

Chapter 2

Evaluation Highlights

- DB collects its data largely from lawyers and accountants deemed knowledgeable about a country's laws and regulations.
- The number of informants who fill in questionnaires on any topic in a country is small.
- DB does not keep track of those who are invited but do not participate; thus participant bias cannot be estimated.
- The DB team validates and adjusts informants' data, which makes the data difficult to verify.
- DB regularly updates its data to reflect changes in methodology or correct errors, but previously published data sets are not made available to users, and the impact of such changes on the overall and indicator rankings is not stated.
- For a given amount of change in an indicator, a country's ranking may change a little or a lot, depending on its initial position in the ranking.



Workers in a furniture factory, Cotonou, Benin. Photo reproduced by permission of Jorgen Schytte/Still Pictures.

Collecting Information and Constructing the Rankings

Doing Business creates its annual country rankings using information supplied by persons deemed knowledgeable about selected laws and regulations in each country covered. The DB team identifies individual lawyers, notaries, officials, and firms and requests that they provide information on one or more specified DB topics. Since the process is not based on a survey sent to a large group, but rather on information solicited from selected individuals, the term “informants” is used in this evaluation instead of “respondents.”

The validity of the DB indicators depends on how representative, reliable, and objective its process is for obtaining, recording, and analyzing information (see Dorbec 2006). This chapter assesses DB’s processes for interviewing informants, reviewing and validating the information they supplied, and constructing the rankings.¹

The Number of Informants

The 2007 and 2008 DB reports note that about 5,000 individuals provided information for the indicators. It is frequently the case that several individuals from the same firm or office help prepare the firm’s response to the questionnaire. For example, junior staff may obtain data for a partner or principal to compile into the firm’s written submission. The DB reports list each of these individuals as informants, but in this evaluation, each completed questionnaire is counted as one informant, irrespective of how many individuals helped to prepare it. Some of the listed individuals completed a DB question-

naire; they are called “questionnaire informants”; others were consulted by DB to confirm or clarify selected points; they are called “supplemental informants.”

The number of informants on each topic in a country is small. For its 2007 report, DB received, on average, *Number of informants for each topic in a country is small.* between one and four completed questionnaires per topic and consulted with up to three supplemental informants per topic for the 5 focus indicators in the 13 randomly selected countries of this evaluation, as summarized in table 2.1.²

The *starting a business* indicator has the most informants in each country—3.5 on average—perhaps because it has been used since 2004 and the questionnaire is relatively simple. But for the *paying taxes* indicator, there is a single survey informant in 142 countries—the local or regional office of the global accounting firm

Table 2.1: Average Number of Completed Questionnaires per Indicator in Each Country Is Low

	DB dimensions					
	Employing workers	Enforcing contracts	Getting credit			Starting a business
Legal rights			Private/public credit bureau	Paying taxes		
Average number of questionnaire informants	1.7	1.8	1.5	1.5	1.0	3.5
Average number of supplemental informants per country	0.2	1.3	0.6	0.0	3.0	1.5

Note: Averages are calculated by dividing total informants by 13 countries, except for *getting credit - public/private credit bureau*, which includes 10 countries because Albania, Moldova, and Tanzania do not have credit bureaus.

PricewaterhouseCoopers LLC (PwC), with several people from each office contributing to one questionnaire for each country. This is because DB has established a partnership with PwC's global tax practice (described in chapter 3), in which PwC is the sole informant on the *paying taxes* indicator (except in the 33 countries where PwC does not participate at all). The information for the *enforcing contracts* and *employing workers* indicators is also based on not more than 2 completed questionnaires in each of the 13 countries.

Where information is factual, the small number of informants may not matter, but subindicators on time and costs include informed estimates.

