

# *Towards a Fresh Start in Measuring Gender Equality: A Contribution to the Debate*

A. GESKE DIJKSTRA

*Geske Dijkstra is with the Faculty of Social Sciences at Erasmus University Rotterdam*

**Abstract** Both the Gender-related Development Index (GDI) and the Gender Empowerment Measure (GEM) represent a “false start” in measuring gender equality. This is because they do not measure gender (in)equality as such, but an odd combination of absolute welfare levels and gender equality that is not easy to interpret. This note argues that the United Nations Development Programme’s Human Development Report Office should take the lead in either constructing a new index for measuring gender equality or elaborating a revised GDI and revised GEM that do measure gender equality. Detailed recommendations are given for both possibilities on how this can be done, partly on the basis of a brief review of alternatives presented in the literature.

**Key words:** Gender equality, Human development, Measurement

## **Introduction**

The Human Development Report Office of the United Nations Development Programme (UNDP) has initiated a debate on improving the existing indices for measuring gender disparities in well-being and agency: the Gender-related Development Index (GDI) and the Gender Empowerment Measure (GEM). These two indices, first published in the 1995 *Human Development Report*, have attracted wide attention among both policy-makers and academics. The UNDP is to be commended for elaborating and publishing these indices. This has stimulated the debate on gender equality and has also stimulated the collection of more and improved data. But it is also timely to discuss a revision of these two indices.

In a way, both the GDI and GEM represent a “false start” in measuring gender equality. Both indices do not measure gender (in)equality *per se*. The GDI is a measure of human development corrected for gender inequality: the higher the gender inequality in the three dimensions of the Human Development Index (HDI)—education, health and income—the greater, in principle, the difference between the HDI and GDI. The GEM is

an odd combination of relative female and male empowerment—but softened by taking a harmonic mean of the female and male scores—and absolute levels of income per capita.

In my view, there is a huge need for an index that measures the position and status of women relative to men, independently of the absolute levels of income per capita or human development.

The need for measuring gender equality as such is obvious from the various attempts that have been made to construct such an index, for example by international institutions such as Social Watch (2005) and the UN Economic Commission for Africa. But there is also an interest in the academic research community, since it is only possible to assess relationships between gender equality and, say, economic growth or overall human development if there is a measure of gender equality that is independent from absolute income levels. Many studies exploring this kind of relationships now mistakenly use the GDI or GEM (see Schüler's contribution in this *Journal of Human Development [JHD]* special issue).

In my view, the Human Development Report Office should take the lead in developing and publishing such a measure of relative gender equality. To this end, a new index could be developed, or the GDI and GEM could be revised. In this article I first explore the first option, and then the second. With respect to the second option, I will explore to what extent it is possible to deduce or create a direct measure of gender equality out of the GDI and GEM, while at the same time preserving as much as possible of the existing indices.

### **On a new measure of gender equality**

The requirements for such a new measure of gender equality are simple:

1. It should cover a limited number of indicators, but these indicators together should cover as many dimensions of gender equality as possible.
2. Data should be available for many countries.
3. It should be simple to calculate and to understand.
4. It should allow comparisons between countries but also over time.

When speaking of *gender* equality it is important to bear in mind that we are not dealing with equality *between the sexes*. As Ferber and Nelson (1993, pp.9–10) put it, “Gender is the social meaning that is given to biological differences between the sexes; it refers to social constructs rather than to biological givens”. This means that it is fully appropriate to adjust the data on the life expectancy variable to account for a “normal” (assumed biological) difference between men and women, as the UNDP has so far done when calculating the GDI. It is a fact that women live, on average, longer than men. Although this may be deemed unequal (see Klasen's contribution in this *JHD* special issue) or even unfair, it should not be regarded as “gender” unequal. Given that it is such a universal

phenomenon (except in situations resulting from an unequal or discriminatory treatment of women, such as the current situation in Africa with regard to HIV/AIDS infection rates), it can be assumed to be the result of biology, not of social relations, culture, or policies/politics. Of course, it is not fully clear what the exact biological difference is; this can only be assessed or approximated empirically.

