

Developing Morally Plausible Indices of Poverty and Gender Equity: a Research Program^{*}

Thomas Pogge

ABSTRACT: *Various indices are used to track poverty, development, and gender equity at the population level. Some of them – the UNDP’s Human and Gender-Related Development Indices, and the World Bank’s Poverty Index associated with the first Millennium Development Goal – have become highly influential. This paper argues that these prominent indices are deeply flawed and therefore distort our moral judgments and misguide resource allocations by governments, international agencies, and NGOs. Examination of these flaws reveals useful pointers toward developing better indices – though much interdisciplinary work will be needed before sound and practicable indices are actually available.*

Why should the discussion of indices – instruments for measuring and tracking phenomena in the world – take up space in a philosophy journal? Beginning to answer this question, let me distinguish three types of indices, or rather three distinct ways in which an index might be understood by its creators and users.

An index might be understood, firstly, as merely delivering convenient aggregates of information. Some baseball statistics are commonly so understood. For example, a batting average aggregates information about what happens during games. It is not understood to be the “right” way of aggregating such information, just one way that has become prominent through use. When so conceived, indices do not pose philosophically interesting problems.

^{*} Many thanks to Kieran Donaghue, Alison Jaggar, Justine Kolata, Matt Peterson, and Elizabeth Weissberg for their good comments and suggestions.

An index might be understood, secondly, as serving an explanatory and/or predictive function. An index of risk factors for heart attacks is an example. Here the index is designed and redesigned – by adding and deleting components perhaps or by changing the formula used for aggregating information from the various components – toward maximizing its predictive power. This is an intellectually more ambitious enterprise, but an empirical, not a philosophical one: one needs to explore what the relevant risk factors are and how they interact with one another in order to fine-tune the index.

An index may also be understood, finally, as a standard of evaluation, typically used in a normative role. To be sure, the preceding case can also be described along these lines: a doctor evaluates a patient's lifestyle and exhorts him to avoid cigarettes or certain foods. In that case, however, there is a clearly defined end – avoiding a heart attack – whose desirability is unproblematic. Matters are different with an index of gender equity, for instance, because here much conceptual and normative work is needed to specify a conception of gender equity that captures what is desirable about it and can serve as the basis for identifying and measuring it in real-life settings. This is philosophical work: furnishing a general basis for ranking as better or worse alternative states of a population as well as alternative changes in such states and for moral judgments and prescriptions based upon such rankings.

Insofar as morally significant indices have actually been constructed and applied, this work has differed from typical philosophical work in two respects. Generally ignored by professional moral philosophers, index construction has received uneven philosophical attention mostly from non-philosophers, including distinguished ones like the economists Joan Robinson and Amartya Sen. And, unlike most philosophical works, some indices have become highly influential: followed and cited by the media, policy makers, and the general public. This is an unfortunate combination – unfortunate for those who suffer the effects of deeply flawed indices and unfortunate also for academic philosophy, which is marginalizing itself and does not love wisdom well enough to understand what matters.

The present papers mark the beginning of a major research effort – funded by the Australian Research Council, Oxfam, and the International Women's Development Agency, among others – aimed at reflecting critically on how gender equity, development, and poverty have been – and on how they ought to be – measured and tracked across a plurality of diverse natural and social environments. This project is important insofar as such indices guide – and often misguide – the efforts of policy makers in governments, NGOs, and international agencies. Taking such indices seriously, political actors will use them as proxies for the values they purport to measure: governments will aim to use the means at their disposal to effect the largest achievable reduction in the poverty index, for example, on the assumption

that they will thereby have brought about – or be perceived to have brought about – the largest achievable reduction in what makes poverty objectionable. Insofar as this assumption is mistaken, money and human effort go to waste.

This danger is real. This is so, first, because index construction has thus far received too little systematic moral attention and reflection. A second reason is that indices tend to have great inertia: once an index is well-introduced and widely known, people tend to go on using it, despite known flaws, because it allows comparisons with the past that would become impossible or at least very difficult if the index were to be replaced or substantially revised. The third and perhaps most important reason is especially relevant to prominent indices used for evaluation or prescription. Because the outputs such indices deliver can have serious political repercussions – for instance by serving as ammunition in debates about WTO globalization – the construction, refinement and application of such indices can be subject to intense lobbying efforts. This is well known on the domestic level but less widely recognized with respect to transnational indices. When governments and corporations have billions at stake in their efforts to shape and empower the emerging global trading regime, then they will take an interest in how the performance of this regime is being monitored. To be successful in their rule-shaping efforts, politicians and corporations need to attract political support and to disarm political opposition. Their ability to do this, in turn, depends in part on whether citizens and other politicians, at home and abroad, believe that WTO globalization “works,” for instance by reducing poverty or by fostering development and gender equity.

The next two sections illustrate these points by discussing one of the most influential indices: the World Bank’s poverty index (section 3) and the goal of “halving world poverty by 2015” which is formulated and tracked in terms of this index (section 2). I then discuss the indices of Human Development and Gender-Related Development as maintained by the United Nations Development Program (section 4).¹ The final section outlines some ideas toward better indices of poverty, development, and gender equity.

2. The Halving of World Poverty by 2015

Our current understanding of world poverty and of its evolution over time is heavily focused on counting the number of poor people. This counting is based on a narrow conception of “absolute” or “extreme” poverty, closely tied to hunger: people are counted as poor only if they fall short of “the income or

¹ Relevant to this exercise, but omitted here for reasons of space, are a detailed critique of the UNDP’s Human Poverty Index and Gender Empowerment Measure, which I hope to present elsewhere.

expenditure level below which a minimum, nutritionally adequate diet plus essential non-food requirements are not affordable.”²

For the last 20 years or so, the World Bank has taken charge of this counting exercise. It employs for this purpose an international poverty line (IPL) widely known as “a dollar a day.” The Bank had initially defined this line in terms of the purchasing power that \$1.02 had in the US in 1985, but has since redefined it three times. The Bank’s latest definition is in terms of the purchasing power that \$1.25 had in the US in 2005. Once the IPL is fixed, poverty is quantified simply by the number of those living below the IPL, and a reduction of poverty as a reduction in this number.

