Extending multidimensional poverty identification: from additive weights to minimal bundles

Sam Jones, UNU-WIDER (Mozambique)

14 February 2022

Agenda

(1) Motivation

Motivation

- Long-standing interest in multidimensional poverty
- Initial focus : problem of robustness
  - Intimately connected with the challenge of choosing weights/cut-off
  - Infinite number of these vectors, but set of outcomes (effectively) finite
  - Led me to the Good-Turing estimator of missing distributional mass for sensitivity analysis – see 'Measuring what's missing: practical estimates of coverage for stochastic simulations.' *Journal of Statistical Computation and Simulation*, 86(9).
- More recently I circled back to the same issues, taking a different track

The Journal of Economic Inequality
https://doi.org/10.1007/s10888-021-09477-8

Extending multidimensional poverty identification: from additive weights to minimal bundles

Sam Jones[1]

Received: 28 May 2019 / Accepted: 19 January 2021 / Published online: 17 May 2021

Abstract

In the popular class of multidimensional poverty measures introduced by Alkire and Foster (2011), a threshold switching function is used to identify who is multidimensionally poor. This paper shows that the weights and cut-off employed in this procedure are generally not unique and that such functions implicitly assume all groups of deprivation indicators of some fixed size are perfect substitutes. To address these limitations, I show how the identification procedure can be extended to incorporate any type of positive switching function, represented by the set of minimal deprivation bundles that define a unit as poor. Furthermore, the Banzhaf power index, uniquely defined from the same set of minimal bundles, constitutes a natural and robust metric of the relative importance of each indicator, from which the adjusted headcount can be estimated. I demonstrate the merit of this approach using data from Mozambique, including a decomposition of the adjusted headcount using a 'one from each dimension' non-threshold function.

Keywords Multidimensional poverty · Switching functions · Mozambique

(2) Identification via weights/cut-off

Terminology

To avoid confusion ...

1. *Dimensions* : refer to underlying indicators of deprivation (here, binary & assumed given by the data)
2. *Domains* : are groups of deprivation indicators (e.g., education, health, housing)
3. *Identification* : is the binary decision process that determines who is counted poor (row-wise operation)
4. *Intensity* : is the degree of poverty experienced by a unit, ranging from zero to one
5. *Aggregation* : is the final step transforming observations on multiple units to a single overall metric (ignore this here)

Alkire-Foster (AF) identification procedure

Inputs:

- D is an $n \times m$ matrix of binary deprivation indicators with n units and m deprivation dimensions
- Elements of D are d_{ij}, such that $d_{ij} = 1$ if unit $i \in \{1, \ldots, n\}$ is deprived in dimension $j \in \{1, \ldots, m\}$
- A normalized vector of weights $\vec{w} = (w_1, w_2, ..., w_m)$; and a cut-off ($k$)

Output: unit i's poverty status is identified from a positive threshold switching function:

$$h_i = 1 \left[\sum_{j=1}^{m} d_{ij} w_j \geq k \right] \quad (1)$$

$$\text{where } \forall j : 0 < w_j < 1, \ \sum_{j=1}^{m} w_j = 1, \ 0 < k \leq 1$$

Limitation 1: non-uniqueness

Distinct choices for $(\vec{w}, k)$ can map to identical **poverty identification** results for a given input matrix D.

Examples:

	w_1	w_2	w_3	w_4	w_5	k	H	M_0
1(a)	0.200	0.200	0.200	0.200	0.200	0.800	0.346	0.307
1(b)	0.179	0.156	0.222	0.235	0.207	0.684	0.346	0.303
1(c)	0.165	0.221	0.184	0.211	0.220	0.673	0.346	0.307
c.v.	0.097	0.171	0.096	0.084	0.048	0.098	0.000	0.007
2(a)	0.328	0.047	0.459	0.068	0.099	0.515	0.490	0.422
2(b)	0.300	0.027	0.483	0.050	0.140	0.514	0.490	0.427
2(c)	0.136	0.114	0.409	0.148	0.194	0.536	0.490	0.398
c.v.	0.408	0.724	0.083	0.588	0.331	0.024	0.000	0.036

Limitation 1: non-uniqueness

Consequences:

- Minimally, differentiating poverty definitions by $(\vec{w}, k)$ may give a false sense of precision and specificity.

- Weights do not always reliably or directly reflect differences in the relative importance of each dimension in the identification process (ranks are not stable) ...

  ... so, what do really mean by "relative importance"?

