2

Selecting the Countries, Part I:
The Methodology

A central tenet of the MCA is that aid can be more effective in achieving development goals if it is focused on nations with governments that are committed to establishing policies and institutions conducive to economic growth and poverty reduction. But putting this idea into practice raises an immediate challenge: what is the best way to identify countries that meet this criterion? The selection methodology proposed by the Bush administration is one way to address several issues central to the MCA, including:

- the income levels that will determine the set of eligible countries,

- the specific indicators that will be used to show commitment of countries in the three broad areas of "ruling justly, investing in their people, and encouraging economic freedom,"[1]

- the passing grades on each indicator, and

- the method of aggregating across the indicators to determine which countries qualify.

This chapter examines the administration's proposed methodology, explores the judgments required, and examines some alternatives. The next chapter goes a step further and, using the administration's methodology

---

1. See the president's March 2002 speech at www.whitehouse.gov/news/releases/2002/03/20020314-7.html.

and the most recent available data, determines which countries would qualify during the first three years.

There is no right or wrong methodology since there are multiple ways to choose countries that meet the president's broad guidelines.[2] The administration's proposed approach is reasonable but could be improved by:

■ dropping countries with per capita incomes between $1,435 and $2,975 from MCA eligibility since they have less pressing needs and more options for financing than the poorest countries;

■ adjusting over time the list of 16 indicators by dropping the weakest indicators, strengthening the others, and adding new ones;

■ using, wherever possible, an absolute standard rather than the median as a passing grade on the indicators;

■ reconsidering the requirement that all countries must pass the corruption hurdle; and

■ considering creating an aggregate score for the 16 indicators rather than using medians to determine qualification.

The first section of this chapter summarizes the debate about aid and country selectivity, which provides the conceptual underpinnings of the qualification process. The second section reviews the main points of the administration's proposal. The third section examines the three proposed country groups eligible for MCA funding during the first three years. The fourth section explores each of the proposed 16 indicators and a few alternatives. The fifth section examines different methods of aggregating the 16 indicators to determine which countries ultimately qualify.

Conceptual Underpinnings: Aid and Country Selectivity

The idea that aid works best in countries committed to sound development policy makes intuitive sense: foreign assistance will go much further in countries where governments are dedicated to building better schools and clinics, creating jobs, and rooting out corruption than it will in countries with dishonest or incompetent governments. Foreign assistance was more effective in Korea and Botswana where governments placed a high priority on growth and development than it was in the Philippines under Ferdinand Marcos or Liberia under Samuel Doe. This is not to argue that aid was a major determinant of growth in Korea or Botswana but it prob-

---

2. For an alternative methodology for choosing the MCA countries, written well before the administration made its recent proposals, see Birdsall et al. (2002).

ably had a stronger impact there than in countries with weaker and more corrupt governments.

The idea of country selectivity—in which aid is focused on a select group of countries with good policies and institutions—has gained much currency in aid programs in recent years. This idea is based to a large extent on the research of World Bank economists Craig Burnside, David Dollar, and Paul Collier who show that aid has a positive relationship with growth in countries with good policies and institutions and little or no effect in others (Burnside and Dollar 2000; Collier and Dollar 2002; and World Bank 1998). This idea has influenced the policies of the World Bank (2002a)—which claims that it has become more selective with its loans in recent years (although the evidence is not clear-cut)—and several bilateral donors.

The research on the relationship between aid and economic growth, however, is hardly unanimous in its conclusions. Some studies find a positive relationship between aid and growth, others find no relationship at all, while still others find that aid has a negative impact on growth.[3] To some extent, these ambiguous findings at the macroeconomic level should not be surprising for several reasons. First, even under the best of circumstances, the impact of foreign aid on economic growth is probably small relative to other factors, making it difficult to detect a clear relationship. Second, data underlying cross-country empirical growth studies are inherently weak, with data missing for many countries or for variables that may strongly influence growth. Third, as discussed in chapter 1, large portions of aid historically were given to countries for political and strategic reasons. Even when this aid was allocated nominally for development, it is hardly surprising that it had a limited impact on growth, since that was not its primary objective. Fourth, some types of aid should be expected to be *negatively* correlated with growth, especially humanitarian aid, disaster assistance, and aid provided in response to terms-of-trade shocks, since by definition all are aimed at countries suffering growth slowdowns.

In recent years, the Burnside, Dollar, and Collier studies have come under attack from two directions. Several studies challenge the finding that the positive aid/growth relationship depends on good policies, showing instead that aid is positively correlated with growth (with diminishing returns) regardless of the policy environment.[4] These studies use differing sample sizes and specifications, cover different time periods, and control for a variety of other factors besides aid, thus making comparisons difficult. A recent study by William Easterly, Ross Levine, and David Roodman (2003), built from the Burnside and Dollar data set, did not find any clear relationship between aid and growth. They found that the initial results were not robust to new data points, different time frames, varying defini-

---

3. See, for example, Bauer (1971); Mosley, Hudson, and Horrell (1987); Chenery and Strout (1966); and Boone (1996). For a review, see White (1994).

4. For a brief list of these studies, see table 7.2.

SELECTING THE COUNTRIES, PART I **21**

tions of aid, and alternative definitions of policy. Much more research is needed on different aid interventions and the varying circumstances under which aid might be more or less effective. The Burnside, Dollar, and Collier results, while perhaps not statistically robust, intuitively seem correct. The idea that aid works best in countries with governments committed to growth and development has caught on partly because of the research results and also because it resonates with development specialists of different backgrounds and comports with the experience on the ground in many developing countries.

However, the focus on the quality of policies in the recipient country, though important, takes us only so far in considering aid effectiveness. Other factors strongly influence the impact of aid, including the quality of the bureaucracy in the donor institutions, donor restrictions (e.g., aid earmarked for specific activities or "tied" to purchases in the donor country), reporting requirements, harmonization across donors, and the capacity of the recipient country to manage aid money. Therefore, country selectivity will not be enough for the MCA to be effective—it must be coupled with changes in aid delivery, as discussed in chapters 4 through 6.

The Administration's Proposal

There are five key parts to the proposed selection process. First, the administration proposes rapid expansion of the group of countries eligible to compete for MCA funding during the program's first three years. During the first year (fiscal 2004), countries with per capita incomes below $1,435,[5] which are also eligible to borrow from the World Bank's concessional lending window, the International Development Association (IDA), will be eligible for MCA funding (box 2.1). There are 74 such countries, including all but 7 of the 81 IDA-eligible countries.[6] In the second year, the IDA-eligibility criterion will be dropped, and the remaining 13 countries with per capita incomes below $1,435 will be added to the group, taking the total number of countries to 87. In the third year, the group will be expanded to include all 28 countries with per capita incomes between $1,435

---

5. All these classifications use per capita incomes converted to US dollars with current exchange rates. An alternative would be to compare incomes based on purchasing power parity (PPP). While the latter technique is widely acknowledged by economists as the superior methodology, in practice there are controversies about the accuracy of the underlying data and conversions in specific countries. These data are also not available for many low-income countries. As these data improve in the coming years, they will offer a sounder basis for comparisons of income across countries.

6. The seven excluded IDA countries are Dominica, Grenada, St. Lucia, St. Vincent, Samoa, Tonga, and the Maldives. These countries are part of a group of ten "small island exceptions" that are deemed eligible for IDA credits even though their incomes exceed the current operational cutoff of $875. For more on IDA eligibility requirements, see the World Bank's IDA Web site, www.worldbank.org/ida.

