indirect effects

through the effect of education on self-reported health. Note that being born in urban areas has an indirect effect, through educational attainment.

The estimation results for the sample of urban residents also support that parental socioeconomic status and parental education have both a direct and an indirect effect. In contrast, in rural areas, the effect of parental socioeconomic status and parental education is realized through years an education (an indirect effect.)

4.3 Dissimilarity Index of Inequality of Opportunity and the Gini-Opportunity Index

I use the predicted probabilities from the estimation of the logistic regression models, given by equations [8] and [9], to calculate the dissimilarity index. Table 5 displays the index value as well as its decomposition for the full sample, and for the rural and urban samples.[10]

The Gini-opportunity index is also tabulated in Table 5. In the calculation of the Gini-opportunity index, I have used two definitions of the health status variable. First, I use the four-category variable where 1 indicates that the health status is poor and 4 that the health status is excellent. Second, I use the dichotomous variable for health status to calculate the Gini-opportunity index. I present the index for the full sample and for the urban and rural subsamples.

I begin with the analysis of the results for the full sample. The dissimilarity index obtained with the LSSM data is about 8.4 percent. The dissimilarity index is usually interpreted as the share of total opportunities for enjoying a better health status that would need to be redistributed from individuals who feel healthier to individuals who feel less healthy for equality of opportunity to prevail.

The Shapley-value decomposition of the dissimilarity index shows that the early life circumstances that have the largest contributions to the dissimilarity index are: household socioeconomic status at age 10 (16 percent), mother's education (10 percent) and father's

[10] For the decomposition of the dissimilarity index, I use the user-written command in Stata -hoishapley-

education (10.2 percent). Once I clean years of education from the influence of circumstances, the decomposition of the index shows a slight increase in the contributions of socioeconomic status at age 10 (22.2 percent), mother's education (12.4 percent) and father's education (13 percent).

The Gini-opportunity index is of 0.10 when the variable for health status with four categories is the outcome of interest. The index is three times larger when the outcome of interest is a dichotomous variable for self-assessed health status (which equals 0.318.) The Gini-opportunity index, likewise the Gini index, ranges between 0 and 1, so that the closer to 1 the most unequal the distribution of health status among the individuals is. The Gini-opportunity index is not easily decomposable or additive and, therefore, I do not provide an estimate of the contribution that each circumstance makes to the index.

The Gini-opportunity index obtained for the full sample is also slightly larger than that calculated for the United Kingdom by Rosa Dias (2009.) In the British household panel, inequality of opportunity in adult health ranges between 0.009 and 0.018. In contrast with Rosa Dias, who only uses parental socioeconomic status as a circumstance, I use the full set of circumstances (except for the demographic variables, gender and age group) to calculate the Gini-opportunity index.

Turning to the results for the urban sample, I calculate a dissimilarity index of 7.9 percent, when I include years of education in the vector of circumstances. That is, 7.9 percent of total of the circumstance-driven opportunities would need to be redistributed from individuals who are healthier to individuals who are less healthy for equality of opportunity to prevail. In rural areas, the index is relatively larger: about 10.1 percent of total opportunities would need to be redistributed from individuals who are healthier to individuals who are less healthy for equality of opportunity to prevail. The calculated indexes do not change considerably once I clean years of education from the influence of circumstances. For urban areas, the decomposition of the index shows a slight increase in the contributions of socioeconomic status at age 10 (from 10.5 percent to 13.7 percent), mother's education (12.9 percent to 16.5 percent) and father's

education (13 percent to 14.6 percent). For rural areas, the decomposition of the index shows a slight change in the contributions of region of birth (from 20.2 percent and 21.1 percent) and socioeconomic status at age 10 (from 35 percent to 40.5 percent), the two circumstances that are most influential in inequality of opportunity in health status in rural areas.

4.4 Additional Checks

As a first additional check, I include variables for self-reported chronic illness and self-reported disability as control variables (results are presented in Table 6 and Table 7.) Self-reported chronic illness is a dichotomous variable that indicates whether the individual suffers from a chronic or long-standing illness like diabetes, heart disease or cancer. Self-reported disability is a dichotomous variable that indicates the presence of a permanent disability.

These objective measures of health status have a negative and significant effect on the likelihood of reporting a good health status. This result is consistent across the full sample and the subsamples of urban and rural areas. Following the results in table 6, the relationships between circumstances and adult health status previously described do not change after including these health variables in my estimations. Note that in the equation for years of education (results available upon request) I did not include the objective health measures. Thus, by construction, the coefficients and standard errors for chronic illness and permanent disability are the same in both the estimation of the non-linear model for health status including years of education and the estimation including years of education purged from the effect of circumstances. These objective measures of health status, however, highly depend on the respondent's access to health care services. The distribution of health services in the country is not necessarily random. For instance, the differential health care use between urban and rural areas may reflect both a major difficulty in securing the availability of health care providers in rural areas and a large concentration of private health care providers in urban areas (Vargas, 2009.) Note that chronic illness and permanent disability are not perfect indicators of health status on their own either. Individuals may experience psychological adjustment and adaptation

to permanent health problems that, in turn, affect how they perceive and report their health status (Graham, 2008.)

Table 7 shows the estimation of the inequality of opportunity indexes. The Gini-opportunity index is below the index presented in Table 5. The index now ranges between 0.042 and 0.077, with rural areas exhibiting the lowest estimate, as in the main results. Note here that the outcome of interest is the health status variable with four categories. The dissimilarity indexes, on the other hand, are now slightly larger than the indexes reported in Table 7.

Regarding the decomposition of the dissimilarity index, note that all circumstances but own education, have a contribution of between 36 percent and 50 percent, with socioeconomic status at age 10 and region of birth being the most important early life circumstances. In urban areas, besides the aforementioned variables, paternal education is perhaps the most important factor in inequality of opportunity, whereas in rural areas, socioeconomic status at age 10 stands out as the most influential variable. Overall, it can be argued that the results are robust to the inclusion of objective measures of health status.

The use of self-reported and retrospective recall data could bias the results here obtained. In order to gauge if there is a systematic bias in how health status is reported, I examine how people perceive their health status based on their economic conditions, after controlling for the set of circumstances and the presence of chronic illness and permanent disability. Self-reported health status and household income per capita (defined in both levels and logs) are strongly correlated, but once I control for circumstances and objective measures of health status this correlation attenuates at conventional significance levels. Thus, the bias created by self-reported measures should be reduced as long as more objective measures are included in the model.

To check for one conceivable source of bias induced by retrospective recall, I analyze whether the age of an individual affects their recall of birth circumstances in a certain direction. In particular, I estimate the logistic regression models for three age cohorts: 25-35, 36-50, and 51-65 years old. The results suggest that self-reported health suffers from reporting bias in view of the substantial differences by age group. Reporting bias constitutes a threat to the analysis in

this study as it compromises the comparisons between individuals with different socioeconomic characteristics.

