

**Building and Running an Effective Policy Index:
Lessons from the Commitment to Development Index**

By David Roodman

March 2006

Stable URL: <http://www.cgdev.org/content/publications/detail/6661>



ABSTRACT

Indexes, which distill large amounts of information into a few numbers, appear to be gaining popularity among policy advocates and researchers. This essay tells the story of one such, the Commitment to Development Index, and draws lessons for others who would also devise indexes. The CDI rates 21 rich countries on how much their government policies help or hurt poorer nations. It has been reasonably successful on the goals of raising awareness of certain ideas, in particular that helping is about more than aid, and embodying and communicating the mission of an institution, the Center for Global Development. Among the lessons: To work well, policy indexes must combine humility with a clear sense of purpose. They must incorporate judicious trade-offs between considerations that range from philosophy to mathematics to science to communications strategy. Outreach is at least as important as design. The Web in particular allows outreach to serve multiple audiences, with detail for those who want it, and overviews for those who do not. Perhaps the greatest design challenge is the tension between the desire for simplicity, and the complexity of policy, which can turn an index into a black box. A suite of materials, paper and electronic, written at different levels of detail, may not eliminate this problem, but can reduce it to the point where the index can have an impact.

The Center for Global Development is an independent think tank that works to reduce global poverty and inequality through rigorous research and active engagement with the policy community. This Essay and the Commitment to Development Index were made possible by funding from the Rockefeller Foundation. In late 2005, the Center for Global Development formed a consortium of governments to support the CDI, which now includes Australia, Canada, Denmark, Finland, France, the Netherlands, Norway, Sweden, Switzerland, and the United Kingdom.

Use and dissemination of this Essay is encouraged, however reproduced copies may not be used for commercial purposes. Further usage is permitted under the terms of the Creative Commons License. The views expressed in this paper are those of the author and should not be attributed to the directors or funders of the Center for Global Development.

Building and Running an Effective Policy Index: Lessons from the Commitment to Development Index¹

David Roodman
Center for Global Development
March 2006

Abstract: Indexes, which distill large amounts of information into a few numbers, appear to be gaining popularity among policy advocates and researchers. This essay tells the story of one such, the Commitment to Development Index, and draws lessons for others who would also devise indexes. The CDI rates 21 rich countries on how much their government policies help or hurt poorer nations. It has been reasonably successful on the goals of raising awareness of certain ideas, in particular that helping is about more than aid, and embodying and communicating the mission of an institution, the Center for Global Development. Among the lessons: To work well, policy indexes must combine humility with a clear sense of purpose. They must incorporate judicious trade-offs between considerations that range from philosophy to mathematics to science to communications strategy. Outreach is at least as important as design. The Web in particular allows outreach to serve multiple audiences, with detail for those who want it, and overviews for those who do not. Perhaps the greatest design challenge is the tension between the desire for simplicity, and the complexity of policy, which can turn an index into a black box. A suite of materials, paper and electronic, written at different levels of detail may not eliminate this problem, but can reduce it to the point where the index can have an impact.

¹ Nancy Birdsall and Moisés Naím conceived of the Commitment to Development Index over lunch in February 2002. Many people have contributed to the CDI project since then, including Alicia Bannon, William Cline, Sarah Dean, William Easterly, Sheila Herrling, Adriana Lins de Albuquerque, Tony Kopetchny, Sarah Lucas, Lawrence MacDonald, Keith Maskus, Gregory Michaelidis, Theodore Moran, Michael O'Hanlon, Yvonne Siu, Scott Standley, Andrew Stober, and John Williamson; Mike Boyer, Travis Daub, Will Dobson, Carlos Lozada, and Laura Peterson of *Foreign Policy*; Jim Cashel and Nam Ho Park of Forum One Communications; Desmond Spruijt of Mapping Worlds; Jeanne Batalova, Elizabeth Grieco, and Kim Hamilton of the Migration Policy Institute; and Amy Cassara and Daniel Prager of the World Resources Institute.

Indexes, which distill large amounts of information into a few numbers, appear to be gaining popularity among policy advocates and researchers. One reason for the trend is the evident success of such examples as the U.N. Development Programme's Human Development Index, Freedom House's political and civil rights variables, and Transparency International's Corruption Perceptions Index in penetrating a competitive media environment and altering how people see the world. The Center for Global Development (CGD), a think tank founded in late 2001 and based in Washington, DC, decided in its early months to create an index too, in cooperation with *Foreign Policy* magazine. The CDI project has been reasonably successful, and perhaps as a result I have been asked many times for advice on how to build and run an effective policy index.