Box 2.1   Countries eligible to compete for MCA funding in the first three years

Year 1: Per capita income less than $1,435 and IDA-eligible

Afghanistan	Congo,	Malawi	Tanzania
Albania	Republic of	Mali	Timor Leste
Angola	Côte d'Ivoire	Mauritania	Togo
Armenia	Djibouti	Moldova	Uganda
Azerbaijan	Eritrea	Mongolia	Uzbekistan
Bangladesh	Ethiopia	Mozambique	Vanuatu
Benin	Gambia, The	Myanmar	Vietnam
Bhutan	Georgia	Nepal	Yemen
Bolivia	Ghana	Nicaragua	Yugoslavia, Federal
Bosnia and	Guinea	Niger	Republic of
Herzegovina	Guinea-Bissau	Nigeria	Zambia
Burkina Faso	Guyana	Pakistan	Zimbabwe
Burundi	Haiti	Papua New Guinea	
Cambodia	Honduras	Rwanda	
Cameroon	India	São Tomé and	
Cape Verde	Indonesia	Príncipe	
Central African	Kenya	Senegal	
Republic	Kiribati	Sierra Leone	
Chad	Kyrgyz Republic	Solomon Islands	
Comoros	Laos	Somalia	
Congo,	Lesotho	Sri Lanka	
Democratic	Liberia	Sudan	
Republic of	Madagascar	Tajikistan	

Year 2: Per capita income less than $1,435

In addition to the 74 countries eligible in year 1:

Belarus	Kazakhstan	Philippines	Turkmenistan
China	Morocco	Swaziland	Ukraine
Ecuador	Paraguay	Syria	West Bank
Equatorial Guinea			

Year 3: Per capita income between $1,435 and $2,975

In addition to the 87 countries eligible in year 2:[1]

Algeria	Fiji	Maldives	South Africa
Belize	Guatemala	Marshall Islands	St. Vincent and
Bulgaria	Iran	Micronesia	the Grenadines
Colombia	Jamaica	Namibia	Suriname
Dominican	Jordan	Peru	Thailand
Republic	Macedonia,	Romania	Tonga
Egypt	Former Yugoslav	Russia	Tunisia
El Salvador	Republic of	Samoa	Turkey

1. Note that these countries will compete separately from the first two groups for MCA funding.

and $2,975. Thus, by the third year 115 countries will be eligible to compete for MCA funding. The increase in the number of eligible countries is designed to correspond with the proposed increase in available MCA funds—the administration has proposed to ramp up the funds to $5 billion between fiscal 2004 and fiscal 2006.

Second, 16 indicators will be used to assess a country's commitment to "ruling justly, investing in their people, and establishing economic freedom." Six indicators are used for "ruling justly," four for "investing in people," and six for "establishing economic freedom." These indicators are examined in detail below.

Third, to aggregate scores across indicators, the administration proposes a "hurdles approach" in which a country must score higher than the median score (relative to other countries in its income group) to get credit on any indicator. To qualify for the MCA, a country must score above the median on half the indicators in each of the three categories. That is, it must be above the median on three of the six "ruling justly" indicators, two of the four "investing in people" indicators, and three of the six "establishing economic freedom" indicators. In addition, to qualify a country *must* score above the median on the corruption indicator.

Fourth, eligible countries will be split into two groups to score on the indicators. Countries with per capita incomes below $1,435 will compete against each other separately from those with incomes between $1,435 and $2,975. This step is meant to partially correct for the fact that most countries with higher incomes will score better than those with lower incomes. If all the countries competed against each other, countries from the higher-income group would be more likely to score above the median on any indicator, which would effectively eliminate many low-income countries.

Fifth, the quantitative process will not be the final determinant of country qualification. Rather, it will be the main input used by the MCC's board of directors, which will determine the final list of qualifying countries. According to the administration's proposal, the board will be guided by the indicators, but in making final decisions it will be "empowered to take account of data gaps, lags, trends, or other material information, including leadership, related to economic growth and poverty reduction." This last step introduces an element of subjectivity that probably is necessary given the weaknesses in the data. However, this discretion must be used carefully and only in a limited set of circumstances to guard against too much political influence on selection.

One implication of this proposed process is that the number of qualifying countries is likely to be small. To make an initial crude estimate, assume that country scores on any indicator are simply random, so that the odds of passing a minimum number of hurdles are the same as getting a minimum number of "heads" on a series of coin flips. It turns out that approximately 18 percent of the countries would pass all four tests (control of corruption, two of the five other "ruling justly" indicators, two of the

four "investing in people" indicators, and three of the six "economic freedom" indicators).[7] While this exercise is very simplistic, it illustrates a basic point: strictly using this system, the number of qualifying countries will not be large and will not change appreciably over time.

Income Levels for Broad Eligibility

There are several ways to define "low income" and thus the universe of countries from which MCA countries will be chosen. The administration has chosen three common definitions, all drawn from the World Bank and based on the Bank's definitions of IDA eligibility, IDA's historical cutoff, and lower-middle-income countries. The advantage of the chosen definitions is that they are internationally recognized; a disadvantage is that the Bank's management decisions on these definitions will determine MCA eligibility of some countries.

It is sensible for the MCA to start with the 74 low-income, IDA-eligible countries in the first year (box 2.1), since poverty is the most extensive there, and they have the greatest development needs; including them is appropriate and generally noncontroversial.

In the second year, 13 countries will be added to the competition. There is a clear trade-off in adding them. On the one hand, as more countries are added, less funds will be available for the poorest countries. Of course, in the second year of the MCA when these countries become eligible, MCA funding will grow, so none of the original countries will necessarily receive less funding. Nevertheless, less funds will be available to the first group than otherwise would have been the case if the second group (and later the third group) were not part of the MCA. Moreover, the second-stage countries will tend to score higher on most indicators, pushing up the median scores and crowding out some of the first-group countries that would have been above the median if they were judged only against other IDA-eligible countries. Thus, in all likelihood, by including the second group, fewer of the poorest countries will qualify and less money will be available to them.

On the other hand, many of the second-stage countries have significant numbers of poor people and also a sufficiently strong policy and institutional environment to put the MCA funds to good use. Moreover, since only 13 countries are added in stage 2, these concerns about the larger eligibility pool are not enormous.

In the third year, the administration proposes adding 28 countries with per capita incomes between $1,435 and $2,975 (box 2.1). These countries

---

7. The probability of passing the corruption indicator is 0.5. If country scores are completely random, the probability of passing 2 or more of the remaining 5 "ruling justly" indicators is 0.8125. The probability of passing 2 or more of the 4 "investing in people" indicators is 11/16, or 0.6875. The probability of passing 3 or more of the 6 "economic freedom" indicators is 0.6562. The joint probability of achieving all four is (0.5*0.8125*0.6875*0.6562) = 0.1833, or 18 percent.

Table 2.1 Development indicators, resource flows, and sources of financing for three MCA country groups (medians)

	IDA-eligible countries with per capita income less than $1,435	Countries with per capita income less than $1,435	Countries with per capita income between $1,435 and $2,975
Development indicators			
GNI per capita, 2001	$380	$460	$1,965
Adult illiteracy rate, 2000 (percent)	36	33	14
Life expectancy at birth, 2000 (years)	54	56	70
Infant mortality rate, 2000			
(per 1,000 live births)	75	69	27
Resource flows and financing			
Aid/GNI, 2000 (percent)	10.8	8.5	1.4
Gross private capital flows/GDP			
(percent)	6.9	8.7	10.3
Tax revenue/GDP (percent)	11.7	12.6	21.8
Gross domestic savings/GDP,			
2000 (percent)	7.3	8.4	16.2
Number of countries	74	87	28

Source: World Bank, *World Development Indicators*, 2000.

will compete separately from the countries with per capita incomes below $1,435, meaning that distinct median scores will be calculated for this group, which they must surpass to qualify. All other qualification requirements will remain the same. Since they will be scored on the indicators separately, the third group will not affect the median scores of the first two groups and will not directly crowd out any of these countries by pushing them below the median. The main argument in favor of adding these countries is that, although their average incomes are higher than the poorest countries, they are by no means rich and they have many people living in poverty.

However, adding this group of countries to the MCA raises significant concerns. There are three strong arguments against including these countries:

■ First, by adding these countries, less funding will be available for the poorest countries. Although these countries are poor, they are far better off than the poorest countries. The top half of table 2.1 compares four development indicators for the three groups. The countries with incomes between $1,435 and $2,975 (column 3) are more than four times richer than the combined low-income group (column 2). They also have substantially lower illiteracy rates, higher life expectancy, and lower infant mortality. In each case, the differences are quite large.

■ Second, the third group of countries has significantly larger alternative sources of financing available to it than do the low-income coun-

tries. One purpose of the MCA is to help prepare poor countries to access private capital markets and generate additional domestic resources, and most of these countries already have achieved progress in these areas. They also can borrow from the World Bank and other international institutions at favorable rates. The bottom half of table 2.1 shows that while the third group of countries (appropriately) receive much less aid, they receive larger flows of international private capital, generate much greater tax revenue, and have significantly higher domestic saving rates than the low-income countries. USAID currently does not have operations in 12 of these 28 countries; in several cases (e.g., Tunisia) they were judged wealthy enough to no longer need aid and so "graduated" from USAID funding.

