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

Jorge Eduardo Pérez Pérez*
Carlos Rodríguez Castelán†
José Daniel Trujillo‡
Daniel Valderrama§

This draft: March 16, 2018¶

*Corresponding author. Department of Economics, Brown University.
Email: jorge_perez@brown.edu
†World Bank - Poverty Global Practice. Email: crodriguezc@worldbank.org
‡Department of Physics, Universidad de Los Andes. Email: jd.trujillo242@uniandes.edu.co
§Department of Economics. Georgetown University. Email: dv348@georgetown.edu

¶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. We thank Viviane Sanfelice for outstanding research assistance, and to Claudia Quintero and Nicolás Fuertes for their assistance with ELCA data. 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. This paper was written while Trujillo was working at DANE and Valderrama was at the World Bank - Poverty Global Practice. Pérez Pérez gratefully acknowledges financial support from Fulbright-Colciencias.

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.

Keywords Multidimensional poverty index, decomposition

JEL Codes I32

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 (Ravallion, 2011). 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[1]. 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. Decompo-

[1]Decomposition approaches are deeply rooted in monetary poverty research. Examples include Ravallion and Huppi (1991), Datt and Ravallion (1992), Bourguignon et al. (2005), Kolenikov and Shorrocks (2005) and Azevedo et al. (2013a)

sitions 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. Even though the adjusted headcount ratio has desirable theoretical properties that the headcount ratio lacks (Alkire and Foster, 2011), the headcount ratio itself has gained prominence in recent poverty analyses. 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 recent literature on decomposition of multidimensional poverty indices. Apablaza and Yalonetzky (2013) propose approaches to decompose multidimensional poverty measures statically or dynamically. They decompose the two components of the adjusted headcount ratio, by decomposing the headcount ratio across groups and the average deprivation share across dimensions. 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. Our approach allows use to decompose changes in the headcount ratio by dimensions.

Roche (2013) implements a Shapley value approach (Shorrocks, 2013) to decompose the adjusted headcount ratio and its components. Garcia-Diaz and Prudencio (2017) adopts this strategy to decompose a multidimensional chronic poverty index. As Apablaza and Yalonetzky (2013), they only decompose the average deprivation share across dimensions, leaving the question of how different dimensions contribute to changes in the incidence of poverty. Our simulation-based approach tackles this issue. 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 could be obtained.

The paper is divided in two parts. In the first part, we describe the problem of decomposing multidimensional poverty measures and outline our methodology. In 2, we summarize the technical aspects of the multidimensional poverty measurement, and set up the analytical framework for the rest of the paper. We outline the technical aspects of the counterfactual simulation methodology in section 3.

In the second part, we apply our methodology on datasets of Colombian households. We use both a panel and repeated cross section datasets, and highlight the differences and pitfalls of using the decomposition methodology in each case. We use the panel dataset, –where our decomposition is exact– to illustrate the methodology and draw lessons for applying it in more general cases. With the repeated cross section dataset, we replicate recent measurements of Colombian multidimensional poverty and highlight the drivers behind its decline, contributing to the literature on poverty in Colombia.

We start by describing our data and calculating multidimensional poverty measures for Colombia in section 4. In section 5, we apply our methodology to a panel dataset and highlight potential issues that may arise when applying the decomposition methodology to the multidimensional headcount and when using it in repeated cross sectional data. We conduct different simulations to identify a method based on stratification that performs well in several scenarios. Then, in section 6, we apply the decomposition analysis to repeated cross sectional data 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. We conclude in section 7.

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 $\sum_{d=1}^{D} w_d = 1$[2]. 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.*

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

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

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

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.

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 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

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

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

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:

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

Combining this equation 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

[4]Notice that Φ need not be invertible.

sorted by the value of random variable

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

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 $R_{x_{d'}}$ 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 $C(x_{d'}, x_{d''})$. 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:

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

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:

[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.