The CDI, like CGD, focuses on how rich-country governments and multilateral institutions affect developing countries. It rates and ranks 21 rich countries on how much their policies in seven areas help or hurt poorer countries—not just foreign aid policy, which gets the most attention in this context, but also policies relating to trade, migration, investment, the environment, military affairs, and technology.² The CDI asks such questions as: How open are countries' borders to goods and workers from poorer countries? How supportive are their governments of healthy investment in developing countries? How active and constructive are they in international security affairs such as peacekeeping? Do their policies impede the flow of new ideas across borders? (For more information, visit www.cgdev.org/cdi.) Since its public launch in 2003, CGD has revised and updated the CDI annually. As with many indexes, its purpose is not only to measure, but to use ranking to draw attention to issues, educate the public and policymakers, stimulate thinking and debate, and serve as a flagship for an institution.

Index-making can be hard to do well—how hard depends in part on the concept to be crystallized. Tricky issues that arise include clear definition of the concept to measure, the relative weight various components deserve, trade-offs between complexity and realism, and the tension between improving the index over time and maintaining comparability with past results, and more. This paper tells the CDI story and draws lessons that may be useful for others devising indexes.

Designing the CDI

As indexes go, the CDI is ambitious. It strives to measure the quality of a huge range of policies with respect to a broad and poorly defined outcome, development in poorer countries, despite major data gaps and limited understanding of how various policies actually affect development. This gives the CDI an edgy and paradoxical character. On the one hand, what it sets out to do is arguably impossible. On the other, CGD believes that it is worth doing, given the purposes already listed. And given that it is to be built, it needs to be built in a way that is intellectually defensible, since CGD wants to be respected as a research institution. The challenge in designing the CDI, then, was to do something that is analytically impossible in a way that is analytically credible.

² Nancy Birdsall and David Roodman, "The Commitment to Development Index: A Scorecard of Rich-Country Policies," Center for Global Development (CGD), April 2003; David Roodman, "The Commitment to Development Index: 2004 Edition," CGD, April 2004; idem, "The Commitment to Development Index: 2005 Edition," CGD, August 2005.

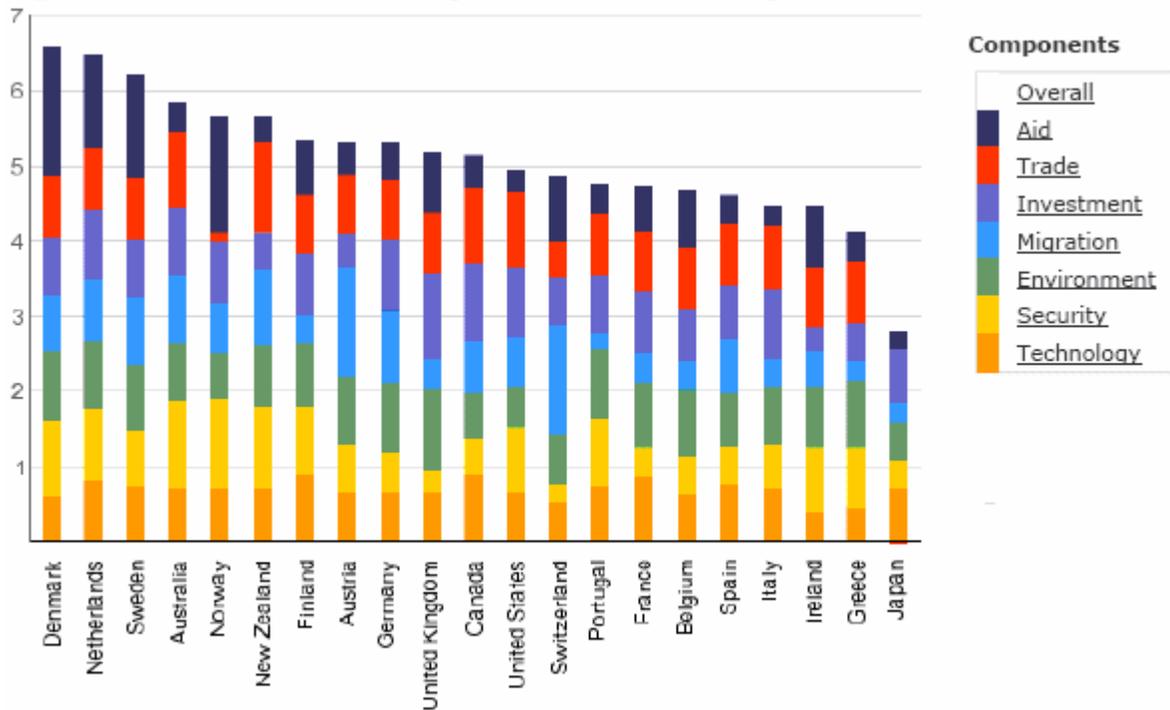
The CDI designers have often had to make tricky choices about how far to go beyond the evidence in order to educate the public. For example, there is no doubt that South-North migration is having a major impact on some developing countries. So as a matter of public education, it is imperative that the CDI cover migration policies. Yet the overall effect of migration policies—even the sign sometimes—is unclear. Is it good for El Salvador that many of its most entrepreneurial people leave for the United States—but send home money? The result, after a lot of internal discussion and external consultation, is a migration component that is reasonable—it rewards all immigration from developing countries, but unskilled immigration more so—but debatable. But debate, after all, is one desideratum of the CDI project.