- ■ Third, and more subtly, adding these countries increases the risk that the allocation of MCA funds will be determined to a greater extent by political and strategic criteria rather than the announced MCA criteria. Political considerations cannot be completely separated from MCA allocation, but the United States will find it especially difficult—in the cases of several countries in this group—to override political and strategic concerns in favor of aid effectiveness. These countries include Colombia, Egypt, Jordan, Turkey, and Russia, among others. Strategic considerations could affect decisions on country qualification (especially in marginal cases), the amounts of money that qualifying countries receive, and funding reductions as a result of poor performance. One cannot help but wonder if the surprise decision to include these countries was partly motivated by the desire of some officials to have MCA funds available for strategic reasons when necessary. Of course, it is perfectly legitimate to use foreign assistance funds for strategic reasons, but it would be far better to use funds outside the MCA for these purposes and focus the MCA on increasing US aid effectiveness in supporting development.

Adding the third group of countries raises another issue. How will MCA funds be allocated between these countries and the lower-income group? There are two choices: (1) establish two pools of money, with a fixed amount of funding allocated to each group every year, or (2) rely on a single pool of money, with the quality of proposals and country performance determining funding allocations. The second choice has merits, and I am generally in favor of competing proposals and performance-based allocations. However, institutional capacity and human-resource skills are so much greater in the third group of countries that they would dominate the proposal process. The richer countries are more likely to show stronger results since they face fewer constraints than the low-income countries (which is one reason they receive larger private capital flows). For these reasons, if the administration sticks with the proposal to

---

Box 2.2   Ideal characteristics of the indicators

Transforming the president's three broad criteria into specific indicators to determine qualification is far from straightforward. There are a large number of indicators to choose from, each with advantages and disadvantages. In making these choices, certain attributes of the data should help guide the choice.[1] The indicators should be:

■ *Simple, transparent, and publicly available, with good country coverage.*

■ *Moderate in number.* Too many indicators can make the selection process cumbersome and opaque; too few could give a misleading perspective on a country's commitment to development. The administration has chosen 16 indicators; one could add a few more if they provided important additional information.

■ *Measures of policies rather than outcomes.* The indicators should focus on policy variables and institutional changes that are within the control of government officials, rather than on outcomes that will only change over time and may be influenced by exogenous factors.

■ *Indicative of broader policies.* Since they are moderate in number, the chosen indicators should capture related policies that are not directly measured. For example, the inflation rate not only gives a direct perspective on monetary policy but also is indicative of a country's overall macroeconomic management and is related to fiscal and exchange rate policies.

■ *Associated with desired development outcomes.* The indicators should be demonstrably empirically related to critical desired outcomes such as faster economic growth, decreased infant mortality, and increased literacy.

■ *Accurate.* All indicators are to some extent simply estimates of the true value, and the more accurate the estimate the better. Indicators estimated with smaller margins

---

include the countries with incomes between $1,435 and $2,975, it would be preferable to allocate to that group a fixed share of the MCA funds, with the remainder to be used in the low-income group. As suggested in Radelet (2002a), a reasonable amount would be $1 billion per year, 20 percent of the total. This amount is loosely based on population shares, since these 28 countries account for 19 percent of the population of the entire group of 115 countries eligible to compete for MCA funds.

Choosing the Indicators

In this section, I describe the 16 indicators proposed by the administration with reference to the ideal characteristics described in box 2.2. I also briefly discuss other indicators that might be considered.

Basic information on the 16 indicators is summarized in table 2.2, including the number of countries covered by the indicator, the source of the data, and the frequency with which the indicator is updated. The table also shows the results of some simple statistical tests on the association

---

Box 2.2   *(continued)*

of error (and with aggregation techniques that minimize rather than multiply those errors) should be preferred to those with larger measurement errors.

■ *Updated frequently with short time lags.*

■ *Not easily subject to targeting or manipulation.* Recipient governments will take great interest in the indicators used to determine MCA eligibility and will naturally try to make sure those indicators are as favorable as possible. Choosing indicators that are less easily manipulated, or choosing a larger set of indicators, would help reduce this problem.

■ *Objective rather than subjective.* To the extent possible, indicators should be objective (based on measurable quantities rather than personal judgments) rather than subjective. Moreover, they should be measured in absolute rather than relative terms. However, many indicators, such as the level of corruption, cannot easily be measured in absolute terms.

In choosing the indicators, it will not be possible to meet all these criteria. For example, it will be difficult to find indicators that are simultaneously simple, indicative of broader policies, and reliably accurate over time. As a result, some trade-offs are inevitable, and any indicator will be stronger on some attributes and weaker on others. The designers of the MCA need to recognize these trade-offs and be willing to revise the methodology over time as data become more refined, as problems in aggregation methods appear, or as improved techniques and indicators are developed.

---

1. For a similar discussion, see Birdsall et al. (2002).

---

between each of the indicators and three important development outcomes: per capita income growth, infant mortality, and literacy. All else being equal, indicators that have a demonstrated correlation with important outcomes (such as lower infant mortality) are preferable to indicators that have no such correlation, since we are interested in policy actions that are associated with better development outcomes. (It would be preferable to show causality between the indicators and the outcomes, a subject left to future research.) In each case, the correlations are measured after controlling for the initial level of income.[8] This procedure controls for the fact

---

8. Specifically, I regress average per capita income growth from 1990–2000 (g) on the initial level of income in 1990 (Y) and the average value of each indicator from 1990–2000 (i) as follows:

$$g = \alpha_0 + \alpha_1 {}^*Y + \alpha_2 {}^*i + \varepsilon$$

where α_0 is a constant, α_1 and α_2 are the estimated coefficients on the initial income level and the indicator, respectively, and ε is an error term. I then repeat the process, substituting infant mortality and literacy for income growth as the left-hand variable. I focus on α_2, the estimated coefficient for each indicator. The results in table 2.2 show both the sign and the statistical significance of the estimate for α_2. Data for the regressions are drawn from all countries where data are available. Note that this process is not meant to demonstrate causality from the indicator to the outcome, only the correlation after controlling for the level of income.

Table 2.2 Characteristics of the proposed indicators

Indicators	Country coverage	Main source	Frequency of data collection by source	Correlation* with					
				Per capita income growth		Log infant mortality		Log literacy	
				Correct sign	Significance (percent)	Correct sign	Significance (percent)	Correct sign	Significance (percent)
Ruling justly									
Control of corruption	115	World Bank Institute[a]	Annual	✓	10	✓	1	✓	
Rule of law	115	World Bank Institute[a]	Annual	✓	5	✓	1	✓	
Voice and accountability	115	World Bank Institute[a]	Annual	✓		✓	1	✓	1
Government effectiveness	115	World Bank Institute[a]	Annual	✓	10	✓	1	✓	1
Civil liberties	115	Freedom House[b]	Annual	✓		✓	1	✓	1
Political rights	115	Freedom House[b]	Annual	✓		✓	1	✓	1
Investing in people									
Immunization rate: DPT and measles	112	World Health Organization[c]	Annual	✓	10	✓	1	✓	1
Primary education completion rate	99	World Bank[d]	Periodic	✓		✓	1	✓	1
Public primary education spending/GDP	83	World Bank[e]	Annual			✓		✓	
Public expenditure on health/GDP	88	World Bank[e]	Annual			✓	5	✓	1

Economic freedom

Indicator								Source	Frequency	Number
Country credit rating	✓	✓	5	✓	1	✓	1	Institutional Investor[f]	Semiannual	85
Inflation	✓	✓	1	✓	5	✓	1	International Monetary Fund[g]	Monthly	97
Regulatory quality	✓	✓	10	✓	1			World Bank Institute[a]	Annual	115
3-year budget deficit	✓	✓	1	✓	1			World Bank[d]	Annual	103
Trade policy	✓	✓		✓	1			Heritage Foundation[h]	Annual	92
Days to start a business	✓	✓		✓	1			World Bank[i]	Annual	63

Other possible indicators

Indicator								Source	Frequency	Number
Political stability	✓	✓	5	✓	1	✓	10	World Bank Institute[a]	Annual	98
Access to essential drugs	✓	✓		✓	5			World Health Organization[j]	Annual	106
Net school enrollment rates	✓	✓		✓	5	✓	1	World Bank[d]	Annual	82
Access to improved water	✓	✓		✓	5			World Bank[d]	Decennial	96
Girls' participation in primary education	✓	✓		✓	1	✓	1	World Bank[d]	Annual	85

* = Correlations control for initial level of income. See text for details.