Invoking the Bank’s tracking method and figures, the world’s governments have repeatedly promised to “halve world poverty by 2015.” One such promise was made prominently at the 1996 World Food Summit (WFS) in Rome, where 186 governments pledged themselves “to achieving food security for all and to an ongoing effort to eradicate hunger in all countries, with an immediate view to reducing the number of undernourished people to half their present level no later than 2015.”³ An “immediate” view that takes 19 years to solve just half of the world poverty problem is not exactly ambitious, to be sure – just imagine Franklin Roosevelt pledging in 1942 to cut the ongoing harm caused by the Nazis down to half by 1961!⁴ But at least the pledge seemed manageable, implying an annual reduction of 3.58% in the number of poor people.

Yet the world’s governments soon decided that they had promised more than they were willing to deliver. In the year 2000, they declared it the first Millennium Development Goal (MDG-1) “to halve, by the year 2015, the proportion of the world’s people whose income is less than one dollar a day and the

² United Nations Development Program, *Human Development Report 1995* (New York: Oxford University Press, 1995), 223.

³ *Rome Declaration on World Food Security*, www.fao.org/docrep/003/w3613e/w3613e00.htm.

⁴ One might think that the pledge of the fictional Roosevelt would have been worse because the harm in question was actively caused in violation of negative duties. But this active harming was done by Germans and their allies, without US involvement. So by doing nothing about it, the US would have violated no negative duties. The US and other affluent states *are* actively involved in designing and upholding global institutional arrangements that foreseeably aggravate poverty worldwide – see Thomas Pogge, *World Poverty and Human Rights*, second edition (Cambridge: Polity Press 2008). If there is an asymmetry, it goes the other way. Other things being equal, it is worse actively to contribute to harms that could be avoided at low cost than passively to allow harms that could be averted only at high cost.

The actual Roosevelt was committed not only to defeating the Nazi menace, but also to ending extreme poverty worldwide. In his State of the Union Address of 6 January 1941, he set forth as a guiding principle for a post-fascist world – to be attained “in our own time and generation” – “freedom from want, which, translated into world terms, means economic understandings which will secure to every nation a healthy peacetime life for its inhabitants – everywhere in the world.”

proportion of people who suffer from hunger.”⁵ This formulation retains the idea of halving the problem by 2015, and yet cleverly diminishes the ambition by focusing not on the *number* but on the *proportion* of poor and hungry people. This proportion is a fraction consisting of the number of poor people in the numerator and world population in the denominator. With world population expected to increase by 2015 to about 120% of what it was in 2000,⁶ a reduction in the number of poor to 60% of what it was in 2000 suffices to cut the proportion in half. By redefining what is to be halved, the world’s governments raised the acceptable number of poor people in 2015 from 836 million (1996 benchmark minus 50%) to 1,004 million (2000 benchmark minus 40%), according to the current World Bank statistics.⁷ They also lowered the promised annual reduction in the number of poor people from 3.58% to 3.35%.

In highlighting this revision, I do not care whether governments focus on the number or the proportion of poor people. My concern is with the dilution of the 2015 goal and with the effort to obscure this dilution. The dilution can be expressed in either idiom: the number of poor is to be reduced by 50% according to the WFS and by only 40% according to MDG-1. Or: the proportion of poor is to be reduced by 58.33% according to the WFS and by only 50% according to MDG-1. Either formulation makes apparent that the goalposts were moved to add 168 million to the number of those whose extreme poverty in 2015 will be deemed morally acceptable – an extra 168 million human beings for whom “a minimum, nutritionally adequate diet plus essential non-food requirements” will be out of reach. This dilution was successfully obscured from the public, and kept out of the media, by opaquely switching from “number” to “proportion” while retaining the language of halving poverty by 2015.

Since its celebrated adoption by the UN General Assembly, the poverty promise has undergone further dramatic dilution through two important revisions. The current UN interpretation and tracking of MDG-1 expresses the poor not as “a proportion of the world’s people,” but as a “proportion of people in the developing world,”⁸ thus taking advantage of a faster-growing denominator. It also back-dates the

⁵ United Nations General Assembly, *United Nations Millennium Declaration* (A/res/55/2, 8 September 2000), article 19(1); available at www.un.org/millennium/declaration/ares552e.htm.

⁶ UN Population Division, *World Population Prospects: The 2006 Revision* (2007), available at <http://esa.un.org/unpp>.

⁷ Shaohua Chen and Martin Ravallion, “The Developing World is Poorer than We Thought, but no Less Successful in the Fight against Poverty,” World Bank Policy Research Working Paper WPS 4703 (2008), 34; available at econ.worldbank.org/docsearch.

⁸ United Nations, *The Millennium Development Goals Report 2008* (New York: UNO, 2008), 6-7; available at www.un.org/millenniumgoals.

baseline: envisioning that the halving should take place “between 1990 and 2015”⁹ rather than between 2000 and 2015. Because the population of the developing countries in 2015 is expected to be about 146% of what it was in 1990,¹⁰ a reduction in the number of poor and in the number of hungry to 73% of what it was in 1990 is now deemed sufficient to “halve poverty by 2015.” Shrinking the reduction to be achieved and also lengthening to 25 years the period in which to achieve it, this revision of MDG-1 lowers the promised annual reduction in these numbers to 1.25%. This revision also raises the accepted number of poor people in 2015 to 1,327 million (1990 benchmark minus 27%), according to the current World Bank statistics¹¹ – thereby adding yet 314 million more to the number of human beings whose extreme poverty in 2015 will be deemed morally acceptable.

One remarkable consequence of the UN’s back-dating of the baseline is that China’s massive poverty reduction in the 1990s – the number of Chinese living in extreme poverty reportedly declined by 265 million during that decade¹² – can now be counted as progress toward MDG-1. The revision of MDG-1 thus led UN Secretary-General Kofi Annan tragically to report to the General Assembly that for the world’s most populous region – East Asia and the Pacific – the 2015 poverty target was met already in 1999, a full year before this goal had even been adopted!¹³

	Baseline Year	Baseline number of poor (millions)	Promised Reduction by 2015	Target for 2015 (millions)	Required Annual Rate of Reduction	Target for 2005 (millions)
World Food Summit	1996	1672	50%	836	3.58%	1204
MDG-1 as adopted	2000	1673	40%	1004	3.35%	1411
MDG-1 as revised	1990	1817.5	27%	1327	1.25%	1505

A Promise Diluted

The nearby table summarizes what the dilution means for the goal of poverty eradication: the number of extremely poor people that the world’s governments deem acceptable in the year 2015 has been

⁹ Ibid.

¹⁰ UN Population Division, *World Population Prospects*.