On the *dimensions* of gender equality, the full range developed in an international workshop in The Hague is still a useful starting point. The eight dimensions are as follows: gender identity, autonomy of the body, autonomy within the household, political power, access to social resources (education and health), access to material resources (land and credit), access to employment and income (including the distribution of unpaid work), and time use (leisure and sleep) (Dijkstra, 2002). Therefore, the search for an index that measures gender equality in all its dimensions does not require to separate the well-being or human development dimensions from the agency or empowerment dimensions, as the UNDP has so far done.

Mentioning all dimensions that should be included is relatively easy, but the availability and accessibility of international *data* restricts the possibilities for their inclusion in a valid index. Unless extensive local research can be sponsored, as has been done for the African Gender and Development Index that is in preparation by the UN Economic Commission for Africa, one has to rely on more or less the same indicators as the UNDP has done. The indicators used to calculate the GDI and GEM can be classified into five categories: education (enrolment and literacy), health (life expectancy), income (although labor market participation might be preferred, see later discussion of the GEM), share in parliament, and share in higher labor market positions. These are the categories used by Dijkstra (2002) to develop a Standardized Index of Gender Equality (SIGE) because they cover the eight dimensions mentioned earlier to a large extent.

A similar index can be obtained by including three composite indicators representing the social, economic and political dimensions of gender equality, as Social Watch (2005) has done in its Gender Equity Index. The social dimension includes only education (as in the HDI and GDI, enrolment and literacy are used to measure it); the economic dimension uses income and share in total paid jobs to detect inequity; and the political or empowerment dimension includes both the share in parliament and the share in higher labor market positions. In comparison with SIGE, the health indicator (i.e. life expectancy) has been dropped, thus avoiding the problem of adjusting for biological differences between men and women. This exclusion can be justified by the close empirical correlation between gender inequality in health and education (Dijkstra, 2002).

Both proposed indices lend themselves for comparisons between countries. The final score for a country, however, depends on the weighting of the different indicators. A persistent problem is that the indicator with the widest variation in scores among countries dominates

the score on the overall index. This can be solved by standardization of the scores: expressing them as deviations from the (arithmetic) mean, as done in the SIGE (for details see Dijkstra, 2002). The cross-country comparison of the scores and ranks becomes then more meaningful. On the other hand, as Klasen's contribution in the current *JHD* special issue rightly points out, it is more difficult to make comparisons over time, as the weighting will be different every year.

Another, more sophisticated, method for solving the problem of unintended weighting is the technique of "data envelopment analysis". In this method, a set of common weights is established in a linear programming model, such that the ultimate score is optimized for all countries. Despotis (2005) used this approach for the variables that constitute the HDI. The new index calculated using these optimal weights is highly correlated with the HDI. Although this is a superior method for establishing the weights of different components, it suffers from the same pitfall as standardization; namely, that comparisons over time become slightly less meaningful.

### **GDI and GEM as direct measures of gender equality**

The GDI includes the same dimensions as the HDI—education (enrolment plus literacy), health (life expectancy) and adjusted income per capita—but penalizes for inequality between female and male achievements by computing a harmonic mean between the two, which gives a higher weight to the lowest achievement. This means that the GDI is basically a *welfare* measure: absolute achievements on the three components have a large influence on the overall score. For example, if gender equality in life expectancy improves due to a general (although more pronounced for men) deterioration in this variable (as has occurred, for instance, in former Soviet Union countries), the GDI score will hardly increase, while an index only taking into account gender equality would. For some this is an advantage. However, it is important to also have a direct measure of gender equality. The question I address now is: to what extent is it possible to deduce or derive such a measure from the GDI itself or from its basic components?

