1. Background and Context

1.1 Since 2003, the annual Doing Business (DB) project has aimed at providing objective measures of business regulations and their enforcement across 190 economies. DB provides comparable, quantitative indicators on business regulations, legal protections, and business processes for 190 economies that can be tracked over time. It also provides an overall ease of doing business” score and ranking by aggregating the indicators for 10 areas, which it claims assesses “the absolute level of regulatory performance over time.”

1.2 From the start, the stated aims of DB were (i) “motivating reforms through country benchmarking”; (ii) “informing the design of reforms”; (iii) “enriching international initiatives on development effectiveness”; and (iv) “informing theory.” Regarding (i), DB intended to mobilize demand for reform of the regulatory environment for business. Regarding (ii), DB wanted to target “what needs to be changed” and disseminate “the experience of countries that perform well according to the indicators.” Regarding (iii), “recognizing that aid works best in good institutional environments,” DB aimed at explicitly allowing donors to base aid on “good-quality data that can be influenced directly by policy reform.” Finally, regarding (iv), DB intended to inform the field of regulatory economics by facilitating “tests of existing theories” and contributing to “the empirical foundation for new theoretical work on the relation between regulation and development” (World Bank 2004).

1.3 DB is recognized as highly influential in business regulatory reform worldwide, and it is the most used set of indicators on business regulation (Besley 2015; Doshi, Kelley, and Simmons 2019; Haidar 2012). Its indicators are widely used and analyzed in the academic literature. They are a component of many other influential indexes, including the World Economic Forum’s Global Competitiveness Index, the Heritage Foundation Index of Economic Freedom, and the Fraser Institute Economic Freedom Index. It is cited by many countries in their reform plans and in many World Bank Group project documents and country strategies.

1.4 Although popular, the DB indicators have also been the subject of controversy regarding their methodology, accuracy, and potential biases and the way they are used in shaping and assessing country policy reforms. The Bank Group and the Independent Evaluation Group (IEG) have been called on several times to review DB, largely to
respond to such criticisms. Each review recommended continuation of the indicators but also identified flaws (box 1.1).

<table>
<thead>
<tr>
<th>Box 1.1. Main Findings on Shortcomings of Doing Business Indicators from Past Independent Evaluation Group and World Bank Group Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Independent Evaluation Group and World Bank Group reviews of Doing Business (DB) to date recommend continuation of the indicators, finding substantial benefits in their motivation of reform. Each, however, also found shortcomings in the connection of the DB indicators to a reform agenda or in capturing reform progress or impact:</td>
</tr>
<tr>
<td>• They focus primarily on compliance costs, with several exceptions (World Bank 2015, 2019b).</td>
</tr>
<tr>
<td>• They do not track reform impacts (World Bank 2008, 2013).</td>
</tr>
<tr>
<td>• They cannot fully capture country-specific nuances and policy idiosyncrasies. Thus, they have had less influence on designing reforms than on spurring debate (World Bank 2008).</td>
</tr>
<tr>
<td>• The practice of aggregating DB indicators is controversial (Doshi, Kelley, and Simmons 2019; World Bank 2013).</td>
</tr>
<tr>
<td>• DB adjustments to indicators generate discontinuities (Morck and Shou 2018; World Bank 2019b).</td>
</tr>
<tr>
<td>• They exclude important legal and regulatory areas (World Bank 2008 2015).</td>
</tr>
</tbody>
</table>

Source: Independent Evaluation Group; see appendix B.

1.5 The scope of the DB report has grown from 5 to 12 topic areas (figure 1.1), including indicators on the cost and quality of business regulation and on the quality of legal frameworks. Country coverage has expanded from 133 countries in 2003 to 190 currently. Data are collected in a standardized way using a simple business case to ensure comparability across economies and over time. Several of the indicators capture the experience of a typical firm, likely a small or medium enterprise, or transaction, requiring assumptions about a business’s legal form, size, and location and the nature of its operations. Data are collected through a questionnaire administered to local experts, including lawyers, business consultants, accountants, freight forwarders, government officials, and other professionals routinely administering or advising on legal and regulatory requirements. The methodology assumes that a business has full information about prevailing procedures. Indicators usually describe conditions in the largest business city, except in 11 countries where the second-largest business city is included. Although initially DB indicators focused only on compliance time, cost, and procedures, several subindicators have been added over time, measuring other qualitative dimensions of law, regulation, or service quality.
1.6 The Committee on Development Effectiveness of the Bank Group Board of Executive Directors requested IEG to assess the way that DB aligns with the Bank Group mandate and a viable development framework for client countries. Recognizing “the importance of DB as a reference for many stakeholders and a valuable tool to benchmark business reforms to attract private investments” and to ensure the credibility and reliability of DB, the committee requested “the Group of Internal Audit (GIA) and IEG to conduct their respective assessment … of the DB, consistent with their respective mandates, and to present findings and recommendations to the Board.” In response, IEG has committed to examine the relevance and effectiveness of the use of DB indicators in guiding client country business environment reforms—both those supported by the Bank Group and those undertaken without its support. This includes an initial stocktaking of literature and existing evaluative evidence to inform an Issues Paper, which will be followed by a Focused Evaluation to assess the DB’s strategic relevance to countries’ reform priorities and to the Bank Group’s strategic agenda. This request came just before the late-August 2020 suspension of the DB report to probe alleged irregularities in the underlying data.

1.7 Two additional assessments of DB are being conducted in parallel by the Bank Group’s Group Internal Audit and Development Economics units. Group Internal Audit will be examining the integrity of processes related to DB including data collection, data review, and controls related to the end-to-end DB report production life cycle. The Development Economics unit (where the team producing DB has been located since 2014) will be commissioning an expert review of many technical aspects of DB, including selection of topics, suitability of its case study approach and data gathering processes, scoring methodologies, and the balance between indicators measuring
different regulatory attributes. It will also conduct a special study on the methodology of the total tax contribution rate component of the paying taxes indicator. IEG intends to maintain ongoing communication with the relevant teams to maximize complementarity.

2. Doing Business Theory of Change

2.1 DB influences country reforms through several channels, informing and motivating a variety of stakeholders (figure 2.1).:

- Government officials may be motivated by considerations of public policy, international standing, or foreign investment attraction to advance their rankings where they are low. DB intends that governments compete to improve their rank among nations.

- Donors may focus projects on areas where indicators are lagging or encourage governments to reform policies and practices that result in low ranking and may also take guidance on reform design. Donors may support or advocate for reforms to improve weak areas as measured by DB. For example, IEG estimates that close to 500 Bank Group operations during the past 10 years were explicitly DB-informed.

- Investors—both domestic and foreign—may encourage or lobby governments to address conditions related to lagging indexes.

- Civil society groups may play a role as well, for example, to argue for more transparent, affordable, or accessible procedures.

2.2 The theory of change posits that as a result of the information and influence created through DB, governments introduce reform measures whose implementation progress is then measured by the DB scores. The choice of reforms may be influenced by those identified as best practices by DB. For example, DB estimates it receives 160 queries a year from governments for policy reform advice and information. The reforms introduced are expected to reduce the time and cost of compliance with regulations and improve the quality of regulatory and administrative procedures in the country. This in turn should improve market functioning—business entry, exit, and operation—and reduce transaction costs to enterprises arising from regulatory compliance. As a result, there is an expectation of private enterprise growth in terms of new investment, employment, and business volume turnover. Although there is no claim that DB reforms are a panacea for social and economic development, since its outset, DB has stated the link between the development outcomes associated with the regulatory reforms it motivates and poverty reduction and prosperity.
Figure 2.1. Preliminary Theory of Change: How Doing Business Is Expected to Influence Country Outcomes


3.1 The Bank Group uses DB in (i) country strategy and policy dialogue; (ii) operations (both financing and advisory); and (iii) research and global knowledge sharing. The Bank Group supports country strategy and programs using DB indicators. Examples include policy development and regulatory streamlining in the Philippines (Country Program Evaluation [CPE], fiscal years [FY]09–18; World Bank 2019c); policy dialogue and advice in Mexico (CPE FY08–17; World Bank 2018a); and policy and project support in Tunisia (CPE FY05–13; World Bank 2014) and Rwanda (CPE FY09–17; World Bank 2018b).