To the extent that DB collects factual information, as distinct from opinion or perception, it arguably does not need a large number of informants to lessen the source of error, as do perception or opinion surveys. But, as will be noted in chapter 3, not all of

DB's information is purely factual. The time and cost subindicators, for instance, require informants to make estimates based on their experience. Increasing the number of informants would reduce both the risk of erroneous factual information from a single informant and the errors inherent in questions requiring informants' judgments. An additional potential risk noted by some Bank Group staff is that interested parties could seek to influence informants' responses to improve their country's ranking. There is no

There are no clear selection criteria for informants.

evaluative evidence suggesting deliberate manipulation of data, but close relationships among lawyers and officials, especially in smaller countries, could

impair informants' objectivity. Informants in Madagascar, for instance, are routinely in direct contact with the government department responsible for enhancing the investment climate. A larger number of informants would dilute the influence of any self-interested responses.

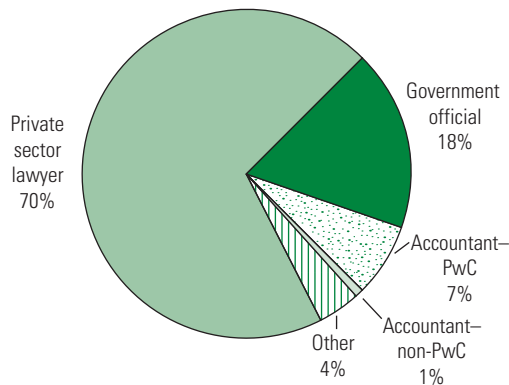
DB actively seeks new informants through referrals from existing informants, IFC and Bank contacts in borrowing countries, and local business directories. The DB Web site now invites prospective informants to register their interest in becoming data contributors. But there are no clearly stated criteria or processes for seeking and selecting informants.³ To ensure that these efforts pay off in progressively more reliable data, DB should consider: (a) targeting them to the countries and indicators most needing increased reliability and (b) establishing selection criteria and numerical goals for new informants. It should report systematically on these activities.

Qualifications and Motivations of Informants

More than two-thirds of DB informants in the 13 countries reviewed by this evaluation are lawyers from private firms, as shown in figure 2.1. Eight percent are accountants, most from PwC accountancy offices.⁴

Informants interviewed for the evaluation were professionally engaged on the topic on which they provided information.⁵ The DB team has noted that in small and/or low-income countries,

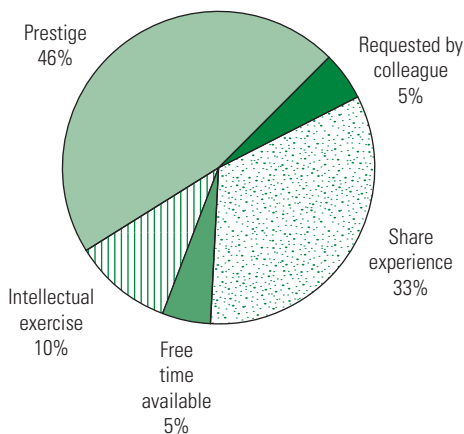
Figure 2.1: The Majority of DB Informants Are Lawyers and Accountants



Note: Based on 141 questionnaire and supplemental informants. "Other" includes members of the private sector and one administrative staff member and graduate student in China.

it sometimes must rely on lawyers in general practice because there are no lawyers with specialized practices in bankruptcy, civil claims, or other relevant areas. In Burundi, for example, one informant was primarily a criminal and family lawyer with minimal experience in corporate law, and some informants for *closing a business* were unaware of the country's bankruptcy law. DB could reduce the risk of error through consulting in advance with Bank and IFC country counterparts in identifying informants and by undertaking greater quality assurance in coun-

Figure 2.2: Why Do You Participate? For Prestige and to Share Expertise



tries with few specialized informants or a weak informant base. DB should also consider developing a systematic procedure or set of criteria for assessing informants' qualifications.

Generally, informants in the case study countries were professionally engaged on the topic they addressed.

DB does not pay its informants. It simply acknowledges the participating individuals in its publications (except for the approximately 10 percent who do not wish to be publicly named).⁶ Assisting DB can require considerable effort; informants interviewed for the evaluation said they spent between one hour and one month on the exercise. The *paying taxes* and *getting credit* indicators are the most demanding of informants' time.

