WP/21/285

IMF Working Paper

When Does Capacity Development Achieve Good Outcomes? Evidence from the IMF Results-Based Management Data

by Antonio Bassanetti

IMF Working Papers **describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

INTERNATIONAL MONETARY FUND

© 2021 International Monetary Fund

WP/21/285

IMF Working Paper

Institute for Capacity Development

When Does Capacity Development Achieve Good Outcomes? Evidence from the IMF Results-Based Management Data*

Prepared by Antonio Bassanetti

Authorized for distribution by Matt Davies

December 2021

IMF Working Papers **describe research in progress by the author(s) and are published to elicit comments and to encourage debate.** The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

Capacity development is one of the IMF's core activities. Its impact is monitored through a Results-Based Management framework. Using for the first time the resulting dataset, the paper investigates how the likelihood of achieving targeted outcomes correlates with macroeconomic conditions and project-specific characteristics. Results indicate a positive correlation with per capita GDP growth and the involvement of resident advisors and regional centers. Results also confirm lower chances of achieving targeted outcomes for fragile, conflict-affected, and small states as well as in complex projects. These findings inform Fund CD strategy, prioritization and delivery to help member countries achieve better outcomes.

JEL Classification Numbers: O19, O22

Keywords: Capacity Development, Evaluation, Results-Based Management

Author's E-Mail Address: abassanetti@imf.org

* The author is grateful to Michaela Erbenova, Matt Davies, Anna Bordon, and Miguel Alves for their valuable comments and guidance on the research project, as well as to Carolina Dyer-Lock and Lina Karaoglanova for their guidance on the RBM framework and data. Helpful discussions were also held with Andrea Presbitero and IMF colleagues who attended an internal seminar.

Contents **Page**

I. Introduction 3

II. Results-Based Management at the IMF 5
CD Project 5
Outcome Ratings 6

III. Dataset 6

IV. Econometric Setup 9

V. Results 10
A. Binary Outcome Ratings 10
Recipient Country's Macroeconomic Conditions 10
CD Project-Specific Features 11
Other Considerations 12
Digging Deeper 13
B. Original Outcome Ratings 14
Recipient Country's Macroeconomic Conditions 14
CD Project-Specific Features 15

VI. Conclusions 15

References 18

Figures

Figure 1. RBM Framework—Results Chain (log frame) 19
Figure 2. CD Project Design—an Example 19
Figure 3. Composition of the RBM Dataset 20
Figure 4. RBM Outcome Ratings—Frequency Distributions 20

Tables

Table 1. RBM Outcome Ratings—Summary Statistics 21
Table 2. Other Variables in the Dataset—Summary Statistics 22
Table 3. Country- and CD Project-Specific Variables—Linear Probability Model 23
Table 4. Country- and CD Project-Specific Variables—Probit Model 24
Table 5. Country- and CD Project-Specific Variables—Ordered Probit Model 25
Table 6. Marginal Effects–Ordered Probit Model 26

Annex 27

I. INTRODUCTION

Capacity development (CD) is one of the IMF's three core activities, alongside surveillance and lending. It supports member countries' efforts to build strong economic institutions and the capacity necessary to formulate and implement sound economic and financial policies. It encompasses both technical assistance (TA) and training. IMF CD has expanded strongly since the Global Financial Crisis, doubling in size and accounting for almost a third of total Fund spending. As recalled in the IMF Policies and Practices on Capacity Development, Fund CD is delivered by either Fund staff or by vetted external experts. Around a quarter of Fund CD is provided through a global network of 11 Regional Technical Assistance Centers (RTACs) and 6 Regional Training Centers (RTCs) that foster proximity to members.[1]

IMF CD is delivered in response to demand received from member country authorities and CD projects are designed with the authorities' active involvement to reflect the country's specific needs and absorptive capacity, as well as promote ownership. As member countries' demand for CD far exceeds the IMF's delivery capacity, a prioritization framework seeks to ensure that CD is targeted to strategically important topics and country groupings. These may not necessarily be consistent with the projects most likely to achieve the best outcomes in the short term, for instance fragile and conflict-affected states (FCS) are prioritized.

Against this backdrop, the five-yearly CD Strategy Review completed in 2018 emphasized country-tailoring and results-orientation. Systematic monitoring of CD results and periodic evaluation of performance are paramount in this as they help focus on goals and learn lessons from experience, thereby fostering accountability, impactful decision-making, efficient project management, and effective project design.

Over the last decade, the IMF has therefore been operationalizing a standardized Results-Based Management (RBM) approach to monitor the capacity enhancements that Fund CD recipients achieve. The regular collection of RBM data began in 2013 and in 2017 it was standardized and extended to all projects. It now covers more than 150 countries, with the number of observations large enough to allow for empirical analysis.

The focus of this paper is on the CD outcomes that are set as targets during project design and in particular on the ratings that project managers assign as projects unfold and progress toward such targets is monitored. The research question we are interested in is how CD outcome ratings are associated with the country-specific macroeconomic conditions at the time of CD delivery (e.g., per capita GDP level and growth, state of fragility) and the project-specific features (e.g., delivery modality, resources, complexity of project design). While the IMF RBM framework applies to both TA and training, we concentrate only on TA projects to ensure internal consistency of the dataset.

Neither the research question nor the empirical strategy that we follow are new. In particular, our approach is inspired by a strand of works in the development assistance literature that simultaneously examined the country- and project-level features associated with project outcomes, mainly through linear probability, probit, and ordered-probit models (see Denizer and others, 2013; and more recently, Presbitero, 2016; and Caselli and Presbitero, 2021).

[1] For an overview of the IMF CD activity, see Capacity Development (imf.org).

This notwithstanding, the application of this approach to the IMF's RBM dataset on CD represents an original contribution.[2] The aim is to carry out exploratory work on what we can potentially learn from the dataset. As such, findings should be interpreted as preliminary and as potential inputs for further in-depth CD evaluations, which could in turn inform the Fund CD strategy and delivery design with the ultimate goal of fostering effectiveness.[3]

Against this backdrop, RBM data suggest that IMF CD is contributing effectively to building economic management capacity in member countries. If we consider outcome ratings for TA completed projects delivered from 2013 onwards, less than 10 percent of targeted outcomes has not been achieved, with the remainder being either partially, or largely, or fully achieved.[4] Relatedly, the main findings of the econometric analysis indicate that:

(i) the likelihood of achieving outcomes is positively correlated with per capita GDP growth, presumably suggesting that the factors that promote growth—including better quality of policies and institutions—are also likely to translate into greater capacity to benefit from CD;

(ii) outcomes are harder to achieve in FCS and in small states;

(iii) IMF CD, lending, and surveillance integration, as well as recipient countries' ownership of CD projects, are associated with higher outcome ratings;

(iv) sustained interactions with CD recipients, particularly through in-country resident advisors and the IMF global network of regional centers, are also correlated with better outcomes;

(v) pursuing multiple reform plans at the same time is associated with lower chances of achieving the targeted results; and

(vi) CD is a complex medium-term process with targeted outcomes taking time to achieve.

Although these findings are broadly in line with intuition, they should be treated with some caution. The explanatory variables identified only account for a limited share of the variation in outcome ratings. This feature, which is common in the related literature, points to the potentially relevant project-specific characteristics that remain unobserved—and thus unaccounted for—despite the notable improvements of the RBM dataset since its launch, in terms of both coverage and consistency. The findings of the analysis should thus be interpreted as partial correlations rather than causal relationships.

The ongoing introduction of new processes and systems for the management and administration of the IMF CD activities will facilitate the gathering of more information on some important project-level features. A richer RBM dataset—also in terms of quantity of observations as more projects are delivered and rated—will allow future analysis to obtain more precise estimates, mitigate the omitted variable bias, explain a larger share of the variation in outcome ratings, better understand the mechanisms underlying some of the results—in particular those related to FCS and small states—and explore other research avenues.

---

[2] Other works have also been concerned with assessing the effectiveness of the IMF CD, although without systematically resorting to the RBM dataset, including IMF, Building Fiscal Policy in Fragile State (2017) on building fiscal capacity in fragile and conflict-affected states and Chami and others (2021) on the impact of technical assistance on revenue mobilization. On training, see Edison and others (2018).

[3] IMF CD evaluations apply the Common Evaluation Framework introduced in 2017 and updated in 2020, adopting the OECD Development Assistance Committee (DAC) evaluation criteria.

[4] As of February 2020, when this analysis was initiated. More recent updates of the dataset confirm that IMF CD is performing well.

The rest of the paper is organized as follows: Section II briefly introduces the main features of the IMF's RBM framework. Section III describes the dataset used for the analysis, composed of both RBM data and a set of macroeconomic variables. Section IV outlines the econometric setup chosen to conduct the investigation. Section V presents the main findings, with a few caveats for their correct interpretation. Section VI concludes by summarizing the main takeaways.

II. RESULTS-BASED MANAGEMENT AT THE IMF

The IMF has piloted results-based monitoring since the mid-2000s, with work on a standardized RBM framework being intensified following the 2013 Fund's CD Strategy. Importantly, since May 2017 the adoption of a harmonized RBM framework has been required for all the Fund's CD operations, thereby further enhancing the systematic gathering and monitoring of information at the project level. The coverage and consistency of the RBM dataset have thus progressively improved since its launch as practices have been standardized; and while there is still some way to go before the system is fully mature, the dataset can now be used for empirical analysis.

Based on the RBM framework, a results chain—or logical framework (log frame)—is specified at the outset of each new CD project, illustrating how the project's inputs (financial resources, staff time) are translated into activities (missions, backstopping, etc.) to produce outputs (advice, reports, workshops, etc.) and achieve the targeted outcomes for the recipient country (i.e., capacity improvements). In a nutshell, log frames illustrate the causal relationships between the resources dedicated to projects and the expected benefits in terms of capacity- and institution-building for the recipient countries (Figure 1).

CD Project

For the purposes of this analysis, we refer to a CD project as a set of CD activities delivered to a single recipient (i.e., beneficiary) country over a specified time frame in order to achieve one or more *objectives* and related *outcomes* (Figure 2; see below for some examples).[5] Each project can be focused on one or more *workstreams* according to the CD needs of the recipient country (e.g., revenue administration, central bank operations, etc.). Such workstreams are selected from a pre-defined catalog that spans the main areas of expertise of the IMF and to which all CD-delivery departments (CDDs) contribute based on their respective competence.