Note on statistical significance: Less than 1 percent (shown as 1 percent), between 1 and 5 percent (shown as 5 percent), between 5 and 10 percent (shown as 10 percent), or not statistically significant at conventional levels (blank).

Sources:

a. Kaufmann, Kraay, and Mastruzzi (2003).
b. Freedom House (2002).
c. WHO Statistical Information System (WHOSIS), World Health Organization.
d. World Bank, *World Development Indicators,* 2002.
e. Bruns, Mingat, and Rakotomalala (2003).
f. Institutional Investor, Country Credit Rankings, September 2002.
g. International Monetary Fund, *International Financial Statistics,* 2002.
h. Heritage Foundation/Wall Street Journal Index of Economic Freedom, 2002.
i. Djankov et al. (2002).
j. World Health Organization, WHO Action Program on Essential Drugs.

that many of the indicators simply improve with income levels. For example, richer countries tend to have both higher immunization rates and lower infant mortality rates. A simple correlation between immunization rates and lower infant mortality rates may simply reflect the effect of income on both variables. By controlling for the level of income, it can be determined, for two countries *with the same income level*, whether the one with the higher immunization rate also has a lower infant mortality rate. These tests are not meant to be deterministic models but simply controlled correlations.

Table 2.2 shows whether the correlation has the "correct" sign, in the sense of whether a better score on an indicator is positively associated with economic growth, negatively associated with infant mortality, and positively associated with higher literacy. (The direction of the sign, positive or negative, is not shown because for some indicators a higher score is a better outcome, while for others a higher score is a worse outcome, so showing the sign could be confusing.) The level of statistical significance of the correlation is then shown: less than 1 percent (highly statistically significant), between 1 and 5 percent (shown as 5 percent), between 5 and 10 percent (shown as 10 percent), or not statistically significant at conventional levels (blank).

Ruling Justly

Control of Corruption.   Corruption is the exercise of public power for private gain. It undermines the rules that govern interactions between public servants and the citizenry, adversely affects business decisions, and can be especially detrimental to the poor. Many different surveys measure aspects of corruption, including DRI/McGraw-Hill, Transparency International, the Economist Intelligence Unit, and the Political Risk Services Group. These surveys draw both on specialists in individual countries and/or experts with knowledge across many countries. They explore many dimensions of corruption, including the frequency of making additional payments, the effectiveness of anticorruption measures, and the impact of corruption on foreign investment.

The administration draws its corruption indicator from a governance database compiled by Daniel Kaufmann and Aart Kraay (hereafter KK) at the World Bank Institute.[9] The methodology and sources used in this database are described in box 2.3. Although the KK indicators are subjective

---

9. The KK indicators can be downloaded from www.worldbank.org/wbi/governance, along with three key papers that describe the data and methodology (Kaufmann, Kraay, and Zoido-Lobatón 1999a, 1999b, and 2002). Details on the last round of indicators can be found in Kaufmann, Kraay, and Mastruzzi (2003). For a paper on data issues in the MCA, see Kaufmann and Kraay (2002a).

Box 2.3   The Kaufmann-Kraay governance indicators

The administration draws 5 of its 16 indicators from a governance research database compiled by Daniel Kaufmann and Aart Kraay at the World Bank (hereafter KK). To create this database, the authors compile 275 governance variables from 20 sources and 18 organizations, including Freedom House, Gallup International, the Economist Intelligence Unit, DRI/McGraw-Hill, the Heritage Foundation, the World Bank, the European Bank for Reconstruction and Development, Latinobarómetro, and others. The authors construct the most comprehensive composite measures of governance available, which they organize into six separate indicators, covering four time periods: 1996, 1998, 2000, and 2002. Note that these data are not World Bank data per se but are aggregate indicators based on data compiled by Bank researchers from independent sources. The administration uses four of these KK aggregate governance measures as "ruling justly" indicators (control of corruption, rule of law, voice and accountability, and government effectiveness) and one as an indicator of "economic freedom" (regulatory quality). It does not use the KK "political stability" variable as an indicator.

Since the KK database draws from so many sources, it covers a much larger set of countries than any individual source. Country coverage for the MCA is complete, with data on all 115 countries with per capita incomes below $2,975. Importantly, research with these indicators reveals more than just an association with development indicators—there is a strong *causal* relationship with these indicators to higher levels of income, lower rates of infant mortality, and higher rates of literacy. Moreover, the method of aggregation provides measures for the precision of the estimates so users know the margin of error associated with any indicator. Most other survey sources do not discuss the sampling error associated with their estimates, giving users a false sense of precision. The authors have written a paper focused on the use of their indicators for the MCA in which they discuss the issue of margins of error at some length, including expressing reservations about using a "hard hurdle" for corruption in the selection process (Kaufmann and Kraay 2002a). In the aggregation process, the KK methodology gives greater weight to survey results with a smaller measurement error and lesser weight to survey results with greater uncertainty. The aggregation methodology, coupled with the multiple sources that the construction of the indicators draws upon, significantly reduces the margins of error, compared with relying solely on a single source.

The KK database, however, has some drawbacks. As with most databases on corruption and other governance concepts, the KK database is primarily based on subjective measures of governance. As a result, a country's score is measured relative to other countries in a particular year.[1] When a country improves from one year to the next, it is difficult to tell the extent to which it improved or others got worse (this is a weakness of most governance-related indicators). However, in their latest update, Kaufmann and Kraay have undertaken to present an "adjustment" factor correcting for changes in world averages over time to ascertain whether there has been an improvement or deterioration on average, and thus estimate the likely difference, if any, between relative and absolute changes over time. Also, the database has a relatively short history, with data available only for 1996, 1998, 2000, and 2002 (although there is now an institutional commitment to periodically generate these estimates). Finally, the statistical methods used to combine various surveys to create the KK database are more complicated than the methodology for any individual survey, making the final scores somewhat less transparent than other measures. The advantages of this method, however, are far more information from many data sources, excellent country coverage, and a better idea of the precision of the estimates. This accuracy and increased information are well worth some

(box continues next page)

Box 2.3 *(continued)*

complexity. All in all, this is the most comprehensive and best-quality database available on governance indicators.

---

1. Specifically, the scores are scaled so that the mean value for the world is set to equal zero, and the scores that are one standard deviation above and below the mean are set equal to +1 and –1, respectively. Thus, each variable is reported on a scale from approximately –2.5 to +2.5 (corresponding to 2.5 standard deviations below and above the mean, respectively).

---

and more complex than most of the other indicators, they show a strong causal relationship with key development outcomes, are based on a wide range of sources, have 100 percent country coverage for the potential MCA countries, and are updated fairly regularly. They are the best set of governance indicators currently available.

In creating their indices, Kaufmann and Kraay draw on nearly 275 indicators from 20 sources to construct six aggregate indicators of dimensions of governance. The KK "control of corruption" indicator draws on surveys by DRI/McGraw-Hill, the Economist Intelligence Unit, the World Bank's business surveys, the Political Risk Services Group, and others. It does not draw directly on Transparency International's (TI) well-known corruption indicator, as TI is itself a compilation of other surveys rather than an original source. The KK indicator draws on all the surveys contained in the TI index and several others. The KK control of corruption index contains data for all 115 countries in the world with incomes below $2,975 in 2002.

The KK control of corruption index is scaled so that the worldwide mean score is 0, and values that are one standard deviation above or below the mean are reassigned values of 1 and –1, respectively. We use a recalibrated version of the data in which each country's score is shown as a percentile rank. As shown in table 2.2, this index performs fairly well statistically: it shows a modestly strong correlation (controlling for initial income) with faster economic growth and a very strong relationship with reduced infant mortality. It is also correlated with improved literacy, although the controlled correlation is not statistically significant at conventional levels.

