The definitive, peer-reviewed and edited version of this book chapter is published and can be cited as

Santos, M. E. and G. Santos (2014), 'Composite Indices of Development', in Currie-Alder, B., Kanbur, R., Malone D. and R. Medhora (eds.), *International Development: Ideas, Experience and Prospects*, Oxford: Oxford University Press, pp. 133-150.

DOI: http://dx.doi.org/10.1093/acprof:oso/9780199671656.003.0009

ISBN paperback: 978-0-19-967166-3
ISBN hardcover: 978-0-19-967165-6

Companion website: http://www.developmentideas.info/

Composite Indices of Development

Maria Emma Santos and Georgina Santos

Abstract

This paper reviews the literature on composite (and multidimensional) indices of development. Composite indices emerged as an alternative to using a portfolio of indicators, whose scattered information is sometimes difficult to grasp, or simply the GNP per capita, which often does not correlate well with development goals. As they emerged, they were also criticized. Points of debate relate to the selection of dimensions and indicators, their correlation (and the trade-off between redundancy and robustness), their type (input vs. output and stock vs. flow), and the normalization procedure, weighting, and aggregation of the components. However, as long as the purpose of the index and its indicators and weights are clearly specified and justified, the direction in which the index will move under specific transformations is axiomatically stated, robustness tests are performed, and the index is open to public scrutiny and revision, composite (and multidimensional) indices can prove invaluable in development studies.

Keywords

composite indices, multidimensional indices, development measures, human development, poverty measures, gender indices

Introduction

The question of how to measure development has attracted the attention of economists and other social scientists as well as non-government organizations (NGOs) and policy-makers for many decades now, especially since the post-war period. Broadly, we can identify three approaches to measuring development that emerged more or less sequentially over time, but now co-exist.

One approach considers that development can be measured with some specification of a monetary indicator: Gross National (or Domestic) Product (GNP and GDP, respectively), usually in per capita terms, and typically with special attention to its growth rate.[1] Most of the proponents of this view do not necessarily regard economic growth as the "end" of development, but consider GNP per capita to be a good enough proxy for well-being, highly correlated with other indicators that are less arguably considered as development goals.

The second approach states that GNP per capita has too many deficiencies as an indicator of well-being and that it does not always correlate well with development goals; therefore, a portfolio or dashboard of social indicators (including but not limited to monetary indicators) should be used to measure development.

A third approach considers that while portfolios of development indicators are informative and necessary, there is also a need for a summary measure that combines a few of these indicators into a single number. This approach has given rise to the construction of composite indices of development. A composite index is a function of variables and weights that maps attainments in a variety of attributes into a single real number, which may have cardinal meaning or be merely ordinal.

In this chapter we examine the motivation for the emergence of composite indices of development and the main grounds on which they have been criticized, which naturally coincide with the decisions involved in their construction. For simplicity, in most of the text we refer to "composite indices" in a broad sense, including multidimensional indices, clarifying below the distinction between the two. There is a discussion of some issues specific to gender-related indices, and we conclude with a review of the trade-offs around composite indices alongside a few recommendations for their design.

The emergence of composite indices of development

In the years immediately following the end of the Second World War, the System of National Accounts was developed with the aim of providing a complete accounting framework for reporting and evaluating the performance of an economy.

This was a natural response to the Great Depression, two devastating wars, and the influence of the Keynesian theory. Attention was focused on obtaining accurate computations of GNP and fostering their increase in per capita terms. Many argue that in those years GNP per capita was the primary—if not the sole—indicator of development (Hicks and Streeten 1979; UNDP 1990; ul Haq 1995, among others). Srinivasan (1994) argues that this was not the case, as policy-makers also considered indicators such as child mortality and life expectancy. Certainly, it would be unfair to say that there was no concern for human well-being. Rather, the underlying idea was that growth was the best instrument to reduce deprivation and warrant human flourishing, and thus it was a good measure of development. We identify this as the **first approach to measuring development**, namely using **GNP per capita**.

Yet in the early 1970s one empirical fact caught the attention of economists and international agencies: thirty years of outstanding economic growth performance had been accompanied by notable rising dualism within nations and a failure to reduce poverty. The limitations of

GNP per capita as an indicator of development and of the power of economic growth as an instrument for poverty reduction started to be exposed. This gave rise to a change of emphasis in the conceptualization of development. It was time to measure development more directly, paying attention to the evolution of unemployment, poverty, and inequality (Seers 1969). An emerging new approach would emphasize the need to refocus development on removing mass deprivation and ensuring that all human beings met their basic needs. [2]

Thus efforts were redirected to push the development of a set of cross-country comparable social indicators, as evidenced by the work of the UN and reports by other international organizations, cited in Hicks and Streeten (1979: 570). These publications considered not just economic indicators, such as the distribution of household income by deciles, but also health and education indicators, such as per capita protein consumption, infant mortality rate, combined primary and secondary enrolment ratio, and literacy rate, to name just a few.

The Basic Needs Approach advocated a parsimonious set of core indicators covering six areas: nutrition, basic education, health, sanitation, water supply and housing, and related infrastructure which would supplement GNP (Hicks and Streeten 1979). We identify this as the **second approach to measuring development**, namely using a **portfolio or dashboard of indicators**, more or less broad depending on the intended focus.

About the same time as efforts to measure development were shifting from GDP per capita to a portfolio of indicators, interest in constructing composite indices started to emerge. In fact, as early as 1964, Harbison and Myers proposed a composite indicator that focused on human resource development (as distinct from development in general). The index was the arithmetic total of enrolment at secondary level of education and enrolment at tertiary level of education, the latter multiplied by a weight of five (Harbison and Myers 1964: 31–2).

A few years later, the United Nations Research Institute for Social Development (UNRISD) proposed an index of socio-economic development designed by McGranahan et al. (1972). The index was composed of nineteen indicators, including several indicators of economic development (defined as GNP per capita) alongside various indicators of structural change (such as manufacturing share of GNP), some of education (such as combined primary and secondary enrolment), and two of health (per capita-per day consumption of animal protein, and life expectancy at birth). The scaling and weighting of the variables was based on a statistical procedure. This index was expressly intended to fill in the gaps of GNP per capita as a measure of development.

