Public Disclosure Authorized

Public Disclosure Authorized

WPS7983

Policy Research Working Paper 7983

Special Study on Benchmarking the Quality of Project Economic Analysis for the South Asia Region

Kene Ezemenari
Xiao Ye

WORLD BANK GROUP
South Asia Region
Office of the Chief Economist
February 2017

Policy Research Working Paper 7983

Abstract

This paper benchmarks the quality of project economic analysis in South Asia against other World Bank regions, using data on project exits between 1975 and 2015. The results show that the South Asia region performs on par with the other regions, in that the share of project documents that report estimated economic rates of return have declined from 70 to 36 percent for South Asia and the other regions. This finding suggests there is less attention to project economic analysis (especially for sectors where this has been a traditional practice, such as energy, transport, water, and agriculture). The finding also indicates that the incidence of reporting rates of return in project documents and the dispersion or difference between rates of return estimated at project appraisal and completion are significantly correlated with project performance (after correcting for country- and project-level variables). For the project-level variables, the task team effect is a key variable that explains project outcomes. The paper discusses the implications of the analyses for strengthening project performance and risk mitigation.

This paper is a product of the Office of the Chief Economist, South Asia Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at KEzemenari@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

SOUTH ASIA REGION

Office of the Chief Economist (SARCE) and Development Effectiveness Unit (SARDE)

Special Study on
Benchmarking the Quality of Project Economic Analysis for the South Asia Region

By

Kene Ezemenari and Xiao Ye

Key Words: Project Performance, Project preparation, Aid Effectiveness, World Bank.

JEL: F35, H43

Table of Content

This paper was prepared under the overall guidance of Martin Rama. The authors gratefully acknowledge comments provided by Aart Kraay and Carolina Monsalve; assistance provided by Kaoru Chikushi, Christian Gonzalez Amador, Kwadwo Kusi, Floyd Goodman and Sadaf Alam in accessing data from the Bank's information systems; and Gloria Kwembe for administrative assistance.

1. Introduction

This paper assesses whether there continues to be a declining trend in the quality of economic analysis in World Bank projects. Project economic analysis (or cost-benefit analysis) became an integral part of applied economic analysis in the 1960s. For the two or so decades that followed, the Bank was known for its strong attention to project economic analysis. However, by the 1980s, the quality and attention to project economic analysis had begun to wane.

An assessment of the declining trends in the quality of economic analysis is important because there is strong evidence to indicate that projects with good economic analysis perform better. This paper therefore examines the links between trends in the quality of project economic analysis and the performance of Bank projects. The paper benchmarks the project performance of the South Asia Region against the other regions of the Bank.

The rest of the paper is organized as follows: section 2 presents a brief overview of the literature on the quality of economic analysis in Bank projects and the determinants of project performance; section 3 outlines the methodology for assessing the trends in quality and links this to project performance; section 4 describes the database used in the analysis. Sections 5 presents an overview of the results and section 6 concludes.

2. Economic Analysis in Bank Projects

Project economic analysis (EA) or cost-benefit analysis (CBA) was central to project preparation in the 1960s, 1970s and 1980s. The assumption was that markets would function efficiently in the absence of government-induced distortions that arose from public ownership of assets, price controls, taxes, subsidies, quotas, etc. Less emphasis was placed on distortions arising from market or government failures such as credit constraints in education, agglomeration and urban congestion. This approach ran counter to the premise for development finance, which is to do what the private sector will not or cannot do.

During this time, various economists developed approaches to account for government-induced distortions (Little and Mirlees 1974, 1969; Dasgupta Marglin, and Sen 1972; Harberger 1973; Squire and van der Tak 1975). For example, Little and Mirrlees (1974) developed methodologies based on border prices for tradeable goods that required converting domestic prices of nontraded inputs and outputs into their border-price equivalent. Standard conversion factors were used in cases where the commodity specific conversions were not possible. The methodology allowed for different weights to be used for different groups. It allowed consumption and investment benefits, and benefits to the private and public sectors to be valued differently. Squire and Van der Tak (1975) further refined the Little and Mirrlees framework in three ways. First, when the goal was income redistribution, weights were to be assigned to project benefits and costs according to whom they accrued. Second, greater weight was placed on changes in investment caused by a project relative to that caused by consumption if there were distortions in capital markets that made the unit cost of decreased consumption less than the unit cost of reduced investment. Third, the cost of finance or maintaining a project was adjusted for the deadweight loss of raising taxes.