The estimation results from the logit models for each age group are shown in Table 8. Being a male is positively associated with reporting a good health for all age-groups. Note for the 25-35 age-group that having a mother who completed primary but not secondary education has a negative association with good health status. In contrast, the opposite is true for the 51-65 age-group. Higher quintiles of household socioeconomic status at age 10 are only statistically significant and positively associated with a good self-assessment of health for individuals between 36 and 50 years of age.

Table 9 shows the estimation of the inequality of opportunity indexes. The Gini-opportunity index ranges between 0.03 and 0.10, with the 50-65 age-group exhibiting the highest coefficient estimate. Note here that the outcome of interest is also the health status variable with four categories. The dissimilarity indexes range between 0.04 and 0.10, with the highest value in the 50-65 group.

Regarding the decomposition of the dissimilarity index, note that all circumstances but own education, have a contribution of between 59 percent and 78 percent. The contribution of each circumstance varies by age cohort. For instance, maternal education seems to be more important for the 50-65 group than for the 35-50 group, for which socioeconomic status at age 10 is the most prominent circumstance in inequality of opportunity. Region of birth and ethnicity are more important for the 25-35 age group than for any other group.

5. Concluding Remarks

This paper measures inequality of opportunity in health using the only dataset publicly available that allows linking early life circumstances to current adult health conditions in Colombia, the 2010 Living Standards and Social Mobility Survey. I have considered self-assessed health status as the outcome of interest as it is effective in predicting health care utilization and mortality. The early life circumstances include parental education and household socioeconomic status at age

10. I also study individual circumstances such as ethnicity, place of birth, and educational attainment, as they are likely to affect reported health status.

The empirical approach relies on the calculation of the dissimilarity and the Gini-opportunity indexes of inequality of opportunity. I applied the Shapley-value decomposition to the dissimilarity index to obtain the relative contributions of early life circumstances. In line with studies that follow a similar approach, the results for the full sample suggest that household socioeconomic status and parental education are the most salient circumstances, but they do not reflect how important circumstances like region of birth or ethnicity may be for specific groups. Parental educational attainment and ethnicity are highly associated with inequality of opportunity in health in urban areas but not so in rural areas. In contrast with urban areas, region of birth is potentially one of the most important circumstances in rural areas.

This study has several limitations. Scholars are usually skeptical with the use of self-reported health status in developing countries. For instance, Sen (2002) argues that socially disadvantaged individuals fail to perceive and report the presence or absence of certain health conditions because they are constrained by their social environment. Moreover, their own understanding and appraisal of their health status may not agree with that of their physicians.

Self-reported health status may suffer from individual reporting heterogeneity. To the best of my knowledge, no study has provided evidence, appropriate for the Colombian context, in favor of or against the use of self-reported health in health research. Objective measures of adult health status are not observed in the LSSM dataset. Unfortunately, surveys like the Demographic and Health Survey do not provide intergenerational information for adults. The study of inequality of opportunity in adult health in Colombia faces the usual problem of data availability.

An additional problem is the use of retrospective questions about circumstances. Household ownership of assets during childhood may not be accurately reported. This misreporting introduces bias in the estimates of the correlation between early life circumstances and adult health. The analysis in this paper does not allow to disentangle the effects of either

genetic inheritance or parental health on investments in child's health capital, which is a weakness also identified in previous research (Trannoy et al., 2010)

The estimation of the dissimilarity index is also likely to be biased due to omitted variables if any of the unobserved circumstances is correlated with any of the observed circumstances included in the analysis. Abras et al. (2013) showed that this problem is potentially mitigated by one of the properties of the dissimilarity index: it can only increase when more circumstances are added. Of course, this property does not imply that the estimated contributions to the index also increase when more circumstances are included.

The inequality of opportunity analysis provides suggestive evidence of the lasting effects of childhood circumstances on adult health. The results presented in this study constitute a first step towards the identification of the potential channels through which health inequalities are transmitted from one generation to the next. The results in this paper also suggest that the transmission channels of health inequality across generations operate differently in rural and urban areas. In order to achieve the goal of equality of opportunity in health, more specific policies should be designed to offset the effects of different circumstances in Colombia as a whole and in both rural and urban areas of that country.

Acknowledgements

I thank Stephen Jenkins and two anonymous reviewers for detailed and helpful comments that have led to a substantially improved manuscript. I am also grateful to Florian Wendelspiess, Marc Bellemare, Paul Glewwe, Elizabeth Davis, Tito Velasco and Juan Chaparro for useful suggestions. I thank participants at the 2015 Meeting of the LACEA Network of Inequality and Poverty (Santa Cruz, Bolivia), 2015 ECINEQ Meeting (Luxembourg), 2015 DIAL Development Conference (Paris), 2014 LACEA Meeting (Brazil), and the 2014 iHEA Congress (Dublin) for their comments. All remaining errors are my own.

References

Abras, A., Hoyos, A., Narayan, A., and Tiwari, S.: Inequality of Opportunities in the Labor Market: Evidence from Life in Transition Surveys in Europe and Central Asia. *IZA* Journal of Labor and Development 2(7), 1-22 (2013)

Anand, P., Roope, L., and Gray, A.: Missing Dimensions in the Measurement of Wellbeing and Happiness. Mimeo (2013)

Arendt, J. N.: Does Education Cause Better Health? A Panel Data Analysis using School Reforms for Identification. Economics of Education Review 24(2), 149-160 (2005)

Cohen, G.A.: On the Currency of Egalitarian Justice. Ethics 99, 906–944 (1989)

Cutler, D. M., Lleras-Muney, A., and Vogl, T.: Socioeconomic Status and Health: Dimensions and Mechanisms. National Bureau of Economic Research Working Paper No. 14333 (2008)

DeSalvo, K., Fan, V.S., McDonnell, M.B., and Fihn, S.D.: Predicting Mortality and Healthcare Utilization with a Single Question. Health Services Research 40(4), 1234-46 (2005)

Fleurbaey, M. and Schokkaert, E.: Unfair Inequalities in Health and Health Care. Journal of Health Economics 28(1), 73-90 (2009)

Flórez C., Soto, V., Acosta, O., Karl, C., Misas, J., and Forero, N.: Avances y Desafíos de la Equidad en el Sistema de Salud Colombiano. Bogotá: Fundación Corona (2007) (in Spanish)

Formby, J., Smith, J., and Zheng, B.: Mobility Measurement, Transition Matrices and Statistical Inference. Journal of Econometrics (120), 181–205 (2004)

Gignoux, J. and Ersado, L.: Labor-market Entry and Opportunity Inequality in Egypt. Mimeo (2012)

Gignoux, J., and Ferreira, F.: The Measurement of Educational Inequality: Achievement and Opportunity, Policy Research Working Paper Series 5873. The World Bank, Washington, DC (2011)

González, J. I., Sarmiento, A., Alonso, C. E., Angulo, R. and Espinosa, F.A.: Efectos del Nivel del Ingreso y la Desigualdad sobre la Autopercepción en Salud: Análisis para el Caso de

Bogotá. Revista Gerencia y Políticas de Salud (9), 120-140. Universidad Javeriana, Bogotá (in Spanish) (2005)