Inspired by the basic needs approach, Morris (1978) proposed a composite index: the Physical Quality of Life Index (PQLI), which was the arithmetic mean of (normalized) life expectancy, infant mortality, and literacy. Notably, the index did not include any measure of economic performance. Ram (1982) proposed two variants of the PQLI, where the main innovation was the technique used to construct the index: the multivariate method of principal components.

The Basic Needs Approach also had a practical influence on poverty measurement in the 1980s. In Latin America, poverty started to be measured with census information as the proportion of people in households that reported one or more Unsatisfied Basic Needs (UBN) of a total of five indicators for housing and education. The UBN Index was the predecessor of multidimensional poverty measures.

Other proposed composite indices were the Index of Social Progress (ISP, Estes 1984) and the Human Suffering Index (HSI, Camp and Speidel 1987). The ISP was composed of over forty indicators grouped in ten sub-indices: education, health status, women status, defense effort, economic, demography, environment, social chaos, cultural diversity, and welfare effort. The IHS was composed of ten indicators representing various dimensions of human suffering, not all of them obvious.

The early work on composite indices of development expressed “the need for a single number which, like GNP per head, can be quickly grasped and gives a rough indication of ‘social’ development” (Hicks and Streeten 1979: 577). Although GNP per capita alone was insufficient as an indicator of development, the information provided by a dashboard of indicators could be, for many, too much to digest.

Each of the aforementioned composite indices of development received some attention. Yet it seems that it was not until the appearance of the Human Development Index (HDI) in the first Human Development Report (HDR) in 1990 that composite indices of development received wider attention, naturally accompanied by deep scrutiny and a host of critiques.

Conceptually, the HDI was rooted in Sen's capability approach, which started with Sen's (1979) Tanner Lecture, "Equality of What?", where he proposed the need to shift the focus of attention from *means* of development, such as income and resources, to *ends* of development: the opportunities a person has. The capability approach differs from the basic needs one in that it broadens the scope of interest to all human beings (not just the poor), and it states the need to change the space of analysis to the capability set. Capabilities are defined as the various combinations of functionings (beings and doings) that the person can achieve. Capability is, thus, a set of vectors of functionings, reflecting the person's freedom to lead one type of life or another, to choose from possible livings (Sen 1992: 40). Human development in this context is the process of expanding the real freedoms that people enjoy (Sen 1999: 3). [3]

Inspired by these ideas, the intention of Mahbub ul Haq in introducing the HDI was to consolidate the concept of development beyond growth in GNP. “The new index would measure the basic concept of human development to enlarge people’s choices . . . at least a few more choices besides income and to reflect them in a methodologically sound composite index” (ul Haq 1995: 177).

The HDI (Anand and Sen 1994) considers a country’s achievements in three dimensions: living standards, health, and education. In its original formulation, the indicator for living standards was the log of the real GDP per capita (in PPP$),[4] the health indicator was life expectancy at birth, and the education indicator was the literacy rate. The index was the arithmetic mean of these (normalized) indicators. Although inspired by the capability approach, the components of the HDI were present in previous indices; the health and education indicators reflected functionings insofar as data permitted at the time, and thus there was no advance on that front.

However, the HDI did have several advantages over its predecessors. First, unlike the PQLI, it included GNP per capita, and thus it was not blind to the relevance of the economic dimension for development. Second, unlike the UNRISD index, the ISP and the HSI, it had a small number of components, which anyone could remember. Third, all these components were intuitive indicators of development. Fourth, although the subject of intense debate, unlike the indices proposed by Ram (1982) and the UNRISD index, the weights as well as the normalization formula for each indicator were transparent, easily understandable, and replicable.[5] These features, together with the fact that UNDP started to publish the HDR annually with continuous updates of the index, placed the HDI as the showcase of what we identify as the **third approach to measuring development**, namely using **composite indices of development**.

The three main criticisms leveled at the HDI are related to (1) the selection of dimensions and indicators, (2) the implicit trade-offs, and (3) the insensitiveness of the HDI to inequalities in the distribution of human development in the population. Many of the critiques were accompanied by proposals of modifications to the HDI or suggestions for completely different alternative indices. Some of these critiques were echoed by UNDP, and over the subsequent years the HDI experienced methodological modifications.[6]

After the HDI, three other composite indices were introduced in the HDR, aiming at capturing more directly certain specific areas related to development, namely gender equality and poverty.[7] The 1995 HDR introduced the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM). In 1996, a Capability Poverty Measure (CPM) was introduced. but was replaced the following year by the Human Poverty Index (HPI) (Anand and Sen 1997), which more accurately expressed failures in the three dimensions of the HDI. The components of the HPI were: for health, the probability at birth of not surviving to age 40; for education, the adult literacy rate; and for living standards, the average of the percentage of the population without access to an improved water source and the percentage of children under weight-for-age. In 1998 a variant of the HPI for developed countries was introduced. These indices were reported in every HDR until 2009.

UNDP has certainly not been the only institution producing composite indices of development. Several other institutions, NGOs, and think tanks have developed composite indices, which usually focus on certain aspects of development. For example, on economic aspects there is the Index of Economic Freedom (Heritage Foundation); on governance, the Ibrahim Index of African Governance (Mo Ibrahim Foundation); on sustainability, the Environmental Performance Index (EPI, Yale University); and we briefly review some gender ones below. Bandura (2008) surveys 178 composite indicators that rank or assess countries according to some economic, political, social, or environmental measure, all of which can be related to development.

The understanding that development is about improving people's lives, which motivated the construction of composite indices, also led, at a political level, to the Millennium Declaration in 2000, by which 189 heads of state committed to eradicating poverty and promoting other fundamental aspects of development by 2015. The Declaration was materialized in eight development goals (Millennium Declaration Goals, or MDGs) with eighteen associated targets and forty-nine quantitative indicators to follow them up (UN 2003). The MDG Indicators pushed the improvement in data collection in many countries, as progress towards *each* goal and target needs to be tracked. Interestingly, while fostering a multidimensional approach, the MDGs counterbalanced the interest on composite indices of development, favoring a dashboard approach. Likewise, the World Development Indicators offer a plethora of indicators related to different areas of development.