By the 1990s, the wide use of the above approaches in the Bank had declined as the extent of government involvement in the production of goods and services lessened (Little and Mirrlees 1991, p. 360). Lower levels of government-induced distortions in the 1990s relative to the 1970s called into question the primary focus on correcting distortions to measure benefits and costs in project economic analysis. Critics noted that more attention was needed on assessing the sustainability of projects and on performance. Specifically, there was insufficient attention to identifying factors that might cause underperformance of the project including political economic consideration and other design issues that ensure resilience to economic, budgetary and policy changes (Jenkins 1997). Devarajan et al. (1995 and 1997) stressed the importance of the counterfactual (or what would happen in the absence of government engagement) and noted that with the declining involvement of government the focus should be on whether the project should be undertaken by the private or public sector. They also highlighted the challenge of fungible financing. For example, if the project being appraised would have been taken up without donor funds, then it was likely that donors were actually indirectly financing some other project that had not been appraised. As a result, they recommended sectorwide expenditure reviews that assessed the likelihood of the project succeeding and linked to the project-specific appraisals. Sectorwide expenditure reviews would help define the private-sector counterfactual and give a sense of the fiscal impact.

In 2010, an IEG evaluation of the quality of project economic analysis reinforced the recommendations put forth by Devarajan et al. (World Bank 2010). The IEG study documented the declining trends in the frequency of project economic analysis and suggested that this stemmed mainly from the shifts in the portfolio toward more programmatic and economywide approaches (away from traditional investment projects). However, the study also documented the within sector decline in the frequency and use of project economic analysis, particularly for those sectors where this has traditionally been strong (i.e. agriculture, energy, transport, water, etc.). This result pointed to other factors at play in determining the declining trends in the frequency of application of project economic analysis. For example, a likely key factor noted by Devarajan et al. (1995, p. 3) pertained to changes in the process for reviewing the quality of project economic analysis that had become slack over time. Drawing on the results of both the IEG evaluation and Devarajan et al., following the IEG evaluation, the Bank issued updated guidance to teams that placed emphasis on: (i) defining the counterfactual and the rationale for public provision; (ii) assessing the fiscal impact and sustainability; (iii) and examining the impact on poverty reduction.

This paper examines more recent trends in project economic analysis since the Bank's updated guidance was issued and links these trends to project performance. A growing literature on the determinants of project performance highlights the importance of controlling for both country and project level characteristics. For example, GDP growth and GDP per capita are commonly used in these regressions and seem to be negatively associated with project outcomes when there are no controls for project level characteristics (Dollar and Leving 2005; Guillamont and Lajaj 2006; Dreher, Klasen, Vreeland and Werker 2010). In contrast, GDP growth and GDP per capita are each positively associated with project outcome measures for studies that control for both country and project level variables (Denizer, Kaufmann and Kraay 2013; Chauvet Collier Duponchel 2010; Kilby 2000). Denizer et al. estimate that country and project level variables explain around 20 percent of the total variation in project performance. In particular, Denizer et al. 2013 include a detailed set of project level variables that capture project length, size, supervision and preparation costs, whether the project was restructured and dummy variables to capture early warning of problems prior to completion. They find that if a project indicates it is problematic in the first half of its life, this is associated with a lower level of performance at the end of the project. Similarly, if a project has been restructured in the first half of its life, this is associated with lower performance at completion (when the Bank has fully disbursed all funds committed for the project). Denizer et al. also introduce a key variable (not used in prior studies) that proxies for the quality of the task team (measured by the average rating over all other projects associated with the team leader, with the exception of the current project). The result is that projects led by task teams that have led well performing projects in the past tend to perform better.