3.2 Regarding operations, a look at FY10–20 suggests that a sample of more than 500 projects belong to the DB-informed portfolio—2.6 percent of the overall portfolio of World Bank and International Finance Corporation operations over the 10 years. This sample consists of investment projects or advisory services that are explicitly informed by DB (table 3.1). A further 6.9 percent of the overall portfolio is concerned with regulatory reforms but is not explicitly informed by DB. The DB portfolio is divided primarily between World Bank advisory and analytic services and International Finance Corporation lending and advisory services (table 3.1). Based on the sample by client country income level, the largest percentage of projects is estimated to be in lower-middle income countries (41 percent); with 32 percent in low income countries, 21 percent in upper-middle income countries, and 5 percent in high income countries.
About 19 percent of this portfolio focused on general business environment, and the rest focused on a single DB indicator, led by starting a business, dealing with construction permits, trading across borders, getting credit, and registering property.

**Table 3.1. Estimate of Doing Business–Informed Portfolio by Instrument, FY10–20**

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Projects (no.)</th>
<th>Projects (percent)</th>
<th>Volume ($)</th>
<th>Volume (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>World Bank lending</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPO</td>
<td>72</td>
<td>51.4</td>
<td>5,781</td>
<td>50.3</td>
</tr>
<tr>
<td>IPF</td>
<td>64</td>
<td>45.8</td>
<td>5,519</td>
<td>48.0</td>
</tr>
<tr>
<td>P4R</td>
<td>4</td>
<td>2.8</td>
<td>202</td>
<td>1.8</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>100</td>
<td>11,502</td>
<td>100</td>
</tr>
<tr>
<td>Subtotal financing</td>
<td>140</td>
<td></td>
<td>11,502</td>
<td></td>
</tr>
<tr>
<td>Analytic and advisory activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World Bank ASA</td>
<td>183</td>
<td>47.4</td>
<td>61</td>
<td>100</td>
</tr>
<tr>
<td>IFC AS</td>
<td>203</td>
<td>52.6</td>
<td>347</td>
<td>100</td>
</tr>
<tr>
<td>Subtotal</td>
<td>386</td>
<td>100</td>
<td>407</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>526</td>
<td></td>
<td>11,909</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Independent Evaluation Group portfolio review and analysis.*

*Note: All figures are preliminary and projected based on population and sample sizes of Doing Business projects and volume in dollar-value projections are the result of multiplying sample projects and commitments by a factor of 372/192 for World Bank lending projects, a factor of 916/225 for World Bank ASA projects, and a factor of 380/159 for IFC AS. Volume/commitment for each of the institutions is as follows: World Bank lending = International Bank for Reconstruction and Development + International Development Association amounts; World Bank ASA = total cumulative cost delivered; IFC investment = total original commitments; advisory services = total funds managed by IFC allocated to Doing Business interventions. ASA = advisory services and analytics; DPO = development policy operation; FY = fiscal year; IFC AS = International Finance Corporation advisory services; IPF = investment project financing; P4R = Program-for-Results.*

3.3 Regarding global knowledge sharing and research as a global public good, DB has proven to be a popular resource widely used by policy makers and researchers. It has been used in guiding or monitoring thousands of country-led reforms and in research (discussed below), which indicates the power and usefulness of its indexes. DB has also inspired new metrics within the Bank Group in related projects, such as Women, Business and the Law and Enabling the Business of Agriculture.

3.4 DB tracks reforms across the topic areas measured by its indicators, from getting credit to getting electricity; from starting a business to trading across borders. For example, the starting a business indicator set tracks changes related to the ease of incorporating and beginning operations of a limited liability company. Changes in a country’s performance in an indicator exceeding a defined threshold are classified as reforms and reported by DB. Changes can be classified as positive or negative reforms depending on whether they make it easier or more difficult to do business. Since DB2005, DB has tracked more than 3,800 business regulatory reforms across the 190 economies measured. Among the reforms that DB measures, the most common relate to
getting credit, starting a business, paying taxes, registering property, and trading across borders (figure 3.1). DB also celebrates “top reformers”—those countries that have achieved the greatest number of measured reforms—partly to encourage competition among countries. During the evaluation period (DB2010–20), Rwanda, Kazakhstan, Indonesia, United Arab Emirates, Azerbaijan, Brunei Darussalam, India, Kenya, Armenia, Ukraine, Uzbekistan, and Vietnam achieved the highest numbers of positive national reforms.

Figure 3.1. Reforms Tracked by Doing Business, 2005–20

3.5 DB also figures prominently in the academic literature as a source of data and as a subject for research. There is an academic literature on the nature and influence of DB indicators overall and lines of research in relation to individual areas tracked by DB. Overall, the academic literature credits DB indicators with influencing the regulatory reform efforts of development institutions and governments. At the same time, DB is faulted for excessive formalism, limited scope, and problems in ranking and aggregation practices (Besley 2015; Doshi, Kelly, and Simmons 2019; Haidar 2012; Marechel and Arlet 2020; Yackee 2016; McCormack 2018; Tang 2017; World Bank 2008, 2015).

3.6 In each of its main topic areas of focus, the DB team has tracked the relevant literature from top journals (table 3.2). According to this database, starting a business, trading across borders, protecting minority investors, enforcing contracts, and resolving insolvency are leading areas of focus of the academic literature on DB.
Table 3.2. Academic Articles Based on Doing Business, by Indicator

<table>
<thead>
<tr>
<th>Doing Business Indicator</th>
<th>Papers (no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting a business</td>
<td>79</td>
</tr>
<tr>
<td>Trading across borders</td>
<td>69</td>
</tr>
<tr>
<td>Protecting minority investors</td>
<td>60</td>
</tr>
<tr>
<td>Enforcing contracts</td>
<td>41</td>
</tr>
<tr>
<td>Resolving insolvency</td>
<td>34</td>
</tr>
<tr>
<td>Employing workers</td>
<td>31</td>
</tr>
<tr>
<td>Getting credit</td>
<td>29</td>
</tr>
<tr>
<td>Paying taxes</td>
<td>28</td>
</tr>
<tr>
<td>Registering property</td>
<td>22</td>
</tr>
<tr>
<td>Dealing with construction permits</td>
<td>17</td>
</tr>
<tr>
<td>Getting electricity</td>
<td>16</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>426</strong></td>
</tr>
</tbody>
</table>

*Sources: Doing Business database of academic literature from top journals, 1989–2020; Independent Evaluation Group calculation.*

4. **Scope and Timing**

4.1 Due to the request for an expedited stocktaking and the current challenges of fieldwork during the coronavirus pandemic (COVID-19), the evaluation will be undertaken as a single exercise delivered in two sequential but interlinked stages:

- **Issues Paper.** The first product, to be delivered by the third quarter of FY21, is an issues paper aimed at taking stock of existing evidence on DB indicators and related reforms, identifying key issues, and elaborating a framework and testable hypotheses for an in-depth, focused evaluation to follow. It will identify key issues regarding the DB indicators in terms of their relevance to client country reform priorities, the issue areas each indicator represents (for example, business start-up, business taxation) and Bank Group priorities (see question 1 in paragraph 5.3 below). It will also look to the literature to map evidence of relevance of DB indicators and related reforms. It will draw on the literature, prior IEG and Bank Group reviews and evaluations, and other sources, laying the groundwork for the focused evaluation. The issues paper will be subject to the normal internal and management review for IEG synthesis notes.

