J Econ Inequal (2016) 14:395–416
DOI 10.1007/s10888-016-9338-2

Inequality of opportunity in adult health in Colombia

Johanna Fajardo-Gonzalez[1]

Received: 31 December 2014 / Accepted: 23 September 2016 / Published online: 29 October 2016

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 early-life circumstances such as parental background, region of origin and ethnicity to inequality of opportunity. The findings suggest that 8 % to 10 % of the circumstance-driven opportunities distinctively enjoyed by those who are healthier should be redistributed or otherwise compensated in order to achieve equality of opportunity. Differences in household socio-economic status during childhood and parental educational attainment appear to be the most salient dimensions of inequality of opportunity in adult health.

Keywords Childhood · Colombia · Health · Inequality · Opportunity

1 Introduction

The 2006 *World Development Report on Equity and Development* highlights 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

Electronic supplementary material The online version of this article (doi:10.1007/s10888-016-9338-2) contains supplementary material, which is available to authorized users.

✉ Johanna Fajardo-Gonzalez
fajar016@umn.edu

1 Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, St. Paul, MN 55108, USA

2006). The traditional focus of policies that aim to reduce health inequity in both developed and developing countries is the reduction of inequality in specific health outcomes as well as in 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. However, 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 et al. 2010; Donni et al. 2014).

Roemer (1998)'s theoretical approach to equality of opportunity suggests 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. 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 to provide 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 2015). 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 each one's circumstances. 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 % 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 the present paper is grounded on Rosa Dias (2009), Trannoy et al. (2010) and Donni et al. (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. 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 % of adult health inequality is due to circumstances.

For developing countries, the literature is very scarce. One of the few studies known is that of Jusot et al. (2014) who investigate 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 striking 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 study 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? This study is one of the first to answer this question using data from a developing country.

Colombia is undergoing rapid demographic changes: The Colombian population predominantly lives in urban areas, is aging as life expectancy at birth has increased from 65 to 75 years in the last 35 years, and its fertility rate has decreased from 4.0 in 1980 to 2.0 births per woman in 2015. Despite overall improvements in many socio-economic outcomes in the last 20 years, health outcomes remain worse in rural areas than in urban areas. Health status varies greatly between rural and urban residents: 32 % of the rural population reports a poor or fair health status whereas 22 % of the urban population reports the same status. Important differences in access to health care services persist between urban and rural areas despite the health system achieved 96 % coverage of the population in 2013 (World Health Organization, 2014). Findings from a few studies (Restrepo et al. 2007; Flórez 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 conditions. 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 et al. (2009). I also compute a dissimilarity index and a Gini-opportunity index as direct measures of inequality of opportunity (Paes de Barros et al. 2008, 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 socio-economic status at age 10 to inequality of opportunity.

The findings suggest that 8 % to 10 % of the circumstance-driven opportunities distinctively enjoyed by those who are healthier should be compensated for or redistributed among those who are less healthy in order to achieve equality of opportunity. Differences in household socio-economic status during childhood and parental educational attainment appear to be the most important dimensions of inequality of opportunity in adult health. Household socio-economic status at age 10 contributes between 15 % and 22 % to the dissimilarity index, whereas parental education between 10 % and 13 %. In contrast, Jusot et al. (2014) suggest ethnicity and region of birth are more important factors for health inequity in Indonesia.

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*) administered by the Colombian Bureau of Statistics (*Departamento Administrativo Nacional de Estadistica* – DANE). This survey provides current and retrospective measures of socio-economic characteristics.

The LSSM is representative of 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 on 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. Table A1 to Table A3 in the Online Supplementary Appendix display a summary of descriptive statistics for the full sample and urban and rural subsamples.

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 (DeSalvo 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 % of the respondents reported a good or an excellent health status whereas 2.2 % reported a poor health status. By area, 78 % of urban residents reported at least a good health status whereas 68 % of rural residents reported the same status.

Self-reported health status has some limitations that have been previously identified in the health literature (Jusot et al. 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 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 socio-economic status at age 10. Parental educational attainment is a categorical variable that indicates whether a parent completed or not a specific level, including primary school, secondary school or higher education. In this sample, approximately 60 % of the heads of household reported that their parents did not attend school or did not complete primary education. In contrast, less than 9 % indicated that their parents completed secondary school or a higher education level. In urban areas, 46 % of fathers and 51 % of mothers did not complete primary education. In rural areas, the percentages for incomplete primary education are even higher: 54 % for fathers and 62 % for mothers.