There are very few papers that explicitly link the quality of project economic analysis to project performance. Deininger, Squire and Basu (1998) focus on the stock of economic analysis defined more broadly. To account for the fact that macroeconomic analysis is likely to affect all projects in a given country while sector-specific analytical work has a more limited impact, they use the sum of all staff-weeks used to produce both macroeconomic analytical work for the country in which the project is implemented and all sector-specific analytical work for the sector in which the project falls. They find that an additional 100 staff-weeks spent on ESW significantly increases the probability of a satisfactory project rating by between 12 and 20 percent. The study adjusts for macro level variables (inflation, openness, fiscal surplus) and staff weeks spent on project supervision. These variables are all found to have an insignificant effect on project performance. Deininger et al. (1998) capture the impact of the overall stock of economic analysis on project performance; but, it is not clear to what extent the concept of quality is being captured, particularly as it applies to project economic analysis.

Belli and Pritchett (1995) (as cited in Vawda, Moock, Gittinger, and Patrinos 2001) define subjective measures of good project economic analysis (based on a desk review and ratings of project documents). Projects are then followed over the cycle to monitor performance. A logit model is used to link project performance to the subjective ratings of the quality of economic analysis. The results show that if the economic analysis was poorly done prior to approval, the probability that the project would perform unsatisfactorily by the third year after implementation was seven times higher than that of a project with good economic analysis. By the fourth year, the probability of failure was 16 times higher than the corresponding probability for a project with a good project economic analysis.

This paper uses two proxy measures for the quality of project economic analysis. The first proxy is based on the frequency or incidence of economic rate of returns (ERRs) in the portfolio and relates the presence of an estimated ERR to the performance of the related project. This measure captures trends in the application of project EA in the portfolio. However, it does not give a sense of the quality of the EA or whether project EA is the most appropriate type of analysis for a particular project. The second proxy is more geared to quality of EA at the project level. It examines the discrepancy or difference between the ERR at appraisal and that estimated at completion (or when the Bank has completed disbursements of funds to a country for a particular project). Some authors have characterized the discrepancy between ERRs estimated at project appraisal and completion as optimism bias (Herrera 2005; World Bank 2010). In contrast, Pohl and Mihaljek (1992) examined the impact of various project level variables related to cost overruns and implementation delays and found that this explained very little of the divergence in ERR between appraisal and completion. They concluded that much of the difference in ERRs estimated at appraisal and completion is due to risk and uncertainty about expected prices and growth rates used in the economic analysis. The IEG 2010 report on the quality of economic analysis noted that "…Downside risks are systematically ignored, and as a result projected ERRs are biased upwards."

In view of the above discussion, it could be argued that some of the differences (in reported ERRs between appraisal and completion) are due to the quality of assumptions, data and estimation techniques which give rise to sustained and large differences in estimated ERRs between project appraisal and completion.

Specifically, World Bank (1995) noted that the sensitivity analysis did not fully capture likely risks. It lacked a consideration of the risk of delays in implementation or achievement of project benefits associated with institutional reforms. The report noted that neglect of this risk in the analyses imparts an upward bias to ERRs at appraisal and is a major cause of the gap between appraisal and completion ERRs. This paper therefore draws on this conclusion to define a project specific measure of the quality of economic analysis based on the discrepancy between ERRs estimated at appraisal and completion.

The current paper contributes to the literature on the determinants of project performance in the following ways. First, to our knowledge it is the first to use the IEG data set on project exits to link proxies for the quality of project economic analysis to project performance. Second, it exploits the difference between the estimated economic rate of return (ERR) at appraisal and completion as an additional measure of quality that is linked to project performance. The methodological approach adjusts for optimism bias and risks arising from assumptions made about trends in prices and economic growth rates, as well as both country and project level effects.

3. Methodology

This section outlines the approach for examining the links between the quality of economic analysis and project performance. It also examines the correlates of the quality of economic analysis. The paper uses regression analysis and draws from the functional forms used by Denizer et al. The quality of project economic analysis (EA) is defined in terms of whether a particular project has estimated a rate of return both at appraisal and completion; and the discrepancy or dispersion between ERR estimates at appraisal and completion, as follows:

- Aj = 1 if an economic rate of return (ERR) has been reported both in the Project Appraisal Document (PAD) and the Implementation Completion Report (ICR); it is 0 otherwise. The presence of an estimated ERR in both the PAD and ICR is used as a proxy to indicate that more attention has been put toward analyzing the economic performance of the project than would be the case if no ERR had been estimated, or if an ERR was reported only in the PAD or ICR. The implication is that there has been attention to ensuring that the needed information is available to conduct the EA in a timely fashion.
- Dj = The absolute percentage change between ERR estimated at appraisal and ERR estimated at completion. It is postulated that appropriate assumptions and proper attention to risks and sensitivity analysis should yield ERRs or expected ERRs that are close in estimated levels between the PAD and ICR. The implication being that an analysis of the expected impact of risks on the estimated ERR should also inform project design, monitoring and implementation and that this may contribute to improved risk management and better project performance.

The above measures are linked to project performance (Pj and Sj) using regressions that control for country, project and task team level variables that may also affect project performance. Specifically, project performance is defined as a function of the quality of economic analysis (Aj or Dj), controlling for project and country characteristics:

Project performance (Pj or Sj) = f (quality of project economic analysis, Xj, Lj, Cj, Tj).

Given the above, measures of project performance, project characteristics, and country characteristics are defined as follows:

Measures of project performance:

Two measures are used.

- Pj = Success of Project Outcome (PO), is rated 1 if the project is rated moderately satisfactory or higher, and 0 otherwise, available from 1985 and onwards.
- Sj= Success of Project Outcome (PO), is rated on a scale from 1 (highly unsatisfactory) to 6 (highly satisfactory), available from 1995 and onwards.

Measures of project characteristics

- Xj = Project characteristics (sector, size of commitment, preparation and supervision costs each as a share of total commitment; project duration as defined by years from appraisal to completion), whether the project was ever flagged as having a problem in the first half of its life, whether a project was flagged as having a potential problem in the first half of its life, whether the project was restructured in the first half of its life, a dummy variable that is assigned one if there was an in-depth evaluation by IEG, etc.).
- Tj = Is a time trend variable based on either the year the project was approved (for regressions that include the dummy variable for whether an ERR is reported in the PAD or in both the PAD and ICR); or the year in which the project exited/completed disbursements (for regressions that include the dummy variable for whether an ERR is reported in the ICR only).
- OBj=Is a proxy for optimism bias represented by a dummy variable which is 1 if the ERR at appraisal is greater than the ERR estimated at completion and 0 otherwise. This is entered as an independent variable for regressions which include the discrepancy between ERR at appraisal and completion.

Country or macro level characteristics

- Cj = Country or macro level variables (GDP per capita growth and CPIA,[1] each averaged over the life of the project, dummy for the region in which the project is located, dummy for the period before 1998 when a change occurred in the CPIA rating scale, dummy for sector in which the project is located, interaction terms between the sector dummy and the approval year dummy, and the project evaluation year dummy, respectively.

Team leader characteristics

- Lj = captures task team leader (TTL) effects, defined as the average ratings for project outcomes over all previous projects that the current TTL managed.

Given the above variable definitions, three key questions are examined:

(i) Is there a continued decline in the frequency of project economic analysis (EA)?

[1] We also tried to use World Wide Government Indicators such as government effectiveness, corruption, regulatory quality, which are, however, highly correlated with each other; these did not seem to explain the probability of the project success/failure. Therefore, they are not included in the reported regressions. GDP per capita is averaged over the years covering project appraisal up to the project exit year. CPI is used to explain the discrepancy between ERR at appraisal and project completion, which is also measured as the average for the project period up to the project exit.

(ii) Is there a trend in the discrepancy or divergence between ERR estimated at appraisal and completion?

(iii) Does the quality of project economic analysis (as defined by the two previous questions) matter to the success of a project, after controlling for country, project and team leader characteristics?

To examine the above questions, three models are estimated, namely:

I. Is there a continued decline in the frequency of project economic analysis (EA)?

Model I (Logit): Aj = f(Xj, Lj, Cj, Tj), where

Aj represents ERR presence (i.e. whether it has been reported in the PAD and/or ICR). The three specifications of the dependent variable are used to examine the determinants of ERR reporting, as follows: (i) Aj set to 1 if an ERR appears in the PAD for a particular project; (ii) Aj set to 1 if an ERR appears in the ICR for a particular project; (iii) Aj set to 1 if an ERR appears in both the PAD and ICR for a particular project; Aj is otherwise 0 across the various specifications.