Graham, C.: Happiness and Health: Lessons and Questions for Public Policy. Health Affairs 27(1), 72-87 (2008)

Hanushek, E. A.: The Economics of Schooling: Production and Efficiency in Public Schools. Journal of Economic Literature 24(3), 1141-1177 (1986)

Howes, S.: A New Test for Inferring Dominance from Sample Data. Discussion Paper, STICERD, London School of Economics, London, UK (1996)

Hoyos, A. and Narayan, A.: Inequality of Opportunities among Children: How Much Does Gender Matter? Working Paper 76511. The World Bank, Washington DC (2011)

Idler, E., and Benyamini, Y.: Self-rated Health and Mortality: A Review of Twenty-Seven Community Studies. Journal of Health Economics 23(6), 1083-99 (1997)

Jusot, F., Mage, S., and Menendez, M.: Inequality of Opportunity in Health in Indonesia. DIAL (Développement, Institutions et Mondialisation) Working Papers DT/2014-6 (2014)

Jusot, F., Tubeuf, S., and Trannoy, A.: Inequality of Opportunities in Health in Europe: Why So Much Difference Across Countries? Health, Econometrics and Data Group (HEDG) Working Papers 10/26 (2010)

Lefranc, A., Trannoy, A., and Pistolesi, N.: Equality of Opportunity: Definitions and Testable Conditions with an Application to Income in France. THEMA Working Paper Series No. 2006-13, University of Cergy-Pontoise (2008)

Lleras-Muney, A.: The Relationship between Education and Adult Mortality in the United States. The Review of Economic Studies 72(1), 189-221 (2005)

Li Donni, P., Checchi, V., and Pignataro, G.: Ex-Ante and Ex-Post Measurement of Equality of Opportunity in Health: A Normative Decomposition, Health Economics 23, 182-198 (2014)

Paes de Barros, R., Ferreira, F., Molinas, J., and Saavedra, J.: Measuring Inequality of Opportunities in Latin America and the Caribbean. The World Bank, Washington, DC (2009)

Paes de Barros, R., Molinas, J., and Saavedra, J.: Measuring Inequality of Opportunities for Children. The World Bank, Washington, DC (2008)

Ramos X., and van de Gaer D.: Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence. IZA DP, 6672 (2012)

Restrepo, JH., Zambrano A., Vélez M., and Ramírez M.: Health Insurance as a Strategy for Access: Streamlined Facts of the Colombian Health Care Reform. Working Paper Documentos de Trabajo (14). Facultad de Economia. Bogotá: Universidad del Rosario (2007)

Roemer, J.E.: Equality of Opportunity. Harvard University Press, Cambridge, MA (1998)

Rosa Dias, P.: Equality of Opportunity in Health. In Culyer, Anthony J. (ed.) Encyclopedia of Health Economics, Vol. 1. San Diego: Elsevier. 282-287 (2014)

Rosa Dias, P.: Inequality of Opportunity in Health: Evidence from a UK Cohort Study. Health Economics 18(9), 1057–1074 (2009)

Rosa-Dias P., Jones A.: Giving Equality of Opportunity a Fair Innings. Health Economics 16(2), 109-112 (2007)

Sen, A. Health: Perception Versus Observation: Self-Reported Morbidity has Severe Limitations and Can Be Extremely Misleading. BMJ: British Medical Journal, 324(7342), 860–861 (2002).

Shorrocks, A. F.: Decomposition Procedures for Distributional Analysis: A Unified Framework Based On the Shapley Value, Journal of Economic Inequality 10(1), 1-28 (2012)

Trannoy, A., Tubeuf, S., Jusot, F., and Devaux, M.: Inequality of Opportunities in Health in France: A first pass. Health Economics 19(8), 921–938 (2010)

van Doorslaer, E., and Gerdtham, U. G.: Does Inequality in Self-assessed Health predict Inequality in Survival by Income?: Evidence from Swedish data. Social Science and Medicine 57(9), 1621-29 (2003)

Vargas I.: Barreras en el Acceso a la Atención en Salud en Modelos de Competencia Gestionada: Un Estudio de Caso en Colombia (doctoral thesis). Barcelona: Universitat Autónoma de Barcelona (2009) (in Spanish)

Vyas, S., and Kumaranayake, L.: Constructing Socio-Economic Status Indexes: How to Use Principal Components Analysis. Health Policy Plan 21(6), 459-68 (2006)

Yalonetzky, G.: Stochastic Dominance with Ordinal Variables: Conditions and a Test. Econometric Reviews, 32(1), 126-163 (2013)

Yalonetzky, G.: A Dissimilarity Index of Multidimensional Inequality of Opportunity, Journal of Economic Inequality, 10(3), 343-373 (2012)

World Bank: World Development Report, Equity and Development. Washington DC: The World Bank (2006)

World Health Organization: Country Cooperation Brief Strategy (2014)

Table 1 Summary Statistics: Full Sample

Heads of Household between 25 and 65 years old. Total Number of Observations: 2,253

Variable	Observations	Mean or Proportion	Std. Dev.
Outcome			
Self-assessed Health Status	2,253	2.78	0.60
Poor	49	2.2%	0.15
Fair	556	24.7%	0.43
Good	1,487	66.0%	0.47
Excellent	161	7.1%	0.26
Early-life Circumstances			
Household Socioeconomic Status at Age 10			
Quintile 1	569	25.3%	0.43
Quintile 2	533	23.7%	0.43
Quintile 3	441	19.6%	0.40
Quintile 4	355	15.8%	0.36
Quintile 5	316	14.0%	0.35
No Information on Assets	39	1.7%	0.13
Education Level of Father			
None or Incomplete Primary	1,258	55.8%	0.50
Complete Primary and Incomplete Secondary	377	16.7%	0.37
Complete Secondary or More	194	8.6%	0.28
Unknown Father's Education	422	18.7%	0.39
No Information on Father's Education	2	0.1%	0.03
Education Level of Mother			
None or Incomplete Primary	1,345	59.7%	0.49
Complete Primary and Incomplete Secondary	447	19.8%	0.40
Complete Secondary or More	171	7.6%	0.26
Unknown Mother's Education	288	12.8%	0.33
No Information on Mother's Education	2	0.1%	0.03
Other circumstances			
Ethnicity			
Indigenous	59	2.6%	0.16
Black, mulato, raizal or palenquero	144	6.4%	0.24
No ethnic minority	2,050	91.0%	0.29
Years of Education	2,253	7.02	4.65
Born in Urban Area	1,103	49.0%	0.50
Born in Rural Area	1,144	50.8%	0.50
No Information on Area of Birth	6	0.3%	0.05
Region of Birth			
Atlantic	507	22.5%	0.42
Eastern	518	23.0%	0.42
Pacific	255	11.3%	0.32
Orinoquia-Amazonia	6	0.3%	0.05
Antioquia	251	11.1%	0.31
Valle del Cauca	160	7.1%	0.26
Bogotá	159	7.1%	0.26
San Andrés islands	2	0.1%	0.03
Central	395	17.5%	0.38
Additional Controls			
Male	1,598	70.9%	0.45
Age	2,253	44.77	11.01
Age group			
25-35	504	22.4%	0.42
35-45	594	26.4%	0.44
45-55	646	28.7%	0.45
55-65	509	22.6%	0.42

