

CHAPTER SEVEN DIRECTIONS FOR FUTURE RESEARCH

THIS RESEARCH PROJECT HAS PRODUCED A MULTIDIMENSIONAL, INDIVIDUAL LEVEL, GENDER SENSITIVE MEASURE OF DEPRIVATION THAT IS GROUNDED IN THE STATED VIEWS OF POOR MEN AND WOMEN. AFTER FOUR YEARS OF RESEARCH, DISCUSSION AND ANALYSIS, AND WITH THE INVOLVEMENT OF THOUSANDS OF PARTICIPANTS ACROSS 18 SITES IN SIX COUNTRIES, WE HAVE COME A CONSIDERABLE WAY TOWARDS PRODUCING A MEASURE THAT CAN BE DEPLOYED TO HELP MEASURE PROGRESS IN THE STRUGGLE AGAINST POVERTY AND GENDER INEQUITY.

Nonetheless, work remains to refine and develop the IDM to the point where it can be easily integrated into national and international systems of social valuation. Future trials in different countries and locations, and continued work with both men and women living in poverty and experts in survey design and measurement will continue to improve the measure. In this chapter we discuss several areas in which further research should be conducted.

Refining dimensions and indicators

The dimensions included in the IDM were developed through participatory processes, academic research into the extensive literature on poverty and poverty measurement and extended dialogue with substantive subject specialists and survey experts. In particular, the participatory research in the first phase—including focus group discussions and individual interviews—helped generate definitions of our candidate dimensions. The second phase of participatory research helped select among those candidate dimensions a final list for inclusion in the IDM. The list of included dimensions and the definition of these dimensions should be further

refined through both participatory and non-participatory methods. For example, the dimension of time-use might best be specified as regarding required hours of labour, actual hours of labour, difficulty of labour, amount of leisure time, quality of leisure time, amount of discretionary time, etc. Each of these specifications gives rise to different priorities in reducing deprivations in the time-use dimension, and might be best captured by different indicator(s) than the one we selected.

The indicators that populate our measure were selected predominantly through an examination of existing methods of information collection, especially multi-topic and single topic surveys, and the secondary literature on their reliability, validity, and cross-cultural applicability. We do not claim that we have identified the best possible indicators for each dimension. There are several ways in which indicators might be improved. First, different indicators may be selected which better cover the dimension in question. For example, in the dimension of food, we currently have a single indicator of hunger during the preceding four weeks. A recent study by USAID's FANTA project⁹² identified the three questions that we used to generate this indicator as the most reliable across different contexts for measuring household hunger. We were not able to include other indicators to track other important deprivations in the area of food, given our commitment to develop a survey that is manageable in terms of technical requirements, costs and time. For example, an indicator of a person's body mass might better reflect long term food deprivation. Or an indicator of micro-nutrients in the blood might better reflect the dietary diversity and nutritional intake a person faces. Similarly, our current indicator regarding time-use simply reflects a person's labour burden in the preceding day. This is because considerable recall bias is introduced when asking about longer periods of time. This indicator therefore may not reflect when a person faces excessive labour burden if they are surveyed on a day that follows a public holiday or a day-off. It also will not capture the seasonal nature of many people's work schedules.

Several dimensions and indicators that were tested in the third phase trial gave us reason to believe that further investigation and refinement is required.

In using a crude asset index to measure financial status for the household, we quickly encounter two problems.

The first is that the indicator is attributed to all household members, even though members may have very different levels of consumption. The second is that countable assets may accumulate over time within a household, and may become cheaper due to lower manufacturing costs, but neither of these changes will necessarily reflect a change in day-to-day consumption of important goods, including food and other necessities. Because we aimed for a short survey, we did not pursue more standard

92. www.fantaproject.org

consumption-expenditure or income measurement at the individual level. Future efforts may advance the indicator(s) for a person's financial status.

In the measurement of education, we attempted to reverse the long-standing practice of measuring enrollment to the exclusion of achievement.

We asked participants to read, write, and perform simple arithmetic. Future iterations may do better to have a longer section of educational testing to reveal greater differences in educational achievement.

In the measurement of access to family planning, we currently treat male and female access equally.

That is, men and women are both considered to be deprived if they do not have access to modern methods of contraception or are restricted in their ability to space pregnancies. In the Philippines this had the odd result of showing men to be more deprived than women. This seems twice mistaken. First, women may have greater access to contraception quite simply because they work harder to procure it. It is odd to mark men as more deprived when they make no such efforts. Second, and more importantly, it is arguably a much greater deprivation for women than men to be without contraception and to face the burdens of unwanted pregnancy. However, the alternative options here are not particularly appealing. One may exclude this indicator for men. But then one must either simply have fewer indicators for men, or find some 'substitute' indicator for men which is specific to men and not women. Doing so also misleadingly suggests that men may not genuinely need, and be deprived of, access to contraception.

