The Research Agenda on Multidimensional Poverty Measurement: Important and As-yet Unanswered Questions

Sabina Alkire*

June 2018

Abstract
The application of multidimensional poverty measures is proliferating, in part due to the emphasis in Goal 1 of the Sustainable Development Goals (SDGs) on ending poverty in all its forms and dimensions. This paper first traces the emergence of a priority for non-monetary poverty measures in key texts that considered then set out the SDGs. It then outlines some vital and feasible research questions on a sub-set of fascinating empirical topics on counting-based multidimensional measures. The topics covered here relate to the SDGs’ focus on measuring the multidimensional poverty of men, women, and children. Building on the existing literature, fascinating questions remain in terms of measurement design (the selection of dimensions, indicators, cutoffs, and weights), the analysis of multidimensional poverty measures, their application to child poverty and their implementation using gendered data. In each of these areas, it is expected that significant breakthroughs are possible.

Keywords: Multidimensional Poverty, Poverty Measurement, Child Poverty, Gender, Robustness

JEL classification: I3, O21, D04, D78

* Oxford Poverty and Human Development Initiative (OPHI), University of Oxford. Email: sabina.alkire@qeh.ox.ac.uk. This study has been prepared within the OPHI theme on multidimensional measurement.
Acknowledgements
I would like to thank Enrica Chiappero Martinetti for suggesting this article and the Editors for their incisive comments, and the participants in an OPHI seminar for feedback. All errors remain my own.

Funding information: This work was supported by the Economic and Social Research Council’s grant number ES/N01457X/1.

Rosa Luxemburg, a socialist and social activist, is said to have observed that ‘[t]hose who do not move, do not notice their chains’. That may be true in some political circumstances. But a lack of noticing applies in other spaces too. Movement can highlight unexpected limitations – such as shortness of breath if one has not been out running recently, or stiff muscles if stretching is not habitual. This is no less the case in intellectual pursuits. When we seek to move towards new applications or new theoretical connections related to measurement or quantitative analysis, we come up against a series of limitations that constrain our current field of practice. In such a case, the limitations are hardly chains – that is, externally imposed and intentional constraints aimed at hampering progress. Rather, they are more akin to where the pavement or ‘sidewalk’ ends. We observe an abrupt but obvious, even conscious, frontier – one that has often been left with the expectation that others will extend work further in that direction. And yet, just as clearly, such a path has not yet been built, so others cannot yet walk upon it to new destinations that might be of great interest to them. So overall, it would be far better if the new pavements were built.

This paper maps the ending of more than a few sidewalks in the literature on multidimensional poverty measurement. The entire literature draws on Amartya Sen’s pioneering work on poverty and capability, and is oriented more recently to the sustainable development goals. In this paper, we focus on the subset of questions related to ending poverty for men, women and children. This paper identifies clear and abrupt frontiers that have not yet been solidly extended – but could and perhaps should be. It provides some enthusiastic mentions of building work that is in progress. It is not a complete map, but it does seek to draw attention to some particularly fascinating frontiers related to measurement design and analysis, to child poverty, and to gendered measures. Yet other active research areas – such as the relationship of measures to economic theory, poverty and inequality, or the poverty-environment frontier – are omitted.

Documenting the ends of sidewalks while many building teams are busily at work is no easy exercise, and, like an early version of Google maps, it is terribly error-prone. There will be advances already published, or already in process, that are unwittingly omitted, and for which this author, despite attempting due diligence, must sincerely apologise in advance. And there will be important questions that are omitted or unimagined. Still, because this is a relatively new field, in which junior scholars as well as senior ones still can create momentous added value in the service of others, it is hoped that this mapping will encourage new and imaginative work.
1. **Multidimensional Poverty in the Sustainable Development Goals**

The literature on multidimensional poverty – which arose empirically out of the basic needs tradition in Latin America, and the social indicators and counting traditions in Europe\(^1\) – drew conceptually on Amartya Sen’s capability approach\(^2\) as well as on other approaches, from rights-based development (CONEVAL, 2010), Ubuntu (Metz and Gaïc, 2010) and catholic social teachings (Deneulin, 2018), to disadvantage (Wolff and De-Shalit, 2007) and social inclusion (Atkinson and Marlier, 2010). Large-scale empirical and participatory studies, including the prominent *Voices of the Poor* studies (Narayan et al., 2000), also propelled its rise, as assisted by a plethora of other local and national studies.