Rule of Law. For economic development to proceed, societies need fair and predictable rules to govern economic and social interactions. Ideally, these rules should govern the enforceability of contracts, dispute settlement, criminal behavior, procedures for the judiciary, the protection of property rights (including intellectual property rights), the extent of tax

evasion, and the extent of black market activity as an impediment to business development. There are several sources that touch on these issues, including DRI/McGraw-Hill, the Economist Intelligence Unit, Heritage Foundation/Wall Street Journal, and others. The administration uses the KK index on rule of law, which compiles information from each of these sources and covered all 115 countries in 2002. As shown in table 2.2, this index shows a strong and statistically significant correlation with faster economic growth and lower infant mortality in the 1990s. It is positively associated with higher literacy, but the controlled correlation is not statistically significant.

Voice and Accountability.   For governments to rule justly, countries require institutions that protect civil liberties; ensure that governments are held accountable for their actions; and allow citizens to participate in the political process, choose and replace their leaders, and freely voice their opinions. Countries with free and fair elections, representative legislatures, fair legal systems, a free press, and minimum role for the military in elections are more likely to be responsive and accountable to their people. Similarly, governments must respect basic freedoms of speech, assembly, and religion. Several surveys focus on these issues, most importantly the one by Freedom House. In addition, the Economist Intelligence Unit and the Political Risk Services Group include questions that touch on these topics, as do several other surveys. The KK indicator on voice and accountability incorporates measures from all these sources, covering all 115 MCA countries in 2002. This measure is positively correlated with economic growth, although the correlation is not statistically significant. The relationships with infant mortality and literacy are both very strong and highly statistically significant.

Quality and Effectiveness of Government Institutions.   Good governance requires effective public institutions. Poor-quality civil service, red tape, ineffective bureaucracies, and weak management impede the government's ability to deliver basic public services and serve the general public. Drawing from a similar set of sources, the KK indicator on government effectiveness compiles data on these and related issues for all 115 MCA countries in 2002. This measure shows a moderately significant relationship with economic growth and a strong relationship with reduced infant mortality. It is also positively associated with literacy, but the relationship is not statistically significant.

Civil Liberties.   The Freedom House civil liberties and political rights indices evaluate the rights and freedoms enjoyed by individuals in countries and territories around the world. Freedom House does not rate governments per se, but rather the extent to which citizens enjoy basic rights.

The civil liberties index focuses on the freedoms for citizens to develop views, institutions, and personal autonomy apart from the state. It is a subjective index, ultimately based on the judgments of the Freedom House survey team, with the ratings subject to several layers of review. The ratings review process involves about 30 outside regional experts, consultants, and in-house staff. Scores are based on a relatively narrow range of 1–7 (whole numbers only), although underlying these numbers is a confidential Freedom House index in which countries are given a score between 1 and 100.

While this scoring system is appropriate for Freedom House's objectives, the fact that there are only seven possible scores raises a statistical problem for the MCA. Since many countries are assigned exactly the same score (e.g., a 4 or 5), they are bunched together around the median score, which is where the administration draws the line between passing or failing on a particular indicator. In this case, there is a big difference between the administration's proposal, in which a score *greater than* the median is given a passing grade, and the alternative, in which a score *greater than or equal to* the median is given a passing grade. For example, of the 87 countries eligible for the MCA in the second year, 21 countries have the median score of 4 on the civil liberties index. An index with a more differentiated scale would be preferable. The decision as to whether these countries should be given a passing grade on this indicator could make a significant difference in the final list of eligible countries. I return to this issue later in the chapter.

The 2001–02 Freedom House survey contains information on 192 countries, including all 115 MCA countries (Freedom House 2002). As mentioned earlier, this information is also included in the KK voice and accountability indicator, so the Freedom House information is actually counted twice in the MCA process. Using our simple statistical analysis, better civil liberties scores are associated with faster growth, although the correlation is not statistically significant. Better scores are strongly associated with both lower infant mortality and higher literacy.

Political Rights.   According to Freedom House, political rights "enable people to participate freely in the political process, which is the system by which the polity chooses authoritative policy makers and attempts to make binding decisions affecting the national, regional, or local community." These rights allow all adults to vote and run for election and for elected officials to have decisive votes on public policies. As with the civil liberties index, political rights are measured on a 1 to 7 scale, raising the same issue about median scores. Again, as with the civil liberties indicator, the political rights indicator has a positive but insignificant correlation with growth, and a very strong relationship to reduced infant mortality and higher rates of literacy.

Investing in People

Immunization Rate.   Immunizations are among the most effective means to prevent the spread of infectious diseases and ensure the basic health of the population. Countries with higher immunization rates against diphtheria, pertussis (or whooping cough), tetanus (DPT), and measles tend to have lower rates of infant mortality and longer life expectancy. Moreover, although immunization rates are not a policy per se, they are within governmental control and can be increased in most countries where governments commit to do so. They are also a good indicator of broader health policies and strategies: governments that establish systems to provide broad-based immunizations tend to take other steps to improve basic health. The United Nations has adopted the measles immunization rate as one indicator of progress toward achieving the Millennium Development Goal of reducing the mortality rate among children below 5 years by two-thirds between 1990 and 2015.[10] The World Health Organization (WHO) provides data on the share of children below 1 year who received immunizations for DPT (three doses) and measles (one dose). The administration uses the average of the two as its indicator. The data are a little uneven, at times varying widely from year to year within one country. Data are available for 112 of the 115 MCA countries, although in some cases they are 3 or even 4 years old. As shown in table 2.2, this variable is very strongly related to lower infant mortality (not surprisingly) and has an equally strong relationship with increased literacy rates. It also shows a moderately positive association with economic growth.

Primary School Completion Rate.   Primary school *enrollment* rates have long been used as a basic indicator of education policy. However, enrollment rates provide little information on achievement of basic standards of competence. Attending just a year or two of school reaps little benefit: a growing body of evidence suggests that students must complete five to six years of school to achieve basic competencies in literacy and numeration (World Bank 2002b). Thus, *completion* rates for primary school are a stronger indicator of student achievement of minimum skill levels than are enrollment rates. One of the 13 Millennium Development Goals adopted by the United Nations is to "ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling."

The World Bank defines the primary school completion rate as "the total number of students successfully completing (or graduating from)

---

10. For more information on the Millennium Development Goals, see www.development goals.org/.

the last year of primary school in a given year, divided by the total number of children of official graduation age in the population." The best source of primary school completion rates is the database compiled by World Bank researchers Barbara Bruns, Alain Mingat, and Ramahatra Rakotomalala (2003). This database is relatively new, as completion rates have only recently been a focus of attention, and as such provides an excellent foundation for a stronger education database. However, there are several drawbacks to this indicator. First, completion rates will tend to increase noticeably only several years after governments initiate a firm commitment to improving primary education. Second, this indicator is either missing or is several years old for many countries. Only about half of the countries have data for 1999 or subsequent years. The World Bank researchers are planning to strengthen this indicator by expanding the number of countries covered and updating it on an annual basis. The school completion rate is very strongly correlated with lower infant mortality and higher literacy. It also is positively correlated with economic growth, but the relationship is not statistically significant at conventional levels.

Public Primary Education Spending as Percent of GDP. Public-sector spending on education is a policy variable that is very much in the government's control. For most of the poorest countries, primary schools are the appropriate focus for government expenditure. At face value, public spending should be indicative of a broader government commitment to improving education. However, greater spending generally does not translate into better schools or better outcomes if it is inefficient or poorly targeted, so this variable probably is a weak indicator of effective government policies on education. The pattern of expenditure (on books, salaries, building maintenance, etc.) is just as important as is a focus on curriculum development and other aspects of the quality of education.

Another concern with this indicator is that it is not directly available: it must be constructed using other variables. The World Bank publishes data on overall spending on education and on primary school spending per student. I constructed the data on primary school spending as a share of GDP from data on spending per student, the number of students, and GDP. Unfortunately, these data are missing for many countries (data are available for only 83 of the 115 MCA countries) and are several years old in other countries. Moreover, the indicator is weakly correlated with development outcomes. Higher primary education spending is correlated with lower infant mortality and higher literacy, but neither relationship is statistically significant. It has a slightly *negative* association with economic growth, although the relationship is not statistically significant. Statistically speaking, this variable is one of the weakest of the 16 indicators. This fact, and that it has to be constructed using other variables, makes it a prime candidate for improvement, if it is to be used for the MCA.

