IDB WORKING PAPER SERIES Nº IDB-WP-787

Organizational Efficiency or Bureaucratic Quagmire:

Do Quality at Entry Assessments Improve Project Performance?

Leonardo R. Corral
Nancy McCarthy

Inter-American Development Bank
Office of Strategic Planning and Development Effectiveness

March 2017

Organizational Efficiency or Bureaucratic Quagmire:

Do Quality at Entry Assessments Improve Project Performance?

Leonardo R. Corral
Nancy McCarthy

Inter-American Development Bank
Lead Analytics

March 2017

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
Corral, Leonardo.
Organizational efficiency or bureaucratic quagmire: do quality at entry assessments improve project performance? / Leonardo R. Corral, Nancy McCarthy.
p. cm. — (IDB Working Paper Series ; 787)
Includes bibliographic references.
1. Economic development projects-Latin America-Evaluation. 2. Economic assistance-Latin America-Evaluation. I. McCarthy, Nancy. II. Inter-American Development Bank. Strategy Development Division. II. Title. III. Series.
IDB-WP-787

OPTIONAL: Type address for correspo

Leonardo Corral Leonardoc@iadb.org

Organizational Efficiency or Bureaucratic Quagmire: Do Quality at Entry Assessments Improve Project Performance?[1]

Leonardo R. Corral [2], Nancy McCarthy [3]

This version: March 2017

Abstract

Do quality at entry assessments enhance the delivery of development projects? In this paper we take advantage of approval and execution systems in place at the Inter-American Development Bank (IDB) to examine whether projects that have higher quality at entry -- captured through grading scores provided by a checklist -- perform better in terms of project implementation performance indicators. Implementation indicators include measures based on actual versus planned schedule of activities and cost outlays, as well as percent of loan disbursed. The analysis suggests higher scores on project logic and economic analyses at entry have had a positive impact on project performance. However, monitoring and evaluation and impact assessment scores had limited impacts on performance. The evidence supports the hypothesis that the use of a checklist can be an effective framework for assessing quality at entry for IDB projects, though there is scope to improve the checklist for certain indicators.

JEL Classification: O1; O12; O19; O22; H43

Keywords: Development Effectiveness; Quality at Entry; Project Performance

[1] The authors would like to thank Shakirah Cossens, Diego Cortes and Ange Liu for supporting data collection efforts, Oscar Mitnik for helpful comments, and Josh Brubaker for excellent research assistance. We are grateful for the support of the Office of Strategic Planning and Development Effectiveness of the Inter-American Development Bank in conducting the study, and to participants of the SPD Half-Baked Lunch seminar series for their helpful comments.

[2] Strategy Development Division, Inter-American Development Bank, Washington, DC, Email: leonardoc@iadb.org

[3] Lead Analytics, Washington, DC, Email: nmccarthy@leadanalyticsinc.com.

I INTRODUCTION

During the last decade, the concept of development effectiveness has progressively gained strategic relevance within development institutions. It encompasses at least three broad concerns: (i) the need to align available resources with country priorities, including meeting the Millennium Development Goals and the Sustainable Development Goals; (ii) improving the impact of interventions in line with a results-based logic fully integrated into the management cycle; and (iii) the need to legitimate the use of resources for development policies, accounting for the results achieved. In this sense, the movement towards development effectiveness strengthens and complements the evaluation agenda that has always accompanied the efforts of development institutions.

This movement has materialized into an important agenda of the international community, of which a major milestone is the 2005 Paris Declaration on Aid Effectiveness, endorsed by the IDB. For the first time, it goes beyond a mere declaration of intentions, since it provides 'a practical, action-oriented roadmap with specific targets to be met by 2010', in seven areas: country ownership, alignment of agendas, harmonization, management for results and mutual accountability (see http://www.oecd.org/dac/effectiveness/34428351.pdf). At the same time, the harmonization efforts carried out by the Multilateral Development Banks (MDB) have been noteworthy. Through the Evaluation Cooperation Group (ECG), MDBs have formulated and agreed upon common standards and good practices for evaluation. Finally, the MDBs have also set in motion several initiatives to support management for results, such as the Common Performance Assessment System (COMPAS) (see http://www.mfdr.org/COMPAS/), which establishes a common framework for MDBs to report results.

At the IDB, these efforts were captured in 2008 with the approval of the Development Effectiveness Framework (DEF). The DEF establishes a set of standards and instruments to ensure accountability for results of all IDB interventions. These instruments apply the MDB's Evaluation Cooperation Group (MDB-ECG) Good Practice Standards for Development Evaluation (GPS) to the Bank's loan and grant development operations. As described more fully below, DEF instruments cover all phases of the project lifecycle, including quality at entry, yearly project implementation performance assessments, and project completion reviews.

The increased emphasis on development effectiveness and the development of frameworks and operational instruments for documenting project development, implementation and completion has generally increased projects teams' costs in developing a successful project proposal, and to a lesser extent, in documenting implementation and project completion. In this paper, our goal is to examine whether the new process leads to measurable benefits. Specifically, we evaluate whether project's that have higher quality at entry scores perform better in terms of project implementation indicators than those projects with lower quality at entry scores. Indicators include percent of loan disbursed, as well as performance measures based on actual versus planned schedule of activities and cost outlays.

There are a few articles that have also looked at various measures of quality at entry and project performance. For instance, Jenkins (1997) describes results from Belli and Pritchett (1995), who found that the probability that a project would perform unsatisfactorily during

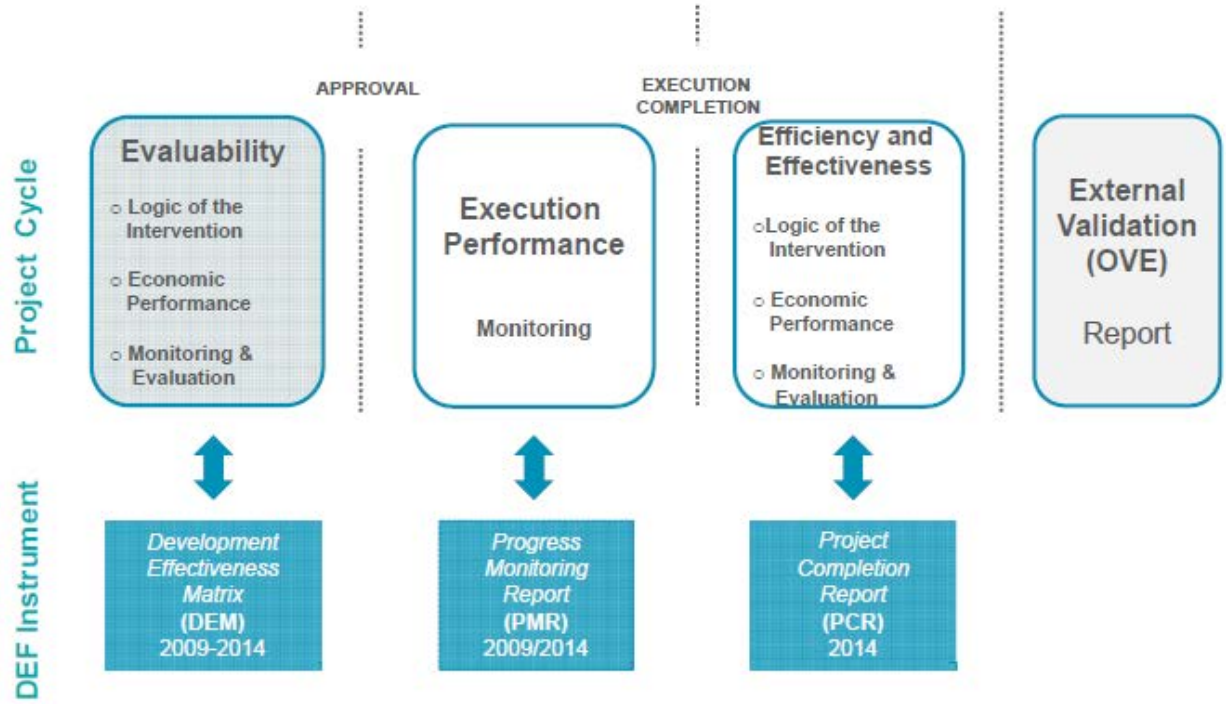
implementation was significantly higher for projects with lower ratings on economic appraisals undertaken at project inception. Similarly, Vawda et al. (2001) find that higher quality economic appraisals at project inception led to a greater probability that World Bank education projects would receive favorable ratings on project performance. Legovini et al. (2012) investigate the effect that project impact evaluations have on World Bank project performance, and find a significant positive impact of higher quality impact evaluations. The impact evaluations combined elements of monitoring and evaluation by integrating impact assessment data collection efforts into project implementation. Raimondo (2016) study looks specifically at whether monitoring and evaluation quality has an impact on World Bank projects, and again finds that it does.

This study contributes to the nascent literature on the impacts of improved quality at entry on project performance in two main ways. First, this is the first study we know of that does not use data on World Bank projects. We believe that it is essential to generate findings from a wide range of institutions in order to identify generalizable lessons. Second, we do not restrict our attention to one dimension of quality at entry as do the previously cited papers. Instead, we include quality at entry scores covering four dimensions: project logic, economic analysis, monitoring, and impact assessment. Our results indicate that higher project logic and economic analysis have positive impacts on IDB project performance. Some of our evidence suggests that higher monitoring scores actually lead to poorer project performance, though the impact is often not significant, as is the case with impact assessment.

II BACKROUND ON IDB's QUALITY AT ENTRY ASSESSMENT PROCESS

The approval of the DEF led to the development and implementation of specific standards and instruments to enhance the objectivity of the quality control functions of the IDB across the project cycle. As per figure 1 below, these included: (i) the Development Effectiveness Matrix (the "DEM") to ensure ex-ante minimum evaluability standards of all of the Bank's interventions; (ii) the reform of the Bank's monitoring and reporting systems (Progress Monitoring Report) to focus on outputs and outcomes and their associated costs; and, (iii) the improvement of the Project Completion Report (PCR), the record of an operation's performance at the end of its execution phase, as the Bank's main instrument for documenting concrete results to its shareholders and disseminating the lessons of a project's experience, focused on accountability and learning. These enhancements where accompanied at each stage by an increased focus on building-up evaluation and monitoring capacity in the Bank and the Region.

Figure 1: Development Effectiveness Framework DEF



Quality at Entry: Development Effectiveness Matrix (DEM)

The DEF requires a minimum evaluability at entry of all IDBs operations. At the IDB, in line with the OECD-DAC definition, evaluability is defined as the "extent to which an activity or a project can be evaluated in a reliable and credible fashion". Checklists are increasingly being used to assess the ex-ante and ex-post quality of public sector interventions. Examples include the "Program Assessment Rating Tool" (PART) of the Office of Management and Budget of the US Government and Mexico's *"Modelo de Términos de Referencia para la Evaluación de Consistencia y Resultados"* developed by the *Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL)*. It is argued that when the standards for a "yes" are clearly specified in a checklist, with criteria based on observable objective characteristics of a project, such assessment tools prove useful in reducing "evaluator bias" and ensuring minimum standards in each of the areas that the evaluator needs to assess. A standardized checklist also serves as a record of the degree to which the project team has been responsive and able to address evaluability issues raised in the project preparation phase, hence providing the elements for a transparent, structured and well documented quality control function.

The IDB has produced Evaluability Standards, which are complemented with templates to support project teams in the preparation of results matrices, economic analysis, and monitoring and evaluation plans. The Standards identify the main dimensions of evaluability, as well as the questions that a loan proposal should answer, and the general content requirements in order to answer these questions

In order to ascertain whether an operation can be rigorously evaluated, the Development Effectiveness Matrix (DEM) was developed as a checklist that adopts all the elements of the "Good Practice Standards for Public Sector Operations", proposed by the Evaluation Cooperation Group. The DEM requires that all projects demonstrate they were designed based on sound theoretical and empirical evidence, in terms of stating the development problem, outlining the causes, and demonstrating the effectiveness of the proposed solutions. This validation involves assessing whether the project diagnosis is empirically sustained, the evidence regarding the proposed solution is sufficient, the intervention is dimensioned appropriately and in accordance with the objectives of the project, the necessary inputs are likely to be provided in a timely manner, the process that transforms inputs into outputs is in place, and that these outputs will be delivered within time and budget. Furthermore, it involves assessing whether the planned outputs are likely to produce the desired outcomes and whether rigorously applying evaluation methods it is probable to attribute these outcomes to the project, thus producing a body of knowledge that is relevant for future project design, execution and policy formulation. Assessing these evaluability dimensions at entry should provide the necessary elements to ensure that these operations will be able to demonstrate their results during implementation and at project completion.

