Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford

OPHI *Working Paper No. 135*

Endogenous Weights and Multidimensional Poverty: A Cautionary Tale

Indranil Dutta,[*] Ricardo Nogales,[**] and Gaston Yalonetzky [***]

January 2021

Abstract
Composite measures such as multidimensional poverty indices depend crucially on the weights assigned to the different dimensions and their indicators. A recent strand of the literature uses endogenous weights, determined by the data at hand, to compute poverty scores. Notwithstanding their merits, we demonstrate both analytically and empirically how a broad class of endogenous weights violates key properties of multidimensional poverty indices such as monotonicity and subgroup consistency. Without these properties, anti-poverty policy targeting and assessments are bound to be seriously compromised. Using real-life data from Ecuador and Uganda, we show that these violations are widespread. Hence, one should be extremely careful when using endogenous weights in measuring poverty. Our results naturally extend to other welfare measures based on binary indicators, such as the widely studied asset indices.

Keywords: Multidimensional poverty, Endogenous weights, Measurement externalities, Principal components analysis (PCA), Frequency weights, Dual-cutoff counting approach

JEL Classification: I32

[*]University of Manchester Email: I.Dutta@manchester.ac.uk.
[**]OPHI, University of Oxford, Email: ricardo.nogales@qeh.ox.ac.uk.
[***]Corresponding author. University of Leeds, Email: G.Yalonetzky@leeds.ac.uk.

This study has been prepared within the OPHI theme on multidimensional measurement.

Acknowledgments
We would like to thank James Foster as well as seminar participans at the Workshop: Recent developments in distributional analysis (Leeds, April 2019) and the OPHI Seminar Series (Oxford, November 2020) for very helpful comments and suggestions.

Citation: Dutta, I., Nogales, R., and Yalonetzky, G. (2021): 'Endogenous weights and multidimensional poverty: A cautionary tale' OPHI Working Paper 135, University of Oxford.

The Oxford Poverty and Human Development Initiative (OPHI) is a research centre within the Oxford Department of International Development, Queen Elizabeth House, at the University of Oxford. Led by Sabina Alkire, OPHI aspires to build and advance a more systematic methodological and economic framework for reducing multidimensional poverty, grounded in people's experiences and values.

Oxford Poverty & Human Development Initiative (OPHI)
Oxford Department of International Development
Queen Elizabeth House (QEH), University of Oxford
3 Mansfield Road, Oxford OX1 3TB, UK

Tel. +44 (0)1865 271915 Fax +44 (0)1865 281801
ophi@qeh.ox.ac.uk http://www.ophi.org.uk

ISSN 2040-8188 ISBN 978-1-912291-26-7

1 Introduction

It has become increasingly common to understand deprivation from a multidimensional perspective. Practitioners undertaking such multidimensional assessments must make several nontrivial methodological decisions, including which dimensions and indicators of deprivation to consider among the several possible and how to combine them into one single composite index of multidimensional poverty. In combining these dimensions and their indicators into a composite index, a natural question to ask is how much weight should be assigned to each of them. This paper examines, both analytically and empirically, the implications of using endogenous (i.e. data-driven) weights on a set of desirable properties for multidimensional poverty indices (see Bourguignon and Chakravarty, 2003; Alkire and Foster, 2011) and demonstrates their failure to satisfy these key properties under endogenous weights.

One of the key methodological challenges in computing composite measures like multidimensional poverty indices is how to weight the observed dimensions of deprivation and their associated indicators. A common approach is to use exogenous weights, which are independent of the dataset and reflect the value judgements of the society, the analyst or the policymaker. In contrast, we focus on a growing body of literature which relies on the alternative of endogenous weights, which are determined by the dataset, as a way to reflect the importance of the different indicators in the composite measure of deprivation (OECD, 2008; Decanq and Lugo, 2013). We consider a general array of endogenous weights, including those based on statistical methods and factorial techniques, such as principal component analysis (henceforth PCA) applied to appropriately defined correlation matrices, and frequency-based weights, which are functions of the frequency of deprivation among the population in the different indicators according to some normative judgements.[1]

The applications of these endogenous weights are widespread. For instance, the Mexican anti-poverty programme 'Prospera' used PCA-based indices to select the localities where the program was implemented (Skoufias et al., 2001; Dávila Lárraga, 2016). Likewise, popular and well-established indices such as the Social Progress Index or the Human Needs Index are constructed using endogenous weights based on PCA.[2] Similarly, a Comprehensive Vul-

[1]Endogenous weights based on data-reduction factorial techniques such as PCA (see e.g. Njong and Ningaye, 2008; Asselin and Anh, 2008; Asselin, 2009; Alkire et al., 2015; Coromaldi and Drago, 2017; Wittenberg and Leibbrandt, 2017) use optimisation procedures applied to statistical concepts such as correlation or variance. Meanwhile, depending on normative judgements, frequency-based weights increase (or decrease) with the proportions of people deprived in a particular indicator.

[2]The Social Progress Index is calculated based on inverted indicators, and hence looks at deprivation. In justi-

nerability Monitoring Exercise recently established by the World Food Program assesses livelihood vulnerabilities of refugees in Turkey based on an index constructed using PCA (WFP, 2019).[3] Influential studies such as Asselin and Anh (2008) claimed that other factorial data-reduction techniques, such as Multiple Correspondence Analyses (henceforth MCA), are particularly well suited to evaluating multidimensional poverty composed of *categorical* indicators. This led to a surge in studies using MCA-based weights (e.g. Ki et al., 2005; Ezzrari and Verme, 2012; Noglo, 2017; Dhongde and Haveman, 2017). A strand of the literature even suggests endogenous weights as benchmarks for multidimensional poverty analyses, claiming that they can be regarded as 'superior' approaches to identify weighting structures for indicators. For instance, Pasha (2017) recently relied on MCA-based weights to criticise the Multidimensional Poverty Index (henceforth MPI) of the United Nations Development Programme (UNDP) and OPHI, concluding that 'equal weighting of the three dimensions cannot be statistically justified' (p. 268). In a similar critical spirit, using Confirmatory Factor Analyses – another factorial technique for data reduction, Nájera Catalán and Gordon (2019) qualify exogenous weights in multidimensional poverty measurement as statistically 'unreliable'. As another example, based on PCA/factor-analysis-like approaches, Heshmati et al. (2008) claim to put forth 'more sophisticated' methods to assess child poverty compared to UNICEF's approaches that are not based on such data-driven weights (UNICEF, 2005, 2007). Likewise, examples of frequency-driven endogenous weights are ubiquitous in the literature on multidimensional poverty measurement (see e.g. Deutsch and Silber, 2005; Njong and Ningaye, 2008; Aaberge and Brandolini, 2014; Whelan et al., 2014; Alkire et al., 2015; Cavapozzi et al., 2015; Rippin, 2016; Datt, 2019; Abdu and Delamonica, 2018).[4]

This paper demonstrates that combining a broad class of endogenous weights with a general class of multidimensional poverty indices based on the popular counting approach (Alkire and Foster, 2011) leads to the violation of two fundamental properties in poverty measurement: **monotonicity** and **subgroup consistency**.[5] Monotonicity states that if the poverty experience of an individual worsens in any indicator, then the overall poverty experience of

fying PCA for the Social Progress Index, Stern et al. (2018, p. 13) state that: 'It essentially assigns each indicator a weight, a method we select over equal weighting to ensure that indicators are meaningfully contributing to a component score, while accounting for similarities between them.'

[3]See https://www.wfp.org/publications/turkey-comprehensive-vulnerability-monitoring-exercise.

[4]Moreover, the broad class of endogenous weights considered in this paper also includes hybrid weights, where data-driven weights are used for some dimensions in combination with exogenous weights for other dimensions (see e.g. Dotter and Klasen, 2014).

[5]The Alkire and Foster (2011) counting approach, which relies on a minimum weighted number of deprived indicators as a threshold to identify the multidimensionally poor, is broadly the path followed by around seventy countries and organisations that measure multidimensional poverty, including the flagship UNDP–OPHI MPI, which is used to evaluate multidimensional poverty globally (see Alkire et al., 2015; MPPN, 2019).

the society to which this individual belongs should not improve (Tsui, 2002). Subgroup consistency requires that changes in overall poverty in a population should coherently reflect the changes in poverty happening at the smaller population subgroup level. For instance, if the population of a country is divided into two subgroups based on regions, say North and South, then if the poverty of the North increases while the poverty of the South remains unchanged, overall poverty in the country should not decrease under fulfillment of subgroup consistency. Moreover, subgroup consistency requires that overall poverty should not decrease whether the North is paired with the South, the East or the West, as long as poverty did not change in the latter three regions.

Failure of the poverty index to satisfy monotonicity implies that we may observe societal poverty falling even when the poverty of some individuals in that society has increased, without any countervailing decrease in any other individuals' poverty. Violation of monotonicity can lead to perverse policies whereby increasing individuals' deprivation in some indicators can be deemed beneficial since it will lead to an overall decrease in multidimensional poverty. Meanwhile, failure of subgroup consistency can lead to a situation where an increase in poverty in some regions or population subgroups, ceteris paribus, may decrease societal poverty, depending on which other regions or population subgroups with unchanged poverty levels are taken into account for the computation. This in turn can lead to policies that ignore increasing poverty in one region or one population subgroup because overall poverty has decreased. Without these key properties, it would be futile to use a poverty index to undertake any kind of comparative exercise, whether across time, regions or population groups, and thus any evaluation of anti-poverty policies would be ineffective (see Sen, 1976; Foster and Shorrocks, 1991).

In our context, endogenous weights generate a *measurement externality* since they depend on the distribution of deprivations across the indicators. Change in one person's deprivation status (e.g. because they are no longer deprived in some indicator) affects the deprivation scores of many other people through its impact on the weighting vector. Our appraisal of other people's poverty is thus altered, despite the absence of any objective change in their deprivation status. By contrast, this measurement externality is nonexistent if weights are set exogenously. This paper examines the implications of this measurement externality in multidimensional poverty measurement. We derive results explaining how, specifically, measurement externalities operate, including how they lead to the violation of monotonicity and subgroup consistency. Moreover, our results can easily be extended to encompass societal welfare measures dependent on binary variables, like the popular asset indices pioneered by

Filmer and Pritchett (2001) and measures of material deprivation in Europe (see e.g. Guio et al., 2016), thus demonstrating that these measures suffer from similar problems. In addition, we illustrate these violations using real-world data from two countries with structurally different patterns of multidimensional poverty: the 2013/14 Ecuador Living Conditions Survey and the 2015/16 Uganda National Panel Survey.

The rest of the paper is organised as follows. Section 2 introduces the notation and discusses the basic poverty measurement framework, including the important properties of monotonicity and subgroup decomposability. In section 3 we introduce a general class of endogenous weights. All we assume for our weights is that if the weight of one indicator is reduced, then the weight(s) of some other indicator(s) is (are) increased; that is, that the weights' values need to be relative to each other in some meaningful sense. Using real-world data from Ecuador and Uganda, section 4 presents an empirical illustration of violation of the properties of monotonicity and subgroup decomposability under commonly used endogenous weighting rules. Then, in order to explain the causes of such violations, section 5 shows how measurement externalities operate once a person's deprivation status in some indicator changes, followed by section 6, which provides the main theoretical results on measurement externality and violation of monotonicity and subgroup consistency properties under endogenous weights. The final section summarises the paper's main message with some concluding remarks.

2 Preliminaries: Multidimensional Poverty Measurement

Consider a deprivation matrix $\mathbf{X_{ND}}$, with each of the $N > 1$ rows representing an individual (or household) and each of the $D > 1$ columns representing a deprivation indicator. We denote any individual as n, where $n = 1,2,...,i,i',...,N$, and any indicator as d, where $d = 1,2,...,j,j',...,D$. Let $\rho_{nd} \in \{0,1\}$ denote the deprivation of person n in indicator d in the deprivation matrix $\mathbf{X_{ND}}$. For any individual n, poverty is determined by the deprivations faced by the individual across the different indicators, which are given by the deprivation vector $\rho_{\mathbf{n}}^{\mathbf{X_{ND}}} : \{\rho_{n1}, \rho_{n2}, ..., \rho_{nD}\}$. Further, let $\mathbf{X_{\bullet d}}$ be the column vector associated with indicator d of $\mathbf{X_{ND}}$, i.e. $\{\rho_{1d}, \rho_{2d}, ..., \rho_{Nd}\}$. Note that, for our purpose, we assume that individuals are either fully deprived in an indicator ($\rho_{nd} = 1$) or not at all ($\rho_{nd} = 0$).