Why, then, do the informants participate? About half of those interviewed for the evaluation, as shown in figure 2.2, said participation would enhance their firm's credibility or prestige (46 percent). Another third said they wanted to share their experience (33 percent), and the rest said they were interested in the intellectual exercise, had free time, or were asked to participate by somebody else. *Informants generally participate for prestige or to share experience.* Several lawyers mentioned they had gained clients who heard of them through the questionnaire, although attracting clients was not their objective in participating. In some Organisation for Economic Cooperation and Development (OECD) countries, lawyers noted that the time spent on DB counts toward the firm's commitment to provide *pro bono* services.

Not all those invited to participate actually do so; some decline, others agree but fail to follow through. DB does not keep track of nonparticipants and why they do not participate. If nonparticipants are systematically different from the actual informants, there is a possible selection bias, whose direction depends on the reasons for participation. For example, if informants are more likely than nonparticipants to be concerned about excessive business regulation in their country, the data may overstate the regulatory burden. If nonparticipants from small firms tend to decline because they lack junior staff to help

complete a complex questionnaire, the data may understate the regulatory burden, because the informants will tend to be those with more capacity to help clients navigate the bureaucracy. Finally, if nonparticipants tend to be those too busy to do *pro bono* work, while informants are those with a lighter workload, the responses may be less reliable to the extent that the informants are those less well established in their field.

DB does not keep track of those who are invited to participate but do not, which makes it difficult to identify and disclose the nature and direction of participant biases.

Like most surveys and polls, DB involves participant bias, and it needs to systematically learn and disclose more about the magnitude and directions of nonresponse. In this regard, it may be useful to systematically collect and track information on the number of informants who were contacted, but who did not qualify as informants, were not interested in participating, or refused to participate for other reasons. DB should also consider diversifying its informant base to include business consulting firms, associations, and think tanks that meet the selection criteria to be developed.

Validating the Data

The DB team validates and adjusts the data based on documents and supplemental informants, but this makes it difficult to verify the data.

Once the informants' questionnaires are received, the DB team validates the information based on documents and consultations with supplemental informants. Typically there are four rounds of interaction between the DB team and the informants, involving conference calls, written correspondence, and in some cases a country visit. When informants' estimates of time differ, DB states that it selects the median value (World Bank-IFC 2006b, p. 61).

The evaluation reviewed the differences between the information provided in the completed questionnaires and the data points as reported.⁷ For *employing workers* and *getting credit*, in all seven case study countries, *Doing Business 2007* published values that differed from at least one questionnaire, based on further consultation by the DB staff

with the informants and/or review of legislation. In the case of time estimates for *enforcing contracts*, the median value did not appear to have been selected.⁸ Informants in Nigeria gave broad estimates of time, so DB staff calculated it based on a review of the current changes in legislation. In Mongolia, one informant's response was discarded in favor of another that reported no change since last year. Cost indicators may also require DB to select from different estimates. *Doing Business 2007* stated that in Spain, enforcing a contract cost 15.7 percent of the disputed debt, while the two informants had estimated 15.1 percent and 18.9 percent. DB explained that the higher estimate was disregarded because it came from a large international law firm. The estimate has subsequently been revised to 17.2 percent, close to the median of the two original responses.

DB's close attention to individual data points and its resolution of differences and anomalies undoubtedly help improve the quality of the database. The process also helps identify and weed out any unreliable informants. The DB team's organization along topic lines permits staff to develop a feel for plausible levels and ranges of the indicator values. At the same time, a risk is created by the considerable reliance placed on the decisions of DB staff to accept, overrule, or select among informants' replies, because this makes it difficult to verify or replicate the data. This risk is partly mitigated by a validation process in which DB sends proposed country data to relevant Bank and IFC staff and to country authorities (through their executive directors) for comment.