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

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 x_1 were distributed as in period 1, but x_2 and the association remained as observed in period 2. 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

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

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. This occurs because the counterfactual $\hat{x}_1$ may have a different ranking that the actual variable in the second period, x_d^2. To restore the constant association, they build a different counterfactual $\tilde{x}_d$ that reorders $\hat{x}_1$ from equation (20) using the ranking of

x_d^2:

$$\tilde{x}_{id}^1 = x_{i^*d}^1$$
$$i^* = C\left(\hat{x}_d, x_d\left(i\right)\right) = R_{\hat{x}_d}^{-1}\left(R_{x_d^2}\left(i\right)\right) \quad (21)$$

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 $\mathbf{y}$ vector and calculate the average of all the contributions obtained[6].

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 . We now move to applying the decomposition to multidimensional poverty in Colombia. Along the process, we highlight several issues that can arise when the methodology is used in applied work. We describe the data and calculate the multidimensional poverty measures –whose change we would like to decompose– in the next section.

4 Multidimensional poverty in Colombia

This section describes the datasets used in the analysis, calculates multidimensional poverty measures and briefly analyses their trends. We use both a panel dataset and a repeated cross section dataset of Colombian households.

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.

The panel dataset provides a useful testing ground for calculating decompositions in an environment where we can track individuals over time. The cross sec-

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

tional dataset allows us to replicate official poverty measures closely, and analyse the drivers behind the decline in Colombian multidimensional poverty.

4.1 Data description

We use two datasets of Colombian Households. The first is the Colombian Longitudinal Survey (Encuesta Longitudinal Colombiana - ELCA), a panel dataset of 10,800 households. About two thirds of the households are rural, and urban households are selected to be low-income. The survey is useful to provide a picture of household's vulnerability to economic shocks. We use the 2010 and 2013 waves, and restrict our analysis to households that can be tracked on both waves.[7]. This leaves a total of 8686 households.

The second dataset is the Colombian Quality of Life Survey for the years 2008, 2010 and 2012 (Encuesta de Calidad de Vida - ECV). 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. Unlike the ELCA, the ECV does not oversample rural households and is representative at the national level.

4.2 Multidimensional poverty indices

We calculate multidimensional poverty indices using both datasets to set the stage for our decomposition methods. We start with the ECV dataset, where we replicate the official poverty measures published by DANE closely. We follow Angulo (2010) to construct the index.

The multidimensional poverty index for Colombia includes 15 dimensions grouped into five broad categories: education, childhood and youth, labor, health and standard of living[8]. Each of the categories has a weight of 0.2, which is distributed evenly 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

[7]For households that split, we keep the primary household

[8]This index is broader that the calculations of Alkire and Santos (2010) for Colombia. Their selection criteria are detailed in Angulo (2010).

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.

For ELCA, we calculate a restricted index using 12 out of the 15 dimensions, due to lack of data on child labour, child care and long-term unemployment. We also restrict our childhood related variables to children under 9 years of age, because older children do not fill the childhood related module. When dimensions are missing, we redistribute the weights to remaining dimensions in the same category.

Panels A1 and B1 of figure 1 show the evolution of monetary and multidimensional poverty measures, along with the measures described in section 2.1, for the years considered. Panel A1 shows official poverty calculations from DANE, which we replicate using ECV data. 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 B1 shows the evolution of these measures as calculated from ELCA data. The focus of ELCA on rural households, the restriction to households that can be linked across waves, and the restrictions placed on the index, result on a higher multidimensional headcount ratio of 47.9 % by 2013. The average deprivation share among the poor also remains relatively constant in this dataset.

Panels A2 and B2 decompose 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 all time periods and across datasets, 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 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 from 2008 to 2012.

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.

After analyzing the overall patterns in these poverty measures, we apply our methodology to find which dimensions drive the overall reduction in the multidimensional poverty headcount in Colombia. We start by applying the decomposition to the ELCA dataset, where we can track the same individuals over time, in the next section.

5 Decomposing H in panel data

In this section we decompose the reduction in poverty from 2010 to 2013 using the ELCA dataset. We take advantage of the panel structure of the data to illustrate the decomposition and highlight issues when applying the decomposition to the multidimensional poverty headcount H.