Public Expenditures on Health as Percent of GDP. Public spending on health has many of the same characteristics as spending on primary education. This policy variable is clearly in the government's control and may be indicative of broader health policies. However, more spending is generally not associated with better health outcomes for the poor, such as if spending is focused on urban cancer hospitals rather than rural clinics. This indicator is drawn from the World Bank's *World Development Indicators* database, which defines public health expenditure as "recurrent and capital spending from government (central and local) budgets, external borrowings and grants (including donations from international agencies and nongovernmental organizations), and social (or compulsory) health insurance funds." Data are available for 88 of the 115 MCA countries, slightly more than for the education-spending indicator. Public expenditure on health (after controlling for income levels) is strongly correlated with lower infant mortality and higher literacy. However, more sophisticated econometric analyses that control for a larger set of variables find little or no relationship between health spending and these outcomes (Pritchett and Filmer 1999). As with public spending on primary education, the correlation of public spending on health with economic growth is of the wrong sign and is not statistically significant. Again, as is the case with primary education spending, the health spending indicator should also be strengthened or replaced over time for MCA selection purposes.

Other Possible Indicators. Several other variables could be used as indicators for health and education policies. It is somewhat surprising that the administration proposed only four "investing in people" indicators—two each for health and education—while they proposed six each for "ruling justly" and "economic freedom." Although adding more indicators may not change appreciably the final list of countries, it will make it more difficult for recipient countries to focus too narrowly and "target" one indicator rather than broader health and education policies. Other possibilities include:

- *Ratio of girls to boys in primary school.* This measure is indicative of both education policies and gender discrimination. It is widely available for most countries and updated regularly. The United Nations uses this ratio as an indicator to achieve the Millennium Development Goal of "promoting gender equality and empowering women." It is very strongly correlated with literacy rates but not strongly correlated with other outcome variables.

- *Primary school enrollment rates.* This indicator measures the ratio of the number of children of primary school age who are actually enrolled in school to the corresponding population. Enrollment rates are

similar to school completion rates but are more widely available, and it is easier for governments to influence enrollment in a short time period than to increase completion rates. Focusing on enrollment rates alone would be a mistake, as discussed earlier, since enrollment in school does not ensure quality education. But the combination of enrollment and completion rates would provide more information about government policies than either indicator alone. However, net enrollment rates are available for just 82 of the 115 MCA countries. This indicator is strongly correlated with lower infant mortality and higher literacy but is not strongly associated with economic growth.

- *Access to improved water sources.* This variable measures the share of the population that has reasonable access to water from an improved source, such as a household connection, public standpipe, borehole, protected well or spring, or rainwater collection. Access to clean water can improve a variety of health indicators. Moreover, government policy can directly impact the share of the population with access to water. The United Nations uses access to improved water sources as an indicator of progress toward one of the Millennium Development Goals. The obstacle to using this variable is that it is recorded just once every 10 years in the World Bank database and even then is missing for many countries. Access to water is strongly correlated with lower infant mortality but is not strongly associated with literacy or economic growth.

- *Access to essential drugs.* Every year the WHO Action Program on Essential Drugs interviews relevant experts in each country about access by the population to essential drugs. The interviewees can choose from four levels: less than 50 percent, between 50 and 80 percent, 80 and 95 percent, and above 95 percent. This variable appears to be too subjective for use in the MCA, as it would be relatively easy for recipient governments to influence its measurement.

Establishing Economic Freedom

Country Credit Ratings.   There are a large number of credit ratings and investor guides for many countries. These ratings usually measure the risk of default on government or private-sector debts and broadly indicate the opinion of private creditors on the economic environment in a country. However, relatively few of these rating agencies report regularly on low-income countries. One exception is *Institutional Investor*, which provides credit ratings, based on the perceived risk of government default, every six months for 145 countries, including 85 MCA countries.[11]

---

11. See data at www.institutionalinvestor.com/premium/rr/index.htm.

Countries are ranked on a scale from 1 to 100 based on information provided by economists and sovereign risk analysts from banks and money management and securities firms. This indicator is strongly correlated with faster economic growth and very strongly correlated with lower infant mortality. It is positively associated with higher literacy, but the relationship is not statistically significant.

Inflation.   High rates of inflation make the environment for new investment more risky, tend to reduce the profitability of most businesses, and are especially harmful to the poor, who are least able to protect themselves from inflation. The administration measures inflation on a year-on-year basis from data available for the most recent 12 months from IMF's monthly publication, *International Financial Statistics*. For inflation, the administration determines the "passing grade" differently than it does for any other indicator. Whereas the passing score for the other indicators is the median, that for inflation is a rate lower than 20 percent. Since the median rates of inflation are under 8 percent for each of the three country groups, the 20 percent standard is much easier to pass than the median. Presumably, this approach is taken because there is strong evidence that inflation rates greater than 20 percent are very harmful, but not much evidence that an 8 percent rate is necessarily so superior to a 10 percent rate in developing countries. Inflation data are available for 97 of the 115 MCA-eligible countries, and 88 of them pass this test, making it by far the easiest hurdle to pass. Lower inflation has a very strong and positive relationship with economic growth. It is also strongly correlated with *higher* infant mortality and *lower* rates of literacy, the opposite of what might be expected. However, since inflation is a *change* in the price level, I checked the controlled correlation between inflation and the *change* in infant mortality and literacy. In this case, higher inflation is associated with a larger increase in infant mortality, and the correlation is significant at the 10 percent level. There is no relationship between inflation and the change in literacy.

Regulatory Policies.   While a certain amount of regulation is necessary to make some markets (such as financial markets) work better, too much regulation, intervention, and government control can undermine the incentives for investment and job creation. Common burdensome regulations include wage and price controls; inadequate bank supervision; excessive controls on trade, investment, and business start-ups; excessive restrictions on international capital flows; ponderous legal restrictions on ownership and equity positions by nonresidents; and red tape. As an indicator of these issues, the administration uses the regulatory quality measure from the KK database, which is probably the most comprehensive measure of these policies available. Other major sources of data on regulatory quality are the Heritage Foundation/Wall Street Journal Index and the DRI/McGraw-Hill dataset. Both these sources are incorporated into

the KK dataset, but individually each has less country coverage and larger standard errors in measurement than the KK composite index. Like the other KK measures, this indicator is a composite of the leading surveys and other data on regulatory issues. The KK indicator covers all 115 MCA countries. It has a modestly strong positive correlation with economic growth and very strong and significant correlations with both lower infant mortality and high literacy.

Budget Deficit.   The budget deficit is a key indicator for overall macroeconomic policy, with larger deficits tending to be associated with macroeconomic instability, inflation, and exchange rate depreciation. It is also the basic measure of a government's propensity to spend beyond its means. Of course, a lower budget deficit is not always better: there are times when running a slightly larger budget deficit is appropriate, especially as a countercyclical policy tool. Moreover, donor funds for particular programs can actually increase the budget deficit (as conventionally measured) because of associated government spending, with donor funds entering the accounting as a below-the-line financing item. Problems can arise, however, when deficits become large, are not funded by grants or highly concessional loans, and persist over time.

Budget deficits can be measured in several ways. In this context, perhaps the most appropriate measure would be the deficit remaining after donor receipts (grants and concessional loans) averaged over three years. In essence, this measure would capture the extent of government spending that is not financed by tax or aid receipts and must be financed by the central bank or by borrowing from domestic or international commercial markets. The biggest problem with this indicator is that budget data are surprisingly incomplete for most countries, with data on concessional loans that finance the budget missing for many countries.

Thus, the administration uses the budget deficit after receipt of grants but not concessional loans. It measures the average over three years to allow for variation in the year-to-year deficit for economic management purposes. The main public source for budget data is the International Monetary Fund's *Government Financial Statistics* as reported in the World Bank's *World Development Indicators*.[12] However, the IMF also keeps a confidential database for a number of countries where the data cannot be made public, and the administration is using this source for the budget deficit indicator. Thus it is the only indicator not publicly available. The public data used here and those used by the administration differ. The budget deficit data are available for 103 of the 115 MCA countries. Lower budget deficits are strongly and positively associated with economic

12. Since data are missing for many countries, in some cases I have augmented the data with information from the World Bank's Country-at-a-Glance tables, available at www.worldbank.org/data/countrydata/countrydata/.

growth. Lower deficits are associated with lower rates of infant mortality, although the relationship is not statistically significant. There is no correlation between deficits and literacy rates.