- Contradiction between *finite* identification outcomes vs. *infinite* $(\vec{w}, k)$ choices ...

  ... but (confusingly?), M_0 *is* sensitive to choice of weights

Limitation 2: functional restrictions

The Alkire-Foster approach assumes perfect substitutability between all subsets of intersecting deprivations of a given size.

What does this mean?

- For a given choice of $(\vec{w}, k)$, order the weights from smallest to largest: $w_{(1)}, w_{(2)}, \ldots, w_{(m)}$
- Now, find the smallest p such that: $\sum_{j=1}^{p} w_{(j)} \geq k$ ...

  ... then all units deprived in at least p deprivations will be identified as poor ...

$\Longrightarrow$ Any such p-tuple of deprivations is a substitute for any other in the production of deprivation.

Limitation 2: functional restrictions

- Means we generally cannot encode functions where the degree of substitutability varies between distinct sub-groups of deprivation indicators (of length $\geq p$)
- Example: how to encode 'one from each dimension'?

	Deprivation indicator	Dimension	Weight
d_1	Inadequate sanitation	A	1/4
d_2	Inadequate housing materials	A	1/4
d_3	No means of transport	B	1/6
d_4	No phone, radio or TV	B	1/6
d_5	No fridge, iron or bed	B	1/6

- $k = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}$? (All A or all B)
- $k > 1/2$? (Requires at least 3 dimensions)

(3) Identification via minimal bundles

Extending identification

The previous results reflect the fact the AF procedure relies on a threshold switching function, which is a restricted (linear) type of positive switching function (Boolean function).

Insight: more general positive switching functions can be used!

Clarification:

- A switching function is just a mapping $f : \{0, 1\}^m \rightarrow \{0, 1\}$, also known as a voting game
- Positive switching functions are monotonically increasing
- Any such function is uniquely defined by its set of true points – all unique combinations of inputs (deprivations) that yield a positive outcome, $\mathcal{T}(f) \subset \mathcal{B}^m$.
- Evident from a truth table, which indexes the true points!

	Bundles						True points of different identification functions						
id.	d_1	d_2	d_3	d_4	d_5	π	1(a)	1(b)	1(c)	2(a)	2(b)	2(c)	3
1	0	0	0	0	0	0.089	0	0	0	0	0	0	0
2	1	0	0	0	0	0.082	0	0	0	0	0	0	0
3	0	1	0	0	0	0.021	0	0	0	0	0	0	0
4	1	1	0	0	0	0.125	0	0	0	0	0	0	0
5	0	0	1	0	0	0.107	0	0	0	0	0	0	0
6	1	0	1	0	0	0.048	0	0	0	1	1	1	1
7	0	1	1	0	0	0.006	0	0	0	0	0	0	1
8	1	1	1	0	0	0.024	0	0	0	1	1	1	1
9	0	0	0	1	0	0.000	0	0	0	0	0	0	0
10	1	0	0	1	0	0.000	0	0	0	0	0	0	1
11	0	1	0	1	0	0.000	0	0	0	0	0	0	1
12	1	1	0	1	0	0.000	0	0	0	0	0	0	1
13	0	0	1	1	0	0.000	0	0	0	1	1	1	0
14	1	0	1	1	0	0.000	0	0	0	1	1	1	1
15	0	1	1	1	0	0.000	0	0	0	1	1	1	1
16	1	1	1	1	0	0.000	1	1	1	1	1	1	1
17	0	0	0	0	1	0.001	0	0	0	0	0	0	0
18	1	0	0	0	1	0.008	0	0	0	0	0	0	1
19	0	1	0	0	1	0.006	0	0	0	0	0	0	1
20	1	1	0	0	1	0.053	0	0	0	0	0	0	1
21	0	0	1	0	1	0.017	0	0	0	1	1	1	0
22	1	0	1	0	1	0.039	0	0	0	1	1	1	1
23	0	1	1	0	1	0.012	0	0	0	1	1	1	1
24	1	1	1	0	1	0.115	1	1	1	1	1	1	1
25	0	0	0	1	1	0.000	0	0	0	0	0	0	0
26	1	0	0	1	1	0.007	0	0	0	0	0	0	1
27	0	1	0	1	1	0.003	0	0	0	0	0	0	1
28	1	1	0	1	1	0.047	1	1	1	1	1	1	1
29	0	0	1	1	1	0.005	0	0	0	1	1	1	0
30	1	0	1	1	1	0.017	1	1	1	1	1	1	1
31	0	1	1	1	1	0.016	1	1	1	1	1	1	1
32	1	1	1	1	1	0.151	1	1	1	1	1	1	1
H							0.346	0.346	0.346	0.491	0.491	0.491	0.552

Extending identification

Straight-forward to see that the output vector associated with f in the truth table is sufficient for poverty identification – i.e. it encodes each unique bundle of deprivations as either zero or one.