5.1 Decomposition by dimension categories

We apply the decomposition outlined in equation (19) by calculating counterfactuals tracking the same individuals over time. The counterfactual variables $\tilde{x}^1$, are just the values of the variables in the first period, x_i^1.

Since the multidimensional headcount H often depends on a large number of dimensions, the Azevedo et al. (2013b) decomposition (ASN hereafter) 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 effects of changes in the interaction across dimensions is small, such that the association between different dimensions remains constant over time. Thus, the ASN methodology disregards changes in the poverty headcount that may arise from changes in the bivariate association of dimensions, even though their univariate distributions are held constant in the counterfactuals. Disregarding changes in the interaction across dimensions may not be reasonable when their relationship may change over time.

We therefore apply the methodology by grouping dimensions into broad categories, and validating our categories such that the interaction between deprivation scores across categories is low and stable over time. 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.

A proper definition of categories should ensure that the interaction across categories is small and constant over time. We use descriptive statistics to assess this. We calculate the Kendall rank correlation coefficients of the deprivation scores across categories for the two years in the sample, as illustrated in table 5. We choose Kendall correlation coefficients since we are concerned about changes in the ordinal association between the deprivation scores across categories. The results are encouraging: the correlation coefficients across categories are small and stable.

5.2 Results in panel data

Table 6 shows the results of the decomposition exercise for the decline in the multidimensional headcount ratio between 2010 and 2013 in the ELCA dataset. All calculations are done with the software provided in Azevedo et al. (2012a). For these households, improvements in standards of living account for more than half of the reduction in the headcount ratio. The dimensions in the education and childhood and youth categories account for most of the remainder of the decrease. The dimensions in the health category actually contribute to a 13 % increase in the headcount ratio, which is offset by the improvements in other categories.

We emphasize the usefulness of the decomposition by comparing it to tracking censored headcounts and incidence separately in tables 2 and 3. Although these tables show that dimensions in the standard of living category play a large role in the reduction of the headcount, because of the large incidence of deprivation in this category andb its large declines, it would be harder to conclude that the education and childhoood and youth categories play such a large role, or that the worsening of access to health services increased the overall headcount ratio. The decomposition provides a succinct measure of these contributions and complements these analyses.

The ELCA decomposition is limited to the 2010-2013 period, and focuses only on a subset of households, which are mostly rural. To provide a wider picture of the drivers of the decline in multidimensional poverty for Colombia as a whole, we turn to decomposing the decline in poverty in the ECV data in the next section.

6 Decomposing H in cross sectional data

In this section, we apply the proposed methodology to decompose the recent decline in the multidimensional headcount ratio in Colombia, replicating official measures using the ECV dataset. We start by highlighting issues that arise when applying the decomposition on repeated cross sectional data. We propose solutions to these issues and assess their performance using simulations with the panel dataset from the previous section. At the end of the section, we present the results of the decomposition using a stratification based method that performs well in simulations. We find that health and education are the major drivers of the decline in the multidimensional headcount from 2008 to 2012.

6.1 Reference variables to build counterfactuals

The first issue faced when applying the decomposition to repeated cross sections is the construction of counterfactual distributions. As shown in the previous section, this is straightforward in panel data. The counterfactuals are obtained by setting the counterfactual deprivation scores by category, $\tilde{x}_i^1$, to their value in the first period for the same household, x_i^1.

In repeated cross-section data, however, the counterfactuals can only be obtained using equation (20) if F is strictly monotone in the bivariate case, or if R is strictly monotone in the multivariate case. This is rarely the case for our application. 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.

To solve this, we follow Azevedo et al. (2012b) and use a continuous variable to build the rank functions R. This amounts to replacing $\hat{x}_d$ with a continuous reference variable in equation (21). The counterfactual for an individual with a x_d^2 value is an individual who has the same rank on the continuous variable corresponding to x_d^2 in period 1. This hinges on assuming rank preservation to track individuals with the same rank across two periods.

In practice, any continuous reference variable could substitute $\hat{x}_d$ if rank preservation is plausible and its ranking function is invertible. Following Azevedo et al. (2012b), we start by assuming rank preservation on income and thus use household income per capita as a reference variable. We call this the "Income" method. We also use expenditure to include households that underreport income, and call this the "Expenditure" method.