A similar problem in measuring the gendered distribution of deprivation arose in measuring deprivation in freedom from violence.

Men's and women's experience of violence counts equally assuming a similar incident of violence occurred. For example, if a man is hit in a public fight, this scores the same as if a woman is hit by her partner. Arguably, in some instances the woman's deprivation is more severe because her exposure to violence is in the home (with implications for her ability to avoid it), because it is more likely to occur again (and fear of this may be ever present), and because it is likely to affect many other aspects of her life. If the man's public fight is not likely to have these similar features (possible future occurrence, affecting many aspects of life) perhaps it should not count equally. Again, the alternatives to this scheme are not particularly attractive. We have deliberately excluded from the questions any information about the location of the violence or the perpetrator, so that respondents who were also perpetrators of violence against other household members would be aware that answers to the survey could not be used to identify them. Changing the questions to identify the location of the violence, in order to attach greater weight to violence in the home, would

increase the risks of responding to the survey given our sampling method seeks data from all adult household members. Alternatively, we could specify some greater weight for women's exposure to violence than men's, regardless of where that violence occurs. But certainly many instances of men being exposed to violence are severe, have enduring multidimensional impacts, may not be readily avoidable, and are equally deserving of a claim to public resources directed towards prevention and mitigation.

Every dimension that requires subjective assessments of objective situations requires further exploration.

Asking individuals to rate their water quantity, or their control over decision-making in the household, permits subjective assessments to influence objective deprivation scores. Individual respondents may understate or overstate the extent of their deprivation, and this may occur systematically on a group basis. For example, women might systematically understate how much control over decision-making they should have, and thus how much they do have. Or geographically remote communities may understate how much water they need to meet their needs, leading to biased assessments of how much they actually use.

Measuring the health status of individuals is difficult, and self-reports of morbidity are not sufficiently reliable.

However, merely reporting on access to health care fails to take account of disproportionate exposure to unhealthy environments. Future iterations of the IDM should adjust the measurement of health status.

We used access to adequate health care during pregnancy as a substitute indicator for women's access to health care.

This decision was intended to be sensitive to unique health needs before and after pregnancy. However, it had the impact of attributing better access to health care to women than had we used information on treatment during a recent illness. This is an odd result, and future iterations of the IDM should consider whether it is valuable to continue taking account of access to health care during pregnancy. It may be that the perceived risks associated with pregnancy encourage women and their families to prioritize maternal care for women to a greater extent than they do other health care. Focusing on this as an indicator of health access for women provides important information about a key event in the lives of many women that involves significant risks; but this comes at the expense of masking ongoing inequity in access to general health care.

Measuring environmental problems by simply counting them may introduce understatement or overstatement of environmental deprivation.

A person who faces a single environmental problem of very severe air pollution may be worse off than a person who suffers several more moderate environmental problems, such as exposure to traffic, pools of water where mosquitoes breed, and other noise during the day. In addition to refining the dimensions and indicators, there is room to refine the scoring system that is applied to each indicator. For example, for the dimension of time use, we identify as deprived a person who works (in both paid and unpaid labour) more than 10 hours in the preceding day. To our knowledge or to the knowledge of experts on time-use we consulted, there are no international standards on the labour burden that individuals ought to face. Because we are the first to introduce scalar assessment within dimensions for the purpose of measuring multidimensional poverty, we recognise that further deliberation and investigation will improve the cut-offs used within each dimension. Similarly, more work needs to be done to test the thresholds that separate different categories of poverty in the composite measure against the perceptions of poor women and men in a variety of contexts and adjust the specifications as needed.

Despite these outstanding challenges, the relevant comparison for the selection of dimensions, indicators, and interval scores is with existing multidimensional measures. On this comparison, the IDM is vastly superior to existing alternatives on offer.

Weighting

Our current weighting scheme is informed by three broad commitments. First, we are broadly prioritarian. More severe deprivations are morally worse than less severe deprivations. The absence of a person's third meal of the day should count for much less than the absence of a second. Second, deprivations in some dimensions of life are more important than other dimensions for a person's physical survival. Food deprivation counts for more than deprivation in the ability to participate in one's community. Third, weighting schemes should, at least in part, reflect the collective preferences of the individuals that they measure. From these commitments, we generate a weighting scheme that assigns different values to each of the intervals on our five point scale within a dimension and assigns different weights to three categories of dimensions.