Household socio-economic status at age 10 is a categorical variable that indicates the quintile group in which a household falls into, based on an asset index following the methodology by Vyas and Kumaranayake (2006).[2] For the full sample, about 25 %[3] of the heads of household are assigned to the first quintile group of the socio-economic index, according to their reports of assets ownership. In urban areas, each of the five quintile groups has approximately the same number of individuals. In rural areas, in contrast, 25 % of individuals belong in the first quintile group. Retrospective data are far from ideal because

[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.

[2]Variables in the socio-economic 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.

[3]The quintile groups 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.

measurement error and recall bias could be also problematic, in particular when income or earnings data are asked. It is still possible to argue that the variables for assets ownership could be remembered with some reasonable accuracy despite observing longer recall intervals for older adults, as suggested by Angulo et al. (2012).

Other variables that are likely to affect individual health status are also considered. In the set of *demographic controls,* I include ethnicity, urban or rural location of birth, and region of birth. About 9 % of heads of household 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.[4] Regarding location of birth, about 28.4 % of current residents in urban areas were born in rural areas, with the younger urban cohorts exhibiting a smaller proportion of rural-born adults. There are substantial socio-economic differences between regions within the country. The World Bank (2015, pp. 45) documents the main regional differences in growth and inequality, which show, in particular, a Gini coefficient of inequality across regions of 0.3 and a large per capita income gap with Bogota. Throughout the analysis, additional controls include gender and age group. In the full sample, about 71 % of household heads are males. The proportion of male household heads is larger in urban (79 %) than rural areas (64 %).

The LSSM does not provide information on individual or parental health-related behaviors, and no effort variables can be measured in these data. The only circumstance that is partly affected by individual effort is years of education. Education is positive and largely associated with health status, health behaviors and preventative service use (Lleras-Muney 2005; Arendt 2005; Cutler et al. 2008). The average number of years of education of the heads of household in this sample is seven years, being larger at 8.4 years for the youngest cohort (25–35 years of age).

3 The measurement of inequality of opportunity in adult health

This section explains the parametric approach used to test for inequality of opportunity following Paes de Barros et al. (2009).[5] 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.

3.1 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

[4]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.

[5]The Online Supplementary Appendix provides a theoretical framework and estimates of inequality of opportunity using a stochastic dominance approach (Lefranc et al. 2009). The empirical tests follow the methodology proposed by Yalonetzky (2013).

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 the following health production function

$$H = f(C, D, e, u) \tag{1}$$

where C is a vector of individual circumstances, D a vector of demographic controls and e a vector of efforts. The residual term u captures luck and other random factors that are not measured by the other variables in the health production function.

Efforts can also be affected by individual circumstances and in most cases are unobserved. In Roemer's definition of equality of opportunity, efforts are assumed orthogonal to circumstances. This assumption suggests that any other determinant of health status that is correlated with circumstances is also understood as a circumstance. One of such variables is educational attainment.

This relationship can be empirically approximated using a non-linear specification:

$$\Pr[H^* = 1|C_i, D_i] = \frac{\exp\{d + C_i a + D_i b\}}{1 + \exp\{d + C_i a + D_i b\}} \tag{2}$$

where H^* denotes a dichotomous health outcome for individual i, C_i the vector of individual circumstances, and D_i demographic characteristics.

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 group of household socio-economic 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*). Therefore, $C_i \equiv \{E_i, FE_i, ME_i, WS_i, LB_i, RB_i, ED_i\}$ and $D_i \equiv \{M_i, AG_i\}$.

In order to estimate the global effect of observed circumstances on health status, I clean years of education of any influence coming from the other observed circumstances. In a related study, Trannoy et al. (2010) propose 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) empirical strategy.

The logistic regression model now takes the following form:

$$\Pr\left[H^* = 1|C_i', \hat{\upsilon}_i^e, D_i\right] = \frac{\exp\{d + C_i' a_1 + \hat{\upsilon}_i^e a_2 + D_i b\}}{1 + \exp\{d + C_i' a_1 + \hat{\upsilon}_i^e a_2 + D_i b\}} \tag{3}$$

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.

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

$$ED_i = k + C_i'g + D_i w + \upsilon_i \tag{4}$$

where υ_i is a disturbance assumed to be normally distributed.

