WPS7514

Policy Research Working Paper 7514

Unpacking the MPI

A Decomposition Approach of Changes in Multidimensional Poverty Headcounts

Jorge Eduardo Pérez Pérez
Carlos Rodríguez Castelán
Jose Daniel Trujillo
Daniel Valderrama

WORLD BANK GROUP
Poverty and Equity Global Practice Group
December 2015

Policy Research Working Paper 7514

Abstract

Multidimensional measures of poverty have become standard as complementary indicators of poverty in many countries. Multidimensional poverty calculations typically comprise three indices: the multidimensional headcount, the average deprivation share among the poor, and the adjusted headcount ratio. While several decomposition methodologies are available for the last index, less attention has been paid to decomposing the multidimensional headcount, despite the attention it receives from policy makers. This paper proposes an application of existing methodologies that decompose welfare aggregates—based on counterfactual simulations—to break up the changes of the multidimensional poverty headcount into the variation attributed to each of its dimensions. This paper examines the potential issues of using counterfactual simulations in this framework, proposes approaches to assess these issues in real applications, and suggests a methodology based on rank preservation within strata, which performs positively in simulations. The methodology is applied in the context of the recent reduction of multidimensional poverty in Colombia, finding that the dimensions associated with education and health are the main drivers behind the poverty decline.

This paper is a product of the Poverty and Equity Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at jorge_perez@brown.edu, crodriguezc@worldbank.org, jdtrujillos@dane.gov.co, and dvalderramagonza@worldbank.org.

Unpacking the MPI: A Decomposition Approach of Changes in Multidimensional Poverty Headcounts*

Jorge Eduardo Pérez Pérez†

Carlos Rodríguez Castelán‡

José Daniel Trujillo§

Daniel Valderrama¶

Keywords Multidimensional Poverty, Decomposition and Non Parametric Methods, Shapley Value, Poverty Transitions

JEL Codes I32, I38, C14

*The authors would like to thank Paola Ballón, Wendy Cunningham, James Foster, and participants of the 2014 Summer Initiative for Research on Poverty, Inequality and Gender of the Poverty Global Practice at the World Bank, and the 2015 LACEA meeting for their useful comments and suggestions. The views, findings, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect those of the World Bank, its Executive Board or member country governments. Pérez Pérez gratefully acknowledges financial support from Fulbright-Colciencias.The do-file code that implement this decomposition could be provided upon request

†Corresponding author. Department of Economics, Brown University.
Email: jorge_perez@brown.edu

‡World Bank - Poverty Global Practice. Email: crodriguezc@worldbank.org

§DANE. Email: jdtrujillos@dane.gov.co

¶World Bank - Poverty Global Practice. Email: dvalderramagonza@worldbank.org

1 Introduction

Calculating multidimensional measures of poverty has become commonplace in many developing countries, as a way to complement traditional monetary indicators. The breadth of multidimensional measures, compared to traditional approaches, is viewed as an advantage by researchers, enabling them to aggregate a large number of welfare-associated variables into a single measure. Multidimensional measures are also attractive to governments, as they are conducive to a more precise understanding of the determinants of individual welfare and quality of life, allowing the identification of those fields where public policy may have a larger impact.

Despite this popularity, multidimensional measures are not free from the criticism faced by other approaches to measuring poverty. As with any indicator that attempts to summarize a complex phenomenon into a single index, multidimensional poverty measures may be hard to interpret if unaccompanied by additional information. In this sense, rather than focusing solely on the measure that aggregates across dimensions, analyses of multidimensional poverty also tend to contain separate information about each dimension. This, however, increases the number of indicators to be tracked, and reduces the usefulness of the summary measures.

These difficulties become compounded when tracking the evolution of multidimensional measures across time. An exact identification of the contribution of each dimension to the evolution of multidimensional measures would require tracking the transition of individuals in and out of poverty, and the dimensions by individual. This exhaustive analysis would defeat the purpose of aggregation of poverty measures.

Decomposition approaches, whereby poverty measures are split up into the contribution of broadly defined determinants of interest, are a reasonable compromise between a comprehensive analysis and a completely aggregated one. As expressed in Ferreira and Lugo (2013), these approaches aim for a "middle ground", which can be informative of multidimensional poverty as well as expeditious. While these approaches are deeply rooted in the traditional monetary poverty research, they have been less developed in the literature of multidimensional measurement. Apablaza and Yalonetzky (2013) propose approaches to decompose multidimensional poverty measures statically or dynamically. Yet, their focus is on two out of the three measures of multidimensional poverty: the average deprivation share among the poor and the adjusted headcount ratio. This is natural since these indicators have additive separability properties that make them easy to decompose, as opposed to the multidimensional headcount, which is nonlinear in terms of deprivation scores. Decompositions of the multidimensional headcount ratio, in particular dynamic ones, appear to be lacking in attention.

Overlooking the multidimensional headcount ratio is conspicuous, considering that this indicator is particularly important on its own. From an academic perspective, focusing on the evolution of the adjusted headcount ratio as opposed to the multidimensional headcount ratio disregards variations in poverty that come only from changes in the number of poor and arise only in the identification step. Using datasets form several countries, Apablaza et al. (2010) and Apablaza and Yalonetzky (2013) show that declines in the adjusted headcount ratio are mostly due to changes in the multidimensional headcount and not to changes in intensity. From a policy perspective, the multidimensional headcount is frequently used by governments as ”the rate of multidimensional poverty”—instead of the adjusted headcount ratio—as it is easily communicable and also comparable to the monetary rate of poverty.

In this paper, we propose an approach to decompose variations in the multidimensional headcount ratio into changes attributed to different categories of dimensions. Our approach builds on counterfactual simulation approaches, which have traditionally been used to decompose poverty and inequality indicators. These approaches were first proposed by Barros et al. (2006), and then extended in a series of papers by Azevedo et al. (2012b, 2013a,b). The approach relies, first, on expressing the indicator of interest as a function of the distribution of its determinants over a finite population. Then it replaces these distributions by counterfactual ones, where one of the determinants has been altered.

Our approach allows us to estimate the extent to which an observed change in a dimension—or category of dimensions—can explain the observed change in the multidimensional headcount ratio. In the presence of panel data, it allows doing this without separately tracking the transitions in and out of poverty of each individual. In repeated cross-section data—when such tracking is impossible—the approach constitutes a good approximation of how much each dimension contributes to the headcount's change. In both cases, it summarizes separate information on the headcount by dimension, weights and incidence.

Our paper contributes to the literature on the decomposition of poverty measures, providing a guideline to split up the changes of the multidimensional poverty headcount into the variation attributed to each of its dimensions. By combining our approach to decompose changes in the multidimensional headcount, and the approach to decompose changes in the average deprivation share among the poor of Apablaza and Yalonetzky (2013), a decomposition of changes in the adjusted headcount ratio by dimension can be obtained.

To illustrate the approach, we apply the decomposition analysis in the context of the recent decline of multidimensional poverty in Colombia, between 2008 and 2012. The results show that education and health were the largest drivers behind the poverty decline. In this sense, the paper also contributes to the literature on poverty in Colombia, providing insights generally absent in the analysis of standard measures.

Section 1 summarizes the technical aspects of the multidimensional poverty measurement, and sets up the analytical framework for the rest of the paper. In Section 2 we review existing approaches to decompose multidimensional poverty measures, covering both static and dynamic analyses. The technical aspects of the counterfactual simulation methodology are reviewed in section 3, in order to identify potential issues that may arise when applying this methodology to the multidimensional headcount. These issues are further analyzed in sub-section 3.2, which provides some practical guidelines to address them.

In section 4 the methodology is used in repeated cross-sections looking at the recent evolution of multidimensional poverty in Colombia. Conducting different simulations, we identify a method based on stratification and income ranks that performs well in several scenarios. Section 5 concludes.

2 Multidimensional Poverty Measures and Their Decompositions

This section provides a review of the existing multidimensional poverty measures. It presents a description of the methodologies used to decompose the measures by dimensions and over time, outlining the difficulties associated with decomposing changes in the adjusted headcount ratio by dimension.