Source: 2010 Colombian LSSM Survey

Table 2 Stochastic Dominance Tests for Inequality of Opportunity

a. Household socioeconomic status at age 10

Full sample

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1		~	~	~	~
Quintile 2	~		~	~	~
Quintile 3	~	>		~	~
Quintile 4	>	>	~		~
Quintile 5	~	>	>	>	

Urban Areas

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1		~	~	~	~
Quintile 2	~		~	~	~
Quintile 3	~	~		~	~
Quintile 4	>	~	~		~
Quintile 5	>	>	>	>	

Rural Areas

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Quintile 1		~	~	~	~
Quintile 2	~		~	~	~
Quintile 3	~	~		~	~
Quintile 4	~	~	~		~
Quintile 5	>	>	~	~	

b. Paternal Education

Full sample

	None	Primary	Secondary and higher
None *		~	~
Primary **	>		~
Secondary and higher	>	~	

Urban Areas

	None	Primary	Secondary and higher
None *		~	~
Primary **	>		~
Secondary and higher	>	~	

Rural Areas

	None	Primary	Secondary and higher
None *		~	~
Primary **	~		~
Secondary and higher	~	~	

c. Maternal Education

Full sample

	None	Primary	Secondary and higher
None *		~	~
Primary **	>		~
Secondary and higher	>	>	

Urban Areas

	None	Primary	Secondary and higher
None *		~	~
Primary **	~		~
Secondary and higher	>	>	

Rural Areas

	None	Primary	Secondary and higher
None *		~	~
Primary **	~		~
Secondary and higher	~	~	

Note: The symbol ">" indicates that the distribution of the type in the row first-order-stochastic dominates the distribution of the type in the column. The symbol "~" indicates that the distributions cannot be ranked using first-order stochastic dominance.

* None or incomplete primary education

** Complete primary or incomplete secondary education

Source: 2010 Colombian LSSM Survey.

Table 3 Purging years of education from circumstances: OLS Results

Dependent variable: years of education

	All Individuals	Urban Areas	Rural Areas
	(1)	(2)	(3)
Male	0.2172	0.6416***	-0.4885*
	(0.1885)	(0.2204)	(0.2690)
Age group (Ref. 25-35 years old):			
35-45 years old	-0.1058	0.0440	-0.7039**
	(0.2245)	(0.2749)	(0.3049)
45-55 years old	-0.2316	-0.3117	-0.8309**
	(0.2394)	(0.2849)	(0.3324)
55-65 years old	-1.1098***	-1.2353***	-1.8467***
	(0.2668)	(0.3243)	(0.3329)
Ethnicity (Ref. Not a minority):			
Indigenous	-0.0621	-0.0304	0.1704
	(0.5613)	(0.8450)	(0.6265)
Black/mulato/raizal/palenquero	0.3016	0.1005	0.2613
	(0.3615)	(0.4651)	(0.4410)
Region (Ref. Atlantic and San Andres islands):			
Eastern	0.0011	-0.3190	-0.1385
	(0.2681)	(0.3290)	(0.3445)
Pacific	0.4841	1.0698*	0.2100
	(0.3596)	(0.5568)	(0.3465)
Orinoquia and Amazonia	-0.5957	-1.0903	-0.2360
	(0.5788)	(0.7468)	(0.9172)
Antioquia	-0.0747	-0.2467	-0.0174
	(0.3158)	(0.3802)	(0.4452)
Valle	0.5982	0.5387	0.3399
	(0.4001)	(0.4505)	(0.5239)
Bogota	-0.3089	-0.5637	2.0025
	(0.3279)	(0.3598)	(1.6562)
Central	0.5395*	0.7487**	0.0573
	(0.2971)	(0.3669)	(0.3522)
Born in urban area	1.0276***	0.4466	0.3522
	(0.2204)	(0.2849)	(0.2865)
Household socioeconomic status at age 10:			
Quintile 2	0.7084***	1.0493***	-0.3497
	(0.2732)	(0.3525)	(0.3114)
Quintile 3	2.0127***	2.1206***	0.4408
	(0.2874)	(0.3614)	(0.3432)
Quintile 4	3.4114***	3.1020***	0.7434**
	(0.3255)	(0.3848)	(0.3549)
Quintile 5	4.5999***	4.2618***	2.2478***
	(0.3554)	(0.4055)	(0.4083)
Paternal education level (Ref. None):			
Complete primary and incomplete secondary	0.9560***	0.7741**	1.2467**
	(0.3064)	(0.3550)	(0.5217)
Complete secondary or more	1.8947***	1.5467***	3.8638***
	(0.4034)	(0.4459)	(0.7869)
Unknown father's level of education	-0.7116**	-0.7402**	-0.5352*
	(0.2907)	(0.3766)	(0.2938)
Maternal education level (Ref. None):			
Complete primary and incomplete secondary	1.0363***	1.1135***	0.6089
	(0.2906)	(0.3392)	(0.4195)
Complete secondary or more	2.5173***	2.5426***	2.4073**
	(0.4135)	(0.4612)	(1.0519)
Unknown mother's level of education	-0.4045	-0.1635	-0.2143
	(0.3553)	(0.4703)	(0.3390)
Constant	4.6050***	5.5638***	4.9071***
	(0.3564)	(0.4646)	(0.4833)
Observations	2,204	1,242	962
R squared	0.430	0.396	0.246

***, **, and * indicate statistical significance at the 1, 5 and 10 percent level, respectively
Robust standard errors in parentheses
Own calculations. Source: 2010 Colombian LSSM

Table 4 Log-odds Ratios for the Correlates of Health Status

Dependent variable: self-reported health status (0=poor or fair, 1= good or excellent)

