Cumulative bias in the new digital opportunity index

James, M.J.

Published in:
Current Science

Publication date:
2007

Cumulative bias in the new Digital Opportunity Index: sources and consequences

Jeffrey James

An influential new ranking of countries in the area of information technology has recently been unveiled. Though the Digital Opportunity Index (DOI) has gathered a large amount of data, there are serious doubts about the variables chosen, as well as the weighting and values that they are assigned. In general, the index lacks an analytical foundation and is found to suffer from cumulative biases of different kinds. These and other problems all suggest that one should be cautious in drawing policy conclusions from the DOI, as it currently stands.

Keywords: Biases in ranking, cross-country index, Digital Opportunity Index, information technology.

In the 2006 edition of its World Information Society Report, the ITU has unveiled a new Digital Opportunity Index (DOI), the composite value of whose 11 indicators is used to rank as many as 180 countries. The DOI purports to be not only 'a tool to measure progress in building the Information Society', but also as a means of enriching and informing policy-making in this context[1]. To its credit, the DOI has compiled an impressively large amount of internationally comparable data across 11 indicators, for rich as well as very poor countries. The problem, however, is that the choice of these particular indicators as well as the values and weights that are assigned to them, give rise to what I describe as a cumulative bias in the overall ranking of countries and hence the validity of the policy conclusions that can be drawn therefrom. In particular, the purpose of this article is to advance two related points of criticism of the DOI. The first is that the index dubiously and without justification assigns equal weights to (clusters of) variables that, on the one hand, represent merely future potential and on the other hand, those that reflect actual achievements. This, I suggest, runs counter to any welfare or ends-based perspective on the issue. Secondly, I find that it is precisely the cluster of potential variables with the least connection to actual outcomes, that turn out (partly due to the specific choice of variables to be included in the index) to have a systematically higher value than the other clusters across the entire sample. Thus the cumulative result of the two individual biases is a clear inflation of the 'least' important variables and a corresponding under-valuation of what ultimately are the true goals of the endeavour. In order to present these arguments and the implications which they give rise to, I need, first, however, to provide a brief explanation of how the DOI itself is created.

Jeffrey James is in the Department of Development Economics, Tilburg University, The Netherlands. e-mail: M.J. James@uvt.nl

Essentials of the DOI

Table 1 shows what 11 variables make up the DOI and how they are clustered into three categories, representing 'opportunity', 'infrastructure' and 'utilization', respectively.

As with the HDI, the DOI first calculates an average score within each of the categories and then weights these values to arrive at an overall score for each country (as shown in ITU[2]). Let us then examine how exactly the three clusters in Table 1 are weighted.

Weighting in the DOI

As is the case with any composite index, the DOI faces the crucial task of weighting the components that comprise it. Indeed, much of the debate concerning the attempts to formulate social indicators in the 1970s and more recently, the HDI, was concerned with the justification of one weighting scheme over another. In the context of basic needs, for example, one influential writer observed that 'If an acceptable system of weights could be developed, it might be possible to combine the core indicators into a composite basic needs index. The chances of doing so are, however, extremely small. Despite considerable research on composite indexes, no one has come close to developing a national weighting system'[3]. Many studies, according to the same author, do not even devote much attention to 'developing a theoretically sound rationale for the weighting system'.

Against this background and indeed the more general literature on the topic, one might have thought that the choice of weights for each cluster of variables in the DOI had been made on an informed and well-reasoned basis. Yet, this is certainly not the impression one gets from the documents which explain how the DOI has been compiled. In fact, there is barely any discussion of the equal

Table 1. Structure of the DOI

Percentage of population covered by mobile cellular telephony	OPPORTUNITY
Internet access tariffs as a percentage of per capita income	
Mobile cellular tariffs as a percentage of per capita income	
Proportion of households with a fixed line telephone	INFRASTRUCTURE
Proportion of households with a computer	
Proportion of households with Internet access at home	
Mobile cellular subscribers per 100 inhabitants	
Mobile Internet subscribers per 100 inhabitants	
Proportion of individuals that used the Internet	UTILIZATION
Ratio of fixed broadband subscribers per 100 inhabitants	
Ratio of mobile broadband subscribers to total mobile subscribers	

Source: ITU[2].
Note: The interested reader can find definitions and methodologies used by the DOI in ITU[2].