2.1 The Multidimensional Poverty Index measures

We follow Alkire and Santos (2010) and Alkire and Foster (2011) in the presentation of the multidimensional poverty index (MPI), with the distinction that we do not focus on the identification, censoring and aggregation steps. Additionally, we depart from the deprivation matrix notation and, instead, describe the index in terms of random variables in order to ease the transition to our discussion in the next section.

Let $i = 1, 2, \ldots, n$ index individuals. Let $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iD})$ be a vector of *achievements* for individual i in dimensions $d = 1, 2, \ldots, D$. Let c_d be the *deprivation cut-off* of dimension d. An individual i is said to be *deprived* in dimension d if $y_{id} < c_d$. Now let $\mathbf{w} = (w_1, w_2, \ldots, w_D)$ be a vector of weights given to each dimension, such that $w_d \geq 0$ and $\sum_{d=1}^{D} w_d = 1$[1]. Individual i is said to be *multidimensionally poor* if

$$p_i \equiv \mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id} < c_d\right) > k\right) = 1 \tag{1}$$

where $\mathbb{1}$ is the indicator function; and k is called the *cross-dimensional cutoff*. Simply put, an individual is multi-dimensionally poor if a weighted sum of deprivation indicators falls below a pre-specified threshold. The amount $\sum_d w_d \mathbb{1}\left(y_i d < c_d\right)$ is called the *deprivation share*.

From this individual measure of poverty, three population wide measures are built. The *multidimensional headcount ratio* is the proportion of the population that is multi-dimensionally poor. It measures the incidence of multidimensional poverty over the population:

$$\begin{aligned} H &\equiv \frac{1}{n}\sum_{i=1}^{n} p_i \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id} < c_d\right) > k\right) \end{aligned} \tag{2}$$

Defining $p \equiv \sum_{i=1}^{n} p_i$, the *average deprivation share among the poor* is

$$A \equiv \frac{1}{p}\sum_{i=1}^{n} p_i \left[\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id} < c_d\right)\right] \tag{3}$$

[1]Note that we define the weights as adding up to 1 and not to the number of dimensions.

which measures the intensity of poverty in the population among the multi-dimensionally poor. Finally the *adjusted headcount ratio* is defined as

$$M_0 \equiv HA \tag{4}$$

which adjusts the headcount by the intensity of poverty.

Alkire and Foster (2011) focus on the M0 measure, due to its desirable properties. These include monotonicity in the number of deprived dimensions and important decomposition properties, which we outline in the next section. The headcount ratio H, however, tends to receive wider attention in policy circles because, as a simple population proportion, its level is immediately comparable with traditional income-based poverty rates.

2.2 Decomposing the measures

The three measures outlined provide a one-dimensional summary of the incidence and intensity of poverty for the population as a whole. In order to examine the particular determinants of poverty, these measures may be decomposed into specific indexes designed to see which factors contribute *more*.

Several decomposition methodologies exist, which can be classified into two broad categories: static and dynamic. Static methodologies decompose a single observation into contributions from cross-sectional determinants, while dynamic ones decompose the time-variation of the measure into the contribution of time-varying components. While the present paper focuses in a particular type of dynamic decomposition, static methodologies are described briefly next.

Static decompositions may be of two types: group decompositions and dimensional decompositions. Group decompositions are customary in the poverty measurement literature and decompose poverty measures into the contributions of particular individual groups. As shown in Alkire and Foster (2011), all the measures considered are decomposable into individual groups. For the headcount ratio, if there are two population groups, 1 and 2, with populations n_1 and n_2, the headcount ratio is:

$$H = \frac{n_1}{n}H_1 + \frac{n_1}{n}H_2$$

Where $H_1$and H_2 are the headcount ratios for each group.

Dimensional decompositions break up measures into the contribution of each dimension. From equation 3, it is clear that the A measure is dimensionally decomposable, with contributions equal to:

$$\frac{1}{p}\sum_{i=1}^{n} p_i w_d \mathbb{1}\left(y_{id} < c_d\right). \tag{5}$$

Alkire and Foster (2011) also show that the M_0 measure is dimensionally decomposable. Defining the *censored headcount ratio* H_d as the proportion of people deprived in dimension d among the poor:

$$H_d = \frac{1}{n}\sum_{i=1}^{n} p_i \mathbb{1}\left(y_{id} < c_d\right) \tag{6}$$

Then, M_0 can be decomposed by dimensions as:

$$M_0 = \sum_{d=1}^{D} w_d H_d \tag{7}$$

Conversely, the headcount ratio H is not decomposable by dimensions, since, by construction, H is a nonlinear function of the contribution of each dimension—as reflected in equation 2.

Dynamic decompositions, on the other hand, focus on splitting up the variation of the measures over distinct periods of time into time-variation from its components. These components may or may not be further decomposed into cross-sectional ones.

As shown in Apablaza et al. (2010), from equation 4, a simple dynamic decomposition of the percent variation in M_0, $\Delta\%M_0$, is:

$$\Delta\%M_0 = \Delta\%H + \Delta\%A + \Delta\%H\Delta\%A \tag{8}$$

Static and dynamic decompositions can be combined. For instance, in the previous equation, one could decompose H and A statically in each period, and then split up the changes in M_0 into changes in the cross-sectional components previously obtained. Such decompositions are available, as long as the indicator is decomposable cross-sectionally. This approach is used by Apablaza et al. (2010), who exploit the fact that the headcount is decomposable across groups, and that the average deprivation share is decomposable across dimensions, to extend this result by decomposing the components of equation 8. They show that the percent variation in H can be further decomposed into population changes within groups, changes in the headcount within groups, and a multiplicative effect. Furthermore, the percent variation in A can be decomposed into the weighted sum of percent variations of each of its dimensional components.

However, if the indicator is not cross-sectionally decomposable, this approach fails. This is the case when attempting to decompose the changes of the headcount ratio H into the variation attributed to each of its dimensions. Due the nonlinearity of H , an explicit closed-form solution for this decomposition is not feasible. However, a counterfactual simulation methodology may be used to work around this. That is the topic of the following section.

3 Methodology

The decomposition approaches described so far are not appropriate to look into the dimensions responsible for changes in the headcount ratio H over time. In this section, we analyze the problem of decomposing changes in H, and describe a counterfactual simulation methodology to address the issue. First, we outline the overall problem of decomposing changes in H. Then we summarize the counterfactual simulation methodology based on Barros et al. (2006), Azevedo et al. (2013a) and Azevedo et al. (2013b). The section concludes by addressing the advantages as well as the potential caveats of applying this method to the headcount ratio H.[2]

Let us assume that there are two observations of the multidimensional headcount H^t, for $t = 1, 2$; the first, 1, corresponds to the initial observation while 2 corresponds to the final observation. We also observe the two associated datasets of information on achievements $\mathbf{y}_i^t : \mathbf{y}_i^1 = (y_{i1}^1, y_{i2}^1, \ldots, y_{iD}^1)\,,\ i = 1, 2, \ldots n^1$ and $\mathbf{y}_j^2 = \left(y_{j1}^2, y_{j2}^2, \ldots, y_{jD}^2\right)\,,\ j = 1, 2, \ldots, n^2$.. Additionally, we observe a set of demographic variables $\mathbf{z}_i^1, \mathbf{z}_j^2$. In the case of panel data, individuals can be tracked across time, in which case variables are indexed by the same index, i, at both periods and $n^1 = n^2$. Otherwise, the datasets refer to repeated cross-sections. The *change in the*

[2]The do-file code that implement this decomposition could be provided upon request

multidimensional headcount ratio is the difference between the two indicators [3]:

$$\begin{aligned}\Delta H &= H^2 - H^1 \\ &= \frac{1}{n^2}\sum_{j=1}^{n^2} \mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{jd}^2 < c_d\right) > k\right) \\ &- \frac{1}{n^1}\sum_{i=1}^{n^1} \mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id}^1 < c_d\right) > k\right) \end{aligned} \tag{9}$$

The goal of decomposing the changes into the different dimensions is to be able to express the change in the headcount ratio as a sum of the changes attributed to each dimension S_d :

$$\Delta H = \sum_{d=1}^{D} S_d \tag{10}$$

Several remarks are in order. Notice, first, that ΔH is not decomposable by dimensions. It is not decomposable in terms of the censored headcount ratios defined in section 2.2. Second, it is a nonlinear function of the contributions of each dimension to the average deprivation share, due to the nonlinearity of the indicator functions that involve y_{id} in equation 2: i.e., only the individuals below the dimension specific cut-off contribute to the headcount. Third, while it is clear that the weights play a role in the determination of the contribution of each dimension to the headcount, they remain constant across time and are not the drivers of the changes in H.