With the advent of the Sustainable Development Goals (SDGs), multidimensional poverty appears to be graduating from the margins to the mainstream. First of all, development itself is framed multidimensionally: ‘The 17 Sustainable Development Goals and 169 targets … are integrated and indivisible and balance the three dimensions of sustainable development: the economic, social and environmental.’

Turning to poverty, the second sentence of the pivotal statement of the SDGs, *Transforming Our World: The 2030 Agenda for Sustainable Development*, reads: ‘We recognise that eradicating poverty in all its forms and dimensions, including extreme poverty, is the greatest global challenge and an indispensable requirement for sustainable development.’ Here, poverty is definitively recognised as having multiple forms and dimensions, which include ‘extreme poverty’ – the words often used to describe the $1.90 a day global poverty measure – as one particularly high-profile component of a wider concept of poverty. This shift of emphasis is sustained: the phrase ‘poverty in all its forms and dimensions’ recurs seven times in this foundational document. And Goal 1 is framed as ‘Ending Poverty in All its Forms’. It has seven targets, of which precisely one (Target 1.1) is devoted to income poverty. Complementing it, the second of 169 targets (Target 1.2) is: ‘by 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions’.

The discussions leading up to the SDGs highlighted the need for improved measures of multidimensional poverty. In his December 2014 Report, the then UN Secretary General, Ban Ki Moon, wrote, ‘Poverty measures should reflect the multi-dimensional nature of poverty.’\(^4\)

---

1 See Nolan and Whelan (2011); Feres and Mancero (2001); and Alkire et al. (2015, Chapter 4), which reviews these literatures, and the references therein.
3 Preamble, *Transforming our World* (UN, 2015b).
4 UN, 2014a, Section 5.1, para 135.
And a resolution of the United Nations General Assembly on 19 December 2014 reasserted the need for multidimensional measures, saying that the General Assembly,

[underlines the need to better reflect the multidimensional nature of development and poverty, as well as the importance of developing a common understanding among Member States and other stakeholders of that multidimensionality and reflecting it in the context of the post-2015 development agenda, and in this regard invites Member States, supported by the international community, to consider developing complementary measurements, including methodologies and indicators for measuring human development, that better reflect that multidimensionality.]

(UN, 2014b, paragraph 5)

Related processes gave rise to similar emphases. For example, the Addis Ababa Action Agenda calls on the United Nations, ‘in consultation with the IFIs [international financial institutions] to develop transparent measurements of progress on sustainable development that complement GDP, building on existing initiatives. These should recognize the multi-dimensional nature of poverty’ (UN, 2015a, paragraph 119).

Given this blossoming of interest in the multidimensionality of poverty, and of improved measurement – as well as other SDG priorities, such as disaggregating statistics by gender, ‘life cycle’ stage (adults and children) and vulnerable group, implementing integrated multi-sectoral policies, and leaving no one behind – it is no surprise that there is a new influx of attention and intellectual effort to engage in improving multidimensional poverty measurement and analysis, in the service of the poorest.

This context of greater demand and attention, of course, carries with it dangers. It is not just that it can turn heads, giving rise to spurious motivations to pose for selfies on newly built (if alarmingly precarious) sidewalks. It is more that work done in haste could be of poor quality and might not stand the test of time, or that the drive to build visible results and ‘impact’ could create a confusing maze of sidewalks aimed at serving those who are out on a self-aggrandising competitive jaunt rather more than poor persons and communities. Worse, the authors of this work could aim to manipulate not empower. Still, there is real work to be done and it is this unfinished work we seek to profile.
2. **Counting-based Measurement Methodology**

Multidimensional measures take many forms. With the greatest of recognition of its shortcomings and limits, we focus here on our counting-based approaches to multidimensional poverty measurement, which reflect overlapping deprivations (the joint distribution of deprivations), and in particular on the adjusted headcount ratio (Alkire and Foster, 2011; Alkire et al., 2015).

This Alkire-Foster (AF) methodology was advocated in the report by the Commission on Global Poverty, chaired by the late and greatly missed Sir Tony Atkinson, as the appropriate methodology for a global measure of multidimensional poverty (World Bank, 2017). It has also been used since 2010 in the global Multidimensional Poverty Index (MPI) published by the United Nations Development Programme and the Oxford Poverty and Human Development Initiative (OPHI) at the University of Oxford. Of equal or greater importance to poor people themselves, a rapidly growing set of countries have adopted official national MPIs based on this methodology. This paper scopes research areas related to this particular type of measure – naturally other papers may trace research areas associated with other measurement approaches.