¹¹ Chen and Ravallion, “The Developing World is Poorer than We Thought,” 34.

¹² Ibid.

¹³ United Nations, *Implementation of the United Nations Millennium Declaration: Report of the Secretary-General* (A/57/270, 31 July 2002), 8 and 22; available at www.undemocracy.com/A-57-270.pdf.

revised upward by 491 million. The reduction in the number of extremely poor people that the world's governments had promised in Rome has correspondingly been cut back by 491 million: once 836 million, the reduction promised for 1996-2015 now stands at 345 million.

There is little doubt that the UN will be able to announce, in 2015, that the target of halving world poverty has been achieved. But this success will depend decisively on having replaced the initial WFS understanding of this target as requiring a 50% reduction in the number of poor from the 1996 level with a new understanding of this goal as requiring merely a 27% reduction in the number of poor from the much higher 1990 level.

The story of how the world's governments are managing the "halving of world poverty by 2015" reflects the political forces that shape how prominent normative indices are designed and applied. The next section on how world poverty is quantified provides further illustration.

3. Tracking Poverty as the Number or Proportion of People Living below some IPL

3.1 The first and most obvious problem with tracking poverty by means of a headcount figure is that it indefensibly prioritizes people living just below the IPL. Moving such people to just above the IPL is the cheapest and easiest way of reducing the poverty headcount. But why should only moves crossing the IPL count as progress? There is no good reason for ignoring changes in people's economic situation when these occur entirely below the IPL or slightly above it.

3.2 A second, related problem is that, when poverty is tracked as a headcount, our picture of the distribution and evolution of poverty may be excessively sensitive to the level at which the IPL is set. The problem is easily illustrated because the World Bank researchers leading the poverty count provide headcounts not only for their own chosen IPL of \$1.25 PPP 2005 (daily consumption), but also for three other poverty lines likewise denominated in 2005 US-dollars.¹⁴ The nearby table displays, for each of these four poverty lines (leftmost column) the headcount change between various base years (top row) and the year 2005.¹⁵ The highlighted column shows how the actual change in the number of poor

¹⁴ Chen and Ravallion, "The Developing World is Poorer than We Thought," 34-35. 'PPP' stands for 'purchasing power parity,' and '\$1.25 PPP 2005' for the purchasing power that \$1.25 had in the US in 2005.

¹⁵ A minus sign in a cell indicates that the poverty headcount has fallen in the column period relative to the row's poverty line.

people during 1990-2005 compares to the 17.2% reduction in this number that would put the world “on track” for achieving the 27% reduction that the diluted MDG-1 promises for 1990-2015.

Period Poverty Line	1981- 2005	1984- 2005	1987- 2005	1990- 2005	Relative to path of diluted MDG-1	1993- 2005	1996- 2005	1999- 2005
\$1.00 PPP 2005	-43%	-35%	-28%	-33%	89% ahead	-29%	-22%	-23%
\$1.25 PPP 2005	-27%	-23%	-19%	-23%	34% ahead	-22%	-16%	-17%
\$2.00 PPP 2005	2%	-1%	-2%	-6%	67% behind	-8%	-7%	-10%
\$2.50 PPP 2005	15%	9%	6%	2%	112% behind	-1%	-3%	-5%

How Success against Poverty Depends of the Level of the IPL

The table shows that the reportable headcount progress depends heavily on the IPL level chosen. No matter what time period one may examine, the change in the poverty count during this period looks better the lower the IPL is set. Choosing the UN’s newly favored 1990 benchmark, for instance, poverty has in the 1990-2005 period fallen by nearly 33% relative to the \$1 PPP 2005 poverty line – far ahead of the 17.2% reduction needed to put us on track for meeting the poverty target of MDG-1. In the very same period poverty has *increased* by 2% relative to the \$2.50 PPP 2005 poverty line – moving us *away* from the MDG-1 target number. If one of the two lower poverty lines is chosen, the world is well on schedule for realizing the 27% headcount reduction required to achieve the diluted MDG-1. If one of the two higher poverty lines were chosen, we would likely get nowhere near this 27% reduction by 2015.

How then does the World Bank decide the level of its IPL? The Bank has defended all four of its successive IPLs as “anchored to what ‘poverty’ means in the poorest countries.”¹⁶ How do they know what “poverty” means in the poorest countries? The Bank infers its answer from one single set of information: the official domestic poverty lines used in poor countries. The details of this “anchoring” have varied, however. In its first exercise around 1990, the Bank chose \$1.02 PPP 1985 per day as its IPL, saying that the domestic poverty lines of eight poor countries – converted at 1985 PPPs – were close to this amount.¹⁷ By 1994 this IPL had been “rounded off” to a more memorable \$1.00 PPP 1985.¹⁸ In 2000, the Bank revised its IPL again, to \$1.08 PPP 1993, saying that this was the median of the ten

¹⁶ Chen and Ravallion, “The Developing World is Poorer than We Thought,” 9.

¹⁷ Martin Ravallion, Gaurav Datt and Dominique Van De Walle, “Quantifying Absolute Poverty in the Developing World” *Review of Income and Wealth* 37 (1991): 345-61 at 348-49. The definition chosen was actually \$31 PPP 1985 per person per month.

¹⁸ Shaohua Chen and Martin Ravallion, “How Did the World’s Poorest Fare in the 1990s?” *Review of Income and Wealth* 47 (2001): 283-300 at 285 n. 7.

lowest domestic poverty lines as converted at 1993 PPPs.¹⁹ For its most recent exercise, unveiled in the summer of 2008, the Bank has chosen \$1.25 PPP 2005 as the IPL, saying that this is the mean of the domestic poverty lines (converted at 2005 PPPs) of the 15 poorest countries.²⁰ It is unclear why political decisions made by rulers or bureaucrats in a few poor countries²¹ should be a reliable indicator of what “poverty” means to poor people all over the world. Nor can the Bank’s involvement in setting domestic poverty lines for many poor countries increase confidence in the exercise: mere consistency of the Bank’s judgments is no substitute for justification of them.

The Bank’s successive substitutions have resulted in a tightening of the poverty criterion in most countries and a broadening in a few.²² The case of the US is not atypical. If we use the US consumer price index²³ to convert the Bank’s four successive IPLs into 2005 dollars, we get:

\$1.02 PPP 1985 = \$1.85 (2005)

\$1.00 PPP 1985 = \$1.81 (2005)

\$1.08 PPP 1993 = \$1.45 (2005)

\$1.25 PPP 2005 = \$1.25 (2005)

The last and now exclusively used IPL comes to roughly \$10 per person per week in 2009 – or \$41 per month or \$500 per year.²⁴ A US resident would count as poor by the Bank’s new standard only if the cost of her entire 2009 consumption of goods and services was under \$500.