Let each indicator of $\mathbf{X_{ND}}$ be weighted; the weight of indicator d is represented as $w_d^{X_{ND}}$. Then we have a weighting vector of strictly positive entries $\mathbf{w}^{\mathbf{X_{ND}}} = (w_1^{X_{ND}}, w_2^{X_{ND}}, ..., w_D^{\mathbf{X_{ND}}})$ such that $\sum_{d=1}^{D} w_d^{X_{ND}} = 1$. As alluded to before, weights can be determined either endogenously or exogenously. The values of exogenous weights can, for instance, remain constant across different deprivation matrices because by definition they are independent from the data, whereas endogenous weights take into account the distribution of deprivation in each indicator. Thus, under endogenous weights, two different deprivation matrices $\mathbf{X}_{\mathbf{ND}}$ and $\mathbf{X}'_{\mathbf{ND}}$ will have different weights for the indicators. Specifically, $\mathbf{w}^{\mathbf{X}_{\mathbf{ND}}} = (w_1^{X_{ND}}, w_2^{X_{ND}}, ..., w_D^{X_{ND}})$ and $\mathbf{w}^{\mathbf{X}'_{\mathbf{ND}}} = (w_1^{X'_{ND}}, w_2^{X'_{ND}}, ..., w_D^{X'_{ND}})$, where for some j and j', $w_j^{X_{ND}} \neq w_j^{X'_{ND}}$ and $w_{j'}^{X_{ND}} \neq w_{j'}^{X'_{ND}}$. We describe the weighting functions with precision in section 3.

2.1 Individual Poverty

In the counting approach proposed by Alkire and Foster (2011), individual poverty is measured through a two-step procedure: (i) identifying whether an individual is multidimensionally deprived based on the number of indicators they are deprived in, and if so, (ii) computing a weighted aggregate of their deprivation over all the indicators (in practice, Alkire and Foster (2011) propose a weighted count of deprivations). Thus, the resulting individual poverty function has two components: a poverty identification function, ψ, and a poverty severity function, s (Silber and Yalonetzky, 2013).

For an individual n, let the total count of deprivations be $t_n = \sum_{d=1}^{D} v_d \rho_{nd}$, where $v_1, v_2, ..., v_D$ are weights adding up to 1, from a vector of weights $\mathbf{v}$ potentially different from $\mathbf{w}^{\mathbf{X}_{\mathbf{ND}}}$ and potentially exogenous. Following Alkire and Foster (2011), a person is considered multidimensionally deprived if their total deprivation count is at least as high as an exogenously specified cutoff. That is, the identification function $\psi : [0,1] \rightarrow \{0,1\}$ compares t_n against a cutoff $0 < k \leq 1$ in order to identify the person as either poor or nonpoor from a multiple-deprivation perspective.[6] Thus,

$$\psi(t_n; k) = \mathbb{I}(t_n \geq k). \tag{1}$$

When $0 < k \leq \min\{v_1, v_2, ..., v_D\}$, the poverty identification function follows a **union approach**, whereby any person with at least one deprivation is deemed poor. On the other extreme, when $k = 1$, poverty identification follows an **intersection approach**, which regards as poor only those who are deprived in all indicators. Hence, under the union and intersection approaches, the weights do not really have a role in identifying whether an individual

[6] We suggest using exogenous weights at the identification stage in order to avoid violating the focus axiom (see Alkire and Foster (2011, p. 480)), which captures the idea that improvements in the wellbeing of the nonpoor should not change the level of societal poverty.

is multidimensionally deprived or not. Between both extremes, several other intermediate approaches exist in a counting framework, corresponding to the other values that k can take (Alkire and Foster, 2011). Within the literature advocating endogenous weights, a notable example of data-driven weights combined with an intermediate approach to poverty identification is provided by Pasha (2017), who essentially uses $\mathbf{v} = \mathbf{w}^{\mathbf{X}_{\mathrm{ND}}}$.[7] In our empirical analysis we use an union approach to identification which is adopted in practice by a swathe of the literature, especially studies using endogenous weights based on data-reduction techniques (e.g. Njong and Ningaye, 2008; Asselin and Anh, 2008; Asselin, 2009; Coromaldi and Drago, 2017).

The severity component, $s : [0,1] \rightarrow [0,1]$, measures the severity of the multiple-deprivation experience among poor people (Chakravarty and D'Ambrosio, 2006; Alkire and Foster, 2011; Silber and Yalonetzky, 2013), where the weighted deprivation score (or counting function) is

$$C_n(\mathbf{X}_{\mathrm{ND}}) = \begin{cases} \sum_{d=1}^{D} w_d \rho_{nd} & \text{if} \quad \psi(t_n;k) = 1 \\ 0 & \text{if} \quad \psi(t_n;k) = 0. \end{cases} \tag{2}$$

The severity component satisfies the following properties: $s(C_i) > s(C_j)$ whenever $C_i > C_j$, $s(0) = 0$ and $s(1) = 1$. We may also include the restriction that $s''(C_n) \geq 0$ (assuming differentiability of s). Thus, the severity function is monotonic in the weighted deprivation count of each individual, and it increases at a nondecreasing rate. Straightforward examples of $s(C_n)$ used in the literature include $s(C_n) = C_n$ (Alkire and Foster, 2011), $s(C_n) = e^{\alpha C_n} - 1$ with $\alpha > 0$ (Chakravarty and D'Ambrosio, 2006) and $s(C_n) = (C_n)^{\beta}$ with $\beta \geq 1$ (Datt, 2019).

Thus, for any deprivation matrix $\mathbf{X}_{\mathrm{ND}}$, the individual poverty function for an individual n, $p_n^{X_{ND}} : \{0,1\} \times [0,1] \longrightarrow [0,1]$, takes the form

$$p_n^{\mathbf{X}_{ND}}(t_n, C_n; k) = \psi(t_n; k) s(C_n). \tag{3}$$

It combines the identification and the severity components to yield a measure of overall deprivation at the individual level. In order to align our empirical illustration in section 4 to some of the most commonly used poverty functions in multidimensional poverty analyses (see e.g. Atkinson, 2019), we calculate individual poverty based on a linear severity function: $p_n = \psi(t_n; k) C_n$.[8]

[7] Hence, the analysis in Pasha (2017) violates the crucial focus axiom.
[8] Our results carry over onto the quadratic case. Available upon request.

2.2 Societal Poverty

The societal poverty function aggregates the individual poverty experiences measured by the individual poverty function in (3). Therefore, we can represent it as $P:[0,1]^N \longrightarrow [0,1]$. In its most general form it could be written as

$$P^{\mathbf{X}_{\mathbf{ND}}}(\mathbf{X}_{\mathbf{ND}};,k) = g(p_1^{\mathbf{X}_{ND}}, p_2^{\mathbf{X}_{ND}}, ..., p_N^{\mathbf{X}_{ND}}),$$

where $\partial P/\partial p_i \geq 0$ and $\partial P^2/\partial p_i \partial p_j = 0$ (assuming differentiability of g). These two assumptions imply that societal poverty P should be at the very least nondecreasing in its constituent parts, which in turn should be strongly separable (Blackorby et al., 1978). There are different methods of aggregation, each satisfying a set of properties. Throughout the rest of the paper, we follow a common additively subgroup-decomposable societal poverty function:

$$P(\mathbf{X}_{\mathbf{ND}};k) = \frac{1}{N}\sum_{n=1}^{N} p_n^{\mathbf{X}_{ND}}. \tag{4}$$

Societal poverty indices like P are expected to fulfill certain desirable properties. Chief among them is **monotonicity**, which requires societal poverty not to decrease if a poor individual suffers from an additional deprivation. For a formal definition, consider a deprivation matrix $\mathbf{X}_{\mathbf{ND}}$ where individual i is deemed poor, i.e. $\psi(t_i;k)=1$. Then let $\mathbf{X}'_{\mathbf{ND}}$ be obtained from $\mathbf{X}_{\mathbf{ND}}$ by a simple increase in deprivation in indicator j of individual i in $\mathbf{X}_{\mathbf{ND}}$; meaning that (i) $\rho_{ij}^{\mathbf{X}'_{ND}} = 1$, (ii) $\rho_{ij}^{\mathbf{X}_{ND}} = 0$ and (iii) $\forall (n,d) \neq (i,j)$, $\rho_{nd}^{\mathbf{X}'_{ND}} = \rho_{nd}^{\mathbf{X}_{ND}}$. Then the monotonicity axiom can be written as follows.

Axiom 1 *Monotonicity (M): Suppose $\mathbf{X}'_{\mathbf{ND}}$ is obtained from $\mathbf{X}_{\mathbf{ND}}$ by a simple increase in deprivation in indicator d of individual i. Then $\Delta P = P^{\mathbf{X}'_{\mathbf{ND}}} - P^{\mathbf{X}_{\mathbf{ND}}} \geq 0$.*

Another important property is **subgroup consistency** (Foster and Shorrocks, 1991; Alkire and Foster, 2011), which requires societal poverty to change (e.g. in a country across time) in the same direction of a change in the poverty levels of a subgroup (e.g. within a region across time), if the poverty levels of all other subgroups remain unchanged. For a formal definition, consider a subgroup-decomposable deprivation matrix $\mathbf{X}_{\mathbf{ND}}$ formed by vertical concatenation of two matrices $\mathbf{X}_{\mathbf{N_1D}}$ and $\mathbf{X}_{\mathbf{N_2D}}$ where $N = N_1 + N_2$. We represent it as $\mathbf{X}_{\mathbf{ND}} = (\mathbf{X}_{\mathbf{N_1D}} \,||\, \mathbf{X}_{\mathbf{N_2D}})$. Then the axiom of subgroup consistency can be stated as follows.

Axiom 2 *Subgroup Consistency (SC): Suppose $\mathbf{X}_{\mathbf{ND}} = (\mathbf{X}_{\mathbf{N_1D}} \,||\, \mathbf{X}_{\mathbf{N_2D}})$ and $\mathbf{Y}_{\mathbf{ND}} = (\mathbf{Y}_{\mathbf{N_1D}} \,||\, \mathbf{Y}_{\mathbf{N_2D}})$ are two subgroup-decomposable deprivation matrices. The societal poverty function P sat-*

isfies subgroup consistency if $[P(\mathbf{X_{N_1D}};\mathbf{w}^{\mathbf{X_{N_1D}}},k) > P(\mathbf{Y_{N_1D}};\mathbf{w}^{\mathbf{Y_{N_1D}}},k) \wedge P(\mathbf{X_{N_2D}};\mathbf{w}^{\mathbf{X_{N_2D}}},k) = P(\mathbf{Y_{N_2D}};\mathbf{w}^{\mathbf{Y_{N_2D}}},k)] \Rightarrow P(\mathbf{X_{ND}};\mathbf{w}^{\mathbf{X_{ND}}},k) > P(\mathbf{Y_{ND}};\mathbf{w}^{\mathbf{Y_{ND}}},k)$.

3 Examples of Endogenous Weights

For a general class of endogenous weights, consider a deprivation matrix $\mathbf{X_{ND}}$ with the weight of indicator d represented as

$$w_d = H_d(\mathbf{h}_1, ..., \mathbf{h}_d, ..., \mathbf{h}_D), \tag{5}$$

where $\mathbf{h}_j = h(\mathbf{X_{\bullet j}})$ represents a transformation of the column vector associated with indicator j of $\mathbf{X_{ND}}$. For PCA, $\mathbf{h}_j = \mathbf{X}^*_{\bullet\mathbf{j}}$ reflects a standardised column vector; for frequency-based weights we shall use $\mathbf{h}_j = \mathbf{X_{\bullet j}}$. Further, for all j, $\Delta\mathbf{h}_j$ implies $\Delta H_d(\mathbf{h}_1, ..., \mathbf{h}_d, ..., \mathbf{h}_D) \gtreqless 0$. That is, we allow for the possibility that changes in the vector of deprivation in indicator j may impact the weight of indicator d. Since these are weights over indicators, a natural constraint requires that the weights should add up to one, i.e.

$$\sum_{d=1}^{D} w_d = 1. \tag{6}$$

Together, equations (5) and (6) characterise a broad class of endogenous weights. Specific examples of frequency-based weights for indicator j are given by

$$w_j^F = \frac{f(\mathbf{X_{\bullet j}})}{\sum_{d=1}^{D} f(\mathbf{X_{\bullet d}})}, \tag{7}$$

where $f'(\mathbf{X_{\bullet j}}) \gtrless 0$ (assuming differentiability of f). One example of f with $f'(\mathbf{X_{\bullet j}}) > 0$ is $f(\mathbf{X_{\bullet j}}) = \sum_{n=1}^{N} \rho_{nj}\psi(t_n;k)/N$, which implies that as more people become deprived in an indicator, it becomes a more important indicator of multidimensional poverty, and hence it should carry a higher weight in the composite index. On the other hand, an example of f with $f'(\mathbf{X_{\bullet j}}) < 0$ is $f(\mathbf{X_{\bullet j}}) = -\ln(\sum_{n=1}^{N} \rho_{nj}\psi(t_n;k)/N)$ (see Deutsch and Silber, 2005, p. 150). Another possibility is $f(\mathbf{X_{\bullet j}}) = 1 - (\sum_{n=1}^{N} \rho_{nj}\psi(t_n;k)/N)$, which essentially captures the intuition that if deprivation in one particular indicator becomes endemic, it may no longer serve as a distinguishing factor and hence should be weighted less in the composite index.