3. **Designing Multidimensional Measures**

The design of counting-based multidimensional measures requires the selection of various parameters. These include:

- **Dimensions**: Conceptual categories in which indicators are grouped for ease of communication, which often are equally weighted.

- **Indicators**: Empirical data reflecting each person or household’s condition in a particular area, such as the material of their flooring, their source of water, their years of schooling or their employment conditions.

- **Deprivation Cutoffs**: Standards for identifying the minimum achievements in each indicator that must be attained in order for a person or household to be considered non-deprived: for example, having completed the years of schooling that are compulsory for that cohort in that country.

- **Relative Values (weights)**: How much should a given deprivation in one indicator contribute to an overall profile of poverty, relative to the other deprivations considered?

---

5 For a review see Alkire et al. (2015, Chapters 3 and 4).

6 See MPPN website or national examples drawn from a network of over 50 participating countries.
Poverty Cutoffs: Standards showing the minimum share of weighted deprivations that a person must experience in order to be identified as poor. Note that several poverty cutoffs are normally reported.

Naturally, there are research questions related to the selection and validation of each of these parameters in multidimensional poverty measures. So we begin with some of the research questions inherent in the design and validation of empirical measures of multidimensional poverty.

Regarding dimensions, it could be useful to include an empirical review of the dimensions currently used in multidimensional poverty measures – expanding upon the recent review of dimensions (and indicators, weights, and cutoffs) present in ‘official’ national MPIs (Zavaleta, 2017 February). The purpose of this would be to observe emerging conventions, habitually ‘missing’ dimensions, and different normative justifications that are being offered. This is an area in which the capability approach – with a prolific literature on the use of participatory methods, Nussbaum’s list, and other approaches to identifying capabilities – could add a great deal.

Regarding indicators there is a need to delineate the survey questions and related indicators which, arguably, are most informative for commonly measured aspects of poverty. In contrast to monetary poverty measures, which draw upon hundreds of survey questions, an MPI relies on far fewer. Thus each indicator must provide an accurate assessment of each person or household’s condition in that area within the given time period. Identifying the ‘best’ indicators of household members’ health, for example, is exceedingly difficult. What is required in each regularly used dimension of multidimensional poverty (such as health and nutrition, education, living standards, work and activities, security, childhood and youth conditions, and social protection) is a set of clearly defined ‘candidate’ indicators and their associated survey questions, informed by the best practices of that discipline, being feasible in terms of data requirements, linked to the Sustainable Development Goals. The strengths and weaknesses of different indicator options should also be discussed.

The indicators must meet the requirements of counting-based measures, namely unit-level accuracy (for a person or household), and accuracy in reflecting deprivation across the relevant period (such as a year, in the case of annual surveys). While universal agreement is unlikely to be obtained, beginning to develop a suite of good practices would generate informative exchanges similar to debates on the optimal ways of constructing consumption surveys (Deaton and Kozel, 2005; Beegle et al., 2012), or of designing the quasi-

---

7 Alkire et al., 2015, Chapters 6 and 7, set out the current normative and empirical considerations for parameter selection and justification.

8 See for example Nussbaum (2000), Alkire (2002), Robeyns (2005), and a burgeoning literature partially cited in Alkire et al. (2015, Chapter 6).
joblessness indicator used by the EUSILC (EU-Statistics on Income and Living Conditions) (Watson et al., 2015) or the European Social Indicators (Atkinson et al., 2002). In time, this would improve the quality of the multidimensional poverty indicator base and begin to establish conventions – for example on the best ways to include ‘decent work’, or knowledge and skills, or the quality of education, into MPIs. It may also support efforts underway in some regions to agree on a core set of MPI indicators for the purposes of regional comparisons, which may be supplemented nationally (UNECE, 2017).

There are also further methodological issues, such as when and why indicator sub-indices are desirable. The Multiple Overlapping Deprivation Analysis (MODA) of Child Poverty (De Neubourg et al., 2012; Chzhen et al., 2016) always aggregates all indicators falling under one ‘dimension’ into a sub-index that reflects the union of deprivations in any component indicators. As a result, MODA can be decomposed, after identification, by dimension – but not by indicator. Dimensional monotonicity in MODA applies only to changes by ‘dimension’ – not to changes by indicator. MODA authors argue that collapsing information into a few dimensional sub-indices is essential for communicating with policy-makers. It also gives very high estimations, which are useful to MODA’s fundamental aim of dramatizing unmet child rights (Evans and Abdurazakov, 2018).