As per figure 2 below, three main evaluability dimensions are assessed and validated through the DEM: (1) Project Logic; (2) Economic Analysis; and (3) Monitoring and Evaluation. Each section sums to ten and has equal weight in the final evaluability score.

Project Logic: This section has 26 checklist questions in three subsections: Program Diagnosis, Proposed Interventions or Solutions, and Results Matrix Quality. The purpose of this section is to analyze whether Project Documentation includes a proper diagnosis, including identification of the development problems to be addressed, their root-causes, and how the intervention will address them. It also assesses whether enough evidence of the effectiveness of the proposed intervention is provided and that a clear vertical logic (i.e. how inputs and activities will be translated into outputs, which will lead to outcomes, which will contribute to final impact) is articulated.

The Results Matrix sub-section requires the definition of the desired medium or long term impacts the project is expected to contribute to, the desired improvements/effects as a result of the project (outcomes), and the project deliverables (outputs), which the project is contractually accountable to provide. For each level, SMART (Specific, Measurable, Attributable/Achievable, Relevant and Time-Bound) indicators must be established, with appropriate baseline values, realistic and relevant targets (preferably anchored by and consistent with the ex-ante economic analysis), and defined sources of data or a clear collection plan as the means of verification. See Annex 1 for the complete list of yes/no questions.

Economic Analysis: The primary objective of conducting an economic analysis is to help design projects that will be effective in promoting development in a country. Economic analysis is a key element in understanding the expected results of an intervention and/or the rationale for choosing one intervention over another. The usefulness of this exercise is greatest when it is done early in the project cycle and contributes to the decisions about whether and how to

proceed with a project. While the outcome of an economic analysis—net present value or economic rate of return for Cost Benefit Analysis (CBA) and cost per unit effect for Cost Effectiveness Analysis (CEA)—are important, it is the process of conducting the analysis and the insights that it provides that are most likely to be useful in designing a better project.

This section of the DEM, in addition to creating incentives for economic analysis to be carried out, aims to measure the adequacy and quality of the analysis performed. The DEM assigns a zero for no economic analysis, a 4 for CBA or CEA having been carried out on the main components of the projects, and from 5 to 10 depending on the quality of the analysis performed and whether the rate is positive and robust to changes in underlying assumptions. The key quality ingredients of an economic analysis are noted in the DEM. Specifically, benefits streams must be identified and valued at the correct price, all resource costs must be identified and calculated, all assumptions specified and supported by evidence, and sensitivity analysis conducted around key parameters where viability hinges. When formal CBA or CEA are not feasible, teams should provide a justification and a thorough discussion of economic costs and benefits. Annex 1 includes the expected standards for ex-ante economic analysis.

Monitoring and Evaluation: The DEM provides elements for projects to include rigorous systems to track implementation and measures results. The DEM recognizes the variety of methodologies that exist to construct the counterfactual project scenario, some of which apply only to some types of projects. While some evaluation methodologies are technically superior to others (internally valid), and thus produce more credible estimates of the impacts attributable to a project, it is not practical or possible to apply such methodologies to every project. However, experience demonstrates that nearly all projects have the potential to include some form of impact evaluation. Furthermore, even though projects with more rigorous evaluation designs will get higher scores in the specific subsection, the DEM is structured in a way that all projects, as long as they have an adequate evaluation plan and conduct sound economic analysis, may achieve a rating that is satisfactory. The DEM includes 16 questions in this section regarding the necessary elements for projects to include systems to track implementation and measure results, such as definition of monitoring and evaluation plans, budget for the activities, and methods. Annex 1 presents the pertinent questions for this dimension.

The IDB has produced guidelines and templates, as well as training workshops, to support project teams in the preparation of results matrices, economic analysis, and monitoring and evaluation plans. This has allowed for the harmonization and common understanding of standards, definitions and expectations across all areas of the Bank.

The DEM produces an evaluability rating that is based on the information provided by teams related to the evaluability dimensions described above. Project teams are required to include the justification for the substantive questions in the DEM, in order to have precise elements to justify the yes/no answer and the validation.

The DEM is validated by an assigned reviewer in the Office of Strategic Planning and Development Effectiveness (SPD) who, based on the standards and sectoral expertise, makes a judgment on whether the dimensions are adequately met.

The specific steps and integration of the DEM into the IDB's quality review process is described below.

1. A member of SPD (i.e. SPD Reviewer) is assigned to each project as soon as project preparation is initiated. The Reviewer follows the project preparation cycle. The Reviewer is not a member of the project team (in cases where an SPD staff is part of the project team, she will be different from the SPD Reviewer assigned to validate the evaluability of the project proposal). The Reviewer coordinates the SPDs processes for DEM validation and acts as liaison between SPD and the project team.

2. SPD provides written comments to the Project Profile (PP) that address, if necessary, potential evaluability concerns at this initial conceptual level, which might be related to diagnosis, project logic and definition of objectives. Early discussions and recommendations on M&E arrangements and economic analysis may also be included in the comments.

3. If deemed necessary SPD can make staff time available to provide direct support to teams in the design process from Project Profile (PP) to Proposal for Operation Development (POD) to ensure that projects meet the evaluability thresholds established by the IDB Governors[4].

4. The SPD Reviewer provides written comments to the Quality and Risk Review (QRR) in the form of an Evaluability Assessment Note (EAN) that includes a qualitative analysis of how well the project meets the evaluability standards set out in the DEM. The EAN and the preliminary DEM are sent by SPD to the QRR and provide an overall assessment of the project's evaluability at this stage of project preparation. The note elaborates on the relevant evaluability dimensions as it pertains to the project under review, with a section of specific recommendations to the project team on how to improve evaluability. The project team prepares a response to comments received at the QRR meetings reflecting and addressing the evaluability issues and recommendations formulated (in addition to other issues raised in the review process by other Departments, i.e. Legal, Safeguards, etc.). Once the project team amends the design and incorporates the necessary changes to project documentation to address all issues, including evaluability, raised at

[4] In the *Report on the General Increase in Resources of the Inter-American Development Bank*, paragraph 4.7, May 2010, the Governors of the IDB established that:*"(i) all Sovereign Guarantee and Non SG projects must be rated for evaluability; (ii) the evaluability score includes only the dimensions of evaluability of the DEM; (iii) the Office of Strategic Planning and Development Effectiveness (SPD) will support teams in meeting evaluability standards from project profile to project proposal, and will validate the final evaluability score for Operation and Policy Committee (OPC) consideration;: (iv) a minimum evaluability threshold of 5 will be required for all operations to be submitted to the Board of Executive Directors; (v) in cases of extraordinary humanitarian or financial crisis (acute payment pressures), and upon the presentation of a properly justified request by the President, the Board of Executive Directors may waive the required threshold prior to approving an operation, so long as the request is accompanied by a timetable giving the shortest possible time period, not to exceed 90 days, for achieving the said threshold."*

QRR, the Draft Loan Proposal is finalized and can proceed to the Operations and Policy Committee (OPC).

5. The SPD reviewer does a final validation of the evaluability score prior to OPC consideration. If the project team leader and the Vice Presidency of Sectors (VPS) managerial team agree with the preliminary DEM score assigned by SPD at the QRR stage and the minutes of the QRR do not record any commitments to make changes related to evaluability, then the POD will not require an additional review by SPD for re-validation.

6. After all changes are made and the final POD's evaluability at entry is scored via the DEM, SPD will update the EAN and make a recommendation to the Executive Vice President (EVP), who chairs the OPC, as to the project meeting minimum evaluability standards. If the DEM evaluability score is below 5, the threshold established by the Board of Governors in document AB-2764, Management will have two possible courses of action: (i) revise the design of the project, adjust the POD and resubmit to SPD for revalidation or (ii) proceed as indicated by the Governors in AB-2764, "*in cases of extraordinary humanitarian or financial crisis (acute payment pressures), and upon the presentation of a properly justified request by the President, the Board of Executive Directors may waive the required threshold prior to approving an operation, so long as the request is accompanied by a timetable giving the shortest possible time period, not to exceed 90 days, for achieving the said threshold"* (See Footnote 1)

7. By the end of the quality review cycle, the final DEM worksheet has been completed by project teams and validated by the SPD Reviewer and, for the most relevant substantial dimensions it includes a justification from the project team and annotations by the SPD Reviewer in case of disagreement. Thus, the worksheet documents the evaluability enhancement process and the degree the project team has been responsive and able to address evaluability issues raised in the project preparation phase, hence providing the elements for a transparent, structured and well documented quality control function.

8. The DEM Summary Table included within the Loan Proposal provides the Board of Executive Directors with a summary matrix of the DEM with SPD's validated scores. It also provides a succinct description of the characteristics of the Loan Proposal that lead to the assigned scores for each of the three main evaluability areas of the Matrix, namely: (i) Program Logic; (ii) Economic Analysis; and (iii) Monitoring and Evaluation.

The objective of this paper is to assess whether higher DEM scores, which are meant to capture higher ex-ante evaluability, are likely to lead to improved performance during project execution. We also assess what type of ex-ante evaluability criterion (project logic, evidence of effectiveness of the intervention, results framework quality, economic analysis) is associated with improved performance.

The outcome performance indicators of interest come from the Bank's monitoring and reporting systems (Progress Monitoring Report) that focuses on outputs, outcomes and related costs. These indicators are described in detail below and vary depending on the project's current stage

of execution: First Stage: after approval and before eligibility; Second Stage: between eligibility and 95% disbursement; and Third Stage: between 95% disbursement and project closure. For the purposes of our analysis, we will primarily focus on the project's performance between eligibility and 95% disbursement.

The first indicator, relevant for our purposes, is the proportion disbursed at the moment the PMR report is prepared in March of each year. This indicator is simply the sum of disbursements until the date of the report divided by the current approved amount. It is easy to obtain and has no measurement errors. The Schedule Performance Index (SPI) and the Cost Performance Index (CPI) use the earn value methodology and therefore compare the actual Earn Value (EV) with the value planned at start up (PV) and the accumulated cost until the date of the report (AV) and planned cost at start up (PC), respectively. The SPI(a) and CPI(a) are analogous but instead of using the plan at start-up they use annual plans that can be reviewed at the beginning of each year. The Planned Value (PV) is the budgeted cost for all the works scheduled to be done until the date of the report; i.e., it is the sum of the planned value of all the outputs until the period of the report. The Planned Value curve includes both IDB financing and counterpart funds. Earned Value (EV) is the percentage of the total budget actually completed at each point in time. EV is calculated by multiplying the budget for an activity/output by the percent progress for that activity. A more detailed description of the creation of these indicators is provided in IDB Document OP-1072-1 (Nov. 2013).

Once the project reaches eligibility, and after annual progress has been entered, performance and classification for the second stage is based on the indicators presented above: (i) Proportion disbursed with respect to the country benchmark, (ii) SPI, (ii) CPI, (iii) SPI(a), and (iv) CPI(a). Each one of these indicators will be assigned a Red, Yellow, or Green "traffic light" according to the status in the particular dimension. To provide a homogeneous scale that allows constructing a synthetic indicator of projects status, the platform gives values of 1, 2, and 3 to Red, Yellow, and Green according the following rules:

Indicator	Red (Indicator = 1)	Yellow (Indicator= 2)	Green (Indicator = 3)
I-Percent Disbursed	Lower than one standard deviation below the country historic disbursement curve	Between half standard deviation and one standard deviation below the country	Half standard deviation below the country historic disbursement curve or higher
I-SPI	Lower than 0.4 or higher than 2	Between 0.4 (inclusive) and 0.8	Between 0.8 (inclusive) and 2 (inclusive)
I-CPI	Lower than 0.4 or higher than 2	Between 0.4 (inclusive) and 0.8	Between 0.8 (inclusive) and 2 (inclusive)
I-SPI(a)	Lower than 0.4 or higher than 2	Between 0.4 (inclusive) and 0.8	Between 0.8 (inclusive) and 2 (inclusive)
I-CPI(a)	Lower than 0.4 or higher than 2	Between 0.4 (inclusive) and 0.8	Between 0.8 (inclusive) and 2 (inclusive)

A Synthetic Indicator (SI) summarizes the five indicators above, as a weighted average. Weights for the synthetic indicator were selected such that those indicators that reflect the history of the project receive 50% of the weight and those indicators with annual information receive the remaining 50%. Therefore, weights are given by:

Indicator	Weights
Proportion disbursed with respect to the country benchmark	0.20
SPI	0.15
CPI	0.15
SPI(a)	0.25
CPI(a)	0.25

Finally, projects are classified using the following rule:

Problem: $SI < 2$

Alert: $2 \leq SI \leq 2.5$

Satisfactory: $SI \geq 2.5$

For the purposes of our analysis some of these indicators need to be transformed to be used as our dependent variables, and this is described in detail in the empirical approach section.