By construction, the residuals $\hat{\upsilon}_i^e$ are orthogonal to circumstances in the equation for health status. These residuals 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 the health status reported by residents in rural areas and respondents in urban areas. To do so, 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 Eqs. 2, 3 and 4.[6] 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.

One contribution of this paper comes from the estimation of Eqs. 3 and 4. 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. A circumstance may also have both direct and indirect effects. In my view, this type of 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 socio–economic 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.

3.2 The dissimilarity index of inequality of opportunity

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 health status ($\hat{p}_i$). In the LSSM sample, 2.2 % of the respondents report a poor health status (category 1), whereas 7.1 % 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 et al. 2008, 2009). The dissimilarity index is a measure proportional to the absolute distance between the distribution of

[6] I retain both significant and insignificant coefficients in the estimation of the dissimilarity index, following Paes de Barros et al. (2009).

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 et al. (2008) show that a consistent estimator for the dissimilarity index for binary outcomes is given by:

$$\widehat{D} = \frac{1}{2\bar{p}} \sum\nolimits_{i=1}^{n} w_i \left| \widehat{p_i} - \bar{p} \right| \tag{5}$$

where $\hat{p}_i$ is the predicted probability of achieving a good or excellent health status for individual $i = 1, \ldots, n$. The estimated conditional probability is $\bar{p} = \sum_{i=1}^{n} w_i \hat{p}_i$, where 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, when an equal proportion of less healthy persons are found in all circumstance groups (Paes de Barros et al. 2008).[7] 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 a number of 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.

Ersado and Aran (2014) also show that the index can only increase when new circumstances are added. Elaborating on the last property, Ferreira and Gignoux (2014) 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.

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 et al. 2009), 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.

[7] 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.

The Gini-opportunity index has been applied to the study in health inequalities by Rosa Dias (2009). The index was first proposed by Lefranc et al. (2009) 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:

$$G_{opp} = \frac{1}{\bar{h}} \sum\nolimits_{i=1}^{k} \sum\nolimits_{i<j} p_i p_j [\bar{h}_j \left(1 - G_j\right) - \bar{h}_i (1 - G_i)] \tag{6}$$

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 et al. (2009) 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 Suarez and Narayan (2011) 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:

$$\Delta D_c = \sum_{M \subset C \setminus \{c\}} \frac{|m|!\,(\kappa - |m| - 1)!}{\kappa!} \left[D(M \cup \{c\}) - D(M)\right] \tag{7}$$

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:

$$S_c = \frac{\Delta \mathrm{D}_c}{D(\kappa)} \text{ where } \sum\nolimits_{i \in C} S_i = 1 \tag{8}$$

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

This section first presents a brief summary of the results obtained using non-parametric statistic tests for stochastic dominance.[8] Lefranc et al. (2009) 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. A non-parametric test suitable for categorical variables was introduced by Yalonetzky (2013). I provide here an extension of this test 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). 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 socio–economic status at age 10 and parental and maternal educational attainment.

The test results, summarized in Table 1, firstly show that the health distribution for the fifth quintile group of socio-economic status at age 10 dominates the health distribution for all but the first quintile group (comparing the fifth and first quintile group, 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 group dominates the distribution for the first and second socio-economic status quintile groups (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 % level. In urban areas, I find that the health distribution for the fifth quintile group dominates each of the distributions for the four remaining quintile groups. In contrast with the urban sample, the statistical test results for rural areas suggest that the only statistically significant dominance relationship is that of the health distribution for quintile group 5 relative to the first and second quintile groups.

Concerning parental education, Table 1 (panel b and panel c) suggests that the higher the levels of paternal and maternal education, the better health opportunities are, especially,

[8] Please see the Online Supplementary Material for further details.