With more time and resources, we would have explored in greater depth the possibility of generating weights directly from participatory exercises. Several different methods exist in the literature for generating participatory weights. Participants may be asked to make comparisons

between sets of pairs of situations in which an individual's status in various dimensions is described. The participant may simply state a preference between the two or state the degree to which they prefer one to the other. From a series of these answers participatory weights can be derived. Participants may also vote on the importance of various dimensions, or increments within those dimensions. Or participants may engage in a budget allocation process, through which they express their willingness to pay for goods in various dimensions.⁹³ Each method can be used to generate dimension weights in composite indicators.

One common shortcoming of most efforts at generating participatory weights is that they apply to dimensions as a whole, rather than to increments within a particular dimension. The weight that a person assigns to food may depend on how much food the person has at the moment, or the increment of food (say, the 2nd meal) they have in mind when assigning weight to the dimension. Future research should seek to develop methods of generating participatory weights for each increment within a dimension.

In addition to generating participatory weights, other methods are available for devising weighting systems.

Data-driven methods require various forms of multi-variate analysis to set weights. Such methods (including factor analysis and principal components analysis) investigate the correlation between the variables in a measure and can be used to set weights based upon the degree to which a particular variable is correlated with other variables in the measure. We do not endorse this method for the following reason. The mere fact that a variable is not correlated with other measured variables is not a sufficient reason to discount its importance. For example, in Klasen (2000), principal component analysis (PCA) revealed violence to be very weakly associated with other dimensions of deprivation in South Africa. Using PCA to set weights would require heavily discounting the importance of violence.⁹⁴

The most sophisticated weighting system would be sensitive to a range of factors, including:

1. Either the individual's or a relevant group's preferences regarding incremental changes in her current status

93. For reviews of various weighting schemes available in multidimensional measurement see Decancq, K & Lugo, A.M. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*. 32(1): 7-34. Also see: Organisation for Economic Cooperation and Development (OECD), European Commission Joint Research Centre. (2008). *Handbook on constructing composite indicators: methodology and user guide*. OECD: Paris.

94. Stephan Klasen notes that there are significant drawbacks to using PCA to set a weighting scheme. "The disadvantage of such an approach is that it implicitly assumes that only components with strong correlations with each other are relevant for the deprivation measure which may be debatable in some cases. For example, the fact that perceptions of safety are not closely correlated with the deprivation index and its other components should not necessarily suggest that safety is not an important indicator of deprivation, as would implicitly be assumed by a principal component analysis." Klasen, S. (2000). Measuring poverty and deprivation in South Africa. *Review of income and wealth*. 46(1): 39.

2. Her position in all other dimensions
3. The relative importance of marginal differences in her current position in a given dimension
4. The interaction of that marginal difference with other possible incremental gains or losses in other dimensions
5. The particular social, environmental, political and economic context in which she finds herself.

We suspect that any system that can actually be implemented in the real world will fall far short of this ideal. However, it is our hope that future research can improve systems of weighting to take account of some of the factors that appear, at least to us, should be taken into account in producing a composite metric of deprivation.

Children

Thus far our proposed measure of deprivation applies only to adults. Children were not included in any of the three phases. Given the ethical concerns that arise from working with children, and the need for distinct methods to involve children in participatory research, we decided from the outset to exclude children under the age of 16.

Among members of our project, there are two divergent views about how one might proceed from the work we have completed to a measure of deprivation that can be applied to children. On one view, important as a process of public reasoning is for setting a system of social valuation, it is not particularly problematic to exclude children from questions of poverty measurement. While there are many things that can be learned from working with children living in poverty about how that poverty may best be measured and combated, it is not necessary to have children living in poverty help construct new measures of deprivation. According to proponents of this view, expanding the Individual Deprivation Measure to children would require a two-step, non-participatory process. First, at the level of dimension selection, one would need to identify those dimensions that are not relevant for children in the current measure and remove them. For example, it may not be necessary to measure a child's ability to participate in community decision-making, as this is not traditionally thought to be a sphere of activity in which children need to participate. Many of the dimensions that adult participants identified as relevant to determining whether a life is free from poverty and hardship, including the importance of control over decision-making in the household and the community, and status in paid and unpaid work,⁹⁵ are entirely or largely irrelevant for children, at least as currently formulated. Second, at the level of indicator selection,

revisions would need to be made to better capture childhood deprivations. For example, in the space of health, it might be more appropriate to discover whether a child had received immunisations, rather than to evaluate the kind and quality of treatment she received when last sick. In the space of education, rather than measuring the final grade completed, it would be better to measure whether the child is enrolled in school, and to test her literacy and numeracy skills against age-appropriate standards.

On an alternate (and perhaps more widely held) view, just as adults should be involved in constructing a publicly justifiable measure of deprivation for adults, so too should children be involved in the construction of child poverty measures. Although different participatory methods might be needed, and heightened ethical scrutiny would be required to ensure the protection of participating children, children can and should be involved in identifying how childhood poverty should be conceived and measured. The IDM should therefore not be applied to children. Rather, a separate measure of childhood poverty should be developed through a participatory process involving, perhaps exclusively, children.