Another common way of constructing composite indices with endogenous weights consists of identifying orthogonal linear combinations of the standardised column vectors of $\mathbf{X_{ND}}$, denoted by $\mathbf{X}^*_{\bullet\mathbf{d}}, \forall d = 1, 2, ..., D$, in such a way as to reproduce their variance and interlinkages as closely as possible. This logic underlies a range of factorial techniques for data reduction, including factor analysis, PCA and MCA (Asselin and Anh, 2008; Krishnakumar and Nagar, 2008).

In our setting, let Σ denote the variance-covariance matrix of $\{\mathbf{X}^*_{\bullet\mathbf{1}}, ..., \mathbf{X}^*_{\bullet\mathbf{D}}\}$. One way to account for the binary nature of the elements in these vectors is to define the off-diagonal elements of Σ as bivariate tetrachoric correlation coefficients (see e.g. Kolenikov and Angeles, 2009; Howe et al., 2012).[9] Let us also write the eigenvalues of Σ as $\lambda_1, ..., \lambda_D$ in descending order, and the corresponding $D \times 1$ eigenvectors as $v_1, ..., v_D$. Then the ℓth principal component, a_ℓ, is given by

$$a_\ell = \mathbf{X}^*_{\mathbf{ND}} v_\ell = \sum_{j=1}^{D} v_{\ell,j} \mathbf{X}^*_{\bullet\mathbf{j}}, \forall \ell = 1, 2, ..., D, \tag{8}$$

and its variance is $V(a_\ell) = \lambda_\ell$, $\forall \ell = 1, 2, ..., D$. Clearly, each principal component is a linear combination of the standardised deprivation indicators where the eigenvectors of Σ define their relative importance, i.e. their weights.

In practice, the first principal component is the most commonly used 'summary' indicator that can be derived from PCA (Asselin and Anh, 2008). The first principal component, $a_1 = \mathbf{X}^*_{\mathbf{ND}} v_1 = \sum_{j=1}^{D} v_{1,j} \mathbf{X}^*_{\bullet\mathbf{j}}$, is the one that reproduces the highest proportion of the 'total' variance in the data (computed as trace$(\Sigma) = \sum_{j=1}^{D} \lambda_j$). Thus, v_1 is the vector of nonnormalised weights for each indicator in the first principal component.

Importantly, being linear combinations of standardised variables, principal components do not have a specific unit of measure or a cardinal interpretation (see e.g. Jolliffe and Cadima, 2016). Once we have identified the first principal component, it can be rescaled in a meaningful way by means of a monotonic transformation that perfectly preserves the ordering of individuals by their scores in that component. In particular, the first principal component can be rescaled such that $\tilde{a}_1 \equiv \frac{a_1}{\sum_{j=1}^{D} v_j}$ with $V(\tilde{a}_1) = \frac{\lambda_1}{(\sum_{j=1}^{D} v_j)^2}$. This implies that

[9] The tetrachoric correlation coefficients can be estimated by taking the observed indicators two-by-two and applying an ML fit to a biprobit model that only includes constants as explanatory variables in each equation (Edwards and Edwards, 1984).

$$\tilde{a}_1 = \mathbf{X}^*_{\mathbf{ND}}\mathbf{w}^{PCA} = \sum_{j=1}^{D} w^{PCA_j}\mathbf{X}^*_{\bullet\mathbf{j}}, \tag{9}$$

where $w_j^{PCA} \equiv \frac{v_j}{\sum_{j'=1}^{D} v_{j'}} \forall j = 1, ..., D$ and $\sum_{j=1}^{D} w_j^{PCA} = 1$. These normalised weights are quantitative representations of the relative importance of each indicator in the first principal component.

Intuitively, the first-principal-component approach gauges the extent to which each indicator contributes to reproducing the largest portion of the total variance in the dataset. This is entirely driven by the indicators' variance-covariance matrix. After standardisation, the indicators that are, overall, highly correlated with the others will receive higher weights. The reason is that these highly correlated indicators form a 'dominant' indicator subset that essentially determines the first principal component. Conversely, those indicators that hold weak correlations with the elements of this 'dominant' indicator subset are regarded as redundant and thus receive lower weights in the principal component.

4 Empirical Illustration

To illustrate the violation of the key properties of monotonicity and subgroup consistency by a multidimensional counting poverty index under endogenous weights, we consider the cases of Ecuador and Uganda. Inherently, as for any endogenous weight procedure, PCA and frequency weights are data-adaptive techniques, which is why our empirical illustration concerns two countries with structurally different multidimensional poverty patterns (see OPHI (2018); Alkire et al. (2020) for a recent description of multidimensional poverty patterns in these countries).

4.1 Data and Indicators

We use nationally representative household-level data for both countries: Encuesta de Condiciones de Vida 2013/14 in the case of Ecuador (N=108,093) and the Uganda National Panel Survey 2015/16 (N=17,465).

Based on the UNDP–OPHI global MPI, our analysis considers ten indicators pertaining to three wellbeing dimensions: health, education and living standards (OPHI, 2018; Alkire and

Kanagaratnam, 2020). We argue that this index is ideal for our illustration due to its wide acceptance in academic and policymaking spheres (Atkinson, 2019; World Bank, 2018). The dimensions and their indicators along with the deprivation thresholds are presented in Table 1.

Deprivation headcount rates are lower in Ecuador compared to Uganda in every single observed indicator (Fig. 1). If one defines a person as being multidimensionally poor by the union approach, then the deprivation rates in Fig. 1 allow the structure of the poverty measure to be intuitively predicted by a wide family of frequency weights. For instance, if rare deprivations are considered particularly important to gauging poverty (Deutsch and Silber, 2005), then deprivations with the lowest frequencies will be assigned higher weights. This is the case for electricity and school attendance in Ecuador, and nutrition and school attendance in Uganda. Conversely, deprivations with the highest frequencies in each country will be assigned lower weights, as they are more commonly observed in the data. This is the case for nutrition and sanitation in Ecuador, and cooking fuel and electricity in Uganda.

Having a sense of the weights that would be assigned to each indicator by a covariance-based technique, such as PCA, is less straightforward. The reason is that they are much more complex transformations of the interlinkages between standardised deprivation vectors. As a starting point, however, it is important to stress that the joint distribution of deprivations in both countries is fundamentally different (Fig. 2). Around half of the population in Ecuador does not face any deprivation, and one out of four people faces one single deprivation. Meanwhile, the modal number of simultaneous deprivations in Uganda is five, and less than 5% of the population face just one deprivation.

A quick visual inspection of the tetrachoric correlation coefficients in the Ecuador data (Table 2) shows that the health indicators (nutrition and child mortality) tend to have relatively low correlations with the rest of the indicators: their correlation coefficients – in absolute value – range from 0.064 to 0.226. In turn, the living standard indicators tend to be more highly correlated with the rest, with coefficients ranging from 0.116 to 0.816. Thus, by a first-principal-component PCA approach, the health indicators would tend to have lower weights than the living standard indicators.

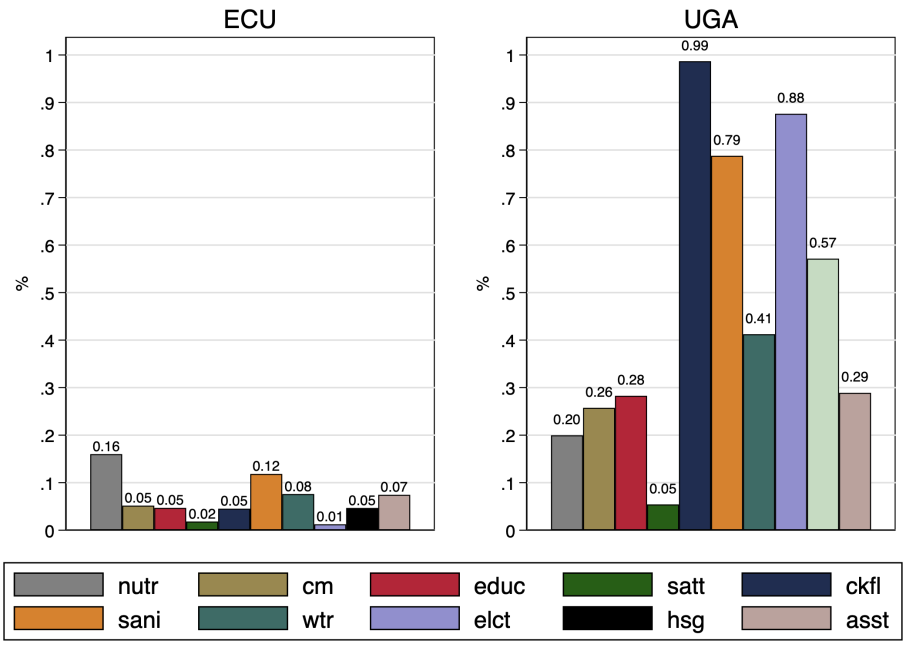
Due to the higher frequency of multiple deprivations in Uganda, clear correlation patterns for each indicator are less discernible at first glance (Table 3). Note, however, that the child mortality indicator appears to be the one most weakly correlated with all the rest; it has tetrachoric coefficients lower than 0.10 with school attendance, cooking fuel, housing and assets. Thus, child mortality would tend to have a relatively low weight in a first-principal-

Table 1: Dimensions, observed indicators and deprivation thresholds

Dimension	Indicators	A person is deprived if...
Health	Nutrition	Any household member under 70 for whom there is nutritional information is undernourished.
	Child mortality	Any child has died in the household in the last five years preceding the survey.
Education	Years of schooling	No household member aged 10 years or older in the household has completed six years of schooling.
	School attendance	Any school-aged child in the household is not attending school up to the age at which he/she would complete class 8.
Living Standards	Cooking fuel	The household cooks with dung, agricultural crop, shrubs, wood, coal or charcoal.
	Sanitation	The household's sanitation facility is not improved (according to SDG guidelines), or it is improved but shared with other households.
	Drinking water	The household does not have access to improved drinking water (according to SDG guidelines, or safe drinking water is at least a 30-minute walk from home, roundtrip.
	Electricity	The household has no electricity.
	Housing	The household has inadequate housing: the floor is of natural meterials or the roof or walls are of rudimentary materials.
	Assets	The household does not own more than one radio, TV, telephone, computer, animal cart, bike, motorbike or refrigerator and does not own a car or truck.