weights that are assigned to the three clusters mentioned above. In part, this neglect of any serious justification for the weighting system adopted by the DOI may reflect confusion over what exactly the index is designed to measure. On the one hand, for example, the DOI is described 'as an objective measurement of individual and household access to ICT'[4]. In another document, however, a much more general goal is adduced, namely of providing 'a comprehensive statistical framework for monitoring the digital divide'[1]. In yet another source, the DOI is described by the ITU as a composite index that measures 'digital opportunity'[2]. In the absence thus, of any clear definition, it may well seem easier to advance the notion of equal weights without any analytical justification or scrutiny. Indeed, in the absence of clear goals, one weighting system tends to be as good as any another (and equal weights seem somehow to be the most intuitively plausible).

Yet, when one views the DOI methodology from an ends-oriented point of view, equal weighting turns out to be based on an extreme view of the relationship between potential and actual variables. Following Sen[5], the former variables can be thought of as purely contingent with a highly variable relationship to the actual ends or goals that are being sought. In the context of welfare economics, for example, Sen argues that what matters to well-being is not the presence of or access to particular goods and services, but rather, what is actually done with them (which determines actual 'functionings', in his terminology). The relationship between commodities (the inputs) and functionings (the goals) may, as Sen emphasizes, be weak or even non-existent, as would be the case, for example, of someone with affordable access to a computer but lacking the ability to do so. Looking at the DOI from this point of view brings out one of its major weaknesses, namely that by adding means and ends-based variables in a single composite index, the myriad ways in which the rate at which the former variables can be more effectively converted into the latter, are ignored. In fact, the DOI can be made to increase by raising the percentage of the population under the mobile footprint, even if this results in no improvement in utilization whatsoever.

If, instead, one were to adopt the view held by a number of those who write on IT and development – that only *current* utilization of the technology matters, not only would this possibility be ruled out, but also, and more dramatically, a large number (more than 30) of developing and several other countries in the DOI rankings would score zero on the index, and many others would be bunched at a negligible value of 0.01 (since these are the utilization scores, unaccompanied by the values of the other two clusters). In effect, this represents a weighting scheme where all other variables apart from utilization (the closest to an ends-based measure) receive a value of zero. It represents, in fact, the polar opposite of the DOI weighting system, if one rules out the possibility of giving *higher* weights to potential than actual variables, as opposed to weighting them equally.

The contrast between these two extremes becomes clearer when they are applied to particular countries in the DOI ranking. Table 2, for example, contains a comparison between two African countries, Botswana and Senegal.

Under the system of equal weights, Botswana obtains a DOI score of 0.35 compared to 0.30 for Senegal, which although performing relatively poorly on opportunity and infrastructure, has a rate of utilization that is 14 times higher than that of Botswana. If, by contrast, only the ends-based measure, current utilization, was allowed to count, the DOI for Senegal would be 14 times higher than that of Botswana.

As noted above, while an extreme ends-based perspective would adopt the second outcome (favouring Senegal rather than Botswana), other considerations argue in favour of a less extreme version of this position. It might be the case, for example, that the potential variables are considered to be important in their own right; that they might be more or less successfully converted into actual achievements at some future date and that the rate of discount be-

Table 2. Comparison between two countries under extreme weighting systems

Country	Opportunity	Infrastructure	Utilization	DOI
DOI weights				
Botswana	0.92	0.12	0.01	0.35
Senegal	0.72	0.06	0.14	0.30
Only utilization counts				
Botswana	0	0	0.01	0.01
Senegal	0	0	0.14	0.14

Source: ITU[2]; Table 1.

Table 3. Ranking reversal of two selected countries

	Opportunity	Infrastructure	Utilization	DOI
Antigua	0.94	0.37	0.05	0.45
Brazil	0.87	0.24	0.16	0.42

Table 4. Absolute values of the three components of the DOI by geographical area (2004/5)

	Opportunity	Infrastructure	Utilization	DOI
World	0.77	0.23	0.11	0.37
Africa	0.52	0.06	0.02	0.20
Americas	0.86	0.23	0.12	0.40
Asia	0.81	0.23	0.10	0.38
Europe	0.97	0.46	0.22	0.55
Oceania	0.71	0.21	0.09	0.33

Source: ITU[2].