- **Focused Evaluation.** The second product, to be delivered by the end of the first quarter of FY22, is an IEG evaluation of DB focusing on the relevance and effectiveness in terms of both immediate and intermediate outcomes. Its questions and methods are discussed in the following section.
4.2 Although this evaluation will answer questions on the relevance and effectiveness of DB indicators as articulated below, the scope is limited. This evaluation will not consider in detail the engineering of indicators or indexes to propose some superior formula or architecture. Some of the more technical aspects of indicator construction and processes will be scrutinized in the parallel Development Economics review (see paragraph 1.7). Neither will it examine the integrity of processes, which is the focus of the parallel Group Internal Audit review. Finally, IEG recognizes that the further one moves from direct outcomes of reforms in the theory of change, the more challenging it is to find evidence of a connection to DB (see figure 2.1).

5. Evaluation Questions and Methods

5.1 This evaluation faces several critical challenges, including the following:

1. **Endogeneity:** There is a potential for endogeneity in measuring results of particular reforms. If a reform is explicitly designed to move an indicator, for example, and the indicator moves, one needs external information to understand whether an economy or society is better off. There are countries that have, for example, streamlined business entry procedures measured by DB without a substantial increase in business entry, due to other unresolved barriers to successful formal business operation. Lowering the tax rate is meaningful only if the rates businesses pay comply with formal rates—and are not negotiated ad hoc by businesses and tax agents.

2. **Data deficits.** Given the challenge of endogeneity, it can be problematic to use DB indicators alone as a measure of successful reform. Yet in many of the areas they measure, there is a dearth of alternative data by which to measure underlying conditions or reform outcomes. Even fairly widespread instruments, such as enterprise surveys, lack the detail and frequency of the DB indicators, and often align only loosely with them.

3. **Identification.** The influence of DB is not always explicit. A country or project may support reforms motivated by a poor DB ranking without explicitly acknowledging this motivation.

4. **Attribution.** Conversely, many of the reforms indicated by DB are also indicated by conventional thinking about business-friendly reform, by other diagnostic instruments, and by businesses themselves. Even if a reform is measured by DB, it can be hard to attribute it to DB. Success has many parents.

5. **COVID-19.** Normal fieldwork involves extensive travel and contact with many stakeholders to better understand how reform initiatives interacted with local
conditions and challenges. COVID-19 necessitates a circumscribed and remote approach to such work.

To address some of these challenges, IEG will build from multiple methods and sources to triangulate in testing key hypotheses.

5.2 The anticipated methods to be applied in the issues paper include the following:

- Portfolio mapping using machine learning and human review will maximize the probability of correctly identifying the Bank Group activities that are informed and motivated by DB.

- A review of country strategy reviews will allow an initial picture of the use of DB in country strategies (reflecting diversity of regions, income levels, types and levels of Bank Group engagement, fragile and conflict-affected situation status, achievement of DB reforms, and so on). This will be complemented by a pilot application of data scraping techniques combined with human review and synthesis to allow a view of country reforms and reform priorities not colored by DB indicators.

- Structured mapping of relevant literature on DB and related reforms will provide an unbiased view of the literature’s treatment of these subjects.

- Synthesis of relevant IEG and Bank Group thematic and country evaluations pertaining to DB will allow full use of the rich analysis that has preceded this exercise.

- Interviews, and panel discussions with internal and external experts will consider a diversity of informed perspectives on DB and its value for reform.

The issues paper is expected to be delivered in the third quarter of FY21.

5.3 The focused evaluation will answer the following questions:

1. *Is DB “doing the right things”? What is the relevance of DB indicators to*

   - The business environment priorities of client countries?
   - The substantive dimensions within each indicator’s focus area?
   - Bank Group strategic and operational priorities?

2. *Are the Bank Group and country client stakeholders “doing things right” when using DB indicators to achieve better development outcomes?*
- Is DB reform achieving desired outcomes?
- Are clients and the Bank Group using DB to achieve improved outcomes for business environment reform (for example, entry, investment, employment) in line with the country’s development policy priorities (for example, public safety, equity, environment)?

5.4 The Focused Evaluation will adopt a theory-based approach that seeks to understand the links among DB, Bank Group interventions, and country-level reforms based on the model elaborated in figure 2.1. It uses mixed methods to understand DB relevance and effectiveness at the global, country, and project or intervention level. The approach will provide information on whether and how Bank Group programs and projects related to DB succeeded, including factors associated with success (or failure). As discussed below, the links to immediate outcomes, such as changes in measured laws or regulations, will be the easiest to demonstrate, with greater caution required in considering the links to intermediate and final outcomes. This Focused Evaluation is expected to be delivered by the first quarter of FY22.

5.5 The mixed methods to be used in this evaluation include those elaborated in appendix A and box 5.1. In general, the hypotheses established in the first product and those arising from work on the second would be tested by triangulating evidence from the multiple sources and methodologies to address the challenges laid out in paragraph 5.1. The intended methodology of the focused evaluation is captured in box 5.1, table 5.1, and appendix A.

**Table 5.1. Evaluation Questions, Methods, and Strengths and Limitations**

<table>
<thead>
<tr>
<th>Evaluation Questions</th>
<th>Data Collection and Analysis Methods</th>
<th>Strengths and Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Doing the right things. What is the relevance of DB indicators to</td>
<td>Literature review; synthesis of available evaluative materials and country case studies; country case studies, PRA, and prior IEG country-based evaluations</td>
<td>• Volume and breadth of available literature</td>
</tr>
<tr>
<td>• The business environment priorities of client countries;</td>
<td></td>
<td>• Access to new tools to enhance identification of correct portfolio</td>
</tr>
<tr>
<td>• The substantive dimensions within each indicator’s focus area; and</td>
<td></td>
<td>• External, independent evaluations of DB use are limited and there are many gaps.</td>
</tr>
<tr>
<td>• World Bank Group strategic and operational priorities.</td>
<td></td>
<td>• Sources on use of DB indicators in client country reforms are inconsistent and there are many gaps.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Limitations on fieldwork due to the coronavirus pandemic</td>
</tr>
</tbody>
</table>
2. **Doing things right.**

Is *DB* effectively achieving desired outcomes?

Is *DB* being used by clients and the Bank Group to achieve the best outcomes for business environment reform (for example, entry, investment, employment) subject to broader policy priorities (for example, public safety, equity, environment)?

**Data Collection and Analysis Methods**

- Structured literature review; PRA of evaluated projects, including PPARs; review of country-based evaluations; country case studies; expert and stakeholder interviews; analysis of *DB* and external indicators of effectiveness

**Strengths and Limitations**

- Consistent triangulation between quantitative and qualitative methodologies to draw findings and conclusions
- Country case studies will be selected based partly on the size and diversity of Bank Group– and non–Bank Group– supported interventions.
- Biases that are inherent in interviews
- Project evaluations rarely show outcomes or whether reforms were sustained.
- Limitations on fieldwork due to the pandemic

**Source:** Independent Evaluation Group.

**Note:** *DB* = Doing Business; IEG = Independent Evaluation Group; PPAR = Project Performance Assessment Report; PRA = portfolio review and analysis.