In some cases it may not be plausible for rank preservation to hold unconditionally, but it may hold *conditional* on a set of 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). 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. We stratify on several demographics including education, income deciles, gender and education of the household head. Details are in table A.2. Some strata may be empty in one of the datasets. By definition, these strata do not satisfy rank preservation, so they are excluded from all calculations. We call the methods based on stratification the "Income Strata" and "Expenditure Strata" methods.

To test if using these reference variables is useful to decompose changes in multidimensional headcount ratios, we carry out a simulation exercise. We treat the ELCA panel dataset as a repeated cross sectional dataset. We then calculated the decomposition using different reference variables. We analyze the performance of these choices by comparing the results of the decomposition to those that would have been obtained using the panel dataset.

Table 7 shows the results of this exercise. We measure the discrepancy as the average difference in the shares of the change in the headcount attributed to each category between each method and the panel method. The method based on income and stratification outperforms the other methods using this measure. Using this method, the values of the contributions attributed to each category are closer to those obtained by tracking individuals over time than using any of the other methods.

6.2 Rescaling of datasets

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.

We test how this method performs in environments when the sample size is reduced. We do so by repeating the exercises of table 7, reducing the sample size in 2013, by taking a random sample of this data and recalculating the decomposition. The results of this exercise are presented in Table 8. We show the difference in the shares of the change in headcount attributed to each category comparing each method to the panel method in the full sample. The "Income Strata" method also outperforms others in this environment of variable sample sizes. As it is only natural, the method 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 a method that best addresses them, outperforming the other ones. We now apply our preferred "Income Strata" method to the multidimensional poverty decline in Colombia using the ECV dataset.

6.3 Results in repeated cross sectional data

We present the results of applying our preferred "Income Strata" 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.

The differences across datasets between the contributions of each category to the decrease after 2010 are worth noting. The decomposition in the ELCA dataset shows improvements in standard of living as a large driver of the decline, while in ECV this accounts to at most 15% of the 2010-2012 decline. Another striking difference is the role of the health dimensions. The decomposition in ELCA attributes increases in poverty to worsening in health dimensions, while the decomposition in ECV points towards health as a large contributor to the decline in the headcount ratio. We attribute this to differences between the rural households in ELCA and the average household in ECV.

7 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 even 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.

BOURGUIGNON, FRANÇOIS; FERREIRA, FRANCISCO and LUSTIG, NORA (2005). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. The World Bank.

DATT, GAURAV and RAVALLION, MARTIN (1992). "Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s". *Journal of Development Economics*, **38(2)**, pp. 275–295.

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.

GARCIA-DIAZ, ROCIO and PRUDENCIO, DANIEL (2017). "A SHAPLEY DECOMPOSITION OF MULTIDIMENSIONAL CHRONIC POVERTY IN ARGENTINA". *Bulletin of Economic Research*, **69(1)**, pp. 23–41.

KOLENIKOV, STANISLAV and SHORROCKS, ANTHONY (2005). "A Decomposition Analysis of Regional Poverty in Russia". *Review of Development Economics*, **9(1)**, pp. 25–46.

RAVALLION, MARTIN (2011). "On multidimensional indices of poverty". *The Journal of Economic Inequality*, **9(2)**, pp. 235–248.

RAVALLION, MARTIN and HUPPI, MONIKA (1991). "Measuring Changes in Poverty: A Methodological Case Study of Indonesia during an Adjustment Period". *World Bank Economic Review*, **5(1)**, pp. 57–82.

ROCHE, JOSÉ MANUEL (2013). "Monitoring Progress in Child Poverty Reduction: Methodological Insights and Illustration to the Case Study of Bangladesh". *Social Indicators Research*, **112(2)**, pp. 363–390.