A US resident evidently cannot purchase even her most minimal requirements locally for \$500 per year.²⁵ It may appear that this is no problem because dollars buy more where very poor people actually live. But this thought cannot save the Bank’s chosen IPL, because the Bank converts all local currency

¹⁹ Ibid., 285; actually using \$32.74 PPP 1993 per person per month.

²⁰ Chen and Ravallion, “The Developing World is Poorer than We Thought,” 10.

²¹ Of the 15 poorest countries just mentioned, nine have very small populations and 13 are located in sub-Saharan Africa (ibid.).

²² Sanjay Reddy and Thomas Pogge, “How Not to Count the Poor,” in Sudhir Anand, Paul Segal, and Joseph Stiglitz, eds.: *Debates in the Measurement of Global Poverty* (Oxford: Oxford University Press, 2009), Table 1; also available at www.socialanalysis.org.

²³ See www.bls.gov/cpi.

²⁴ Ibid.

²⁵ Even \$2.50 PPP 2005 per day, or roughly \$1000 per year in 2009, is still a pitifully small amount of money – far less than a person requires to meet her basic needs in the US. See www.cnpp.usda.gov/USDAFoodCost-Home.htm for cost estimates in the US for a nutritionally adequate diet for short-term or emergency use.

amounts not at market exchange rates but at PPPs. And the Bank cannot say that, while \$500 per annum is too little to live on in the US, the PPP equivalent of \$500 in less developed countries is adequate. Making this claim, the Bank would be undermining the very methodology it is using, which assumes that PPPs provide suitable conversion rates for comparing incomes and consumption expenditures of poor people worldwide and for fixing the level of the IPL.

The Bank also cannot very well say that its low line delivers, as we have seen, much prettier trend figures and shows us on track toward achieving the poverty target of MDG-1. This fact cannot *justify* the choice of a low line.

There is one more possible justification for the very low IPL chosen by the Bank and employed by the UN for tracking MDG-1. This justification invokes the claim that the magnitude of the world poverty problem, as measured relative to some IPL, should be reasonably related to the world's resources: we should not define poverty so broadly that its eradication becomes wholly impractical. Doing so might even be counterproductive by discouraging efforts to reduce poverty to the extent that we can.

Poverty line expressed in US-dollars PPP 2005 per person per day	Poor people in 2005		Aggregate shortfall from the poverty line in percent of the global product	
	Number in billions	Average shortfall from the poverty line	at PPPs	at market exchange rates
\$1.25	1.40	28%	0.33%	0.15%
\$2.00	2.58	40%	1.3%	0.6%
\$2.50	3.14	45%	2.2%	1.1%

How the Size of the Global Poverty Gap Varies with the Level of the IPL

The nearby table²⁶ shows that the worry is overblown. Using the IPL favored by the Bank and the UN, we find that the aggregate shortfall of the 1.4 billion people living below this IPL amounts to merely 0.15% of the global product – or about \$70 billion per annum or one-ninth of US military spending.²⁷ It is this poverty problem that the world's privileged see themselves as committed to reduce at the stately pace of 1.25% per annum.

Were we to double the Bank's IPL, both the number of poor and their average shortfall from the poverty line would greatly increase. A seven times larger 1.1% change in the distribution of the global product

²⁶ Figures based on Chen and Ravallion, "The Developing World is Poorer than We Thought," 23 and 32-36.

²⁷ See [www.cdi.org/PDFs/What is the Defense Budget.pdf](http://www.cdi.org/PDFs/What%20is%20the%20Defense%20Budget.pdf).

would then be required to eradicate poverty so defined. Yet, even this shift would hardly be discouragingly large – not when one considers that poverty causes some 18 million premature deaths annually,²⁸ thus killing at over twice the rate of World War II during its worst years. Winning that war cost the lives of 23 million allied soldiers as well as half or more of the gross domestic products of the US, UK, and USSR during the war years.²⁹ Ending world poverty would require no young people to be sent into battle. And it would – even on as broad a definition as \$2.50 PPP 2005 – not noticeably affect the economies of the affluent countries.

3.3 A third problem with the Bank’s counting of the poor arises from the fact that the exclusive focus on monitoring income and consumption expenditure leaves out factors that are relevant to poverty as intuitively understood. Someone who survives by working in a cold climate, which requires additional expenses for staying warm, is poorer than another who can earn the same income without having to bear such extra costs. Likewise for someone who faces additional expenses due to local disease vectors or other challenges of the natural or social environment.

Another crucial factor is labor effort: someone who reaches the IPL by working 70 hours a week is poorer than another who can maintain the same consumption expenditure by working only 20 hours per week. Any plausible measure of poverty must take account of necessary labor time – including here not merely paid work, but also the work on family food production, housekeeping, and child rearing, which is typically unpaid and very predominantly performed by women and girls.

3.4 A fourth problem arises from the Bank’s reliance on household surveys that seek to value, in local currency, the totality of a household’s consumption or income. This amount is then simply divided by the number of household members to determine the members’ economic position. This simple division ignores variations in course-of-life needs: as persons’ needs vary with age, so do the requirements of poverty avoidance. The Bank should not then ignore a household’s age composition when assessing the adequacy of its consumption expenditure. The simple division also ignores how consumption and labor are distributed within each household, and thereby systematically overlooks, in particular, gender biases in the distribution of work, food, education, and health care. A closer look would often reveal what the

²⁸ World Health Organisation, *The World Health Report 2004* (Geneva: WHO, 2004), 120-25; available at www.who.int/whr/2004/en/.

²⁹ Mark Harrison, “Resource Mobilization for World War II: the U.S.A., U.K., U.S.S.R., and Germany, 1938-1945,” *Economic History Review* 41 (1988): 171-192, at 184. See also Mark Harrison, ed., *The Economics of World War II: Six Great Powers in International Comparison* (Cambridge: Cambridge University Press, 2000), 287 et passim.