Bank Group staff and in-country stakeholders in 6 of the 13 countries reported dissatisfaction with the DB's process for validating the data.⁹ Albanian officials, for instance, considered that the rankings for *starting a business*, *enforcing contracts*, and *protecting investors* do not square with the facts on the ground and found the response of the DB team to their rebuttals "not satisfactory." DB received 115 specific challenges and clarifications to its 2007 report from 50 country teams; 21 percent of these

Country stakeholders and Bank Group staff in 6 out of 13 countries reported dissatisfaction with DB's process for validating data.

challenges were accepted.¹⁰ Some Bank Group staff noted that despite recent improvements, DB still gives them insufficient time (just a few days in some cases) to review draft DB data. Some interviewees considered that their team's challenges were rejected by DB without due consideration. DB should consider how to devise a more open and in-depth validation process to help increase both the quality and the credibility of the data.

Publishing and Revising the Data

DB publishes its data and country rankings in its annual report each autumn and on its Web site. It makes ongoing changes to previously published data, and the Web site indicates that it contains the most current version. The DB 2007 data presented on the Web site in October 2007 had 2,284¹¹ differences (on the total 5,600 data points used to calculate the EODB ranking) from the data originally published in the DB 2007 report (see appendix C). The Web site does not provide nor link to the original data set. In the DB 2008 report, the revised 2007 data were used as the comparator for the previous year.

The DB 2008 report (World Bank-IFC 2007b, pp. 67–69) notes that data changes have been made and gives three reasons:

- Changes in methodology for three indicators: *enforcing contracts*, *dealing with licenses*, and *employing workers*. (Separately, the report indicates that the methodology for *paying taxes* was also changed [World Bank-IFC 2007b, pp. 78–79].)
- Corrections in 47 data points.
- Addition of three new countries.

The evaluation team's review of the changes to the DB 2007 data found that:

- Changes in methodology for the four indicators account for 1,284 changes in the DB 2007 data (56 percent of total changes). DB has not indicated how the 2008 methodology was retroactively applied to information that informants supplied in prior years on the basis of different assumptions or definitions. The DB team has ex-

plained that for three of these indicators (*paying taxes*, *enforcing contracts*, and *dealing with licenses*), some of the changes may also reflect data corrections, but that “it is difficult to separate corrected errors from methodology revisions.”¹²

- Corrections account for 1,000 changes in the DB 2007 data (44 percent of total changes). The DB team has not indicated the reasons why information supplied by *DB regularly revises previously published data . . .* team has explained that “minor” changes to data for the prior year—that is, changes amounting to 10 percent or less of the original value—are made without further investigation. Of the 1,000 corrections, 222 (22 percent) were “minor.”¹³
- The addition of three new countries affected the rankings but not the underlying data being discussed here.

The practice of changing previously published data can be helpful in improving the reliability and consistency of a data set. . . . *but does not make available previously published data sets.* At the same time, to fulfill its objective of facilitating research and informing theory, DB should disclose all such corrections and changes and explain their effects on the rankings as explained below, and make available previously published data sets.¹⁴

Effects of data changes on country rankings and top reformers

The DB Web site presents 2007 EODB rankings and a top reformers list derived from the revised data. It does not state how the changes in data affected these rankings. The evaluation finds that the 2,284 changes resulted in changes to the rankings for 106 countries (even after accounting for the addition of three new countries in 2008).¹⁵ Twenty-four countries improved 10 or more positions and another 24 dropped 10 or more positions on the EODB ranking. The most significant changes are listed in table 2.2. The roster of top reformers also changed. Latvia entered by moving up from eleventh to tenth, and Ghana exited by falling from ninth to nineteenth. *Some data changes have had significant impact on ratings.*

Table 2.2: Large Changes in 2007 Rankings Resulting from Data Revisions

Five biggest winners				Five biggest losers			
Country	EODB October 2007	EODB August 2007	Change	Country	EODB October 2007	EODB August 2007	Change
Guyana	96	136	+40	Nicaragua	85	67	-18
Italy	49	82	+33	Samoa	59	41	-18
Turkey	64	91	+27	Tajikistan	151	133	-18
St. Kitts & Nevis	61	85	+24	Papua New Guinea	79	57	-22
Bhutan	120	138	+18	Uruguay	87	64	-23

DB should fully explain the nature and extent of periodic data changes and their implications for the rankings.