The rest of the paper is organized as follows: in the next section, we present our empirical approach. We then discuss our choice of dependent and explanatory variables, followed by our empirical strategy, descriptive statistics, and present results. The paper closes with a discussion of results and concluding remarks.

II EMPIRICAL ANALYSIS

2.1 Dependent variables

As detailed above, the cost performance index (CPI) and schedule performance index (SPI) are based on the ratio of actual cost expenditures and schedule of activities to planned cost expenditures and schedule of activities. A ratio of 1 then indicates that actual cost and schedule performance are perfectly aligned to planned performance. Data is collected on planned cost and schedule performance at the outset of the project, as well as on cost and schedule targets that are adjusted annually. Because there are more observations available for the annually adjusted targets, we use the CPI(a) and SPI(a) indicators in our analysis.

The ratios are also categorized according to a "traffic light" system to provide a simple summary of project status, following the "red", "yellow" and "green" categories given above[5]. Since relatively few programs fall into the alert category, we run probits for whether or not a project was considered satisfactory versus alert or problem according to their cost and schedule performance scores.

In addition to the probits, we also construct indicators of performance using the ratios. We cannot use the ratios themselves, as both high and low ratios indicate potential problems in project implementation. In particular, low cost and schedule performance ratios indicate situations where projects are falling behind, whereas high ratios indicate that the project is over-spending relative to its targets or undertaking more actions than anticipated at the beginning of the year. Thus, both high and low values can indicate a problem with project performance.

There are a number of ways the ratios can be transformed to capture potential problems that may occur when actual performance deviates from planned performance. The simplest transformation is to take the absolute value of the deviation from the perfect actual:planned ratio of 1. However, such a transformation gives equal weighting to deviations both close to one and far from one. The SPD traffic light categories instead capture the supposition that the impact of deviations increases as the ratio moves further away from one. A priori, we do not have any good theoretical reason to prefer any particular transformation that weights deviations farther from one more heavily than those closer to one. We thus generated a number of different weighting functions, but present the results of just one of those here in the text.[6] The first is based on the traffic light categorization, with some alteration. Letting "Y" stand for C (cost) and S (schedule), the weighting proceeded as follows:

.1* absolute value (1-"Y"PIa)	if "Y"PIa> .8 and "Y"PIa<=2
.5* absolute value (1-" Y"PIa)	if "Y"Pia> .4 and "Y"Pia<=.8
1 * absolute value (1-" Y"PIa)	if "Y"PIa<= .4 or "Y"PIa>2

In addition to cost and schedule performance, we also have data on the percent of project funds disbursed. While this is not likely to be as good a measure of project performance as the cost performance index, which accounts for differences from planned expenditures, it is a measure that has been used in other studies because of its availability.

2.2 Explanatory Variables.

[5] Beginning in 2015, SPD has now categorized values greater than 2 as "outliers". For our analysis, we use the original categories to develop our cost and schedule performance indicators.

[6] To highlight the robustness of results to alternative weighting specifications, in Appendix 2, we provide results for two alternative weighting transformations and a simple, unweighted, absolute deviation.

The primary explanatory variables of interest include measures of "quality at entry" discussed above, including project logic and economic analysis scores. Additionally, though SPD combines monitoring and evaluation into a single score for reporting purposes, given that our focus is on project implementation, we have kept scores for more traditional monitoring activities separate from the evaluation design scores, as the latter focus more on evaluating hoped-for impacts at the end of the project.

Since 2008, there have been a number of changes to the DEM framework, and thus the comparability of scores over time. In Gibbons and Mitnik (2016), the authors developed and applied a methodology to ensure comparability in the measures over time, primarily by re-weighting and harmonizing across all measures. We use these transformed measures of project logic, economic analysis, monitoring, and impact evaluation in our empirical analysis.

We also have additional information on the projects, which we include as control variables. In particular, we include a dummy for whether the DEM format used was 2010 or later, to capture any bias that may be remaining even after the variables were transformed for comparability. We also include a measure of initial project delays, which is the number of days from project approval to first disbursement. Initial delays may indicate potential problems that result in poor project performance throughout the duration of the project. Certain projects face longer delays because they must be ratified by the legislative body; we include a dummy for ratification to separate this effect from other influences on project delays. Next, we include dummies for whether the project is financed by concessional funds from the Fund of Special Operations (FSO) or by Ordinary Capital and by the type of loan, i.e. whether it is a specific or multiple works type operation, an innovation loan, or some other kind. In particular, specific investment operations are more fully defined at approval, including for example final designs and feasibility studies, and therefore we might expect that this contributes to better performance during execution. We also include the natural logarithm of the amount of the loan. Larger projects may make it more difficult to set cost and schedule targets. In addition, we also control for the age of the project.

Finally, we also include a number of country-specific controls, which vary by project within the country depending on the projects' start dates. These include real GDP growth rates and population densities. We also include two measures of economic volatility, the first is the variance in the real GDP per capita growth rate, and the second is a measure of currency instability, calculated as the number of consecutive months where foreign exchange declined by more than .25 standard deviations from the historical mean. In addition to all of these variables, we also include a set of sectoral division dummies and country dummies.[7]

2.3 Empirical Strategy

The dataset includes project progress scores in both 2014 and 2015. Thus, some projects have two scores, while others only have scores in 2014 or 2015, depending on the project cycle.

[7] We could not, however, include dummies for project leaders, as did Legovini et al. (2015) and Raimondo (2016). Our dataset contains missing entries for project leaders for nearly 1/3rd of the projects, and even for those projects with this information, nearly 40% of project leaders led only one project.

Additionally, all of our critical variables of interest are fixed across both years, as are all of the project-specific control variables. Only the country-level control variables vary over time.

To deal with these issues, we follow two paths. First, we created average scores for those projects with two observations, as well as "latest" scores using the most recent score available. In these cases, we do not have multiple observations for the same project, and thus run simple OLS regressions for the continuous performance indicators and probits for the cost and schedule performance binary variables that take the value of one when performance is satisfactory[8]. In all cases, we control for clustered errors, where the cluster is the IDB sectoral division by country.

Next, in order to use all of the observations, we run Generalized Estimating Equations (GEE) using the entire set of observations. The GEE is a semi-parametric model first introduced by Liang and Zeger (1986) that treats the error structure as a nuisance, allowing for unknown correlation between repeated observations. The GEE requires the researcher to select a correlation matrix structure; we specified the unstructured correlation matrix, with robust standard errors. Selecting the unstructured matrix imposes the least structure, though it can also lead to some efficiency loss due to the larger number of parameter estimates than other specifications (Agresti, 2013).[9] Given our unbalanced panel, this selection is arguably the best choice.

Finally, for project proposals beginning in 2010, there are two sets of "quality at entry" scores. The first set of scores is given at QRR, and the second set is given at OPC. The first set of scores (QRR scores) is determined by SPD reviewers based on project documentation prepared by teams at the end of the first stage of project preparation. As noted above, SPD reviewers provide substantive comments and recommendations to improve all dimensions of project evaluability at this stage. In the second stage, a more detailed final proposal is submitted, and the SPD reviewer generates a second set of scores, the OPC scores, which, if minimum thresholds are met, allows the project to be presented to the Operation and Policy Committee. A priori, we expect the OPC scores to more accurately reflect the quality of the project, both because project team members work to improve their QRR scores and because more information is then known to the SPD reviewer to determine the final score.

---

[8] We tested for normality of the residuals using the "sktest" command in STATA, which tests for skewness and kurtosis in the error distributions. We ran the tests because due to the fact that number of observations are bunched at one, that is, where actual and planned performance coincide. Test results lead us to not reject normally distributed errors except in the case of the squared CPI(a) deviation. In the latter case, the test for skewness was significant at the 10% level. To probe the robustness of results, we normalized all of the cost and schedule deviation variables, and ran generalized linear models (GLMs). The results for all of GLMs are qualitatively very similar to the OLS results, including for the squared cost deviation, thus we include only the OLS results here in the paper. Results can be obtained from the corresponding author.

[9] We also ran the estimations using the "exchangeable" structure, which produced similar results that can be obtained from the corresponding author.

However, because of the importance that many division managers place on OPC scores, there is anecdotal evidence of project members "lobbying" their SPD reviewers, which can lead to inflated scores. We can test the different impacts of QRR versus OPC scores by using the data available on the subset of projects whose proposals were submitted from 2010 onwards. To do so, we run only the GEE specifications, which gives us the maximum number of useable observations.

Finally, we note that, given our sample size, we have limited ability to employ more sophisticated empirical strategies to address potential endogeneity between quality at entry scores and project performance, such as the propensity score matching employed by Raimondo (2016). Additionally, one of the main concerns expressed in Raimondo (2016) and Legovini et al. (2015) is potential endogeneity when the capacity of project team members is omitted, since such capacity may well be related both to quality at entry scores as well as better project performance. Both of those authors use project member fixed effects, which we cannot do given our dataset.

Nonetheless, we do have information on scores from QRR and OPC that can provide some information on the link between capacity of project members and quality at entry scores. If this link is positive, then we would expect that QRR and OPC scores on each of the four dimensions would be highly correlated, and that the four quality at entry scores themselves would be highly correlated at QRR and at OPC, since high capacity teams should perform relatively better across all quality at entry dimensions, and they should obtain higher scores both at QRR and OPC. As shown in the Descriptive Statistics section below, most correlation coefficients point to modest correlation across quality at entry scores on the four dimensions within each stage and for the same score between QRR and OPC. Taken together, the correlations suggest that the quality at entry process, reflected in the scores, has a separate impact from team capacity, though both are likely captured in final OPC scores. Thus, estimated impacts likely capture an upper-bound impact of the scores on project performance.

2.4 Descriptive Statistics

Before moving on to the econometric results, in this section we present some descriptive statistics for the DEM quality at entry scores and for project implementation indicators. Particularly in the first few years, 2008-2010, IDB staff from most divisions participated in trainings that focused mainly on developing project logic, conducting economic analyses, and understanding the range of methodologies that can be used to conduct rigorous impact, and to a lesser extent, developing more traditional monitoring and internal project evaluation frameworks. Figure 2 below shows the trends over time for the averages of these four scores, using the OPC data. Note that because we have separated monitoring and impact assessment scores, the maximum these scores can take is 2.5 and 7.5 respectively whereas the project logic and economic analysis maximum scores are 10. To facilitate comparison, we have thus normalized all four scores.