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

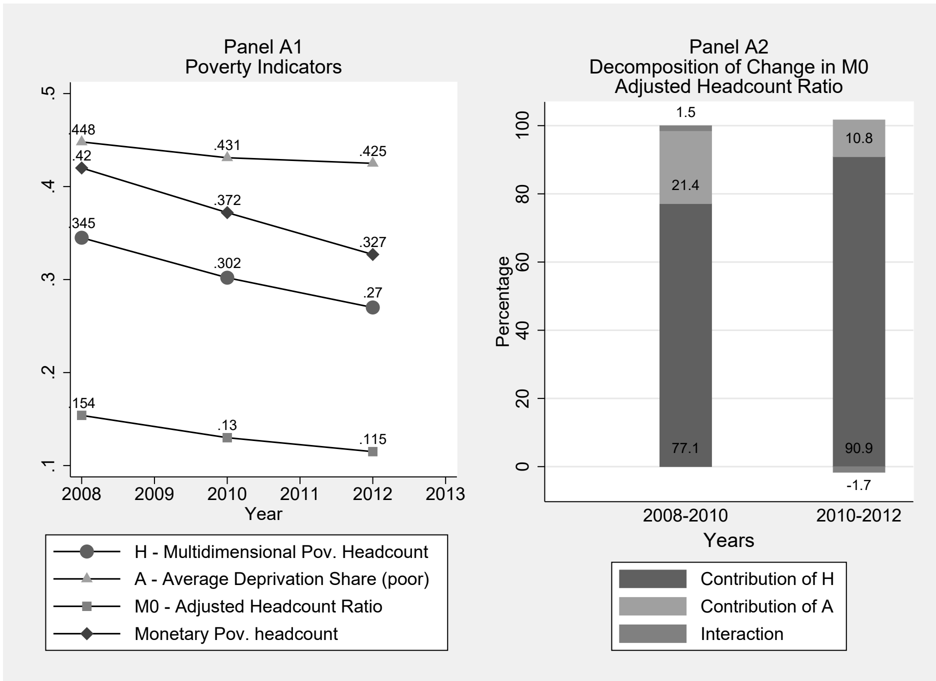
Table 1: Categories and dimensions of the Colombian multidimensional poverty index.

Category	Dimension	Deprived if (ECV)	Deprived if (ELCA)
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.	Any child 6 to 9 years old does not attend school.
	Children behind grade	Any child 7 to 17 years old is behind the normal grade for his age.	Any child 7 to 9 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.	Not measurable
	Child labour	Any child 12 to 17 years old works.	Not measurable
Employment	Long term unemployment	Any economically active member has been unemployed for 12 months or more.	Not measurable in 2010
	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.	

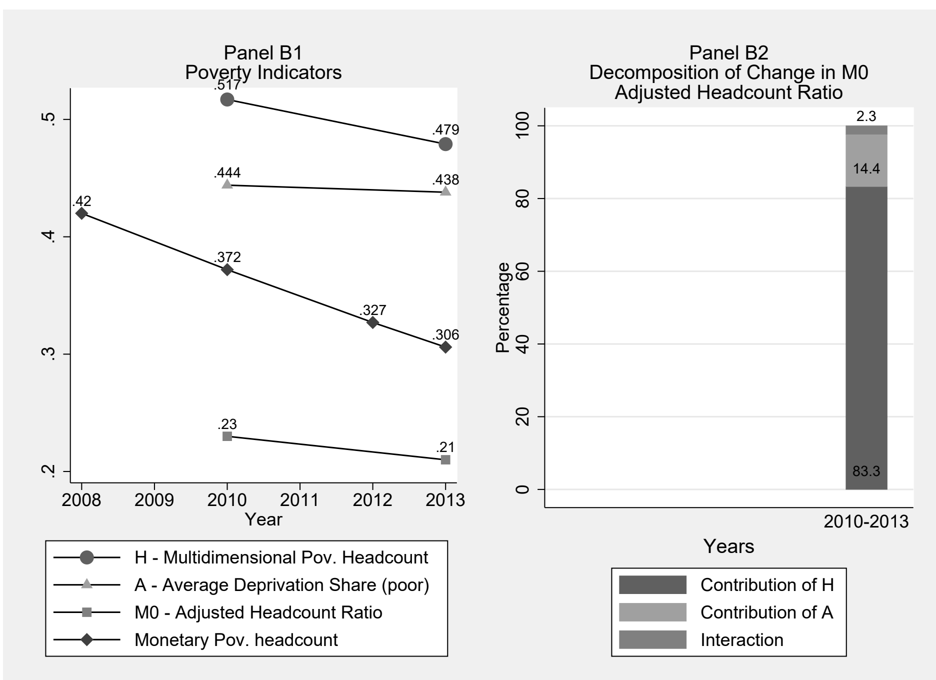
Source: Angulo (2010). Deprivations are measured at the household level: all members of the household are considered deprived if one of the members is deprived in a dimension.