Within each workstream, projects are designed to achieve one or more *objectives*, which are also drawn from the pre-defined catalog and specified at the outset in agreement with the CD recipient country (Figure 2). Objectives are high-level, medium-term goals and are not directly observable (for example, possible objectives for the workstreams on revenue administration and central bank operations are, respectively, strengthening the core tax administration functions and enhancing the central bank's decision-making capacity). Hence, in order to assess the country's progress towards the objectives, *outcomes* are used, with one or more targeted outcomes associated with each objective (Figure 2; for example, possible outcomes for the objectives of strengthening the core tax administration functions and

[5] Each country may benefit from more than one project. A project is always uniquely identified with a financing source (e.g., core IMF funds or a specific externally financed vehicle such as an RTAC or thematic fund).

enhancing the central bank's decision making capacity are, respectively, that a larger portion of taxpayers meet their filing obligations as required by law and that the central bank has operational independence).[6] Targeted outcomes are strictly linked to the corresponding objectives, they are also drawn from the catalog at the outset of the project, and—importantly—they are observable through a set of pre-defined indicators.[7] This means that the project manager is able to assess and rate the achievement of targeted outcomes throughout the lifetime of the project, with ratings being subject to potential revisions until the project is completed and as new or updated information on the recipient country's capacity improvements is gathered.[8] Although outcome ratings may also partially reflect a subjective component by project managers and cannot be interpreted as the final assessment of how impactful a project has been, they incorporate important information to be analyzed and will be the focus of the empirical analysis presented in the rest of the paper.[9]

The IMF RBM catalog currently includes more than 30 workstreams covering about a hundred possible objectives, almost 600 outcomes, and more than a thousand indicators.[10]

Outcome Ratings

Outcome ratings are integers in the 1–4 range on the following scale: 1-*not achieved*, 2-*partially achieved*, 3-*largely achieved*, and 4-*fully achieved*. In line with the literature, in the empirical analysis we have also alternatively relied on a binary transformation of the original 1–4 rating scale by assigning a value of zero to outcomes originally rated either 1 or 2, and a value of 1 to outcomes originally rated either 3 or 4; this will allow to complement the analysis with the estimation of models with binary dependent variable, which offer a simpler interpretation of results.[11]

III. DATASET

As of February 2020, when this analysis was initiated, the RBM dataset included 2,957 TA rated outcomes—from both completed and ongoing projects—associated with 1,641 objectives, involving 152 recipient countries. The projects composing the dataset started between 2013 and 2019. Although the dataset used in the analysis is not the most up-to-date, it has the advantage of being internally consistent, as it excludes the structural break caused

[6] Like the objectives, the targeted outcomes are also decided in agreement with the recipient country. Log frames also include milestones to measure interim steps toward the achievement of outcomes.

[7] The value or status of the pre-defined indicators is regularly monitored and assessed against the baseline defined at the beginning of the project. Each targeted outcome is associated with at least one indicator.

[8] The aggregate (i.e., average) rating of the outcomes informs the project manager's assessment of the associated objective. The project manager can decide whether to assign equal or unequal weights to the outcomes, depending on whether they equally contribute or not to the achievement of the objective. The information on outcome weights had not yet been recorded in the dataset when this analysis was initiated and therefore could not be used to discriminate between outcomes according to their relevance; however, the ongoing introduction of new processes and systems for the management and administration of the IMF CD activities will allow to overcome this gap, making the information available for future work.

[9] Outcome ratings are assigned under the supervision of the departmental CD portfolio managers, limiting the risk of a substantial bias in the subjective component.

[10] The catalog is updated as needed.

[11] Admittedly, the binary transformation of the rating scale implies equating the outcomes that were partially achieved to those that were not achieved, thereby introducing a potential measurement error in the assessment of CD. However, as described in Section V, the results of the analysis are similar regardless of the chosen rating scale, be it the binary or the original one.

by the COVID-19 pandemic during which the only possible CD delivery modality has been the virtual one, with many of the recipient countries preoccupied with managing the crisis, potentially hampering their CD absorptive capacity.

The rated outcomes in the dataset are particularly concentrated in the workstreams of *revenue administration* and *public financial management* which constitute the largest share of IMF CD. Several other workstreams from the IMF RBM catalog are also well represented and span a wide range of fiscal, monetary, financial, statistical, legal, and macroeconomic topics. They include *tax policy, debt management, financial supervision and regulation, central bank operations, national account statistics, government finance statistics, anti-money laundering and counter financing of terrorism, financial and fiscal law reform, macroeconomic analysis*, among others. Around 63 percent of the rated outcomes comes from projects delivered by the Fiscal Affairs Department of the IMF, 21 percent by the Monetary and Capital Markets Department, and 13 percent by the Statistical Department, while the shares pertaining to the Legal Department and to the Institute for Capacity Development are relatively smaller (Figure 3).

In terms of geographical distribution, almost half of the dataset relates to CD delivered to countries in Sub-Saharan Africa, followed by the Western Hemisphere and Asia-Pacific regions (20 and 15 percent, respectively). Middle Eastern and Central Asian as well as European recipient countries are relatively less represented (11 and 8 percent). As far as income groups are concerned, low-income countries are the main recipients of IMF CD, followed by the emerging markets; the share of advanced economies is almost negligible.[12] Importantly, 35 percent of total observations in the dataset relates to FCS, which are highly ranked within the IMF CD prioritization framework; 18 percent relates to small states (which partly overlap with FCS). Finally, the majority of outcome ratings is from projects delivered through the IMF's global network of regional centers (61 percent).

According to the outcome ratings, the IMF has indeed contributed to capacity gains in CD recipient countries: since 2013 only less than 10 percent of the targeted outcomes of completed TA projects has not been achieved, with the remainder being either partially, or largely, or fully achieved (Figure 4). While the corresponding distribution for ongoing projects is slightly less favorable, it should be borne in mind that since targeted outcomes take time to achieve, ratings are subject to potential revisions—generally upward, based on experience—until projects are completed and a final assessment can be done. The sample average outcome rating—including both completed and ongoing projects—is 2.35 when using the standard 1–4 scale (2.64 and 2.31, respectively, for completed and ongoing projects; Table 1). Among income groups, average ratings are lower for low-income countries than for emerging markets and advanced economies. FCS and, to a less extent, small states also show lower average ratings. The average rating for recipients in the Western Hemisphere region is the highest among macro-regions.

Besides outcome ratings, other project specific variables are used in the analysis. They include the project completion share—proxied by the time elapsed between the project start date and the latest date in which some of the main project features have been subject to change, expressed as a share of project duration—to account for the possibility of rating

[12] Income groups and macro-regions are based, respectively, on the IMF WEO classification and on the composition of the IMF area departments.

revisions for ongoing projects;[13] the amount of CD services provided to the recipient country by short-term experts and long-term advisors—i.e., in-country resident advisors—as indicators of the resource intensity of the project;[14] a dummy variable identifying projects delivered through the network of regional centers (RTACs); and a proxy for the complexity (or ambition) of the CD project. The latter has been calculated based on three possible alternatives: (1) the number of workstreams in the project associated with the CD-delivery department responsible for the relevant targeted outcome; (2) the number of objectives in the project associated with the workstream within which the relevant targeted outcome falls; (3) the number of outcomes in the project associated with the objective within which the targeted outcome falls. Based on these definitions, different targeted outcomes within the same project may be associated with possibly different degrees of complexity, depending on how complex the design leading to each single targeted outcome is (see Figure 2).

The RBM dataset has been integrated with a set of macro variables specific to each CD recipient country. In line with the literature, they include per capita GDP—both levels and growth rates—and growth volatility (see Table 2 for data descriptions and sources). Besides, two dummy variables have been used to identify the outcome ratings associated with, respectively, FCS and small states.[15] Another dummy variable was used to identify the recipient countries that, while being delivered a CD project, were also engaged in a program with the IMF (either with or without borrowing arrangement).[16] Finally, we considered including in the dataset the World Bank's Country Policy and Institutional Assessment ratings (CPIA-IDA, ranging from 1=low quality to 6=high quality), as they may provide valuable information on the capacity of the recipient country of achieving the CD targeted outcomes. However, the CPIA-IDA ratings are publicly available for far fewer countries than those included in the RBM dataset; their use would thus result in a drastic reduction in the observations available for the estimates (by about 40 percent), suggesting to exclude them from the analysis.[17] Alternatively, we resorted to the Worldwide Governance Indicators (WGIs) which report on six broad dimensions of governance for a large number of countries (the indicators range from approximately -2.5=weak to 2.5=strong).[18]

---

[13] An alternative way of calculating the proxy for completion share could have relied on the outcome rating date instead of the latest date of project modification; however, this information was not available in the data dump used for this analysis. Also, should we use the date of the data dump instead of the latest date of project modification to calculate the completion share, the main findings of the analysis would be broadly confirmed.

[14] Short-term experts include both Fund staff and external experts selected from the IMF's roster. The amount of CD services provided by both short-term and long-term advisors is measured in *full-time equivalents*—FTEs—and is taken from the IMF's internal Travel Information Management System. The information currently available in the RBM dataset on the financial resources poured into CD projects is not granular enough to be used for this analysis.

[15] To this purpose, the IMF internal classifications for both FCS and small states were used. FCS are identified as countries facing particularly challenging economic problems, reflecting their limited administrative capacity and vulnerability to political, security, and other shocks. Small states are countries with a population of under 1.5 million and exclude advanced economies—defined by the IMF World Economic Outlook—and fuel-exporting countries.

[16] For the purposes of this analysis, a CD recipient country is identified as a program country as long as it has been engaged in a program with the IMF for at least 10 percent of the duration of the CD project.

[17] We also considered the possibility of including in the dataset the country-level measure of human capital from the Penn World Table (see Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" American Economic Review, 105(10), 3150-3182, available for download at www.ggdc.net/pwt), but also in this case the merge with the RBM dataset would result in a considerable reduction of the sample size.

[18] See www.govindicators.org. On the WGI methodology, see Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010), "The Worldwide Governance Indicators: A Summary of Methodology, Data and Analytical Issues", World Bank Policy Research Working Paper No. 5430. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1682130. The six dimensions of governance covered by the WGIs are voice and accountability; political stability and absence of violence/terrorism; government effectiveness; regulatory quality; rule of law; and control of corruption.

Following the merger with the macro variables, the sample size decreased slightly (by a hundred outcome ratings) mainly due to the drop of observations related to some overseas dependencies and just a few FCS. Notwithstanding, the composition of the dataset as well as the outcome rating sample distributions by workstream, income group, macro region, analytical group (i.e., FCS; small states), and project status (i.e., completed; ongoing) remained broadly unchanged (i.e., no relevant sample selection bias has been introduced).

IV. ECONOMETRIC SETUP

In line with the literature, in order to estimate the relationship between CD outcome ratings, on one hand, and the country-specific macroeconomic conditions as well as the project-specific features, on the other, we resort to traditional econometric models.

As a first step, we adopt the binary transformation of outcome ratings as a dependent variable and estimate the conditional probability of achieving—either largely or fully—the targeted outcomes through a linear probability model (LPM):

$$\Pr\left(OR_BIN_{kijt} = 1 | CF_{jt}, PF_{ijt}, CO_{kijt}\right) = \phi\left(\beta CF_{jt} + \gamma PF_{ijt} + \omega CO_{kijt}\right) \quad (1)$$

where ϕ is the identity function; OR_BIN_{kijt} is the binary rating assigned to outcome *k* of project *i* delivered to country *j* starting in year *t*; CF_{jt} is a set of macroeconomic variables of country *j*; PF_{ijt} is a set of specific features of project *i* delivered to country *j*; CO_{kijt} is a set of controls; β, γ, and ω are the parameters of the model. The recipient country's macroeconomic variables and the project-specific features are those described in the previous section. The set of controls include: (i) the initial year of the project; (ii) the latest year in which some of the main project features have been subject to change (i.e., project-change year); (iii) the income group and the macro-region of the recipient country; and (iv) a dummy variable for each CD workstream to account for potentially different rating practices, both within and across CDDs.