Figure 2: OPC Scores over time

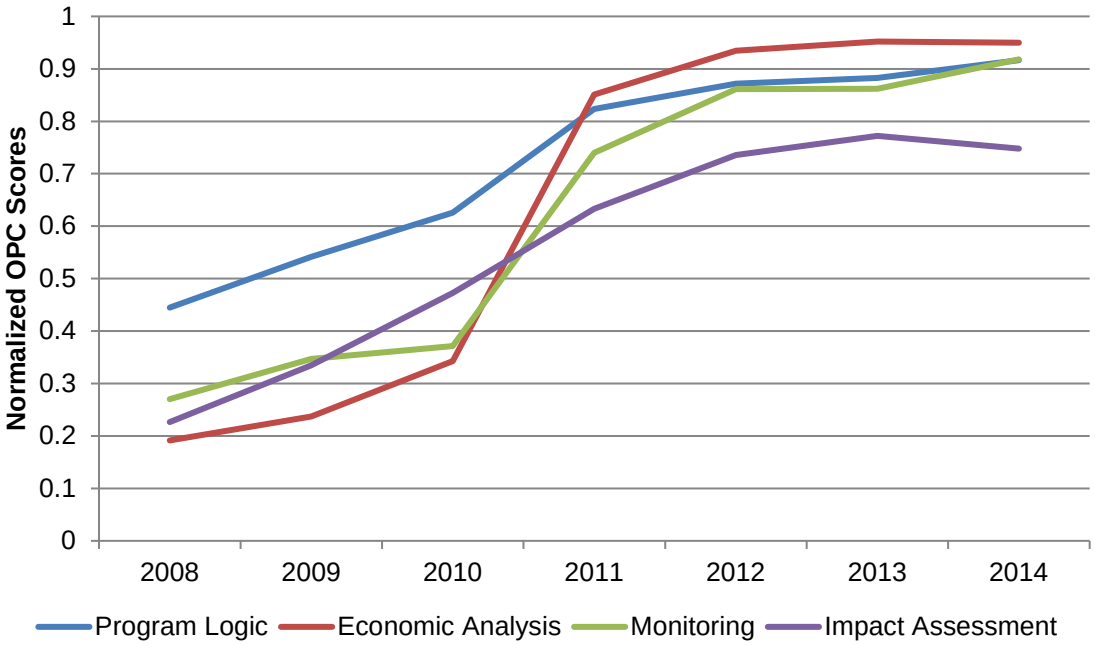
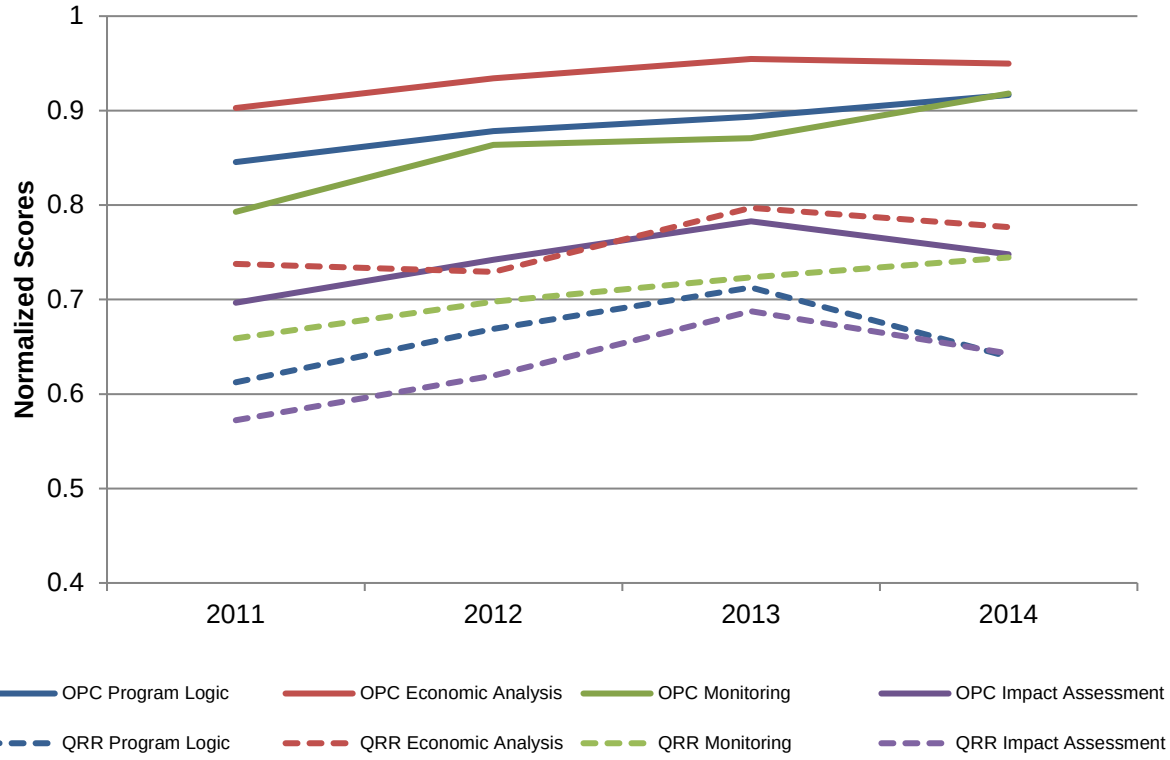


Figure 2 shows substantial increases for all scores, particularly for economic analysis and for monitoring[10]. Average program logic scores are high, but also started at a higher level. The economic analysis scores jumped steeply after 2010. We note here that projects generally hire consultants to perform the economic analysis, and the jump indicates that project proposers were able to quickly transmit new requirements for scoring the economic analysis to consultants, and that those consultants could indeed make the changes required. Impact assessment scores also increased but less so than the other scores. In part this is due to the fact that not all projects can undertake rigorous statistical analyses of project impacts for a number of reasons, including program design and costs of such analyses. As noted above, SPD trainings and workshops focused relatively less on traditional monitoring frameworks, though scores increased quite steeply nonetheless.

We can also evaluate the trends over time for the QRR scores. Figure 3 captures these trends from 2011, and for ease of comparison, we also include the OPC scores as well.

[10] Gibbons and Mitnik (2016) describes the evolution of DEM scores in detail by sectors and specific functional divisions, both at QRR and OPC. Overall, they find a general increase from QRR to OPC with heterogeneity across sectors, particularly at QRR stage.

Figure 3: OPC and QRR Score Trends



As might be expected all QRR average scores are below their respective OPC scores. For the most part QRR trends were increasing similarly to OPC scores until 2013, though three of four QRR scores trended down in 2014. Somewhat obscured in the graph above is the fact that program logic scores increase the most between QRR and OPC, and impact assessment scores increase the least.

As discussed above, the difference between QRR and OPC may be due to factors associated with higher quality proposals as well as the ability of the SPD reviewer to more accurately assess quality in the second stage. However, it may also be due to lobbying on the part of project proposers. If the latter were true, we would expect differences between QRR and OPC to become more pronounced over time for a number of reasons. First, it took some time for division managers to stress the importance of these scores to project staff. Second, as scores became more important to managers, project staff might attempt to learn how to influence their reviewers. Thus, we would expect that the impact of successful lobbying would grow over time. The data, however, is not consistent with such an hypothesis, as shown in Table 1 below.

Table 1: Difference in OPC and QRR Scores over time

	(OPC-QRR) Program Logic	(OPC-QRR) Economic Analysis	(OPC-QRR) Monitoring	(OPC-QRR) Impact Assessment
2011	0.23	0.17	0.13	0.12
2012	0.21	0.21	0.17	0.12
2013	0.18	0.16	0.15	0.10
2014	0.28	0.17	0.17	0.10

As discussed above, we are also concerned with determining the extent to which scores are correlated with underlying team capacity. If the scores are capacity-driven, we would expect that the correlation along the four dimensions would be high at both OPC and QRR. Tables 2a-2c show the correlation coefficients across project scores at QRR, at OPC and across OPC and QRR, for the subset of projects with both scores, N=442. As shown in Table 2a and 2b below, the correlation across project scores for OPC and QRR are modest, ranging from .1 to .32 for OPC, and .17 to .3 for QRR. We would also expect high capacity teams to have higher scores across both the QRR and OPC stages. As shown in Table 2c below, there is moderate positive correlation across the stages except for impact assessment, which has a relatively high correlation. Finally, in almost all cases, the economic analyses are undertaken by external consultants. We would thus expect these scores to be less influenced by team capacity. But in all three tables, we see that the economic analysis score correlations are quite similar to the other pairwise correlations. Altogether, the descriptive analysis suggests that project scores are not dominated by team capacity, though it is likely they reflect team capacity in part.

Table 2a: Correlation across Project Scores for OPC

	OPC			
OPC	Project Logic	Economic Analysis	Monitoring	Impact Assessment
Project Logic	1			
Economic Analysis	0.3218	1		
Monitoring	0.1398	0.1255	1	
Impact Assessment	0.2405	0.1474	0.1047	1

Table 2b: Correlation across Project Scores for QRR

	QRR			
QRR	Project Logic	Economic Analysis	Monitoring	Impact Assessment
Project Logic	1			
Economic Analysis	0.2981	1		
Monitoring	0.1767	0.2378	1	
Impact Assessment	0.2724	0.174	0.2687	1

Table 2c: Correlation between OPC and QRR project scores

	OPC			
QRR	Project Logic	Economic Analysis	Monitoring	Impact Assessment
Project Logic	0.4586	0.1568	0.0843	0.1161
Economic Analysis	0.2595	0.4841	0.1593	0.0629
Monitoring	0.0941	0.0959	0.5967	0.1505
Impact Assessment	0.2244	0.2194	0.1439	0.7312

III RESULTS

Table 2 gives results for the specifications using OPC quality at entry scores as regressors. Coefficients are given for the OLS regressions, and average marginal effects are given for the probit and GEE models. In Table 2, we only present results for our regressors of interest, the OPC quality at entry variables and the most consistently significant control variable, delay from approval to first disbursement. Full results tables are provided in Appendix 1. For the weighted versions of the CPI(a) and SPI(a) deviation measures, negative coefficients indicate lower deviations and thus improvements toward a perfect score. For the percent disbursed, positive coefficients indicate better performance. For the CPI(a) and SPI(a) traffic light based binary measure, a positive result indicates a marginal increase in the probability of receiving a satisfactory rating. Finally, to summarize the goodness of fit for the GEE models presented, we adopt the measure proposed by Zheng & Agresti (2000) and present the correlation between our outcome variable Y and our predicted value for the variable $\hat{Y}$; we also present the adjusted R^2 for the OLS results.

From Table 2, we note that the most robust result is the impact of the OPC Project Logic score on indicators of schedule performance, and for certain cost performance specifications. Better

OPC Project Logic scores on all of these indicators lead to better project performance. A higher OPC Economic Analysis score also tends to have a statistically significant and positive impact on project performance, particularly for schedule performance.

Table 2: Project Performance Estimations

	Cost Performance, Satisfactory=1		Cost Performance, Weighted Indicator		
	Last Score (Probit)	Full Sample (GEE)	Avg. Score (OLS)	Last Score (OLS)	Full Sample (GEE)
OPC Program Logic	0.166 ***	0.027 **	-0.028 **	-0.040 **	-0.004
OPC Economic Analysis	-0.000	0.008	-0.020	-0.023	-0.006 *
OPC Monitoring	-0.170	-0.025	0.063	0.096	0.014 *
OPC Evaluation	-0.101	-0.001	-0.002	0.014	-0.001
Initial Project Delay	-0.014 ***	-0.002 ***	0.004 ***	0.003 ***	0.001 ***
Number of observations	490	846	497	497	853
Adjusted R2					
cor(Y,E(Y|X))	0.442	0.506	0.572	0.537	0.483

	Schedule Performance, Satisfactory=1		Schedule Performance, Weighted Indicator		
	Last Score (Probit)	Full Sample (GEE)	Avg. Score (OLS)	Last Score (OLS)	Full Sample (GEE)
OPC Program Logic	0.114 **	0.031 **	-0.025 ***	-0.020 **	-0.008 **
OPC Economic Analysis	-0.007	0.007	-0.017 **	-0.013 *	-0.011 ***
OPC Monitoring	-0.058	-0.011	0.029	0.034	0.015
OPC Evaluation	-0.024	-0.002	0.009	0.011	0.004
Initial Project Delay	-0.017 ***	-0.004 ***	0.004 ***	0.004 ***	0.001 ***
Number of observations	497	853	497	497	853
Adjusted R2					
cor(Y,E(Y|X))	0.525	0.481	0.631	0.594	0.6

Table 2: Project Performance Estimations, continued

	Percent Disbursed		
	Average Score	Last Score	Full Sample
OPC Program Logic	0.005	0.010	0.005
OPC Economic Analysis	0.006	0.005	0.008 *
OPC Monitoring	-0.031	-0.049 **	-0.028
OPC Evaluation	-0.012	-0.011	-0.007
Initial Project Delay	-0.004 ***	-0.004 ***	-.0008 ***
Number of observations	434	434	853
Adjusted R2			
cor(Y,E(Y|X))	0.801	0.795	0.798

The OPC Monitoring score is significant in a few specifications, though where significant, it has an unexpected sign. Specifically, it is positive and significant in one of the cost deviation measures, and it is negative and significant in one of the "percent disbursed" measures.

The OPC Impact Assessment score is never significant. The score is based on the quality of the proposed impact assessment of the project that takes place at the end of the project. As argued by Legovini et al. (2015), data collection at baseline can be useful for project implementation. Our results suggest that baseline data, when collected, have not yet been successfully used for implementation purposes.