Figure 1: Trends in monetary and multidimensional poverty measures

(a) Panel A: ECV



(b) Panel B: ELCA



Source: DANE, and author's calculations based on ECV and ELCA data.

Table 2: Multidimensional poverty measures, censored headcount ratios by dimension and average deprivation shares by category. 2008-2013.

Indicator	Weight	2008 ECV	2010 ECV	2010 ELCA	2012 ECV	2013 ELCA
H: Multidimensional headcount ratio (%)		34.5	30.2	51.7	27.0	47.9
A: Average deprivation share among the poor		0.45	0.43	0.44	0.43	0.44
M_0: Adjusted headcount ratio		0.15	0.13	0.23	0.11	0.21
Education	.2					
Educational Achievement (%)	0.1	96.3	94.3	95.3	94.2	95.6
Literacy (%)	0.1	43.4	45.0	30.6	44.6	30.6
At least 1 component (%)		96.4	94.8	95.3	95.0	95.6
Average deprivation share in category		0.70	0.70	0.63	0.69	0.63
Childhood and youth	.2					
School Attendance (%)	0.05	20.1	20.2	1.1	18.5	1.1
Children behind grade (%)	0.05	71.3	72.7	19.8	71.7	16.0
Access to child care services (%)	0.05	29.4	28.0	.	21.4	.
Child labour (%)	0.05	17.7	16.5	.	14.6	.
At least 1 component (%)		81.9	82.3	20.1	79.0	16.3
Average deprivation share in category		0.35	0.34	0.10	0.32	0.09
Labour	.2					
Long term unemployment (%)	0.1	10.1	10.6	.	11.1	.
Formal employment (%)	0.1	99.0	99.3	99.7	99.2	99.8
At least 1 component (%)		99.1	99.4	99.7	99.2	99.8
Average deprivation share in category		0.55	0.55	1.00	0.55	1.00
Health	.2					
Health insurance (%)	0.1	53.3	47.9	20.0	45.0	11.8
Health services (%)	0.1	23.0	17.0	18.0	19.8	30.0
At least 1 component (%)		63.5	57.5	33.3	55.9	38.2
Average deprivation share in category		0.38	0.32	0.19	0.32	0.21
Standard of living	.2					
Water system (%)	0.04	30.3	27.5	35.4	30.3	38.5
Sewage (%)	0.04	32.0	29.0	31.9	29.6	27.7
Floors (%)	0.04	23.4	20.6	36.5	19.5	29.2
Walls (%)	0.04	7.8	7.6	6.3	5.8	4.8
Overcrowding (%)	0.04	40.0	36.9	41.1	35.3	34.8
At least 1 component (%)		68.6	64.5	81.0	66.7	77.5
Average deprivation share in category		0.27	0.24	0.30	0.24	0.27

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. denotes censored headcount ratios of individuals deprived in at least one dimension within the category. is the average over poor individuals of the number of deprivations divided by the number of dimensions in the category. Childhood and youth related dimensions are calculated for chidren 6-9 years old in ELCA, and for children up to 17 years old in ECV. When dimensions are missing in a category, weights are redistributed among the remaining dimensions. See table 3 for the changes of censored headcounts over time.