Note that by combining a decomposition of ΔH and the decomposition of ΔA that arises from equation 3, a dimensional decomposition of $\Delta\% M_0$ can be achieved. As Apablaza and Yalonetzky (2013) show, $\Delta\% A$ can be decomposed as:

$$\Delta A = A_2 - A_1 = \sum_{d=1}^{D} \Delta\left[\sum_{i=1}^{n} \frac{p_i}{p} w_d \mathbb{1}\left(y_{id} < c_d\right)\right] = \sum_{d=1}^{D} S_A \tag{11}$$

[3] In the case of panel data, this simplifies to

$$\Delta H = \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c} \mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id}^2 < z_d\right) < k\right) \\ -\mathbb{1}\left(\sum_{d=1}^{D} w_d \mathbb{1}\left(y_{id}^1 < z_d\right) < k\right)\end{array}\right]$$

Combining equation this with equations 8 and 10 yields:

$$\begin{aligned}\Delta\%M_0 &= \Delta\%H + \Delta\%A + \Delta\%H\Delta\%A \\ &= \sum_{d=1}^{D}\frac{S_d}{H} + \sum_{d=1}^{D}\frac{S_A}{A} + \Delta\%H\Delta\%A \end{aligned} \tag{12}$$

This decomposition breaks changes in the adjusted headcount ratio into changes in the number of poor attributed to each dimension, S_d, changes in the intensity of multidimensional poverty attributed to each dimension, S_A, and an interaction effect. The logic behind larger contributions of dimensions in explaining the changes in intensity is very clear: if the majority of poor are no longer deprived in a particular dimension, this dimension will contribute more to the decrease in intensity.

In the case of panel data, larger contributions of a dimension to changes in H are also intuitive. As an illustrative example, consider a case where the change in the headcount ratio occurs over a short period of time so that the demographic variables z remain constant. Assume that many dimensions change, but each individual only experiences changes in one dimension. If all individuals whose deprivation share crossed the multidimensional cut-off experienced change in the same dimension, then this dimension would be the unique contributor to the change. Indeed, Apablaza and Yalonetzky (2013) show that the change in the headcount in the case of panel data can be decomposed into a weighted difference in the transition probabilities of moving in and out poverty. Individuals at the margin of transition may be more prone to changes in particular dimensions, which would influence the transition probabilities and would contribute more to the variation in the headcount.

In the case of repeated cross-section data, changes in the headcount biased towards a particular dimension may arise due to the inclusion of people with different deprivation profiles in the sample, at one of the points of time. Thus, changes would be biased by the inclusion of larger shares of people deprived in one dimension. In applied work, dimensions do not change one at a time for each individual; dimensions may be correlated; and cross-sectional data is the rule and not the exception, especially in the surveys used to measure poverty in developing countries. A general framework to decompose changes in H should thus take into account the notion that dimensions are jointly distributed, that certain demographic profiles are more likely to experience changes in particular dimensions, and that only individuals with similar demographic profiles should be compared. This is the topic of the following section.

3.1 The decomposition methodology

We propose the use of a counterfactual simulation methodology, first suggested by Barros et al. (2006), to decompose changes in H additively across dimensions. This section describes the methodology, following closely Barros et al. (2006), and Azevedo et al. (2012b, 2013a,b).

From equation 2, H can be written as function[4] of the joint distribution of the vector of achievements $\mathbf{y}$, the weights $\mathbf{w}$ and the cut-offs $\mathbf{z}$ across the population: $H = \Phi\left(F_{\mathbf{y},\mathbf{w},\mathbf{z}}\right)$. However, since the weights and the cut-offs do not change across individuals, H can be considered a function only of the *deprivation score by dimension*, defined as:

$$x_d = w_d \mathbb{1}\left(y_{id} < c_d\right) \tag{13}$$

With $\mathbf{x} = (x_1, \ldots, x_D)$, H can be written as

$$H = \Phi\left(F_{\mathbf{x}}\right) = \Phi\left(F_{(x_1,\ldots,x_D)}\right) \tag{14}$$

Barros et al. (2006) show that, in finite populations, bivariate joint distributions can be characterized by three functions: the two marginal distributions of each variable, and a function that describes their association. If we define the *ranking of individual $\tilde{\imath}$ according to the random variable x_d* as the position of the individual in a list sorted by the value of random variable

$$R_{y_d}\left(i\right) = \#\left(i \in \{1, \ldots, n\} : x_{id} \leq x_{\tilde{\imath}}\}\right) \tag{15}$$

Then, according to Barros et al. (2006), for the two variables $x_{d'}$ and $x_{d''}$, their joint distribution is completely characterized by

$$F_{x_{d'},x_{d''}} = \left(F_{x_{d'}}, F_{x_{d''}}, R_{x_{d''}}\left(R^{-1}_{x_{d'}}\right)\right) \tag{16}$$

Where $R_{x_{d''}}\left(R^{-1}_{x_{d'}}\right)$ is the ranking according to $x_{d''}$ of a observation with of rank

[4] Notice that Φ need not be invertible.

$R_{x_{d'}}$ according to $x_{d'}$. The function characterizes the rank dependence between the two variables[5]. We call this function the *association* between $x_{d'}$ and $x_{d''}$ and denote it as $C\left(x_{d'}, x_{d''}\right)$. In the multivariate case, the joint distribution can be characterized by all the marginals, along with either all the pairwise associations between the variables, or simply with the association of each variable to a reference variable r, from which the pairwise associations can be obtained. The reference variable may either be one of the deprivation scores by dimension or a demographic variable:

$$F_{\mathbf{x}} = \left(F_{x_1}, F_{x_2}, \ldots, F_{x_D}, C\left(x_1, r\right), C\left(x_2, r\right), \ldots, C\left(x_D, r\right)\right) \tag{17}$$

With this representation in hand, Barros et al. (2006) show that decomposing changes in a welfare aggregate into two components can be achieved by sequentially changing the marginals and the association in the joint distribution. In the case of H, the random variables considered are the deprivation scores by dimension, and the change in equation 9 can be rewritten, using equations 14 and 17, as:

$$\begin{aligned} \Delta H = {} & H^2\left(F_{x_1^2}, F_{x_2^2}, \ldots, F_{x_D^2}, C\left(x_1^2, r^2\right), C\left(x_2^2, r^2\right), \ldots, C\left(x_D^2, r^2\right)\right) \\ & - H^1\left(F_{x_1^1}, F_{x_2^1}, \ldots, F_{x_D^1}, C\left(x_1^1, r^1\right), C\left(x_2^1, r^1\right), \ldots, C\left(x_D^2, r^1\right)\right) \end{aligned} \tag{18}$$

A Barros et al. (2006) decomposition for ΔH in the case of two dimensions, $D = 2$, would then be:

$$\begin{aligned} \Delta H &= S_1 + S_2 + S_{12} \\ S_1 &= H\left(F_{x_1^2}, F_{x_2^2}, C\left(x_1^2, x_2^2\right)\right) - H\left(F_{x_1^1}, F_{x_2^2}, C\left(x_1^2, x_2^2\right)\right) = H^2 - H\left(F_{x_1^1}, F_{x_2^2}, C\left(x_1^2, x_2^2\right)\right) \\ S_2 &= H\left(F_{x_1^1}, F_{x_2^2}, C\left(x_1^2, x_2^2\right)\right) - H\left(F_{x_1^1}, F_{x_2^1}, C\left(x_1^2, x_2^2\right)\right) \\ S_{12} &= H\left(F_{x_1^1}, F_{x_2^1}, C\left(x_1^2, x_2^2\right)\right) - H\left(F_{x_1^1}, F_{x_2^1}, C\left(x_1^1, x_2^1\right)\right) = H\left(F_{x_1^1}, F_{x_2^1}, C\left(x_1^2, x_2^2\right)\right) - H^1 \end{aligned} \tag{19}$$