5.6 IEG anticipates multiple limitations to its proposed mix of methods and will seek to mitigate the risks they impose, including by triangulating from multiple evidence sources to test hypotheses. These limitations include gaps and inconsistencies in the literature and country social and economic data sets; the limited number of relevant evaluated projects and country programs with *DB* relevance (especially in light of the lack of IEG-validated evaluations of advisory services and analytics); the limited ability of a few case studies and deep dives to represent the universe; limited data on intermediate outcomes and sustainability of reforms and difficulties in establishing attribution to specific reforms; and difficulties imposed by COVID-19 in conducting fieldwork (even remotely). There may also be some false negatives or false positives in identifying relevant literature, project activities, and reforms due to the multifaceted nature of areas covered by some *DB* indicators, lack of standard terminology, and the sometimes indirect nature of influence. Establishing causal connections between *DB* indicators and observed reforms, and between reforms and observed outcomes, is complicated by the likelihood of multiple causal factors.

**Box 5.1. Planned Methodologies for Doing Business Evaluation**

- **Portfolio review and analysis.** The team will conduct a systematic desk identification, review, and assessment of projects informed by *Doing Business (DB)* along the evaluation’s analytical framework to identify design features and characteristics, results indicators, and drivers of success and failure. The entire portfolio will be initially captured and classified, assisted initially by a machine-learning approach supplemented by human review. This will be followed by detailed codification and analysis.
• **Structured literature review.** The team will commission a structured review after
the Independent Evaluation Group’s protocol covering academic and World Bank Group
literature on the DB indicators and related reforms in general and for the indicators
selected for deep dives. Initially, the exercise will map the DB literature broadly. As time
and resources permit, additional indicators may be covered.

• **Indicator-specific deep dives.** Deep dives will include a focused structured literature review
and portfolio analysis on the topics of DB focus for up to five of the areas covered by DB.
It is anticipated that selection will be driven by the importance of areas in recorded
projects, country reforms, and literature. Each deep dive will discuss the area’s indicator
features and examine construct and content validity (respectively, the degree to which
the index or indicators measure the concept they intend to and the extent to which the
index or indicators are representative of the area they aim to measure). For each area,
the deep dives will provide detail on the relevant portfolio’s design features and draw on
portfolio review and analysis evidence on relevance and effectiveness, including factors
that are associated with success and failure. Each will draw on a structured literature
review for that area. Deep dives will also draw from the draft case studies and will
consider the role of stakeholders (other than the Bank Group) at the country or global
level.

• **Country case studies (with embedded indicator-level case studies).** In five purposively
selected countries, the Independent Evaluation Group will follow a nested structure with
a country-level case study of the relevance and effectiveness of DB indicators and related
reforms. Cases will involve both desk-based and remote field-based assessments; they
will aim to identify to what extent the Bank Group’s efforts were effective and how and
why specific interventions were or were not successful in delivering the intended results.
Countries will be selected to represent both those with substantial Bank Group support
and those that carried out relevant reform efforts independent of such support.

• **Country-level reviews of policy, strategy, and diagnostics.** The evaluation will conduct a
series of systematic document reviews for a sample of countries that will include
Systematic Country Diagnostics, Country Partnership Frameworks, and Country Private
Sector Diagnostics to better understand the level of alignment and coherence of Bank
Group country-level strategies with DB-related concerns. The evaluation may also draw
evidence from Country Program Evaluations and Completion and Learning Report
Reviews.

• **Semistructured interviews.** Semistructured interviews with subject matter experts within
the Independent Evaluation Group, the broader Bank Group, and external stakeholders
such as governments, donors, nongovernmental agencies, academia, and private sector
entities. It is anticipated that a panel of experts will be convened to assist in identifying
key issues and challenges.

• **Statistical and econometric review.** The validity and robustness of DB indicators and
indexes will be examined using (i) other indicators and measures of the business
enabling environment (including enterprise surveys) and (ii) indicators and measures of
outcome indicators, such as country-level data on business entry and productivity, and
private sector investment (domestic and foreign), employment, and growth.
6. Internal and External Audiences, Dissemination, and Communication

6.1 The primary audience of this evaluation is the Bank Group Board of Executive Directors and Bank Group management and staff. However, the findings of the evaluation will be relevant to a broader audience, including multilateral and bilateral agencies, donors, investors, and government officials; practitioners in client countries will share the findings and stimulate discussion among practitioners to encourage exchange of ideas across Global Practices, departments, and Bank Group institutions. A blog will be posted on the IEG website to reach a broader audience. The team will also explore internal and external forums for further dissemination.

6.2 The synthesis report will be led by Andrew Stone (Financial, Private Sector, Infrastructure, and Sustainable Development) under the guidance and overall direction of José Carbajo Martínez (director, Financial, Private Sector, and Sustainable Development). The team will also include Melvin Vaz (senior evaluation officer), Ana Belén Barbeito (senior evaluation officer), and Ayooluwa Olusola Adewole, Regina Legarreta, Anna Mortara, Nadia Ramirez, and Rasec Niembro Urista. Paul Holden, a senior consultant, will contribute to the analysis and drafting. Other expertise will be contracted as required. Emelda Cudilla will provide administrative support. Other IEG staff will be consulted as needed.

6.3 This evaluation will be subject to IEG’s standard internal quality assurance and an enhanced external quality review process. Three peer reviewers will provide guidance and quality assurance at the concept stage. At the concept stage they were Vijaya Ramachandran, a senior fellow at the Center for Global Development and expert on enterprise productivity and the business environment; Alvaro Gonzales, principal economist, Jobs Cross-Cutting Solutions Area; and Judith Kelley, dean and professor in the Sanford School of Public Policy and Terry Sanford Distinguished Professor of Public Policy. In addition to the peer reviewers, an advisory group representing internal and external stakeholders will offer guidance throughout the evaluation. This advisory group will consist of experts on private sector development and country programs (internal), and representatives of academia, government, and country management (external).

6.4 The evaluation (both products) has an estimated budget of approximately $675,000.
1 For more information about the ease of doing business score, see https://www.doingbusiness.org/en/data/doing-business-score.

2 The 11 economies for which data on the second-largest city are collected are Bangladesh, Brazil, China, India, Indonesia, Japan, Mexico, Nigeria, Pakistan, the Russian Federation, and the United States. Scores for the economy are calculated as the population-weight average for the two cities.

3 Examples include the expansion of the indicators on dealing with construction permits to include an index of the quality of building regulation and its implementation; the addition to the getting electricity indicators of an index of the reliability of electricity supply and transparency of tariffs; the addition to enforcing contracts indicators of an index of the quality and efficiency of judicial processes; and the addition of a gender dimension to the indicators for starting a business, registering property, and enforcing contracts.

4 The theory of change presented here is preliminary and may be refined after completion of the Issues Paper and synthesis.

5 Note that in most cases, this is compliance with regulations as they apply to a prototypical small or medium enterprise and may not apply to firms substantially larger or smaller—that is, microenterprises or large corporations. Nonetheless, the claim for Doing Business (DB) indicators is that they capture important constraints to economic and private sector performance: “Cumbersome red tape holds back more than individual businesses or investors: an economy’s ability to grow sustainably may suffer. Economic freedom to do business goes hand in hand with economic development and a thriving private sector, and these in turn underpin poverty elimination and the pursuit of shared prosperity” (World Bank 2019a).

6 “A vibrant private sector—with firms making investments, creating jobs, and improving productivity—promotes growth and expands opportunities for poor people…. Although macro policies are unquestionably important, there is a growing consensus that the quality of business regulation and the institutions that enforce it are a major determinant of prosperity” (World Bank 2004, viii).

7 Defined as having one or more DB indicators in the objectives or monitoring indicators, or justifying the project using DB indicators as evidence.