Table 3: Changes of censored headcount ratios over time

Indicator	2008-2010 ECV	2010-2012 ECV	2008-2012 ECV	2010-2013 ELCA
Education				
Educational Achievement (% Change)	-2.06	-0.07	-2.13	0.23
Literacy (% Change)	3.86	-1.09	2.73	-0.14
Childhood and youth				
School Attendance (% Change)	0.87	-8.67	-7.88	-4.54
Children behind grade (% Change)	1.99	-1.36	0.60	-18.90
Access to child care services (% Change)	-4.83	-23.54	-27.23	.
Child labour (% Change)	-6.71	-11.53	-17.47	.
Labour				
Long term unemployment (% Change)	5.15	4.67	10.07	.
Formal employment (% Change)	0.29	-0.14	0.15	0.12
Health				
Health insurance (% Change)	-10.14	-6.05	-15.57	-40.92
Health services (% Change)	-26.19	16.74	-13.83	66.59
Standard of living				
Water system (% Change)	-9.27	10.18	-0.03	8.76
Sewage (% Change)	-9.51	2.33	-7.40	-13.23
Floors (% Change)	-12.07	-5.60	-16.99	-19.96
Walls (% Change)	-2.50	-23.83	-25.74	-23.90
Overcrowding (% Change)	-7.78	-4.44	-11.88	-15.22

Source: Author's calculations. Childhood and youth related dimensions are calculated for chidren 6-9 years old in ELCA, and for children up to 17 years old in ECV.

See table 2 for the censored headcount levels.

Table 4: Multidimensional measures and average deprivation shares by category

Indicator	**2008** ECV	**2010** ECV	**2010** ELCA	**2012** ECV	**2013** ELCA
H: Multidimensional headcount ratio (%)	34.5	30.2	51.7	27.0	47.9
A: Average deprivation share among the poor	0.45	0.43	0.44	0.43	0.44
M_0 :Adjusted headcount ratio	0.15	0.13	0.23	0.11	0.21
Education	14.0	13.9	12.6	13.9	12.6
Childhood and youth	6.9	6.9	2.1	6.3	1.7
Labour	10.9	11.0	19.9	11.0	20.0
Health	7.6	6.5	3.8	6.5	4.2
Standard of living	5.3	4.9	6.1	4.8	5.4

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. ELCA.

			2010		
	Education	Childhood and youth	Labour	Health	Standard of living
Education	1.00	-0.01	-0.04	-0.18	-0.03
Childhood and youth	-0.01	1.00	-0.08	-0.03	-0.02
Labour	-0.04	-0.08	1.00	-0.05	-0.00
Health	-0.18	-0.03	-0.05	1.00	-0.15
Standard of living	-0.03	-0.02	-0.00	-0.15	1.00
			2013		
	Education	Childhood and youth	Labour	Health	Standard of living
Education	1.00	-0.05	-0.03	-0.20	0.00
Childhood and youth	-0.05	1.00	-0.08	-0.02	0.03
Labour	-0.03	-0.08	1.00	-0.06	-0.01
Health	-0.20	-0.02	-0.06	1.00	-0.25
Standard of living	0.00	0.03	-0.01	-0.25	1.00

Source: Author's calculations. Numbers are τ_b rank correlation coefficients between deprivation scores by category.

Table 6: Decomposition of the change in the multidimensional headcount ratio. Colombia, ELCA 2010-2013

Category	**2010-2013**
Education	
Change due to category (Percentage points)	-0.95
Percentage contribution of category (%)	25.15
Childhood and youth	
Change due to category (Percentage points)	-0.79
Percentage contribution of category (%)	20.86
Labour	
Change due to category (Percentage points)	-0.45
Percentage contribution of category (%)	11.89
Health	
Change due to category (Percentage points)	0.51
Percentage contribution of category (%)	-13.58
Standard of living	
Change due to category (Percentage points)	-2.10
Percentage contribution of category (%)	55.68
Total	-3.77

Source: Author's calculations. Columns show the share of the change in the multidimensional headcount attributed to each row category.