The CDI has one component for each major policy area covered. Originally, there were six: aid, trade, investment, migration, environment, and peacekeeping. In the second edition, the peacekeeping component was broadened and renamed “security,” and a seventh component on technology was added. The subject population for the index is the membership of the Development Assistance Committee, the official donor club, sans Luxembourg, because it is so small. Each of the 21 countries receives a score in each of the seven areas, and overall scores are simple averages of component scores. The index aims to measure not the absolute impact of individual rich countries, but the degree to which they are realizing their potential to help. In other words, it controls for size. Thus small countries such as Denmark and the Netherlands can score high despite modest aggregate impact. (See Table 1 and Figure 1.)

Table 1. Commitment to Development Index 2005: Table

Rank	Country	Aid	Trade	Investment	Migration	Environment	Security	Technology	Average
1	Denmark	12.3	5.9	5.5	5.3	6.5	7.2	4.4	6.7
2	Netherlands	8.7	5.9	6.8	5.7	6.6	6.8	5.8	6.6
3	Sweden	9.8	5.8	5.5	6.4	6.4	5.2	5.3	6.4
4	Australia	2.5	7.3	6.5	6.5	5.4	8.5	5.0	6.0
5	Norway	10.8	1.0	5.8	4.9	4.2	8.5	5.2	5.8
5	New Zealand	2.1	8.8	3.4	7.1	5.9	7.8	5.1	5.8
7	Finland	4.9	5.8	5.9	2.5	6.1	6.6	6.3	5.4
7	Austria	3.0	5.8	3.0	10.5	6.5	4.7	4.6	5.4
7	Germany	3.4	5.7	6.7	6.8	6.7	3.8	4.7	5.4
10	United Kingdom	5.6	5.9	8.1	2.8	7.9	2.1	4.6	5.3
10	Canada	2.9	7.3	7.6	4.9	4.3	3.5	6.3	5.3
12	United States	1.9	7.2	6.7	4.7	4.0	6.2	4.7	5.0
13	Switzerland	6.0	3.3	4.6	10.5	4.7	1.6	3.8	4.9
13	Portugal	2.8	5.9	5.5	1.4	6.7	6.4	5.3	4.9
15	France	4.1	5.8	6.0	2.7	6.2	2.8	6.2	4.8
15	Belgium	5.4	5.8	5.0	2.9	6.3	3.6	4.5	4.8
17	Spain	2.6	5.8	5.2	5.1	5.1	3.6	5.4	4.7
18	Italy	1.6	6.1	6.8	2.5	5.4	4.1	5.3	4.5
18	Ireland	5.6	5.6	2.5	3.2	5.9	6.1	2.8	4.5
20	Greece	2.8	5.8	3.6	1.8	6.3	5.8	3.2	4.2
21	Japan	1.4	-0.2	5.1	1.8	3.7	2.8	5.0	2.8

Figure 1. Commitment to Development Index 2005: Graph



The broad topical coverage necessitated going outside of CGD for collaborators for most components. But our Washington location meant that we usually did not need to search far. Washington is for policy analysis what Detroit once was for cars, a place where producers congregate, compete, and interact intimately with suppliers and “customers.” CDI collaborators have included scholars at the Brookings Institution (Michael O’Hanlon and Adriana Lins de Albuquerque), Georgetown University (Theodore Moran and B. Lindsay Lowell), the Migration Policy Institute (Jeanne Batalova, Elizabeth Grieco, and Kim Hamilton), and the World Resources Institute (Amy Cassara and Daniel Prager), as well as one institution more than two miles from CGD, the University of Colorado at Boulder (Keith Maskus). The most important collaborator, *Foreign Policy* magazine, is published by the Carnegie Endowment for International Peace across the street.

In each annual design cycle, collaborators had to choose indicators available for all 21 countries that could be collected on a relatively low budget (for the whole project, about \$200,000/year) and could be defensibly included in the index. Their written rationales were sent to colleagues for review, and after revisions, CGD made final design decisions. The design team members, especially those at CGD making the final calls, faced many challenging design issues:

Lack of a theoretical model. There is no overall model for the development process and the role of rich-country policies within it. No overarching theory describes how aid, trade, environmental, and other policies affect the development in various parts of the world, which matter most, nor how they interact. As a result, for example, the top-level structure of the index—the simple averaging of seven component scores—is atheoretical.