Importantly, while there can be multiple rated outcomes assigned to the recipient country within each project, the project features available in the dataset are not granular enough to allow for the inclusion of outcome-specific explanatory variables in the regressions other than, to some extent, the proxy for complexity (particularly when the number of outcomes per objective is used). All the other project-specific explanatory variables are at the project level (i.e., they apply to all the outcomes in the project). That is, all the outcome ratings assigned to the recipient country within a project will be regressed against the same set of project-specific features, as well as the same set of macroeconomic variables, plus one outcome-level covariate. This implies that we will be able to explain only a fraction of the variation of outcomes ratings associated to the recipient country. While this is an important shortcoming of the dataset, and therefore of the analysis, the estimates can nevertheless provide interesting preliminary findings.

The decision to start the analysis with the LPM reflects some of the characteristics of this model, including its simplicity and ease of interpretation, with the estimated coefficients representing the marginal probabilities of the corresponding explanatory variables.[19] This

[19] Marginal probabilities measure the change in the probability of achieving the targeted CD outcomes (i.e., the dependent variable equals 1) for a one-unit change of the explanatory variable of interest, everything else being equal.

notwithstanding, the LPM is also subject to well-known limitations, which led us to estimate a probit model to check for the robustness of results. In the latter case, the model specification is unchanged from what already described, with the only exception of ϕ now being the cumulative normal distribution.

Finally, we complete the analysis by switching to the original outcome rating scale as a dependent variable. The ordered categorial nature of the new dependent variable suggests to use an ordered-probit model to estimate the probability of realization of each of the four alternative ratings (*OR;* from *1-not achieved* to *4-fully achieved*), conditional on the usual sets of recipient country's macroeconomic variables, project specific features, and controls:

$$(2) \qquad \Pr\left(OR_{kijt} = z | CF_{jt}, PF_{ijt}, CO_{kijt}\right) = \phi\left(\alpha_z - \beta CF_{jt} - \gamma PF_{ijt} - \omega CO_{kijt}\right) - \phi\left(\alpha_{z-1} - \beta CF_{jt} - \gamma PF_{ijt} - \omega CO_{kijt}\right)$$

with *z=1, ..., 4* and where ϕ is the cumulative normal distribution, while α_z represent the thresholds (cutoffs) that allow to separate the probabilities of adjacent ratings.

V. RESULTS

A. Binary Outcome Ratings

Recipient Country's Macroeconomic Conditions

The results obtained using the LPM model are reported in Table 3.[20] Based on the estimates, the probability of achieving the targeted CD outcomes (i.e., *OR_BIN=1*) is positively correlated with per capita GDP growth (averaged over the year of project start and the four years preceding; see column 1 for the baseline regression). While the marginal effect is rather small—a one percentage point increase in growth rates translates into a 1.5 percentage point increase in the probability of a favorable outcome rating—it is strongly significant, indicating that the factors that promote growth—including better quality of policies and institutions—are also likely to translate into greater capacity to benefit from CD (in this regard, see the findings related to the WGIs illustrated later in this section). On the other hand, growth volatility and the level of per capita GDP turned out not to be statistically correlated with CD outcomes.

Results also confirm that both FCS and, less robustly, small states are associated with a lower probability of achieving the targeted outcomes compared to other countries (by around seven and five percentage points, respectively, other things being equal), underscoring the importance of continuing to strengthen the country-tailored approach to CD. To shed some light on this result and verify if it reflects some sort of absorptive constraints, we introduced in the regressions a variable accounting for the number of CD projects simultaneously delivered to each recipient country, either in total or focused on the same workstream within which the relevant targeted outcome falls.[21] The evidence is, however, inconclusive: while there is some weak indication—not statistically robust—of a negative correlation between outcome ratings and the number of projects delivered, there is no indication of a faster

[20] Based on the estimated LPM, more than 98 percent of the predicted probabilities of achieving the CD targets are within the 0−1 interval, pointing to the reliability of this model for the analysis at hand.

[21] In order to consider two different projects as simultaneously delivered, we required that they overlap for at least either one or two years (the results we obtain are broadly similar in both cases). The number of projects simultaneously delivered was then scaled with the recipient country's population size.

decline for FCS and small states. A better understanding of the mechanisms underlying the lower chances for FCS and small states of achieving the targeted results—as well as, more generally, of the relationship between volume of projects and ratings—is thus left for future research, when a larger dataset will be available.

The interpretation of the results associated with countries that, while being engaged in a CD project, have also been engaged in an IMF program—either with or without borrowing arrangement—provides interesting insights. In fact, though not robustly significant, the correlation signs indicate that the CD recipient country ownership—proxied by the successful completion of the IMF program—or lack thereof—proxied by programs that went off track—can be potential factors associated with the probability of achieving the targeted outcomes.[22]

CD Project-Specific Features

Based on the LPM estimates (Table 3), CD projects delivered through the global network of regional centers (RTACs)—as part of the broader integrated IMF CD delivery—are positively correlated with the probability of achieving the CD targets, indicating the importance of continued engagement with the recipient countries to better tailor CD to their specific needs and circumstances, adjust nimbly to the rapidly evolving conditions on the ground, and offer hands-on advice. This finding is strongly robust across specifications. However, it may be influenced by differences in the type or level of ambition of targeted outcomes between IMF headquarters CD and that delivered through regional centers.[23]

The importance of continued engagement and field presence is further reinforced by the positive and statistically significant coefficient associated with the amount of CD services provided by in-country resident advisors, where present. On the other hand, there does not seem to be a significant correlation with the amount of services provided (either in-country or remotely) by short-term experts, at least in regressions based on binary outcomes.

The positive coefficient associated with the completion share of the project confirms that outcome ratings tend to be higher at later stages of the project lifetime (i.e., the older the project, the higher the ratings tend to be).[24] This result is in line with the fact that, as outlined by the 2018 CD Strategy Review, capacity building is a complex process, which requires a medium-term approach, with targeted outcomes taking time to achieve, and with possible setbacks along the way.

Finally, the coefficient associated with the number of workstreams included in the project and delivered by the same CDD indicates that pursuing multiple reform plans at the same time is negatively correlated with the likelihood of achieving the targeted results.

[22] Completed programs are those whose reviews have been all completed during the program period, including if they were completed with delays, after rephasing, or during a program extension. Off track programs include both programs for which at least two reviews were completed and at least two reviews were not completed at the end of the program (off track mid-program), and programs for which at most one review was completed and at least two reviews were not completed at the end of the program (quickly off track). Unfortunately, we do not have a proxy for ownership for CD recipient countries that have not been engaged in an IMF program.

[23] It is often the case, particularly for multiyear complex reform projects, that projects delivered by headquarters-based staff and RTACs are designed in a complementary fashion so the nature of targeted outcomes may not be identical for the two groups.

[24] It should be recalled that, while past practices across CDDs have varied, based on the IMF RBM governance framework, outcomes should be rated annually, with the possibility of subsequent revisions throughout the project lifetime to properly reflect the results achieved by the CD recipient country.

Other Considerations

Overall, the results presented so far seem broadly in line with intuition. Notwithstanding, some considerations need to be kept in mind when interpreting them.

Firstly, the presence of FCS and small state dummies in all our specifications, in addition to the low number of rated outcomes available for several countries in the sample, prevented from introducing country fixed effects in the regressions.[25] Even if we were to account for them, while excluding the other macroeconomic covariates, the share of explained variation in outcome ratings would increase only modestly compared to the levels reported in our tables (from around 13 to 20 per cent).[26] This seems to indicate that most of the variation in outcome ratings is likely coming from within countries rather than between countries, [27] which is in line with what Denizer and others (2013) found for the outcomes of development assistance projects. The limited share of outcome ratings variation that we are able to explain with the set of project-specific features at our disposal is also in line with the related literature and points, as Denizer and others (2013) emphasized, to the possibly large set of relevant project-specific factors that remain unobserved and thus unaccounted for, including at the outcome level.

Relatedly, while sensible, the results should be interpreted with caution as they may be subject to the omitted variable bias to the extent that some of the unobserved factors are correlated with both the dependent and independent variables.

Among them—and most importantly—Denizer and others (2013) pointed to the degree of difficulty inherent in the project. For example, in our context the more difficult and challenging the project, (1) the lower the probability of achieving the targeted outcome (dependent variable); and (2) possibly the greater the amount of CD services poured into it by resident advisors and short-term experts (independent variables). While the absence of suitable instruments prevents from addressing the ensuing endogeneity issue, a few considerations on the direction of the potential bias may offer some insights on the interpretation of the estimated coefficients.

Based on the correlation signs described in the example above, the potential bias in the estimated coefficients for the CD services provided by both resident advisors and short-term experts would be downward. As a result, such estimated coefficients could be considered as possible lower bounds.

There are uncertainties regarding the direction of the bias affecting project complexity, whose correlation with the degree of project difficulty could potentially go either way: positive (generating a downward bias)—if the larger challenges are addressed through a more complex (all-encompassing) project—or negative (generating an upward bias)—if they are addressed through a coordinated set of smaller (more targeted and, possibly, sequenced) projects.

[25] As with other issues, going forward the availability of a larger dataset, together with different ways of accounting for FCS and small states, may allow to overcome this limitation.

[26] In this case, the findings related to the project-specific covariates would still hold.

[27] Both between and within projects.

The other CD project-specific variables included in the regression—those controlling for RTAC projects and the project completion shares—are likely uncorrelated with the omitted variable and thus not subject to the ensuing bias.

By the same line of reasoning just illustrated for the set of project-specific variables, the omission of the degree of project difficulty can potentially bias also the estimated coefficients of the recipient country's macroeconomic conditions, particularly of per capita GDP growth (upward) and the dummy variable controlling for FCS (downward). Notwithstanding, the signs of the estimated coefficients are consistent with the findings available in the related literature (Denizer and others, 2013; Presbitero, 2016; Caselli and Presbitero, 2021), mitigating—at least in part—the reasons for concern.

More generally, however, besides the degree of project difficulty, one could probably think of other omitted variables that might as well impact the estimates in our model. It therefore remains appropriate, also in view of some remaining limitations of the dataset, to interpret the results presented in this paper with caution and as partial correlations rather than causal relationships.