8 DB defines this as “a change of 0.5 points or more in the score and 2% or more on the relative score gap, except where the result of indexation)” (World Bank 2019a).

9 Independent Evaluation Group calculation based on database of reforms shared by DB team. This includes national reforms DB2010–20, excluding subnational reforms and employing workers reforms. N = 3,290. India, North Macedonia, Russian Federation, Togo, and Uzbekistan tied for 12th place.
Bibliography


Appendix A. Methodological Approach and Design Matrix

Evaluation Objective

The evaluation aims to delve into the accuracy and usefulness of the Doing Business (DB) report and indicators in identifying reform priorities, guiding the design of reforms, and monitoring the progress of reforms. It will probe available sources for evidence of the benefits of DB-related reforms for improving investment and employment, inequality, and other country development objectives. It will also gather and report available information on the social and environmental consequences of DB-related reforms, the comprehensiveness of DB as a guide to investment climate reform, and the value (and challenges) of aggregated DB indexes, including the ease of doing business score.

The evaluation objective inspired two lines of inquiry, which will guide the collection of data and analysis (box A.1). Evaluation questions were designed to break the topic into answerable components in the areas of DB influence, relevance, effectiveness, and factors affecting outcomes of DB support (see section on Evaluation Questions).

---

Box A.1. Two Lines of Inquiry That Will Guide the Evaluation

Doing the right things. This line of inquiry examines the relevance of indicators to country contexts and priorities, substantive dimensions of the areas they cover, and World Bank Group strategic and operational priorities.

Doing things right. This line of inquiry examines whether Doing Business is being used effectively by the Bank Group and client countries to achieve intended immediate and intermediate business environment reform outcomes subject to broader policy priorities.

Source: Independent Evaluation Group.

---

Evaluation Questions

The evaluation seeks to answer the questions in table A.1 under an overarching question about the lessons from the use of DB in guiding business environment reforms in client countries during fiscal years (FY)10–20. The evaluation questions were informed by a review of literature, previous Independent Evaluation Group (IEG) evaluations, and consultations with subject matter experts in business environment reform.
Table A.1. Evaluation Questions

<table>
<thead>
<tr>
<th>Evaluation Questions</th>
<th>Data Collection and Analysis Methods</th>
<th>Strengths and Limitations</th>
</tr>
</thead>
</table>
| Is DB doing the right things in terms of specific indicators and country contexts? | Literature review; synthesis of available evaluative materials and country case studies; review of country strategies, diagnostics, and evaluations; PRA of approved projects; prior IEG country-based evaluations; deep dives for individual DB areas, statistical and econometric analysis | • Volume and breadth of available literature  
• Access to new tools to enhance identification of correct portfolio  
• External, independent evaluations of DB use are limited and there are many gaps.  
• Sources on use of DB indicators in client country reforms are inconsistent and there are many gaps.  
• Limitations on fieldwork due to the coronavirus pandemic |

1. **Doing the right things.**  
What is the relevance of DB indicators to  
   a. The business environment priorities of client countries;  
   b. The substantive dimensions within each indicator’s focus area; and  
   c. World Bank Group strategic and operational priorities.  

2. **Doing things right.**  
   a. Is DB effectively achieving desired outcomes?  
   b. Is DB being used by clients and the Bank Group to achieve the best outcomes for business environment reform (for example, entry, investment, employment) subject to broader policy priorities (for example, public safety, equity, environment)?  

---

*Source: Independent Evaluation Group.  
Note: DB = Doing Business; IEG = Independent Evaluation Group; PPAR = Project Performance Assessment Report; PRA = portfolio review and analysis.*
Overarching Principles and Methods Design

Three central principles motivated the evaluation design: theory-based evaluation and mixed methods and multilevel analysis. First, the evaluation will adopt a theory-based approach that seeks to understand the links among DB, Bank Group– or client government–supported interventions, and country-level reforms based on the influence model elaborated. Second, the evaluation will also apply a mixed methods approach that combines an array of complementary methods for data collection and analysis (for example, internal project-level data, external country data sets, project performance data, semistructured interviews, case studies, sector deep dives, and structured literature reviews), then triangulates to ensure robust findings. Third, the evaluation will use three levels of analyses, namely, global, country, and project or intervention level, which will provide information on whether and how Bank Group programs and projects that are related to DB succeeded, including on factors associated with success (or failure; figure A.1).

The evaluation will produce two products:

- First, an issues paper and synthesis that takes stock of existing evidence on DB relevance, identifies key issues, and elaborates a framework and testable hypotheses for the deeper analysis of a full evaluation to follow.

- Second, a focused evaluation that applies the theory of change and hypotheses generated in the first product. This product will answer questions about relevance and effectiveness in terms of both immediate and intermediate outcomes.

The global level of analysis will be supported by the structured literature review, analysis of the Bank Group strategies and policies, deep dives into a subset of areas measured by DB, an expert panel discussion, and a quantitative analysis of the DB indicators or indexes. At the country level, the evaluation will collect data and evidence through the following methods: (i) Bank Group country strategy and diagnostics analysis; (ii) desk- and field-based country case studies; (iii) portfolio review and analysis for case study countries; (iv) comparative case analysis; and (v) structured interviews of key informants (figure A.1). At the portfolio or intervention level, there will be a portfolio review and analysis to understand the breadth of interventions undertaken by the Bank Group to support DB reforms.
Evaluation Components

Portfolio Review and Analysis

The evaluation will conduct a systematic desk review and assessment of projects to identify design features and characteristics, achievement of objectives, and drivers of success and failure. The evaluation identified the DB portfolio approved between FY10 and FY20 by performing targeted keyword searches in text-based data sets and searches for DB indicators in project result matrixes, and using matching learning techniques. The evaluation will also explore the use of semiautomated portfolio identification in delineating and classifying contributions related to business areas DB indicators. Building on the learning sample already coded by the team, the evaluation will use multilabel classification to identify projects that reference DB areas or relevant dimensions in their rationale, objectives, or indicators. A combination of manual taxonomy building and supervised learning methods will be used to achieve this objective. Some DB indicators will be harder to classify in taxonomy than others, and this may generate a high false-positive rate. For those that may generate a high false-positive classification rate, additional (manual) inputs and iterative intercoder reliability
testing will be realized to ensure that the coding method accurately captures the prevalence of the indicators in the underlying data.

The team will extract information through reviewing Project Appraisal Documents, Program Documents, Project Papers, Implementation Completion and Results Reports, Implementation Completion and Results Report Reviews, and Project Performance Assessment Reports for World Bank projects. For International Finance Corporation (IFC) investment services, this evaluation will rely on Board reports and Extended Project Supervision Reports and their Evaluative Notes. For IFC advisory services (AS), the team will review AS Concept Notes, AS Implementation Plans, and AS Project Completion Reports and their Evaluative Notes. Finally, for the Multilateral Investment Guarantee Agency project the team will extract information from the President’s Report, Project Evaluation Reports, and IEG Validation Notes. The forthcoming issues note will elaborate the findings of the preliminary portfolio review.

**Preliminary portfolio review.** To inform the Approach Paper and issues paper, the team undertook a preliminary portfolio review using a sampling methodology. To identify a preliminary *DB* portfolio, the IEG team followed these steps:

1. First, the IEG conducted systematic keyword searches of project objectives and titles and searched for *DB* indicators and subindicators in project result matrixes. IEG isolated those that contained at least one of the targeted keywords described in table A.2.