Finally, we note that initial project delays have a negative and significant impact on project performance in every equation presented in Table 2. The longer it takes for first disbursement, the worse a project will do on cost performance, schedule performance, the percent disbursed, and then of course, the synthetic indicator of performance shown in Appendix 1. Initial delays signal significant problems that affect project performance throughout its duration, suggesting there is wide scope for improving mechanisms to specifically address the problems causing initial delays.

In Table 3, we present results using OPC and QRR scores, with the smaller dataset comprised of projects that began after 2010. As noted above, to use the maximum number of observations possible, we ran GEE's that account for potential correlation between error terms for projects with more than one observation in the dataset. We present results for the CPI(a), SPI(a) measures, noting that there were no statistically significant parameter estimates in the percent disbursed equations.

We first note that Program Logic, using the smaller dataset and OPC scores, no longer has any statistically significant impact on cost performance, but it remains statistically significant and positive on all measures of schedule performance. The OPC Economic Analysis score is positively associated with project performance in six of eight equations. OPC Monitoring scores are no longer significant predictors in any performance equation. However, OPC Impact Assessment scores are negative and significant in two of the cost performance equations, indicating that project's may have begun to use baseline data to inform project implementation relatively recently.

Table 3: Project Performance Estimations, Comparing OPC and QRR Quality at Entry Scores

	Cost Performance, Satisfactory=1		Cost Performance, Weighted Indicator	
	OPC	QRR	OPC	QRR
Program Logic	0.017	-0.006	-0.001	0.003
Economic Analysis	0.049 **	0.017 **	-0.015 **	-0.004 *
Monitoring	-0.006	-0.043	0.001	0.010
Evaluation	0.006	0.018	-0.010 *	-0.010 ***
Initial Project Delay	-0.004 ***	-0.004 ***	0.001 ***	0.001 ***
Number of observations	410	410	413	413
cor(Y,E(Y|X))	0.544	0.506	0.529	0.534

	Schedule Performance, Satisfactory=1		Schedule Performance, Weighted Indicator	
	OPC	QRR	OPC	QRR
Program Logic	0.051 *	-0.008	-0.020 **	0.003
Economic Analysis	0.025	0.008	-0.021 **	-0.008 **
Monitoring	-0.024	-0.011	0.024	0.022 *
Evaluation	-0.004	-0.019	0.003	0.002
Initial Project Delay	-0.005 ***	-0.005 ***	0.002 ***	0.002 ***
Number of observations	410	410	413	413
cor(Y,E(Y|X))	0.525	0.481	0.631	0.594

With respect to comparing QRR scores and OPC scores, results suggest that OPC scores are modestly better measures of quality at entry. This is particularly true for Project Logic, as the QRR score does not have a significant impact on any measure of project performance. The QRR Economic Analysis is associated with improved cost and schedule performance, but in those cases where it is significant, OPC scores have statistically higher coefficients. The latter indicates either that there were real improvements in the economic analysis between QRR and OPC, or that reviewers provided more accurate scoring at OPC, or both. The QRR Monitoring score is only significant in the weighted schedule performance equation, and it has an unexpected positive sign. Finally, the QRR Impact Assessment score is negative and significant in the cost performance deviation specifications, with coefficients very similar to OPC score coefficients. Though Impact Assessment scores increase from QRR to OPC, results suggest that the increase is similar across projects indicating little differentiation across projects when moving from QRR to OPC.

IV CONCLUDING COMMENTS

The good news is that the evidence suggests that an emphasis on project logic and economic analyses has had a positive impact on project performance. And, evidence also suggests that while OPC Project Logic scores have a positive impact on schedule performance, QRR scores do not. And, OPC scores for Economic Analyses have positive and statistically larger marginal effects on cost performance vis-à-vis QRR scores. Working with SPD can help project proposers improve their scores during the period from QRR to OPC, and having SPD revise scores in the second stage should lead to more precise, less noisy scores. With the data at hand, it is not possible to disentangle the effects of these two different mechanisms. The fact that OPC scores are more robust predictors of project performance also lends weight to the argument that it is the DEM process itself that leads to better project performance, and not simply a reflection of unobserved differences in team capacity.

Our one odd variable is the OPC Monitoring score. This score should be most closely related to the project implementation process, but nonetheless, when significant, it has a negative impact on project implementation. The evidence thus suggests a need to reconsider the monitoring scoring section of the DEM. As shown in Figure 2, there were sharp increases in the Monitoring scores after 2010, but these increases have not translated into better performing projects, either when looking at results using all years in the dataset or when looking just at results for those projects with scores starting in 2011. A number of reasons may be behind why, as captured, monitoring plans are not leading to improved decision making during project execution and thus to better project performance. First, the sub-components might not adequately capture what really needs to be done in order to effectively monitor the most important variables of interest. In particular, it may be that the sub-components are too coarse to enable SPD reviewers to distinguish high vs. low quality monitoring. For instance, of the five sub-components, the three related to project costs are fairly trivial, including “total project costs are grouped by each output”, “costs for each output have annual amounts”, and “total costs equal to the total project amount”. Finally, none of the sub-components are directly related to evaluating the data collected for monitoring, or how such an evaluation would feedback into program implementation. These results differ from Raimondo (2016) who found that higher quality

monitoring led to better project performance; in that analysis, the monitoring rating is a function of more substantive components such as quality of the monitoring design, and utilization of the monitoring outputs during the course of the project. To date, SPD has not focused on more traditional project monitoring mechanisms in its trainings and documentation, but doing so could lead to improved monitoring, particularly if accompanied by more meaningful sub-components for the DEM evaluation.

Overall, the analyses in this paper support the hypothesis that the DEM is an effective framework for assessing quality at entry for IDB projects, though there is room for improvement on assessing traditional monitoring. It also supports the hypothesis that it is important for SPD to continue working with project team members through the stages to improve quality at entry. The most important measures of project quality, however, are its final impacts. Assessing project completion reports is an important task for the future, when sufficient data has been collected.

References

Belli, Pedro and L. Pritchett. 1995. Does Good Economic Analysis Improve Project Success? Mimeo, Operations Policy Division. World Bank Group.

Gibbons, A. and O. Mitnik. 2016. Development Effectiveness Matrix (DEM) Scores Harmonization: 2008-2015. Inter-American Development Bank IDBDocs #39835062.

Gibbons, A. and O. Mitnik. 2016. Evaluability of IDB Sovereign Guarantee Loans in Perspective: 2011-2015. Inter-American Development Bank SPD Mimeo.

Jenkins, Glenn, P. 1997. Project Analysis and the World Bank. AER, Vol. 87, No.2, pp.38-42

Legovini, A., V. Di Maro, and C. Piza. 2015. Impact Evaluation Helps Deliver Development Projects. Policy Research Working Paper 7157. Washington, DC: World Bank Group.

Papke, L.E. and J. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics, 11: 619-632.

Raimondo, E. 2016. *What difference does good monitoring & evaluation make to World Bank project performance ?*. Policy Research Working Paper 7726. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/180811468197076970/What-difference-does-good-monitoring-evaluation-make-to-World-Bank-project-performance

Vawda, Y.A., P. Moock, P. J. Gittinger, and A.H. Patrinos. 2003. Economic Analysis of World Bank Education Projects and Project Outcomes. *International Journal of Educational Development*, 23(6): 645-660.

Appendix 1: Full Results for Performance Indicators

	Cost Performance Probit, Satisfactory = 1					Cost Performance, Weighted Indicator				
	MLE Mean	MLE Last	GEE, f(bin) l(probit)	GEE, f(bin) l(probit)	GEE, f(bin) l(probit)	OLS Mean	OLS Last	GEE, f(bin) l(logit)	GEE, f(bin) l(logit)	GEE, f(bin) l(logit)
OPC Program Logic Score	0.136 *	0.166 ***	0.027 **			-0.028 **	-0.040 **	-0.004		
OPC Economic Analysis Score	0.008	-0.000	0.008			-0.020	-0.023	-0.006 *		
OPC Monitoring Score	-0.013	-0.170	-0.025			0.063	0.096	0.014 *		
OPC Evaluation Score	-0.053	-0.101	-0.001			-0.002	0.014	-0.001		
QRR Program Logic Score				-0.006					0.003	
QRR Economic Analysis Score				0.017 **					-0.004 *	
QRR Monitoring Score				-0.043					0.010	
QRR Evaluation Score				0.018					-0.010 ***	
OPC Program Logic Score					0.017					-0.001
OPC Economic Analysis Score					0.049 **					-0.015 **
OPC Monitoring Score					-0.006					0.001
OPC Evaluation Score					0.006					-0.010 *
Project complies with FOE	3.142	5.235 **	0.402	1.676	1.952	0.364	-2.182	-0.002	0.254	0.237
Ratification occurred for this project	1.156 ***	0.453	0.064	0.017	-0.005	-0.221 *	-0.316	-0.001	-0.028	-0.013
Delay from project approval to first disbursement	-0.015 ***	-0.014 ***	-0.002 ***	-0.004 ***	-0.004 ***	0.004 ***	0.003 ***	0.001 ***	0.001 ***	0.001 ***
Project approved amount [ln(millions of USD)]	-0.315 ***	-0.221 ***	-0.044 ***	-0.028	-0.031	0.030	0.033	0.007	0.011	0.011 *
Operation Sub-Type = ESP or INO or TCR	0.122	-0.014	-0.008	-0.039	-0.050	-0.048	-0.031	0.002	0.018 *	0.026 **
DEM format is 2010 or later	0.709 **	0.368	0.069		*-omitted-*	-0.119	-0.121 *	-0.012		0.000
Number of years between approval and PMR		0.530 ***	0.152 ***	0.225 ***	0.237 ***		-0.106 ***	-0.031 ***	-0.046 ***	-0.052 ***
Mean number of years between approval and PMR	0.659 ***					-0.142 ***				
% Change in Real GDP Per Capita	0.084	-1.463	-0.348	-0.914	-0.804	-0.936	-0.421	-0.188	0.358	0.375
Variance in Real GDP Per Capita Growth Rate	-0.047	-0.078 **	-0.010 **	-0.037 ***	-0.039 ***	0.002	-0.005	0.002 *	0.011 ***	0.012 ***
Land area per population (km2/person)	02.202	-89.872	-0.338	9.500	-2.811	1.845	12.143	-2.518	-38.618	-39.137
Foreign exchange instability (consecutive months)		2.179 ***	0.048	-0.203	-0.153		-0.190 **	-0.016	0.060	0.047
Mean foreign exchange instability (consecutive months)	0.000					-0.127 **				
Division = CMF	-0.674 **	-0.903 ***	-0.076	-0.105	-0.092	-0.020	0.114	-0.010	-0.000	-0.007
Division = CTI	0.892 *	-0.255	-0.006	-0.002	-0.023	-0.019	0.066	-0.001	-0.009	-0.006
Division = EDU	0.357	-0.088	0.063	0.014	0.032	-0.027	0.034	-0.020	-0.060 **	-0.064 **
Division = ENE	0.289	0.000	0.051	0.142	0.111	-0.098	0.036	-0.034	-0.074 **	-0.072 **
Division = ICS	0.072	-0.269	0.026	0.061	0.047	-0.055	0.067	-0.014	-0.032	-0.028
Division = LMK	0.986	1.122 *	0.204	0.220	0.238	-0.116	-0.082	-0.038	-0.040	-0.047
Division = RND	-0.477 *	-0.637 **	-0.065	-0.087	-0.122	0.108	0.230 *	0.014	0.032	0.038
Division = SPH	0.707	0.511	0.170 **	0.254 **	0.287 **	-0.126	-0.006	-0.047 ***	-0.096 ***	-0.105 ***
Division = TIU	-0.481	-1.853 ***	-0.206 ***	-0.214 *	-0.202 *	0.402 *	0.632 **	0.048 ***	0.060 ***	0.056 ***
Division = TSP	0.152	-0.161	0.019	-0.024	-0.041	-0.024	0.079	-0.008	0.002	0.001
Division = WSA	0.005	-0.157	0.029	-0.042	-0.065	-0.078	-0.019	-0.016	-0.004	-0.005
Argentina	3.179	3.111	0.129	-0.041	0.258	-0.107	-0.431	0.054	0.877	0.894
Bahamas, The	-4.783	-5.383 *	-0.262	0.175	-0.384	0.168	0.782	-0.050	-1.433	-1.442
Barbados	0.000	-1.944	0.061	0.219	-0.030	-0.180	0.011	-0.102 *	-0.638 *	-0.635
Belize	0.000	1.467	-0.105	-0.681	-0.357	-0.091	-0.416	0.037	1.063	1.094
Bolivia	2.536	0.384	-0.396	-2.215	-1.726	-0.254	1.553	0.164	2.189 *	2.229 *
Chile	-2.936 ***	-2.133 ***	-0.415 ***	-0.328	-0.349	0.495 **	0.272	0.080 ***	0.142 **	0.143 **
Colombia	-2.571	-2.344	-0.123	0.047	-0.142	0.367	0.652	0.003	-0.634	-0.655
Costa Rica	-3.747	-2.209	-0.018	0.246	-0.157	0.280	0.472	-0.062	-1.115	-1.133
Dominican Republic	-5.280	-3.509	-0.016	0.513	0.042	0.445	0.835	-0.068	-1.433	-1.461
Ecuador	-2.915	-2.382	-0.082	0.109	-0.205	0.107	0.213	-0.043	-0.962	-0.975
El Salvador	-5.913	-4.514	-0.148	0.197	-0.278	0.476	0.905	-0.057	-1.401	-1.427
Guatemala	0.000	-13.169 ***	-0.692	*-omitted-*	*-omitted-*	-0.046	2.885	-0.016	0.000	0.000
Guyana	0.000	14.523	-0.086	-3.206	-0.852	0.000	0.000	0.694	7.968	8.088
Haiti	-4.923	-4.229	-0.140	0.247	-0.173	0.276	0.778	-0.072	-1.431	-1.479
Honduras	-6.965	-7.731 *	-0.481	-1.504	-2.070	-0.132	2.701	-0.054	-1.244	-1.255
Jamaica	-4.362	-4.347	-0.269	-0.122	-0.584	0.334	0.491	-0.021	-1.298	-1.312
Mexico	-1.752	-1.871	0.084	0.245	-0.110	-0.124	0.220	-0.066	-0.972	-0.971
Nicaragua	-5.890	-6.395 *	-0.319	-1.453	-1.978	-0.062	2.706	-0.047	-1.019	-1.026
Panama	-0.388	0.134	0.034	0.281 **	0.305 **	0.043	-0.022	-0.002	-0.038	-0.047
Paraguay	0.000	-1.831	-0.076	0.272	-0.039	0.509	0.534	0.006	-0.791	-0.809
Peru	0.000	0.000	*-omitted-*	*-omitted-*	*-omitted-*	-0.026	2.601	0.000	0.000	0.000
Suriname	24.691	22.057	-0.145	-2.612	0.518	-0.195	-2.961	0.717	9.841	9.964
Trinidad and Tobago	-5.649	-4.765	-0.318	0.288	-0.206	0.868	1.261	-0.017	-1.432	-1.445
Uruguay	-0.211	0.545	-0.071	-0.159	-0.052	0.182	-0.026	0.072 *	0.417	0.433
Venezuela, RB	0.000	0.000	*-omitted-*	*-omitted-*	*-omitted-*	-0.364 *	-0.114	-0.191 ***	-0.971 ***	-0.997 ***
Number of observations	440	490	846	410	410	497	497	853	413	413
Adjusted R-Square						0.251	0.208			
Pseudo R-Square	0.301	0.303								
corr(Y,E(Y|X))			0.506	0.544	0.554			0.483	0.529	0.534