Digging Deeper

As a next step, we dug deeper into the findings and—firstly—introduced in the baseline specification a proxy for the quality of policies and institutions for each of the countries in the sample, while excluding the FCS dummy variable.[28] In particular, we used the WGI on regulatory quality, which reflects perceptions of the ability of the government to formulate and implement sound policies and regulations. The estimates of this slightly modified specification broadly confirm the relevant findings discussed so far and indicate that recipient countries suffering from a less severe quality gap in policies and institutions are associated with higher chances of achieving the CD targets (Table 3, column 2).[29]

Subsequently, while returning to our baseline specification, we looked into the macro-regional disaggregation of some of the main findings. This kind of analysis calls for an even more cautious interpretation of results, given the still relatively limited size of the dataset; at the same time, it may provide some interesting insights and indicate potential areas for deeper future analysis and evaluations. We started by interacting the small-state dummy variable with a set of regional dummies in order to estimate separate effects for each macro-region. The dataset available for the estimates includes 532 outcome ratings related to small states, largely concentrated in the Western Hemisphere region and, to a slightly lesser extent, in the Asian-Pacific region and in Sub-Saharan Africa; observations for the European region are few, while the sample does not include any outcome ratings for small states in the Middle-East and Central Asia region (both the latter regions feature only one small state). Based on the results (Table 3, columns 3, 5–7), it appears that the negative small-state effect estimated for the sample aggregate is actually attributable to small states in the Western Hemisphere region, with the Asian-Pacific small states—on the opposite—featuring more

[28] Although the spectrum of fragilities can be very broad and not be limited to the quality of policies and institutions, in the absence of an all-encompassing measure we believe that the latter variable represents a reasonable approximation for the task at hand and has the advantage of being available for both fragile and non-fragile countries.

[29] We would have obtained the same result if, instead of regulatory quality, we had used the WGIs on voice and accountability or on rule of law. The correlation between the probability of achieving the CD targets and the remaining WGIs (i.e., government effectiveness, political stability and absence of violence/terrorism, control of corruption) is also positive, although not always statistically significant.

favorable ratings than the rest of the sample. In Sub-Saharan Africa and in Europe the small-state effect is negative but not robustly significant.

Next, we conducted a similar exercise for FCS whose observations in the sample are 1,010, largely coming from the Sub-Saharan region and, to a much smaller extent, from the Asia-Pacific and Middle-East and Central Asia regions; outcome ratings from the European and Western-Hemisphere regions, which include only one FCS each, are few. Based on the estimates, the aggregate negative FCS effect seems to be largely attributable to Sub-Saharan Africa and, less robustly, to the Asia-Pacific region (Table 3, columns 4–7) where most of the FCS are also small states. The positive coefficients estimated for FCS in the Western Hemisphere and Europe should be viewed with great caution given the very limited number of outcome ratings available for these regions.

Finally, we experimented with alternative proxies for the complexity—or ambition—of the CD project. The negative correlation estimated between the probability of achieving the CD targeted outcomes and the number of workstreams included in the project and delivered by the same CDD is confirmed—albeit with smaller magnitude—when the proxy is replaced by either the number of objectives per relevant workstream or the number of outcomes per relevant objective, with the latter—however—not being statistically significant (Table 3, columns 6 and 7, respectively).

Overall, the main findings of the analysis illustrated so far seem to be rather robust. They are broadly confirmed when, instead of the LPM, we estimate a probit model (Table 4), with signs and sizes of the marginal probabilities associated to each covariate being consistently aligned. Results for both LPM and probit models broadly holds also when we cluster the standard errors at the country level, with some weakening of the findings related to IMF program countries (Tables A1 and A2 in the Annex).

B. Original Outcome Ratings

In the second part of the analysis, we replaced the binary dependent variable with the original RBM ratings and estimated the ordered-probit model (2) to calculate the probability of realization of each of the four alternative outcome ratings (*OR=1-not achieved; 2-partially achieved; 3-largely achieved; 4-fully achieved*), conditional on the usual set of covariates and controls. The quantitative interpretation of the order-probit is not as straightforward as that of a binary model, as it requires transforming the estimated coefficient for each of the covariates into four different marginal effects, one for each of the alternative outcome ratings.[30] Overall, however, the ordered-probit results broadly confirm the main findings obtained with the LPM and probit models, with some qualifications (see Table 5 for the ordered probit estimates; Table 6 for the marginal effects associated with the baseline specification).[31]

Recipient Country's Macroeconomic Conditions

As regards the recipient country's macroeconomic conditions, based on the marginal effects of the baseline specification a one percentage point increase in per capita GDP growth rates

[30] As shown above, in binary models each covariate is associated with only one marginal effect. In any case, the interpretation of the marginal effects is the same for both binary and ordered-probit models, as they measure the change in the probability of realization of each alternative outcome rating associated with a unit change in the covariate.

[31] Marginal effects for the remaining specifications are available from the author upon request.

is associated with a small increases in both probabilities of higher ratings (by around one percentage point for *OR* equal to *3* and 0.5 for *OR* equal to *4*; Tables 6) and into similar decreases in both probabilities of lower ratings (i.e., *OR* equal to *1* or *2*), thereby confirming previous findings.

Results also confirm that both FCS and small states feature lower probabilities of achieving the two higher ratings and higher probabilities of achieving the two lower ones compared to the rest of the sample, other things being equal (Table 6). Furthermore, ordered-probit results disaggregated at the macro-regional level broadly corroborate the findings obtained with the LPM and probit models, with the negative correlation associated with FCS largely attributable to Sub-Saharan Africa and that associated with small states attributable to the Western Hemisphere (Table 5).

The correlations between outcome ratings and the contextual engagement of the CD recipient countries in IMF programs are also in line with the indications emerged from the regressions based on binary ratings. The positive and robustly significant correlation between ratings and successfully completed programs supports the importance of traction and ownership, as well as of a close integration between the Fund's CD, lending, and surveillance activities to foster impact and effectiveness.

CD Project-Specific Features

Moving to the project-specific variables, the ordered-probit models corroborate all previous findings on the positive correlation between, on one side, the probability of achieving higher outcome ratings and, on the other, RTAC projects and the services provided by resident advisors. Furthermore, based on the estimates, higher ratings are now positively correlated also with the amount of services provided by short-term experts. The correlation between the complexity of the project and outcome ratings remains negative and significant as long as complexity is proxied by the number of workstreams included in the project and delivered by the same CDD, while it becomes not significant when complexity is proxied by either the number of objectives per relevant workstream or the number of outcomes per relevant objective.

Finally, as in the case of the regressions based on binary ratings, also with ordered-probit models the main findings are broadly confirmed when we cluster the standard errors at the country level (Table A3 in the Annex).

VI. CONCLUSIONS

The analysis presented in this paper provides a first empirical insight into when IMF CD is most likely to help member countries building capacity and institutions. In particular it looks at how outcomes are correlated, on the one hand, with macroeconomic conditions at the time of CD delivery and, on the other, with the specific features of the project. To conduct the analysis, we used, for the first time in the literature, the IMF's RBM dataset.

Importantly, the analysis focuses on correlations, rather than causal relationships. As such, findings are preliminary and should be interpreted with caution and as potential inputs for further in-depth evaluations that can inform Fund CD strategy, prioritization processes and delivery, with the ultimate goal of helping member countries achieve better outcomes.

As for the macroeconomic conditions, results indicate that:

- The likelihood of achieving targeted results is positively correlated with per capita GDP growth, presumably suggesting that the factors that promote growth—including better quality of policies and institutions—are also likely to foster greater capacity to benefit from CD.
- Fragility issues, severe quality gaps in policies and institutions, as well as small country size are associated with lower outcome ratings, other things being equal.
- Recipient country's strong ownership of the CD project is associated with higher chances of an impactful CD, as suggested by the correlations between outcome ratings and the country's concomitant engagement in IMF programs (whose successful completion, or lack thereof, provides indications on ownership). Relatedly, close integration between the IMF CD, lending, and surveillance activities is also associated with higher probabilities of achieving the targeted outcomes.

Moving to the CD-project specific features, findings indicate that:

- Continued engagement with the recipient country can be an important ingredient for impactful CD, as illustrated by the positive correlation between outcome ratings and the services provided by both regional centers and in-country resident advisors as part of the broader integrated IMF CD delivery. Based on experience, continued engagement favors country-tailoring, enhances responsiveness to changing needs and circumstances, facilitates hands-on support, and fosters the IMF's role as a trusted advisor.
- Pursuing multiple reform plans at the same time is negatively associated with the likelihood of achieving the targeted results, as indicated by the estimated coefficients of the proxies for project complexity, particularly when based on the number of workstreams delivered by the same CDD.
- Completed projects tend to feature higher ratings than ongoing ones, confirming that CD is a complex medium-term process with targeted outcomes taking time to achieve and possible setbacks along the way.

It is important to note that this set of results should not be interpreted as questioning whether CD should be directed to countries facing economic headwinds or that are fragile and/or small. Rather, it should be considered as a starting point to investigate the reasons behind weak results (e.g., overly-ambitious project design, under-resourcing, weak commitment from the recipient authorities) and consequently undertake remedial actions to enhance impact and effectiveness.

The size of the IMF RBM dataset continues to increase rapidly as more CD projects are delivered and rated. Additionally, the ongoing introduction of new processes and systems for the management and administration of the IMF CD activities will facilitate the gathering of more information on some important project-level characteristics, such as more granular and precise data on the amount of resources dedicated to each project and the assessment of potential risks that might hamper impact on the ground. The quality of the data will also improve as a result of increased consistency in rating practices and targeted results, both across and within CDDs as well as between IMF headquarters and regional centers. A richer dataset, in terms of both quantity and quality of observations and project-level characteristics, may allow future research to obtain more precise estimates, mitigate the omitted variable bias impacting on the analysis, explain a larger share of the variation in outcome ratings, better

understand the mechanisms underlying some of the results—in particular those related to FCS and small states—and deepen the investigation in several directions, including at a more disaggregated level. It will also allow to pursue other research avenues, such as investigating how RBM outcome ratings—which are based on the CD project manager's assessment informed by pre-defined and verifiable indicators—relate to long-term impact.

As this preliminary analysis has shown, the RBM dataset can deliver potentially relevant information and contribute, together with broader evaluations, to (1) define and implement the IMF CD strategy; (2) inform the prioritization and resource allocation processes as CD demand exceeds delivery levels; (3) enhance CD effectiveness through better country-tailoring; (4) support CD project management during project execution; and (5) foster accountability, knowledge-sharing and reporting on CD results.

References

Caselli, Francesca G., and Andrea F. Presbitero, 2021, *Aid Effectiveness in Fragile States*, in Chami, Ralph, Raphael Espinoza, and Peter J. Montiel, eds., *Macroeconomic Policy in Fragile States*, Oxford Scholarship Online. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198853091.001.0001/oso-9780198853091?rskey=l4GGlW&result=1

Chami, Ralph, Elorm Darkey, and Oral H. Williams, 2021, *A Time to Build: Does Technical Assistance Matter for Revenue Mobilization?* IMF Working Paper WP/21/35 (Washington: International Monetary Fund).

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay, 2013, *Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance*, Journal of Development Economics, Vol. 105, pp. 288-302.

Edison, Hali, Dimitre Milkov, and Josephat Koima, 2018, *Evaluating IMF Training: What Can We Learn?* IMF Working Paper WP/18/151 (Washington: International Monetary Fund).