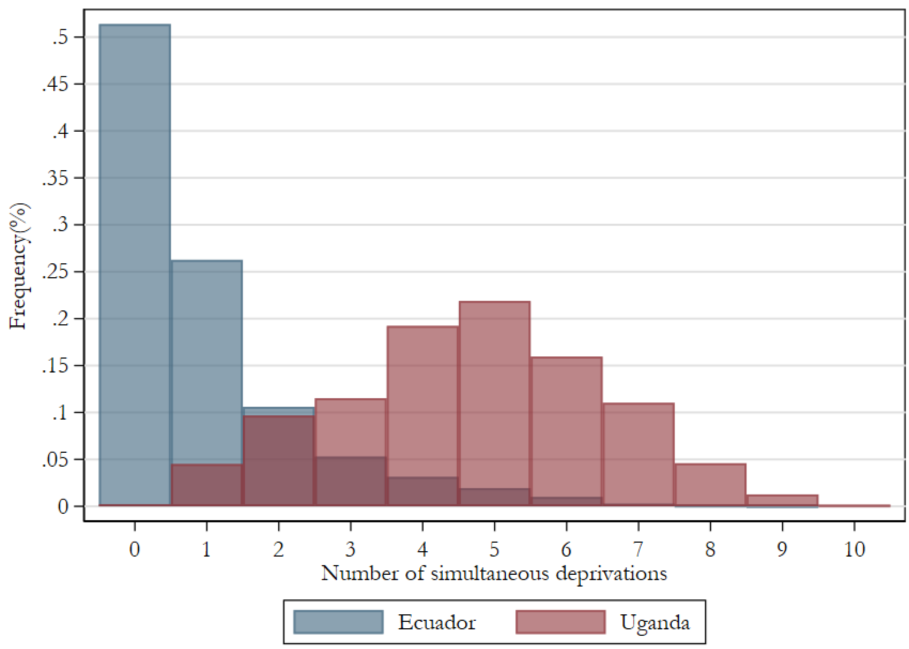
Source: Adapted from OPHI (2018).

Figure 1: Deprivation headcount ratios



Note: *Nutr*: Nutrition; *Cm*: Child mortality, *Educ*: Years of schooling; *Satt*: School attendance; *Ckfl*: Cooking fuel; *Sani*: Sanitation; *Wtr*: Drinking water; *Elct*: Electricity; *Hsg*: Housing; *Asst*: Assets

Figure 2: Distribution of simultaneous deprivations



component PCA-based poverty measure.

4.2 Weighting Vectors, Monotonicity and Subgroup Consistency: A Simulation Analysis

Let us now turn to an empirical assessment of the response of multidimensional poverty indices to changes in the deprivation matrices. We will compare indices based on exogenous and endogenous weights in order to show that the latter weighting structures violate monotonicity and subgroup consistency. By focusing on Ecuador and Uganda, we illustrate that these violations take place in contexts of both high and low overall structural poverty. We omit country labels for the sake of notational simplicity.

Our baseline scenario is defined by the deprivation matrices $\mathbf{X}_{\mathbf{ND}}$ effectively observed in both datasets. We simulate changes to these matrices by adding random deprivations in three indicators, one at a time: nutrition (*nutr*), access to safe drinking water (*wtr*) and access to electricity (*elct*). These indicators roughly cover the spectrum of relative deprivation prevalence in both countries as well as interlinkages with the other deprivations. The aim of our simulation analysis is not to perform an exhaustive scrutiny of all possible situations, but rather to offer a compelling illustration of axiom violations in a wide array of empirical heterogeneity.

For our simulations, for each country we assign random identifiers to the population that is nondeprived in each of the three indicators, one at a time. We use these identifiers to form random ventiles of nondeprived people in each indicator. Then, in a cumulative process for one indicator at a time, we gradually assign simulated deprivations to each random ventile; we first assign simulated deprivations to 5% of the relevant population, then to an additional 5% (for a total of 10%), and so on. At the higher end, we assign the simulated status of deprived to 95% of the part of the population that is originally nondeprived in the indicator that is the object of the simulation.

We will denote the ensuing deprivation matrices as $\mathbf{X}^{\mathbf{s_i}}_{\mathbf{ND}}$, with $s_i = \{0, 5\%, 10\%, ..., 95\%\}$ representing the proportion of nondeprived individuals that are assigned simulated deprivations in indicator i, where $i = \{nutr, elct, wtr\}$. Note that $\mathbf{X}^{0}_{\mathbf{ND}} = \mathbf{X}_{\mathbf{ND}}$ and that $\forall s_i' > s_i$, $\mathbf{X}^{\mathbf{s_i'}}_{\mathbf{ND}}$ objectively represents a Pareto-inferior state of affairs where nobody has fewer deprivations and at least one person has more deprivations vis-a-vis $\mathbf{X}^{\mathbf{s_i}}_{\mathbf{ND}}$. Only endogenous weighting procedures (i.e. data-adaptive techniques) will capture these changes and readjust the weights accordingly. Naturally, an exogenous weighting procedure is data-agnostic, thereby yielding

Table 2: Ecuador – Tetrachoric correlation coefficients

	Nutr	Cm	Educ	Satt	Ckfl	Sani	Wtr	Elct	Hsg	Asst
Nutr	1.000	0.112	-0.205	0.226	0.176	0.116	0.153	0.190	0.064	0.141
Cm	0.112	1.000	-0.109	0.167	0.225	0.149	0.134	0.222	0.152	0.168
Educ	-0.205	-0.109	1.000	0.056	0.369	0.248	0.209	0.252	0.375	0.435
Satt	0.226	0.167	0.056	1.000	0.342	0.297	0.253	0.369	0.206	0.361
Ckfl	0.176	0.225	0.369	0.342	1.000	0.537	0.542	0.721	0.587	0.714
Sani	0.116	0.149	0.248	0.297	0.537	1.000	0.500	0.632	0.406	0.587
Wtr	0.153	0.134	0.209	0.253	0.542	0.500	1.000	0.662	0.165	0.513
Elct	0.190	0.222	0.252	0.369	0.721	0.632	0.662	1.000	0.289	0.816
Hsg	0.064	0.152	0.375	0.206	0.587	0.406	0.165	0.289	1.000	0.472
Asst	0.141	0.168	0.435	0.361	0.714	0.587	0.513	0.816	0.472	1.000

Note: *Nutr*: Nutrition; *Cm*: Child mortality, *Educ*: Years of schooling; *Satt*: School attendance; *Ckfl*: Cooking fuel; *Sani*: Sanitation; *Wtr*: Drinking water; *Elct*: Electricity; *Hsg*: Housing; *Asst*: Assets

Table 3: Uganda – Tetrachoric correlation coefficients

	Nutr	Cm	Educ	Satt	Ckfl	Sani	Wtr	Elct	Hsg	Asst
Nutr	1.000	0.146	0.191	0.171	-0.070	0.226	0.109	0.346	0.162	0.134
Cm	0.146	1.000	0.154	0.092	0.019	0.142	0.100	0.286	0.076	0.085
Educ	0.191	0.154	1.000	0.392	0.126	0.410	0.211	0.628	0.431	0.546
Satt	0.171	0.092	0.392	1.000	0.118	0.316	0.136	0.321	0.365	0.433
Ckfl	-0.070	0.019	0.126	0.118	1.000	0.072	0.239	0.334	0.189	0.174
Sani	0.226	0.142	0.410	0.316	0.072	1.000	0.136	0.503	0.588	0.425
Wtr	0.109	0.100	0.211	0.136	0.239	0.136	1.000	0.471	0.273	0.135
Elct	0.346	0.286	0.628	0.321	0.334	0.503	0.471	1.000	0.700	0.609
Hsg	0.162	0.076	0.431	0.365	0.189	0.588	0.273	0.700	1.000	0.507
Asst	0.134	0.085	0.546	0.433	0.174	0.425	0.135	0.609	0.507	1.000

Note: *Nutr*: Nutrition; *Cm*: Child mortality, *Educ*: Years of schooling; *Satt*: School attendance; *Ckfl*: Cooking fuel; *Sani*: Sanitation; *Wtr*: Drinking water; *Elct*: Electricity; *Hsg*: Housing; *Asst*: Assets a simulation-invariant set of weights.

Throughout the simulation, we adopt a union approach to identify the multidimensionally deprived.[10] Hence, the weighting choices are only relevant for the individual poverty severity functions. We follow the global MPI (OPHI, 2018) to select exogenous weights corresponding to a nested-weight structure by which each dimension is assigned an equal weight (1/3). This reflects a consideration of health, education and living standards being equally important in gauging poverty. In turn, indicators within poverty dimensions are also assigned equal weights; the education and health indicators are assigned a 1/6 (=0.1667) weight, and each living standards dimension is assigned a 1/18 (=0.0667) weight. This reflects a judgment in which each observed indicator is deemed equally important within its respective dimension.

For endogenous weights, we take into account two weighting procedures. The first is PCA based on the tetrachoric correlations, and the second is frequency weighting operationalised by $f(\mathbf{X}_{\bullet \mathrm{j}}) = -\ln(\sum_{n=1}^{N} \rho_{nj}/N)$, for any observed indicator j (see Eq. 7). In line with the intuitive scrutiny of the data presented above, this configuration of frequency weights amounts to considering deprivations with low prevalence as more important in the poverty assessment.

For the sake of parsimony, let us focus on nutrition deprivation simulations as an example. In Tables 4 and 5, we show the weighting vectors for selected s_{nutr} values.

The intuition gained from the inspection of tetrachoric correlation matrices is corroborated. In the baseline PCA weight structure for Ecuador (PCA, $s_{nutr} = 0$ column in Table 4), the living standard indicators tend to have a higher weight, followed by the education indicators and then the health indicators. In Uganda (see PCA, $s_{nutr} = 0$ column in Table 5), child mortality has the lowest weight by this approach, while living standard indicators tend to have the highest. In both countries, when nutrition deprivations increase, the nutrition indicator is assigned lower weights. That is, the simulation reconfigures the correlation patterns between this indicator and the rest in such a way that nutrition becomes redundant in the first principal component. In compensation, all the other indicators in Ecuador are given higher weights in a relatively uniform manner. In Uganda, higher weights tend to be assigned to education, housing and assets.

Turning now to the frequency weights, once more we corroborate the intuition gained from the patterns of relative deprivation prevalence in each country. In the baseline frequency weighting vectors (Freq., $s_{nutr} = 0$ columns in Tables 4 and 5), electricity and school atten-

[10] This ensures the fulfillment of the focus axiom. We could easily extend our analysis to different identification cutoffs based on equation (1).

Table 4: Ecuador – Indicator weights by structure and proportion of people assigned simulated deprivations in nutrition

Indicator	Exog.	PCA			Freq.		
		s_{nutr}=0	s_{nutr}=50	s_{nutr}=95	s_{nutr}=0	s_{nutr}=50	s_{nutr}=95
Nutr	0.167	0.037	0.014	0.005	0.061	0.019	0.001
Cm	0.167	0.045	0.045	0.046	0.099	0.103	0.105
Educ	0.167	0.072	0.077	0.078	0.103	0.108	0.110
Satt	0.167	0.080	0.080	0.081	0.135	0.141	0.144
Ckfl	0.056	0.143	0.146	0.147	0.104	0.108	0.110
Sani	0.056	0.124	0.127	0.128	0.072	0.075	0.077
Wtr	0.056	0.114	0.117	0.118	0.086	0.090	0.092
Elec	0.056	0.144	0.147	0.148	0.148	0.154	0.157
Hsg	0.056	0.097	0.100	0.101	0.103	0.107	0.109
Asst	0.056	0.145	0.148	0.149	0.088	0.092	0.094
Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes:
1. *Nutr*: Nutrition; *Cm*: Child mortality, *Educ*: Years of schooling; *Satt*: School attendance; *Ckfl*: Cooking fuel; *Sani*: Sanitation; *Wtr*: Drinking water; *Elct*: Electricity; *Hsg*: Housing; *Asst*: Assets.
2. s_{nutr} reflects the percentage of households with simulated deprivations in nutrition.

Table 5: Uganda – Indicator weights by structure and proportion of people assigned simulated deprivations in nutrition

Indicator	Exog.	PCA			Freq.		
		s_{nutr}=0	s_{nutr}=50	s_{nutr}=95	s_{nutr}=0	s_{nutr}=50	s_{nutr}=95
Nutr	0.167	0.062	0.022	0.008	0.158	0.057	0.005
Cm	0.167	0.046	0.046	0.047	0.133	0.149	0.157
Educ	0.167	0.127	0.133	0.135	0.124	0.138	0.146
Satt	0.167	0.098	0.102	0.103	0.285	0.320	0.337
Ckfl	0.056	0.054	0.059	0.061	0.001	0.001	0.002
Sani	0.056	0.119	0.123	0.124	0.023	0.026	0.028
Wtr	0.056	0.075	0.078	0.079	0.087	0.097	0.102
Elec	0.056	0.155	0.160	0.162	0.013	0.014	0.015
Hsg	0.056	0.136	0.142	0.144	0.055	0.061	0.065
Asst	0.056	0.128	0.135	0.137	0.121	0.136	0.143
Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes:
1. *Nutr*: Nutrition; *Cm*: Child mortality, *Educ*: Years of schooling; *Satt*: School attendance; *Ckfl*: Cooking fuel; *Sani*: Sanitation; *Wtr*: Drinking water; *Elct*: Electricity; *Hsg*: Housing; *Asst*: Assets.
2. s_{nutr} reflects the percentage of households with simulated deprivations in nutrition.

dance have the highest weights in Ecuador, as do school attendance and nutrition in Uganda. In turn, the lowest weights are assigned to nutrition and sanitation in Ecuador and to cooking fuel and electricity in Uganda.