On the other hand, all national, regional, and global applications of the AF methodology to household or child poverty present and use indicator-level breakdown extensively for two reasons. First, the MPI is then sensitive to monotonicity across indicators (if a poor person who was deprived in two indicators in one dimension becomes non-deprived in one indicator, poverty decreases for the MPI, but not for MODA). Secondly, the indicator detail is vital for policy design because policies to address indicators differ. For example, a deprivation in the dimension of health in MODA for children aged 0–4 may mean that the child did not receive DPT vaccinations or that there was no skilled assistant at birth. The policy actions required to remove these deprivations are quite different. The MODA dimensional breakdown cannot show the proportion of the population who are deprived in vaccination only, or assisted delivery only, or both. So a second layer of indicator-level analysis is required to equip this measure for policy use – which adds complexity. Further probing of this purported tradeoff between ‘communication’ and ‘policy relevance’ would be welcome.

Parallel discussions regarding the benefits of a sub-index versus a larger suite of indicators emerge not only in relation to MODA, but also in other areas. For example, roof, wall, and flooring materials all indicate housing quality, but is the ‘best’ way of incorporating these three indicators to create a housing index (for parsimony of indicators) or to enter each individually (for sensitivity to improvements in any single housing material, and to guide policies)?
There are also diverging views on technical issues such as the extent to which the coefficient of variation should be used to exclude certain indicators. Also, although Alkire et al. (2015), proposed a ‘redundancy’ indicator that maps the strength of association across indicators (in Chapter 7), we have not yet proposed clear ‘standards’ by which to assess whether redundancy is excessive – standards that must take into account the marginal headcount ratios of each indicator and their difference – though it is essential to start developing reasoned conventions.

Moving to weights or deprivation values, Alkire et al. (2015) distinguish these from precision weights, but that distinction is very often overlooked or misunderstood, so a method for communicating this very important difference is required. Otherwise, the early debate about the role of relative values in multidimensional poverty measures subsided with (a) a clarification of the nature and function of relative values (which differ from the ‘precision weights’ that are applied to continuous variables – as in the Human Development Index – and create marginal rates of substitution across all indicators at different levels); and (b) the application of robustness tests that require measures that guide policy to be robust to a plausible range of weights, given that there will be an ongoing plurality of views (Alkire and Santos, 2014). However fascinating questions remain. First of all, while various robustness tests such as rank correlations or pairwise comparisons considering (analytical and/or bootstrapped) standard errors have been applied (Alkire and Santos, 2014), there is no ‘standard’ has yet been set for what is ‘robust enough’. Nor is it easy or clear how to set such a standard (it would need to reflect the number of comparisons being undertaken and perhaps their population shares). Furthermore, when there are few subnational groups or regions that can be compared, or in situations of considerable equality across regions taken as a whole, such tests may be of limited value. In these cases, it may be useful to assess the robustness of the ranking by considering the ranking of all respondents in the data set (replicating sampling weights), though a clear methodology for doing so is not yet available (see Alkire and Fang, 2018).

A related observation, raised in fact by the Atkinson Commission Report (World Bank, 2017), is that when a ranking is relatively stable to a range of weights but the component indicators are not redundant or highly associated (which is also a requirement discussed below), then it would be interesting to map the shape of weights across the entire range. At some point, if there is low redundancy and each single indicator in turn obtains a weight of 100%, the rankings will greatly diverge. But where are the points of inflection, and what can be learned across many data sets? Linking such empirical work to the setting of robustness standards would be extraordinarily useful.

---

9 Ravallion (2011).
10 See Alkire et al. (2015, Chapter 7) for tests of redundancy.
Also, to date, relative values or weights have been normative, with some explorations into statistical weights (Pasha, 2017), and into weights based on responses to community-based surveys or surveys on topics such as such as socially perceived necessities (Statistics South Africa, 2016). Still, it would be desirable to continue to explore whether a survey module (or other intervention such as a discrete choice experiment) could accurately obtain relevant information on the permanent value of not experiencing a deprivation, and thus be used to justify or question weights proposed by other methods. Also, how often should such a module be used, given that updating weights on every survey makes changes in measured poverty difficult to interpret. It would likewise be interesting to explore ‘scientific’ weights – for example by exploring analogies to the Framingham risk score, or developing a multidimensional measure of nutritional inputs combining caloric intake, protein, micronutrients and dietary diversity, and food security, and weighting each insofar as it is assessed by nutrition experts to contribute to overall nutrition.