International Monetary Fund, 2017, *New Common Evaluation Framework*, IMF Policy Papers (Washington). https://www.imf.org/en/publications/policy-papers/issues/2017/04/27/pp040717new-common-evaluatioin-framework-for-imf-capacity-developement

International Monetary Fund, 2017, *Building Capacity in Fragile States*, IMF Policy Papers (Washington). https://www.imf.org/en/Publications/Policy-Papers/Issues/2017/06/14/pp041817building-fiscal-capacity-in-fragile-state

International Monetary Fund, 2018, *2018 Review of the Fund's Capacity Development Strategy*, IMF Policy Papers (Washington). https://www.imf.org/en/Publications/Policy-Papers/Issues/2018/11/20/2018-review-of-the-funds-capacity-development-strategy

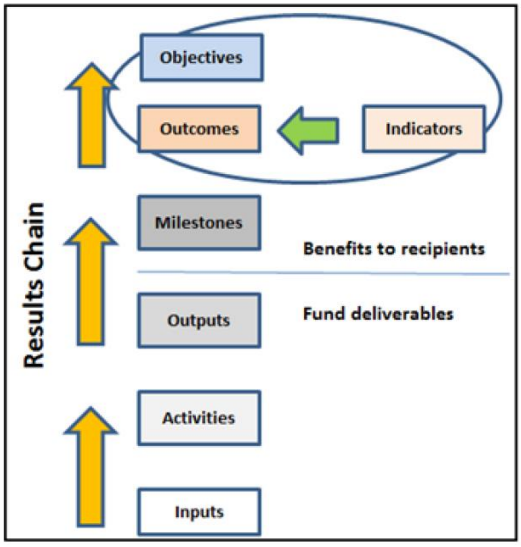
International Monetary Fund, 2019, *IMF Policies and Practices on Capacity Development*, IMF Policy Papers (Washington). https://www.imf.org/en/Publications/Policy-Papers/Issues/2019/11/14/IMF-Policies-and-Practices-on-Capacity-Development-48811

International Monetary Fund, 2020, *Updated Common Evaluation Framework for Capacity Development and Guidance Note*, IMF Policy Papers (Washington). https://www.imf.org/en/Publications/Policy-Papers/Issues/2020/09/24/Updated-Common-Evaluation-Framework-For-IMF-Capacity-Development-And-Guidance-Note-49779

Presbitero, Andrea F., 2016, *Too Much and Too Fast? Public Investment Scaling-Up and Absorptive Capacity*, Journal of Development Economics, Vol. 120, pp. 17-31.

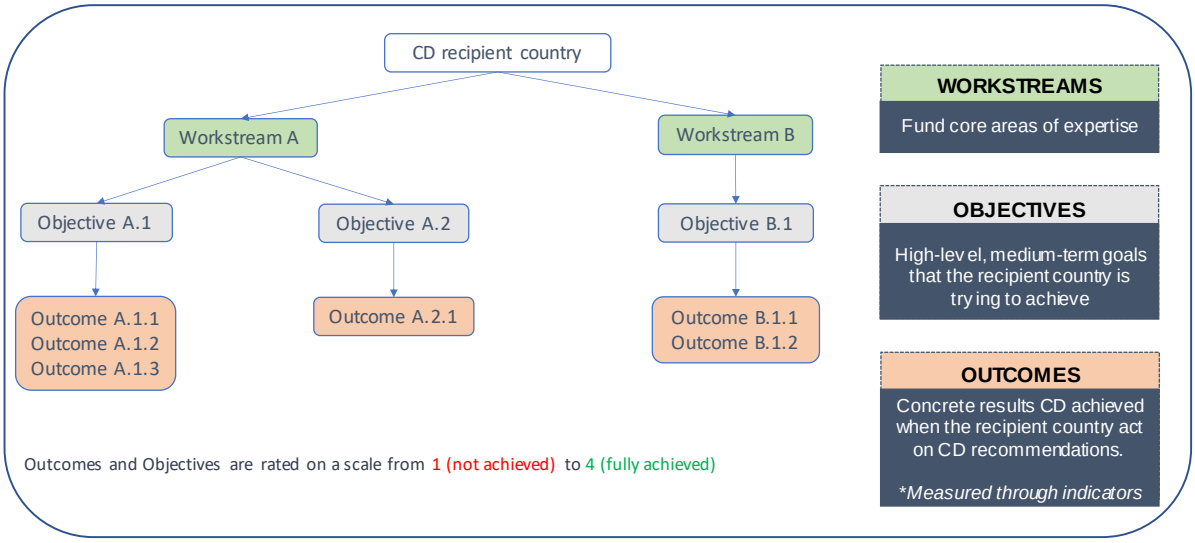
Figures

Figure 1. RBM Framework—Results Chain (log frame)



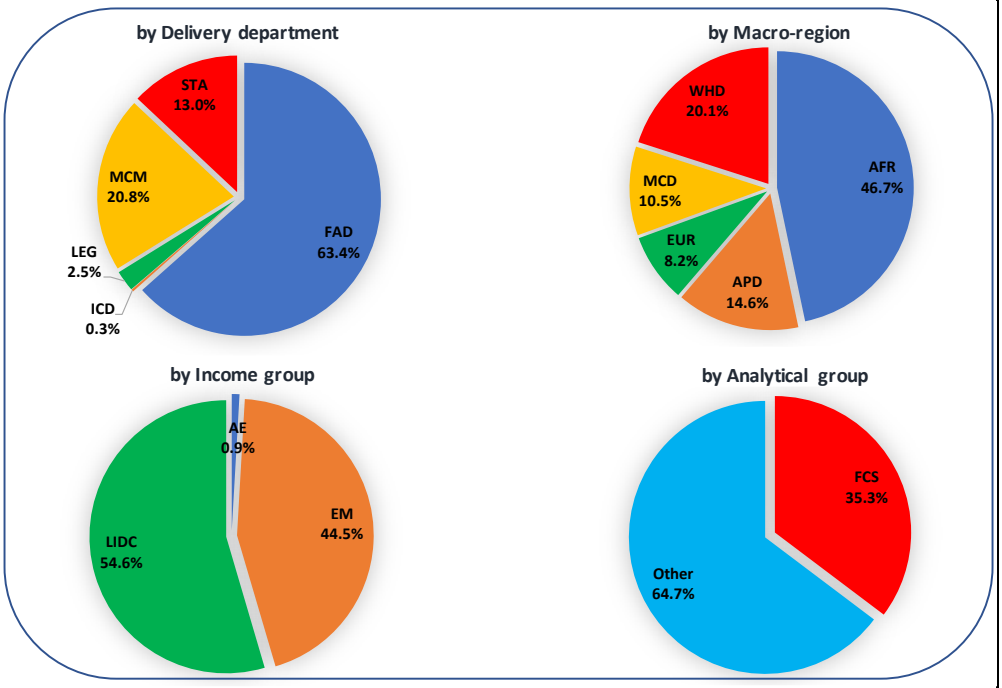
Source: 2018 IMF CD Strategy Review, Box 2.

Figure 2. CD Project Design—an Example



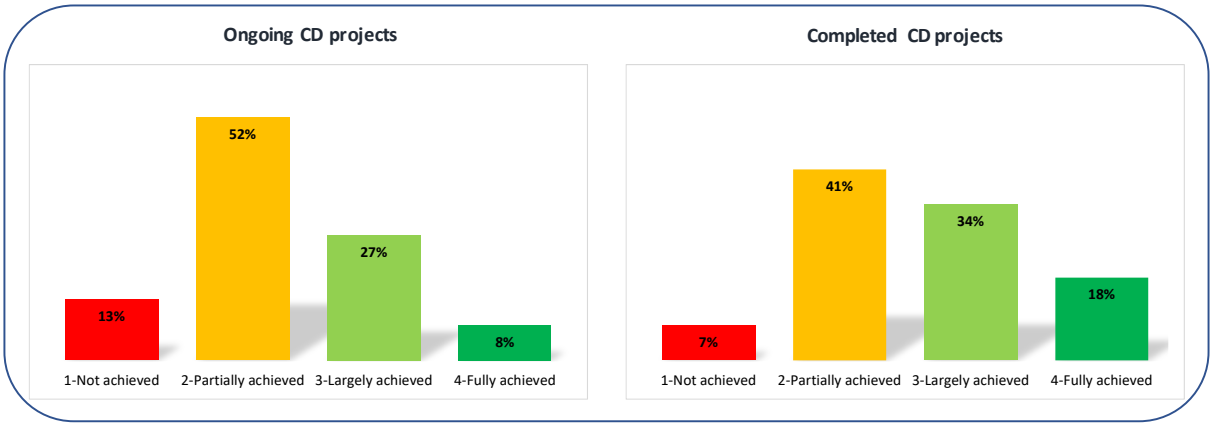
Note: Examples of workstreams, objectives, and outcomes from the IMF RBM catalog are reported in Section II of the main text.

Figure 3. Composition of the RBM Dataset



Note: As of February 2020. Author's calculation. FAD: Fiscal Affairs Department; MCM: Monetary and Capital Markets Department; STA: Statistics Department; LEG: Legal Department; ICD: Institute for Capacity Development; AFR: African Department; WHD: Western Hemisphere Department; APD: Asia & Pacific Department; MCD: Middle East & Central Asia Department; EUR: European Department.

Figure 4. RBM Outcome Ratings—Frequency Distributions



Note: As of February 2020. Author's calculation. Outcome ratings for ongoing projects are subject to potential revisions until projects are completed and a final assessment can be done.

Tables

Table 1. RBM Outcome Ratings—Summary Statistics

Dataset composition	Number of rated outcomes	Original ratings (1-4)		Binary ratings (0-1)	
		Mean	Std. Dev.	Mean	Std. Dev.
Total dataset	2957	2.35	0.81	0.37	0.48
of which:					
Completed projects	321	2.64	0.90	0.51	0.50
Ongoing projects	2636	2.31	0.79	0.36	0.48
Income groups					
Advanced economies	27	2.44	0.97	0.41	0.50
of which:					
Completed projects	14	2.86	1.10	0.64	0.50
Ongoing projects	13	2.00	0.58	0.15	0.38
Emerging markets	1317	2.41	0.89	0.42	0.49
of which:					
Completed projects	159	2.77	0.94	0.59	0.49
Ongoing projects	1158	2.36	0.87	0.40	0.49
Low-income countries	1613	2.29	0.73	0.34	0.47
of which:					
Completed projects	148	2.49	0.80	0.41	0.49
Ongoing projects	1465	2.27	0.72	0.33	0.47
Analytical groups					
Fragile and conflict affected states (FCS)	1043	2.22	0.71	0.28	0.45
of which:					
Completed projects	67	2.37	0.71	0.33	0.47
Ongoing projects	976	2.21	0.70	0.28	0.45
Small states	537	2.31	0.75	0.37	0.48
of which:					
Completed projects	23	2.39	0.66	0.39	0.50
Ongoing projects	514	2.31	0.76	0.37	0.48
Macro-regions					
Asia-Pacific	431	2.29	0.84	0.35	0.48
of which:					
Completed projects	37	2.11	0.91	0.35	0.48
Ongoing projects	394	2.30	0.84	0.35	0.48
Europe	241	2.29	0.97	0.32	0.47
of which:					
Completed projects	57	2.96	0.96	0.63	0.49
Ongoing projects	184	2.08	0.88	0.22	0.41
Middle East and Central Asia	311	2.36	0.83	0.38	0.49
of which:					
Completed projects	17	3.24	0.75	0.82	0.39
Ongoing projects	294	2.31	0.81	0.36	0.48
Sub-Saharan Africa	1381	2.27	0.69	0.32	0.47
of which:					
Completed projects	116	2.31	0.68	0.29	0.46
Ongoing projects	1265	2.26	0.69	0.32	0.47
Western Hemisphere	593	2.59	0.91	0.54	0.50
of which:					
Completed projects	94	2.96	0.87	0.71	0.45
Ongoing projects	499	2.52	0.91	0.51	0.50

Note: As of February 2020. Author's calculation.