<table>
<thead>
<tr>
<th>Areas of Business Regulation</th>
<th>Keywords Used</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>General</em></td>
<td>doing business, business environment, business regulation, business operation, regulatory reform</td>
</tr>
<tr>
<td><em>Starting a business</em></td>
<td>procedure/time/cost with start/starting</td>
</tr>
<tr>
<td><em>Dealing with construction permits</em></td>
<td>procedure/time/cost with construction permit(s)</td>
</tr>
<tr>
<td><em>Getting electricity</em></td>
<td>procedure/time/cost with electricity</td>
</tr>
<tr>
<td><em>Registering property</em></td>
<td>property with procedures/time/cost and land with administration</td>
</tr>
<tr>
<td><em>Getting credit</em></td>
<td>bureau/registry/information with credit and movable collateral with laws</td>
</tr>
<tr>
<td><em>Protecting minority investors</em></td>
<td>minority with rights</td>
</tr>
<tr>
<td><em>Paying taxes</em></td>
<td>paying taxes, tax payment, time to comply</td>
</tr>
<tr>
<td><em>Trading across borders</em></td>
<td>trading across borders and time/cost/documents with import/export</td>
</tr>
<tr>
<td><em>Enforcing contracts</em></td>
<td>time/cost with dispute</td>
</tr>
<tr>
<td><em>Resolving insolvency</em></td>
<td>time/cost/recovery rate/legal framework with insolvency</td>
</tr>
</tbody>
</table>
Areas of Business | Regulation | Keywords Used
---|---|---
Employing workers | employment with regulation | 
Contracting with the government | time/procedure with procurement and procurement with regulatory framework | 

Source: Independent Evaluation Group portfolio review and analysis.

2. Next, for World Bank projects, if the project objective was not available in the business warehouse project list, IEG searched relevant themes for alignment with DB areas (see table A.3). World Bank advisory services and analytics (ASA) included the economic and sector work and technical assistance product lines from operations approved between FY10 and FY18 because these are the product lines that identify sectors and thematic codes. But because the five ASA product lines (economic and sector work, technical assistance, impact evaluation, programmatic approach and external training) were replaced with the code AA after FY18, IEG used this code for the operations approved in FY19 and FY20. For IFC projects, if the project objective was not available, IEG used relevant sectors names or business lines, summarized in table A.3.

Table A.3. World Bank Group System Codes to Identify the Doing Business Portfolio

<table>
<thead>
<tr>
<th>World Bank Lending and ASA</th>
<th>IFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theme codes:</td>
<td>Business line products and sector names:</td>
</tr>
<tr>
<td>• Business enabling environment (21)</td>
<td>• Economic and private sector development- climate business area</td>
</tr>
<tr>
<td>• Energy and policy reform (862)</td>
<td>• A2F - credit bureaus (D)</td>
</tr>
<tr>
<td>• Access to energy (863)</td>
<td>• FAM – collateral registries/secured transactions (I-D)</td>
</tr>
<tr>
<td>• Personal and property rights (423)</td>
<td>• EFI – credit infrastructures (ENT)</td>
</tr>
<tr>
<td>• Financial infrastructure and access (32),</td>
<td>• IC-investment policy (I-D)</td>
</tr>
<tr>
<td>• Tax policy (114)</td>
<td>• IC - discontinued product- access to land (EXT)</td>
</tr>
<tr>
<td>• Trade (14)</td>
<td>• IC - discontinued product- business taxation (I-D)</td>
</tr>
<tr>
<td>• Legal institutions for a market economy (422)</td>
<td>• IC - debt resolution and business exit (ENT)</td>
</tr>
<tr>
<td>• Judicial and other dispute resolution mechanism (421)</td>
<td>• Investment climate - business regulation (D)</td>
</tr>
<tr>
<td>• Labor markets institutions (662)</td>
<td>• SBA - sustainable and inclusive investing (ENT)</td>
</tr>
<tr>
<td>• Transparency, accountability and good governance, procurement (432)</td>
<td>• SBA - environmental, social and trade standards (ENT)</td>
</tr>
</tbody>
</table>

Source: Independent Evaluation Group Review

Abbreviations: A2F = Access to Finance IC = Investment Climate, EFI = Equitable Growth, Finance and Institutions, FAM = Finance and Markets, SBA = Sustainable Business Advisory, TAC = Trade and competitiveness, ASA = advisory services and analytics; IFC = International Finance Corporation.
3. The team also explored the use of semiautomated portfolio identification. Based on an initial learning sample compiled by the team, the evaluation used a multilabel classification to identify World Bank lending projects that reference DB indicators or relevant dimensions in their rationale, objectives, or results matrix. This process was especially useful in identifying qualifying projects that were excluded from initial human search (possible false negatives).

A typology of the intervention dimensions was developed to understand how World Bank projects use DB indicators. The DB report has expanded to cover 12 areas of business regulation with 194 subindicators (table A.2). Ten of these areas—starting a business, dealing with construction permits, getting electricity, registering property, getting credit, protecting minority investors, paying taxes, trading across borders, enforcing contracts, and resolving insolvency—are included in the ease of doing business score and ranking. DB also measures regulation on employing workers and contracting with the government, which are not included in the ease of doing business score and ranking.

4. The initial review identified 1,669 potential DB projects (372 World Bank lending projects, 916 World Bank ASA projects, 380 IFC AS projects, and 1 IFC investment project) that had been approved between (FY10 and FY20). Of the 1,669 projects, IEG reviewed 577 projects (35 percent sample), composed of 192 World Bank lending (52 percent of its identified portfolio), 225 World Bank ASA (25 percent), 1 IFC investment services (100 percent), and 159 IFC AS (42 percent). To expedite early understanding, a random sample stratified by institution, income level, and region for the preliminary portfolio review was drawn from this population of projects initially identified and was assigned to team members to review.

The rate of correctly identified DB-informed projects was 35 percent. From the 577 projects, 202 were classified as referencing DB areas or relevant dimensions in their rationale, objectives, or indicators. The identification success rate was 38 percent (72 projects) for lending projects, 20 percent (45) for World Bank ASA operations, and 53 percent (85) for IFC AS projects—the IFC investment services project was not confirmed as DB (table A.4). The high rate of false positives was due to the broad range of search terms and topics and the fact that many DB terms are also phrases in common use outside of DB report, such as “paying taxes.”
Table A.4. Identification Strategy Results Summary of Doing Business Portfolio

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>DB-Identified Projects</th>
<th>Coded Projects</th>
<th>Positives</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(no.)</td>
<td>(percent)</td>
<td>(no.)</td>
<td>(percent)</td>
</tr>
<tr>
<td>World Bank lending</td>
<td>372</td>
<td>192 (52)</td>
<td>72 (38)</td>
<td>120 (63)</td>
</tr>
<tr>
<td>World Bank ASA</td>
<td>916</td>
<td>225 (25)</td>
<td>45 (20)</td>
<td>180 (80)</td>
</tr>
<tr>
<td>IFC IS</td>
<td>1</td>
<td>1 (100)</td>
<td>0 (0)</td>
<td>1 (100)</td>
</tr>
<tr>
<td>IFC AS</td>
<td>380</td>
<td>159 (42)</td>
<td>85 (53)</td>
<td>74 (47)</td>
</tr>
<tr>
<td>Total</td>
<td>1,669</td>
<td>577 (35)</td>
<td>202 (35)</td>
<td>375 (51)</td>
</tr>
</tbody>
</table>

Source: Independent Evaluation Group portfolio review and analysis.
Note: The Positives columns show projects that were classified as using Doing Business indicators in their rationale, objectives, or results matrix after manual review; the False Positives columns show those projects classified as not having such interventions. The table shows all projects reviewed: a random sample of the projects approved in the evaluation period FY10–20. ASA = advisory services and analytics; DB = Doing Business; IFC AS = International Finance Corporation advisory services; IFC IS = International Finance Corporation investment services.