So we can directly use this to derive the poverty headcount:

$$
\begin{aligned}
H(D; f) &= \sum_{x \in \mathcal{B}^m} f(x)\pi(D; x) && \text{(2a)} \\
&= \sum_{x \in \mathcal{T}(f)} \pi(D; x) && \text{(2b)}
\end{aligned}
$$

where $\pi(D; x)$ is the (sample weighted) proportion of observations from matrix D with bundle x.

Extended identification in practice

1. Choose a set of *m* deprivation indicators (perhaps grouped into broader dimensions);
2. Collapse the observed data, *D*, into the collection of at most 2^m unique deprivation bundles, collecting sample or frequency weights as desired (by subgroup);
3. Select an identification rule, *f* which can be *any* positive Boolean function or monotonic game;
4. Identify all true points of *f* (describing which of all feasible bundles define a unit as poor); and
5. Calculate the multidimensional poverty headcount as per equation (2b).

Representing switching functions

Advantage of the AF approach is that threshold switching functions are easy & compact to write-down.

What about non-threshold functions?

Not necessarily much more complicated:

- *Nested AF-type formulation* : any positive Boolean function (monotonic voting game) can be expressed as the intersection of $c \geq 1$ threshold functions
- *Minimal bundles* : from monotonicity, we only need the set of bundles that is minimally sufficient to classify a unit as poor (i.e., bundles in which all dimensions are 'swing') ...Then, the underlying Boolean function is given by the union of these minimal deprivation bundles. *Example*:

$$f_3 = 1[(d_1 + d_2) \cdot (d_3 + d_4 + d_5) \geq 1] \quad (3)$$

Intensity (adjusted headcount)

Without explicit (*ex ante*) weights, how can one calculate the adjusted headcount?

Game-theory provides an answer – the 'power' of each dimension indicates which dimensions are (relatively more) decisive in switching the function from zero to one.

Choice of $f \implies$ true points $\implies$ power (implicit weights)

Different power concepts:

- *Banzhaf power*: relative frequency of each dimension in the collection of minimal bundles.
- *Shapley value*: relative frequency each dimension is pivotal among all true bundles (sequential criteria; not here).

Metrics of influence for alternative identification functions

Example:

	f_1			f_2			f_3	
	$\vec{w}$	BZ	mBZ	$\vec{w}$	BZ	mBZ	BZ	mBZ
d_1	0.200	0.200	0.200	0.328	0.200	0.184	0.250	0.219
d_2	0.200	0.200	0.200	0.047	0.100	0.163	0.250	0.219
d_3	0.200	0.200	0.200	0.459	0.300	0.286	0.167	0.188
d_4	0.200	0.200	0.200	0.068	0.200	0.184	0.167	0.188
d_5	0.200	0.200	0.200	0.099	0.200	0.184	0.167	0.188
M_0	0.307	0.307	0.307	0.422	0.388	0.387	0.426	0.420

(Various decompositions follow naturally)

(4) Summary

Summary

The paper builds on a long tradition of scholarship.

Most of the ideas are not new.

AF approach extremely powerful. ... But, does have limitations that may be relevant in some situations:

- Weights/cut-off choices are non-unique
- Only certain poverty definitions can be encoded

Main contributions:

- Showed how to extended poverty identification to other switching functions, represented by the set of minimal bundles, which nests AF approach as a special case
- Suggested use of game-theoretic weights, which are derived uniquely from the set of minimal bundles

(I have a bunch of `R` functions that implement the analysis.)

(5) Future directions?

Future directions?

Welcome ideas and thoughts on how this work can be applied and/or extended.

Possible paths:

1. Clarify game-theoretic weight concepts: which is most appropriate?
2. Place global MPI within this framework (apply game-theoretic weights; use a ‘one from each dimension’ definition)

   ... suggestive evidence in the paper, from Mozambique, indicates material differences, esp. for poverty decompositions
3. Subjective poverty definitions via choice of bundles?
4. Robustness analysis, by varying (minimal) bundles