*Table report marginal effects. Asterisks denote significance, * p<0.10, ** p<0.05, *** p<0.01.*

Appendix 1, Continued

	Schedule Performance, Satisfactory = 1					Schedule Performance, Weighted Indicator				
	MLE Mean	MLE Last	GEE, f(bin)	GEE, f(bin)	GEE, f(bin) l(probit)	OLS Mean	OLS Last	GEE, f(bin)	GEE, f(bin)	GEE, f(bin)
OPC Program Logic Score	0.159 ***	0.114 **	0.031 **			-0.025 ***	-0.020 **	-0.008 **		
OPC Economic Analysis Score	0.046	-0.007	0.007			-0.017 **	-0.013 *	-0.011 ***		
OPC Monitoring Score	-0.006	-0.058	-0.011			0.029	0.034	0.015		
OPC Evaluation Score	-0.037	-0.024	-0.002			0.009	0.011	0.004		
QRR Program Logic Score				-0.008					0.003	
QRR Economic Analysis Score				0.008					-0.008 **	
QRR Monitoring Score				-0.011					0.022 *	
QRR Evaluation Score				-0.019					0.002	
OPC Program Logic Score					0.051 *					-0.020 **
OPC Economic Analysis Score					0.025					-0.021 **
OPC Monitoring Score					-0.024					0.024
OPC Evaluation Score					-0.004					0.003
Project complies with FOE	-1.490	1.107	-0.240	1.572	1.433	4.456	4.774	0.177	0.154	0.258
Ratification occurred for this project	0.965 **	0.171	0.125	0.025	-0.048	-0.210 **	-0.154	-0.022	-0.017	0.010
Delay from project approval to first disbursement	-0.012 ***	-0.017 ***	-0.004 ***	-0.005 ***	-0.005 ***	0.004 ***	0.004 ***	0.001 ***	0.002 ***	0.002 ***
Project approved amount [ln(millions of USD)]	-0.119	0.029	0.008	0.028	0.024	0.000	0.002	-0.001	-0.010	-0.010
Operation Sub-Type = ESP or INO or TCR	-0.037	-0.102	-0.044	-0.039	-0.031	-0.035	-0.025	0.001	0.011	0.009
DEM format is 2010 or later	0.141	-0.260	0.014		0.000	-0.133 ***	-0.106 **	-0.027		0.000
Number of years between approval and PMR		0.339 ***	0.132 ***	0.206 ***	0.214 ***		-0.135 ***	-0.064 ***	-0.111 ***	-0.115 ***
Mean number of years between approval and PMR	0.506 ***					-0.160 ***				
% Change in Real GDP Per Capita	-0.988	-0.797	-0.217	-1.269	-0.839	-0.187	0.359	-0.168	0.535	0.413
Variance in Real GDP Per Capita Growth Rate	0.017	-0.016	-0.000	-0.004	-0.000	-0.000	0.002	0.002	0.007 **	0.006 *
Land area per population (km2/person)	18.090	16.033	10.916	-55.218	-55.778	19.829	22.476	-8.676	-26.500	-29.338
Foreign exchange instability (consecutive months)		-0.118	0.000	-0.067	-0.022		-0.061	-0.006	0.077	0.058
Mean foreign exchange instability (consecutive mon	0.049					-0.080 **				
Division = CMF	-0.018	-0.279	0.012	0.061	0.084	0.071	0.191 **	0.021	0.039	0.026
Division = CTI	-0.193	0.105	0.004	0.233 **	0.128	0.097	0.045	0.033	-0.042	-0.003
Division = EDU	0.623 *	0.634 *	0.047	0.158	0.025	-0.028	-0.054	-0.015	-0.031	0.012
Division = ENE	0.361	0.584 *	0.089	0.181	0.058	-0.022	-0.018	-0.024	-0.034	0.024
Division = ICS	0.249	0.233	0.018	0.075	0.068	0.020	0.007	0.009	0.013	0.023
Division = LMK	1.181 *	1.586 ***	0.225	0.331	0.214	-0.128	-0.138 *	-0.062	-0.060	-0.036
Division = RND	-0.381 *	-0.389 **	-0.091	0.105	-0.061	0.126 ***	0.163 ***	0.036 *	0.027	0.085 **
Division = SPH	0.515	0.445 *	0.090	0.204 *	0.106	-0.062	-0.062	-0.030	-0.039	-0.018
Division = TIU	-0.873 *	-1.039 *	-0.163	0.076	0.041	0.134	0.167	0.024	0.038	0.052
Division = TSP	0.067	-0.076	-0.050	-0.005	-0.083	0.073	0.061	0.027	0.031	0.059 **
Division = WSA	-0.063	0.289	0.007	-0.005	-0.075	0.019	-0.039	0.019	0.066 **	0.095 ***
Argentina	-0.027	-0.010	-0.215	1.366	1.367	0.348	0.429	0.174	0.626	0.694
Bahamas, The	-0.264	-0.128	0.079	-2.143	-2.200	-0.718	-0.772	-0.289	-1.062	-1.146
Barbados	-1.054	-0.739	-0.294	-1.353	-1.366	-0.200	-0.226	-0.070	-0.298	-0.319
Belize	0.000	-0.046	-0.016	1.465	1.478	0.664	1.086	0.305	1.065	1.146
Bolivia	-0.101	-1.955	-0.532	2.047	2.283	-2.995	-3.242	0.401	1.511	1.578
Chile	-2.194 **	-0.602	-0.254	0.083	0.100	0.328	0.257	0.110 *	0.061	0.068
Colombia	0.403	0.480	0.154	-1.010	-1.025	-0.309	-0.438	-0.137	-0.446	-0.505
Costa Rica	-0.150	0.830	0.137	-1.596	-1.587	-0.408	-0.669	-0.230	-0.755	-0.852
Dominican Republic	-0.055	0.915	0.395	-1.899	-1.946	-0.497	-0.708	-0.303	-1.031	-1.121
Ecuador	0.302	0.357	0.180	-1.451	-1.499	-0.473	-0.655	-0.185	-0.637	-0.702
El Salvador	0.003	0.000	0.324	-2.175	-2.205	-0.488	-0.660	-0.289	-0.971	-1.076
Guatemala	2.911	0.233	0.619	0.000	0.000	-5.173	-5.649	-0.512	0.000	0.000
Guyana	0.000	-3.635	-1.853	11.511	11.475	0.000	0.000	1.786	5.274	5.889
Haiti	-0.457	0.158	0.190	-2.282	-2.179	-0.652	-0.815	-0.302	-0.999	-1.148
Honduras	1.753	0.319	0.602	-2.819	-2.614	-4.851	-5.343	-0.380	-0.873	-1.087
Jamaica	0.101	0.274	0.186	-2.333	-2.389	-0.658	-0.678	-0.260	-0.901	-0.982
Mexico	1.452	1.332	0.419	-1.311	-1.336	-0.643	-0.734	-0.238	-0.691	-0.733
Nicaragua	1.788	0.309	0.648	-2.574	-2.394	-4.757	-5.223	-0.382	-0.732	-0.915
Panama	0.203	0.388	0.034	0.211	0.173	-0.016	-0.059	0.001	-0.058	-0.046
Paraguay	0.000	0.928	0.291	-0.866	-0.845	-0.299	-0.504	-0.169	-0.665	-0.736
Peru	0.000	0.000	0.000	0.000	0.000	-3.717	-4.284	0.000	0.000	0.000
Suriname	-5.437	-4.571	-2.911	13.738	13.832	5.359	6.058	2.334	6.826	7.526
Trinidad and Tobago	0.180	0.199	0.347	-2.003	-2.070	-0.415	-0.447	-0.241	-1.001	-1.084
Uruguay	-0.418	0.043	-0.050	0.612	0.597	0.190	0.173	0.110	0.263	0.307
Venezuela, RB	0.586	1.091	0.344	-0.316	-0.317	-0.573 ***	-0.556 **	-0.290 ***	-0.692 **	-0.717 **
Number of observations	472	497	853	413	413	497	497	853	413	413
Adjusted R-Square						0.331	0.280			
Pseudo R-Square	0.207	0.249								
corr(Y,E(Y|X))			0.481	0.476	0.483			0.6	0.628	0.635

*Table report marginal effects. Asterisks denote significance, * p<0.10, ** p<0.05, *** p<0.01.*