Bank's method renders a priori impossible: that some – typically female – household members are poorer than others.³⁰

3.5 A fifth problem arises from how the results of household consumption surveys, expressed in the currencies of many countries and years, are compared against the IPL. How does the Bank determine whether the consumption of an Indian household in 1987, say, valued at some amount of 1987 Rupees per person per day, is above or below \$1.25 PPP 2005?

The Bank makes such determinations through two transformations. It uses PPPs of the base year to convert its chosen IPL into base-year amounts of all other currencies (including 2005 Rupees). It then uses national consumer price indices (CPIs) to convert the results further into national currency amounts of other years (including 1987 Rupees).

This conversion method is unreliable on account of the CPI and PPP conversions it involves. Price comparisons across years (within the same country) and across countries (within the same year) yield different results for diverse commodities: the price of computers may be falling even while the price of rice is rising, and dollars in the US may buy more bread but less maid services than yuans do in China. CPIs and PPPs aggregate such diverse price data, weighting the prices of the various commodities according to their role in national (CPI) and international (PPP) consumption expenditure. This fact renders CPIs and PPPs unreliable within a poverty measurement exercise, because national and international consumption patterns do not come close to reflecting the consumption needs of the very poor. The prices of basic foodstuffs, for example, have a significance in the lives of poor people that is vastly greater than their influence on CPI and PPP calculations reflects. The doubling of these prices during 2006-08 made much less of a difference to national CPIs, and hence to the official poverty statistics, than to the poor who typically must spend well over half their income on basic foodstuffs. Conversely, prices of cars or services play a much larger role in CPI and PPP calculations than in the lives of poor people. There is no assurance, then, that plausible poverty comparisons across space and time can be generated by using CPIs and PPPs in order to map any amount, denominated in any currency of any year, onto the common currency of the IPL.

To illustrate, imagine a simple world with three commodities: *necessities*, *discretionaries*, and *services* (always in this order). If their prices do not move in lockstep, the CPI will reflect a weighted average of

³⁰ See Susan M. Okin, "Poverty, Well-Being, and Gender: What Counts, Who's Heard?" *Philosophy and Public Affairs* 31 (2003): 280-316 at 284-86, 305.

their price movements, based on the national spending pattern. By relying on the CPI, the World Bank thus loses track of the price of necessities. Falling prices of discretionary (e.g., consumer electronics) may lead to a falling CPI even while rising biofuel demand is driving up food prices. Poor people on constant incomes become poorer relative to what they need to buy (necessities), yet by the Bank's calculations they become richer (relative to what their country's population is consuming).

Similar errors are introduced through the use of PPPs, as can be shown with a numerical example. Suppose the prices of our three commodities, in pesos, are 5, 6 and 1 in some poor country and \$3, \$4 and \$9 in the US. What is the PPP? The answer depends on the spending patterns in *both* countries. Suppose this pattern is 30%/50%/20% in the poor country and 10%/50%/40% in the US. This yields a bilateral PPP of 1.55 – each peso is deemed to be equivalent to \$1.55. (Here the hypothetical peso is valued so highly in dollars because it buys nine times more services and because services constitute so large a portion – 40% – of US consumption expenditure.) But this bilateral PPP is inadequate for assessing the incomes of the very poor, who do and must concentrate their consumption expenditure on necessities – relative to which each peso is worth only \$0.60 (5 pesos buy as much necessities in the foreign country as \$3 buy in the US.) Again, many who are extremely poor, relative to what they really need to buy, may not show up in the Bank's poverty statistics.

This distortion is compounded once third countries enter the picture. Bilateral PPPs calculated without regard to other countries would not satisfy transitivity, that is, the condition:

$$\text{for all countries A, B, C: } \mathbf{PPP(A,B) \times PPP(B,C) = PPP(A,C)}$$

Transitivity fails to hold because the left side of the equation is substantially influenced by the spending pattern in country B, while the right side is not so influenced at all. It is highly desirable, however, that PPPs be transitive because the Bank's poverty measurement exercise would otherwise not be robust with respect to the choice of base country (the relation between the domestic poverty lines of any two countries would change depending on which currency they are converted into and compared in). To ensure transitivity, the calculation of PPPs conventionally involves a final step that adjusts all preliminary bilateral PPPs to one another in a way that guarantees transitivity. As a result, the PPP assigned to any local currency is affected by the prices and spending patterns not only of its home country and the United States (base country), but also by those of every other country. In the Bank's method, then, the classification of any person as poor or non-poor is influenced not merely by the money she has and the

prices she faces, but also by the prices and spending patterns of all countries included in the PPP exercise.

The dependence of PPP calculations on the spending patterns of all other countries and on the prices of commodities that are irrelevant to poverty avoidance distorts the World Bank's poverty assessment exercise. This distortion could be reduced by leaving affluent countries out of account entirely. When counting the poor the Bank is looking only at less developed countries – on the plausible assumption that virtually no one in the affluent countries lives below \$1.25 PPP 2005. But then the Bank includes the affluent countries in its calculation of PPPs, thereby allowing the prices of necessities to be drowned out by the prices of discretionaries and services, which play a much larger role in the consumption of affluent than in that of poor countries. The Bank would do better if it excluded affluent countries also from its PPP calculations, using the currency of some developing country rather than the US dollar as the common measure – though this revision would only reduce, not eliminate, the error.

The fifth problem with the Bank's method is then that it is hugely over-inclusive in its informational base. To determine whether a person is poor or not, the Bank rightly takes account of the local prices of necessities: nutrition, shelter, water, clothing, utilities, and basic medical care. But it marginalizes this information by also incorporating much additional information that is irrelevant: the local prices of other commodities, the prices of all commodities in other countries, and national spending patterns in all countries. By making all this information relevant through its CPI and PPP conversions, the Bank distorts our picture of world poverty in unknown ways.

3.6 Perhaps the best evidence one can have against any method is that its applications can deliver massively divergent results. The two notions of equivalence invoked in CPI and PPP calculations rely on very different (national and global) spending patterns. As a consequence, the comparison of two amounts in different years and countries varies with the base year chosen for PPP conversion. For example, one can use the CPIs of the two countries to convert into 1993 amounts and then compare via 1993 PPPs – or one can use these national CPIs to convert into any other year and then do the comparison in PPPs of that year. One can get as many different results as there are available PPP base years.