Table 2. Other Variables in the Dataset—Summary Statistics

Variable	Description	Source	Number of observations	Mean	Std. Dev.	Min.	Max.
Log of real per capita GDP	Logarithm of real per capita GDP (constant 2010 USD); average over year one of CD project plus four years preceding.	Author's calculation on World Development Indicators - World Bank.	2857	7.68	1.09	5.44	11.42
Real per capita GDP growth	Annual real per capita GDP growth (percentage); average over year one of CD project plus four years preceding.	Author's calculation on World Development Indicators - World Bank.	2857	2.05	2.69	-9.91	13.93
Volatility of real per capita GDP growth	Standard deviation of real per capita GDP growth over year one of CD project plus nine years preceding.	Author's calculation on World Development Indicators - World Bank.	2856	0.03	0.03	0	0.47
Fragile and conflict affected state (FCS)	Dummy equal to one if the CD recipient country is a fragile and conflict afflected state based on the IMF internal definition.	Author's calculation on IMF classification.	2957	0.35	0.48	0	1
Small state	Dummy equal to one if the CD recipient country is a small state based on the IMF internal definition.	Author's calculation on IMF classification.	2957	0.18	0.39	0	1
IMF completed program	Dummy equal to one if the CD recipient country has been contextually engaged in a successfully completed IMF program, either with or without a borrowing arrangement.	Author's calculation on IMF internal data.	2957	0.14	0.35	0	1
IMF off-track program	Dummy equal to one if the CD recipient country has been contextually engaged in an IMF program, either with or without a borrowing arrangement, that went rapidly off-track.	Author's calculation on IMF internal data.	2957	0.06	0.24	0	1
Completion share of CD project	Share of project length already passed (percentage).	Author's calculation on RBM data - IMF.	2957	47.21	30.15	0.21	100
RTAC project	Dummy equal to one if the CD project is delivered through an RTAC.	Author's calculation on RBM data - IMF.	2957	0.61	0.49	0	1
Resident Advisor	CD services provided by an in-country resident advisor associated with the project and the recipient country; measured in full-time equivalents (FTEs).	Author's calculation on IMF internal data.	2937	0.18	0.63	0	5.28
Short-term expert	CD services provided by short-term experts associated with the project and the recipient country, including both Fund staff and external experts selected from the IMF's roster; measured in full-time equivalents (FTEs).	Author's calculation on IMF internal data.	2937	0.57	0.72	0	6.87
N. of workstreams	Number of workstreams per recipient country in the CD project.	Author's calculation on RBM data - IMF.	2957	1.16	0.45	1	4
N. of objectives	Number of objectives per workstream in the CD project.	Author's calculation on RBM data - IMF.	2957	2.53	1.56	1	8
N. of outcomes	Number of outcomes per objectives in the CD project.	Author's calculation on RBM data - IMF.	2957	2.35	1.20	1	7
Regulatory quality	Average over year one of CD project plus two preceding years.	Author's calculation on Worldwide Governance Indicators.	2896	-0.44	0.57	-2.24	1.83

Note: Author's calculation.

Table 3. Country- and CD Project-Specific Variables—Linear Probability Model

Dependent variable: binary outcome rating (0-1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.00609	-0.00464	-0.00163	-0.0106	-0.00426	0.000138	0.00296
	(0.0170)	(0.0166)	(0.0171)	(0.0176)	(0.0177)	(0.0177)	(0.0178)
Per capita real GDP growth	0.0148***	0.0145***	0.0153***	0.0140***	0.0146***	0.0150***	0.0148***
	(0.00379)	(0.00377)	(0.00383)	(0.00387)	(0.00392)	(0.00396)	(0.00394)
Volatility of per capita real GDP growth	-0.316	-0.332	-0.270	-0.196	-0.193	-0.240	-0.221
	(0.325)	(0.322)	(0.330)	(0.337)	(0.341)	(0.348)	(0.343)
FCS	-0.0662***		-0.0659***				
	(0.0247)		(0.0247)				
Small state	-0.0490*	-0.0456		-0.0418			
	(0.0288)	(0.0288)		(0.0288)			
IMF completed program	0.0355	0.0313	0.0260	0.0410	0.0306	0.0383	0.0354
	(0.0300)	(0.0301)	(0.0302)	(0.0303)	(0.0305)	(0.0306)	(0.0305)
IMF off-track program	-0.0710*	-0.0845**	-0.0786**	-0.0539	-0.0635	-0.0603	-0.0600
	(0.0382)	(0.0373)	(0.0391)	(0.0392)	(0.0403)	(0.0402)	(0.0405)
Completion share of CD project	0.00365***	0.00375***	0.00377***	0.00370***	0.00380***	0.00359***	0.00363***
	(0.000733)	(0.000729)	(0.000737)	(0.000737)	(0.000740)	(0.000748)	(0.000743)
RTAC project	0.0923***	0.0885***	0.0813***	0.0962***	0.0857***	0.0885***	0.0859***
	(0.0295)	(0.0293)	(0.0296)	(0.0298)	(0.0300)	(0.0300)	(0.0301)
Resident advisor	0.0368**	0.0356**	0.0392**	0.0367**	0.0392**	0.0306*	0.0313*
	(0.0179)	(0.0179)	(0.0178)	(0.0178)	(0.0178)	(0.0179)	(0.0180)
Short-term expert	-0.00559	-0.00926	-0.00320	-0.00617	-0.00258	-0.0224	-0.0274
	(0.0174)	(0.0177)	(0.0177)	(0.0176)	(0.0179)	(0.0172)	(0.0168)
N. of workstreams per country	-0.118***	-0.108***	-0.122***	-0.118***	-0.122***		
	(0.0302)	(0.0303)	(0.0305)	(0.0302)	(0.0304)		
N. of objectives per workstream						-0.0145*	
						(0.00840)	
N. of outcomes per objective							-0.0131
							(0.00837)
Regulatory quality		0.0565**					
		(0.0234)					
Small state in Africa			-0.0329		-0.0350	-0.0289	-0.0256
			(0.0440)		(0.0441)	(0.0438)	(0.0440)
Small state in Asia-Pacific			0.151***		0.150***	0.146**	0.135**
			(0.0580)		(0.0579)	(0.0585)	(0.0578)
Small state in Western Hemisphere			-0.194***		-0.180***	-0.182***	-0.182***
			(0.0512)		(0.0518)	(0.0520)	(0.0522)
Small state in Europe			-0.133*		-0.0859	-0.0829	-0.0848
			(0.0707)		(0.0735)	(0.0729)	(0.0728)
FCS in Africa				-0.0975***	-0.0877***	-0.0720**	-0.0717**
				(0.0317)	(0.0317)	(0.0317)	(0.0318)
FCS in Asia Pacific				-0.0726	-0.0905*	-0.0821	-0.0729
				(0.0524)	(0.0516)	(0.0514)	(0.0515)
FCS in Middle East and Centr. Asia				-0.0959	-0.0878	-0.0754	-0.0907
				(0.0624)	(0.0629)	(0.0630)	(0.0632)
FCS in Western Hemisphere				0.340**	0.260	0.266*	0.288*
				(0.153)	(0.158)	(0.159)	(0.159)
FCS in Europe				0.0726	0.0658	0.0815	0.0683
				(0.0642)	(0.0666)	(0.0675)	(0.0672)
Observations	2,836	2,836	2,836	2,836	2,836	2,836	2,836
R-squared	0.134	0.134	0.142	0.138	0.144	0.140	0.140
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the estimated coefficients and, in brackets, the associated robust standard errors. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Table 4. Country- and CD Project-Specific Variables—Probit Model

Dependent variable: binary outcome rating (0-1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.0102	-0.0081	-0.0066	-0.0157	-0.0096	-0.0038	-0.0007
	(0.0185)	(0.0185)	(0.0187)	(0.0191)	(0.0193)	(0.0193)	(0.0192)
Per capita real GDP growth	0.0166***	0.0165***	0.0174***	0.0157***	0.0166***	0.0172***	0.0169***
	(0.0043)	(0.0044)	(0.0044)	(0.0044)	(0.0045)	(0.0045)	(0.0045)
Volatility of per capita real GDP growth	-0.3056	-0.3531	-0.2688	-0.1671	-0.1861	-0.2594	-0.2286
	(0.3620)	(0.3688)	(0.3626)	(0.3706)	(0.3728)	(0.3778)	(0.3729)
FCS	-0.0799***		-0.0815***				
	(0.0279)		(0.0279)				
Small state	-0.0502	-0.0458		-0.0419			
	(0.0316)	(0.0317)		(0.0317)			
IMF completed program	0.0392	0.0342	0.0292	0.0470	0.0360	0.0434	0.0408
	(0.0320)	(0.0321)	(0.0323)	(0.0325)	(0.0327)	(0.0328)	(0.0326)
IMF off-track program	-0.0787*	-0.0943**	-0.0858*	-0.0575	-0.0677	-0.0643	-0.0625
	(0.0449)	(0.0440)	(0.0458)	(0.0460)	(0.0470)	(0.0467)	(0.0471)
Completion share of CD project	0.0041***	0.0043***	0.0043***	0.0043***	0.0044***	0.0042***	0.0042***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
RTAC project	0.1078***	0.1047***	0.0935***	0.1122***	0.0977***	0.1026***	0.0981***
	(0.0332)	(0.0331)	(0.0337)	(0.0335)	(0.0340)	(0.0342)	(0.0341)
Resident advisor	0.0396**	0.0383**	0.0418**	0.0379**	0.0405**	0.0319*	0.0322*
	(0.0192)	(0.0192)	(0.0193)	(0.0192)	(0.0193)	(0.0194)	(0.0194)
Short-term expert	-0.0077	-0.0101	-0.0082	-0.0081	-0.0076	-0.0289	-0.0363*
	(0.0206)	(0.0211)	(0.0213)	(0.0210)	(0.0217)	(0.0214)	(0.0211)
N. of workstreams per country	-0.1329***	-0.1208***	-0.1397***	-0.1332***	-0.1401***		
	(0.0342)	(0.0342)	(0.0346)	(0.0342)	(0.0345)		
N. of objectives per workstream						-0.0178*	
						(0.0093)	
N. of outcomes per objective							-0.0158*
							(0.0093)
Regulatory quality		0.0639**					
		(0.0260)					
Small state in Africa			-0.0389		-0.0412	-0.0352	-0.0312
			(0.0491)		(0.0494)	(0.0491)	(0.0492)
Small state in Asia-Pacific			0.1744***		0.1776***	0.1713***	0.1596***
			(0.0618)		(0.0621)	(0.0625)	(0.0617)
Small state in Western Hemisphere			-0.1882***		-0.1724***	-0.1736***	-0.1725***
			(0.0523)		(0.0528)	(0.0530)	(0.0530)
Small state in Europe			-0.2179*		-0.1519	-0.1478	-0.1420
			(0.1203)		(0.1237)	(0.1218)	(0.1209)
FCS in Africa				-0.1180***	-0.1069***	-0.0853**	-0.0861**
				(0.0353)	(0.0353)	(0.0349)	(0.0351)
FCS in Asia Pacific				-0.0870	-0.1157**	-0.1038*	-0.0954*
				(0.0585)	(0.0579)	(0.0575)	(0.0575)
FCS in Middle East and Centr. Asia				-0.1052	-0.0964	-0.0813	-0.1005
				(0.0674)	(0.0679)	(0.0679)	(0.0680)
FCS in Western Hemisphere				0.3676	0.2887	0.2990	0.3264
				(0.2464)	(0.2507)	(0.2531)	(0.2524)
FCS in Europe				0.1082	0.0939	0.1150	0.0963
				(0.0854)	(0.0877)	(0.0883)	(0.0875)
Observations	2,829	2,829	2,829	2,829	2,829	2,829	2,829
Pseudo R-squared	0.107	0.106	0.114	0.110	0.116	0.112	0.112
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the marginal effects calculated with all covariates set to their mean values; the associated robust standard errors are reported in brackets. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Table 5. Country- and CD Project-Specific Variables—Ordered Probit Model