Household-based measures of deprivation provide an easy solution to the problem of measuring child poverty. They attribute the (usually financial) poverty of the household to all its members. This is mistaken for obvious reasons. Children in poor households may be better off or worse off than their parents. Effective programs targeting children may reduce their deprivations, in education, health care, or nutrition, for example, without making this progress for adults. In fact, many gains by children in these areas will not be reflected in most poverty measurement unless they raise the living standards of the entire household.⁹⁶

Child deprivation measures must face several additional challenges not faced by adult measures.⁹⁷ They must find a way to select indicators that retain some degree of comparability as children age. The indicators needed for measuring deprivation among children under five will be quite different than those for children who are entering their teenage years. Ideally, measures of child poverty would also allow some degree of comparability between

95. Whereas adults ought to be treated with respect in paid and unpaid work, a child that is free from poverty would be involved in no or minimal levels of paid or unpaid work.

96. Income poverty lines generally treat children as poor if the household counts as poor. Similarly, multidimensional poverty measures generally treat children as poor if the household counts as poor. This has the effect of making invisible the impacts of any anti-poverty programs directed at children which do not raise the living standards of the household more generally. For example, improved quality and access to education, nutritional programs targeting children, or immunisation programs would have no impact on reducing poverty according to most measures even though they clearly reduce child poverty.

97. On extant child poverty measures see Roelen, K. & Gassmann, F., (2008). Measuring child poverty and well-being: A literature review. Maastricht Graduate School of Governance. Working Paper Series 2008/WP001. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1105652. Also see: Gordon, D., Nandy, S., Pantazis, P., Pemberton, S. & Townsend, P. (2003). Child poverty in the developing world. The Policy Press: Bristol, UK.

children and adults, so that programs seeking to reduce poverty amongst both groups are judged by equivalent standards. They also must find a way of collecting the data from individuals who may not be fully capable of responding to survey questions. Much of the information about young children in existing multi-topic surveys is provided by parents. Depending on the nature of the question, parents may have reason to hide or deceive sensitive information about the child. For example, parents may not respond honestly if the child has been a victim of physical or sexual abuse, especially if that abuse was perpetrated by a family member. Similarly, a family may be hesitant to identify deprivations for which they might fear they will be held responsible, such as a lack of schooling. These are challenges that must be faced as child poverty measurement is integrated into pro-poor, individual-based measures of deprivation.

Context sensitivity

Finally, as discussed previously, there is a tension between developing a measure of deprivation that is comparable across context and over time, and having a measure be sensitive to the specific context in which poverty is being measured. The need for context-sensitivity arises at three stages in the process of generating the IDM. First, in existing survey questions, some questions and candidate answers need to be tailored to a particular context. For example, in assessing whether housing materials are natural, rudimentary, or finished, the selected answers must be revised in distinct natural contexts as the materials from which individuals construct their homes change a great deal. Second, indicator selection sometimes varies from case to case across countries. For example, indicators on shelter should take account of housing materials in Angola, but more appropriately take account of access to heat in Russia, due to very different environmental contexts. Third, it may be that new dimensions and indicators should be added in some contexts. For example, in many countries Female Genital Mutilation or Cutting is not prevalent, and therefore not relevant to assessing individual deprivation. But in other countries, these practices may be prevalent, and warrant a distinct dimension in a gender-sensitive deprivation measure.

The IDM can form the common core of a system of multidimensional measurement, but additional dimensions and corresponding indicators can be added in different contexts when needed. This would allow for both the measurement of the IDM (globally comparable), and a national IDM, relevant only within the country in question. Further research is needed in diverse national and subnational contexts if the IDM is to provide the groundwork for both internationally comparable but locally usable measurement of poverty and gender equity.

Conclusion

It has been a great privilege to join with thousands of participants in Angola, Fiji, Indonesia, Malawi, Mozambique, and the Philippines to develop a new measure of deprivation that is genuinely gender sensitive and, in our view, a considerable improvement upon existing methods of poverty measurement and a necessary complement to existing measures of development and progress. Much more work remains to be done. In addition to further research to address the questions identified above, policy makers must provide the resources to refine and adopt new measures to guide policy making and poverty eradication in the decades to come. The measurement of poverty and gender disparity is a necessary component of any successful program to eradicate poverty. We hope the efforts of this project go some way toward improving measurement to the benefit of poor men and women. Members of this project will continue to work on the measurement of poverty, gender equity, and social progress, and hope that the findings included in this report will inform ongoing efforts at local, national, and international levels to improve the measurement of poverty and gender disparity.

SEE UPDATES