An intuitive way of displaying the problem involves taking a national currency amount – denominated in 1993 Chinese yuan, say – on a “round trip”: using 1993 PPPs to convert it into Bangladeshi taka, then the Bangladeshi CPI to convert into 1985 taka, then 1985 PPPs to convert into 1985 yuan, and then finally

the Chinese CPI to convert it back into 1993 yuan. The final amount resulting from such a round trip is typically quite different from the initial amount, in this case 31% higher.³¹ This is so because the CPI and PPP conversions performed involve very different notions of purchasing power equivalence, based on different consumption patterns. And the comparison of a 1993 yuan amount and a 1985 taka amount will then vary depending on whether this comparison is made via 1993 PPPs and the Bangladeshi CPI or via 1985 PPPs and the Chinese CPI (or in some other way). Using 1993 rather than 1985 as the base year raises all Chinese amounts — prices, incomes, consumption expenditures — in all years by 31% relative to all Bangladeshi amounts in all years. And conversely, using 1985 rather than 1993 as the base year raises all Bangladeshi amounts in all years by 31% relative to all Chinese amounts in all years.

The choice of base year thus has a large impact on the relative position of national poverty lines — often much greater than 31%.³² Bound to be arbitrary, the Bank's choice of base year affects the classification — as poor or non-poor — of hundreds of millions of people, though some of these effects do cancel out, of course, as poverty is shifted from one country or continent to another.

3.7 Intuitively, whether people are poor and how poor they are depends on whether and to what extent they can, at reasonable cost in necessary labor, gain access to the goods and services they need to meet their basic human needs within their actual natural and social environment. To assess their ability to do this, one needs to know their age-specific human needs, the relevant environmental factors (such as climate and disease vectors) that co-determine what nutrients and other necessities persons require to satisfy their needs, the cost of the locally cheapest way of meeting these requirements in a culturally acceptable way, and the amount of labor necessary to gain access to these required goods and services.

The World Bank's poverty monitoring method bases its binary output (poor/non-poor) on the per capita cost of household consumption. It completely ignores household composition (age and gender), intra-household distribution, necessary labor, culture and other environmental factors. And it marginalizes the local prices of necessities by swamping them with an enormous wealth of irrelevant data: about prices of commodities that are irrelevant to poverty avoidance, about the cost of all commodities in all other countries and other years, and about the proportions in which all national populations all over the world allocate their spending over commodities. There are strong reasons to doubt that this method can

³¹ Reddy and Pogge, "How Not to Count the Poor," Table 1.

³² Ibid.

reliably identify the poor, assess how poor they are, or reveal whether world poverty is in decline and, if so, at what rate.³³

Given the great political importance that tracking the evolution of poverty, and the possible feminization of poverty, have assumed in the wake of MDG-1 and in the heated debates about WTO globalization, it is high time to specify and try out an alternative approach that would draw on the right information base and would combine this information in a plausible way. The rich data produced by such an approach would not merely give us a more comprehensive picture of world poverty and its evolution, but would also enrich our empirical understanding of poverty, thus enabling more effective poverty avoidance.

4. Tracking Development through the HDI and Gender Equity through the GDI

The UNDP's Human Development Index (HDI) is meant to measure the degree of development a population enjoys. It is calculated as a weighted average of four components, each normalized to a scale of 0-1. These components are the national population's *life-expectancy at birth* (L), its *adult literacy rate* (A), its *primary, secondary and tertiary school enrollment rate* (E), and its *gross domestic product (GDP) per capita* (P). Each country (C) receives a score in each of these components, and its overall HDI score is then the weighted average of these four scores, with the weights being 1/3, 2/9, 1/9, and 1/3, respectively.³⁴ In a simple equation:

$$HDI_C = (3L_C + 2A_C + E_C + 3P_C)/9$$

The Gender Development Index (GDI) uses the same four components weighted in the same way. The difference is that these four scores are now calculated in a more complicated way. First, the population is divided into its male and female subpopulations. Each subpopulation is then scored just like the entire population would be scored for HDI purposes. The GDI then uses, in each component, the harmonic

³³ One further empirical reason for questioning the Bank's optimism about the decline in poverty derives from the evolution of the number of chronically undernourished people as tracked by the UN Food and Agriculture Organization. This number – stated as “more than 800 million” in the *Rome Declaration* – has recently been estimated to have been 963 million in 2008 (www.fao.org/news/story/en/item/8836/icode/) – even *before* the current economic crisis. The number of those who experience hunger intermittently is substantially higher. The reduction of hunger is the third target within MDG-1.

³⁴ All indices used by the UNDP are explained in a Technical Note in United Nations Development Program, *Human Development Report 2007/2008* (Houndmills: Palgrave Macmillan, 2007), 355-61; also available at hdr.undp.org/en/media/HDR_20072008_Tech_Note_1.pdf.

mean of the male and female scores. When the male and female scores in some component are equal, then the harmonic mean is simply this equal score. The more unequal they are, the more the harmonic mean sinks below the average, reflecting a penalty for gender inequity.

To illustrate, suppose adult literacy in some country is 50% for women and 75% for men. The HDI adult literacy score is then simply the weighted average of the two gendered scores:

$$A_C = PS_F \times A_{CF} + PS_M \times A_{CM}$$

(where PS_F and PS_M denote respectively the female and male population share). A_C is unaffected by how literacy is distributed. If each gender constitutes half the population, $A_C = 0.5 \times 0.5 + 0.5 \times 0.75 = 0.625$.

For the GDI, the adult literacy score is calculated in a gender-sensitive way: by aggregating not through the weighted average but through the weighted harmonic mean defined as

$$A_{CG} = 1 / (PS_F / A_{CF} + PS_M / A_{CM})$$

In our example, $A_{CG} = 1 / (0.5 / 0.5 + 0.5 / 0.75) = 0.600$. The GDI thus is similar to the HDI except that the component scores get marked down a bit for gender inequity – in our example from 0.625 to 0.600. Let us call this difference – a country's HDI score minus its GDI score – the gender inequity penalty.

The HDI/GDI method usefully complements income with other factors, though it aggregates them in a crude way (through normalization of the index components and the weight then assigned to them) that is not backed by any reflective justification. For example, the indices imply that certain trade-offs between income and life-expectancy are neutral with regard to development, but these implications are not explored and examined for their plausibility. In addition, the HDI and GDI are gravely defective in various ways.³⁵

4.1 The first and most general problem is that both indices are focused on countries and also marketed as standards of achievement for countries. This is a step backward from the World Bank's poverty headcount measure which is focused mainly on the *global* poverty count (though headcounts disaggregated by country and region are also provided). This focus on the global poverty count strongly suggests that what matters is the impoverishment of *individual human beings* and that the responsibility of avoiding such poverty is shared worldwide. The HDI and GDI, by contrast, focus on the ranking of

³⁵ Some of the following criticisms, and others, are fruitfully discussed in a special issue of the *Journal of Human Development* 7 (2006): 145-290, edited by Stephan Klasen.

countries and thereby sustain the common view – well-liked in affluent countries – that poor countries themselves bear primary responsibility for their underdevelopment.