Trade Policy.   Trade policy is one of the most important determinants of economic growth and poverty reduction. Almost all economists would agree that at least at a broad level, trade openness is good for growth and poverty reduction. Controversy abounds, however, on exactly what kind of trade policy is best for developing countries and how to measure it. Many economists argue that lower tariff rates and quota coverage are essential to open trade and growth; some others believe that *modest* and time-bound import substitution in selected industries can be conducive to long-term growth. Overall tariff averages can be misleading because the composition of the items protected can matter a great deal. Moreover, other institutions and policies such as export processing zones and directed credit to exporters make statutory tariff rates less meaningful, thus making it difficult to summarize overall trade policy in a single index. Also, data on average tariffs and quota coverage are surprisingly sparse for most countries. Perhaps the most appropriate indicators—if the data were available—would be the average tariffs and quota coverage for capital and intermediate goods (and not consumer goods).

As substitutes, some surveys include a trade policy component in an attempt to capture business and expert opinions on overall trade policy. The administration uses one of these as its indicator: the trade component of the Heritage Foundation/Wall Street Journal Index of Economic Freedom.[13] The 2003 index includes information gathered during the last half of 2001 and the first half of 2002. This index is subjective—the Heritage Foundation/Wall Street Journal authors rate the countries from 1 to 5, basing their judgments primarily on tariff and quota rates where available. Missing entries are filled in with data that, in some cases, unfortunately may not be indicative of trade openness. For example, as a substitute for tariffs the authors use government tariff revenues as a share of imports. But in many low-income countries, low tariff collections are more indicative of corruption in the customs department than of open trade policy.

Partly because of these issues, this indicator is only weakly correlated with development outcomes. A better score is positively associated with economic growth and higher literacy, but the relationships are weak and not statistically significant. The index is strongly correlated with lower rates of infant mortality. This index also suffers from the same problem as the Freedom House indices when working around the median. Since there are only five possible scores a country can receive, many countries receive the median score. Very few low-income countries receive a score

---

13. The data are available at www.heritage.org/research/features/index/.

of 1 or 2: of the 92 MCA-eligible countries with available data, 82 receive a score of 3, 4, or 5. In effect, this becomes a three-point scale, with countries either passing at the median, or failing. This makes the judgments on whether a passing grade is *greater than* the median or *greater than or equal to* the median very important, since 33 of the 92 countries receive the median score of 4. Therefore, much greater differentiation in the country scores would be preferable. While this index might be acceptable for trade policy at the moment, the administration should work toward identifying or creating a stronger index based on actual tariff and quota rates.

Days to Start a Business. The procedures, time, and costs of starting a business can be serious detriments to entrepreneurial energies in many countries. The administration uses the number of days to start a business as the indicator, drawing on data compiled by Simeon Djankov et al. (2002).[14] This database counts the number of days required for companies to complete all procedures necessary to legally start a business. The time required is very high in most countries, especially in low-income countries. The average number of days required in low-income countries is 66, compared with two days in Canada. Djankov et al. find that heavy start-up regulations are not correlated with better social outcomes (such as lower pollution or fewer accidental deaths) but are correlated with higher levels of corruption. (However, the latter relationship is not statistically significant in low-income countries.)

The major difficulty with this database is that it is available for only 110 countries and only 63 MCA countries. Thus 52 of the MCA countries—45 percent—fail this hurdle simply because of missing data. By definition, only half the remaining countries can surpass the median, making this indicator by far the hardest to pass. Also, since this database is relatively new, it has not been thoroughly tested over time, and a process to update it annually is not yet in place. Using a simple correlation controlling for level of income, there is no statistically significant relationship between this variable and economic growth, infant mortality, or literacy in the 1990s.

Although this indicator has great merit in principle and the authors have done a commendable job of creating this relatively new indicator, more work is necessary to improve the coverage of the indicator and its relationship to growth. One possibility (suggested by the authors themselves) is to combine this indicator with other measures of costs and procedures necessary to start a business (available in their database) to build a more comprehensive measure of barriers to start a business. In the near future, the authors also hope to expand the scope of their database to include bureaucratic harassment, protection of property rights, the quality of infrastructure services, and other related issues.

---

14. A slightly older version of the data can be downloaded from http://econ.worldbank. org/files/2379_wps2661.pdf. Most recent data were directly received from Simeon Djankov.

Aggregating the Indicators

Once the indicators are chosen, there are several alternatives to determine the precise standards that countries should be expected to meet on each indicator, how much weight each indicator should be given in the final determination of eligibility, and how the indicators should be aggregated into the final ranking of countries. In turn, the answers to these questions are partly related to the question of how many countries should qualify, since the final number will directly depend on how high or low the standards are set for each indicator. Does the United States want the MCA to focus on the top 10 countries according to the president's criteria? The top 20? Or the top 30? Analytically, there is no right answer for how many countries should be chosen, since we do not have precise empirical evidence on the exact point at which policy and institutions become so weak that aid is not effective (Clemens and Radelet 2003). There are two broad methodologies for setting standards, aggregating the data, and choosing the final set of qualifying countries: (1) establishing specific "hurdles" for each indicator or (2) adding together the scores (appropriately rescaled) for each indicator.

Hurdles Approach

In the hurdles approach, countries are expected to meet a specific standard on each indicator. The administration's proposal uses this approach. It requires that countries score above the median (the hurdle) on half the indicators in each of the three groups of criteria. In addition, the administration adopted a "hard" hurdle for corruption: a country must be above the median on corruption to qualify, regardless of how well it does on the other indicators. Apparently, the president insisted on a hard hurdle for corruption.

There are several advantages to this approach. First and foremost, countries do not have to do well on absolutely every indicator to qualify. For example, a country can have the highest tariff rates in the world and still qualify for the MCA as long as it passes enough of the other hurdles. Indeed, a country can have both high trade barriers and a large budget deficit and still qualify, as long as it makes three of the remaining four "economic freedom" indicators. Second, the system is transparent and easy to understand: all one needs to know about a country are its score on an indicator and the median. Third, it helps countries to quickly identify where they need to improve if they want to qualify, since it is clear which indicators they have missed. Fourth, it helps partially alleviate the missing data problem. Presumably, missing data counts as a missed hurdle, but since a country needs to pass only half the hurdles, it can still qualify even if it is missing some data (yet incentives remain for countries to collect more data).

There is nothing particularly magical about the administration's choice for using the median as the hurdle, or for the requirement that a country must pass half the hurdles in each area, or for insisting that countries make the corruption hurdle. The hurdles could have been set at lower or higher levels (e.g., the 40th or 60th percentile) or at specific numbers such as a 70 percent immunization rate. The required number of hurdles could have differed as well. There are clear trade-offs in making these decisions. The higher the standard on each indicator or the more hurdles required, the fewer the countries that will qualify. Thus, these three variables must be decided jointly: the standard on each indicator, the number of hurdles required, and the approximate number of desired qualifying countries.

Implicitly, the hurdles approach takes the view that there are critical values on each indicator that a country must surpass in order to achieve growth and development, or at least to make foreign aid effective. Whereas there is evidence that lower corruption or higher immunization rates are associated with better development outcomes, there is very little evidence on specific minimum levels that must be achieved for better outcomes. Perhaps the one exception is inflation, where strong evidence suggests that inflation rates over 20 percent are particularly detrimental. As mentioned previously, consistent with this finding, the administration has adopted 20 percent as the inflation hurdle.

One weakness of the hurdles approach is that it limits the incentives for countries to continue to improve on the indicators once they have passed the hurdle. A country either meets the hurdle or not—it does not receive additional credit for passing the hurdle by a large margin, nor does it receive a penalty for missing by a large margin. Thus, once a country is above the hurdle, it does not need to improve to continue to be eligible (unless the hurdle rises over time). Moreover, countries with scores well below the hurdle receive no credit until they actually pass it—there is no benefit for improving, say, from the 10th percentile to the 40th. For some countries starting from a very poor, resource-constrained situation (e.g., Mozambique or Rwanda), it may be many years before they are able to pass some of the hurdles, especially the "investing in people" indicators.