4.2.1 Violations of Monotonicity and Subgroup Consistency

We now assess the empirical response of three common additively decomposable societal poverty functions with linear severity components to a gradual increase in deprivations in both countries. Each function corresponds to one of the weighting procedures, namely, exogenous (denoted by EX), PCA and frequency-based (denoted by F). Omitting the country index, the poverty functions are, respectively: $P_1^{s_i} = \frac{1}{N}\sum_n^N C_n^{EX,s_i}$, $P_2^{s_i} = \frac{1}{N}\sum_n^N C_n^{PCA,s_i}$ and $P_3^{s_i} = \frac{1}{N}\sum_n^N C_n^{F,s_i}$. The poverty identification functions ψ^{s_i} are omitted purposefully, as we are adopting a union approach to poverty identification, meaning $\psi^{s_i}(t_n;0) = 1$ for all n, in order to secure fulfillment of the focus axiom.

Our simulations yield compelling, unequivocal results: (i) monotonicity and subgroup consistency are never violated under exogenous weights, but (ii) these axioms are bound to be violated under endogenous weights. For conciseness, we discuss only the results of simulated nutrition deprivations ($i = nutr$); all the other results lead to the same qualitative conclusions and can be found in Appendix B.

Let us first focus on violations of monotonicity. In both countries (see Fig. 3 for Ecuador and Fig. 4 for Uganda), the poverty measure constructed with exogenous weights, $P_1^{s_i}$, is in theory a nondecreasing function of s_i. In our empirical illustration it always increases as societal welfare deteriorates with a gradual increase in nutrition deprivations. That is, $P_1^{s_i'} - P_1^{s_i} \geq 0, \forall s_i' \geq s_i$. This is not true for $P_2^{s_i}$ and $P_3^{s_i}$. We can clearly see in Fig. 3 that in Ecuador, $P_2^{s_i} - P_2^{40_i} < 0$ for some $45\% \leq s \leq 95\%$ and $P_3^{s_i} - P_3^{35_i} < 0$ for all $40\% \leq s \leq 95\%$. Similarly, in Uganda (see Fig. 4), $P_2^{s_i} - P_2^{70_i} < 0$ for some $75\% \leq s \leq 95\%$ and $P_3^{s_i} - P_3^{50_i} < 0$ for all $55\% \leq s \leq 95\%$. Thus, there are many instances where these endogenous-weights-based poverty measures *decrease* when there has been an unequivocal *increase* in the proportion of people suffering nutrition deprivations, ceteris paribus. These are flagrant violations of monotonicity. As shown in Appendix B, we observe similar violations of monotonicity for other indicators such as water and electricity.

To assess possible violations of subgroup consistency, we perform another set of deprivation simulations in an identical way, except that the nondeprived population eligible for simulation deprivations is solely concentrated in one specific subnational region in each country.

Figure 3: Violations of monotonicity: Ecuador, simulated nutrition deprivations

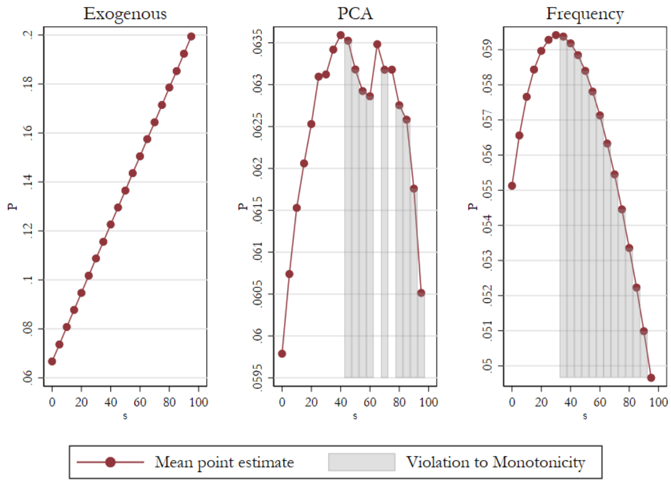
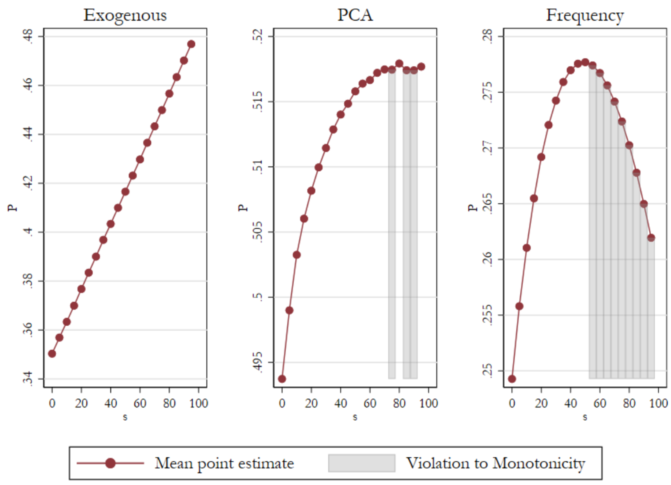


Figure 4: Violations of monotonicity: Uganda, nutrition indicator



Our data in Ecuador allows for a representative disaggregation of the additively decomposable poverty functions at the level of four geographical regions: Mountains (47.7% of the sample), Coast (33.1%), Amazon (17.6%) and Galapagos Island (1.6%). In Uganda, this is possible for four regions as well: Central (25.9%), Eastern (26.4%), Northern (23.9%) and Western (22.8%). The regional poverty measures take an identical form to their national-level counterparts. Omitting the country index, the poverty functions for a generic region R are given by $P^{s_i}_{1,R} = \frac{1}{N_R} \sum_{n \in R}^{N_R} C^{EX,s_i}_{n \in R}$, $P^{s_i}_{2,R} = \frac{1}{N_R} \sum_{n \in R}^{N_R} C^{PCA,s_i}_{n \in R}$ and $P^{s_i}_{3,R} = \frac{1}{N_R} \sum_{n \in R}^{N_R} C^{F,s_i}_{n \in R}$.

For our analysis, we focus on the Eastern region in Uganda and the Coast in Ecuador. We verified that choosing other regions does not alter our main qualitative results: subgroup con-

Figure 5: Violations of subgroup consistency: Ecuador, nutrition indicator

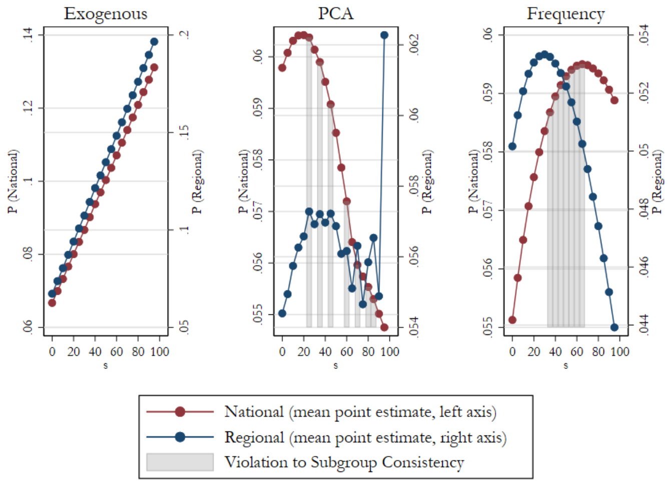
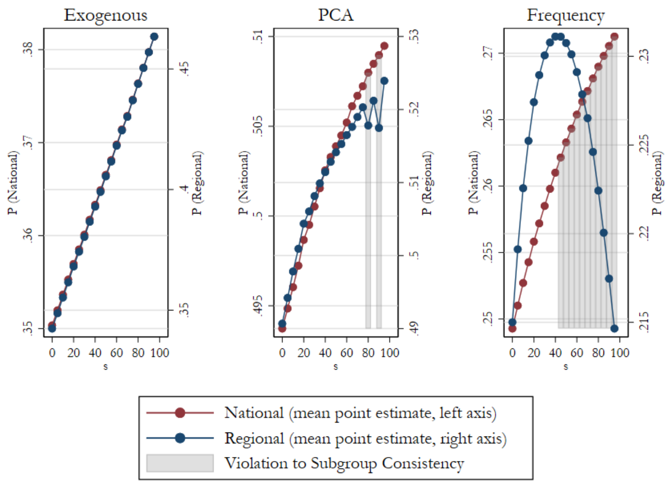


Figure 6: Violations of subgroup consistency: Uganda, nutrition indicator



sistency is never violated under exogenous weights, but it can be violated under endogenous weights.

In both countries, the regional and national poverty measures under exogenous weights increase (monotonically) with additional simulated nutrition deprivations in that region (see Fig. 5 for Ecuador and Fig. 6 for Uganda). That is, $(P^{s_i'}_{1,R} - P^{s_i}_{1,R} \geq 0) \implies (P^{s_i'}_{1} - P^{s_i}_{1} \geq 0), \forall s' \geq s$. However, there are several instances in which the subnational poverty measures based on PCA or frequency-based weights increase (respectively decrease) while the ensuing national poverty measures decrease (respectively increase), even though deprivations in all the other regions within each country are held constant. In Fig. 5, we can see that in Ecuador,

$(P_{2,R}^{s'_i} - P_{2,R}^{s_i} > 0) \implies (P_2^{s'_i} - P_2^{s_i} < 0)$ for some $s' > s$ with $s, s' > 20\%$, and $(P_{3,R}^{s'_i} - P_{3,R}^{s_i} < 0) \implies (P_3^{s'_i} - P_3^{s_i} > 0), \forall\ 30\% \leq s, s' \leq 65\%$ such that $s' > s$. We can clearly see violations of subgroup consistency in Uganda as well (see Fig. 6): $(P_{2,R}^{s'_i} - P_{2,R}^{s_i} < 0) \implies (P_2^{s'_i} - P_2^{s_i} > 0)$ for some $s' > s$ with $s, s' > 75\%$, and $(P_{3,R}^{s'_i} - P_{3,R}^{s_i} < 0) \implies (P_3^{s'_i} - P_3^{s_i} > 0), \forall s, s' > 40\%$ such that $s' > s$. These violations under endogenous weights do not seem to follow any particular pattern. In some cases we observe several violations, while in others, just a few. The important point is that, under these circumstances, any comparative exercise will be difficult because even if poverty in the affected region decreases, poverty at the national level might increase, ceteris paribus.

5 Endogenous Weights and Measurement Externalities

Why do we observe these violations of basic properties when using endogenous weights? In this section, we investigate this issue in greater depth. Our focus will be on the weighted deprivation score (or counting function) C_n (see (2)); that is where the endogenous weights come into play, because the severity functions depend on these scores under the counting approach to poverty measurement.[11] We establish that the change in one person's deprivation status in one indicator changes the counting functions for everyone else too. Thus, there are clear measurement externalities among individuals, which, as will be shown later, lead to situations where fundamental properties of the poverty functions are violated.

Consider a deprivation matrix $\mathbf{X_{ND}}$ where individual i is deemed poor. Let $\mathbf{X'_{ND}}$ be obtained from $\mathbf{X_{ND}}$ by a simple increase in deprivation in indicator j of individual i in $\mathbf{X_{ND}}$ (as defined in section 2.2). Then the change in the counting function of any individual $n \neq i$, who is also identified as poor, is

$$\Delta C_n = C_n^{X'_{ND}} - C_n^{X_{ND}} = \rho_{nj} \Delta w_j + \sum_{\substack{d=1 \\ d \neq j}}^{D} \rho_{nd} \Delta w_d, \tag{10}$$

where $\Delta C_n = C_n^{\mathbf{X}'_{ND}} - C_n^{\mathbf{X}_{ND}}$ and $\Delta w_d = w_d^{\mathbf{X}'_{ND}} - w_d^{\mathbf{X}_{ND}}$, $\forall d \in \{1, 2, ..., D\}$,. For simplicity of notation we write $\rho_{nj}^{\mathbf{X}_{ND}} = \rho_{nj}$ and $\rho_{nd}^{\mathbf{X}'_{ND}} = \rho'_{nd}$.