With ordinal data, the dichotomisation of each indicator means that technically every measure is robust to monotonic changes in the underlying data and deprivation and poverty cutoffs – hence AF measures respect the property of ‘ordinality’ (Alkire and Foster, 2016). However, it is not possible then to distinguish between ‘deeper’ levels of deprivation in the same multidimensional poverty measure. For this reason two vectors of deprivation cutoffs are often applied, as in the case of ‘destitution’ (Alkire and Seth, 2016). Sometimes there could be a requirement to include some ‘depth’ information in a single poverty measure. Technically, it is possible, for example, to include undernutrition and severe undernutrition as two separate indicators (or, if monetary poverty measures are included in the MPI, to include an extreme poverty and a basic needs poverty indicator that use different income poverty lines, or adequate sanitation vs flush toilet). This is rather difficult to justify, because the relative weights applied to different indicators drawn from the same variable, in effect, ‘cardinalise’ the difference between deprivations. Such cardinalisation is not easy to justify. But it is not completely different from the exercise of setting relative values across deprivations drawn from different data points. Hence this would be an interesting topic for further self-critical exploration.

Also, given the interest in not dropping information unnecessarily, it could be interesting to construct ‘hybrid’ measures in which ordinal data are dichotomised but cardinal data reflect their ‘normalised shortfall’ from a deprivation cutoff. The challenge of doing so is that ordinal measures will always have the full normalised gap of 1, whereas the cardinal measures’ gap will change. This creates problems in comparisons, because the relative weights between ordinal and cardinal variables change as the profiles of deprivation change. The algorithm required to adjust relative weights to this environment has not yet been developed.
The cross-dimensional poverty cutoff used to identify who is poor is a critical component in fixing the reported incidence of multidimensional poverty. To date, empirical poverty orderings across subnational groups are routinely tested for robustness considering standard errors, and often seem relatively robust to changes in the poverty cutoff across a range of values. This finding requires further probing and – similar to the issue of weighting – the setting of robustness standards, and/or of alternative robustness tests when few comparisons are available.

It is also interesting that the poverty cutoff serves multiple purposes at the same time. First, if data are inaccurate – so a person ‘appears’ to have a deprivation (like lacking schooling), but in fact either there is a data error or their qualification is not recognised (for example, they were home schooled), then if a non-union identification is not used, the censoring process will rightly disqualify a ‘false positive’ indicator from identifying poverty. Similarly, if individual preferences are such that an apparent deprivation exists but is not a real deprivation (for example, a hard-packed dirt floor in a leading indigenous household that has an eco-lodge), or a real deprivation is deliberately endured (for example, low body mass, for fashion reasons), these will likewise be censored so the final poverty measure will be relieved of spurious deprivations (though if single deprivations are real, they will be overlooked). There also may be contextual features – for example, in a particular region, water from an unprotected spring may in fact be safe although the conventions would code it differently. Furthermore, substantively the poverty cutoff enables poverty measures and the policies they inform to focus limited resources on the multiply deprived. In practice, however, the function the poverty cutoff is playing, and the relative share of censored cases that belong to each ‘type’ is not at all clear. If we are to follow the advice of the Atkinson Commission (World Bank, 2017) and document ‘total error’ for multidimensional poverty measures as well as monetary ones (Groves and Lyberg, 2010), then it would be very useful if qualitative work, as well as the revisiting of sampled households within a short period to test the stability of responses, could be used to document the magnitude of different purported reasons that ‘inaccurate’ deprivations are visible in data sets. Needless to say, adjusting surveys to clear false positives would enable the deprivation scores of the whole population to be used for additional studies such as multidimensional inequality.

In terms of properties, clearly it would be desirable to characterise the AF class of measures. In addition, there are different ways of constructing the property of dimensional breakdown (Alkire and Foster, 2016), and it would be interesting to consider this further and to explore stricter and more general ways of constructing it.

As mentioned above, dichotomisation was used in order to permit measures to draw rigorously upon binary and ordinal data. However, dichotomisation is recognisably severe in not permitting any recognition of different ‘depths’ of deprivation unless alternative deprivation cutoffs are applied. Are there any
alternative empirical strategies that still respect the features of ordinal data but make better use of the ordered responses in ordinal data that might be explored?