Table 1 Stochastic dominance tests for inequality of opportunity

Quintile group	1 (lowest)	2	3	4	5 (highest)
a. Household socioeconomic status at age 1					
Full sample					
1 (lowest)		~	~	~	~
2	~		~	~	~
3	~	>		~	~
4	>	>	~		~
5 (highest)	~	>	>	>	
Urban Areas					
1 (lowest)		~	~	~	~
2	~	~	~	~	~
3	~	~		~	~
4	>	~	~		~
5 (highest)	>	>	>	>	
Rural Areas					
1 (lowest)		~	~	~	~
2	~		~	~	~
3	~	~		~	~
4	~	~	~		~
5 (highest)	>	>	~	~	
b. Paternal Education					
Level	None	Primary	Secondary and higher		
Full sample					
None *		~	~		
Primary **	>				
Secondary and higher	>	~	~		
Urban Areas					
None *		~	~		
Primary **	>		~		
Secondary and higher	>	~			
Rural Areas					
None *		~	~		
Primary **	~		~		
Secondary and higher	~	~			
c. Maternal Education					
Level	None	Primary	Secondary and higher		
Full sample					
None *		~	~		
Primary **	>		~		
Secondary and higher	>	>			

Table 1 (continued)

c. Maternal Education			
Level	None	Primary	Secondary and higher
None *		∼	∼
Primary **	∼		∼
Secondary and higher	>	>	
Rural Areas			
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

in urban areas. The distribution of the health status of individuals whose fathers have some degree of education dominates the health distribution of individuals whose fathers have no education at all, which is suggestive of inequality of opportunity. 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 Eq. 8, where the variable for individual years of education is purged from the effect of circumstances. The coefficients reported in Table 2 on household socio-economic status at age 10 and parental education are all statistically significant at the 5 % level. In particular, the coefficient on socio-economic status is positive and increasing with quintile group. This result suggests how relevant the capacity of richer households is 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 male and born in the Central region is positively associated with higher educational attainment in the urban subsample, while the opposite is observed in rural areas. There is an important cohort effect in educational attainment in Colombia: younger cohorts in rural areas have had better access to primary and secondary schooling in the past thirty years. A similar trend was documented for Guatemala and other developing countries by UNESCO (2012, Chapter 6).

Table 2 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 (Ref. Quintile group 1):			
Quintile group 2	0.7084***	1.0493***	−0.3497
	(0.2732)	(0.3525)	(0.3114)
Quintile group 3	2.0127***	2.1206***	0.4408
	(0.2874)	(0.3614)	(0.3432)
Quintile group 4	3.4114***	3.1020***	0.7434**
	(0.3255)	(0.3848)	(0.3549)
Quintile group 5	4.5999***	4.2618***	2.2478***
	(0.3554)	(0.4055)	(0.4083)

Table 2 (continued)

Dependent variable: Years of education	All Individuals	Urban Areas	Rural Areas
	(1)	(2)	(3)
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	0.43	0.39	0.24

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

4.2.1 Correlates of health status in the full sample

The first two columns in Table 3 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 suggested by Eq. 6. 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 years of education. Regional differences are also important. Being born in the Pacific or in 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 3 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 Eq. 7. The variable for years of education purged from circumstances has the same

Table 3 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 (Ref. Quintile group 1):						
Quintile group 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 group 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 group 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 group 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)

Table 3 (continued)

Dependent variable: Self–reported health status (0 = poor or fair, 1 = good or excellent)	All Individuals (1)	(2)	Urban Areas (3)	(4)	Rural Areas (5)	(6)
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
Log-likelihood	−4.477e+0	−4.477e+0	−3.328e+0	−3.328e+0	−1.085e+0	−1.085e+0
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

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 socio–economic status during childhood, which becomes highly significant and increasing with the quintile group 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 quintile groups of the socio-economic status variables.

4.2.2 Correlates of health status in the rural and urban subsamples

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

Table 4 Gini-opportunity index and decomposition of the dissimilarity index of inequality of opportunity

	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		22.5
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	

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

I find a positive relationship between reporting a good health status and coming from the fifth quintile group of the socio-economic 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 to 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.

Using the sample for rural residents, I only find a positive and significant relationship between reporting a good health status and high socio-economic status during childhood, only in the comparison of quintile groups 3, 4 and 5 against quintile group 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 2 and 3. Parental socio-economic status and parental education attainment have both direct and indirect effects through the effect of education on self-reported health. Being born in urban areas has an indirect effect, through educational attainment.

The estimated results for the sample of urban residents also support that parental socio-economic status and parental education have both a direct and an indirect effect. In contrast, in rural areas, the effect of parental socio-economic status and parental education is realized through years an education, which is 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 Eqs. 3 and 4, to calculate the dissimilarity index. Table 4 displays the dissimilarity index value as well as its decomposition for the full sample, and for the rural and urban samples.[9] The Gini-opportunity index is also tabulated in Table 4. 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 indicates 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 in Table 4. The dissimilarity index obtained with the LSSM data is about 8.4 %. 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 socio-economic status at age 10 (16 %), mother's education (10 %) and father's education (10.2 %). Once I clean years of education from the influence of circumstances, the decomposition of the index shows an increase in the contributions of socio-economic status at age 10 (22.2 %), mother's education (12.4 %) and father's education (13 %).