More deeply, the evaluation of policy is sometimes a matter of political philosophy, or deep ideas about how the world works. One person suggested that the CDI reward every dollar of U.S. defense spending because the United States, it was argued, is the sole hegemon, the guarantor of global stability and protector of democracy. Another has submitted that every dollar of U.S. defense spending should be penalized. CGD's response to such fundamental disputes has often been to seek common ground in order to earn credibility with the largest possible audience. Thus the security component takes no stand on defense spending generally, and, notably, it is neutral on the invasion of Iraq, in the sense that this is neither rewarded nor penalized. (But this lack of reward does not look neutral to supporters of the invasion.) It *does* reward participation in peacekeeping and warmaking with an official international mandate, such as the NATO-led war in Kosovo and the UN-endorsed Australian intervention in East Timor—but not the Iraq invasion, which lacked such a mandate. Of course, every design choice in the CDI implies a stand, each one no doubt with its dissenters; and the decision to take these stands must reflect to some extent the biases of the designers.

Scaling and weighting. A substantial set of issues, again as with most indexes, surrounds how to scale and weight scores on various indicators and composites thereof. The principle driving our decisions was that in the face of ignorance about the “true” parameters, not to mention the functional form of the development “production function,” it is best to be transparent and minimally arbitrary to avoid seeming biased.

With regard to normalization of scores, one early question was how to present the results. Should they be numbers? Or letter grades like those used in American universities? Should countries be grouped, or strictly ordered? We chose to present numerical scores and to create a full ranking because doing otherwise would have added another level of inherently arbitrary processing and distract from the real issues, and because grouping would have dulled the provocative effect of ranking.

There remained the question of how to normalize readings on various indicators onto a common scale in order to combine them. In general, there are four axioms it would be nice for any normalization system to satisfy: 1) normalized scores should fall in an intuitive range such as 0–10; 2) they should have the same average, say 5, so that, for example, 6 reliably means “above average”; 3) they should have the same standard deviation, like z scores; and 4) 0 should map to 0, so that, for example, a country that gives no aid cannot earn an aid score of 3 or –2. On indicators of “bads” such as trade barriers, axiom 4 would analogously require that a complete absence of the thing scored corresponds to an intuitive maximum such as 10. If one also requires for simplicity that normalizations are linear, then one has two degrees of freedom per indicator, not enough to ensure that all four axioms are satisfied.

As a result, index designers must decide which axioms matter most. In the first edition of the CDI, we favored 1 and 4. After, we switched to 2 and 4, with few exceptions. Specifically, most indicators are now normalized by first dividing by the 21-country average, then multiplying by 5. For “bad” indicators such as of pollution, where a lower number is good, normalized scores are then subtracted from 10. This guarantees that normalized scores average 5 (axiom 2), and that 0 maps to 0 for “good” indicators and to 10 for “bad” ones (axiom 4). Since countries that are average on a “good” indicator get a 5, those that are twice the average get a 10, and those doing

even better score above 10. Similarly, a country twice as bad as average on a “bad” indicator gets a 0—and one even worse gets a negative score. Thus the intuitive 0–10 scale (axiom 1) is sometimes violated, which can confuse audiences. The benefit is the elimination of confusing situations such as that of Switzerland in the first CDI. Because all scores were forced into the 0–10 range (axiom 1) then, Switzerland scored higher on trade than aid (4.0 versus 3.3) but was actually below-average on trade and above average on aid.

It is worth emphasizing that almost none of the CDI indicators is normalized relative to some ideal score. It has been suggested that the measure of aid quantity be scaled against the ideal of 0.7% of GDP, a rate of giving embodied in several official international documents. But that number is arbitrary and there are few comparable ideals for indicators of migration, security, technology, and other policies. As a result, the only indicators that are normalized relative to ideal maxima are binary ones on whether countries have ratified the Kyoto Protocol and certain other international treaties.

Intimately related to the scaling issues are ones of weighting. Indeed, changing the normalization of an indicator, to the extent that it affects the dispersion of normalized scores, is equivalent to changing the weight on that indicator. Nevertheless, scaling and weighting are distinct notions in the minds of most readers and raise distinct issues. Chief among these is whether the “equal weighting” of the CDI—the simple averaging of its seven components—appropriate. Attacking the equal weighting, Robert Picciotto, director of the Global Policy Project, wrote cleverly, “To be approximately right is better than to be precisely wrong.”³ However, there has been no consensus among those opposed to equal-weighting the CDI on the right weights. Some have favored aid and trade, others want extra weight for migration since total earnings gains for émigrés dwarfs aggregate aid and trade flows, others point to the tremendous importance of climate change and technological innovation, and so on.

Shyamal Chowdhury and Lyn Squire take an interesting approach to the weighting question, surveying development experts in the global North and South to elicit their preferred weights.⁴ In their results, average preferred weights on several of the components are statistically different from equal weighting. However, the differences are not great and the correlation between the standard CDI and the re-weighted one is 0.992.