	All Individuals		Urban Areas		Rural Areas	
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.5690***	0.5932***	0.6489***	0.7217***	0.5281**	0.4781**
	(0.1277)	(0.1280)	(0.1560)	(0.1566)	(0.2104)	(0.2089)
Age group (Ref. 25-35 years old):						
35-45 years old	-0.5462***	-0.5579***	-0.5281*	-0.5231*	-0.5544**	-0.6264**
	(0.2005)	(0.2005)	(0.2748)	(0.2748)	(0.2481)	(0.2474)
45-55 years old	-0.7550***	-0.7808***	-0.7587***	-0.7941***	-0.8692***	-0.9542***
	(0.1948)	(0.1946)	(0.2650)	(0.2647)	(0.2516)	(0.2527)
55-65 years old	-1.3172***	-1.4406***	-1.3481***	-1.4882***	-1.4127***	-1.6015***
	(0.1964)	(0.1967)	(0.2663)	(0.2669)	(0.2608)	(0.2626)
Ethnicity (Ref. Not a minority):						
Indigenous	-0.2143	-0.2213	-0.7064	-0.7099	0.5513	0.5687
	(0.4386)	(0.4386)	(0.5983)	(0.5983)	(0.4468)	(0.4469)
Black/mulato/raizal/palenquero	-0.2408	-0.2073	-0.3739	-0.3625	-0.0548	-0.0281
	(0.2386)	(0.2385)	(0.2945)	(0.2944)	(0.3495)	(0.3493)
Region (Ref. Atlantic and San Andres islands):						
Eastern	-0.2613	-0.2612	-0.2041	-0.2403	-0.5537**	-0.5679**
	(0.1826)	(0.1826)	(0.2370)	(0.2371)	(0.2488)	(0.2494)
Pacific	-0.6624***	-0.6086***	-0.7622**	-0.6409**	-0.7878***	-0.7663***
	(0.2119)	(0.2107)	(0.3131)	(0.3099)	(0.2704)	(0.2693)
Orinoquia and Amazonia	0.3799	0.3136	0.8195	0.6959	-0.6004	-0.6246
	(0.5176)	(0.5175)	(0.7804)	(0.7804)	(0.7997)	(0.7999)
Antioquia	0.0858	0.0775	0.2955	0.2676	-0.6974**	-0.6992**
	(0.2213)	(0.2214)	(0.2864)	(0.2868)	(0.3055)	(0.3055)
Valle	0.1610	0.2275	0.2359	0.2970	-0.3386	-0.3038
	(0.3232)	(0.3235)	(0.3939)	(0.3942)	(0.4189)	(0.4185)
Bogota	-0.4860*	-0.5203*	-0.4415	-0.5054*		
	(0.2795)	(0.2801)	(0.3047)	(0.3060)		
Central	-0.2169	-0.1569	-0.1171	-0.0322	-0.4650*	-0.4591*
	(0.2017)	(0.2010)	(0.2678)	(0.2664)	(0.2543)	(0.2542)
Born in urban area	-0.0722	0.0420	-0.1611	-0.1105	0.1597	0.1957
	(0.1371)	(0.1360)	(0.1794)	(0.1793)	(0.2370)	(0.2366)
Household socioeconomic status at age 10:						
Quintile 2	0.1220	0.2008	0.1109	0.2299	0.1291	0.0934
	(0.1618)	(0.1604)	(0.2248)	(0.2211)	(0.2500)	(0.2498)
Quintile 3	0.3300*	0.5538***	-0.0288	0.2117	0.7877***	0.8328***
	(0.1831)	(0.1796)	(0.2331)	(0.2282)	(0.2552)	(0.2559)
Quintile 4	0.1149	0.4943**	-0.2175	0.1342	0.7065***	0.7825***
	(0.2148)	(0.2044)	(0.2707)	(0.2540)	(0.2576)	(0.2564)
Quintile 5	0.4963*	1.0078***	0.3021	0.7854**	0.7044**	0.9343***
	(0.2986)	(0.2846)	(0.3614)	(0.3426)	(0.2864)	(0.2786)
Paternal education level (Ref. None):						
Complete primary and incomplete secondary	0.3043	0.4106*	0.4688*	0.5566**	-0.2181	-0.0906
	(0.2216)	(0.2217)	(0.2618)	(0.2628)	(0.3625)	(0.3596)
Complete secondary or more	-0.0745	0.1362	-0.0144	0.1610	0.4579	0.8531
	(0.3773)	(0.3788)	(0.4069)	(0.4085)	(0.7744)	(0.7731)
Unknown father's level of education	0.1135	0.0344	0.3437	0.2597	-0.3095	-0.3642
	(0.1950)	(0.1948)	(0.2674)	(0.2668)	(0.2480)	(0.2464)
Maternal education level Ref. None):						
Complete primary and incomplete secondary	-0.0212	0.0940	0.0231	0.1493	-0.3439	-0.2816
	(0.2117)	(0.2109)	(0.2558)	(0.2546)	(0.3187)	(0.3173)
Complete secondary or more	0.5116	0.7915*	0.7245	1.0128**	-1.1600*	-0.9138
	(0.4441)	(0.4398)	(0.5181)	(0.5139)	(0.6946)	(0.6867)
Unknown mother's level of education	-0.0382	-0.0831	-0.0705	-0.0891	0.0485	0.0266
	(0.2310)	(0.2307)	(0.3211)	(0.3210)	(0.2663)	(0.2664)
Years of education	0.1112***		0.1134***		0.1023***	
	(0.0174)		(0.0219)		(0.0262)	
Years of education purged from circumstances		0.1112***		0.1134***		0.1023***
		(0.0174)		(0.0219)		(0.0262)
Constant	0.6589***	1.1709***	0.7384**	1.3694***	0.6988**	1.2006***
	(0.2437)	(0.2368)	(0.3416)	(0.3290)	(0.3528)	(0.3408)
Observations	2,204	2,204	1,242	1,242	956	956
Region of Birth Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	-4.477e+06	-4.477e+06	-3.328e+06	-3.328e+06	-1.085e+06	-1.085e+06
Pseudo R squared	0.126	0.126	0.136	0.136	0.113	0.113

***, **, and * indicate statistical significance at the 1, 5 and 10 percent level, respectively.

Robust standard errors in parentheses

Own calculations. Source: 2010 Colombian LSSM

Table 5 Gini-Opportunity index and Dissimilarity Index of Inequality of Opportunity, with its Decomposition

	All individuals		Residents in Urban Areas		Residents in Rural Areas	
Gini-Opportunity Index (1)	0.1019		0.1148		0.0720	
Gini-Opportunity Index (2)	0.3182		0.3550		0.2604	
Dissimilarity Index (3)	0.0838	0.0839	0.0793	0.0793	0.1016	0.1016
			Decomposition of the Dissimilarity Index (in %)			
Educational Attainment	46.59		45.25		30.13	
Education purged from circumstances		33.31		36.76		22.53
Circumstances	53.41	66.69	54.75	63.24	69.87	77.47
Early Life Circumstances	35.80	47.71	36.42	44.85	44.13	49.99
Mother's Education	10.04	12.93	12.90	16.50	3.54	2.20
Father's Education	10.21	12.49	12.98	14.57	5.64	7.30
Household Socioeconomic Status at age 10	15.56	22.28	10.53	13.77	34.96	40.49
Demographics	17.61	18.98	18.33	18.39	25.73	27.49
Region of Birth	11.64	11.95	13.13	13.17	20.19	21.10
Born in Urban Area	4.56	5.61	1.00	0.97	3.87	4.71
Ethnicity	1.42	1.42	4.20	4.25	1.67	1.69
Observations	2,204		1,242		962	

Bootstrapped standard errors in parentheses. 100 replications.
Own calculations. Source: 2010 Colombian LSSM
Notes:
(1) The Gini-opportunity index is calculated using a self-assessed health status variable in which 1=poor, 2=fair, 3=good, and 4=excellent. A categorical variable for the individual's years of education has also been used in this calculation. Gender and age group are not included.
(2) The Gini-opportunity index is calculated using a self-assessed health status variable in which 0=poor or fair, and 1=good or excellent.
(3) The index in the first, third and fifth columns include years of education as a circumstance, whereas the second, fourth, and sixth columns include years of education purged from circumstances.