tween the present and the future is such as to favour the latter (or, in more technical terms, that there is a relatively high social discount rate). My intention in this article is not to answer these important questions (which, in any event, will vary from one case to another), but rather to emphasize that they should be taken seriously and made explicit in the choice of weights for any index that is likely to influence policy-making (as is certainly the case with the DOI). By the same token, policy-makers are entitled to a no less analytical justification for the use of equal weights in the DOI, though, as noted above, no such analytical foundation is actually presented. Nor, apparently, is there thought to be a need to apply different sets of weights to the various components of the index, in order to assess the degree to which the country rankings might change. Let us, for example, return to the comparison between Botswana and Senegal. As shown in Table 2, the former has a DOI of 0.35 as opposed to 0.30 for the latter. Instead of equal weighting, let us for example now apply weights of $\frac{1}{6}$, $\frac{2}{3}$ and $\frac{3}{6}$ to opportunity, infrastructure and utilization respectively. The result is a reversal of rankings: Senegal now has a higher DOI than Botswana. Or again, compare the rankings of Antigua and Brazil under equal weights and those used in the previous example. Table 3 shows that according to the DOI the former country receives a value of 0.45, as opposed to the latter, which only scores 0.42.

Under the new weighting system $(\frac{1}{6}, \frac{2}{3}, \frac{3}{6})$, the ranking is again reversed because of the higher weight it assigns to the utilization component, which is more than three times higher in Brazil than in Antigua.

Values of clusters to which weights are assigned

Thus far, I have been concerned purely with the issue of how the three components of the DOI should be weighted, as against the way that this actually takes place. I have suggested that on the basis of an ends-based, achievement-oriented framework, the potential variables of the index tend to be overvalued (sometimes substantially) relative to variables that reflect actual achievements (or functionings in Sen's terminology). But what also needs to be recognized is that in deriving actual values of the DOI, the absolute scores that are assigned to each variable cluster also need to be considered. If, somehow these scores are also biased in the same direction as the weights, then the DOI will suffer from multiple bias and become even more of a questionable basis for making policy towards IT. For then, to an even greater extent than before, it will overstate the values of opportunity variables at the expense of variables reflecting actual achievement.

Let us initially, however, simply record the absolute values of opportunity, infrastructure and utilization across the sample, disaggregated by major geographical areas.

Table 4 shows that across all geographical areas, the value of the opportunity component is orders of magnitude higher than the value of utilization, especially in Africa, where the ratio between the two is 25. In one sense, of course, it is inevitable that only a fraction of a potential variable is converted into an actual achievement (not all those who enroll in primary school in Africa, for example, ultimately finish the programme). And it also bears emphasis that the specific variables included in the opportunity and utilization categories did not emerge out of the blue, but were chosen instead on the basis of international agreement (and in particular by the Partnership

Table 5. Variables included in the opportunity and utilization components of the DOI

Opportunity	Utilization
Percentage of population covered by mobile cellular telephony	Proportion of households that used the Internet
Internet access tariffs as a percentage of per capita income	Ratio of fixed broadband subscribers to total Internet subscribers
Mobile cellular tariffs as a percentage of per capita income	Ratio of mobile broadband subscribers to total mobile subscribers

Source: ITU[2].

on Measuring ICT for Development, 2005). Yet, neither of these recognitions is enough to rule out the type of bias mentioned at the beginning of this section, as we shall now seek to demonstrate (focusing, for the sake of argument, only on opportunity and utilization, the variables that most closely represent potential and actual respectively). Indeed, there is no inherent reason why international agreement makes one set of variables analytically more appropriate than another set.

In the interest of expositional simplicity, Table 5 reproduces from Table 1 the variables that are included in the two categories under consideration.

Let us begin by examining the variables that are included under the opportunity column in Table 5, representing 'the basic access and affordability needed to participate in the Information Society in mobile population coverage, Internet access prices and mobile prices'[2]. As shown in Table 5, the access measure is about mobile coverage and more specifically 'the potential usage of telecom services that could be achieved *if users had a mobile phone and subscription*'[2] (emphasis added). Because of rapid growth in recent years, mobile coverage is now estimated to be in the region of 90% of the entire world population and above 50% even in the poorest region, Sub-Saharan Africa. Being under the mobile footprint is thus hardly a demanding index of access to mobile phones, which nevertheless comprises fully one-third of the value of the opportunity cluster. In the case of China, for instance, the value of mobile coverage is 0.8 (reflecting an 80% coverage of the population by mobile telephony), which counts just as much to the overall DOI, as the 7% of the population that uses the Internet (with a value of 0.07). This, plainly, is nonsensical.