The 20th anniversary edition (2010) of the HDR introduced a number of changes in the measures of development. First, the living standard and educational indicators, the goalposts, and the aggregation formula of the HDI were modified. Second, the Inequality-Adjusted HDI was introduced. Third, the Gender Inequality Index (GII) replaced previous gender indices. Fourth, the HPI was replaced by the Multidimensional Poverty Index (MPI; see Alkire and Santos 2010). The MPI covers the three HDI dimensions (health, education, and living standards) using the ten indicators (nutrition and child mortality, school attendance and years of education, access to drinking water, improved sanitation, electricity, clean cooking fuel, non-dirt floor, and two small assets or one big one). The MPI was designed to reflect acute multidimensional poverty in a cross-country comparable way.[8] A distinctive feature of the MPI is that it looks at *joint deprivations*, requiring that the data come from the same source. In fact, the MPI is the product of two intuitive sub-indices: the **incidence** of poverty and the

intensity of the deprivation that the poor experience. The 2010 HDR changes rekindled the debate on composite (and multidimensional) indices of development.

To sum up, there are currently three approaches to measuring development. Growth is still regarded by many as the key measure of development: "growth is a necessary, if not sufficient, condition for broader development, enlarging the scope for individuals to be productive and creative" (Commission on Growth and Development 2008: 1). Others explicitly favor a multidimensional approach as long as each indicator is kept separate, and they cast some doubt about the value-added of composite indices and their policy relevance (Ravallion 2010a). Finally, other researchers and institutions, such as UNDP, favor composite and/or multidimensional indices of development, as "they have a stronger impact on the mind and draw public attention more powerfully than a long list of many indicators combined with a qualitative discussion" (Streeten 1994: 235) and, in the case of multidimensional indices, allow looking at the joint distribution of achievements (or deprivations).

Critical issues in designing composite indices

Basic notation

In order to discuss the main issues of debate around composite indices of development, it is useful to introduce some common mathematical notation.

Given m indicators across a population of n individuals, let $X = \left[x_{ij} \right]$ denote the $n \times m$ matrix of achievements. The typical entry $x_{ij} \geq 0$ represents individual i's achievement in indicator j. A composite index of development can be written in a general form as:

$$I(X) = \varphi\left[w_j \eta(A_j(x_{ij})) \right] \qquad (1)$$

where A_j is a function that aggregates individual achievements in an attribute j across the population to obtain an indicator of that achievement, for example at country level. η is a normalization function that expresses all indicators in the same unit of measurement (typically between 0 and 1). w_j is the (explicit) weight attached to attribute j (with $\sum_{j=1}^{m} w_j = 1$), and φ is an aggregation function.

A typical aggregation function A_j of individual data in attribute *j* is given by the generalized mean of order *β*:

$$\mu_\beta(x_{ij}) = \begin{cases} \left[\frac{1}{n}\sum_{i=1}^{n}(x_{ij})^\beta\right]^{1/\beta} & \beta \neq 0 \\ \prod_{i=1}^{n}(x_{ij})^{1/n} & \beta = 0 \end{cases} \qquad (2)$$

When *β=1*, the expression is reduced to the arithmetic mean. When *β*<1, higher weight is given to lower x_{ij} values. In this range, the general means capture inequality in a distribution of achievements. For a given value of *β*, the more unequal a distribution is, the lower the *β*-mean will be with respect to the arithmetic mean. For a given distribution, the more one wants to penalize inequality, the lower the *β* value chosen should be. Two cases are frequently used: when *β*=0, the *β*-mean is called the *geometric mean*, and when *β*=-1, it is called the *harmonic mean*.[9] Most commonly, the arithmetic mean is used to obtain the indicator of achievement to construct composite indices, as it is the case of GDP per capita, the literacy rate, and the gross enrolment ratio, to name a few.

A typical normalization function is that used by the HDI: $\eta(A_j) = (A_j - A^{\min}) / (A^{\max} - A^{\min})$. The normalized indicator expresses the proportion achieved by a country of a total potential achievement.[10]

A typical aggregation function φ has most commonly been the arithmetic mean over the normalized weighted indicators, but other members of the general means are frequently employed as well.

Equation (1) can also be used as a general expression for multidimensional (rather than composite) indices, but with some important differences. First, although the equation presents the composite index as a function of the matrix of achievements, access to micro-data is not necessary when computing them. Only information on the indicators A_j, already aggregated across the population, is required, and thus, information on each indicator can come from different data sources. Moreover, different indicators may have different *base populations.* For example, while income per capita is computed over the total population *n*, the literacy rate is typically computed over the population age 15 years and older.

When computing multidimensional indices, on the other hand, all data need to come from the same source. This is because aggregation is performed first across attributes for each individual, and then across individuals. The A_j is in this case the identity function, and all aggregation is performed with the φ function. In other words, in multidimensional measures, the *joint distribution* matters. This is clear in multidimensional indices of poverty, where first the poor need to be identified, generally based on the number of deprivations experienced by each person. For example, the MPI identifies as "poor" anyone living in a household deprived in 33.33 percent of the weighted indicators. Even when an identification step is not required, as in a well-being index, aggregation across attributes is still the first step, as there is usually interest in accounting for relations of substitutability or complementarity across attributes.[11] Note, then, that in multidimensional indices the base population of all attributes needs to be the same.

A second difference, specific to multidimensional indices of poverty, is that the normalization function usually transforms achievements into shortfalls with respect to a desired threshold z_j. That is: $\eta(x_{ij}) = (z_j - x_{ij}) / z_j$ for those with $x_{ij} < z_j$ *and* who have been identified as poor, and $\eta(x_{ij}) = 0$ otherwise. When variables are of ordinal nature (that is, the magnitude of the distance between categories is meaningless, as in "sanitation facility"), a robust normalization procedure is simply to dichotomize the variable with reference to the deprivation cutoff: 1 being deprived, and 0 non-deprived. This is the procedure followed in the MPI.

Selection of dimensions and indicators

One of the key debates about composite indices relates to the selection of the relevant dimensions and the indicators to measure them—the x_{ij} of equation (1). These are obviously linked to the purpose of the measure, which can be to track development across countries, or to monitor national poverty reduction, or to target a poverty reduction program, or some other.

In practice, the selection of dimensions and indicators for composite indices has usually been based on (a) existing data or convention, (b) theory, (c) public “consensus,” (d) ongoing deliberative participatory processes, or (e) empirical evidence or analysis (Alkire 2008); but also (f) pragmatism or intuitive appeal, or some combination thereof (Booysen 2002: 119). Methods (b) to (e) seem preferable to (a), and especially to (f). Robeyns (2005) recommends a thorough justification of the reasons and methods used in the selection as well as an explanation for any omissions, and Sen (2009) argues that any list of basic capabilities should enjoy a high degree of consensus built upon a process of public reasoning.