Table 6 Log-odds Ratios, controlling for presence of chronic illness or permanent disabilities.

Dependent variable: self-reported health status (0=poor or fair, 1= good or excellent)

	All Individuals		Urban Areas		Rural Areas	
	(1)	(2)	(3)	(4)	(5)	(6)
Any chronic illness (1=Yes)	-1.9755***	-1.9755***	-2.0409***	-2.0409***	-1.7436***	-1.7436***
	(0.1761)	(0.1761)	(0.2068)	(0.2068)	(0.2678)	(0.2678)
Any permanent disability (1=Yes)	-1.4031***	-1.4031***	-1.5184***	-1.5184***	-1.2382**	-1.2382**
	(0.3701)	(0.3701)	(0.5360)	(0.5360)	(0.5053)	(0.5053)
Male	0.4603***	0.4863***	0.5621***	0.6401***	0.4354**	0.3888*
	(0.1373)	(0.1375)	(0.1685)	(0.1686)	(0.2188)	(0.2171)
Age group (Ref. 25-35 years old):						
35-45 years old	-0.5017**	-0.5144**	-0.4825*	-0.4772	-0.5309**	-0.6043**
	(0.2094)	(0.2093)	(0.2914)	(0.2914)	(0.2540)	(0.2535)
45-55 years old	-0.4342**	-0.4619**	-0.3978	-0.4357	-0.6642**	-0.7509***
	(0.2071)	(0.2072)	(0.2857)	(0.2859)	(0.2588)	(0.2600)
55-65 years old	-0.8310***	-0.9638***	-0.8056***	-0.9556***	-1.0912***	-1.2735***
	(0.2108)	(0.2108)	(0.2921)	(0.2924)	(0.2767)	(0.2772)
Ethnicity (Ref. Not a minority):						
Indigenous	-0.1588	-0.1663	-0.4919	-0.4956	0.4388	0.4555
	(0.3975)	(0.3975)	(0.5402)	(0.5402)	(0.4583)	(0.4583)
Black/mulato/raizal/palenquero	-0.1288	-0.0927	-0.2521	-0.2399	-0.0422	-0.0178
	(0.2604)	(0.2606)	(0.3303)	(0.3304)	(0.3843)	(0.3837)
Region (Ref. Atlantic and San Andres islands):						
Eastern	-0.1640	-0.1639	-0.1203	-0.1590	-0.4932*	-0.5061**
	(0.1922)	(0.1922)	(0.2514)	(0.2512)	(0.2546)	(0.2552)
Pacific	-0.5767**	-0.5188**	-0.6139*	-0.4840	-0.7038**	-0.6838**
	(0.2287)	(0.2277)	(0.3561)	(0.3534)	(0.2801)	(0.2790)
Orinoquia and Amazonia	0.2593	0.1880	0.7666	0.6341	-0.7848	-0.8086
	(0.4692)	(0.4690)	(0.6800)	(0.6798)	(0.7630)	(0.7633)
Antioquia	0.1878	0.1788	0.3785	0.3486	-0.5712*	-0.5727*
	(0.2334)	(0.2334)	(0.3046)	(0.3047)	(0.3144)	(0.3145)
Valle	0.3126	0.3842	0.3487	0.4141	-0.1455	-0.1129
	(0.3235)	(0.3236)	(0.3891)	(0.3891)	(0.4842)	(0.4839)
Bogota	-0.5127*	-0.5496*	-0.4760	-0.5445*		
	(0.2826)	(0.2831)	(0.3103)	(0.3110)		
Central	-0.0846	-0.0201	0.0448	0.1358	-0.4093	-0.4044
	(0.2104)	(0.2102)	(0.2829)	(0.2821)	(0.2592)	(0.2591)
Born in urban area	-0.1281	-0.0052	-0.2701	-0.2159	0.1596	0.1944
	(0.1451)	(0.1434)	(0.1928)	(0.1924)	(0.2469)	(0.2465)
Household socioeconomic status at age 10:						
Quintile 2	0.0974	0.1821	0.0538	0.1812	0.1404	0.1063
	(0.1696)	(0.1682)	(0.2357)	(0.2323)	(0.2598)	(0.2598)
Quintile 3	0.4048**	0.6455***	0.0342	0.2918	0.8708***	0.9125***
	(0.1983)	(0.1955)	(0.2609)	(0.2567)	(0.2621)	(0.2625)
Quintile 4	0.2750	0.6830***	-0.0029	0.3738	0.7075***	0.7801***
	(0.2261)	(0.2180)	(0.2850)	(0.2712)	(0.2688)	(0.2674)
Quintile 5	0.8770***	1.4271***	0.7342*	1.2518***	0.9375***	1.1488***
	(0.3115)	(0.2959)	(0.3773)	(0.3572)	(0.3138)	(0.3063)
Paternal education level (Ref. None):						
Complete primary and incomplete secondary	0.3285	0.4428**	0.5145*	0.6086**	-0.2968	-0.1746
	(0.2219)	(0.2219)	(0.2646)	(0.2657)	(0.3712)	(0.3693)
Complete secondary or more	-0.1788	0.0478	-0.0889	0.0990	0.0872	0.4411
	(0.3903)	(0.3908)	(0.4313)	(0.4318)	(0.7347)	(0.7313)
Unknown father's level of education	0.0902	0.0051	0.3461	0.2562	-0.3840	-0.4351*
	(0.2038)	(0.2033)	(0.2810)	(0.2802)	(0.2546)	(0.2528)
Maternal education level (Ref. None):						
Complete primary and incomplete secondary	-0.1319	-0.0079	-0.1069	0.0283	-0.3793	-0.3155
	(0.2109)	(0.2096)	(0.2582)	(0.2559)	(0.3129)	(0.3113)
Complete secondary or more	0.4682	0.7693*	0.6236	0.9324*	-0.7971	-0.5783
	(0.4583)	(0.4541)	(0.5500)	(0.5452)	(0.6692)	(0.6610)
Unknown mother's level of education	-0.1725	-0.2209	-0.2442	-0.2641	0.0170	-0.0030
	(0.2360)	(0.2360)	(0.3259)	(0.3259)	(0.2771)	(0.2772)
Years of education	0.1196***		0.1215***		0.0961***	
	(0.0182)		(0.0231)		(0.0278)	
Years of education purged from circumstances		0.1196***		0.1215***		0.0961***
		(0.0182)		(0.0231)		(0.0278)
Constant	0.7647***	1.3154***	0.8864**	1.5622***	0.8300**	1.3060***
	(0.2635)	(0.2576)	(0.3727)	(0.3609)	(0.3667)	(0.3511)
Observations	2,204	2,204	1,242	1,242	956	956
Region of Birth Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	-4.044e+06	-4.044e+06	-2.964e+06	-2.964e+06	-1.018e+06	-1.018e+06
Pseudo R squared	0.211	0.211	0.230	0.230	0.168	0.168