Anand and Sen (1995, p. 7), who developed the methodology of the GDI, wrote that the GDI "incorporates implicitly something like a gender equality index". This underlying "index of equality" for each component is given by the ratio of the harmonic mean (as in the GDI) and the arithmetic mean (as in the HDI). White (1997) uses it to construct his Gender Equality Index (GEQ), defined as GDI/HDI. If  $I$  is the index of average achievement used to calculate the HDI components,  $I^*$ , the index used to calculate the GDI components, is given by (White, 1997):

$$I_{i,j}^* = \left[ \frac{p_{m,j}}{I_{i,m,j}} + \frac{p_{f,j}}{I_{i,f,j}} \right]^{-1} \quad (1)$$

where  $I_{i,m,j}$  is the indexed value of variable  $i$ , for men and for country  $j$ ; and  $p_{m,j}$  is the proportion of males in the population (and similarly for female with  $f$ ). It can be seen in equation (1) that if the population shares  $p_{m,j}$  and  $p_{f,j}$  are both 0.5 and if  $I_m = I_f$  (equality of achievement), then  $I^* = I$ .

Following up on equation (1), the components of the GDI and also (arguably, see later) the GDI itself can be decomposed into a component measuring overall welfare, the HDI, and a component measuring gender equality.

$$GDI = HDI \times \frac{GDI}{HDI} = HDI \times GEQ \quad (2)$$

Then:

$$\hat{GDI} \approx \hat{HDI} + \hat{GEQ} \quad (3)$$

where “^” indicates a percentage change in the variable. However, in practice there are two reasons why the GEQ is not an accurate measure of gender equality. First, the individual components of the GDI penalize gender inequality in both directions, and the overall GDI accumulates these different types of gender inequality. Countries that treat women better than men are also penalized in their GDI score, and countries cannot compensate higher relative female scores in, for example, education with lower female scores in income. This means the score on GDI/HDI does not say anything about the relative position of women in societies relative to men.

Second, the index of the income variable of the GDI is not a linear combination of female and male achievements as the other components are. The HDI uses *adjusted* income per capita, which means that higher incomes are discounted. The GDI follows the HDI in this. Since 1999, the UNDP has used a simple discounting procedure: the natural logarithm of income per capita. For the GDI, the adjusted female and male incomes per capita are first computed, and then the population-weighted harmonic mean is taken of the two, as in equation (1). Finally, the resulting score is indexed to get a scale from 0 to 100.<sup>1</sup> But since this income component is not a linear transformation of female and male achievements, there is no simple relationship anymore between the income component of the GDI and the income component of the HDI. In practice, the logarithmic transformation also discounts the gap between female and male achievements. As Table 1 shows, not only is the income component of the GDI ( $I^*$ ) higher for countries with higher income and the same gender gaps, as could be expected (see the sixth column of data, and compare countries B and C with Ecuador, or countries D and E with each other), but this also holds for the *ratio* of the income component of the GDI ( $I^*$ ) and the income component of the HDI ( $I$ ) (see final column). This means that the overall GDI/HDI index favors countries with higher incomes, and thus is *not* an accurate measure of gender equality *per se*.

It has also been proposed to use the difference between the HDI and GDI, defined as  $(HDI - GDI)/HDI$ —as, for example, by Forsythe *et al.*

Table 1. Calculations of the income component of the HDI ( $I$ ), the GDI ( $I^*$ ) and the GEQ ( $I^*/I$ ) for Ecuador and for four hypothetical countries with the same population and population shares, but different levels of income per capita and/or income shares

	Income share	Income per capita	Population share	Index of adjusted (ln) income	$I$	$I^*$	$I^*/I$
Ecuador							
Female	0.19	1173	0.50	0.41	0.57	0.50	0.89
Male	0.81	4818	0.50	0.65			
Total/average		3003		0.57			
Country B (lower income, same gender gap)							
Female	0.19	390	0.50	0.23	0.38	0.31	0.79
Male	0.81	1605	0.50	0.46			
Total/average		1000		0.38			
Country C (higher income, same gender gap)							
Female	0.19	7795	0.50	0.73	0.88	0.83	0.94
Male	0.81	32095	0.50	0.96			
Total/average		20000		0.88			
Country D (same income as Ecuador, lower gender gap)							
Female	0.40	2413	0.50	0.53	0.57	0.56	0.99
Male	0.60	3587	0.50	0.60			
Total/average		3003		0.57			
Country E (higher income, same gender gap as country D)							
Female	0.40	16073	0.50	0.85	0.88	0.88	0.99
Male	0.60	23892	0.50	0.91			
Total/average		20000		0.88			

Source: Author's own elaboration. Data for Ecuador are from UNDP (2000).