Dependent variable: original outcome rating (1-4)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.0055	-0.0100	0.0034	-0.0185	-0.0057	0.0085	0.0141
	(0.0407)	(0.0397)	(0.0412)	(0.0424)	(0.0429)	(0.0430)	(0.0429)
Per capita real GDP growth	0.0402***	0.0391***	0.0406***	0.0387***	0.0391***	0.0400***	0.0395***
	(0.0085)	(0.0085)	(0.0086)	(0.0087)	(0.0088)	(0.0088)	(0.0088)
Volatility of per capita real GDP growth	-0.3230	-0.2638	-0.2015	-0.0490	-0.0020	-0.1310	-0.0829
	(0.6258)	(0.6328)	(0.6357)	(0.6601)	(0.6690)	(0.6812)	(0.6719)
FCS	-0.1023*		-0.1026*				
	(0.0604)		(0.0609)				
Small state	-0.1324*	-0.1265*		-0.1224*			
	(0.0678)	(0.0677)		(0.0677)			
IMF completed program	0.1475**	0.1373**	0.1272*	0.1581**	0.1352*	0.1557**	0.1491**
	(0.0675)	(0.0680)	(0.0682)	(0.0688)	(0.0693)	(0.0694)	(0.0690)
IMF off-track program	-0.1264	-0.1428	-0.1473	-0.0876	-0.1127	-0.1027	-0.1004
	(0.0921)	(0.0894)	(0.0938)	(0.0933)	(0.0951)	(0.0940)	(0.0946)
Completion share of CD project	0.0135***	0.0136***	0.0137***	0.0137***	0.0139***	0.0133***	0.0134***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
RTAC project	0.4552***	0.4505***	0.4384***	0.4609***	0.4458***	0.4504***	0.4419***
	(0.0756)	(0.0749)	(0.0762)	(0.0759)	(0.0766)	(0.0763)	(0.0762)
Resident advisor	0.1053***	0.1040***	0.1125***	0.1029***	0.1101***	0.0846**	0.0849**
	(0.0384)	(0.0384)	(0.0387)	(0.0383)	(0.0387)	(0.0389)	(0.0388)
Short-term expert	0.0864**	0.0804*	0.0891**	0.0838*	0.0889**	0.0265	0.0123
	(0.0421)	(0.0424)	(0.0430)	(0.0429)	(0.0438)	(0.0418)	(0.0411)
N. of workstreams per country	-0.3554***	-0.3406***	-0.3620***	-0.3564***	-0.3636***		
	(0.0734)	(0.0730)	(0.0746)	(0.0735)	(0.0746)		
N. of objectives per workstream						-0.0352	
						(0.0220)	
N. of outcomes per objective							-0.0150
							(0.0203)
Regulatory quality		0.1129**					
		(0.0555)					
Small state in Africa			-0.0621		-0.0677	-0.0488	-0.0434
			(0.0951)		(0.0958)	(0.0942)	(0.0945)
Small state in Asia-Pacific			0.2508*		0.2439*	0.2271*	0.2005
			(0.1393)		(0.1373)	(0.1377)	(0.1369)
Small state in Western Hemisphere			-0.4799***		-0.4584***	-0.4653***	-0.4679***
			(0.1255)		(0.1280)	(0.1284)	(0.1284)
Small state in Europe			-0.2295		-0.1342	-0.1273	-0.1366
			(0.2078)		(0.2171)	(0.2133)	(0.2147)
FCS in Africa				-0.1777**	-0.1579**	-0.1109	-0.1139
				(0.0721)	(0.0724)	(0.0719)	(0.0722)
FCS in Asia Pacific				-0.0576	-0.1006	-0.0747	-0.0615
				(0.1288)	(0.1269)	(0.1262)	(0.1265)
FCS in Middle East and Centr. Asia				-0.1655	-0.1575	-0.1232	-0.1484
				(0.1560)	(0.1577)	(0.1567)	(0.1573)
FCS in Western Hemisphere				0.5181**	0.3229	0.3511	0.4011
				(0.2579)	(0.2735)	(0.2762)	(0.2732)
FCS in Europe				0.1670	0.1694	0.2084	0.1768
				(0.2009)	(0.2094)	(0.2109)	(0.2108)
Observations	2,836	2,836	2,836	2,836	2,836	2,836	2,836
Pseudo R-squared	0.0698	0.0699	0.0728	0.0707	0.0734	0.0702	0.0698
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the estimated coefficients and, in brackets, the associated robust standard errors. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Table 6. Marginal Effects—Ordered Probit Model

Outcome ratings	Log of per capita real GDP	Per capita real GDP growth	Volatility of per capita real GDP growth	FCS	Small state	IMF completed program	IMF off-track program	Completion share of CD project	RTAC project	Resident advisor	Short-term expert	N. of workstreams per country
1- Not achieved	0.000939	-0.00685***	0.0550	0.0174*	0.0226*	-0.0251**	0.0215	-0.00231***	-0.0776***	-0.0179***	-0.0147**	0.0606***
	(0.00693)	(0.00147)	(0.107)	(0.0103)	(0.0115)	(0.0115)	(0.0157)	(0.000310)	(0.0132)	(0.00659)	(0.00722)	(0.0127)
2 - Partially achieved	0.00114	-0.00831***	0.0668	0.0212*	0.0274*	-0.0305**	0.0261	-0.00280***	-0.0941***	-0.0218***	-0.0179**	0.0735***
	(0.00841)	(0.00180)	(0.129)	(0.0126)	(0.0141)	(0.0141)	(0.0191)	(0.000397)	(0.0163)	(0.00798)	(0.00870)	(0.0155)
3 - Largely achieved	-0.00132	0.00960***	-0.0771	-0.0244*	-0.0316*	0.0352**	-0.0302	0.00323***	0.109***	0.0251***	0.0206**	-0.0849***
	(0.00971)	(0.00205)	(0.149)	(0.0145)	(0.0162)	(0.0161)	(0.0220)	(0.000442)	(0.0186)	(0.00921)	(0.0101)	(0.0177)
4 - Fully achieved	-0.000762	0.00556***	-0.0447	-0.0142*	-0.0183*	0.0204**	-0.0175	0.00187***	0.0629***	0.0146***	0.0120**	-0.0492***
	(0.00562)	(0.00120)	(0.0867)	(0.00836)	(0.00943)	(0.00940)	(0.0128)	(0.000256)	(0.0106)	(0.00533)	(0.00581)	(0.0104)

Notes: the table reports the marginal effects, calculated with all covariates set to their mean values, associated with the ordered probit estimates of Table 5, column 1; the associated robust standard errors are in brackets. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Annex

Table A1. Country- and CD Project-Specific Variables—Linear Probability Model with Clustered s.e.

Dependent variable: binary outcome rating (0-1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.00609	-0.00464	-0.00163	-0.0106	-0.00426	0.000138	0.00296
	(0.0220)	(0.0242)	(0.0224)	(0.0226)	(0.0233)	(0.0241)	(0.0243)
Per capita real GDP growth	0.0148***	0.0145***	0.0153***	0.0140***	0.0146***	0.0150***	0.0148***
	(0.00474)	(0.00467)	(0.00479)	(0.00512)	(0.00511)	(0.00524)	(0.00517)
Volatility of per capita real GDP growth	-0.316	-0.332	-0.270	-0.196	-0.193	-0.240	-0.221
	(0.311)	(0.297)	(0.353)	(0.348)	(0.378)	(0.411)	(0.387)
FCS	-0.0662**		-0.0659**				
	(0.0322)		(0.0301)				
Small state	-0.0490	-0.0456		-0.0418			
	(0.0391)	(0.0394)		(0.0386)			
IMF completed program	0.0355	0.0313	0.0260	0.0410	0.0306	0.0383	0.0354
	(0.0279)	(0.0303)	(0.0287)	(0.0284)	(0.0290)	(0.0304)	(0.0307)
IMF off-track program	-0.0710	-0.0845	-0.0786	-0.0539	-0.0635	-0.0603	-0.0600
	(0.0662)	(0.0682)	(0.0654)	(0.0662)	(0.0660)	(0.0615)	(0.0639)
Completion share of CD project	0.00365***	0.00375***	0.00377***	0.00370***	0.00380***	0.00359***	0.00363***
	(0.00110)	(0.00109)	(0.00109)	(0.00111)	(0.00110)	(0.00112)	(0.00111)
RTAC project	0.0923**	0.0885*	0.0813*	0.0962**	0.0857*	0.0885*	0.0859*
	(0.0461)	(0.0457)	(0.0470)	(0.0464)	(0.0472)	(0.0483)	(0.0482)
Resident advisor	0.0368*	0.0356	0.0392*	0.0367*	0.0392*	0.0306	0.0313
	(0.0216)	(0.0218)	(0.0207)	(0.0208)	(0.0200)	(0.0203)	(0.0200)
Short-term expert	-0.00559	-0.00926	-0.00320	-0.00617	-0.00258	-0.0224	-0.0274
	(0.0200)	(0.0207)	(0.0198)	(0.0205)	(0.0202)	(0.0184)	(0.0194)
N. of workstreams per country	-0.118***	-0.108***	-0.122***	-0.118***	-0.122***		
	(0.0399)	(0.0409)	(0.0388)	(0.0396)	(0.0386)		
N. of objectives per workstream						-0.0145	
						(0.0105)	
N. of outcomes per objective							-0.0131
							(0.0112)
Regulatory quality		0.0565*					
		(0.0328)					
Small state in Africa			-0.0329		-0.0350	-0.0289	-0.0256
			(0.0585)		(0.0586)	(0.0580)	(0.0598)
Small state in Asia-Pacific			0.151**		0.150**	0.146**	0.135*
			(0.0681)		(0.0701)	(0.0712)	(0.0712)
Small state in Western Hemisphere			-0.194***		-0.180***	-0.182***	-0.182***
			(0.0588)		(0.0581)	(0.0580)	(0.0595)
Small state in Europe			-0.133***		-0.0859	-0.0829	-0.0848
			(0.0440)		(0.0545)	(0.0520)	(0.0549)
FCS in Africa				-0.0975**	-0.0877**	-0.0720*	-0.0717
				(0.0421)	(0.0419)	(0.0432)	(0.0445)
FCS in Asia Pacific				-0.0726	-0.0905*	-0.0821*	-0.0729
				(0.0655)	(0.0512)	(0.0485)	(0.0489)
FCS in Middle East and Centr. Asia				-0.0959	-0.0878	-0.0754	-0.0907
				(0.0825)	(0.0810)	(0.0896)	(0.0855)
FCS in Western Hemisphere				0.340***	0.260***	0.266***	0.288***
				(0.0526)	(0.0627)	(0.0675)	(0.0667)
FCS in Europe				0.0726	0.0658	0.0815	0.0683
				(0.0588)	(0.0651)	(0.0563)	(0.0629)
Observations	2,836	2,836	2,836	2,836	2,836	2,836	2,836
R-squared	0.134	0.134	0.142	0.138	0.144	0.140	0.140
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the estimated coefficients and, in brackets, the associated standard errors clustered at the country level. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Table A2. Country- and CD Project-Specific Variables—Probit Model with Clustered s.e.