I have already seriously questioned the equal weighting of opportunity and utilization variables. In the rest of this section, I present a number of reasons why the bias in weights assigned to opportunity and utilization are compounded by differential values of these clusters that work in the same direction. For one thing, it can plausibly be argued that access to mobile telephony should be measured by a more sensible and demanding index than simply being under the mobile footprint. Would it not be more discriminating, for example, to estimate what percentage of the population covered by a mobile signal, is able, within a certain time or distance, to use an actual phone, be it a payphone or a phone belonging to someone who is prepared to lend it to family and friends? (Or, if this is impossible, to measure the total number of phones under the mobile footprint). In fact, any access measure that is more demanding than simply being covered by a mobile signal, would reduce what currently appears to be an artificial inflation of the value assigned to this variable. Further inflation arises from a fundamental asymmetry in the opportunity index, which, for no apparent reason, includes the Internet in affordability, but not in access. Adding a measure such as the percentage of the population served by Internet providers, would almost certainly lower the value of the opportunity index, because its value generally lies far below those of the three existing components of the index (see the *CIA Handbook* for estimates of the ratio being proposed here).

If, therefore, I believe that the value of the opportunity index turns out to be overstated, I also wish to advance the notion that the utilization index is rather heavily understated (especially from the point of view of developing and certain other transition countries as well). My first reason for thinking as I do about utilization, is the opposite side of the coin regarding the asymmetry between mobile phones and Internet, referred to earlier. For, whereas access to the latter technology is neglected, the former is absent from the *utilization* index, which concerns only the number of Internet users. Adding a measure of mobile use would clearly raise the value of the utilization component, since the constraints on using this form of IT are much less demanding than those required by the Internet (such as literacy, computer capabilities and language skills). The Partnership document that underlies the DOI does make mention of mobile usage as an 'extended core' indicator and even proposes a provisional way of measuring it in the future. But the DOI fails to mention how much the absence of this crucial measure may undervalue the current value of the utilization index.

Even if one looks only at the existing components of this index, it is clear that the inclusion of a technology as advanced as mobile broadband, implies that the vast majority of countries will register low or zero absolute values of the ratio of mobile broadband subscribers to total mobile subscribers. In fact, by the end of 2005, there were only 51 countries in which mobile broadband subscribers could be located. 'Many developing countries have yet to launch broadband mobile networks'[1]. Including a variable that conveys so little information seems to me difficult to justify and the assignment to it, of the same weight granted to Internet use, all the more difficult to understand.

Conclusions

The DOI focuses mainly on the measurement and addition of different categories of variables into an overall ranking of countries. I have argued that the specific way in which this exercise has been conducted is subject to important biases. In addition, the process itself may divert attention from a major policy issue. In particular, the focus on adding variables that are related to one another diverts attention from how to convert the potentially oriented variables more efficiently into variables that reflect actual achievements. For, there may be many cases in which it is preferable from a policy point of view to improve the conversion rates of certain variables than merely to increase their values. There is already some evidence, for example, that the way in which infrastructure is used matters far more to growth than increased infrastructural investment.

1. ITU, World Information Society Report, Geneva, 2006.
2. ITU, *The Digital Opportunity Index: A User's Guide*, Digital Opportunity Forum, 31 August – 1 September, Seoul, Republic of Korea, 2006 (available at www.itu.int/digitalopportunity).
3. Streeten, P. *et al.*, *First Things First*, Oxford University Press, USA, 1977.
4. Minges, M., *The Digital Opportunity Index as a Tool for Policy Analysis*, Digital Opportunity Forum, 31 August – 1 September, Seoul Republic of Korea, 2006 (available at www.itu.int/osg/spu/digitalbridges/forum/phtml).
5. Sen, A., *Commodities and Capabilities*, North-Holland, Amsterdam, 1985.
6. UN, Partnership on Measuring ICT for Development. *Core ICT Indicators*, 2005.

Received 6 October 2006; revised accepted 20 November 2006