However, it is frequently the case in the construction of composite indices that while the “ideal” list of dimensions and indicators follows some of the preferred methods, the actual indicators included in the index are constrained by data availability (method [a]). This usually happens with indices that intend to make cross-country comparisons, such as the HDI or the MPI. Nonetheless, the binding constraint of internationally comparable data is not exclusive of composite indices.[12] Thus, rather than deterring their use for cross-country comparisons, this should foster international homogenization of data collection. When the intended purpose of the measure is restricted to a country, there is scope for the use of better data.

In the construction of most composite indices, researchers have offered a justification for the selection of dimensions and indicators. Yet McGillivray and Noorbakhsh (2007) argue that the selection is ultimately always ad hoc. In fact, no composite index so far has escaped criticism in this respect. Discussions on this matter typically reflect the trade-off between parsimony, with its inherent risk of omitting relevant variables, and comprehensiveness, with the risk of being redundant. In the case of the HDI, for example, the intention was to "keep it simple and manageable" (ul Haq 1995: 182). Other indices reflect the other option, such as the Mo Ibrahim Index of African governance, which contains fifty-seven indicators.

Statistical techniques can be of aid in dealing with this problem. Provided a selection of the dimensions has already been done, the selection of indicators to represent those dimensions can be guided by correlations analysis, Principal Components Analysis, or Factor Analysis. The use of these techniques can help reduce the set of indicators.

Correlation analysis has also been used to scrutinize some popular composite indices of development, fuelling arguments both for and against composite indices. Hicks and Streeten (1979), for example, oppose the use of composite indices by arguing that if the individual indicators are highly correlated any of them alone would serve as an adequate index, and if they are not correlated but move in different directions across countries and in time, averaging would only conceal important issues. Larson and Wilford (1979) find that the three PQLI's components are closely correlated, and McGillivray (1991) finds the same for the HDI. Thus using just one component of these indices would yield similar findings as using the index itself. However, Noorbakhsh (1998) shows that the correlation coefficients between the component indicators of the HDI are much lower and often insignificant for sub-samples of countries grouped, for example, by level of human development. In this case, using different components of the index would result in different rankings within these groups.

Moreover, although higher correlation between indicators is often criticized as redundancy, Foster, McGillivray, and Seth (2012) show that the more correlated the component indicators of a composite index, the more robust the weighting, something usually desired for a composite index. They suggest that the trade-off between redundancy and robustness needs further research (pp. 51–2).

Aside from the choice of specific dimensions and indicators, there are also choices to be made in terms of the type of indicators to be used. One of them is whether to use indicators of inputs or means (such as resources), indicators of outputs or ends (such as functionings), or a combination of both. Interestingly, variables such as literacy represent measures of both ends and means (Booysen 2002: 120). Statistics on inputs are more available than on outputs, but outputs tend to be better measures (Atkinson et al. 2002: 20; Booysen 2002: 120 and 144–45). For example, Das et al. (2008: 3) argue that higher investment in health care infrastructure in many low income countries has not translated into improved health. Composite indices usually have some combination of both means and ends variables.

Another decision to be made is whether to use indicators of stock (such as assets or wealth) or flow (such as income). As remarked by Atkinson et al. (2002: 32), flow measures are easier to change through policy; moreover, flows may have impacts on stocks: higher qualifications of those *entering* the labor market may have an impact on the qualifications of those in the *existing* labor market. Concentrating on either flow or stock variables might yield a misleading picture. For example, Klasen (2007) notes that focusing on life expectancy (flow) for women and men in countries such as China might indicate that gender bias is being reduced, as women's mortality rates are decreasing. However, this conceals an increase in sex-selective abortions (Sen 2003). Including both stock and flow variables would uncover this issue (Klasen 2007).

In sum, choosing dimensions and indicators is a critical step in the construction of composite indices. Given the outlined difficulties, it seems worth clearly *delimiting* the purpose of the measure, fully justifying the selection of dimensions and indicators, and exploiting the index to its full potential while recognizing that such selection may still fall short of fulfilling the intended purpose.

Normalization, weights, and aggregation

Another line of significant debate over composite indices of development has been the embedded trade-offs between their indicators. This is given by the relative weights between components, called in economics the marginal rate of substitution: for a given level of development or well-being, how much of one achievement must be resigned in order to obtain an extra unit of another achievement? It can be verified that these trade-offs are typically determined by three elements expressed in equation (1): the normalization procedure η, the explicit weights attached to each component (w_j), and the aggregation function φ (Decancq and Lugo 2012). Ravallion (2010b, 2011) argues that while the explicit weights attached to components are usually transparent, the trade-offs between them are not.

The normalization procedure η involves selecting goalposts or deprivation cutoffs or some other normative reference. Higher maximum goalposts or deprivation cutoffs lead to an implicit higher relative weight of that indicator as it becomes more difficult to achieve the same relative increase in that particular variable (Booysen 2002: 125).[13]

As for the weights of explicit attributes, the approaches that have been used for setting them can be grouped into *statistical*, *normative* or *hybrid*, which combines statistical and normative (Decancq and Lugo 2012).

Statistical approaches include principal components analysis, correlation and regression coefficients of the variables with some selected variable not included in the index (hedonic weighting). The problem with these approaches is that the weights depend on the particular dataset used, making comparisons over time difficult. They are also less transparent to and understandable by non-scholars.

Normative approaches imply setting weights based on explicit value judgments. Within this category, Decancq and Lugo (2012) include using prices, an option advocated by many (Srinivasan 1994; Ravallion 2010a, 2011, for example). However, markets are imperfect and prices depend on the distribution of income (Lustig 2011; Seers 1969). Thus, meaningful prices may not be available across all dimensions of relevance to poverty (Alkire, Foster, and Santos 2011: 503), and—we could add—to development.

The limitations of prices as weights have led to the frequent use of equal weighting. This has been the case of the HDI as “there was no a priori rationale for giving higher weight to one choice than to another” (ul Haq 1995). The MPI also follows an equal weighting approach, a structure criticized by Ravallion (2010a, 2011). Yet equal weighting has been favored by experts, following a wide consultation (Chowdhury and Squire 2006), and it has also been recommended by Atkinson et al. (2002).