***, **, and * indicate statistical significance at the 1, 5 and 10 percent level, respectively.

Robust standard errors in parentheses

Own calculations. Source: 2010 Colombian LSSM

Table 7 Gini-Opportunity index and Dissimilarity Index of Inequality of Opportunity, with its Decomposition, controlling for presence of chronic illness or permanent disability

	All individuals		Residents in Urban Areas		Residents in Rural Areas	
Gini-Opportunity Index (1)	0.0777		0.0735		0.0429	
Dissimilarity Index (2)	0.1033	0.1034	0.0990	0.0999	0.1227	0.1226
			Decomposition of the Dissimilarity Index (in %)			
Educational Attainment	50.87		42.19		26.09	
Education purged from circumstances		36.12		41.41		19.83
Circumstances	49.13	63.88	57.81	58.59	73.91	80.17
Early Life Circumstances	45.00	31.27	38.80	38.76	53.70	58.94
Mother's Education	8.99	6.15	10.43	11.96	4.16	3.05
Father's Education	10.14	7.74	12.71	13.56	8.57	9.35
Household Socioeconomic Status at age 10	25.86	17.38	15.65	13.24	40.97	46.54
Demographics	18.89	17.85	19.01	19.82	20.21	21.22
Region of Birth	13.46	13.14	16.35	17.07	17.33	17.86
Born in Urban Area	4.32	3.64	0.56	0.90	1.82	2.26
Ethnicity	1.11	1.07	2.11	1.85	1.06	1.11
Observations	2,204		1,242		962	

Bootstrapped standard errors in parentheses. 100 replications.
Own calculations. Source: 2010 Colombian LSSM

Notes:
(1) The Gini-opportunity index is calculated using a self-assessed health status variable in which 1=poor, 2=fair, 3=good, and 4=excellent. A categorical variable for the individual's years of education has also been used in this calculation. Gender and age group are not included.
(2) The index in the first, third and fifth columns include years of education as a circumstance, whereas the second, fourth, and sixth columns include years of education purged from circumstances.

Table 8 Log-odds Ratios for the Correlates of Self-Assessed Health Status by Age Group
Dependent variable: self-reported health status (0=poor or fair, 1= good or excellent)

	Age group: 25-35		36-50		51-65	
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.7924**	0.7271**	0.5218***	0.5611***	0.5171***	0.5651***
	(0.3150)	(0.3162)	(0.1979)	(0.1990)	(0.2003)	(0.2018)
Ethnicity (Ref. Not a minority):						
Indigenous	-1.2907	-1.2854	0.1894	0.2864	0.0161	-0.1839
	(0.7920)	(0.7921)	(0.5639)	(0.5636)	(0.7473)	(0.7466)
Black/mulato/raizal/palenquero	-0.4458	-0.4976	-0.4391	-0.3624	0.1345	0.1635
	(0.4735)	(0.4720)	(0.3821)	(0.3827)	(0.4120)	(0.4118)
Region (Ref. Atlantic and San Andres islands):						
Eastern	-0.3581	-0.3333	-0.2892	-0.3258	-0.1536	-0.0746
	(0.5251)	(0.5248)	(0.2749)	(0.2757)	(0.2780)	(0.2771)
Pacific	-0.9042*	-0.8942*	-0.6281*	-0.5816*	-0.7038**	-0.6137*
	(0.4620)	(0.4616)	(0.3425)	(0.3406)	(0.3515)	(0.3490)
Orinoquia and Amazonia	0.0000	0.0000	0.2286	0.0964	-0.0296	-0.2552
	(0.0000)	(0.0000)	(0.6933)	(0.6922)	(0.9000)	(0.8989)
Antioquia	0.6988	0.7516	-0.0448	-0.0612	-0.0082	0.0004
	(0.6109)	(0.6142)	(0.3545)	(0.3551)	(0.3351)	(0.3349)
Valle	-0.5004	-0.4549	0.6391	0.6139	-0.1454	0.0859
	(0.7554)	(0.7549)	(0.5005)	(0.5008)	(0.4494)	(0.4488)
Bogota	-0.4951	-0.6110	-0.4700	-0.4831	-0.4874	-0.4970
	(0.6106)	(0.6154)	(0.4525)	(0.4526)	(0.4520)	(0.4521)
Central	0.0089	0.0556	-0.1130	-0.0816	-0.4387	-0.3035
	(0.5189)	(0.5184)	(0.3295)	(0.3291)	(0.3021)	(0.2990)
Born in urban area	0.1192	0.2132	-0.2122	-0.0309	0.0884	0.1610
	(0.4015)	(0.3989)	(0.2100)	(0.2057)	(0.2100)	(0.2093)
Household socioeconomic status at age 10:						
Quintile 2	0.9255*	0.9853*	0.2990	0.4309*	-0.2145	-0.1829
	(0.5268)	(0.5236)	(0.2433)	(0.2407)	(0.2479)	(0.2469)
Quintile 3	0.1625	0.5102	0.8799***	1.1013***	-0.0481	0.1371
	(0.4791)	(0.4726)	(0.2919)	(0.2902)	(0.2784)	(0.2701)
Quintile 4	-0.1975	0.2942	0.5725*	0.9566***	-0.0799	0.2757
	(0.5514)	(0.5258)	(0.3080)	(0.2996)	(0.3666)	(0.3514)
Quintile 5	0.4275	1.0926*	0.9503**	1.4916***	0.0380	0.5081
	(0.6903)	(0.6312)	(0.4699)	(0.4481)	(0.4653)	(0.4386)
Paternal education level (Ref. None):						
Complete primary and incomplete secondary	0.3920	0.5783	0.4352	0.4840	0.0960	0.2276
	(0.4682)	(0.4693)	(0.3598)	(0.3595)	(0.3887)	(0.3890)
Complete secondary or more	0.4664	0.8151	0.3590	0.4415	-0.6995	-0.3575
	(0.6931)	(0.6845)	(0.5559)	(0.5554)	(0.6188)	(0.6162)
Unknown father's level of education	-0.3563	-0.3242	0.3718	0.2627	0.2458	0.1583
	(0.4181)	(0.4205)	(0.3394)	(0.3375)	(0.2855)	(0.2850)
Maternal education level (Ref. None):						
Complete primary and incomplete secondary	-0.9342**	-0.8795*	-0.1845	-0.0424	0.7081*	0.8547**
	(0.4580)	(0.4582)	(0.3050)	(0.3055)	(0.3920)	(0.3913)
Complete secondary or more	1.2847	1.5335	-0.2113	0.1268	1.0861	1.4160*
	(1.0338)	(1.0314)	(0.6177)	(0.6086)	(0.7485)	(0.7471)
Unknown mother's level of education	0.4241	0.4291	-0.5432	-0.7347*	0.1256	0.1816
	(0.5115)	(0.5113)	(0.3760)	(0.3779)	(0.3222)	(0.3224)
Years of education	0.1433***		0.1158***		0.1042***	
	(0.0461)		(0.0259)		(0.0264)	
Years of education purged from circumstances		0.1433***		0.1158***		0.1042***
		(0.0461)		(0.0259)		(0.0264)
Constant	0.4941	1.2478**	-0.0921	0.4363	-0.4141	-0.1043
	(0.5416)	(0.5181)	(0.2871)	(0.2748)	(0.2963)	(0.2885)
Observations	541	541	918	918	735	735
Region of Birth Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	-716710	-716710	-1.823e+06	-1.823e+06	-1.816e+06	-1.816e+06
Pseudo R squared	0.151	0.151	0.113	0.113	0.0817	0.0817