This argues for the minimally arbitrary choice of simple averaging in the CDI. And there are additional, powerful arguments. One is that for all the debate over equal weighting, it is not a well-defined concept as long as scalings have an atheoretical component that affects score dispersion, since increasing the dispersion of normalized scores on an indicator effectively increases its weight. Consider: the CDI changed its scaling system between 2003 and 2004, which changed the standard deviations of normalized scores on various components, thus their effective weights. Yet the CDI was a simple average of component scores both years. So which year truly gave equal weight to the seven policy areas? On reflection, equal weighting is essentially a chimerical notion. What one *could* ask is that *any two CDI-measured policy changes in a given*

³ Robert Picciotto, “Giving Weight to the CGD Rankings: A Comment on the Commitment to Development Index,” Global Policy Project, 2003, <http://www.globalpolicyproject.org/indexComment.html>.

⁴ Shyamal Chowdhury and Lyn Squire, “Setting Weights for Aggregate Indices: An Application to the Commitment to Development Index and Human Development Index,” *Journal of Development Studies*, forthcoming.

country that have an equal effect on development have an equal effect on the CDI. But “equal effect on development” is not well-defined since development is a process of societal transformation affecting billions of people in many ways that are fundamentally incomparable. So in a deep sense, there is no such thing as equal weighting.

A final argument for simple-averaging is presentational. Simple-averaging at the top level of the CDI makes the score table easier to understand than weighted averaging would, for those not mathematically inclined. Moreover, if the average is weighted, it is difficult to represent CDI results in the iconic stacked bar graph in Figure 1.

The “black box problem.” Another issue the design group faced is the tradeoff between precision and transparency. The CDI is sometimes cast as a counterpart of the Human Development Index (HDI), but the two are quite different. The HDI measures an outcome, development, with a small collection of intuitive indicators such as GDP/capita and life expectancy; this makes the HDI simple. The CDI measures policies, which are complex and diverse in themselves and have equally complex and diverse relationships with development outcomes around the world. The CDI evolved substantially in response to comments during its first three editions, and many of these comments pointed out things the CDI excluded or important distinctions it did not make. Thus there is an inherent tendency toward complexity in policy indexes.

Moreover, theory sometimes argues for formulas that are simple in concept but complex in practice. For example, the CDI’s foreign aid component assesses both quantity and quality of aid. The initial design draft, by William Easterly, ranked donors on aid/GDP and a few quality indicators, then took an average of the ranks. This was easy to understand but contrary to sensible theory since quantity and quality ought to interact multiplicatively, not additively. Otherwise, a donor that gave a penny of high-quality aid could outrank one that gave \$10 billion in medium-quality aid. The current aid component design works quite differently. Each quantum of aid that a donor gives a country is discounted for quality factors such as the apparent appropriateness of the recipient for aid (a “selectivity” weight based on poverty and governance quality); then the discounted quanta are summed for each donor. The result is an aid component that is more theoretically sound—but also opaque, as it involves thousands of calculations in a custom database.

We have found some ways to minimize this “black box problem.” One is to keep the top-level structure of the index, which receives the vast majority of the attention, simple. This buttresses an earlier argument for simple averaging. Another is to make available to the reader plain-language summaries of what the components reward and penalize, as well as country reports that summarize the sources of each nation’s performance. Last is to fully document the calculations, in technical papers and, where possible, in spreadsheets. In our experience, these steps allow the index promulgator to limit, though certainly not eliminate, the trade-off between precision and transparency.

Comparability over time. CDI design has been a public learning exercise. Each edition has provoked commentary that influenced the subsequent one. After three design cycles, the pace of change has slowed markedly, so the index now seems to have reached a kind of maturity. Nevertheless, major year-to-year methodological changes posed a transitional communications

challenge because changes in measurement have dwarfed changes in what is measured. Between the first and second editions seven countries saw their ranks change by more than 11, out of 21. Some asked how the index could be taken seriously as a policy metric if readings jumped that much from year to year. (Most of these critics also pointed to what they saw as additional design problems, which pointed to the need for more methodological changes!)

In response, outreach for the third edition gave much more prominence to back-calculations of the latest methodology to previous years. These are the basis for multi-year graphs on the web site and multi-year tables in *Foreign Policy*, and are included in the public spreadsheet. Notably, these back-calculations use the same parameters in normalizing scores for all indicators, in order to make inter-temporal comparisons meaningful. That is, scores on indicators are normalized in all years so that those for 2003, the first year, average exactly 5. With these constant scalings, the average overall CDI score rose to 5.1 in 2004 and 5.2 in 2005, indicating slow improvement over time. As for communications strategy, CGD put its best foot forward, explaining that it was proud of the improvements that had been made. On balance, we feel that the benefits of public learning—a more credible index—will outweigh the costs provided that the index methodology stabilizes in due course.

Data availability. With a small budget, the CDI project must rely primarily on standard data sources such as the World Bank and Organisation for Economic Co-operation and Development (OECD). Where the design team does go beyond such sources to collect information country by country, costs rise by several orders of magnitude. As a result, what is visible to the index designers is substantially determined by international organizations and the rich-country governments that finance them. Where data are relatively good, such as for foreign aid and trade, the CDI can be most sophisticated. Where the data are poor, as with migration, the CDI is coarser. We hope that by highlighting data gaps, the CDI is stimulating international organizations to close them.