[11] As previously discussed, a combination of endogenous weights in the identification function with an intermediate approach to poverty identification leads to violation of the focus axiom. Meanwhile, the extreme identification approaches, union and intersection, never violate the focus axiom but neither do they rely on deprivation weights.

For the ith individual who became deprived in the jth indicator, we know that $\rho'_{ij}w'_j - \rho_{ij}w_j = w'_j$. Thus, ΔC_i due to a change in the deprivation of person i with respect to indicator j is given by

$$\Delta C_i = w'_j + \sum_{\substack{d=1 \\ d\neq j}}^{D} \rho_{id}\Delta w_d. \tag{11}$$

As long as person i is also deprived in some other indicator, the changes in the other weights produced by the change in i's status regarding j (i.e. $\Delta w_d, \forall d \neq j$) also affect the total change in C_i. These same changes in weights led by the change in deprivation status of person i in indicator j produce, in turn, changes in the counting function of every other person.

Note that in (10), the change in the counting function will depend on how the endogenous weights change (the signs of Δw_d $\forall d$), which, in turn, depends on the weighting rule. Hence, a priori, the change in any person's counting function (which in turn affects their individual poverty measure, p_n) is ambiguous. Proposition 1 captures how changes in ρ_{ij} can impact weights in each indicator and, through that channel, the counting function of everybody besides person i.

Proposition 1 *Suppose $\mathbf{X}'_{\mathrm{ND}}$ is obtained from $\mathbf{X}_{\mathrm{ND}}$ by a simple increase in deprivation in indicator j for individual i. For all $n \neq i$:*

(i) if $\rho_{nd} = 0, \forall d \in \{1,...,D\}$ or $\rho_{nd} = 1, \forall d \in \{1,...,D\}$ then $\Delta C_n = 0$

(ii) if $0 < \sum_{d=1}^{D} \rho_{nd} < D$ then $\begin{cases} \Delta C_n \lesseqqgtr 0 \Longleftrightarrow \Delta w_j \lesseqqgtr 0 & \text{if } \rho_{nj} = 1 \\ \Delta C_n \gtreqqless 0 \Longleftrightarrow \Delta w_j \lesseqqgtr 0 & \text{if } \rho_{nj} = 0 \end{cases}$.

Proof: See Appendix A.

Proposition 1 states that the impact of the increased deprivation of individual i in indicator j on the counting function of other individuals, $n \neq i$, depends only on the direction of change in the weight of indicator j, in combination with the deprivation status of n in indicator j. Note that indicator j is the only one where the number of deprived people changes. It plays a central role in understanding the changes in the counting function and, as a result, changes in individual poverty levels. If individual n is deprived in j, then an increase (respectively decrease) in the weight of j leads to an increase (respectively decrease) in n's counting function. If n is not deprived in j, then an increase (respectively decrease) in the weight of

j reduces (respectively increases) n's counting function. Naturally, there is no impact if n is either not deprived in any indicator or deprived in all indicators.

6 Endogenous Weights and Societal Poverty

In this section, we discuss how endogenous weights impact on the fulfillment of important properties of the societal poverty measures.

6.1 Monotonicity

One of the main implications of Proposition 1 for societal poverty indices based on endogenous weights (at least those of the form (6)) is that they can violate the desirable axiom of monotonicity (Axiom 1). This violation implies, inter alia, that when poor individuals in a society become less deprived, societal poverty may increase, and vice versa. In order to understand how this situation comes about, it is important to derive the impact produced by this change in the deprivation status of person i on the societal poverty index.

For any $\mathbf{X_{ND}}$, let $\Pr[\rho_{nj}=1|n\neq i]\equiv\frac{1}{N-1}\sum_{n=1,n\neq i}^{N}\mathbb{I}(\rho_{nj}=1|n\neq i)$ (and a similar definition for $\Pr[\rho_{nj}=0|n\neq i]$). Suppose $\mathbf{X}'_{\mathbf{ND}}$ is obtained from $\mathbf{X_{ND}}$ by a simple increase in deprivation in indicator j of individual i. For any individual i', let the change in poverty be denoted by $\Delta p_{i'}=p_{i'}^{\mathbf{X}'_{ND}}-p_{i'}^{\mathbf{X}_{ND}}$. Then,

$$\begin{aligned}\Delta P = & \tfrac{1}{N}\Delta p_i \\ & +\tfrac{N-1}{N}\Pr[\rho_{nj}=1|n\neq i]\tfrac{1}{(N-1)\Pr[\rho_{nj}=1|n\neq i]}\textstyle\sum_{n=1,n\neq i}^{N}\mathbb{I}(\rho_{nj}=1|n\neq i)\Delta p_n \\ & +\tfrac{N-1}{N}\Pr[\rho_{nj}=0|n\neq i]\tfrac{1}{(N-1)\Pr[\rho_{nj}=0|n\neq i]}\textstyle\sum_{n=1,n\neq i}^{N}\mathbb{I}(\rho_{nj}=0|n\neq i)\Delta p_n,\end{aligned} \tag{12}$$

where

$$\Delta p_n=\psi(t_n;k)s(C_n^{X_{ND}}+\Delta C_n)-\psi(t_n;k)s(C_n^{X_{ND}}). \tag{13}$$

That is, the change in societal poverty, ΔP, depends on (i) the change in person i's individual poverty (Δp_i), (ii) the total change in deprivation of other individuals deprived in j (captured in (12) as the average change in the poverty of other people deprived in j $\left(\frac{1}{(N-1)\Pr[\rho_{nj}=1|n\neq i]}\sum_{n=1,n\neq i}^{N}\mathbb{I}(\rho_{nj}=1|n\neq i)\Delta p_n\right)$ multiplied by the proportion of other people deprived in j ($\Pr[\rho_{nj}=1|n\neq i]$)), and (iii) the total change in deprivation of other individuals who are not deprived in j (shown in (12) as the average change in the poverty of other people not deprived in j $\left(\frac{1}{(N-1)\Pr[\rho_{nj}=0|n\neq i]}\sum_{n=1,n\neq i}^{N}\mathbb{I}(\rho_{nj}=0|n\neq i)\Delta p_n\right)$ multiplied by the proportion of other people not deprived in j ($\Pr[\rho_{nj}=0|n\neq i]$).

In the following discussion, we show how the three components above react to an increase in one person's deprivation. First, we show that an increase in deprivation in any one indicator for any individual does not decrease their poverty.

Corollary 1 *Let $\mathbf{X}'_{\mathrm{ND}}$ be obtained from $\mathbf{X}_{\mathrm{ND}}$ by a simple increase in deprivation in indicator j of individual i. Then individual i's poverty function does not decrease; that is, $\Delta p_i \geq 0$.*

Proof: See Appendix A.

Since an increase in a person's deprivation does not decrease their individual poverty function, the main problem with counting poverty functions relying on endogenous weights lies elsewhere with the presence of measurement externalities.

Next, we investigate how the poverty of other individuals changes as a result of the change in i's deprivation. Two helpful corollaries stem from (10) combined with Proposition 1 and the definition of individual poverty (3).

Corollary 2 *Let $\mathbf{X}'_{\mathrm{ND}}$ be obtained from $\mathbf{X}_{\mathrm{ND}}$ by a simple increase in deprivation in indicator j of individual i. Suppose $\Delta w_j > 0$. For any individual $n \neq i$,*

$$\begin{array}{ll}\Delta p_n \geq 0 \iff \Delta w_j > |\sum_{d=1,d\neq j}^{D}\rho_{nd}\Delta w_d| & \text{if } \rho_{nj}=1,\\ \Delta p_n \leq 0 \iff \sum_{d=1,d\neq j}^{D}\rho_{nd}\Delta w_d < 0 & \text{if } \rho_{nj}=0.\end{array}$$

Similarly, we get the following result when $\Delta w_j < 0$.

Corollary 3 *Let $\mathbf{X}'_{\mathrm{ND}}$ be obtained from $\mathbf{X}_{\mathrm{ND}}$ by a simple increase in deprivation in indicator j of individual i. Suppose $\Delta w_j < 0$. For any individual $n \neq i$,*

$$\begin{array}{ll}\Delta p_n \leq 0 \iff \Delta w_j > \sum_{d=1,d\neq j}^{D}\rho_{nd}\Delta w_d & \text{if } \rho_{nj}=1,\\ \Delta p_n \geq 0 \iff \sum_{d=1,d\neq j}^{D}\rho_{nd}\Delta w_d > 0 & \text{if } \rho_{nj}=0.\end{array}$$

Corollary 2 and Corollary 3 demonstrate that, with endogenous weights, $\Delta\rho_{ij} \neq 0$ is bound to produce changes in the poverty of other individuals, i.e. $\Delta p_n \neq 0$ where $n \neq i$, which will differ based on their deprivation status regarding j. Therefore, the aforementioned average changes (among those deprived in j and among those not deprived in j) will bear *opposite signs*. Hence, a priori, expression (12) may be positive, negative or even nil. Thus, we can deduce the following result.

Proposition 2 *Let $\mathbf{X}'_{\mathrm{ND}}$ be obtained from $\mathbf{X}_{\mathrm{ND}}$ by a simple increase in deprivation in indicator j of individual i. Then $\Delta P = P^{\mathbf{X}'_{\mathrm{ND}}} - P^{\mathbf{X}_{\mathrm{ND}}} \gtreqless 0$, thereby violating monotonicity (Axiom 1).*

This is a general result, not relying on any particular functional form of the weighting function or any particular parameters or data. It demonstrates that the change in societal poverty, ΔP, resulting from a change in deprivation in any one indicator experienced by any one poor individual would be ambiguous, thereby violating monotonicity (Axiom 1). Therefore, under endogenous weights it is quite possible that if the deprivation of an individual increases in some indicator, overall poverty will decline (or vice versa).[12]

From (13) we know the magnitude of change depends on k; therefore, the same change in the deprivation status of i (regarding j) may generate different values and signs for ΔP, depending on the choice of k. Likewise, the specific functional forms chosen for the weights and the severity function, s, also influence the total effect.

Finally note that, by contrast, with exogenous weights, the score of everybody except i remains unaltered: $\Delta C_n = 0, \forall n \neq i$. Consequently, $\Delta p_n = 0, \forall n \neq i$. Hence, finally, $\Delta P = \frac{1}{N}\Delta p_i$. That is, with exogenous weights, societal poverty changes coherently with the change in person i's individual poverty, as the latter does not affect the poverty measurement of anybody else. Hence, monotonicity is fulfilled.

6.2 Subgroup Consistency

A key implication of Proposition 1 is that societal poverty indices based on endogenous weights (respecting the form (6)) can also violate the desirable property of subgroup consistency (Axiom 2). In other words, we may find that poverty in a subgroup of the population has declined, with the poverty of all the other subgroups remaining unchanged, yet the poverty of the whole society has increased. However, this does not mean that subgroup consistency is only relevant when the poverty of *one* subpopulation changes while the poverty of the other subpopulations remains unaltered. Through repeated application of the property

[12] Note that this result also holds for any hybrid weighting rule where endogenous weights have been used alongside exogenous weights.

of subgroup consistency, we can compare situations when the poverty of *one or more* of the subpopulations changes (Foster and Szekely, 2008). Violation of subgroup consistency, on the other hand, will essentially imply that societal poverty may increase even if the poverty of *all* the subpopulations has declined. Thus, this is a powerful axiom which ensures that changes in the poverty of the total population is consistent with the changes happening at the subpopulation level. We claim the following.

Proposition 3 *Suppose for any deprivation matrix* $\mathbf{X}_{\mathbf{ND}}$*, the societal poverty measure is given by an additively decomposable poverty function* $P(\mathbf{X}_{\mathbf{ND}}; \mathbf{w}, k)$*, where* $\mathbf{w}$ *represents the class of endogenous weights respecting equation (6). Then* $P(\mathbf{X}_{\mathbf{ND}}; \mathbf{w}, k)$ *fails to satisfy subgroup consistency (Axiom 2).*

Proof: See Appendix A.

Proposition 3 thus demonstrates, in general terms, that endogenous weights will lead to the violation of subgroup consistency (Axiom 2). Note that the class of endogenous weights used is very general.[13]

7 Conclusions

The use of endogenous weights in multidimensional poverty measurement has enjoyed some popularity, yet the implications of letting weights depend on the dataset have not been studied in depth, above and beyond some reflections and sensible warnings (e.g. Alkire et al., 2015). In this paper we focus on a broad class of endogenous weights based on several instances of policy applications. We find that endogenous weighting leads to violations of monotonicity and subgroup consistency for a general class of multidimensional poverty indices based on the counting approach. Changes in the deprivation status of a household (or individual) generate measurement externalities in the form of changes in the counting function of many other households (or people), despite the absence of any changes in the latter's deprivation profiles. These transformations operate through the effect of the original changes in deprivation status on the weights.