4.2 The shift in focus away from individuals is most obvious in the first and last components which, together, constitute two thirds of both indices. Life expectancy is simply averaged across a national population, or across its male and female subpopulations, and all information about its distribution within such (sub)populations is discarded as irrelevant. This is implausible, because inequalities in life expectancy – not merely between men and women, but generally – must be considered relevant for any morally plausible conception of development. Thus an increase in a society's life expectancy cannot be celebrated as development when it derives from a gain to the rich beyond 80 that slightly outweighs losses to the poor below 50.³⁶

Analogous thoughts apply also to the last component – more so, because inequalities related to class, race, or religion tend to be much greater in income than in life expectancy. The HDI uses gross domestic product per capita, which averages over an entire country, while the GDI uses average earned income per capita computed separately for the male and female subpopulations. In both cases, internal inequalities are disregarded completely. Thus, the HDI is indifferent between a society's income being heavily concentrated at the top so that the poor majority lives in life-threatening poverty and this income being widely dispersed so that no one is compelled to endure severe poverty. A given increase in per capita GDP is considered by the HDI to make the same contribution to development, regardless of whether the vast majority of the population is sharing in this increase or is left behind.

Similarly, what matters for the GDI is the average earned income of females relative to the average earned income of males. It makes no difference to gender-related development, as the GDI conceives it, whether overall gender income parity exists because a small minority of women at the top have enormous pay packages or because women in all strata of society enjoy employment opportunities and pay packages that roughly match those of their male counterparts. This is obviously morally implausible.

³⁶ Remarkably, the GDI defines gender equity in life expectancy as attained when female life expectancy is 5 years above male. It thus marks down a society in which life expectancy is 26 for males and 30 for females on account of a gender inequity against *females*. The GDI thereby encourages such a society to shift resources from males to females in order to approach the ideal 5 year difference. Sen has not explained how this is supposed to be consistent with his capability approach which, in the name of equalizing capabilities, demands that additional resources be devoted to those who are worse endowed by nature in order to raise them to the same level of freedom to function. For some discussion of this issue, see Stephan Klasen, "UNDP's Gender-related Measures: Some Conceptual Problems and Possible Solutions" *Journal of Human Development* 7 (2006): 243-74, and A. Geske Dijkstra, "Towards a Fresh Start in Measuring Gender Equality: A Contribution to the Debate" *Journal of Human Development* 7 (2006): 275-283.

More importantly, this total neglect of distribution also provides perverse incentives. When policy makers are concerned to improve their country's HDI, they will seek to enhance its GDP without regard to whether the additional income goes to the rich or to the poor. They will emulate the example of Angola, which achieved huge gains in per capita GDP between 2001 and 2007 by expanding crude oil production, without worrying whether the extra income was earned by foreigners or embezzled by a small ruling clique. Likewise, insofar as policy makers are concerned to reduce their country's gender inequity penalty, they will seek to shift the income distribution in favor of females – but again without regard to whether the additional income goes to rich or to poor women. The GDI thus encourages such policy makers to boost their country's GDI score through policies that double to \$200,000 the incomes of 10,000 women executives if this is even slightly more convenient to them than it would be to achieve the same female-income boost by doubling to \$2,000 the incomes of 1 million female domestics. The GDI here values doubling an affluent woman's income *100 times* more highly than doubling a poor women's income – even though poor women have more at stake, in terms of both human needs and gender disadvantage.

FIRST SCENARIO				SECOND SCENARIO		
Female Average	Male Average	Female Share of Earned Income		Female Average	Male Average	Female Share of Earned Income
0	43	0.0000	Decile 1	44	43	0.5057
0	96	0.0000	Decile 2	99	96	0.5077
13	148	0.0807	Decile 3	147	148	0.4983
77	202	0.2760	Decile 4	204	202	0.5025
129	294	0.3050	Decile 5	292	294	0.4983
189	383	0.3304	Decile 6	385	383	0.5013
288	519	0.3569	Decile 7	520	519	0.5005
449	731	0.3805	Decile 8	711	731	0.4931
979	1293	0.4309	Decile 9	1192	1293	0.4797
6180	6441	0.4897	Decile 10	4710	6441	0.4224
830	1015	0.4500	AVERAGE	830	1015	0.4500

How the GDI Can Fail to Register Great Gender Inequity Affecting Most Women

The nearby table illustrates two hypothetical scenarios that attract identical gender inequity penalties from the GDI. Looking at the female-to-male earned-income ratios decile by decile, we see that the first scenario displays large disadvantages suffered by a large majority of women while the second displays much smaller disadvantages suffered by a minority of women at the top of the socioeconomic distribution. It is intuitively obvious that gender inequity is much worse in the first scenario than in the second. Nonetheless, the GDI ranks them as equals and thus guides policy makers toward the first if

achieving it is even slightly less costly for them than achieving the second. This evidently wrong conclusion arises from the GDI here giving equal weight to each unit of income rather than to each woman. It thereby implicitly assigns much greater importance to disadvantages and percentage gains in the incomes of affluent women than to disadvantages and percentage gains in the incomes of poor women.

4.3 The shift in focus away from individuals manifests itself also in how the HDI and GDI aggregate information across the four index components. They do so by first aggregating within each component indicator across the population, and then aggregating the results across these components. This renders both indices insensitive to how different aspects of disadvantage are correlated. This is problematic because disadvantages persons suffer are morally more problematic when they aggravate than when they mitigate one another.

The GDI inherits these problems. And it adds an additional one by discarding information about whether gender inequities aggravate or mitigate one another. The GDI calculates the gender inequity penalty separately in each component and simply adds up the penalties – regardless of whether women are disadvantaged in all four components, say, or each gender is disadvantaged in two out of the four.

In order to promote development, governments, international agencies, and NGOs should, other things equal, prioritize the neediest. Yet, in order to pinpoint those who are in greatest need overall, any multidimensional index must be constructed to aggregate in the opposite order: aggregating first intra-personally across the component indicators, and then aggregating the results across the population.