4. Analysis

The SDGs have changed the landscape of poverty and wellbeing analysis irrevocably, but methods have not kept up. Whereas the SDGs focus on multiple dimensions, recognise interconnections, and advocate integrated policies, tools of analysis still use rather crude approaches to controlling endogeneity. Yet even without a methodological revolution, the analysis of multidimensional poverty also would benefit from a series of research inputs. Using micro data from the same household survey, and controlling for endogeneity, one can regress different potential determinants on the poverty status or the deprivation score of the person or household. But what are the most interesting hypotheses that can be tested by such work? Do empirical regularities emerge, in terms of the determinants of the MPI, when micro-level analyses are performed across a large number of data sets? Or does analysis vary by country and even sub-population? This area is, as yet, very much a work in progress.

Following the methodologies used to analyse income poverty, one can also explore measured or regression-based growth elasticities (Santos et al., 2016), as well as the determinants of the level or trend of multidimensional poverty in terms of public expenditure, infrastructure, institutions, governance, income per capita, and so on. Multi-level models can also be applied (Jindra and Vaz, 2018). Such analyses can be done, with suitable caution, across multiple subnational units or countries. Preliminary work suggests that the growth elasticity of multidimensional poverty is lower than that of income poverty, but much more work is to be done. The data available for the global MPI make large-scale empirical analyses not only possible, but also certain to generate a new empirical body of evidence.

A related but methodologically very different question concerns the mismatches between who is identified as poor according to monetary measures and who is identified as poor if multidimensional measures are applied. When both measures are constructed from the same data sets (whether cross-sectional or panel), and the results compared, the mismatches have tended to be consistently high. For example, in Bhutan in 2012, whereas 12% of people were income-poor and 12.6% were multidimensionally poor, only 3.2% were poor by both measures (RGOB-NSB, 2012). Further (and slightly counter-intuitively), the mismatches have tended to increase rather than match more perfectly when more ‘extreme’ poverty cutoffs are applied to both measures. The percentage of people or households who are poor according to the MPI and whose monetary poverty levels fall in the higher and highest quintiles is also normally non-

---

negligible, and patterns differ across household size, with MPI values being less sensitive to household size than monetary poverty values if rates are reported in per capita terms. The mismatch between consumption poverty and income poverty measures has received considerable attention (Meyer and Sullivan, 2008). Furthermore, a literature in Europe explores mismatches extensively, as does another on child poverty (Roelen et al., 2012). Further attention is now required to understand the mismatches between monetary and multidimensional poverty in developing countries. In particular, Q-squared information, in the form of follow-up visits to households whose poverty statuses are mismatched, is required to provide definitive insights into what share of mismatches is due to intra-household inequalities and what share to structural causes, for example. It would also be useful for mapping ‘total error’ as commended (for monetary poverty) in the Atkinson Commission report.\(^\text{12}\)

Naturally a very pertinent policy question for multidimensional poverty, as for monetary poverty, is whether it is possible to predict or simulate the rates of poverty reduction accurately under different policy scenarios and different budget allocation scenarios. Relatedly, the extension of public expenditure models, benefit incidence, and other tools to adjust allocations according to the levels of multidimensional poverty are also required.

Another growing area of policy application involves the use of the MPI for individual, household, or geographic targeting, and the accuracy and quantity of type I and II errors (Azevedo and Robles, 2013). It is claimed by the governments that have officially switched to the MPI that this targeting is more precise than proxy means targeting (PMT), but two issues bear reflection: first, if the observed mismatch of MPI and monetary poverty is real (and not due, as it might be, to the volatility of consumption data), then two different targeting measures are required; if the mismatch is spurious the MPI may target the monetary poor. Second, PMT is often used to target households in need of non-monetary benefits; here an MPI may be demonstrably more accurate, particularly if the observed mismatch is real.

Another huge area of work is the use of the MPI for policy coordination. For example, in Colombia, McKinsey & Co. worked with the government to link the MPI and its component indicators to a set of management structures. Further documentation of existing experience is needed on MPI and management or ‘deliverology’ – in particular, how to set targets, how to join the MPI-related information platform with management practices, how to establish key performance indicators, and how to use the MPI within a larger information management system (Zavaleta and Angulo, 2017).

In terms of impact evaluation methodologies, one criticism has been that standard methods only reflect the average change – not the distribution – and further that they only reflect changes in a single variable.