Where $H\left(F_{x_1^1}, F_{x_2^2}, C\left(x_1^2, x_2^2\right)\right)$ is the counterfactual headcount that would be observed if y_1 were distributed as in period 1, but x_2 and the association remained the

[5] In practice, there can be ties in these rankings. This is inconsequential since, as Barros et al. (2006) notes, this can be solved by randomizing the ranking for the tied cases.

same. It is important to note that counterfactuals are purely exercises to examine changes that occur *ceteris paribus*, and do not intend to reflect equilibrium outcomes (Azevedo et al., 2012a). To compute this counterfactual, we can calculate H from the distribution of $(\hat{x}_1^1, x_2^2)$, where

$$\hat{x}_1 = F^{-1}_{x_1^1}\left(F_{x_1^2}(x_1)\right) \tag{20}$$

Other counterfactuals may be obtained accordingly, requiring inversion of the distribution functions at each step. From here on, we refer to this method as the Barros decomposition.

Azevedo et al. (2012b, 2013a,b) have proposed several improvements to the Barros decomposition to have a broader applicability. Their applications focus mainly on the decomposition of poverty and inequality indicators, but the same improvements could be applied to other measures.

For both panel data and repeated cross-section applications, these studies consider the multivariate case rather than the Barros bivariate one. They propose to keep the associations C constant and they add each variable sequentially, such that no effect is attributed to the interactions. In terms of equation 19, this allows decomposing ΔH into just two components S_1 and S_2 , as stated in the original problem in equation 10. Barros et al. (2006) show that the counterfactuals need to be changed in the multivariate case in order to hold the associations constant. Thus, instead of $\hat{x}_1$ from equation 20, they build different counterfactuals $\tilde{x}_d$, as:

$$\begin{aligned}\tilde{x}^1_{id} &= x^1_{i^*d} \\ i^* &= C\left(\hat{x}_d, x_d(i)\right) = R^{-1}_{\hat{x}_d}\left(R_{x_d^2}(i)\right)\end{aligned} \tag{21}$$

Additionally, these studies note that the Barros decomposition is path dependent: the order in which the variables are replaced by their counterfactuals shifts the result of the decomposition. This is addressed by using a Shapley value decomposition approach based on Shorrocks (2013) (for details see Azevedo et al. (2012b)). Basically, the authors compute the decomposition along each permutation of the y vector and calculate the average of all the contributions obtained[6].

[6]This has the disadvantage of making the method dependent on whether the variables are added together. See Azevedo et al. (2012a) for details.

These papers also address specific problems that arise in the presence of repeated cross-section data. The first is the construction of the counterfactual variables $\hat{x}_d$. In the case of panel data, this is straightforward, for instance in the previous example: $\tilde{x}_i^1 = x_i^1$.. In repeated cross-section data, however, $\tilde{x}_d$ can only be obtained using equations 20 and 21 if F and R are strictly monotone, but this is rarely the case. To work around this, Azevedo et al. (2012b) assume rank preservation in income, and then use income rank as a reference variable to track individuals across the two periods. This amounts to replacing $\hat{x}_d$ with income in equation 21, and avoiding using equation 20 completely. The counterfactual for an individual with a x_d^2 value is an individual who has the same income rank corresponding to x_d^2 in period 1.

In practice, any reference variable could substitute $\hat{x}_d$ if rank preservation is plausible and its ranking function is invertible. While in some cases it may not be plausible for rank preservation to hold unconditionally, it may hold *conditional* on the demographic variables $\mathbf{z}$. In this instance, it is necessary to stratify the data, computing income ranks within strata previously defined by the demographic variables before applying equation 21.

A second problem associated with repeated cross-section data refers to unequal sizes in the available data across time. This issue is pervasive in applied work, since survey samples typically become larger over time. To address this, Azevedo et al. (2012b) suggest rescaling the ranking in the larger dataset to match the smaller dataset, which necessarily generates observations with the same ranking. Then, each observation in the smaller dataset is matched with a randomly selected observation with the same ranking in the rescaled larger dataset.

A third problem that may arise in this line refers to variation in the survey sampling schemes across cross-section samples. This can be addressed by making the weights compatible across samples through a reweighting procedure.

So far we have described the counterfactual simulation methodology, outlining how it can be applied to a decomposition of the multidimensional headcount. We have also reviewed several issues that can arise when the methodology is used in applied work. We have yet to explore how likely it is that these issues appear when decomposing the headcount ratio H, instead of the indicators studied in the literature so far. The next section examines these practical issues more closely.

3.2 Applying the decomposition methodology to H

This section discusses the practical difficulties of applying the methodology described—based on counterfactual simulations—to the multidimensional headcount H. Specifically, it focuses on the problems reviewed in the previous section. Our purpose is to provide a practical guide that can be followed when applying the methodology suggested in this paper. From here on, we will refer to the Barros decomposition with the modifications by Azevedo et al. (2012b, 2013a,b) as the ASN decomposition.

The practical difficulties of applying the methodology can be summarized in four issues. These are listed next, followed by a brief description of the approaches suggested to address them.

1. The ASN methodology disregards interactions across dimensions. To address this, dimensions can be grouped into categories with small interactions between them.

2. By disregarding interactions, the ASN methodology assumes that interaction across dimensions remains constant over time. This too may be addressed if dimensions are grouped into categories where interaction remains constant over time.

3. The ASN methodology is not applicable with discrete random variables as reference variables. A continuous variable, such as income, needs to be used as reference, through the assumption of rank preservation by income. Stratification may be an option to ensure rank preservation.

4. With repeated cross-section data, rescaling datasets of unequal sizes may lead to matching individuals across strata. To avoid this, rescaling needs to be done within strata; while assessing the sensitivity of the methodology to the rescaling process.

Each issue, and proposed approach, is presented in further detail below.

3.2.1 Disregarding interactions across dimensions

Since the multidimensional headcount H often depends on a large number of dimensions, the ASN decomposition is much better suited than the Barros one to decompose it, as it abstracts from calculating the effects of pairwise calculations, which can be quite large. Moreover, as noted by Azevedo et al. (2013b), due to the path dependence of the Barros decomposition, not all pairwise combinations would be calculated, only those of the variables that are consecutive in the path chosen. Using the ASN decomposition comes, however, at the cost of assuming that the interaction across dimensions is small.

Disregarding the interaction across dimensions may not be reasonable when dimensions are highly related to each other. For example, two dimensions may be based on education variables and may be very likely to vary together. It is then necessary to group the dimensions into broad categories such that the interaction between different categories is small.

We therefore apply the methodology by grouping dimensions into broad categories, and validating our categories such that the interaction between them is low. This is done without loss of generality, by simply partitioning the dimensions into disjoint categories. Formally, let us assume that we partition the dimensions into two categories: $\{1, 2, \ldots, d_1\}$ and $\{d_1, d_1 + 1 \ldots, D\}$. We can then redefine the deprivation scores by the dimension of equation 13 as deprivation scores by category:

$$\bar{x}_1 \equiv \sum_{d=1}^{d_B} x_d$$
$$\bar{x}_2 \equiv \sum_{d=d_{A+1}}^{D} x_d \tag{22}$$

And carry out the decomposition over $\bar{x}_1$ and $\bar{x}_2$.

The fact that the ASN methodology may be generalized to categories of dimensions, should by no means be interpreted as stating that building categories is costless. As Azevedo et al. (2012b) note, the ASN methodology is sensitive to aggregating the dimensions into categories, and results may vary depending on the aggregation.

In our empirical application of section 4 below, we propose the use of descriptive statistics in order to assess whether interactions are indeed small across categories.