(1998), who take it as a measure of gender inequality. However, this gender inequality does not measure gender inequality *per se*, but instead *the reduction in welfare* due to gender inequalities. Furthermore, the same objections apply as against the GEQ; namely, the peculiar measurement of the income variable and the accumulation of gender gaps in different directions.

Given the impossibility to derive either a direct measure of gender equality from the GDI or a GDI–HDI comparison, it is necessary to drop the methodology of the GDI altogether for the purpose of creating a gender inequality index. It might still be possible to use its components to create a new composite index by using a different aggregation methodology. This has been proposed, for example, by Dijkstra and Hanmer (2000), who have developed the Relative Status of Women (RSW) index. The RSW uses exactly the same components of the GDI but does not include a relation with absolute levels of achievement. It computes an arithmetic average of three relative scores: the ratio of the female and male index for education ( $E_m$  and  $E_f$ ), the ratio of the female and male index for life expectancy ( $L_f$  and  $L_m$ ),

and the relative female and male returns to labor ( $w_f^*$  and  $w_m^*$ ):

$$RSW = \frac{1}{3} \left\{ \frac{E_f}{E_m} + \frac{L_f}{L_m} + \frac{w_f^*}{w_m^*} \right\} \quad (4)$$

Since female achievements are always in the numerator, the RSW does not accumulate inequalities in different directions, but allows for compensation.

The indices for the three components are computed in exactly the same way as for the GDI, including the income variable, which means that it is a combination of relative labor market participation and relative non-agricultural wages. Since the latter is not known for most countries and instead a 75% is taken, it is recommended to drop the relative wages from the formula, and just include relative labor market participation. It does not make sense to include a variable that does not vary for most countries. This RSW is the most promising improvement on the GDI if (something like) the GDI is to be maintained. However, the RSW also suffers from the shortcoming that the variable with the highest variation, which is the income variable, has the highest actual weight in the overall score. As shown above, this can be solved by standardizing the three different scores, or by applying data envelopment analysis. It is worthwhile further exploring this latter technique, either for the RSW or for a (modified) SIGE or Gender Equity Index.

To summarize the points made in this section, Table 2 provides an overview of the advantages and disadvantages of the GDI and GDI-related indices.

For the GEM, my recommendations are relatively simple but very important. As long as there are no internationally comparable and reliable data on gender-based violence, and on relative representation of women in local governments, the three indicators now used in the GEM could remain the same: share of women in parliament, share of women in higher labor market positions (professional and technical positions, legislators, senior officials and managers), and income. However, the GEM can be made a simple and direct measure of relative power of women vis-à-vis men by eliminating the methodological linkages with the HDI and GDI. As in the SIGE earlier, for the first two components the female shares can be taken directly without taking population-weighted harmonic means of them, which currently softens inequality without any conceptual justification (other than “similarity with the GDI”; see Dijkstra, 2000). The percentage shares can just be multiplied by two in order to get an index from 0 (absolute inequality, a share of 0) to 100 (absolute equality, corresponding to a share of 50%).<sup>2</sup>

For income, the female share in total income could be used. In order to get a direct measure of (in)equality, this income share should no longer be multiplied by average income per capita, as is now done in the GEM. The female income share should then be multiplied by two to get an index from 0 to 100. However, given the shortage of data on the female/male wage ratio (see earlier in the discussion of the RSW), it is better to use the