A promising hybrid method is that of setting weights based on participatory studies where people express their relative valuations of different attributes. Although this has not been implemented in the widely used composite indices of development (Ravallion 2011), there are some small-scale exercises (see Decancq and Lugo 2012) that may shed some light on the design of broader-scale indices in the future.

In terms of the aggregation function φ, ideally, one would like to know the functional form of a well-being production function (HDR 1993: 109). Yet this function remains unknown (McGillivray and Noorbakhsh 2007). The typically used arithmetic mean implies that the normalized attributes are considered perfect substitutes: failures in one area of development can be compensated by achievements in another. Many see this as conceptually problematic. An alternative is to use a member of the general means with $\beta<1$. This penalizes uneven development across attributes. The use of such an aggregation function with $0<\beta<1$ has been proposed by Chakravarty (2003). It has also been proposed by Foster, Lopez-Calva, and Szekely (2005) in the range of $\beta<1$, together with using a general mean of the same order β for constructing the aggregate indicators A_j, so that the index also captures inequality in the distribution of each attribute across people. This last idea was taken by UNDP to construct the Inequality-Adjusted HDI, introduced in the 2010 HDR.

The HDI used a linear form until 2009. Because GDP per capita was logged, the trade-off between longevity and income was dependent on the country's income level. Thus, the HDI's implicit monetary valuation of an extra year of life went from very low levels in poor countries to very high levels in rich ones (Ravallion 1997: 633). In 2010, UNDP changed the aggregation formula of the HDI to the geometric mean, in order to capture "how well rounded a country's performance is *across* the three dimensions" (HDR 2010: 15, emphasis added). The goalposts for the living standards and the life expectancy indicator were also changed. Given the data, these changes led in practice to an even lower (higher) valuation of longevity in poor (richer) countries, and a similar problem with every extra year of schooling, something Ravallion (2010b) calls "troubling trade-offs." Klugman, Rodriguez, and Choi (2011) responded to this critique, arguing that in rich countries income contributes very little to further expanding capabilities and this is why the "value" of anything in terms of income appears very high. However, it should not be concluded that more resources should be devoted to increasing longevity in rich countries than in poor ones.

The challenges faced by defining a normalization function, a weighting structure, and an aggregation function, with their implied trade-offs alongside the controversies in the particular selection of dimensions and indicators discussed in Section 3.2, lead some to lean towards a dashboard approach to measuring development (Hicks and Streeten 1979; Ravallion 2011). However, dashboards also suffer from several problems as detailed in Alkire, Foster, and Santos (2011: 503–04). First of all, policy-makers prefer a summary statistic to show how overall poverty (and development) has changed. Second, dashboards leave trade-offs completely open; thus they do not catalyze public scrutiny on these trade-offs, nor encourage transparency. Third, dashboards are blind to the joint distribution of achievements (or deprivations), something also acknowledged by Lustig (2011) and even Ravallion (2011). There is increasing consensus on the importance of considering the joint distribution.[14] Finally, in the particular case of poverty measurement, a dashboard is unable to answer the basic question of who is poor.

Thus, despite the challenges, many still argue for summary indices of development. One possible “way out” of potential controversies is to make each choice as transparent, explicit, and justifiable as possible. A common practice in this respect is to follow an axiomatic framework that clearly states the direction of change of the summary measure (or the requirement not to change) under different possible transformations of the achievements. This has been the tradition followed in the unidimensional inequality and poverty measurement. It is also being followed by multidimensional indices of poverty, as evidenced by several papers in this area, including Alkire and Foster (2011), which provides the mathematical structure of the MPI. This has been less common in the area of composite indices of development, but not totally absent, as shown by Chakravarty (2003) and Foster, Lopez-Calva, and Szekely (2005). Axiomatic frameworks are helpful in selecting indices and in identifying trade-offs between the desired properties. For example, for policy relevance, it is useful for a composite index to be subgroup consistent, such that if the development level of one subgroup (say a geographical region) rises and the rest are unchanged, the value of the overall index rises. Yet not all indices satisfy this property (see Foster, Lopez-Calva, and Szekely 2005). Similarly, also for policy convenience, one may want to break the index down into the contributions of its dimensional components; however, this requires assuming independence of the considered attributes.

A second important “way out” of controversies is to perform thorough robustness checks of the particular selection of normalization procedures and weighting structures, as done by Slottje (1991) and stressed by Ravallion (2010a), among others. In the particular cases of the HDI and the MPI, these checks suggest that country rankings are highly robust to alternative weighting structures. (For the HDI, see Foster, McGillivray, and Seth 2012; for the MPI, see Alkire and Santos 2010.)

Issues specific to gender indices

Gender equality is central to development. It is well documented that in some regions of the world there is anti-female bias that starts as early as before birth, with female children more likely to be aborted; they are also more likely to die earlier because of neglect (Sen 1990, 2003). Gender bias also exists in other dimensions such as food consumption, education, employment, and income-earning opportunities (Anand and Sen 1994: 11). Aside from the obvious unfairness of this bias, enhancing women's achievements in the mentioned dimensions is instrumental to other development goals.

Naturally, there have been efforts to capture gender bias in composite indices of development. In particular, the 1995 HDR, which focused on Gender and Human Development, introduced the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM). The GDI considered gender equality in the same three dimensions as the HDI and with very similar indicators (life expectancy by gender; adult literacy and enrolment by gender; female and male income shares). The objective was to discount the level of human development by the degree of inequality between these two groups, and thus was similar in spirit to the IHDI (Seth 2011). This was achieved using the harmonic mean to aggregate achievements of males and females within each dimension, and then the three indicators were aggregated with an arithmetic mean.

The GEM was designed to capture women's participation in economic and political decisions through women's representation in parliaments, women's share of managerial and professional positions, women's participation in the active labor force, and women's share of national income. As with GDI, achievements of each gender were aggregated using the harmonic mean within each dimension and then aggregated across dimensions with an arithmetic mean.[15]

Both GDI and GEM had important limitations.[16] A key one was that because of data limitations they had to rely significantly on imputations, especially for the income component. Additionally, since they aggregated across dimensions with an arithmetic mean, they were not penalizing uneven development across dimensions.