Table 7: Results of the decomposition treating panel as repeated cross section. Colombia: ELCA 2010-2013

	Panel	**Income**		**Income Strata**		**Expenditure**		**Expenditure Strata**	
	Value	Value	abs. diff.	Value	abs. diff.	Value	abs. diff.	Value	abs. diff.
Education	-0.95	-1.54	0.59	-0.92	0.02	-1.32	0.37	-0.84	0.11
Childhood and Youth	-0.79	-0.86	0.07	-1.08	0.29	-0.86	0.07	-0.97	0.18
Labour	-0.45	-0.68	0.23	-0.50	0.05	-0.33	0.12	-0.46	0.01
Health	0.51	0.88	0.37	0.57	0.05	0.49	0.03	0.13	0.38
Standard of living	-2.10	-1.57	0.53	-1.83	0.27	-1.75	0.35	-1.64	0.46
Average			0.36		0.14		0.19		0.23
As % of total change			0.09		0.04		0.05		0.06

Source: Author's calculations. Column "Panel" show the results of the decomposition tracks the same households over time. The remaining columns show the result of the decomposition using different methods. "Value" columns show the contribution of each category in each method and "abs diff" columns show the deviation from the contribution of each category in the panel method. Method "Income" uses income per capita by household as the reference variable. Method "Expenditure" uses expenditure per capita by household as the reference variable. Methods "Income Strata" and "Expenditure Strata" use income and expenditure, respectively, stratifying first on variables for: urban household, gender of household head, members of household, number of kids, education of the household head and decile of income or expenditure.

Table 8: Results of the decomposition treating panel as repeated cross section. Unequal sample sizes. Colombia: ELCA 2010-2013

	Income	**Income Strata**	**Expenditure**	**Expenditure Strata**
		15 % sample reduction		
Education	0.13	0.06	0.04	0.02
Childhood and Youth	0.00	0.01	0.03	0.06
Labour	0.02	0.00	0.05	0.02
Health	0.02	0.07	0.07	0.13
Standard of living	0.17	0.03	0.19	0.15
Average	0.07	0.04	0.07	0.08
		30 % sample reduction		
Education	0.24	0.04	0.25	0.13
Childhood and Youth	0.05	0.10	0.00	0.07
Labour	0.09	0.02	0.01	0.04
Health	0.04	0.08	0.02	0.12
Standard of living	0.16	0.08	0.24	0.29
Average	0.12	0.06	0.11	0.13

Source: Author's calculations. Values are differences in shares attributed to each category compared to the panel method. The columns show these differences in shares across different decomposition methods. Method "Income" uses income per capita by household as the reference variable. Method "Expenditure" uses expenditure per capita by household as the reference variable. Methods "Income Strata" and "Expenditure Strata" use income and expenditure, respectively, stratifying first on variables for: urban household, gender of household head, members of household, number of kids, education of the household head and decile of income or expenditure.

Table 9: Decomposition of the change in the multidimensional headcount ratio. Colombia. ECV 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

Source: Author's calculations. Columns show the share of the change in the multidimensional headcount attributed to each row category using the "Income Strata" method of section 6.1.

Appendix

Table A.1: Uncensored headcount rations by dimension.

Indicator	**2008** ECV	**2010** ECV	**2010** ELCA	**2012** ECV	**2013** ELCA
Education					
Educational Achievement(%)	62.7	58.7	71.8	56.2	67.5
Literacy(%)	17.4	16.0	16.3	14.2	15.0
Childhood and youth					
School Attendance(%)	8.0	7.1	0.7	6.2	0.5
Children behind grade(%)	45.1	47.0	11.4	44.6	8.9
Access to child care services(%)	17.1	16.3	.	12.9	.
Child labour(%)	7.8	6.6	.	5.4	.
Labour					
Long term unemployment(%)	6.7	6.8	.	6.6	.
Formal employment(%)	83.7	83.6	81.1	82.9	81.9
Health					
Health insurance(%)	27.7	24.3	12.1	21.1	7.1
Health services(%)	10.8	7.8	11.0	7.9	17.6
Standard of living					
Water system(%)	14.6	12.6	20.1	13.2	20.8
Sewage(%)	15.7	13.0	17.9	13.5	14.8
Floors(%)	9.2	7.5	19.7	6.8	14.6
Walls(%)	3.4	3.1	3.5	2.4	2.6
Overcrowding(%)	22.2	19.2	25.2	18.4	20.4

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