Sensitivity analysis. Sensitivity analysis should be a matter of course when developing composite indexes in a research context. But in a communications context, in a sound bite world, it can undermine the credibility of an index. And the CDI is above all a communications tool. Moreover, in an index so ambitious, one should take as given that results depend significantly on assumptions, including assumptions that go quite deep into political philosophy. Accepting this, sensitivity analysis will not be very enlightening. Finally, there are easily a hundred parameters to test in the CDI; analyzing the importance of all of them would quickly overwhelm a limited budget and human cognition.

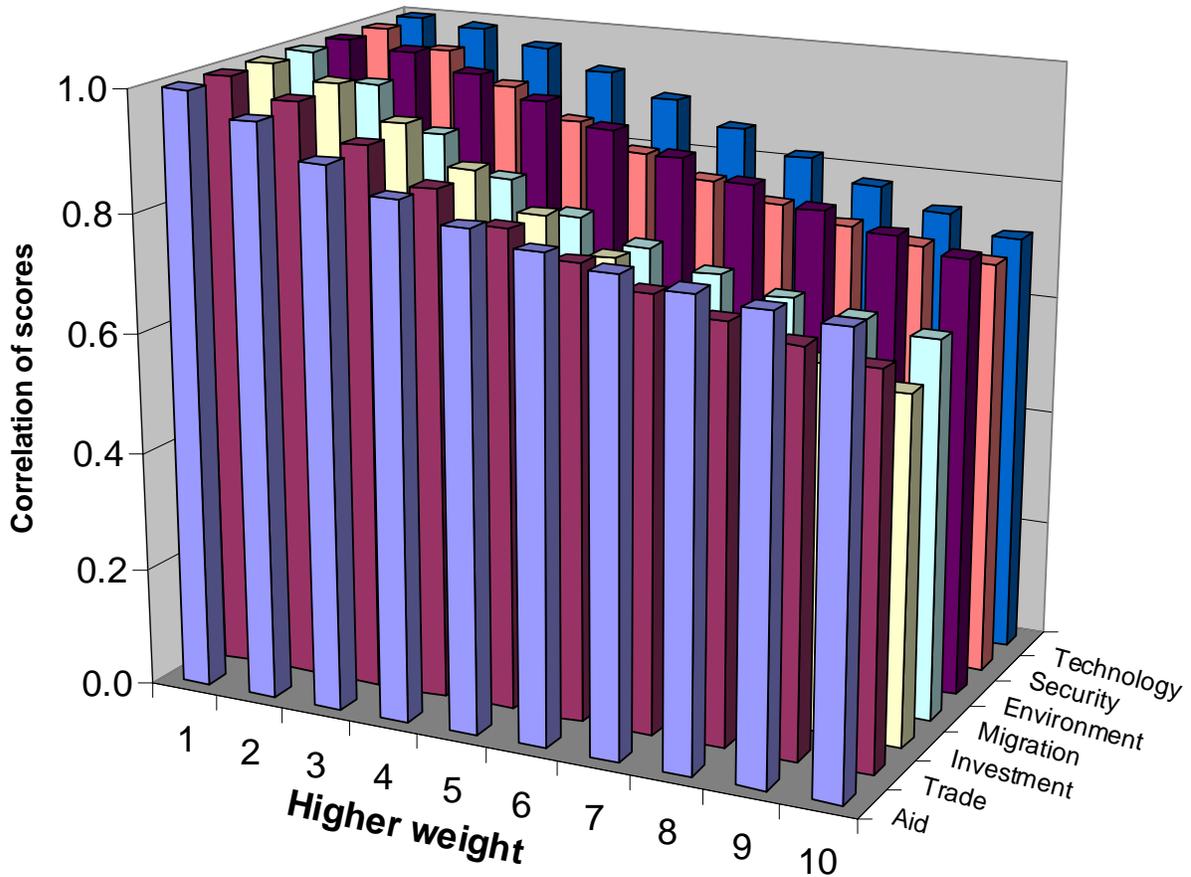
For these reasons, the CDI designers have not analyzed the sensitivity of the index to a great extent. Put simply, it would not greatly advance the goals of the project. An exception has been an analysis of the sensitivity of the overall results to deviations from “equal weighting” of the seven components. As reported in the overall CDI technical paper⁵, 63 non-standard versions of the CDI are generated: first with the weight on aid raised to 2, then 3, and so on up to 10, while weights on the other components were held at 1, then the same for trade, and then the other components.⁶ Figure 2 shows the correlations between the official overall 2005 CDI scores and

⁵ Roodman (2005), op. cit.