In 2010, these two indices were replaced by the Gender Inequality Index (GII), which followed the methodology proposed by Seth (2009). The index considers three dimensions: women's reproductive health, as measured by the (inverse of the) maternal mortality ratio and the (inverse of the) adolescent fertility rate; empowerment, as measured by the share of parliamentary seats held by each sex and attainment at secondary or higher educational level; and labor market participation, as measured by the labor market participation rate for each gender. By not including income as an indicator, GII avoids the need to impute values for a significant number of countries, and thus the estimates are more reliable. Like the GDI and GEM, GII captures inequality *across* genders by aggregating the male and female indices with the harmonic mean. However, GII also captures inequality in achievements across dimensions *within* each gender using the geometric mean.[17]

Although the GII is an improvement over the GDI and GEM, it has a number of shortcomings (summarized in Seth 2011: 16). One in particular is that the health indicators are only applicable to women; thus men are given a value of “one” as if they had achieved the best possible outcome. However, this does not really provide a basis for comparison of achievements across genders. Another is that the indicators of empowerment have a bias towards elites. To avoid that, they should include participation at the local government level and elsewhere in the community and public life.

There are several other gender indices. Van Staveren (2012) compares five of them: the GII, the Global Gender Gap Index (GGGI, used by the World Economic Forum), the Social Institutions and Gender Index (SIGI, OECD), the Gender Equality Index (GEI, Indices of Social Development database of the Institute of Social Studies of Erasmus University Rotterdam), and the Women’s Economic Opportunities Index (WEOI, Economist Intelligence Unit). She finds that the Pearson correlations between the indices are relatively high (between 0.50 and 0.81). However, each index yields quite different ranking results.

This is because they focus on different dimensions of human development, which she classifies into resources, capabilities, functionings, and institutions. For example, while GII uses mostly indicators of functionings, SIGI focuses on institutional gender equality and WEOI uses mainly indicators of resources. The rankings at the bottom are more similar than those at the top, suggesting that the different dimensions of human development do not automatically move together as countries develop.

The discussion above suggests that the construction of gender-related composite indices faces additional specific challenges on top of those that affect the design of composite indices in general. Klasen (2007) reviews some of these issues. Essentially, he notes that a great deal of gender inequality is generated within the household, be it allocation of income, food, or educational opportunities. This is obviously influenced by women's empowerment, a variable difficult to measure as women are typically willing to forgo resources for their children and to delegate decision making power. In order to uncover these intra-household allocation dynamics, Klasen highlights the need for better quality data on functionings such as nutrition and health status, cognitive abilities, time use, and so on. Additionally, he notes that combining stock and flow indicators becomes particularly crucial in gender indices, where life expectancy should be complemented by some stock measure in order to capture gender-selective abortions.

Conclusion

Composite indices stand as one of the options to measure development, as an alternative to using a portfolio of indicators or simply the GNP (or GDP) per capita. Most of what we have discussed in this chapter can be expressed as a trade-off. The very first one is whether or not to use composite indices. With composite indices we broaden the considered dimensions of development, an advantage with respect to GNP, and we gain power of synthesis such that countries or regions can be ranked, an advantage with respect to portfolios of indicators. Yet by aggregating "incomparables" we compromise on not expressing all components in market prices. Additionally, we cannot sensibly integrate *all* the relevant dimensions present in a dashboard.

While no single measure will suffice for measuring development, different measures may suit different purposes. Composite (and multidimensional) indices offer a flexible framework that can prove helpful in many instances. However, their construction faces a number of trade-offs. In selecting dimensions and indicators one can be comprehensive, albeit possibly redundant, or parsimonious, at risk of omitting significant variables. Furthermore, avoiding redundancy by including lowly correlated indicators can compromise the robustness of the index to alternative weighting structures. The index components can be "universal" and available for a wide number of countries, but possibly not as relevant to specific contexts.

Also, the index can be a composite one, overlooking joint deprivations, or a multidimensional one, imposing high data requirements.

None of these decisions is trivial for development policy, and none is free from value judgments. As argued by Atkinson (2001), there is a need to bring back into economics the study of normative principles that underlie our welfare assessments. In line with this observation, a number of practices related to their use seem recommendable. First, to fully specify the purpose of a measure. Second, to select dimensions, indicators and weights (in the broad sense of implicit trade-offs of components) with as much public scrutiny and justification as possible. In this respect, it is desirable to extend the practice of participatory processes to perform legitimate choices. Third, to elicit at least some of the alternative roads not taken. Fourth, to choose the mathematical structure of an index in view of alternative axiomatic frameworks, so that the direction in which the index will move under specific transformations is transparent. Fifth, to perform robustness tests in order to make explicit the scope of the conclusions that may be derived from the index values. Finally, to leave composite indices open to revision, as feedback from the different stakeholders may offer new insights into measurement.

The design and use of composite indices would also benefit from improvements in data collection, in particular at least a core set of outcomes (functionings) measurements in a cross-country comparable way.

---

[1] From now onwards we refer to GNP generically, which may in fact be GNP or GDP. The implications of each of these specifications for measuring development are not addressed in this chapter.

[2] See UNEP/UNCTAD (1974) for a first articulation of the Basic Needs Approach. Other relevant reports were those of the Dag Hammarskjöld Foundation (1976); Herrera et al. (1976); and ILO (1976). In 1978, the World Bank started to foster the approach (Streeten et al. 1981).

[3] Incidentally, there are important coincidences between the capability approach and the Christian view of economic development, as presented by the Social Doctrine of the Roman Catholic Church: "Development . . . cannot be restricted to economic growth alone . . . it must foster the development of each man and of the whole man" (Pope Paul VI, 1967, *Populorum Progressio*, point 14).

[4] The use of the log of the real GDP per capita was to reflect the diminishing returns of income on development.

[5] Slottje (1991) showed empirically that the choice of the weighting technique is quite critical.

[6] See OPHI (2011).

[7] We concentrate on indices that were reported annually for several years.

[8] Acute poverty is understood as a person's inability to meet minimum internationally agreed standards in a set of core human functionings and rudimentary services *simultaneously*.

[9] When $\beta>1$, higher weight is given to higher x_{ij} values; this is used to penalize for inequality when the x_{ij} arguments are deprivations rather than achievements. The general means with $\beta<1$ correspond to what Atkinson (1970) called the "equally distributed equivalent (EDE) income" (Y_{EDE}), which is core to his inequality measure defined as: $I_\beta = 1 - Y_{EDE}/\mu = 1 - \mu_\beta/\mu$.

[10] See OECD (2008: 83–88) for other normalization functions.

[11] See, for example, Seth (2009).

[12] The construction of cross-country comparable unidimensional indices also requires imputations, intra- and extrapolations, and other assumptions regarding the comparability between countries.