3.2.2 Interactions remaining constant over time

The ASN methodology also assumes that the interactions among variables remain constant over time, as these are not changed when building the counterfactuals. Therefore, in addition to being small, our empirical application requires that the interaction between categories is constant over time. This can be achieved, as before, by a proper definition of categories. This requirement is therefore assessed in our empirical application, as will be shown below.

3.2.3 Income as reference variable in repeated cross-section data

Another difficulty is associated with the inability to use discrete random variables as reference variables in repeated cross-section data. Given that the achievement variables y_d are indicator variables, the deprivation scores by dimension x_d have discrete distributions. This implies that their distribution functions are not invertible, so that the Barros decomposition may not be calculated. The issue does not come up when using panel data as, in such case, the counterfactuals are simply built by tracking the same individuals. The ranking functions R_x are also stepwise, non-invertible functions, so that the ASN methodology cannot be applied, in principle, if the reference variables are chosen from the deprivation scores by dimension.

However, as discussed in the previous section, the reference variable could be any continuous variable. Following Azevedo et al. (2012b), the ASN methodology can be applied, assuming rank preservation on income and thus using income as a reference variable. As noted, if it is assumed that rank preservation holds only after stratification, then the rank function of income needs to be invertible within strata. Narrowly defined strata may not satisfy this assumption, so we define strata broadly enough to have enough income variation within strata. Some strata may be empty in one of the datasets. By definition, these strata do not satisfy rank preservation, so they should be excluded from all calculations.

3.2.4 Rescaling of datasets

When faced with unequal sizes in the different cross-sectional data available across time, the ASN methodology suggests rescaling the ranking of the larger dataset to that of the smaller dataset. Similarly ranked observations are then matched to build the counterfactuals, but if the rankings are rescaled across strata, individuals from different strata may be matched. To avoid this, rankings need to be rescaled within strata. If possible, the sensitivity of the method to the rescaling should be assessed, for example, by recalculating the decomposition with subsamples of the original data.

This section has outlined the practical issues of applying the ASN methodology to the multidimensional headcount H. We now turn to two different empirical applications to illustrate the performance of the methodology, which highlight the importance of the issues described.

4 Application to repeated cross section data: The case of Colombia

We apply the proposed methodology in the context of the decline in multidimensional poverty observed in Colombia between 2008 and 2012. The Colombian case is interesting for a couple of reasons. Being a middle-income developing country, a large share of Colombians is still deprived in the dimensions considered for cross country calculations of MPI measures. On the other hand, the pace of decline has varied over the years. Monetary and multidimensional poverty declined sharply from 2003 to 2008; while over the last five years, the decline has been less sizable though steady.

4.1 Description of the data, the MPI and trends

We use data from the Colombian Quality of Life Survey for the years 2008, 2010 and 2013. This is the survey that the National Administrative Department of Statistics (*Departamento Administrativo Nacional de Estadística*, DANE) uses for multidimensional poverty calculations. The survey included around 50,000 households in 2008; while the sample size increased by about 5 percent and 38 percent in the next two rounds, respectively. We were able to replicate the official poverty measures published by DANE closely. For a detailed summary, see Angulo (2010).

The multidimensional poverty index for Colombia includes 15 dimensions grouped into five broad categories: education, childhood and youth, labor, health and standard of living[7]. Each of the categories has a weight of 0.2, which is distributed evenly

[7]This index is broader that the calculations of Alkire and Santos (2010) for Colombia. Their selec-

across the dimensions within each category. Table 1 shows all the dimensions of the index. Many dimensions are household-based: if the household is deprived in any of the dimensions, all household members are considered deprived. The cross-dimensional cut-off k is 1/3; that is households are considered multi-dimensionally poor if the weighted sum of deprivation scores is larger than 1/3. For example, a household deprived in all the dimensions within two categories receives a score of 0.4, and is considered poor. We use the categories to group the 15 dimensions, as described in section 3.2.1.

Figure 1, panel A, shows the evolution of monetary and multidimensional poverty measures, along with the measures described in section 2.1, for the three years considered. Monetary and multidimensional poverty headcounts have been declining over time, almost at the same pace. By 2012, around 27 percent of the Colombian population was estimated to be multi-dimensionally poor, while 32.7 percent is considered poor by the monetary measure. Although the headcount ratio has fallen, the average deprivation share among the poor has remained almost constant. This implies that the adjusted headcount ratio, which adjusts for the intensity of poverty, has not declined as quickly as the headcount.

Panel B decomposes the decrease in the adjusted headcount ratio into the contributions of changes in the multidimensional headcount –due to changes in the number of poor– and changes in the average deprivation share among the poor –due to changes in the intensity of poverty–. For both years, changes in intensity account for less than a quarter of the change in the adjusted headcount ratio, while most of the decrease is attributed to changes in the headcount. This is consistent with Apablaza et al. (2010) and Apablaza and Yalonetzky (2013), who find this pattern for different countries and datasets. The large contribution of the change headcount to the overall change in the adjusted headcount ratio, highlights the importance of analyzing its determinants separately.

Table 2 shows the evolution of censored headcounts by dimension H_d, that is, the headcounts of those deprived in at least one dimension within each category, and the average number of deprivations in each category. This table shows the percentage of the population deprived in an specific dimension within those that are multidimensional poor, as opposed to uncensored headcounts, which are shown in table A.1. Since deprivations are calculated at the household level, the numbers in table 2 are

tion criteria are detailed in Angulo (2010).

higher that national aggregates at the individual level.

Almost all poor individuals have low educational attainment and are not employed in the formal sector. The average poor individual is deprived in both education dimensions, but usually only in one dimension of standard of living. The average number of deprivations within each category declines slightly over time for the poor in most categories. The sharpest decline in the percentage of people with at least one deprivation occurs in the health category, where this headcount falls by about 10 percentage points.

In Table 3, we examine the percentage evolution of the censored headcounts by dimension over time. The sharpest declines occur in deprivation in access to childcare services and on the number of households living in homes built with low-quality materials. Although most headcounts declined between 2008 and 2012, many of them underwent a sharp decline from 2008 to 2010, followed by a rebound in the next two year; for example, lack of access to health services loses more than half of its initial decrease. Furthermore, long-term unemployment headcounts do not decrease but rather experience a steep increase over the period. The trends in the uncensored headcounts, shown in table A.1, are similar to the censored ones, but the U-shaped patterns in some dimensions are not as stark since the number of multidimensionally poor declines over time. Together, tables 2, 3 and A.1 show that the largest movers are in the categories of health, standard of living, and childhood and youth, although we cannot conclude that these are the biggest contributors to the change in the multidimensional headcount.

Table 4 examines the contribution of each category to the intensity of poverty. It decomposes the average deprivation share as in equation 5, grouping dimensions over categories. The largest contributors to the intensity of poverty are the education and labor categories. The contributions of each category are stable over time, with the exception of health, for which the contribution is declining.

4.2 Assessing the plausibility and performance of the methodology

Before applying the ASN methodology to the Colombian case, it is necessary to address the issues described in section 3.2, concerning repeated cross-section data. To address the first two issues—limited and stable interaction between categories—we need to confirm that in the categorization of dimensions defined so far, dimensions are not strongly associated across categories, and that associations remain constant over time. To do this, we calculate the Kendall rank correlation coefficients of the deprivation scores across categories for the three years, as illustrated in table 5. We choose Kendall correlation coefficients since we are concerned about changes in the ordinal association between the deprivation scores and the reference variable, which turn into changes in the association between deprivation score across categories. The results are encouraging: the correlation coefficients across categories are small and remain stable over time.

The third issue—the inability to use discrete random variables as reference variables with repeated cross-sections—requires significant attention. We need to address whether the continuous variable, income, is a proper reference variable for the methodology; if rank preservation is reasonable at the sample level; and, if not, we need to stratify the sample in order to make rank preservation more plausible. The observational data is uninformative in these matters, as it is not longitudinal data. A simulation exercise is carried out to address this issue, instead.