⁶ My colleague Michael Clemens suggested these tests.

the modified ones. For all the components, tenfold overweighting yields a score correlation of 0.58–0.75 with the original. Whether this constitutes robustness is in the eye of the beholder.

Figure 2. Correlation of standard CDI with versions with higher weight placed on one component, 2005



Outreach

The CDI aims to bring certain facts and ideas to new audiences. So outreach is central to the project. The CDI outreach approach has been multi-pronged, involving print publications, a web site, and in-person events.

The starting point for outreach was the *Foreign Policy* article, which gave the CDI instant stature and continues to bring the index to a large and prime audience. The stable of CDI publications has expanded from there over the years and can be thought of as being organized in a pyramid from the most detailed and technical to the simplest and most accessible. At the base are the background technical papers for each component. Built upon them is the overall technical paper, and the public spreadsheet that includes most of the calculations. All these are essential for transparency. Then, in a large step up, is an accessible brief that summarizes the methodology in a few pages, and reports for each country. At the highest level of generality and accessibility are the *Foreign Policy* article and a postcard that features a graph like Figure 1 and key messages.

These include that aid is about more than quantity (quality matters too), that helping is about more than aid, and that all countries can do better. [Do you want to mention the poster?]

We have increasingly come to appreciate and exploit the special strengths of the Web as an outreach tool for composite indexes. The CDI has an elegant top-level structure in three dimensions—time, policy areas, and countries—and CGD's CDI web site, developed with Forum One Communications, allows visitors to switch between views along the various dimensions with a click. It also accommodates the spectrum from accessibility to specifics by allowing visitors to drill down to detail. It includes all the resources listed above except the magazine article, which is on the *Foreign Policy* site, and additional content to explicate the results, including a Flash-based animated map developed by Mapping Worlds in Amsterdam. These web-based applications have an important cost advantage: they are available around the clock, around the world, and the marginal cost of reaching people is effectively zero.

At the other end of the spectrum in cost per person reached have been events in Washington, DC, to publicize the index among the many locally based journalists and development professionals. Moderated panel discussions, kick off with a presentation of the latest CDI, and help us connect with officials at the U.S. Treasury, U.S. aid agencies, the World Bank, IMF, and Inter-American Development Bank, rich-country embassies, and other institutions. Notable panelists have included Wesley Clark, then a candidate for U.S. President, Gene Sperling, former advisor to Bill Clinton, and Angel Gurría, who has since taken the helm at the OECD. Similarly, press conferences or teleconferences raise our profile with media outlets around the world, many of which have reporters in Washington. In order to raise our international profile further, we intend to run such events in other rich-country capitals, such as Brussels and Tokyo, in conjunction with local partners.

We have found that the CDI and its outreach machinery, once built, are a useful platform for rapid-response outreach around current events. For example, within days of the Asian tsunami and the subsequent debate over whether rich countries are stingy, CGD issued a press release drawing specifics from the CDI on how much government and private charitable aid each rich country gives, and reminding the public that true generosity is about more than aid. The release led to a lot of media interest in the work of CGD. Similarly, work on the trade component positioned CGD to provide specifics on levels of trade barriers against developing-country exports in the run-up to the Hong Kong trade conference in December 2005.

Impacts

The ranking device has worked well for CGD and *Foreign Policy*. The CDI has been cited in the *Economist*, the *New York Times*, the *Washington Post*, the *Los Angeles Times*, the *Financial Times*, and many other newspapers worldwide via major wire services. These have led to radio and television interviews from Denmark to Australia. University professors are teaching from the CDI—how many, we do not know. The CDI is probably what CGD is best known for, so it is serving its intended purpose of introducing people to the institution, its mission, and its work.

The most engaged audience has been officials at bilateral aid agencies. They are the people most responsible for thinking about how their governments' policies affect developing countries. The

Dutch and Finnish governments, as part of their efforts to frame development policy as being about more than foreign aid, have adopted the CDI as an official metric of development policy performance. The Australian government commissioned two high-quality papers critiquing the CDI in its first year.⁷ The CDI has influenced development policy white papers in Australia, Canada, Finland, and Norway. And aid agencies officials from many of the CDI countries have provided written or oral comments informally, or have asked questions as they report to their ministers on their country's CDI performance.

The strongest response has come from Japan, a country that once took pride in being the world's largest donor (the U.S. has since surpassed it). Japan finishes firmly in last place on the CDI because of its tight barriers to workers and goods from developing countries, minimal contributions to peacekeeping, and an aid program that is actually modest for the country's size. In general, it seems that Japanese are proud of their economic accomplishments over the last 60 years, and proud of having made them in a distinctive Japanese way. Yet along with the understandable pride comes, somewhat paradoxically, a sensitivity to how the rest of the world sees their country. For many Japanese, who see their country as leading the economic expansion in Asia, and as a moral exemplar with its peace constitution, the CDI seems implausible. Thus the strong Japanese reaction. Yet even most Japanese recognize that Japan is more closed than the other 20 rated countries. And a core belief embedded in the CDI is that openness in rich countries is good for poorer ones. Certainly, the ability to export to the big U.S. market helped Japan develop rapidly.