[13] This issue was early raised by Hicks and Streeten (1979) and later by Kelley (1991: 319) and McGillivray and Noorbakhsh (2007).

[14] See, for example, Stiglitz, Sen, Fitoussi (2009: 15); Deaton (2011: 14–15).

[15] For details on the computation of both GDI and GEM, see UNDP (1995, Technical Note).

[16] There are many reviews of the GDI and GEM indices. Klasen (2007) is one.

[17] Note that, in contrast to GDI and GEM, GII is an index of gender *in*equality. For a detailed description of the steps to compute GII, see UNDP (2010) or Seth (2011).

References

Alkire, Sabina (2008). "Choosing Dimensions: The Capability Approach and Multidimensional Poverty," in Nanak Kakwani and Jacques Silber (eds.), *The Many Dimensions of Poverty*. Basingstoke: Palgrave Macmillan.

Alkire, Sabina, and James E. Foster (2011). "Counting and Multidimensional Poverty Measurement," *Journal of Public Economics*, 95(7/8): 476–87.

Alkire, Sabina, James E. Foster, and Maria Emma Santos (2011). "Where Did Identification Go?", *Journal of Economic Inequality* 9(3): 501–05.

Alkire, Sabina, and Maria Emma Santos (2010). "Acute Multidimensional Poverty: A New Index for Developing Countries," OPHI Working Paper No. 38, Oxford Poverty & Human Development Initiative (OPHI), Department of International Development, University of Oxford. (http://www.ophi.org.uk/acute-multidimensional-poverty-a-new-index-for-developing-countries)

Anand, Sudhir, and Amartya K. Sen (1994). *Human Development Index: Methodology and Measurement.* Human Development Report Office, United Nations Development Programme (UNDP), New York. (http://hdr.undp.org/en/media/Anand_and_Sen_HDI.pdf)

——— (1997). "Concepts of Human Development and Poverty: A Multidimensional Perspective," Human Development Papers, United Nations Development Programme (UNDP), New York. (http://ias7.berkeley.edu/Academics/courses/center/fall2007/sehnbruch/UNDP%20Anand%20and%20Sen%20Concepts%20of%20HD%201997.pdf)

Atkinson, Anthony B. (1970). "On the Measurement of Inequality," *Journal of Economic Theory*, 2(3): 244–63.

——— (2001). "The Strange Disappearance of Welfare Economics," *Kiklos*, 54(2/3): 193–206.

Atkinson, Anthony B., Bea Cantillon, Eric Marlier, and Brian Nolan (2002). *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Bandura, Romina (2008). "A Survey of Composite Indices Measuring Country Performance: 2008 Update," UNDP/ODS Working Paper, United Nations Development Programme (UNPD), Office of Development Studies (ODS), New York. (http://web.undp.org/developmentstudies/docs/indices_2008_bandura.pdf)

Booysen, Frederik (2002). "An Overview and Evaluation of Composite Indices of Development," *Social Indicators Research* 59(2): 115–51.

Camp, Sharon L., and J. Joseph Speidel (1987). *The International Human Suffering Index*. Washington: Population Crisis Committee.

Chakravarty, Satya R. (2003). "A Generalized Human Development Index," *Review of Development Economics*, 7(1): 99–114.

Chowdhury, Shyamal, and Lyn Squire (2006). "Setting Weights for Aggregate Indices: An Application to the Commitment to Development Index and Human Development Index," *Journal of Development Studies*, 42(5): 761–71.

Commission on Growth and Development (2008). *The Growth Report: Strategies for Sustained Growth and Inclusive Development*.

Dag Hammarskjöld Foundation (1975). *What Now? Another Development*. The 1975 Dag Hammarskjöld Report on Development and International Cooperation, Seventh Special Session of the United Nations General Assembly (New York, September 1975). Motala, Sweden: Motala Grafiska.

Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard (2008). "The Quality of Medical Advice in Low Income Countries," *Journal of Economic Perspectives*, 22(2): 93–114.

Deaton, Angus (2011). "Measuring Development: Different Data, Different Conclusions?", in *Measure for Measure: How Well Do We Measure Development? Proceedings of the 8th AFD–EUDN Conference*, Paris, December 1, 2010. (http://www.afd.fr/webdav/site/afd/shared/PUBLICATIONS/RECHERCHE/Scientifiques/Conferences-seminaires/03-VA-Conferences-seminaires.pdf)

Decancq, Koen, and Maria Ana Lugo (2012). "Weights in Multidimensional Indices of Wellbeing: An Overview," *Econometric Reviews*, 32(1): 7–34.

Estes, Richard J. (1984). *The Social Progress of Nations*. New York: Praeger.

Foster, James E., Mark McGillivray, and Suman Seth (2012). “Composite Indices: Rank Robustness, Statistical Association, and Redundancy,” *Econometric Reviews*, 32(1): 35–56.

Foster, James E., Luis F. Lopez-Calva, and Miguel Szekely (2005). “Measuring the Distribution of Human Development: Methodology and Application to Mexico,” *Journal of Human Development*, 6(1): 5–30.

Harbison, Frederick, and Charles A. Myers (1964). *Education, Manpower, and Economic Growth*. New York: McGraw-Hill.

Herrera, Amílcar O., Hugo D. Scolnik, Graciela Chichilnisky, Adolfo Chorni, Gilberto C. Gallopin, Isabel Gómez, Jorge E. Hardoy, Diana Mosovich, Enrique Oteiza, Rafael Pastoriza, Victor H. Ponce, Gilda L. de Romero Brest, Juan V. Santiere, Abraam Sonis, Juan V. Sourrouille, Carlos E. Suarez, and Luis Talavera (1976). *Catastrophe or New Society? A Latin American World Model*. Ottawa: International Development Research Centre.

Hicks, Norman, and Paul Streeten (1979). “Indicators of Development: The Search for a Basic Needs Yardstick,” *World Development*, 7(6): 567–80.

International Labour Organization (ILO) (1976). *Employment, Growth and Basic Needs: A One World Problem.* Geneva.

Kelley, Allen C. (1991). “The Human Development Index: Handle With Care,” *Population and Development Review*, 17(2): 315–24.

Klasen, Stephan (2007). “Gender-Related Indicators of Well-Being,, in M. McGillivray (ed.), *Human Wellbeing: Concept and Measurement*. Basingstoke: Palgrave Macmillan.