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 ranges between 0 and 1, so that the closer to 1, the most unequal the distribution of health status among individuals is. Although the Gini-opportunity index could be decomposed using the Shapley-value, I do not provide estimates of the contribution that each circumstance makes to the index as this paper focuses on the dissimilarity index.

The Gini-opportunity index obtained for the full sample is also larger than that calculated for the United Kingdom by Rosa Dias (2009). In the British Household Panel Survey, inequality of opportunity in adult health ranges between 0.009 and 0.018. In contrast with Rosa Dias, who only uses parental socio-economic status as a circumstance, I use the full set

[9] For the decomposition of the dissimilarity index, I use the user-written command in Stata *hoishapley* (Hoyos Suarez 2013).

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, shown in Table 4, I calculate a dissimilarity index of 7.9 %, when I include years of education in the vector of circumstances. That is, 7.9 % 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 % 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 an increase in the contributions of socio-economic status at age 10 (from 10.5 % to 13.7 %), mother's education (12.9 % to 16.5 %) and father's education (13 % to 14.6 %). For rural areas, the decomposition of the index shows a change in the contributions of region of birth (from 20.2 % and 21.1 %) and socio–economic status at age 10 (from 35 % to 40.5 %), the two circumstances that are most influential in inequality of opportunity in health status in rural areas.

I present two additional sets of results in the Online Supplementary Material. The first set of results include chronic illness and disability as control variables in the logistic regression model. 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. The addition of these measures does not change the association between circumstances and adult health status previously described.

The use of self-reported and retrospective recall data could bias the results obtained here. In order to gauge if there is a systematic bias in how health status is reported, I examine how people perceive their own 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.

In the second set of results, I analyze whether the age of an individual affects their recall of early-life circumstances in a certain direction. I estimate the logistic regression models for three age cohorts: 25–35, 36–50, and 51–65 years old. There are substantial differences by age group. For instance, maternal education seems to be more important for the 50–65 group than for the 35–50 group, for which socio-economic 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 the degree of inequality of opportunity in adult health in Colombia by employing stochastic dominance tests and a decomposition of a dissimilarity index. The empirical results suggest that household socio-economic status and parental education are the most salient early-life circumstances that affect health inequality in adulthood. These circumstances, however, do not reflect how important region of birth or ethnicity may be for different socio-economic groups. Ethnicity, for instance, is 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.

Even though this study provides suggestive evidence on the various sources of adult health inequality, it 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 the country.

Further research on inequality of opportunity in health in Colombia and Latin America should be based on novel longitudinal and administrative data that collect comprehensive information on the parents of tomorrow's children. Recall bias, a limitation of the data used in this study, could be minimized through a proper combination of administrative records and longitudinal information.

Acknowledgments 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 Marc Bellemare, Juan Camilo Chaparro, Elizabeth Davis, Paul Glewwe, Tito Velasco and Florian Wendelspiess for useful suggestions. I thank Alejandro Gaviria for sharing the 2010 Colombian Living Standards and Social Mobility Survey. I acknowledge 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. Replication materials (code and data) are available from the author. All remaining errors are my own.

References

Abras, A., Hoyos Suarez, A., Narayan, A., Tiwari, S.: Inequality of opportunities in the labor market: Evidence from life in transition surveys in Europe and Central Asia. IZA J. Labor Dev. **2**(7), 1–22 (2013)

Angulo, R., Azevedo, J., Gaviria, A., Paez, G.: Movilidad social en Colombia. Documentos CEDE **43** (in Spanish) (2012)

Arendt, J.N.: Does education cause better health? A panel data analysis using school reforms for identification. Econ. Educ. Rev. **24**(2), 149–160 (2005)

Cutler, D.M., Lleras-Muney, A., Vogl, T.: Socio-economic status and health: dimensions and mechanism. National Bureau of Economic Research Working Paper No. 14333 (2008)

DeSalvo, K., Fan, V.S., McDonnell, M.B., Fihn, S.D.: Predicting mortality and healthcare utilization with a single question. Health Serv. Res. **40**(4), 1234–46 (2005)