---

at a time. Hence an active frontier of research is the application of standard impact evaluation methodologies to interventions with multiple intended outcomes for the same person or household (Vaz and Malaeb, 2018; Robano and Smith, 2014). In such cases, a project MPI can be constructed that contains the intended outcomes, and one can therefore assess the impact of the intervention jointly across a set of outcomes for the same person or household. It will be interesting to compare the outcome of such multidimensional impact evaluations with impact evaluations of the same intervention on each component indicator, to clarify in what circumstances there is added value in the multidimensional evaluation. Some papers underway suggest this may become an interesting field.\(^\text{13}\)

Similarly, it will be useful to assess the value-added of multidimensional analyses of natural experiments, by re-analysing existing data sets with established results, and exploring the extent to which a multidimensional framework of analysis adds value in synthesising the joint results. In such a way, the incremental extension of familiar methodologies to multidimensional metrics will be field-building exercises.

5. Gendered MPIs

To date, most MPIs have used the household as the pivotal unit for identifying who is poor. While this has great strengths in terms of feasibility and also supports caring and sharing among household members, as well as permitting informative disaggregation by female-headed household (Chant, 1997), there are also drawbacks. In particular, a household-based MPI does not probe intra-household inequities, to see whether and in what ways women are more deprived than men, or girls than boys, or toddlers than teenagers. Also, the set of indicators will need to be adjusted to reflect comparable deprivations across genders and age cohorts.

A number of individual poverty measures have been implemented using the AF methodology. Batana (2008) implemented a women’s MPI. Bhutan’s Gross National Happiness Pilot (2008) and measures (2012, 2015) are implemented at the individual level; women are significantly less happy, and their happiness is differently constituted to men’s (Ura et al., 2012, 2015). The Women’s Empowerment in Agriculture Index (WEAI) (Alkire et al., 2013) uses individual-level data, and the linked Gender Parity Index reflects inequalities across women’s and men’s deprivation scores within the same household. Alkire et al. (2014) designed and implemented an exploratory individual-level MPI for 31 European countries over six waves of data using EU-SILC data sets, which found that in no case were women significantly less poor than men, and in many cases, they were significantly poorer. Vijaya et al. (2014) and Klasen and Lahoti (2016) also develop individual measures. Together with colleagues, Pogge and Wisor (2016) explored extensively

\(^{13}\) A set of papers and presentations are under preparation; see OPHI Research Workshops (link).
the design of deeply contextual gendered poverty measures and elucidated the ways that participatory consultations can inform the design and uses of gendered measures.

With a few exceptions, gendered MPIs have tended to be designed using indicators that are present in standard survey instruments. Normative challenges remain. In order to create gendered MPIs, in which people’s poverty can be compared across gender and age or the life cycle, it is necessary, arguably, to develop ‘comparable’ definitions of capability deprivation that pertain to people in different age cohorts or different life situations.

Comparing women and men within a household, using intra-household measures of assets and income, is difficult but relatively well studied. Constructing indicators approximating decent work is more difficult, as is specifying an attainment that is comparable to decent work for those involved in unpaid domestic or caring work, or not in the labour force – but there is an active literature on these questions. Even more problematically, salient indicators of health and reproductive health differ by gender, change across the life cycle, and vary across family structures and disability status. Yet little work proposes comparable indicators of health deprivation (including reproductive health) for men and women at different ages and in different family situations. Furthermore, decision-making powers are difficult to measure, given high mismatches between men and women from the same household reporting who makes the same decision. Yet if gendered measures of poverty are to go beyond educational attainment and nutritional status, comparable deprivations in health, work, assets, autonomy and so on need to be defined, and new measurement conventions proposed.

6. Child Poverty

A similar though not identical array of issues affect the measurement of child poverty, covering boys and girls aged 0–17. Cooper and Stewart’s (2013) systematic review found that monetary factors were important but did not capture all aspects of child poverty. So how is child poverty to be measured? It is an interesting question, on which there are many views.14

UNICEF defines childhood as stretching from birth through 17 years. This presents a stark problem for those working on child poverty measures: either one needs to design several age-specific child poverty measures (for example, for children aged 0–5, 6–14, and 15–17), or a single multidimensional measure for children aged 0–17. If multiple measures are used, and it is found, for example, that 30% of children aged 0–5 are poor, but only 15% of those aged 6–14, the question is whether the definition of poverty overall

is ‘comparable’ to the extent that policy-makers can confidently shift resources to the younger children – or not.

Many studies draw on the Convention on the Rights of the Child to select and justify the indicators and cutoffs for child poverty measures insofar as data in existing surveys cover those rights. Complementing this, participatory studies such as Biggeri et al. (2006; cf Biggeri, 2007; Biggeri et al., 2011; and Callender et al., 2012) use surveys, focus group discussions and case studies with children in different age groups, to include the views of children as participants in the selection of their own capabilities. The main result of such processes is the definition of a list of capabilities which are relevant for children aged 0–17 or for specific groups. In a parallel initiative, the Young Lives study has conducted qualitative longitudinal research on 200 children in each of four different countries (see for example van der Gaag, 2016; Chuta and Morrow, 2015; Tafere, 2014). Based on these results, other studies have analysed the levels of child poverty in developing countries (Trani et al., 2013; Trani and Cannings, 2013; García and Ritterbusch, 2015).