Table 2. Advantages and disadvantages of the GDI and GDI-related indices

Index	What it is	Advantages	Disadvantages
UNDP's GDI	Human development corrected for gender equity	Overall welfare measure  Directly comparable with HDI	No direct measure of gender equality  Penalties for gender inequalities in three indicators are accumulated, even if inequalities are in different directions Gender differences in income dominate
White's GEQ=GDI/HDI	An approximate measure of gender equality in same dimensions as HDI	More informative on gender equality than GDI	No measure of overall welfare Penalties for gender inequalities in three indicators are accumulated, even if inequalities are in different directions Gender differences in income dominate
Forsythe <i>et al.</i> 's GI= (HDI -GDI)/HDI	The extent of lower human development due to gender inequality	More informative on gender equality than GDI (but less than GEQ or RSW)	No measure of overall welfare No direct measure of gender equality Penalties for gender inequalities in three indicators are accumulated, even if inequalities are in different directions Gender differences in income dominate
Dijkstra and Hanmer's RSW	Gender equality in same dimensions as HDI	Direct measure of equality, same dimensions as HDI No longer accumulation of penalties for inequalities in three indicators, but compensation allowed	No measure of overall welfare Gender differences in income dominate

relative female/male labor market participation rate as an approximation of relative incomes. This is even more justified for a revised GEM than for the RSW or a revised GDI, since the GEM is meant to measure the empowerment dimension of income. Labor market participation is an important indicator of this as it embodies formal recognition of women's work, be it in formal employment, unemployment, self-employed, or work



in the informal or subsistence sector.<sup>3</sup> The ratio of female-to-male labor market participation is already on a scale between 0 and 100, so it does not need further indexing. A simple average can be taken of the three scores. This “new GEM” or “GEM\*” would be a direct measure of gender equality.

## Notes

- 1 Before 1999, the calculation was a bit different. First, the harmonic mean was taken of female and male shares in earned income, and the resulting figure was multiplied by the adjusted income per capita, and then indexed. This also discounts the gap between female and male incomes, but to a slightly lesser extent. In my earlier paper (Dijkstra, 2002), I showed larger differences between the old and new method of computing the GDI, but I then overlooked that the largest part of that difference stemmed from the new way of calculating the adjusted income for the HDI: from a complicated Atkinson transformation that only reduces incomes above a middle income level, to a log transformation that produces a more continuous scale (see technical note in UNDP, 1999).
- 2 This assumes that women account for 50% of the population, or in fact 50% of the working population and of the passive electorate. Although this is seldom fully true, the differences are so small that this assumption is justified.
- 3 This is the official definition of the UN System of National Accounts, but in practice not all countries include the latter two categories.

## References

- Anand, S. and Sen, A. (1995) ‘Gender inequality in human development: theories and measurement’, *Occasional Papers No. 19*, UN Human Development Report Office, New York.
- Despotis, D.K. (2005) ‘A reassessment of the human development index via data envelopment analysis’, *Journal of the Operational Research Society*, 56, pp. 969–980.
- Dijkstra, A.G. (2000) ‘A larger pie through a fair share? Gender equality and economic performance’, *ISS Working Paper Series No. 315*, Institute of Social Studies, The Hague.
- Dijkstra, A.G. (2002) ‘Revisiting UNDP’s GDI and GEM: Towards an alternative’, *Social Indicators Research*, 57, pp. 301–338.
- Dijkstra, A.G. and Hanmer, L.C. (2000) ‘Measuring socio-economic gender equality: towards an alternative for UNDP’s GDI’, *Feminist Economics*, 6(2), pp. 41–75.
- Ferber, M.A. and Nelson, J.A. (1993) *Beyond Economic Man: Feminist Theory and Economics*, Chicago University Press, Chicago and London.
- Forsythe, N., Korzeniewick, R.P. and Dunant, V. (1998) ‘Gender inequalities, economic growth, and structural adjustment: a longitudinal evaluation’, paper presented to the *XXI Conference of LASA*, Washington, DC.
- Social Watch (2005) *Roars and Whispers Gender and Poverty: Promises versus Action*, Social Watch, Montevideo.
- UNDP (1999) *Human Development Report 1999*, Oxford University Press for UNDP, New York and Oxford.
- UNDP (2000) *Human Development Report 2000*, Oxford University Press for UNDP, New York and Oxford.
- White, H. (1997) ‘Patterns of gender discrimination: an examination of the UNDP’s Gender Development Index’, *mimeo*, Institute of Social Studies (ISS), The Hague.