Some data changes nullify reforms cited in the text of the DB 2007 report. For example, 23 countries earlier identified as having reduced their corporate taxes show no changes in tax rates using the revised data. DB highlights reforms to *getting credit* in Italy and *trading across borders* in China, but the revised data revealed that the relevant subindicators actually deteriorated (World Bank-IFC 2006b, pp. 3, 30). DB should make clear that its rankings are subject to change and fully explain the extent, nature, and implications of these changes on country rankings.

Constructing the Rankings

The DB process first establishes cardinal values for each subindicator: time, costs, number of procedures, and the like. Countries are ranked on each subindicator. The subindicator percentiles are averaged to come up with an indicator-level ranking. The 10 indicator percentiles are then averaged to generate the overall EODB ranking.

DB's reliance on successive stages of ordinal rankings obscures the underlying cardinal values. The magnitude of the difference between the countries is not the same on all points of the distribution. For example, on total tax rate, there is a 5.1 percentage point difference between the top two performers, Maldives and Vanuatu, and a 4.7 percentage point difference between the bottom two, Gambia and Burundi. But the countries

ranked fifty-ninth and sixtieth (Israel and Mozambique) are separated by just 0.1 of a percentage point.

A given change in a cardinal value (for example, a reduction in the number of days needed for a procedure) is more likely to advance a country's rank, holding other countries' actions constant, if the country starts from a more concentrated segment of the distribution than if it starts from a more dispersed section. This arithmetic means that countries at the more dispersed parts of the distribution have to work harder to change their overall ranking. Countries can make significant changes, yet fail to improve their rankings, if they are at the dispersed sections of the distribution for that indicator. The following three examples illustrate this asymmetry by simulating the change in rankings for a subindicator, holding the actions of the other countries constant (see appendix B).

- *How does reducing the minimum capital requirement affect ranking on starting a business?* The DB 2008 report notes that Egypt drastically reduced its minimum capital requirement from 695 percent of income per capita to just 13 percent. Holding other countries' actions constant, this reduction would have boosted its ranking by 33 positions.¹⁶ Although Gambia, Macedonia, and Saint Kitts and Nevis all reduced the minimum capital requirement much less than Egypt in absolute terms, they would have boosted their simulated rankings much more than Egypt. By eliminating the minimum capital requirement, these

A country's location in the distribution affects how a given reform will change its ranking . . .

3 countries would tie with the 66 others for first place on this subindicator.

- *How much does the tax rate have to fall to improve ranking on paying taxes?* With a 43 percentage point reduction in its total tax rate, Sierra Leone would improve only one position in the simulated ranking for *paying taxes*. But Latvia, by reducing the total tax rate by just 10 percentage points, could improve 17 positions because it is situated in the most concentrated segment of the distribution.
- *How does reducing the time to open a business improve starting a business?* *Doing Business 2008* notes that the Republic of Lao reduced the time to start a business by 60 days (36 percent of its starting value). Yet this change would not affect its simulated ranking for *starting a business*. Mauritius, by contrast, reduced the time by 41 days (89 percent of its starting value), and would thereby advance 20 positions on the simulated ranking for *starting a business*.

It has been suggested that DB's use of rankings might create an incentive for a country to reform

the areas where it can most improve its ranking for the least reform effort. If this were the case, one would expect the highly concentrated subindicators to be associated with more reforms.¹⁷ But the correlation between tightness of distribution and frequency of reforms is almost nonexistent (0.01).¹⁸ The total tax rate is the third most frequent area of reform, and it has the tightest distribution of all the subindicators, with 94 percent of the countries' rankings within one standard deviation from the mean. But the two most popular areas for reform—number of procedures to start a business and legal rights of creditors and debtors—are not among the most tightly distributed. This quantitative analysis is fully consistent with the finding reported in chapter 4—that IEG did not find evidence of countries making superficial changes for the sole purpose of improving their rankings. Nevertheless, the DB team may wish to consider ways of making the rankings more informative, perhaps by establishing country groupings that reflect the cardinal values of each indicator.

... but IEG did not find evidence that countries use this characteristic of the ranking system to manipulate their DB indicator ranking.