Appendix 1, Continued

	Percent Disbursed					Synthetic Indicator				
	OLS Mean	OLS Last	GEE, f(bin)	GEE, f(bin)	GEE, f(bin) l(logit)	OLS Mean	OLS Last	GEE, f(bin)	GEE, f(bin)	GEE, f(bin)
OPC Program Logic Score	0.005	0.010	0.005			0.032 **	0.043 ***	0.015 **		
OPC Economic Analysis Score	0.006	0.005	0.008 *			0.018	0.008	0.011 *		
OPC Monitoring Score	-0.031	-0.049 **	-0.028			-0.108 *	-0.130 **	-0.062 **		
OPC Evaluation Score	-0.012	-0.011	-0.007			-0.008	-0.018	-0.004		
QRR Program Logic Score				0.007					-0.030 *	
QRR Economic Analysis Score				-0.006 *					0.029 ***	
QRR Monitoring Score				-0.007					-0.077	
QRR Evaluation Score				-0.005					0.012	
OPC Program Logic Score					0.013					0.017
OPC Economic Analysis Score					-0.000					0.074 ***
OPC Monitoring Score					-0.025					-0.132 **
OPC Evaluation Score					-0.012					0.007
Project complies with FOE	-0.264	-0.368	-0.076	-0.477	-0.319	-9.299 *	-10.609 **	-0.289	33.299	19.218
Ratification occurred for this project	0.072	0.065	0.059	0.092	0.092	-0.238	-0.178	-0.135 *	-0.322 *	-0.509 **
Delay from project approval to first disbursement (Da	-0.004 ***	-0.004 ***	-0.004 ***	-0.007 ***	-0.008 ***	-0.006 ***	-0.008 ***	-0.003 ***	-0.005 ***	-0.006 ***
Project approved amount [ln(millions of USD)]	0.004	0.005	0.002	0.009	0.009	-0.033	-0.011	-0.011	0.029	0.030
Operation Sub-Type = ESP or INO or TCR	-0.069 ***	-0.065 ***	-0.052 ***	-0.041 *	-0.048 **	-0.106 **	-0.107 **	-0.060 **	-0.067	-0.094
DEM format is 2010 or later	-0.002	0.002	0.049		0.000	-0.273 ***	-0.227 ***	-0.132 ***	*-omitted-*	*-omitted-*
Number of years between approval and PMR		0.107 ***	0.142 ***	0.127 ***	0.126 ***		-0.059	-0.024 **	-0.070	-0.050
Mean number of years between approval and PMR	0.106 ***					-0.058				
% Change in Real GDP Per Capita	0.592 **	0.519 *	0.102	-0.411	-0.199	-0.179	-0.102	-0.040	3.726	3.843
Variance in Real GDP Per Capita Growth Rate	0.002	0.002	-0.001	-0.001	-0.002	0.007	0.005	0.001	-0.116 ***	-0.144 ***
Land area per population (km2/person)	1.806	5.623	-0.993	24.866	17.423	42.856 *	47.935 *	17.835 *	63.109	99.480
Foreign exchange instability (consecutive months)		0.040	0.051	0.101 *	0.095		0.150	0.037	*-omitted-*	*-omitted-*
Mean foreign exchange instability (consecutive mont	0.038					0.111				
Division = CMF	0.025	0.012	0.041	0.163 ***	0.160 ***	-0.015	-0.172	-0.026	-0.024	0.031
Division = CTI	0.027	0.022	0.012	0.086	0.116 **	0.029	0.006	0.009	*-omitted-*	*-omitted-*
Division = EDU	0.070	0.080	0.082 **	0.059	0.052	0.229 **	0.235 **	0.135 **	*-omitted-*	*-omitted-*
Division = ENE	0.030	0.015	0.038	0.053	0.033	-0.037	-0.092	-0.030	*-omitted-*	*-omitted-*
Division = ICS	0.044	0.052	0.059	0.059	0.067	-0.060	-0.045	-0.020	0.051	0.007
Division = LMK	-0.049	-0.073	-0.060	-0.050	-0.037	0.132	0.062	0.076	*-omitted-*	*-omitted-*
Division = RND	-0.037	-0.035	-0.053 *	0.001	0.009	-0.104	-0.144	-0.046	*-omitted-*	*-omitted-*
Division = SPH	0.040	0.034	0.052	0.120 ***	0.129 ***	0.023	-0.011	0.032	*-omitted-*	*-omitted-*
Division = TIU	-0.023	-0.024	-0.051	0.095 **	0.098 **	-0.320 *	-0.427 **	-0.118 **	-0.177	-0.217
Division = TSP	0.012	0.012	0.020	0.089 **	0.080 *	0.090	0.027	0.055	-0.029	-0.085
Division = WSA	-0.038	-0.039	-0.023	0.036	0.021	-0.028	-0.034	-0.026	-0.124	-0.211 *
Argentina	-0.033	-0.103	0.080	-0.451	-0.270	-1.060 *	-1.155 *	-0.439 *	*-omitted-*	*-omitted-*
Bahamas, The	0.027	0.181	-0.137	1.175	0.884	1.474	1.566	0.615	*-omitted-*	*-omitted-*
Barbados	0.085	0.158	0.025	0.348	0.258	0.346	0.372	0.139	*-omitted-*	*-omitted-*
Belize	0.068	-0.014	0.066	-0.528	-0.295	-1.209 *	-1.288	-0.502	*-omitted-*	*-omitted-*
Bolivia	0.084	-0.027	0.132	-1.129	-0.816	6.873 *	7.811 **	-0.683	*-omitted-*	*-omitted-*
Chile	-0.013	-0.004	-0.011	0.032	0.037	-0.288	-0.190	-0.078	0.353	0.377
Colombia	0.045	0.113	0.011	0.499	0.369	0.333	0.467	0.128	-3.550	-2.493
Costa Rica	-0.084	0.051	-0.173	0.732	0.499	1.972 **	2.074 **	0.885 **	*-omitted-*	*-omitted-*
Dominican Republic	-0.006	0.149	-0.025	1.003	0.721	1.881 **	2.023 **	0.829 **	*-omitted-*	*-omitted-*
Ecuador	0.145	0.252	0.090	0.792	0.618	0.923	1.142 *	0.366	-4.516	-3.019
El Salvador	0.171	0.377	0.080	1.150	0.856	1.991 **	2.131 **	0.846 **	*-omitted-*	*-omitted-*
Guatemala	0.359	0.612	0.024	0.000	0.000	10.962 *	12.436 **	0.970	*-omitted-*	*-omitted-*
Guyana	-0.431	-1.127	0.318	-4.518	-3.199	0.000	0.000	-3.601 *	*-omitted-*	*-omitted-*
Haiti	0.069	0.221	-0.023	0.985	0.709	1.518	1.781 *	0.659	-6.417	-3.788
Honduras	0.468	0.728	0.210	1.314	0.952	10.855 *	12.386 **	0.993 *	*-omitted-*	*-omitted-*
Jamaica	0.088	0.209	-0.068	0.902	0.613	1.164	1.427	0.466	*-omitted-*	*-omitted-*
Mexico	0.301	0.431 *	0.213	0.819	0.637	1.054 *	1.215 *	0.426	-4.461	-2.964
Nicaragua	0.303	0.523	0.091	1.056	0.734	10.538 *	11.876 **	0.841 *	*-omitted-*	*-omitted-*
Panama	-0.010	0.009	0.020	0.130 **	0.129 ***	-0.028	-0.017	-0.012	*-omitted-*	*-omitted-*
Paraguay	-0.098	0.020	-0.066	0.595	0.438	1.189 **	1.302 **	0.505 *	*-omitted-*	*-omitted-*
Peru	0.000	0.000	0.000	0.000	0.000	8.330 *	9.531 **	0.000	*-omitted-*	*-omitted-*
Suriname	-0.472	-1.475	0.252	-6.153	-4.259	11.467 *	12.790 **	-4.755 *	*-omitted-*	*-omitted-*
Trinidad and Tobago	0.038	0.157	-0.172	1.043	0.765	1.578	1.713 *	0.661	*-omitted-*	*-omitted-*
Uruguay	-0.020	-0.066	0.047	-0.139	-0.071	-0.431 *	-0.471 *	-0.191	*-omitted-*	*-omitted-*
Venezuela, RB	-0.079	-0.052	-0.334 ***	0.389	0.343	0.323	0.405	0.170	*-omitted-*	*-omitted-*
Number of observations	434	434	787	402	402	460	460	770	104	104
Adjusted R-Square	0.595	0.584				0.176	0.171			
Pseudo R-Square										
corr(Y,E(Y|X))			0.798	0.827	0.825				0.169	0.188

*Table report marginal effects. Asterisks denote significance, * p<0.10, ** p<0.05, *** p<0.01.*

Appendix 2: Full Results for Two Alternative Weights for Cost and Schedule Performance Indexes

	Cost Performance, Absolute Difference From 1					Cost Performance, Alternate Weight				
	OLS Mean	OLS Last	GEE	GEE	GEE	OLS Mean	OLS Last	GEE	GEE	GEE
OPC Program Logic Score	-0.025 **	-0.041 ***	-0.005			-0.050	-0.159	0.000		
OPC Economic Analysis Score	-0.014	-0.016	-0.005 *			-0.178	-0.329 *	-0.002		
OPC Monitoring Score	0.059	0.104 *	0.013			0.887	1.118 *	0.009 **		
OPC Evaluation Score	0.001	0.018	-0.000			-0.128	0.075	-0.002 **		
QRR Program Logic Score				0.002					0.002	
QRR Economic Analysis Score				-0.002					-0.001	
QRR Monitoring Score				0.017 *					0.005	
QRR Evaluation Score				-0.007 **					-0.002 ***	
OPC Program Logic Score					-0.000					0.002
OPC Economic Analysis Score					-0.012 **					-0.003 **
OPC Monitoring Score					0.002					0.005 *
OPC Evaluation Score					-0.008					-0.004 ***
Project complies with FOE	1.481	-0.891	0.035	-0.092	-0.157	22.170	35.847	0.025	-0.010	-0.012
Ratification occurred for this project	-0.238 **	-0.354 *	-0.016	-0.054 **	-0.041 *	-1.221	-3.447	-0.004	-0.016 **	-0.007
Delay from project approval to first disbursement	0.003 ***	0.003 ***	0.001 ***	0.001 ***	0.001 ***	0.013	0.007	0.000 *	0.000	0.000
Project approved amount [ln(millions of USD)]	0.016	0.027	0.004	0.006	0.007	0.016	0.037	0.001	0.002	0.002
Operation Sub-Type = ESP or INO or TCR	-0.028	-0.007	0.003	0.025 **	0.029 ***	-0.674	-0.873	-0.000	0.005 *	0.006 **
DEM format is 2010 or later	-0.057	-0.052	-0.009		0.000	-0.293	-1.064	0.000		0.000
Number of years between approval and PMR		-0.053	-0.025 ***	-0.033 ***	-0.038 ***		0.335	-0.002	0.000	-0.002
Mean number of years between approval and PMR	-0.092 **					0.368				
% Change in Real GDP Per Capita	-0.912	-0.630	-0.113	0.307	0.287	16.120	-14.261	-0.038	0.040	0.057
Variance in Real GDP Per Capita Growth Rate	-0.003	-0.010	0.000	0.008 ***	0.009 ***	-0.039	-0.177	0.000	0.001 *	0.001 **
Land area per population (km2/person)	-3.501	6.540	-1.008	14.517	12.055	46.349	-118.995	-1.455	-1.629	-2.242
Foreign exchange instability (consecutive months)		-0.180 **	-0.021 **	0.018	0.004		-1.296	-0.005 *	-0.001	-0.007
Mean foreign exchange instability (consecutive mon	-0.125 **					-0.387				
Division = CMF	0.003	0.151	-0.007	0.000	-0.003	-1.042	-0.513	-0.012 **	-0.008 **	-0.010 **
Division = CTI	-0.058	0.007	-0.013	-0.023	-0.017	-0.905	-0.276	-0.008 *	-0.008	-0.004
Division = EDU	-0.028	0.011	-0.022	-0.060 **	-0.058 **	2.898	3.671	0.004	-0.025 ***	-0.025 ***
Division = ENE	-0.128	-0.024	-0.029	-0.070 **	-0.072 **	-0.734	0.876	-0.013 **	-0.026 ***	-0.026 ***
Division = ICS	-0.025	0.062	-0.008	-0.031	-0.027	-1.635	-0.350	-0.015 **	-0.021 **	-0.017 **
Division = LMK	-0.070	-0.011	-0.020	-0.035	-0.030	-0.608	-0.319	-0.009	-0.009	-0.007
Division = RND	0.108	0.232 *	0.018	0.026	0.037	-0.187	1.047	-0.001	0.003	0.005
Division = SPH	-0.090	0.002	-0.026	-0.065 **	-0.069 ***	-0.223	0.636	-0.009 **	-0.023 ***	-0.021 ***
Division = TIU	0.361 *	0.608 **	0.049 **	0.034	0.038	3.269	3.199	0.007	0.004	0.006
Division = TSP	-0.019	0.094	-0.008	-0.001	0.001	-0.780	0.897	-0.010	-0.000	-0.002
Division = WSA	-0.086	-0.038	-0.019	-0.013	-0.011	-1.522	-0.338	-0.016 *	-0.012 *	-0.015 **
Argentina	0.091	-0.227	0.030	0.316	0.257	2.497	3.282	0.044	0.029	0.048
Bahamas, The	-0.029	0.601	0.004	-0.508	-0.397	-3.180	-6.442	-0.052	-0.048	-0.067
Barbados	-0.126	0.022	-0.024	-0.207	-0.168	-1.562	-2.819	-0.033 *	-0.038	-0.050
Belize	-0.010	-0.370	0.009	0.364	0.306	1.310	2.128	0.040	0.017	0.041
Bolivia	-1.037	0.608	0.036	1.022	0.927	16.491	-24.514	0.073	0.121	0.155
Chile	0.462 **	0.313	0.070 **	0.102 **	0.093 *	2.721	4.525	0.019 **	0.016	0.010
Colombia	0.256	0.521	0.030	-0.209	-0.179	2.360	2.171	-0.007	-0.014	-0.031
Costa Rica	0.081	0.275	-0.021	-0.383	-0.308	0.038	-1.192	-0.040	-0.031	-0.055
Dominican Republic	0.232	0.618	-0.007	-0.540	-0.454	1.288	0.287	-0.047	-0.061	-0.084
Ecuador	-0.065	0.042	-0.018	-0.375	-0.313	0.502	-2.402	-0.030	-0.053	-0.066
El Salvador	0.207	0.619	-0.002	-0.498	-0.410	0.218	-1.488	-0.039	-0.046	-0.073
Guatemala	-1.343	1.325	-0.001	0.000	0.000	25.028	-40.946	-0.063	0.000	0.000
Guyana	0.000	0.000	0.283	3.040	2.578	0.000	0.000	0.329	0.333	0.465
Haiti	0.038	0.546	-0.015	-0.506	-0.431	0.133	-0.734	-0.047	-0.048	-0.084
Honduras	-1.404	1.295	-0.049	-0.262	-0.145	22.096	-36.429	-0.062	-0.019	-0.039
Jamaica	0.084	0.251	0.026	-0.419	-0.323	3.138	-5.827	-0.028	-0.028	-0.050
Mexico	-0.232	0.071	-0.029	-0.363	-0.294	-1.046	-3.096	-0.034	-0.041	-0.055
Nicaragua	-1.306	1.312	-0.043	-0.138	-0.036	20.320	-34.099	-0.049	-0.010	-0.026
Panama	0.011	-0.012	-0.009	-0.074 **	-0.085 **	1.658	1.500	0.001	-0.011	-0.015
Paraguay	0.411	0.496	0.048	-0.257	-0.208	4.231	2.966	-0.010	-0.009	-0.026
Peru	-0.894	1.619	0.000	0.000	0.000	16.012	-19.375	0.000	0.000	0.000
Suriname	1.147	-1.519	0.313	3.752	3.134	13.020	30.851	0.384	0.437	0.591
Trinidad and Tobago	0.627	1.017	0.050	-0.524	-0.432	12.577	9.248	-0.022	-0.057	-0.078
Uruguay	0.249	0.088	0.051	0.162	0.148	2.966	2.718	0.029 *	0.017	0.026
Venezuela, RB	-0.222	0.012	-0.060	-0.525 **	-0.521 **	-0.177	0.824	-0.025	-0.075	-0.086
Number of observations	497	497	853	413	413	497	497	853	413	413
Adjusted R-Square	0.191	0.188				0.052	0.074			
corr(Y,E(Y|X))	0.522	0.519	0.433	0.489	0.498	0.384	0.409	0.586	0.643	0.675