We build a panel dataset on the basis of the 2010 observations, adding an additional simulated period of data. For this added period, we simulate changes in the deprivations of individuals, in order to replicate the actual changes that took place in the uncensored headcount ratios by dimension between 2010 and 2012. We also change household income to reflect the income growth over the period. To do this, we calculate the variation in mean income within income deciles between the two years from the observational data, and add a normal random variable centered in this mean change to each individual within the income decile. The variance of this increase is chosen to match the change in within-decile variances that occurred between both years. We change deprivations in all the categories, initially one category at a time, and then all at once. We then assess different choices in the ASN methodology. We address the performance of these choices by comparing the results of the decomposition to those that would have been obtained using the panel dataset. Some of our choices are made with the purpose of highlighting the potential pitfalls of applying the method without examining the issues. All calculations are made using the software in Azevedo et al. (2012a).

Table 6 presents the results of the decompositions when deprivations in all categories are changed: this is the scenario closer to the observational data. For each method, the value on the left corresponds to the share that each category contributes to the change in the multidimensional headcount. The value on the right is the absolute difference between each column and the panel method. We describe each of the columns below.

- The "Panel" method corresponds to the decomposition using the panel structure of the simulated data. The third issue discussed does not arise here, so this column is provided as reference.
- The "No Panel" method applies the decomposition ignoring the panel structure, and using the deprivation share as the ranking variable. We know that, despite the fact that there are many dimensions in the index, the distribution of this score is discrete, and so its ranking function is not invertible. Because of this, it is not surprising that this method performs poorly.
- The "Strata" method is equivalent to the "No Panel" method, but it stratifies the sample into groups defined by categorical demographic variables. The categories are defined in Table A.2. This method shows that the ranking functions of the deprivation score are also non invertible within strata, so the method performs at least as poorly as the "No-Panel" one.
- The "Income Within" method is our preferred one. It uses income as a reference variable within strata, by stratifying the sample, sorting it within strata by income, and building the ranking functions based on income-within-strata. A potential pitfall of this is that income may not have enough variation within strata, and the ranking function may become non invertible. This may happen if strata are too narrow. However, this does not appear to be the case here, since the method performs about as well as an artificial method that builds the ranking variable by strata (across) and by incomes.

Overall, we find that the last method outperforms all remaining choices in pretty much all cases, as expected from the discussion in section 3.2. We also compute simulations changing categories one at a time. The full results are reported in Table A.3, and summarized in table 7. We calculate the difference of each method with the panel one, as in Table 6. Then, we average these discrepancies across categories. The result is then scaled by the total change in the headcount to arrive at "average discrepancies as a share of the total change in the headcount". The results confirm what we learned from the previous, more realistic simulation that changed all categories. The "Income Within" method outperforms the other ones, except for the

"Strata" method, which performs rather well. This is, however, unrealistic: since the changes in the deprivation score in this case are small, the ranks of the deprivation score are similar across years, which results in individuals being matched almost as if the dataset had a panel structure.

So far, our simulations have not addressed the fourth issue—rescaling datasets of unequal sizes within strata. We do so by repeating the exercises of Tables 6 and 7, reducing the sample size in period 1, by taking a (stratified on demographics) random sample of this data and repeating the exercises. The results of this exercise are presented in Table 8. The "Panel" method is no longer applicable here, so we compute the discrepancy against the panel method with equal sample sizes of the previous exercises. As in the previous case, the "No Panel" method performs poorly. The "Strata" method performs well when the changes are isolated into one category, but does poorly if changes occur in all categories.Our preferred method, the "Income Within" one, outperforms the others with sample reduction, although, as it is only natural, performs worse when the sample in the first period is smaller compared to the second period.

Having examined the potential issues with the methodology, we have found the method that best addresses them, outperforming the other ones. In the next section, we apply our preferred "Income Within" method to the case of the multidimensional poverty decline in Colombia.

4.3 Results of the decomposition for the Colombian case

We present the results of applying our preferred "Income Within" method to the case of the multidimensional poverty decline in Colombia in Table 9. These results present several highlights that would be absent in a more standard analysis focusing solely on the evolution of censored or uncensored headcounts.

The largest contributors to the decrease in the Colombian multidimensional headcount ratio are the 'education' and 'health' categories. Together, they account for approximately five percentage points out of a 7.5 percent reduction between 2008 and 2012; that is, more than 60 percent of the decline. Their contribution is similar between the 2008-2010 period and the 2010-2012 one. The next contributor, 'childhood and youth', is responsible for about one percentage point of the decline. The 'labor' category does not contribute much: this result could be expected from the analysis of the censored and uncensored headcounts, which do not present large reductions in Tables 3 and A.1. It is also intuitive that labor does not contribute much, given the sample period analyzed, as Colombia experienced an economic slowdown due to the global financial crisis over these years. Nevertheless, from those same tables, it would not have been intuitive to conclude that education was a large driver, instead, more weight would have been attributed (erroneously) to standard of living. The childhood and youth category is the only one responsible for a larger reduction of poverty in 2010-2012 than over the previous two years.

5 Conclusions

This paper analyzes the problem of decomposing changes in the multidimensional headcount ratio into the contributions from dimensions, or categories of dimensions. We examine the potential use of decompositions based on counterfactual simulations to break up changes in the multidimensional headcount; outlining potential issues with the methodology.

We propose and examine different options to address the caveats of the methodology, identifying a method to address these issues that performs well in simulations. The paper presents the application of this method to decompose the recent decline of poverty in Colombia, finding that health and education are the largest contributors to the decline.

Our proposed decomposition provides a useful way to estimate the extent in which each category contributes to the change in the headcount in the absence of panel data, without tracking which individuals cross the multidimensional poverty cut-off and which dimensions changed for each of those individuals. This methodology can be a useful complement to the analysis of multidimensional poverty that focuses on a wide range of indicators, such as those suggested by Ferreira and Lugo (2013). The exploration of further tools to decompose multidimensional poverty measurements based on non-scalar indexes, such as multidimensional distributions, appears as a fruitful avenue for future research.

References

ALKIRE, SABINA and FOSTER, JAMES (2011). "Counting and multidimensional poverty measurement". *Journal of Public Economics*, **95(7-8)**, pp. 476–487.

ALKIRE, SABINA and SANTOS, MARIA EMMA (2010). "Acute Multidimensional Poverty: A New Index for Developing Countries". *Working Paper 038*, OPHI.

ANGULO, ROBERTO (2010). "Índice multidimensional de pobreza para Colombia". In: *Seminario Internacional "Colombia en las nuevas tendencias de medición de la pobreza y la igualdad de oportunidades"*, .

APABLAZA, M.; OCAMPO, J.P. and G., YALONETZKY (2010). "Decomposing Changes in Multidimensional Poverty in 10 Countries". Mimeo, Oxford Poverty and Human Development Initiative.

APABLAZA, M. and YALONETZKY, G. (2013). "Decomposing Multidimensional Poverty Dynamics". *Working Paper 101*, Young Lives.

AZEVEDO, JOAO PEDRO; INCHAUSTE, GABRIELA; OLIVIERI, SERGIO; SAAVEDRA, JAIME and WINKLER, HERNAN (2013a). "Is Labor Income Responsible for Poverty Reduction? A Decomposition Approach". *Policy Research Working Paper 6414*, The World Bank..

AZEVEDO, JOAO PEDRO; INCHAUSTE, GABRIELA and SANFELICE, VIVIANE (2013b). "Decomposing the Recent Inequality Decline in Latin America". *Policy Research Working Paper 6715*, The World Bank.

AZEVEDO, JOAO PEDRO; NGUYEN, MINH CONG and SANFELICE, VIVIANE (2012a). "ADECOMP: Stata module to estimate Shapley Decomposition by Components of a Welfare Measure". Statistical Software Components, Boston College Department of Economics.

AZEVEDO, JOAO PEDRO; SANFELICE, VIVIANE and NGUYEN, MINH CONG (2012b). "Shapley Decomposition by Components of a Welfare Aggregate". Mimeo, he World Bank.