Even though we focus on poverty indices, our analysis is equally relevant to societal welfare indices. For example, in the case of asset indices, each binary indicator could denote owner-

[13] As before, this result will also apply to any hybrid weights which include both endogenous and exogenous weights, so long as the changes in deprivation status are happening in indicators where the weights are endogenous.

ship of a specific asset (e.g. equal to one) or lack thereof (e.g. equal to zero). The asset score for each individual or household could be the weighted sum of the binary indicators. This is equivalent to the counting function (equation 2) in our paper. Since we are measuring welfare, we can implicitly adopt a union approach to identification. The societal poverty index in our paper can then be interpreted as the societal asset index, which would essentially be the average of the individual or household asset scores. Monotonicity in that context would require that a loss of ownership of any asset (e.g. losing livestock) by an individual should not be accompanied by an increase in the overall societal asset index. Similarly, subgroup consistency would imply that if the asset score of a subgroup decreases, with asset scores of the other subgroups unchanged, then the overall societal asset index should not increase. In all such cases where the properties of monotonicity and subgroup consistency are required, our results will hold and the use of a broad class of endogenous weights would be problematic.

Likewise, our analysis also bears important implications for individual living standard indices based on binary indicators, such as the burgeoning asset indices pioneered by Filmer and Pritchett (2001) and prominent measures of material deprivation in Europe (see e.g. Guio et al., 2016). In these cases, if individual living standards are measured by a weighted sum of binary indicators using endogenous weights, then a change in one household's vector of binary indicators representing asset holdings would generate measurement externalities in the form of changes in the living standard scores of other households despite any objective changes in their asset holdings, and with the concomitant rerankings among many households not necessarily involving the household whose objective living conditions changed.

By contrast, societal composite measures based on exogenous weights will satisfy monotonicity and subgroup consistency, so that if one household's deprivation (respectively achievement) worsens, then the societal measure of poverty (respectively wellbeing) will not increase, ceteris paribus. At the individual level, a change in a household's asset holdings will produce many fewer rerankings, always involving that same household, if the living standards score is computed using exogenous weights.

Of course, resorting to exogenous weights involves tricky, even potentially arbitrary, choices. Best-practice suggestions for choosing exogenous weights are in their infancy but are certainly emerging. For instance, Esposito and Chiappero-Martinetti (2019) monitor multidimensional poverty in the Dominican Republic using exogenous weights generated from a field experiment (i.e. information independent from their main dataset).

One way of using endogenous weights while satisfying monotonicity and subgroup consistency could be to compute endogenous weights with one particular dataset and then leave them fixed for future comparisons. This is precisely what Asselin and Anh (2008) do in their application to poverty comparisons in Vietnam with weights derived from MCA. However, this option would not really simplify the complexity of the decision regarding weight selection, since one would still need to decide *which* dataset to use in order to compute the weights for poverty comparisons (e.g. should one use a particular dataset or pool datasets?). Moreover, as pointed out by Alkire et al. (2015, p. 99), if datasets are pooled to compute weights based on data-reduction techniques (e.g. MCA, principal component analysis, factor analysis), there is no guarantee that a poverty comparison will be robust to sample updating, e.g. adding new time periods and including the new datasets in a recalculation of weights. Clearly, the latter decisions are hardly less arbitrary than choosing a vector of exogenous weights.

References

Aaberge, R. and A. Brandolini (2014). Multidimensional poverty and inequality. In A. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2A, Chapter 3, pp. 141–216. North-Holland.

Abdu, M. and E. Delamonica (2018). Multidimensional child poverty: From complex weighting to simple representation. *Social Indicators Research 136*, 881–905.

Alkire, S. and J. Foster (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics 95*(7-8), 476–487.

Alkire, S., J. Foster, S. Seth, M. Santos, J. Roche, and P. Ballon (2015). *Multidimensional Poverty Measurement and Analysis*. Oxford University Press.

Alkire, S. and U. Kanagaratnam (2020). Revisions of the global multidimensional poverty index: indicator options and their empirical assessment. *Oxford Development Studies*, 1–15.

Alkire, S., U. Kanagaratnam, R. Nogales and N. Suppa (2020). Revising the global Multidimensional Poverty Index: Empirical insight and robustness. Oxford University Press. *Research in Progress 56a*. Oxford Poverty and Human Development Initiative, University of Oxford

Asselin, L.-M. (2009). *Analysis of Multidimensional Poverty: Theory and Case Studies*. Springer.

Asselin, L.-M. and V. T. Anh (2008). Multidimensional poverty and multiple correspondence analysis. In N. Kakwani and J. Silber (Eds.), *Quantitative Approaches to Multidimensional Poverty*. Palgrave Macmillan.

Atkinson, A. B. (2019). *Measuring poverty around the world.* Princeton University Press.

Blackorby, C., D. Primont, and R. Russell (1978). *Duality, Separability, and Functional Structure: Theory and Economic Applications*. North-Holland.

Bourguignon, F. and S. Chakravarty (2003). The measurement of multidimensional poverty. *Journal of Economic Inequality 1*, 25–49.

Cavapozzi, D., W. Han, and R. Miniaci (2015). Alternative weighting structures for multidimensional poverty assessment. *Journal of Economic Inequality 13*(3), 425–447.

Chakravarty, S. and C. D'Ambrosio (2006). The measurement of social exclusion. *Review of Income and Wealth 52*(3), 377–398.

Coromaldi, M. and C. Drago (2017). An analysis of multidimensional poverty: Evidence from Italy. In R. White (Ed.), *Measuring Multidimensional Poverty and Deprivation. Incidence and Determinants in Developed Countries*, Chapter 4, pp. 69–86. Palgrave MacMillan.

Datt, G. (2019). Multidimensional poverty in the Philippines, 2004–13: Do choices for weighting, identification and aggregation matter? *Empirical Economics 57*, 1103–1128.

Datt, G. (2019). Distribution-sensitive multidimensional poverty measures. *World Bank Economic Review 33*(3), 551–572.

Dávila Lárraga, L. G. (2016). How does prospera work?: Best practices in the implementation of conditional cash transfer programs in Latin America and the Caribbean. *Inter-American Development Bank IDB-TN-971*.

Decanq, K. and M. A. Lugo (2013). Weights in multidimensional indices of well-being: an overview. *Econometric Reviews 1*, 7–34.

Deutsch, J. and J. Silber (2005). Measuring multidimensional poverty: An empirical comparison of various approaches. *Review of Income and Wealth 51*(1), 145–174.

Dhongde, S. and R. Haveman (2017). Multi-dimensional deprivation in the U.S. *Social Indicators Research 133*(2), 477–500.

Dotter, C. and S. Klasen (2014). The concept of relative multidimensional poverty: An illustration using Indian DHS data. Paper presented at the IARIW 33rd General Conference, Rotterdam.

Edwards, J. and A. Edwards (1984). Approximating the tetrachoric correlation coefficient. *Biometrics 40*, 563.

Esposito, L. and E. Chiappero-Martinetti (2019). Eliciting, applying and exploring multidimensional welfare weights: Evidence from the field. *Review of Income and Wealth*. https://doi.org/10.1111/roiw.12407.

Ezzrari, A. and P. Verme (2012). A multiple correspondence analysis approach to the measurement of multidimensional poverty in Morocco, 2001–2007. World Bank Policy Research Working Paper 6087.

Filmer, D. and L. Pritchett (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India. *Demography 38*(1), 115–132.

Foster, J. and A. Shorrocks (1991). Subgroup consistent poverty indices. *Econometrica 59*(3), 687–709.

Foster, J. and M. Szekely (2008). Is economic growth good for the poor? Tracking low income using general means. *International Economic Review 49*(4), 1143–1172.

Guio, A.-C., E. Marlier, D. Gordon, E. Fahmy, S. Nandy, and M. Pomati (2016). Improving the measurement of material deprivation at the European Union level. *Journal of European Social Policy 26*(3), 219–333.

Heshmati, A., C. S. Bajalan, and A. Tausch (2008). Measurement and analysis of child well-being in middle and high income countries. *European Journal of Comparative Economics 5*(2), 187–249.

Howe, L. D., B. Galobardes, A. Matijasevich, D. Gordon, D. Johnston, O. Onwujekwe, R. Patel, E. A. Webb, D. A. Lawlor, and J. R. Hargreaves (2012). Measuring socio-economic position for epidemiological studies in low-and middle-income countries: A methods of measurement in epidemiology paper. *International Journal of Epidemiology 41*(3), 871–886.

Jolliffe, I. T. and J. Cadima (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374*(2065), 20150202.

Ki, J., S. Faye, and B. Faye (2005). Multidimensional poverty in Senegal: A non-monetary basic needs approach. PEP working paper series 2005-05.

Kolenikov, S. and G. Angeles (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth 55*(1), 128–165.

Krishnakumar, J. and A. L. Nagar (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, mimic and structural equation models. *Social Indicators Research 86*(3), 481–496.

MPPN (2019). Multidimensional Poverty Peer Network. https://www.mppn.org/ Accessed March 2019.

Nájera Catalán, H. E. and D. Gordon (2019). The importance of reliability and construct validity in multidimensional poverty measurement: An illustration using the multidimensional poverty index for Latin America (MPI-LA). *The Journal of Development Studies*, 1–21.

Njong, A. and P. Ningaye (2008). Characterizing weights in the measurement of multidimensional poverty: An application of data-driven approaches to Cameroonian data. OPHI Working Paper 21.

Noglo, Y. (2017). Nonmonetary poverty in Togo: A multidimensional approach. *Review of Social Economy 75*(2), 181–211.

OECD (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD.

OPHI (2018). Global Multidimensional Poverty Index 2018: The most detailed picture to date of the world's poorest people. *Oxford Poverty & Human Development Initiative, University of Oxford, UK*.

Pasha, A. (2017). Regional perspectives on the multidimensional poverty index. *World Development 94*, 268–285.

Rippin, N. (2016). Multidimensional poverty in Germany: A capability approach. *Forum for Social Economics 45*(2-3), 230–255.

Sen, A. (1976). Poverty: An ordinal approach to measurement. *Econometrica 44*(2), 219–231.

Silber, J. and G. Yalonetzky (2013). Measuring multidimensional deprivation with dichotomized and ordinal variables. In G. Betti and A. Lemmi (Eds.), *Poverty and Social Exclusion: New Methods of Analysis*, Chapter 2, pp. 9–37. Routledge.

Skoufias, E., B. Davis, and S. de la Vega (2001). Targeting the poor in Mexico: An evaluation of the selection of households into PROGRESA. *World Development 29*(10), 1769–1784.

Stern, S., A. Wares, and T. Epner (2018). 2018 Social Progress Index: Methodology summary. Technical report, Social Progress Imperative.

Tsui, K.-Y. (2002). Multdimensional poverty indices. *Social Choice and Welfare 19*, 69–93.

UNICEF (2005). Child poverty in rich countries. *Innocenti Report Card* (6).

UNICEF (2007). Child poverty in perspective: An overview of child well-being in rich countries. *Innocenti Report Card* (7).

WFP (2019). Comprehensive Vulnerability Monitoring Exercise Multidimensional Poverty Index. *World Food Program, Turkey*.

Whelan, C., B. Nolan, and B. Maitre (2014). Multidimensional poverty measurement in Europe: An application of the adjusted headcount approach. *Journal of European Social Policy 24*(2), 183–197.

Wittenberg, M. and M. Leibbrandt (2017). Measuring inequality by asset indices: A general approach with application to South Africa. *Review of Income and Wealth 63*(4), 706–730.

World Bank (2018). Poverty and shared prosperity 2018: Piecing together the poverty puzzle.

A Appendix: Proofs

Proof of Proposition 1:

Case (i): Suppose individual n is not deprived in any indicator. In that case, $\forall d \in \{1,2,...,D\}$, $\rho_{nd} = 0$. Thus, from (10), we know that $\Delta C_n = 0$. Now suppose individual n is deprived in all D indicators. Since $\sum_{d=1}^{D} w_d = 1$, we can deduce that

$$\sum_{d=1}^{D} \Delta w_j = 0. \tag{A1}$$

Thus,

$$\Delta w_j = -\sum_{\substack{d=1\\ d\neq j}}^{D} \Delta w_d. \tag{A2}$$

Hence, from (10), $\Delta C_n = 0$.