Xj (project characteristics) and Cj (country characteristics) are as defined above. Sector dummies are specified for high-CBA projects with low-CBA projects as the default/benchmark. Regression results are summarized in Tables A.2 to A.4 below.

II. Is the discrepancy or divergence between ERR estimated at appraisal and completion worsening?

Model II (OLS): Dj = f(Xj, Lj, Cj, OBj, Tj), where

Dj is the ratio of the absolute difference between ERR reported at completion (ICR) and at appraisal (PAD), divided by the ERR reported at appraisal[2]: $R_j = log \left|\frac{ERR_{ICR} - ERR_{PAD}}{ERR_{PAD}}\right|_j$;

For both Models I and II above, the following subset of variables are used in the regression:

Xj (project level characteristics) = log commitment, project length, closing extension, and preparation cost as a share of total commitment; the regression also attempts to correct for optimism bias by including a dummy variable that is 1 if the ERR reported in the PAD is greater than that reported in the ICR.

Cj (country characteristics) = Institutional characteristics (captured through the average annual CPIA rating; the CPIA provides an assessment of country-level policies and institutions that may have a bearing on the country's overall level of development which in turn may have a bearing on project performance). A dummy variable captures the change in the CPIA scale from a 5-point scale (ranging from 1 to 5) to a 6-point scale (ranging from 1 to 6); it also includes a time trend. Sector dummies indicate whether the project is in the agriculture, water, energy, transport or urban development sectors that traditionally have

[2] For regressions with the team leader effect as one of the explanatory variables, the sample size for the regressions is limited to the projects that have the information on team leaders. The team leader track record is defined as the average ratings of World Bank project outcomes for a team leader over all his/her projects in the sample, excluding his/her current project rating. Therefore, the TTL who had only one project in the sample is dropped out from the regression analysis because he/she did not have a track record. The absolute difference in the value of ERR estimated at Appraisal and Completion is used to measure divergence because the focus is on the degree of divergence between the two estimates, and not the sign of the divergence. Estimation of Model II using OLS is limited to the High-CBA projects that report ERRs at Appraisal and Completion only, resulting in a significant reduction in the sample size.

a well-developed methodology that is consistently applied for project economic analysis, or "high-CBA" sectors. Regression results for this model are summarized in Table A.11. Model 2 consists only of high-CBA projects which account for over 90 percent of the projects that report an ERR in both the PAD and ICR. The default sector for the sector dummies is the transport sector.

Lj = the average rating for other projects previously led by the same task team leader.

III. Does EA matter to the success of a project, after controlling for country, project and team factors?

Model III regressions are estimated using both logit (when project outcomes are rated as successful/unsuccessful) and OLS (when project outcomes are rated on a scale from 1 to 6). Models III (a) and (b) include two definitions of the quality of economic analysis (Aj=1 if ERR is reported in both the PAD and the ICR for a particular project and 0 otherwise; and Dj= the absolute difference between ERR in the PAD and ICR as defined above):

Model (Logit) III a. Pj = f(Xj, Lj, Cj, Tj, Aj)

Model (Logit) III b. Pj = f(Xj, Lj, Cj, Tj, Dj)

Similarly, Models III (c) and (d) include the above two definitions of the quality of economic analysis, and the same functional form but using the OLS estimation method:

Model (OLS) III c. Sj = f(Xj, Lj, Cj, Tj, Aj)

Model (OLS) III d. Sj = f(Xj, Lj, Cj, Tj, Dj).

One shortcoming of the above measures is that they assume that a CBA is conducted for all investment projects (or IPFs); or that a CBA is the most appropriate form of economic analysis for all projects. This assumption also does not examine whether a CBA (and the ERR) has been done in a way that is most appropriate for a particular IPF. The indicator therefore does not account for those projects for which estimation of an ERR would not be the most appropriate form of economic analysis. This note attempts to partially adjust for this by distinguishing between the high CBA sectors and the low CBA sector projects (i.e. education, health, governance, etc.), by using dummy variables to categorize low-CBA sectors as default sectors. The share of low CBA projects has been increasing gradually from 34 percent in the full sample to 45 percent in the sample that includes projects completed in 1995 and thereafter.