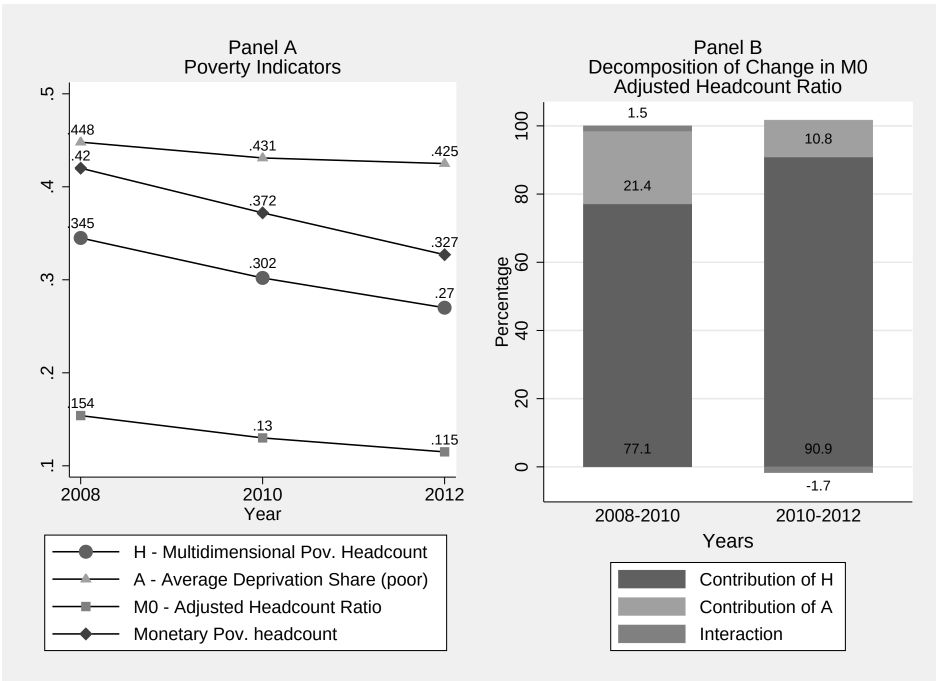
BARROS, RICARDO; DE CARVALHO, MIRELA; FRANCO, SAMUEL and MENDOÇA, ROSANNE (2006). "Uma Análise das Principais Causas da Queda Recente na Desigualdade de Renda Brasileira". *Revista Econômica*, **8(1)**, pp. 117–147.

FERREIRA, FRANCISCO H. G. and LUGO, MARIA ANA (2013). "Multidimensional Poverty Analysis: Looking for a Middle Ground". *World Bank Research Observer*, **28(2)**, pp. 220–235.

SHORROCKS, ANTHONY F. (2013). "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value". *The Journal of Economic Inequality*, **11(1)**, pp. 99–126.

Figures and tables

Figure 1: Trends in monetary and multidimensional poverty measures



Source: DANE. Author's calculations

Table 1: Categories and dimensions of the Colombian multidimensional poverty index.

Category	Dimension	Deprived if
Education	Educational achievement	Any person older than 15 years has less than 9 years of schooling.
	Literacy	Any person older than 15 years and illiterate.
Childhood and youth	School attendance	Any child 6 to 16 years old does not attend school.
	Children behind grade	Any child 7 to 17 years old is behind the normal grade for his age.
	Access to child care services	Any child 0 to 5 years old doesn't have access to health, nutrition or education.
	Child labour	Any child 12 to 17 years old works.
Employment	Long term unemployment	Any economically active member has been unemployed for 12 months or more.
	Formal employment	Any employed household members is not affiliated to a pension fund.
Health	Health insurance	Any person older than 5 years does not have health insurance.
	Health services	Any person who fell sick or ill in the last 30 days did not look for specialized services.
Standard of living	Water system	Urban: Household not connected to public water system. Rural: Household obtains water used for cooking from wells, rainwater, spring source, water tanks, water carriers or other sources.
	Sewage	Urban: Household not connected to public sewer system. Rural: Household uses a toilet without a sewer connection, a latrine or simply does not have a sewage system.
	Floors	Households has dirt floors.
	Walls	Rural: The household's exterior walls are made of vegetable, zinc, cloth, cardboard or waste materials or if no exterior walls exist. Urban: Walls made of rural materials or untreated wood, boards or planks.
	Overcrowding	Urban: There are 3 people or more per room. Rural: More than 3 people per room.

Source: Angulo (2010). Deprivations are measured at the household level: all members of the household are considered deprived if one of the members is deprived in a dimension.

Table 2: Multidimensional poverty measures, censored headcount ratios by dimension and average deprivation shares by category

Indicator	Weights	2008	2010	2012
H: Multidimensional headcount ratio (%)		34.47	30.38	26.97
A : Average deprivation share among the poor		0.448	0.432	0.425
M_0: Adjusted headcount ratio		0.154	0.131	0.115
Education	.2			
Educational Achievement (%)	0.1	96.30	94.26	94.24
Literacy (%)	0.1	43.37	44.76	44.56
At least 1 component (%)		96.36	94.78	94.96
Average deprivation share in category		0.698	0.695	0.694
Childhood and youth	.2			
School Attendance (%)	0.05	20.06	20.23	18.48
Children behind grade (%)	0.05	71.27	72.51	71.69
Access to child care services (%)	0.05	29.38	27.91	21.38
Child labour (%)	0.05	17.67	16.42	14.58
At least 1 component (%)		81.90	82.12	79.00
Average deprivation share in category		0.346	0.343	0.315
Labour	.2			
Long term unemployment (%)	0.1	10.12	10.61	11.14
Formal employment (%)	0.1	99.03	99.25	99.18
At least 1 component (%)		99.12	99.31	99.19
Average deprivation share in category		0.546	0.549	0.552
Health	.2			
Health insurance (%)	0.1	53.25	47.78	44.96
Health services (%)	0.1	22.99	16.97	19.81
At least 1 component (%)		63.51	57.39	55.93
Average deprivation share in category		0.381	0.324	0.324
Standard of living	.2			
Water system (%)	0.04	30.27	27.56	30.26
Sewage (%)	0.04	32.01	29.21	29.64
Floors (%)	0.04	23.44	20.75	19.46
Walls (%)	0.04	7.75	7.62	5.76
Overcrowding (%)	0.04	40.04	38.38	35.29
At least 1 component (%)		68.65	66.00	66.70
Average deprivation share in category		0.267	0.247	0.241

Source: Author's calculations. Headcounts are censored, i.e. calculated over the poor. See table A.1 for the uncensored headcount levels, calculated over the whole sample. "At least 1 component" denotes censored headcount ratios of individuals deprived in at least one dimension within the category. "Average deprivation share in category" is the average over poor individuals of the number of deprivations divided by the number of dimensions in the category. See table 3 for the changes of censored headcounts over time.

Table 3: Changes of censored headcount ratios over time

Indicator	2008-2010	2010-2012	2008-2012
Education			
Educational Achievement (% Change)	-2.11	-0.02	-2.13
Literacy (% Change)	3.20	-0.46	2.73
Childhood and youth			
School Attendance (% Change)	0.87	-8.67	-7.88
Children behind grade (% Change)	1.74	-1.12	0.60
Access to child care services (% Change)	-4.99	-23.40	-27.23
Child labour (% Change)	-7.07	-11.19	-17.47
Labour			
Long term unemployment (% Change)	4.79	5.03	10.07
Formal employment (% Change)	0.22	-0.07	0.15
Health			
Health insurance (% Change)	-10.27	-5.91	-15.57
Health services (% Change)	-26.20	16.75	-13.83
Standard of living			
Water system (% Change)	-8.96	9.81	-0.03
Sewage (% Change)	-8.75	1.48	-7.40
Floors (% Change)	-11.49	-6.22	-16.99
Walls (% Change)	-1.69	-24.46	-25.74
Overcrowding (% Change)	-4.15	-8.06	-11.88

Source: Author′s calculations. See table 2 for the censored headcount levels.

Table 4: Multidimensional measures and average deprivation shares by category

Indicator	**2008**	**2010**	**2012**
H: Multidimensional headcount ratio (%)	34.47	30.38	26.97
A: Average deprivation share among the poor	0.448	0.432	0.425
M_0 :Adjusted headcount ratio	0.154	0.131	0.115
Education	0.140	0.139	0.139
Childhood and youth	0.069	0.069	0.063
Labour	0.109	0.110	0.110
Health	0.076	0.065	0.065
Standard of living	0.053	0.049	0.048

Source: Author's calculations. Average deprivation shares by category are calculated by calculating the average deprivation shares by dimension as in equation 5, then adding over categories.