Structured Literature Review

IEG will commission and supervise a structured literature review, after IEG’s protocol covering academic and Bank Group literature on the DB indicators and related reforms in general and for the indicators selected for deep dives. The objective is to shed light from the literature on the degree to which the DB indicators accurately capture reform priorities, provide comprehensive framework for the guidance of reform design, and track reform progress. Particular attention will be paid to the relevance of DB indicators in measuring country development outcomes and any evidence of development impact or social and environmental consequences for the reforms themselves. The structured literature review is expected to provide empirical evidence on the links between the DB indicators and reform outcomes and impact. The review will follow IEG’s structured literature review framework for major evaluations, with a clearly defined search strategy, inclusion and exclusion criteria, and extraction and synthesis. Before this more rigorous review, the team will conduct a desk review of relevant literature identified by the DB team from top journals to understand the pattern of evidence.

Country Case Studies

The evaluation will conduct five case studies in purposively selected countries. Case selection will reflect a diversity of country conditions and contexts (including regional diversity, income level, and International Development Association and fragility, conflict, and violence classification). The country selection criteria will include the presence of DB reforms, whether supported by the Bank Group or government, and the presence of evaluation evidence (Implementation Completion and Results Reports, Expanded Project Supervision Reports, Project Completion Reports, Project Evaluation Reports) in countries with Bank Group support. The evaluation will adopt a template for data
collection and follow standard protocols to facilitate comparison across case studies. These methods include (i) a review of literature on the DB indicators and related reforms in the country; (ii) a review of Bank Group country strategies, diagnostics, and analytical works; (iii) a review of the Bank Group’s portfolio to support DB in the country; (iv) structured interviews with stakeholders (government, multilateral development banks, the private sector, nongovernmental organizations, academics, and so on). The evaluation team will also consult with internal and external experts to identify the countries that offer the richest opportunities for learning. Cases will involve both desk-based and (remote) field-based assessments, and they will aim to identify to what extent the Bank Group’s efforts were effective and how and why specific reforms or reform interventions were or were not successful in delivering the intended results.

Country studies will be undertaken to understand reform priorities from independent (especially press) sources, supported by machine-assisted data scraping and other tools to better understand the extent to which government reform processes are being informed by DB indicators. The analysis will draw on publicly available information to gauge whether the DB indicators have had any influence on or in some way informed a set of reforms or policies. This process can be broken down into a search of news aggregates (for example, Factiva) to explore government reform and an exploration of a custom-coded data set of DB-related reforms to analyze the extent to which government reforms on selective dimensions of the framework can be tangibly traced to inputs or influences from the DB project.

This approach presents risk of gaps between substantive reforms and reporting in the media, and the coverage of some areas will be sparser than of others. The team initially considered a systematic search of individual government ministry websites, but this approach was deemed challenging because of the inherent data issues intrinsic to the proposed scraping process. Although a search of ministry websites may work as a pilot, a number of such websites are not updated regularly, diminishing the usefulness of the exercise. Furthermore, individual scraping of each ministry website may make such an endeavor prohibitively time-consuming. These shortcomings may be addressed through the combined use of news aggregators and a bespoke data set of DB projects generated by the data science team.

**Indicator-Specific Deep Dives**

Indicator-specific deep dives will provide an opportunity for the evaluation to study DB areas in a structures and focused manner. The deep dives methodology, guided by a common template, will include (i) a focused structured literature review and profound analysis on five of the areas covered by DB; and (ii) a review of indicator-relevant portfolio and full consideration of evidence (from multiple sources) describing their
relevance and effectiveness. It is anticipated that the selection of deep dives will be based on prevalence of areas in recorded projects, country reforms, and literature. Each deep dive into an area of DB will discuss indicator and subindicator features and will examine construct and content validity (the degree to which the index or indicators measure the concept they intend to, and the extent to which the index or indicators are representative of the area they aim to measure). For each area, the deep dives will provide detail on the relevant portfolio’s design features and draw on portfolio review and analysis evidence on relevance and effectiveness, including factors that are associated with success and failure. Each will be drawn on a structured literature review for that area. Deep dives will also draw from the draft case studies and will consider the role of stakeholders (other than the Bank Group) at the country or global level.

**Semistructured Interviews**

Semistructured interviews will be conducted with subject matter experts and practitioners within IEG, the broader Bank Group, and external bodies including governments, donors, nongovernmental agencies, academia, and private sector entities. It is anticipated that, at an early stage, the evaluation will convene an expert panel to assist in identifying key issues and challenges, better understanding the underlying theory, and developing a set of preliminary hypotheses. The interviews with Bank Group staff engaged in DB projects will help in understanding institutional priorities, program features, achievement of objectives, and lessons derived from the experiences. They will also attempt to gauge stakeholders’ perspectives on the accuracy and usefulness of DB for identifying reform priorities, guiding the design of reforms, and monitoring reforms outcomes. The evaluation team will develop an interview guide to ensure key questions are asked consistently across interviews while maintaining the flexibility needed to follow response trajectories that may not adhere to the guide.

**Country-Level Reviews of Policy, Strategy, and Diagnostics**

The evaluation will conduct a series of systematic document reviews for a sample of countries to complement the evaluation’s portfolio review. The evaluation envisions carrying out a systematic review of Bank Group country strategy documents (for example, Country Partnership Frameworks) to better understand the level of alignment and coherence of Bank Group country-level strategies and DB-related concerns. A similar review will be carried out for those countries that have been subject to Systematic Country Diagnostics and Country Private Sector Diagnostics. A categorical array will be developed to systematically assess evaluation questions across strategy documents and diagnostics. The evaluation may also draw evidence from Country Program Evaluations and Completion and Learning Report Reviews.
Statistical and Econometric Analysis

The evaluation will apply statistical and econometric methods to examine the robustness of DB indicators and indexes. Statistical and econometric analysis of analysis of DB and external indicators would compare the DB indicators with other measures of the business enabling environment that are not based primarily on DB (including those based on enterprise surveys), and examine the link of movements in DB indicators (reforms) to outcome indicators, such as country-level data on business entry and productivity, private investment (domestic and foreign) and employment, and private sector growth.

Design Matrix

Table A.5 aligns evaluation questions with the evaluation design. Each check mark represents the anticipated strength of the method to answer the questions: one check mark suggests that the method will provide some data to answer the evaluation question, and three check marks suggest the method is expected to provide a great deal of data to answer the evaluation question.

Table A.5. Evaluation Questions and Methods Applied

<table>
<thead>
<tr>
<th>Evaluation Questions</th>
<th>Case Studies</th>
<th>Portfolio Review and Analysis</th>
<th>Deep Dives</th>
<th>Literature Review</th>
<th>Country-Level Reviews</th>
<th>Interviews</th>
<th>Statistical and Econometric Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Doing the right things. Is DB doing the right things in terms of specific indicators and country contexts?</td>
<td>✓✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓✓</td>
<td>✓✓✓</td>
<td>✓✓✓</td>
</tr>
<tr>
<td>a. The business environment priorities of client countries;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. The substantive dimensions within each indicator’s focus area; and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. World Bank Group strategic and operational priorities.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Doing things right. Is DB being used by clients and the Bank Group in ways to achieve the best effect for business environment reform?</td>
<td>✓✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
</tr>
<tr>
<td>a. Is DB effectively achieving desired outcomes?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Is DB being used by clients and the Bank Group to achieve the best outcomes for business</td>
<td>✓✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓</td>
<td>✓✓✓</td>
<td>✓✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
## Design Limitations

IEG anticipates multiple limitations to its proposed mix of methods and will seek to mitigate the risks they impose (including by triangulating from multiple evidence sources to test hypotheses). These limitations include gaps and inconsistencies in the literature and country social and economic data sets; the limited number of relevant evaluated projects and country programs with DB relevance (especially in light of the lack of IEG-validated evaluations of ASA); the limited ability of a small number of case studies and deep dives to represent the universe; limited data on intermediate outcomes and sustainability of reforms, difficulties in establishing attribution to specific reforms; and difficulties imposed by the coronavirus pandemic in conducting fieldwork (even remotely). There may also be some false negatives or false positives in identifying relevant literature, project activities, and reforms due to the multifaceted nature of areas covered by some DB indicators, lack of standard terminology, and the sometimes indirect nature of influence. Establishing causal connections between DB indicators and observed reforms, and between reforms and observed outcomes, is complicated by the likelihood of multiple causal factors.