Donni, P.L., Peragine, V., Pignataro, G.: Ex-Ante and ex-post measurement of equality of opportunity in health: A normative decomposition. Health Econ. **23**(2), 182–198 (2014)

Ersado, L., Aran, M.: Inequality of opportunity among Egyptian children. World bank policy research working paper 7026. The world bank, Washington (2014)

Ferreira, F., Gignoux, J.: The measurement of educational inequality: Achievement and opportunity. World Bank Econ. Rev. **28**(2), 210–246 (2014)

Fleurbaey, M., Schokkaert, E.: Unfair inequalities in health and health care. J. Health Econ. **28**(1), 73–90 (2009)

Flórez, C., Soto, V., Acosta, O., Karl, C., Misas, J., Forero, N.: Avances y desafíos de la equidad en el sistema de salud colombiano. Fundación Corona, Bogotá (in Spanish) (2007)

Hoyos Suarez, A.: Hoishapley: Stata module to perform Shapley decomposition of the human opportunity index. Statistical software components archive. Boston College, Boston, MA. https://ideas.repec.org/c/boc/bocode/s457618.html (2013)

Hoyos Suarez, A., Narayan, A.: Inequality of opportunities among children: How much does gender matter? Working Paper 76511. The world bank, Washington (2011)

Idler, E., Benyamini, Y.: Self-rated health and mortality: A review of twenty-seven community studies. J. Health Econ. **23**(6), 1083–99 (1997)

Jusot, F., Tubeuf, S., 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)

Jusot, F., Mage, S., Menendez, M.: Inequality of opportunity in health in Indonesia. DIAL (Développement, institutions et mondialisation) working papers DT/2014–6 (2014)

Lefranc, A., Pistolesi, N., Trannoy, A.: Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France. J. Public Econ. **93**(11), 1189–1207 (2009)

Lleras-Muney, A.: The relationship between education and adult mortality in the United States. Rev. Econ. Stud. **72**(1), 189–221 (2005)

Paes de Barros, R., Molinas, J., Saavedra, J.: Measuring inequality of opportunities for children. The world bank, Washington (2008)

Paes de Barros, R., Ferreira, F., Molinas, J., Saavedra, J.: Measuring inequality of opportunities in Latin America and the Caribbean. The world bank, Washington (2009)

Ramos, X., Van de Gaer, D.: Empirical approaches to inequality of opportunity: Principles, measures, and evidence (2015)

Restrepo, J.H., Zambrano, A., Vélez, M., 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. Universidad del Rosario, Bogotá (2007)

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

Rosa Dias, P. In: Culyer Anthony, J. (ed.) Equality of opportunity in health, Encyclopedia of health economics, vol. 1, pp. 282–287. Elsevier, San Diego (2014)

Rosa Dias, P.: Inequality of opportunity in health: Evidence from a UK cohort study. Health Econ. **18**(9), 1057–1074 (2009)

Rosa-Dias, P., Jones, A.: Giving equality of opportunity a fair innings. Health Econ. **16**(2), 109–112 (2007)

Sen, A.: Health: Perception versus observation: Self-reported morbidity has severe limitations and can be extremely misleading. Br. Med. J. **324**(7342), 860–861 (2002)

Shorrocks, A.F.: Decomposition procedures for distributional analysis: A unified framework based on the Shapley-value. J. Econ. Inequal. **10**(1), 1–28 (2012)

Trannoy, A., Tubeuf, S., Jusot, F., Devaux, M.: Inequality of opportunities in health in France: A first pass. Health Econ. **19**(8), 921–938 (2010)

UNESCO: World Atlas of Gender Equality in Education. UNESCO Publishing (2012)

van Doorslaer, E., Gerdtham, U.G.: Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. Soc. Sci. Med. **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). Universitat Autónoma de Barcelona, Barcelona (in Spanish) (2009)
Vyas, S., 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. Econ. Rev. **32**(1), 126–163 (2013)
Yalonetzky, G.: A dissimilarity index of multidimensional inequality of opportunity. J. Econ. Inequal. **10**(3), 343–373 (2012)
World Bank: Colombia: Systematic country diagnostic. The World Bank, Washington (2015)
World Bank: World development report: Equity and development. The World Bank, Washington (2006)
World Health Organization: Country cooperation brief strategy. The World Health Organization, Geneva (2014)