The choice of using the median as the hurdle raises two major concerns. First, the median will change over time. Countries with an immunization rate that is too low in one year could pass in the next year with the same immunization rate if the median falls. Conversely, a country that meets the standard in one year could find that it does not meet it in the next year if the median rises. Moreover, countries that nearly qualify in one year and work to raise their scores may be disappointed to find that their improved scores fall below new, higher medians. Note that this problem would occur if any percentile score is used, not just the 50th percentile (the median). Moving scores are inconsistent with the idea of choosing countries that meet a minimum standard of policy quality, which should

not change arbitrarily from year to year (although it could be gradually increased by a preannounced amount over time).

Second, using medians as benchmarks severely limits the potential for the number of MCA countries to expand over time. As "near miss" countries raise their scores to try to qualify, the medians will rise, so other countries will be bumped off the list. Under the methodology proposed by the administration, it is highly unlikely for the number of MCA countries to expand much beyond 20, even if over time dozens of countries improve their scores above the first-year medians.

The best way to address these concerns is to set absolute standards for hurdles where possible (as the administration has done with inflation), perhaps determined by the median in the first year. Thus, if the median immunization rate in the first year is 70 percent, the hurdle in each subsequent year would be 70 percent (or as a variant, the hurdle could gradually increase over time). However, this change can be made immediately for all four "investing in people" indicators, the budget deficit, and days to start a business. But this approach cannot be used for the subjective indicators that are always measured on a relative scale, such as corruption (at least as those indicators are currently measured). Nevertheless, using absolute instead of relative hurdles where possible would improve the system.

Missing data pose a different issue for the hurdles approach. The administration counts missing data as below the hurdle in determining scores, which makes sense. But how should missing data be treated in calculating the median? Omitting these data assumes that the missing entries are normally distributed around the median—that is, the median would be the same whether the missing entries were included or not. However, missing data tend to come either from very small countries where surveys are not completed or from poorly performing countries. To the extent that most missing data come from poor performers (which is probably the case), omitting their low scores tends to increase the median. Thus, it is quite possible that a country with a score just above the *true* median would end up below the *observed* median when data from some countries are omitted. In effect, this country is penalized (and could miss qualifying) because of missing data from other countries.

A final and important concern with the hurdles approach is errors in the data. Margins of error in estimating the indicators can be a significant problem, as highlighted in a recent paper by Daniel Kaufmann and Aart Kraay (2002a). Many of the indicators are based on survey data, including all six of the "ruling justly" indicators and several of the "economic freedom" indicators (regulatory quality, credit risk, and even inflation, which is based on price surveys). Survey results are always estimated with margins of error. Indeed, even the nonsurvey-based indicators are estimated with margins of error, albeit from different sources (e.g., immunization rates are estimated by vials of vaccine distributed, which is an imperfect

gauge of actual immunizations). The problem is that for a country with an *observed* score just below the median on any indicator, we cannot have a high degree of confidence that the *true* level is below the median. Margins of error in the estimation could be the difference between passing a hurdle or not. By contrast, some countries that have *observed* scores above the hurdle may have *true* levels below the hurdle, and thus receive passing grades when they are not warranted.

The administration reduces the potential problems stemming from measurement errors by requiring that countries make only half the hurdles. Thus, even if a country misses a hurdle because of bad data it can still qualify for the MCA (and might even benefit from bad data on a different hurdle).

There is one indicator, however, in which measurement errors remain a major concern: corruption. This is not because the data on corruption are any less reliable than for the other indicators—there is no evidence to suggest this is the case—but rather that a country that scores below the median on corruption is eliminated from qualifying for the MCA, regardless of its scores on other indicators. This procedure raises the possibility that a country could be eliminated from the MCA simply because of bad corruption data. As Kaufmann and Kraay (2002a) point out, for many countries with estimated levels of corruption near the median, one can be only 90 percent certain that the actual level is somewhere between the 40th and 60th percentiles. Using the 2000 corruption data (not the most recent version), they examine the corruption indicator for 61 countries with available data from the pool of 74 countries eligible for the MCA in the first year. For 21 of these countries, there is a 75 percent or higher probability that the actual score is above the median, and for 17 countries there is a 75 percent or higher probability that the actual score is below the median. But there are 23 intermediate cases in which there is much less certainty about whether they actually fall above or below the median. Thirteen of these countries have estimated scores that fall below the observed median and are therefore eliminated from the MCA, despite this uncertainty.

Although I have great sympathy for a high standard on corruption, the make-or-break requirement may unnecessarily eliminate some countries. One alternative would be to eliminate only the 17 countries with a 75 percent or higher probability of an actual corruption score below the median. Other countries would remain eligible for the competition, following the other rules set by the administration. Thus, if a country is not one of the 17 eliminated but scores just below the median on corruption, it would not get credit for the corruption hurdle but could qualify for the MCA so long as it passes half the hurdles in each of the three categories. Another alternative would be to fully eliminate only the countries with corruption scores in the bottom quartile. Changing this approach, however, is bound to be difficult, as no one will want to appear to be "soft" on corruption because of what may appear to be an arcane statistical problem.

Aggregate Ranking Approach

An alternative to the hurdles approach is to rescale each indicator and then add the scores to create a final tally. Countries can then be ranked from the highest to lowest score, and the administration can choose, say, the top 15 or 20 to qualify for the MCA. The simplest way to do this is to rescale each indicator so that the mean score is reassigned a value of zero, and the values that are one standard deviation above or below the mean are reassigned values of 1 and –1 respectively. All other scores are converted accordingly. This is a common statistical approach in aggregating numbers with different scales. It is used, for example, in compiling the KK measures that are used for 5 of the 16 MCA indicators. Once the 16 indicators are rescaled, they can be added, giving each country an aggregate score. Different weights can be assigned to different indicators or to groups of indicators. Countries can then be ranked from best to worst scores.

One advantage of this approach is that it avoids the need to establish (rather arbitrary) hurdles that a country either passes or fails. Also, countries are given more credit for a higher score on any indicator, so they continually have an incentive to improve even if they are above the hurdle. It also significantly reduces (though does not eliminate) the issues surrounding margins of error discussed previously, as it is of no concern if a country barely misses one or more hurdles. This method could be combined with the hard hurdle for corruption, if desired. That is, a country's final ranking would determine its MCA qualification, but if it scored too low on the corruption index, it could be eliminated, regardless of its overall ranking.

One drawback is that a particularly high or low score could significantly alter a country's overall score. A country with a very high score on one indicator could qualify, even if it has poor scores on most of the other indicators. Moreover, missing data are a concern in this approach, as it is not clear what value to add to a country's score. One approach would be to add the lowest score achieved by any other country. Another would be to give the country its average on other indicators, so that the missing value does not affect its overall score. Both of these approaches, however, are problematic.

Once an overall ranking is tabulated, the administration would have to choose where to draw the line between the qualifiers and the nonqualifiers. This choice would be arbitrary, similar to the arbitrary choice of standards on each indicator in the hurdles approach. This approach requires only one line to be drawn, but it is an important line and will be seen as obviously arbitrary. Thus, while this approach has the advantage of allowing a country to see exactly where it ranks vis-à-vis other countries, it could create diplomatic pressures on those administering the MCA. For example, if the top 15 countries were chosen to qualify, there would be immense pressure from the governments of the next several countries to include them and draw the line at 18 or 20 countries. It would

be very difficult for the administration to defend the choice of 15 countries rather than 18 or 20 on analytical or technical grounds.

Moreover, this approach makes it difficult to compare country performance over time. By combining as described, this approach only measures a country's performance relative to its peers. It cannot show if all countries are getting better or worse. In other words, if a country moves up from 21st to 19th, is it because it got better or because the others got worse? This is a difficult issue for *both* the aggregate ranking and the hurdles approaches (since the latter uses relative scores like medians as hurdles). When scores are measured relative to other countries, it is difficult to observe all scores rising so that more countries can qualify over time.

In sum, there is no perfect way to aggregate across indicators. Either method requires arbitrary judgments and raises some difficulties with measurement errors, relative rankings, and other issues. Either method will lead to some surprises, both in terms of countries that miss qualifying and some that qualify. The imperfections inherent in the underlying data are magnified when combing across such different indicators. The administration should continue to examine the data with *both* the methodologies and improve them over time, even if only one method is the official procedure.