In direct response to the CDI, the director-general of the Economic Cooperation Bureau of Japan's Ministry of Foreign Affairs wrote an opinion piece in 2003 in *Asahi Shimbun*, a leading Japanese newspaper, out of fear of the "misunderstandings that may result from the publication of the ranking worked out by a well-known think tank in an authoritative political journal."⁸ It has been a hot topic in an e-mail discussion group involving a thousand or more Japanese aid officials around the globe. In my experience, Japanese have been quite critical of the CDI, but nearly always polite and constructive in discussing it. Some have argued that rather than creating an incentive for improvement, the CDI is undermining support in Japan for foreign aid by casting the country as a hopeless failure.

This raises the interesting empirical question of when policy indexes create incentives for reform, and when they backfire. Prime Minister Junichiro Koizumi's plan to help the poorest countries reap more benefits from international trade, announced in just before the Hong Kong trade ministerial in December 2005, offers hope that even in Japan, external criticism of development policy can inspire constructive reform.⁹

Lessons

The challenges and successes of the Commitment to Development Index offer lessons for others who would devise indexes.

⁷ Mark McGillivray, "Commitment To Development Index: A Critical Appraisal," AusAID, November 2003; Ian Castles, "Evaluation of Donor Performance Monitoring Initiatives," AusAID, February 2004.

⁸ Hajime Furuta, "ODA Report Unfairly Portrays Japan's Efforts," *Asahi Shimbun*, September 8, 2003.

⁹ Junichiro Koizumi, "Joint Effort is Needed to Eradicate Poverty," *Financial Times*, December 11 2005.

Clarity of concept and purpose is essential. It is important to understand from the start the concept being measured and the ultimate goals of the index project. In the case of the CDI, goals include spurring research and policy reform and building the reputation of a new institution, specifically a reputation for blending serious analysis with practical creativity. These goals influence the details of the index design. Such goals should also be communicated to the audience. In particular, if goals beyond measurement itself are made clear, this helps people take the project in the right spirit and not hold the index itself to the standard of perfection. While remaining aware of the index's defects, one should publicly criticize the index as a measurement tool only enough to demonstrate humility and then explain that it is a means to greater ends. An index that pretends to be more than it is loses credibility.

Big, simple ideas get attention. Transparency International's Corruption Perceptions Index and the U.N. Development Programme's Human Development Index are examples of indexes that not only embody big, easily grasped ideas (corruption, human development as more than money income), but have promoted those ideas in the public consciousness worldwide. They demonstrate that grand indexes can have real impact. Of course, it may be entirely appropriate for a given index project to focus on a narrower idea, such as technology use in business or investment in primary education in Latin America, and aim at a narrower audience.

Top-level accessibility is invaluable for a complex index. A reader who can easily understand the idea and overall structure of an index will feel oriented and more prepared to buy into the whole construct. Most readers will not explore beyond the top level of structure. It is also important for plausibility to summarize in plain language the details that the reader cannot see. This strategy allows for an index that reflects at its roots the complexities of policy while catering to the busy lay reader.

Public learning need not be fatal. CDI designers made substantial improvements in the second and third editions. This understandably engendered criticism because a few countries jumped up or down in the ranking. But it had several benefits. Any change adds interest, and the worst enemy of a communications strategy is boredom. Change can actually add credibility by signaling that the designers do not claim to hold a monopoly on the truth, thus that the ultimate purpose of the index is to provoke, not measure. Finally, it makes for a better index in the long run. Excellence is normally achieved through continual feedback and learning. It is important, however, to guard against misleading inter-temporal comparisons when the methodology is changing substantially. It is probably also best to stabilize the design after a few years.

An attack on an index is a victory if one purpose of the index is to raise awareness. When most people turn to the question of how rich countries affect poorer ones, they think of foreign aid. The Center for Global Development aims to change that, in part through the CDI, by showing people that trade, migration, and other policies matter at least as much, and need reform. We therefore consider most criticism of the CDI to be successes, since it means that people are talking about things we believe deserve more attention. That is better than being ignored.

Conclusion

Composite indexes are useful tools for raising awareness of public policy issues. To work well, they must combine humility with a clear sense of purpose, and incorporate judicious trade-offs between considerations that range from philosophy to mathematics to science to communications strategy. Outreach is at least as important as design. The Web in particular allows outreach to serve multiple audiences, with detail for those who want it, and overviews for those who do not. Perhaps the greatest design challenge is the tension between the desire for simplicity and the complexity of policy, which can lead turn an index into a black box. Proper documentation and outreach can minimize, though not eliminate, this problem. In the end, the test of an index is not whether it is right, but whether it draws people in to the work of institution and a world of ideas.