Klugman, Jeni, Francisco Rodriguez, and Hyung-Jin Choi (2011). “Response to Martin Ravallion,” *Journal of Economic Inequality*, 9(3): 497–99.

Larson, David A., and Walton T. Wilford (1979). “The Physical Quality of Life Index: A Useful Social Indicator?”, *World Development*, 7(6): 581–84.

Lustig, Nora (2011). “Multidimensional Indices of Achievements and Poverty: What Do We Gain and What Do We Lose?” CGD Working Paper 262, Center for Global Development (CGD), Washington, DC.

McGillivray, Mark (1991). “The Human Development Index: Yet Another Redundant Composite Development Indicator?”, *World Development*, 19(10): 1461–68.

McGillivray, Mark, and Farhad Noorbakhsh (2007). “Composite Indices of Human Well-Being—Past, Present, and Future,” in Mark McGillivray (ed.), *Human Wellbeing: Concept and Measurement*. Basingstoke: Palgrave Macmillan.

McGranahan, Donald V., Claude Richard-Proust, N. V. Sovani, and M. Subramanian (1972). *Contents and Measurement of Socioeconomic Development: A Staff Study of the United Nations Research Institute for Social Development (UNRISD)*. New York: Praeger.

Morris, Morris David (1978). “A Physical Quality of Life Index,” *Urban Ecology*, 3(3): 225–40.

Noorbakhsh, Farhad (1998). “A Human Development Index: Some Technical Issues and Alternative Indices,” *Journal of International Development*, 10(5): 589–605.

Organization for Economic Cooperation and Development (OECD) (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris: OECD Publications. (http://www.oecd.org/std/leadingindicatorsandtendencysurveys/42495745.pdf)

Oxford Poverty & Human Development Initiative (OPHI) (2011). “Training Material for Producing the Human Development Index (HDI),” OPHI Research Paper 29a, Department of International Development, University of Oxford. (http://www.ophi.org.uk/wp-content/uploads/OPHI-RP-30a.pdf-2011.pdf)

Pope Paul VI (1967). *Populorum Progressio: Encyclical Letter on the Development of Peoples*. Washington, DC: United States Catholic Conference (now United States Conference of Catholic Bishops).

Ram, Rati (1982). “Composite Indices of Physical Quality of Life, Basic Needs Fulfilment, and Income: A ‘Principal Component’ Representation,” *Journal of Development Economics*, 11: 227–47.

Ravallion, Martin (1997). “Good and Bad Growth: The Human Development Reports,” *World Development*, 25(5): 631–38.

——— (2010a). "Mashup Indices of Development," World Bank Policy Research Working Paper No. 5432, World Bank, Washington, DC. (http://elibrary.worldbank.org/docserver/download/5432.pdf?expires=1350925722&id=id&accname=guest&checksum=F894255F85CC25BD883EC0088BDB37DC)

——— (2010b). "The Troubling Tradeoffs of the Human Development Index," World Bank Policy Research Working Paper No. 5484, World Bank, Washington, DC. (http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2010/11/29/000158349_20101129085041/Rendered/PDF/WPS5484.pdf)

——— (2011). "On Multidimensional Indices of Poverty," *Journal of Economic Inequality*, 9(2): 235–48.

Robeyns, Ingrid (2005). "Selecting Capabilities for Quality of Life Measurement," *Social Indicators Research*, 74(1): 191–215.

Seers, Dudley (1969). "The Meaning of Development," Institute of Development Studies Communication 44, Institute of Development Studies (IDS), University of Sussex, Brighton, UK. (http://www.ids.ac.uk/files/dmfile/themeaningofdevelopment.pdf)

Sen, Amartya K. (1979). “Equality of What?”, in Sterling M. McMurrin (ed.), *Tanner Lectures on Human Values*. Cambridge, UK: Cambridge University Press.

——— (1990). “More Than 100 Million Women Are Missing,” *The New York Review of Books*, 37(20), December 20.

——— (1992). *Inequality Reexamined.* New York, NY and Cambridge, MA: Harvard University Press.

——— (1999). *Development as Freedom*. Oxford: Oxford University Press.

——— (2003). “Missing Women—Revisited,” *British Medical Journal*, 327(6): 1297–98.

——— (2009). *The Idea of Justice*. London: Allen Lane.

Seth, Suman (2009). “Inequality, Interactions, and Human Development,” *Journal of Human Development and Capabilities*, 10(3): 375–96.

——— (2011). “Training Material for Producing National Human Development Reports: The Gender Inequality Index (GII),” OPHI Research Paper 28a, Oxford Poverty & Human Development Initiative (OPHI), Department of International Development, University of Oxford. (http://www.ophi.org.uk/wp-content/uploads/OPHI-RP-28a.pdf-2011.pdf)

Slottje, Daniel J. (1991). “Measuring the Quality of Life Across Countries,” *Review of Economics and Statistics*, 73(4): 684–93.

Srinivasan, T. N. (1994). “Human Development: A New Paradigm or Reinvention of the Wheel?”, *American Economic Review* (Papers and Proceedings), 84(2): 238–43.

Stiglitz, Joseph E., Amartya K. Sen, and Jean-Paul Fitoussi (2009). *Report of the Commission on the Measurement of Economic Performance and Social Progress*. (http://www.stiglitz-sen-fitoussi.fr/)

Streeten, Paul (1994). “Human Development: Means and Ends,” *American Economic Review* (Papers and Proceedings), 84(2): 232–37.

Streeten, Paul, Shahid Javed Burki, Mahbub ul Haq, Norman Hicks, and Frances Stewart (1981). *First Things First: Meeting Basic Human Needs in Developing Countries*. New York: Oxford University Press.

ul Haq, Mahbub (1995). *Reflections on Human Development*. New York: Oxford University Press.

United Nations (2003). *Indicators for Monitoring the Millennium Development Goals*. New York.

United Nations Development Programme (UNDP) (various years). *Human Development Report*. Human Development Report Office, UNDP, New York.

United Nations Environment Programme/United Nations Conference on Trade and Development (UNEP/UNCTAD) (1974). *The Cocoyoc Declaration*. *International Organization*, 29(3), International Responses to Technology (Summer 1975): 893–901.

van Staveren, Irene (2012). “To Measure is to Know? A Comparative Analysis of Gender Indices,”, *Review of Social Economy*, 0(0): 1–34.