***, **, and * indicate statistical significance at the 1, 5 and 10 percent level, respectively.
Robust standard errors in parentheses
Own calculations. Source: 2010 Colombian LSSM

Table 9 Gini-Opportunity index and Dissimilarity Index of Inequality of Opportunity, with its Decomposition, by Age Group

	Age group: 25-35		35-50		50-65	
Gini-Opportunity Index (1)	0.0331		0.0920		0.1029	
Dissimilarity Index (2)	0.0473	0.0473	0.0720	0.0720	0.1018	0.1018
			Decomposition of the Dissimilarity Index (in %)			
Educational Attainment	21.97		22.88		28.38	
Education purged from circumstances		28.60		38.14		41.30
Circumstances	78.03	71.40	77.12	61.86	71.62	58.70
Early Life Circumstances	50.32	45.51	55.52	42.12	58.42	46.80
Mother's Education	20.47	19.43	9.53	6.86	26.24	21.08
Father's Education	8.78	6.85	9.62	7.61	13.73	12.50
Household Socioeconomic Status at age 10	21.07	19.23	36.37	27.64	18.44	13.23
Demographics	27.71	25.89	21.60	19.74	13.20	11.89
Region of Birth	19.32	18.53	14.88	13.99	6.59	6.71
Born in Urban Area	0.80	0.31	5.54	4.42	6.04	4.76
Ethnicity	7.59	7.04	1.18	1.34	0.58	0.42
Observations	541		918		735	

Bootstrapped standard errors in parentheses. 100 replications.
Own calculations. Source: 2010 Colombian LSSM

Notes:
(1) The Gini-opportunity index is calculated using a self-assessed health status variable in which 1=poor, 2=fair, 3=good, and 4=excellent. A categorical variable for the individual's years of education has also been used in this calculation. Gender and age group are not included.
(2) The index in the first, third and fifth columns include years of education as a circumstance, whereas the second, fourth, and sixth columns include years of education purged from circumstances.

Appendix

Table 10. Summary Statistics: Urban Subsample

Variable	Observations	Mean or Proportion	Std. Dev.
Outcome			
Self-assessed Health Status	1,263	2.85	0.60
Poor	25	2.0%	0.14
Fair	258	20.4%	0.40
Good	856	67.8%	0.47
Excellent	124	9.8%	0.30
Early-life Circumstances			
Household Socioeconomic Status at Age 10			
Quintile 1	265	21.0%	0.41
Quintile 2	252	20.0%	0.40
Quintile 3	253	20.0%	0.40
Quintile 4	243	19.2%	0.39
Quintile 5	237	18.8%	0.39
No information on assets available	13	1.0%	0.10
Education Level of Father			
None or Incomplete Primary	585	46.3%	0.50
Complete Primary and Incomplete Secondary	289	22.9%	0.42
Complete Secondary or More	177	14.0%	0.35
Unknown Father's Education	210	16.6%	0.37
No information on father's education	2	0.2%	0.04
Education Level of Mother			
None or Incomplete Primary	647	51.2%	0.50
Complete Primary and Incomplete Secondary	333	26.4%	0.44
Complete Secondary or More	151	12.0%	0.32
Unknown Mother's Education	130	10.3%	0.30
No information on mother's education	2	0.2%	0.04
Other circumstances			
Ethnicity			
Indigenous	22	1.7%	0.13
Black, mulato, raizal or palenquero	80	6.3%	0.24
No ethnic minority	1,161	91.9%	0.27
Years of Education	1,263	8.83	4.54
Born in Urban Area	899	71.2%	0.45
Born in Rural Area	359	28.4%	0.45
No information on area of birth	5	0.4%	0.06
Region of Birth			
Atlantic	259	20.5%	0.40
Eastern	325	25.7%	0.44
Pacific	74	5.9%	0.23
Orinoquia-Amazonia	5	0.4%	0.06
Antioquia	146	11.6%	0.32
Valle del Cauca	102	8.1%	0.27
Bogotá	153	12.1%	0.33
San Andrés islands	2	0.2%	0.04
Central	197	15.6%	0.36
Additional Controls			
Male	811	64.2%	0.48
Age	1,263	45.13	10.96
Age group			
25-35	275	21.8%	0.41
35-45	315	24.9%	0.43
45-55	385	30.5%	0.46
55-65	288	22.8%	0.42

Source: 2010 Colombian LSSM Survey

Table 11. Summary Statistics: Rural Subsample

Variable	Observations	Mean or Proportion	Std. Dev.
Outcome			
Self-assessed Health Status	990	2.69	0.58
Poor	24	2.4%	0.15
Fair	298	30.1%	0.46
Good	631	63.7%	0.48
Excellent	37	3.7%	0.19
Early-life Circumstances			
Household Socioeconomic Status at Age 10			
Quintile 1	246	24.8%	0.43
Quintile 2	158	16.0%	0.37
Quintile 3	181	18.3%	0.39
Quintile 4	194	19.6%	0.40
Quintile 5	185	18.7%	0.39
No information on assets available	26	2.6%	0.16
Education Level of Father			
None or Incomplete Primary	673	68.0%	0.47
Complete Primary and Incomplete Secondary	88	8.9%	0.28
Complete Secondary or More	17	1.7%	0.13
Unknown Father's Education	212	21.4%	0.41
Education Level of Mother			
None or Incomplete Primary	698	70.5%	0.46
Complete Primary and Incomplete Secondary	114	11.5%	0.32
Complete Secondary or More	20	2.0%	0.14
Unknown Mother's Education	158	16.0%	0.37
Other circumstances			
Ethnicity			
Indigenous	37	3.7%	0.19
Black, mulato, raizal or palenquero	64	6.5%	0.25
No ethnic minority	889	89.8%	0.30
Years of Education	990	4.71	3.66
Born in Urban Area	204	20.6%	0.41
Born in Rural Area	785	79.3%	0.40
No information on area of birth	1	0.1%	0.03
Region of Birth			
Atlantic	248	25.1%	0.43
Eastern	193	19.5%	0.40
Pacific	181	18.3%	0.39
Orinoquia-Amazonia	1	0.1%	0.03
Antioquia	105	10.6%	0.31
Valle del Cauca	58	5.9%	0.23
Bogotá	6	0.6%	0.08
Central	198	20.0%	0.40
Additional Controls			
Male	787	79.5%	0.40
Age	990	44.31	11.06
Age group			
25-35	229	23.1%	0.42
35-45	279	28.2%	0.45
45-55	261	26.4%	0.44
55-65	221	22.3%	0.42

Source: 2010 Colombian LSSM Survey