### Evaluation Questions

<table>
<thead>
<tr>
<th>Case Studies</th>
<th>Portfolio Review and Analysis</th>
<th>Deep Dives</th>
<th>Literature Review</th>
<th>Country-Level Reviews</th>
<th>Interviews</th>
<th>Statistical and Econometric Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>environment reform (for example, entry, investment, employment) subject to broader policy priorities (for example, public safety, equity, environment)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Source: Independent Evaluation Group.*
Appendix B. Past Findings and Recommendations on the Relationship of the Doing Business Indicators to Real Reform Progress

This short note is based on three Independent Evaluation Group (IEG) evaluations with findings relevant to the Doing Business (DB) indicators plus one panel review and one external audit commissioned by the World Bank Group. All studies to date recommend continuation of the DB indicators and find substantial benefits to the indicators in motivating reform. Each, however, find some flaws in the connection of the indicators to a reform agenda, capturing reform progress and effects, or both. This note identifies the issues raised regarding the relationship of DB indicators to real reform progress, but does not cover critiques of DB practices, such as the identity, diversity and number of its informants and data or source transparency.

DB indicators do not capture the full range of legal and regulatory priorities. IEG’s original 2008 evaluation suggested that the DB team should use World Bank analyses such as enterprise surveys and investment climate assessments to drive the choice of new DB indicators and to periodically reassess its current set of indicators. IEG’s 2015 investment climate evaluation finds that DB indicators “are incomplete” because they “do not cover all areas of regulation as identified in the best practice list” (World Bank 2015, p. 153). It recommends that the Bank Group expand “the coverage of current diagnostic tools and integrate them to produce comparable indicators so that these can capture the areas of the business environment not yet covered by existing tools” (Ibid.). Specific missing areas include accounting and auditing, contract laws, competition policy, consumer protection, environmental laws, intellectual property (including privacy, copyright, patent, trademark, unfair business practices), investment policy and promotion, employment law, labor safety and health, and alternative dispute resolution. The DB indicators have not changed their coverage since this recommendation. The 2013 panel review recommended that DB change its name to reflect its limited focus: “The report’s title, Doing Business, implies that it provides a comprehensive measure of the business environment, rather than just a measure of business regulations” (World Bank 2013, p. 4). Further, a 2018 external audit, “On the Integrity of the ‘Ease of Doing Business’ Indicators” (again, not IEG) finds “the World Bank may wish to expand the ease of doing business database with new indicators to cover additional aspects of the business environment, rather than with methodological changes to existing indicators” (Morck and Shou 2018, p. 2).

The DB indicators focus primarily on compliance costs with some exceptions. All studies have observed that DB indicators look primarily at compliance costs (in time, money, or procedural steps) to businesses as the core metric, with several exceptions.
include the extend of protection of property, the protection of minority shareholder rights, and the range of assets that can be used as collateral.) They have noted that under such metrics, less is always better, even if regulations are inadequate to pursue public policy objectives such as protection of public health and safety. The 2015 investment climate evaluation explicitly found that the Bank Group needed “separate measures … to capture a wider range of benefits and costs (social, economic, and environmental) if existing regulations are changed” (World Bank 2015, p. xl). No evidence has been provided of the generation of such systematic metrics. This recommendation was echoed regarding the trading across borders indicator in the 2019 IEG evaluation on trade facilitation. It found “Many trade regulations are intended to serve socially beneficial purposes such as enhancing public health, safety, and the environment, or reducing informality and corruption. However, insufficient attention has been paid to such objectives and only compliance costs are routinely monitored” (World Bank 2019, p. ix).

DB indicators do not track reform impacts. IEG’s original 2008 evaluation recommended that the Bank Group trace the impact of DB reforms at the country level. IEG is aware of individual studies but no comprehensive effort to do so. The 2013 independent panel review (not by IEG) critiques DB’s implied claim that the reforms it promotes cause economic development and growth:

Empirical evidence on the results of business-regulation reforms captured by the report is mixed and suggestive at best. Correlations between the report’s topics and developmental outcomes often do point to a negative association between the regulatory burden and economic development and growth. However, such correlations do not justify a causal interpretation…. The evidence in favour of specific country reforms is contingent on many auxiliary factors not captured by Doing Business report topics. (World Bank 2013, p. 2)

Findings on specific indicators recommend changes:

- **Paying taxes.** IEG’s original DB2008 evaluation recommended that DB “revise the paying taxes indicator to include only measures of administrative burden” by excluding the tax rate, which was an issue of public policy rather than compliance burden (World Bank 2008, p. xxi). This finding was echoed in the 2013 independent panel report (World Bank 2013). DB continues to include tax rates above a 15th percentile benchmark in determining the paying taxes indicator.

- **Trade facilitation.** The 2019 trade facilitation evaluation found that the Bank Group produces two leading sets of indicators on trade—the DB trading across borders indicators and the Logistics Performance Index indicators (World Bank
It found that although each had strengths, they frame their subjects differently, resulting in gaps and inconsistencies. The evaluation recommends that the Bank Group should rationalize its two-trade facilitation to assure they are effective benchmarks of performance that are useful to assess and monitor reforms, while maintaining continuity of subindicators that have proven accurate in tracking reform. The Bank Group partially agreed, but has not committed to any action based on this.

The DB practice of aggregating indicators is controversial. The 2013 independent panel review of the DB Report (produced by a panel selected by the president of the Bank Group) finds that aggregate rankings are challenging because they involve aggregating across indicators (topics)—a process that explicitly or implicitly involves a value judgment of what is “better” for DB and how much better it is (World Bank 2013). It also finds that “small revisions or inaccuracies in primary data can significantly change a country’s rankings.” (World Bank 2013, p. 3) Therefore, it recommended that the World Bank continue to publish the report but “without the overall aggregate rankings (the ease of doing business ranking). Rather, the scores (cardinal values) for each of the indicators should be emphasized.” (World Bank 2013, p. 4) In fact, DB continues to produce a country ease of doing business score.

DB adjustments to indicators generate discontinuities. The 2018 external audit “On the Integrity of the ‘Ease of Doing Business’ Indicators” found that “frequent methodology changes reduce the value of the indicators to researchers, policy makers and the media” (Morck and Shou 2018, p. 1). It recommended that “the World Bank may wish to minimize methodology changes except to fix confirmed problems with existing methodology” (Morck and Shou 2018, p. 2). However, the same study concluded that “concerns that World Bank staff implement methodology changes to manipulate the Ease of Doing Business indicators of specific economies or to sway domestic politics in affected economies are entirely without evidence” (Morck and Shou 2018, p. 1).

---

1 This appendix was prepared by Andrew Stone, lead evaluation officer, Independent Evaluation Group Financial, Private Sector, Infrastructure, and Sustainable Development.
References