Dependent variable: binary outcome rating (0-1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.0102	-0.0081	-0.0066	-0.0157	-0.0096	-0.0038	-0.0007
	(0.0234)	(0.0270)	(0.0237)	(0.0240)	(0.0245)	(0.0254)	(0.0256)
Per capita real GDP growth	0.0166***	0.0165***	0.0174***	0.0157***	0.0166***	0.0172***	0.0169***
	(0.0055)	(0.0054)	(0.0056)	(0.0059)	(0.0059)	(0.0060)	(0.0059)
Volatility of per capita real GDP growth	-0.3056	-0.3531	-0.2688	-0.1671	-0.1861	-0.2594	-0.2286
	(0.3460)	(0.3310)	(0.3846)	(0.3866)	(0.4131)	(0.4396)	(0.4153)
FCS	-0.0799**		-0.0815**				
	(0.0362)		(0.0335)				
Small state	-0.0502	-0.0458		-0.0419			
	(0.0431)	(0.0439)		(0.0428)			
IMF completed program	0.0392	0.0342	0.0292	0.0470	0.0360	0.0434	0.0408
	(0.0294)	(0.0319)	(0.0300)	(0.0301)	(0.0303)	(0.0314)	(0.0320)
IMF off-track program	-0.0787	-0.0943	-0.0858	-0.0575	-0.0677	-0.0643	-0.0625
	(0.0779)	(0.0809)	(0.0762)	(0.0771)	(0.0761)	(0.0704)	(0.0731)
Completion share of CD project	0.0041***	0.0043***	0.0043***	0.0043***	0.0044***	0.0042***	0.0042***
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
RTAC project	0.1078**	0.1047**	0.0935*	0.1122**	0.0977*	0.1026*	0.0981*
	(0.0512)	(0.0512)	(0.0528)	(0.0515)	(0.0531)	(0.0542)	(0.0540)
Resident advisor	0.0396*	0.0383	0.0418*	0.0379*	0.0405*	0.0319	0.0322
	(0.0237)	(0.0238)	(0.0228)	(0.0228)	(0.0221)	(0.0221)	(0.0218)
Short-term expert	-0.0077	-0.0101	-0.0082	-0.0081	-0.0076	-0.0289	-0.0363
	(0.0248)	(0.0259)	(0.0247)	(0.0250)	(0.0251)	(0.0233)	(0.0243)
N. of workstreams per country	-0.1329***	-0.1208***	-0.1397***	-0.1332***	-0.1401***		
	(0.0440)	(0.0452)	(0.0427)	(0.0437)	(0.0423)		
N. of objectives per workstream						-0.0178	
						(0.0113)	
N. of outcomes per objective							-0.0158
							(0.0124)
Regulatory quality		0.0639*					
		(0.0366)					
Small state in Africa			-0.0389		-0.0412	-0.0352	-0.0312
			(0.0653)		(0.0660)	(0.0662)	(0.0681)
Small state in Asia-Pacific			0.1744**		0.1776**	0.1713**	0.1596**
			(0.0711)		(0.0740)	(0.0753)	(0.0754)
Small state in Western Hemisphere			-0.1882***		-0.1724***	-0.1736***	-0.1725***
			(0.0608)		(0.0599)	(0.0596)	(0.0615)
Small state in Europe			-0.2179***		-0.1519*	-0.1478*	-0.1420*
			(0.0644)		(0.0804)	(0.0783)	(0.0796)
FCS in Africa				-0.1180**	-0.1069**	-0.0853*	-0.0861*
				(0.0464)	(0.0459)	(0.0466)	(0.0481)
FCS in Asia Pacific				-0.0870	-0.1157**	-0.1038**	-0.0954*
				(0.0729)	(0.0539)	(0.0508)	(0.0509)
FCS in Middle East and Centr. Asia				-0.1052	-0.0964	-0.0813	-0.1005
				(0.0906)	(0.0882)	(0.0969)	(0.0930)
FCS in Western Hemisphere				0.3676***	0.2887***	0.2990***	0.3264***
				(0.0571)	(0.0673)	(0.0740)	(0.0732)
FCS in Europe				0.1082	0.0939	0.1150	0.0963
				(0.0875)	(0.0932)	(0.0811)	(0.0886)
Observations	2,829	2,829	2,829	2,829	2,829	2,829	2,829
Pseudo R-squared	0.107	0.106	0.114	0.110	0.116	0.112	0.112
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the marginal effects calculated with all covariates set to their mean values; the associated standard errors clustered at the country level are reported in brackets. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.

Table A3. Country- and CD Project-Specific Variables—Ordered Probit Model with Clustered s.e.

Dependent variable: original outcome rating (1-4)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita real GDP	-0.0055	-0.0100	0.0034	-0.0185	-0.0057	0.0085	0.0141
	(0.0527)	(0.0573)	(0.0551)	(0.0552)	(0.0576)	(0.0595)	(0.0593)
Per capita real GDP growth	0.0402***	0.0391***	0.0406***	0.0387***	0.0391***	0.0400***	0.0395***
	(0.0108)	(0.0114)	(0.0109)	(0.0114)	(0.0114)	(0.0116)	(0.0115)
Volatility of per capita real GDP growth	-0.3230	-0.2638	-0.2015	-0.0490	-0.0020	-0.1310	-0.0829
	(0.6964)	(0.7082)	(0.7750)	(0.7978)	(0.8553)	(0.9362)	(0.8798)
FCS	-0.1023		-0.1026				
	(0.0703)		(0.0697)				
Small state	-0.1324	-0.1265		-0.1224			
	(0.0986)	(0.0997)		(0.0985)			
IMF completed program	0.1475**	0.1373**	0.1272*	0.1581**	0.1352**	0.1557**	0.1491**
	(0.0632)	(0.0681)	(0.0675)	(0.0628)	(0.0669)	(0.0701)	(0.0692)
IMF off-track program	-0.1264	-0.1428	-0.1473	-0.0876	-0.1127	-0.1027	-0.1004
	(0.1520)	(0.1553)	(0.1539)	(0.1509)	(0.1549)	(0.1397)	(0.1436)
Completion share of CD project	0.0135***	0.0136***	0.0137***	0.0137***	0.0139***	0.0133***	0.0134***
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0029)	(0.0029)
RTAC project	0.4552***	0.4505***	0.4384***	0.4609***	0.4458***	0.4504***	0.4419***
	(0.1172)	(0.1158)	(0.1197)	(0.1169)	(0.1190)	(0.1198)	(0.1200)
Resident advisor	0.1053**	0.1040**	0.1125**	0.1029**	0.1101**	0.0846	0.0849
	(0.0528)	(0.0524)	(0.0523)	(0.0521)	(0.0517)	(0.0540)	(0.0528)
Short-term expert	0.0864	0.0804	0.0891*	0.0838	0.0889	0.0265	0.0123
	(0.0531)	(0.0531)	(0.0536)	(0.0546)	(0.0548)	(0.0484)	(0.0510)
N. of workstreams per country	-0.3554***	-0.3406***	-0.3620***	-0.3564***	-0.3636***		
	(0.0991)	(0.0997)	(0.1005)	(0.0991)	(0.1006)		
N. of objectives per workstream						-0.0352	
						(0.0264)	
N. of outcomes per objective							-0.0150
							(0.0278)
Regulatory quality		0.1129					
		(0.0832)					
Small state in Africa			-0.0621		-0.0677	-0.0488	-0.0434
			(0.1437)		(0.1448)	(0.1438)	(0.1470)
Small state in Asia-Pacific			0.2508		0.2439	0.2271	0.2005
			(0.1531)		(0.1503)	(0.1599)	(0.1600)
Small state in Western Hemisphere			-0.4799***		-0.4584***	-0.4653***	-0.4679***
			(0.1493)		(0.1516)	(0.1520)	(0.1554)
Small state in Europe			-0.2295		-0.1342	-0.1273	-0.1366
			(0.1648)		(0.2274)	(0.2200)	(0.2251)
FCS in Africa				-0.1777**	-0.1579*	-0.1109	-0.1139
				(0.0892)	(0.0914)	(0.0962)	(0.0965)
FCS in Asia Pacific				-0.0576	-0.1006	-0.0747	-0.0615
				(0.1339)	(0.1161)	(0.1229)	(0.1233)
FCS in Middle East and Centr. Asia				-0.1655	-0.1575	-0.1232	-0.1484
				(0.2037)	(0.2034)	(0.2266)	(0.2200)
FCS in Western Hemisphere				0.5181***	0.3229**	0.3511**	0.4011**
				(0.1166)	(0.1515)	(0.1689)	(0.1640)
FCS in Europe				0.1670	0.1694	0.2084	0.1768
				(0.2088)	(0.2407)	(0.2223)	(0.2308)
Observations	2,836	2,836	2,836	2,836	2,836	2,836	2,836
Pseudo R-squared	0.0698	0.0699	0.0728	0.0707	0.0734	0.0702	0.0698
Project-start year dummies	Y	Y	Y	Y	Y	Y	Y
Project-change year dummies	Y	Y	Y	Y	Y	Y	Y
Workstream dummies	Y	Y	Y	Y	Y	Y	Y
Macro-region dummies	Y	Y	Y	Y	Y	Y	Y
Income group dummies	Y	Y	Y	Y	Y	Y	Y

Notes: the table reports the estimated coefficients and, in brackets, the associated standard errors clustered at the country level. ***, **, and * indicate significance at 1, 5, and 10 percent, respectively.