5. Toward New Indices of Development, Poverty and Gender Equity

Today's leading indices of development, poverty and gender equity suffer from one or more of the following deficiencies:

- (1) lack of a moral rationale, which results in some aspects of development or gender inequity being selected in preference to others without a sound justification of the former as being of greater constitutive or instrumental importance;
- (2) false universality – a focus on deprivations (e.g. income poverty) that anyone may suffer combined with relative neglect of deprivations that differentially affect persons according to gender, age, and environment;

(3) bias towards the better-off, which manifests itself in

(3a) a focus on indicators that are relevant mostly to the more privileged, such as women in parliament³⁷ or women in higher education, which matter but are less important than gender inequities burdening much larger numbers of more disadvantaged women and girls, and in

(3b) implausible aggregation that ignores correlations among different aspects of deprivation, for instance, or compares male and female income totals and thereby implicitly gives much greater weight to income inequalities at the top; and

(4) lack of integration, as when several partial indices are presented side-by-side without any guidance for how to resolve conflicts among them about the ranking of programs and policies.

Because indices thus flawed provide faulty guidance and incentives to policy makers, it is important to explore the possibility of constructing better indices. This is the point of the research project I am commencing with a group of scholars (including our editor, Alison Jaggar) and officials from five NGOs from five countries

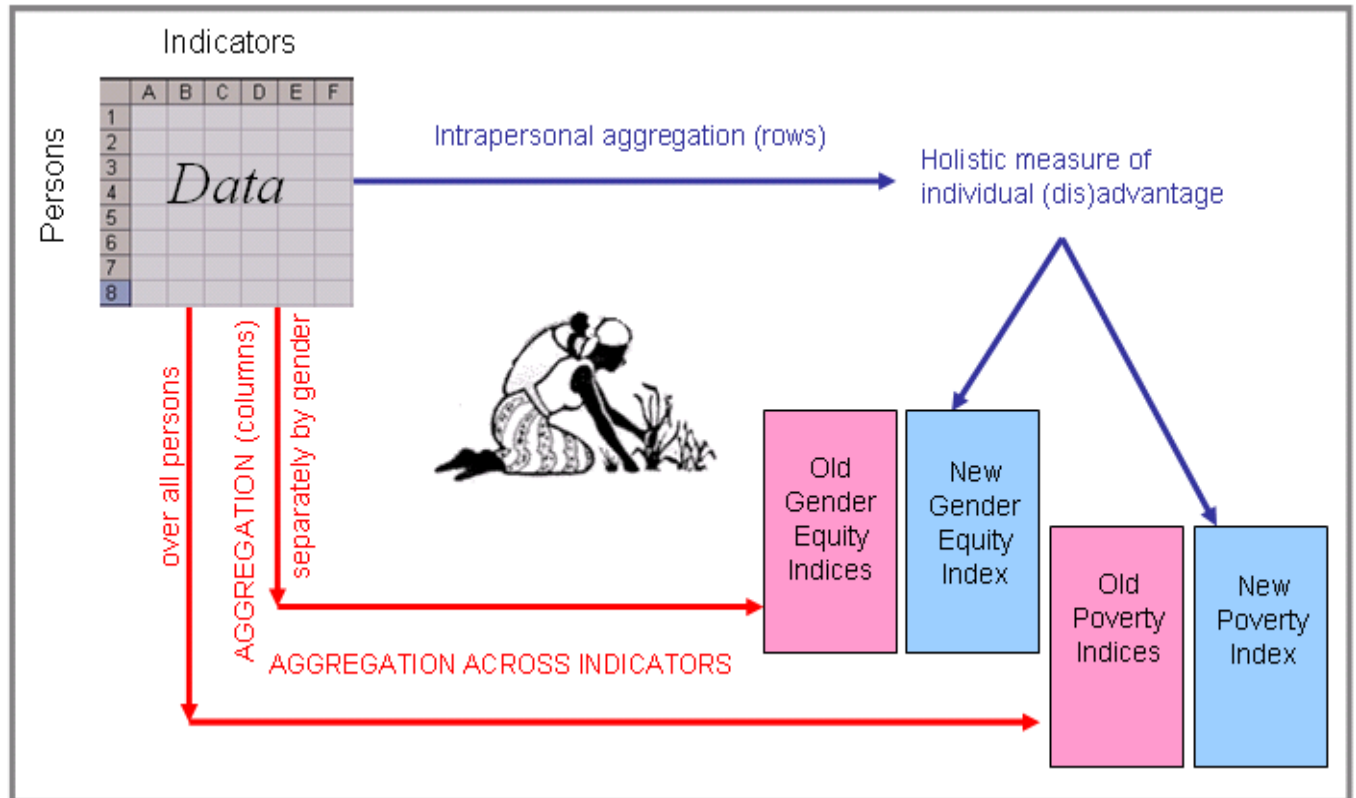
We have two broad, constructive ideas that serve as a starting point. First, handling distributional issues correctly requires a holistic measure of individual deprivation. An index is supposed to provide summary information about a group. To do so, it must track how group members are doing. This purpose is much better served when the interpersonal aggregation is performed last, after the relevant aspects of each person's situation have been holistically assessed. This is so because the significance of these situational aspects depends on the age and gender of the person and because these aspects are also interrelated in their significance. If individual lives are what ultimately matters, then we must attend to these interdependencies. For example, a credible index of development must be sensitive to whether an increase in literacy goes to landowners or the landless, an improvement in medical care goes to children or to the aged, an increase in enrolment to privileged university students or to children in slums, an increase in life expectancy to the elite or to the marginalized, enhanced physical security to males or to females.

The second idea is that a holistic measure of individual deprivation, grounded in a sound conception of basic human needs or requirements or capabilities, can serve within different aggregation exercises. Thus, the poverty of a population, or its level of development, can be defined as the mean level of

³⁷ This point alludes to the UNDP's Gender Empowerment Measure, which I could not discuss here for reasons of space.

individual deprivation. And gender inequity in a group can be defined as the mean difference across population fractiles (averaging over male/female ratios as calculated for each fractile).

Obviously, these ideas – illustrated in the nearby diagram – are rough and preliminary. But they show in general terms at least how better indices might be specified – indices that would support plausible judgments about whether, for instance, WTO globalization has fostered development or fueled the feminization of poverty.



New Indices Based on a Holistic Measure of Individual (Dis)Advantage