Technically, differently structured child poverty measures are feasible.\textsuperscript{15} For example, with co-authors from the National Statistics Bureau of Bhutan, we trialled a 0–17 child poverty measure using the MICS survey in Bhutan (Alkire et al., 2016). This Q-squared study includes indicators with age-specific definitions that aim to capture deprivations for each age cohort that are comparable to the deprivations defined for other age cohorts. The dimension cognitive development, for example, was defined as cognitive stimulation for babies, educational activities for toddlers and school attendance for school-aged children. It was not possible to extend all indicators across all child cohorts using this survey, so a dimension ‘childhood conditions’ included indicators such as immunisation and undernutrition for children aged 0–5, child labour with age-specific cutoffs for those aged 6–14, and pregnancy, early child marriage and domestic violence for girls aged 15–17. The report showcased the problem of data limitations, and made explicit the uncertainties and child development questions which must be addressed in order to design such measures so that they are useful for policy. Now many countries include in their national MPI report a chapter on child poverty, in which at a minimum age disaggregation is undertaken and, more recently, child deprivations are added as a separate dimension or dimensions, and individual children’s poverty is studied by age and gender. Nevertheless, the problem of indicators remains.

The issues are complex and will not be easily solved. A particularly tricky issue is this: because child development stages vary by individual child as well as by cultural and social factors, measured deprivations may or may not be accurate for each particular child. So interdisciplinary collaboration that includes child development specialists, child psychologists, and pediatricians is necessary, as are cognitive and qualitative

\textsuperscript{15} Alkire and Roche (2012).
studies (including those that define total error), and analysis of longitudinal surveys for children, such as the Young Lives data sets.

7. Closing Remarks

As this paper has demonstrated, even for the limited topics covered here, there are many potential cutting-edge and field-building studies yet to be undertaken and published. Furthermore, a great number of the topics that could have been developed in a question-mongering paper such as this one were omitted. Other ‘ends of sidewalks’ include, perhaps most prominently, the incorporation of environmental and natural resource indicators into MPIs or their joint analyses using spatial techniques (Thiry and Alkire, 2018). The development of domain-specific MPIs – such as a health MPI using EQ5D surveys, or an educational MPI that includes serious variables on the quality of education – has yet to be completed, although an innovative work-related MPI has been built (IADB, 2017). Inequality could be measured multi-dimensionally using the vector of deprivation scores, as could related statistics such as growth in the bottom 40%; also the relationship between the MPI and income inequality is open for empirical exploration. Similar empirical studies are needed with respect to institutions, not only official and government bodies but also private sector and civil society institutions. Both of the latter are using the MPI and require bespoke studies of their impact and their cost-effectiveness in reducing MPI indicator scores, and as well as of how the MPI is best applied in their institutional setting. One of the next big steps forward will be experimentation in the merging of administrative and survey data sources in countries where this can be performed with high accuracy. This raises the question of how accurately ‘big data’ sources can be merged with household surveys. There are notable missing or contested conceptual links to utility and preferences – which are essential if in turn an MPI is to be linked to welfare economics and thus take a central place in economic theory. There are maps to be overlaid and analysed, and structural equation and latent variable models to be explored for MPI analysis (Di Tommaso, 2007). While a number of MPIs are constructed from census data and analysed for policy at the country level, academic research on census-based MPIs is lacking. There is also a significant research agenda using panel data sets – both descriptively to assess the volatility/stability of indicators (and verify which are stock vs flow), and analytically to probe the drivers and dynamics of MPI reduction. Thus, while the field of multidimensional poverty measurement seems to be gaining some ground, these topics, singly and together, reflect the ‘chains’ against which those seeking to use multidimensional poverty measures with the same dexterity as other indicators, chafe. It is hoped that the next wave of academic research will build many ‘sidewalks to somewhere’ and create new ways by which the overlapping deprivations that batter poor people’s lives in so many serious ways may be sharply overcome.
References


Zavaleta, D. (2017, February). ‘What are the dimensions and indicators most commonly used by countries in their national MPIs?’ *Dimensions*, 2: 13-17.