Case (ii): Suppose for n, $\rho_{nj} = 1$ and $\exists d \neq j$ such that $\rho_{nd} = 0$. Then, from (A2), we can infer

$$\Delta w_j > \sum_{\substack{d=1\\ d\neq j}}^{D} \rho_{nd} \Delta w_d ,$$

since the right-hand side of the inequality aggregates over only those indicators in which individual n is deprived, except j. Thus,

$$\Delta C_n \gtreqless 0 \Longleftrightarrow \Delta w_j = \gtreqless 0.$$

On the other hand, if, for n, $\rho_{nj} = 0$, then from (10) we get

$$\Delta C_n = \sum_{\substack{d=1\\ d\neq j}}^{D} \rho_{nd} \Delta w_d.$$

Then,

$$\sum_{\substack{d=1\\ d\neq j}}^{D} \rho_{nd} \Delta w_d \gtreqless 0 \text{if } \Delta w_j \lesseqgtr 0.$$

Thus, $\Delta C_n \gtreqqless 0 \Longleftrightarrow \Delta w_j = \lesseqqgtr 0$.

Proof of Corollary 1:
First we prove that $\Delta\rho_{ij} > 0$ leads to $\Delta C_i > 0$. From equation (11) we can get

$$\Delta C_i = w'_j + \sum_{\substack{d=1\\ d\neq j}}^{D} \rho_{id}\Delta w_d, \tag{A3}$$

where w'_j is the weight of indicator j in $\mathbf{X}'_{\mathbf{ND}}$. Since $\sum_{d=1}^{D}\Delta w_d = 0$, $\Delta w_j \gtrless 0$ implies $\sum_{d=1,d\neq j}^{D}\Delta w_d \lessgtr 0$. Thus,

$$\Delta w_j \;=\; \sum_{\substack{d=1\\ d\neq j}}^{D}\Delta w_d \;\geq\; \sum_{\substack{d=1\\ d\neq j}}^{D}\rho_{id}\Delta w_d \;. \tag{A4}$$

Suppose $\Delta w_j > 0$. Thus, from (A4), $|w'_j| > |\sum_{d=1,d\neq j}^{D}\rho_{id}\Delta w_d|$, which from (A3) implies $\Delta C_i > 0$. Likewise, if $\Delta w_j < 0$, we know from (A4) that $\sum_{d=1,d\neq j}^{D}\Delta w_d > 0$. Given $w'_j > 0$, we can deduce from (A3) that $\Delta C_i > 0$.

Let t_i be the (exogenously weighted) number of indicators in which individual i is deprived and k is the cutoff for the (weighted) number of indicators one has to be deprived in to be identified as poor. Then if $t_i \geq k$, given $\Delta C_i \geq 0$ and the definition of p_n, we can infer that $\Delta p_n \geq 0$. Likewise, if $t_i < k$ and $t'_i \geq k$ given $\Delta C_i > 0$, then again $\Delta p_n > 0$. Otherwise, $\Delta p_n = 0$.

Proof of Proposition 3:
Consider a deprivation matrix decomposed by subgroups $\mathbf{X}_{ND} = (\mathbf{X}_{N_1D} \,||\, \mathbf{X}_{N_2D})$ where $N = N_1 + N_2, \forall n \in \mathbf{X}_{N_1D}, \rho_{nj} = 1$ and $\forall n \in \mathbf{X}_{N_2D}, \rho_{nj} = 0$. Suppose $\mathbf{X}'_{ND} = (\mathbf{X}_{N_1D} \,||\, \mathbf{X}'_{N_2D})$, where $\mathbf{X}'_{N_2D}$ is obtained from $\mathbf{X}_{N_2D}$ by a simple increase in deprivation of person i in indicator j, i.e. $\Delta\rho_{ij} = 1$, $i \in X_{N_2D}$. Suppose for $\mathbf{X}'_{N_2D}$: $\Delta w_j = w_j(\rho_{ij} = 1) - w_j(\rho_{ij} = 0) < 0$.

To be subgroup consistent, it must be the case that $\Delta P^{\mathbf{X}'_{ND}-\mathbf{X}_{ND}} \lesseqqgtr 0$ if and only if $\Delta P^{\mathbf{X}'_{N_2D}-\mathbf{X}_{N_2D}} \lesseqqgtr 0$. Applying (12), we get

$$\Delta P^{\mathbf{X}'_{N_2D}-\mathbf{X}_{N_2D}} = \frac{1}{N_2}\Delta p_i^{\mathbf{X}'_{N_2D}-\mathbf{X}_{N_2D}} + \frac{1}{N_2}\sum_{n\neq i}^{N_2}\mathbb{I}(\rho_{nm} = 0)\Delta p_n^{\mathbf{X}'_{N_2D}-\mathbf{X}_{N_2D}}. \tag{A5}$$

In (A5), $\Delta p_i^{\mathbf{x}'_{N_2D}-\mathbf{x}_{N_2D}} \geq 0$ from Corollary (1). Also, $\Delta p_n^{\mathbf{x}'_{N_2D}-\mathbf{x}_{N_2D}} \geq 0$, $\forall n \neq i$ due to Corollary 3. Therefore, $\Delta P^{\mathbf{x}'_{N_2D}-\mathbf{x}_{N_2D}} \geq 0$.

Now,

$$\Delta P^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} = \frac{1}{N}\Delta p_i^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} + \frac{1}{N}[\sum_{n\neq i}^{N}\mathbb{I}(\rho_{nm}=0)\Delta p_n^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} + \sum_{n\neq i}^{N}\mathbb{I}(\rho_{nm}=1)\Delta p_n^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}}] \tag{A6}$$

Again, in (A6), $\Delta p_i^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} \geq 0$. Likewise, $\sum_{n\neq i}^{N}\mathbb{I}(\rho_{nm} = 0)\Delta p_n^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} \geq 0$. However, from Corollary 3, $\sum_{n\neq i}^{N}\mathbb{I}(\rho_{nm} = 1)\Delta p_n^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} \leq 0$. Therefore, $\Delta P^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} \lesseqqgtr 0$, unlike $\Delta P^{\mathbf{x}'_{N_2D}-\mathbf{x}_{N_2D}} \geq 0$. In fact, with $N_1 \to \infty$, we can obtain $\Delta P^{\mathbf{x}'_{ND}-\mathbf{x}_{ND}} \leq 0$.

B Appendix: Additional Deprivation Simulation Results

Figure 7: Violations of monotonicity: Ecuador

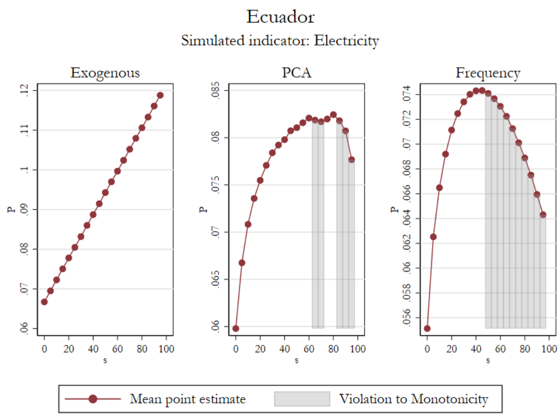


Figure 8: Violations of monotonicity: Uganda

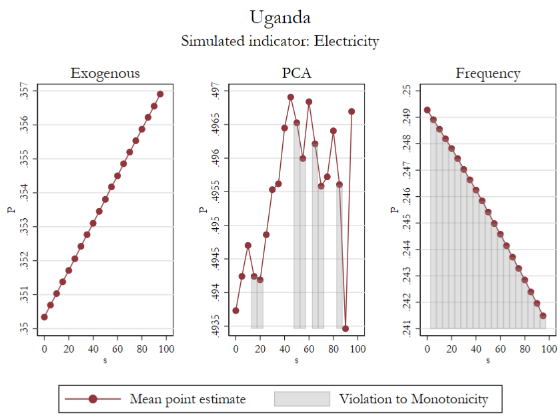


Figure 9: Violations of monotonicity: Ecuador

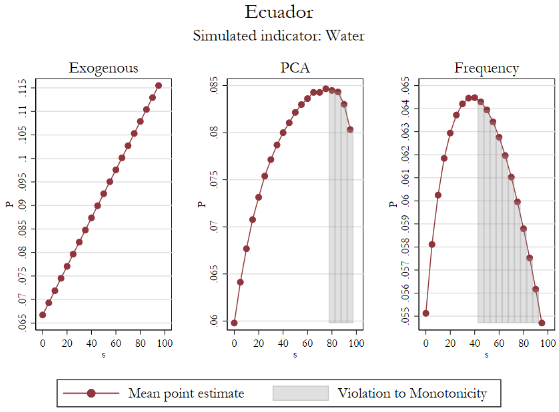


Figure 10: Violations of monotonicity: Uganda

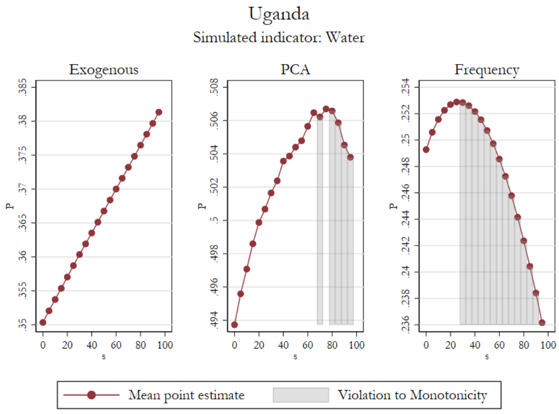


Figure 11: Violations of subgroup consistency: Ecuador

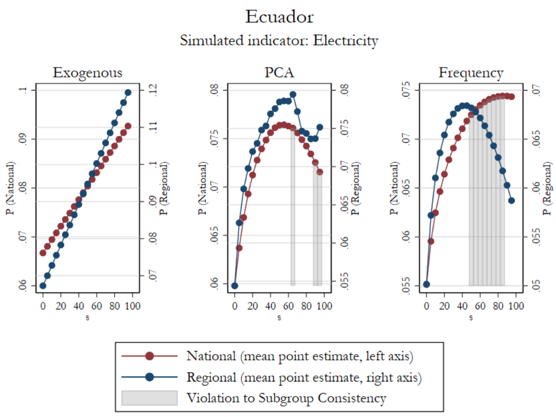


Figure 12: Violations of subgroup consistency: Uganda

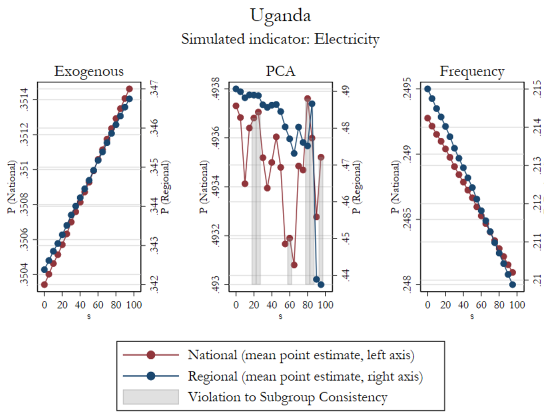


Figure 13: Violations of subgroup consistency: Ecuador

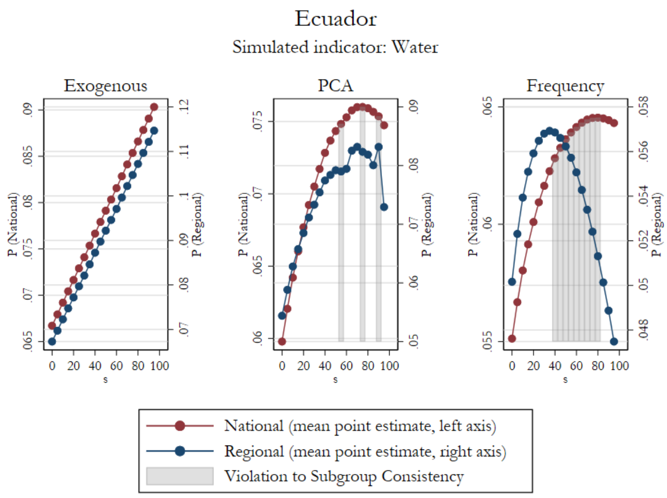


Figure 14: Violations of subgroup consistency: Uganda