*Table reports marginal effects. Asterisks denote significance, * p<0.10, ** p<0.05, *** p<0.01.*

Appendix 2, Continued

	Schedule Performance, Absolute Difference From 1					Schedule Performance, Alternate Weight				
	OLS Mean	OLS Last	GEE	GEE	GEE	OLS Mean	OLS Last	GEE	GEE	GEE
OPC Program Logic Score	-0.025 ***	-0.021 **	-0.009 **			-0.032 **	-0.020	-0.002		
OPC Economic Analysis Score	-0.008	-0.004	-0.006 *			-0.039 *	-0.030	-0.003 ***		
OPC Monitoring Score	0.024	0.039	0.012			0.081	0.133	0.007 **		
OPC Evaluation Score	0.008	0.011	0.003			0.032	0.022	0.002		
QRR Program Logic Score				0.003					0.000	
QRR Economic Analysis Score				-0.005 *					-0.002 **	
QRR Monitoring Score				0.013					0.004	
QRR Evaluation Score				-0.001					-0.000	
OPC Program Logic Score					-0.019 **					-0.003 *
OPC Economic Analysis Score					-0.014 *					-0.005 **
OPC Monitoring Score					0.016					0.004
OPC Evaluation Score					-0.000					0.000
Project complies with FOE	5.788	5.525	0.240	-0.202	-0.159	22.287 **	31.244	0.163	0.186	0.205
Ratification occurred for this project	-0.217 **	-0.146	-0.043	-0.027	-0.008	-0.158	-0.135	-0.000	-0.005	-0.000
Delay from project approval to first disbursement (Da	0.004 ***	0.004 ***	0.001 ***	0.002 ***	0.002 ***	0.004 ***	0.003 ***	0.000 ***	0.000 ***	0.000 ***
Project approved amount [ln(millions of USD)]	-0.004	-0.006	-0.002	-0.010	-0.009	0.003	-0.005	-0.000	-0.001	-0.001
Operation Sub-Type = ESP or INO or TCR	-0.039	-0.037	-0.001	0.001	0.000	-0.193	-0.171	-0.007 *	-0.005	-0.005
DEM format is 2010 or later	-0.098 *	-0.069	-0.033 *		0.000	-0.222	-0.233	-0.009		0.000
Number of years between approval and PMR		-0.102 ***	-0.053 ***	-0.088 ***	-0.092 ***		-0.103 **	-0.010 ***	-0.017 ***	-0.018 ***
Mean number of years between approval and PMR	-0.118 ***					-0.119 **				
% Change in Real GDP Per Capita	-0.272	0.274	-0.075	0.578	0.481	-0.883	0.412	-0.044	0.086	0.078
Variance in Real GDP Per Capita Growth Rate	-0.002	0.001	0.000	0.005	0.004	0.005	0.009	0.001	0.001	0.001
Land area per population (km2/person)	25.841	25.879	-9.292	-5.540	-4.952	98.506 **	43.119	-8.645 *	11.014	11.901
Foreign exchange instability (consecutive months)		-0.045	-0.013	0.032	0.013		-0.045	-0.000	0.008	0.005
Mean foreign exchange instability (consecutive mon	-0.066 *					-0.083 *				
Division = CMF	0.054	0.154 **	0.013	0.027	0.016	-0.006	0.177	-0.001	0.010	0.007
Division = CTI	0.037	-0.028	0.009	-0.033	-0.003	0.006	0.037	0.001	0.001	0.006
Division = EDU	-0.016	-0.039	-0.019	-0.013	0.021	0.292	0.335	0.005	0.002	0.005
Division = ENE	-0.034	-0.038	-0.013	-0.036	0.012	-0.004	0.009	-0.008	-0.004	0.005
Division = ICS	-0.015	-0.037	-0.005	0.004	0.009	-0.091	-0.141	-0.006	0.006	0.008
Division = LMK	-0.180 **	-0.198 **	-0.082 *	-0.039	-0.025	-0.278 *	-0.197	-0.020	-0.004	-0.004
Division = RND	0.086 *	0.129 ***	0.026	0.026	0.072 *	0.110	0.240	0.005	0.021 *	0.028 **
Division = SPH	-0.093 *	-0.075	-0.044 *	-0.042	-0.021	-0.144	-0.099	-0.010	-0.001	0.001
Division = TIU	0.122	0.187 **	0.025	0.028	0.034	-0.014	0.057	-0.004	0.010	0.012
Division = TSP	0.026	0.019	0.008	0.018	0.040	0.064	0.161	0.003	0.005	0.008
Division = WSA	-0.012	-0.067	-0.001	0.047	0.068 **	-0.147	-0.206	-0.008	0.010	0.016 **
Argentina	0.502	0.503	0.189	0.104	0.093	2.191 *	3.241	0.197 *	0.256	0.277
Bahamas, The	-0.879	-0.791	-0.302	-0.199	-0.156	-3.963 **	-5.590	-0.335 *	-0.429	-0.457
Barbados	-0.320	-0.385	-0.096	-0.018	0.003	-1.233 *	-1.935	-0.111	-0.153	-0.161
Belize	0.773	1.104	0.296	0.445	0.431	4.314 *	7.215	0.306 **	0.433 **	0.458 **
Bolivia	-3.910	-3.783	0.399	0.538	0.449	15.418 **	21.458	0.402 *	0.529 *	0.565 *
Chile	0.328	0.232	0.107	-0.012	-0.002	0.621 *	0.519	0.035 **	0.023	0.027
Colombia	-0.418	-0.510	-0.152	-0.087	-0.082	-1.575 *	-2.531	-0.145 *	-0.186	-0.203
Costa Rica	-0.612	-0.848	-0.246	-0.154	-0.144	-2.870 *	-4.427	-0.259 *	-0.321	-0.349
Dominican Republic	-0.724	-0.848	-0.313	-0.249	-0.216	-3.284 *	-5.076	-0.311 *	-0.407	-0.439
Ecuador	-0.622	-0.726	-0.211	-0.120	-0.097	-2.372 **	-3.654	-0.208 *	-0.270	-0.293
El Salvador	-0.746	-0.833	-0.308	-0.188	-0.164	-3.138 *	-5.172	-0.304 *	-0.410	-0.444
Guatemala	-6.755	-6.591	-0.588	0.000	0.000	25.849 **	36.330	-0.482 *	0.000	0.000
Guyana	0.000	0.000	1.792	0.966	0.864	0.000	0.000	1.747 *	2.157	2.340
Haiti	-0.869	-0.944	-0.314	-0.192	-0.204	-3.689 *	-5.569	-0.325 *	-0.417	-0.459
Honduras	-6.326	-6.163	-0.448	0.025	-0.024	24.741 **	34.970	-0.387 *	-0.481	-0.530
Jamaica	-0.923	-0.890	-0.300	-0.140	-0.101	-3.664 **	-5.240	-0.314 *	-0.394	-0.423
Mexico	-0.755 *	-0.794	-0.243	-0.135	-0.102	-2.541 **	-3.754	-0.213 *	-0.277	-0.291
Nicaragua	-6.187	-6.037	-0.418	0.097	0.055	24.104 **	34.122	-0.342 *	-0.412	-0.453
Panama	-0.006	-0.073	-0.002	-0.075	-0.067	0.208	0.148	0.010	-0.002	-0.000
Paraguay	-0.407	-0.603	-0.164	-0.189	-0.180	-1.870 *	-3.124	-0.178 *	-0.245 *	-0.264 *
Peru	-4.891	-4.967	0.000	0.000	0.000	19.930 **	28.389	0.000	0.000	0.000
Suriname	6.920 *	6.945	2.482	1.579	1.427	25.883 **	37.405	2.218 *	2.814	3.031
Trinidad and Tobago	-0.690	-0.632	-0.262	-0.217	-0.175	-1.793	-3.369	-0.275	-0.405	-0.434
Uruguay	0.313	0.265	0.121 *	0.055	0.059	1.154 **	1.452	0.100 **	0.115	0.127 *
Venezuela, RB	-0.619 ***	-0.576 **	-0.270 ***	-0.398	-0.385	-1.383 ***	-1.768 *	-0.140 **	-0.204 **	-0.216 **
Number of observations	497	497	853	413	413	497	497	853	413	413
Adjusted R-Square	0.290	0.264				0.129	0.129			
corr(Y,E(Y|X))	0.601	0.581	0.54	0.553	0.563	0.465	0.465	0.512	0.535	0.532

*Table reports marginal effects. Asterisks denote significance, * p<0.10, ** p<0.05, *** p<0.01.*