Table 5: Rank correlations between deprivation scores across categories

	Education	**Childhood and youth**	**Labour**	**Health**	**Standard of living**
			2008		
Education	1.000	0.209	0.201	0.121	0.288
Childhood and youth	0.209	1.000	0.063	0.150	0.240
Labour	0.201	0.063	1.000	0.112	0.100
Health	0.121	0.150	0.112	1.000	0.081
Standard of living	0.288	0.240	0.100	0.081	1.000
			2010		
Education	1.000	0.177	0.217	0.061	0.241
Childhood and youth	0.177	1.000	0.024	0.112	0.189
Labour	0.217	0.024	1.000	0.104	0.074
Health	0.061	0.112	0.104	1.000	0.055
Standard of living	0.241	0.189	0.074	0.055	1.000
			2012		
Education	1.000	0.170	0.205	0.066	0.248
Childhood and youth	0.170	1.000	0.021	0.108	0.190
Labour	0.205	0.021	1.000	0.095	0.082
Health	0.066	0.108	0.095	1.000	0.040
Standard of living	0.248	0.190	0.082	0.040	1.000

Source: Author's calculations.

Table 6: Results of the decomposition in a simulated panel with changes in all of the categories

	Panel	No Panel		Strata		Income within	
	Value	Value	$\|\Delta\|$ vs Panel	Value	$\|\Delta\|$ vs Panel	Value	$\|\Delta\|$ vs Panel
Education	-0.618	-0.571	0.046	-0.513	0.104	-0.597	0.020
Chilhood and Youth	-0.142	-0.485	0.343	-0.450	0.308	-0.082	0.060
Labour	-0.034	0.023	0.057	-0.085	0.050	-0.066	0.032
Health	-0.812	-0.738	0.074	-0.977	0.165	-0.943	0.131
Standard of living	-0.117	0.049	0.166	0.302	0.419	-0.035	0.082

Source: Author's calculations. See the text for a description of the table.

Table 7: Performance of simulations when changing categories one by one: average discrepancy as share of change in headcount

	Panel	No Panel	Strata	Income within
Education	0.000	0.749	0.886	0.346
Chilhood and Youth	0.000	3.070	1.833	1.507
Labour	0.000	11.083	0.499	4.245
Health	0.000	0.816	0.707	0.408
Standard of living	0.000	3.089	1.430	1.965
All	0.000	0.398	0.607	0.188

Source: Author's calculations. See the text for a description of the table.

Table 8: Performance of simulations with samples of unequal sizes

	No Panel	**Strata**	**Income within**
85 % subsample in first period			
Education	1.035	0.715	0.526
Chilhood and Youth	3.778	1.167	3.646
Labour	8.179	3.440	8.511
Health	0.566	0.656	0.588
Standard of living	4.528	1.623	3.983
All	0.475	0.588	0.244
70 % subsample in first period			
Education	1.102	1.031	0.796
Chilhood and Youth	4.641	3.899	4.602
Labour	12.3	10.52	11.75
Health	1.068	0.823	0.773
Standard of living	5.648	4.613	5.318
All	0.544	0.704	0.379

Source: Author's calculations. See the text for a description of the table.

Table 9: Decomposition of the multidimensional headcount ratio in Colombia 2008-2012

Indicator	**2008-2010**	**2010-2012**	**2008-2012**
Education			
Change due to category (Percentage points)	-1.124	-1.034	-2.465
Percentage contribution of category (%)	27.51	30.24	32.85
Childhood and youth			
Change due to category (Percentage points)	-0.484	-0.663	-1.152
Percentage contribution of category (%)	11.85	19.40	15.35
Labour			
Change due to category (Percentage points)	-0.334	-0.100	-0.405
Percentage contribution of category (%)	8.18	2.93	5.39
Health			
Change due to category (Percentage points)	-1.418	-1.007	-2.372
Percentage contribution of category (%)	34.72	29.46	31.61
Standard of living			
Change due to category (Percentage points)	-0.725	-0.614	-1.111
Percentage contribution of category (%)	17.74	17.97	14.80
Total	-4.09	-3.42	-7.50

Appendix

Table A.1: Evolution of uncensored headcount ratios by dimension

Indicator	**2008** (%)	**2010** (%)	**2012** (%)	**2008-2010** (Δ %)	**2010-2012** (Δ %)	**2008-2012** (Δ %)
Education						
Educational Achievement	62.70	58.67	56.24	-6.43	-4.13	-10.30
Literacy	17.42	16.04	14.19	-7.92	-11.52	-18.53
Childhood and youth						
School Attendance	8.04	7.09	6.23	-11.80	-12.08	-22.45
Children behind grade	45.10	47.00	44.64	4.21	-5.03	-1.03
Access to child care services	17.06	16.28	12.92	-4.57	-20.67	-24.29
Child labour	7.78	6.60	5.39	-15.13	-18.35	-30.71
Labour						
Long term unemployment	6.73	6.77	6.63	0.53	-2.07	-1.55
Formal employment	83.72	83.55	82.90	-0.20	-0.79	-0.99
Health						
Health insurance	27.72	24.34	21.13	-12.18	-13.20	-23.77
Health services	10.79	7.81	7.89	-27.61	0.98	-26.90
Standard of living						
Water system	14.56	12.84	13.22	-11.86	3.00	-9.22
Sewage	15.67	13.33	13.51	-14.93	1.31	-13.81
Floors	9.22	7.68	6.78	-16.76	-11.72	-26.52
Walls	3.43	3.22	2.37	-6.11	-26.50	-30.99
Overcrowding	22.17	20.76	18.36	-6.38	-11.52	-17.16

Source: Author's calculations. Headcounts are uncensored, i.e. calculated over the whole sample. See table 2 for the censored headcount levels, calculated among the poor.

Table A.2: Demographic variables used for stratification

Variable	**Description**
Decile	Income Decile
Household head gender	1 if the household head is male, 0 otherwise
Education	1 if the household head has no education, 2 if he has primary, 3 secondary and 4 tertiary.
Household size	1 if household has more than 4 people, 0 otherwise.
Kids	1 if household has kids, 0 otherwise.
Urban	1 if household is located in a urban area, 0 otherwise

Table A.3: Results of the decomposition in a simulated panel with changes in each one of the categories

	Panel	**No Panel**	**Strata**	**Income within**
Simulated change in education				
Education	-0.678	-0.551	-0.574	-0.722
Chilhood and Youth	0.000	-0.240	-0.123	0.050
Labour	0.000	-0.013	-0.006	0.091
Health	0.000	0.079	-0.096	-0.098
Standard of living	0.000	0.047	0.120	0.001
Simulated change in chilhood and youth				
Education	0.000	0.005	0.016	-0.009
Chilhood and Youth	-0.115	-0.291	-0.155	-0.108
Labour	0.000	0.067	0.011	0.098
Health	0.000	0.095	-0.027	-0.099
Standard of living	0.000	0.009	0.039	0.003
Simulated change in labour				
Education	0.000	0.149	0.005	-0.019
Chilhood and Youth	0.000	-0.144	-0.001	0.037
Labour	-0.047	0.035	-0.053	0.052
Health	0.000	-0.114	-0.005	-0.110
Standard of living	0.000	0.027	0.007	-0.007
Simulated change in health				
Education	0.000	0.206	0.094	0.016
Chilhood and Youth	0.000	-0.315	-0.137	0.045
Labour	0.000	-0.019	0.019	0.102
Health	-0.818	-0.724	-0.905	-0.981
Standard of living	0.000	0.035	0.112	0.000
Simulated change in standard of living				
Education	0.000	0.111	0.005	-0.012
Chilhood and Youth	0.000	-0.135	-0.025	0.031
Labour	0.000	0.029	-0.005	0.095
Health	0.000	-0.047	-0.020	-0.110
Standard of living	-0.087	-0.046	-0.043	-0.092