4. Data Description

Project data used for this study are taken from the Independent Evaluation Group (IEG) World Bank Project Performance Ratings data set, which covers 1975 to the present. The data set contains project variables on IEG evaluated/validated project outcome ratings, year of approval, exit fiscal year, net commitment, sector, and estimated ERRs reported in the PAD and ICR. Other project level variables such as preparation and supervision costs, TTLs' identification numbers, potential problems, actual problems and restructuring of a project during operations are taken from the World Bank's management information system (or Business Intelligence –BI- database). Data on GDP per capita and inflation are taken from World Development Indicators (WDI) for the respective years and were averaged over the period when each project was in operation. The World Bank's Country Policy and Institutional Assessment (CPIA) is used as a proxy measure for the extent of governance or institutional development and is also averaged over the years in which the project was in operation.

Although the IEG data set on project exits covers the period 1975 to 2015, this paper focuses primarily on the years 1995 to 2015 (and in some instances goes back to 1985), to coincide with major reforms to the project assessment system that took place after the mid-1980s. Specifically, the project rating system was introduced in the 1980s. Initially the rating system was based on a 2 point scale (i.e. satisfactory/not satisfactory). The ratings were then expanded to a four-point scale in 1993, including "Highly unsatisfactory" and "Highly satisfactory". In 1994 this scale was further expanded to the current six point scale including Moderately Unsatisfactory and Moderately Satisfactory. Prior to 1980, project performance reviews did not include outcome ratings. However, imputed values for project performance (Successful/Unsuccessful) prior to 1980 were derived in the mid-1980s when the IEG database on project exits was established. During this time, "early warning" variables or flags to monitor the risk of the project not meeting its stated objectives were also introduced.

Over the period 1970-1980, project performance reviews were summarized in Project Performance Audit Reports (PPARs), which were prepared by IEG. In 1982, Project Completion Reports (PCRs) were introduced, which were prepared by task teams. IEG reviewed these PCRs, and submitted them to the Board with an IEG cover note that summarized its independent assessment and performance ratings. In addition, during the period 1994-1995, the Implementation Completion Report (ICR), a self-assessment instrument currently in use, was introduced, wherein task teams assessed the degree to which a project had achieved its intended objectives. Also at this time, IEG began to review and validate all ICRs and for a random sub-sample would also conduct a more detailed performance audit (from which Project Performance Audit Reports are produced) that typically include data collection at the project site and are usually completed 3 to 4 years after the initial ICR has been completed by the task team. As a result, this paper uses the PPAR for a particular project, to measure performance; the ICR rating is used in cases where no PPAR has been completed for a particular project.

Denizer, Kaufmann and Kraay (2013) provide more details on the data generating process of the IEG data set and also note some methodological issues that may have a bearing on the regressions that link the quality of economic analysis to project performance. The first point to note, however, is that unlike Denizer et al., who include budget support operations in their analysis, this paper is limited to investment financing projects, which account for 82 percent of total projects and 61 percent of total financing in our sample between 1995 and 2015, because these are the operations that undertake the type of economic analysis that is the focus of this paper.

Denizer et al. make note of three key issues that may affect the regression results: First, project ratings are based on success in attaining the stated objective of the project rather than a common standard across all projects and over time. The rationale for this may be partly due to the difficulty in setting common standards across projects in different sectors (i.e. roads relative to teacher training, or civil service reform projects). However, the lack of a common standard across projects and over time may lead to bias in the regression estimates. In addition, the standards for setting development objectives and evaluating success relative to a given objective may have changed over the past 30 years. To account for these potential effects, we follow Denizer et al. to construct a set of dummy variables that correspond to the five-year period over which the project was evaluated (i.e. 1980-84, 1985-89, 1990-94). These dummies are included along with their interactions with a set of sector dummies.

Another set of concerns pertains to the objectivity of the ICRs which may reflect the views of the task team leader who may not be completely candid about the project shortcomings. Denizer et al. show that there is no significant difference in ratings on average between the team leader's assessment and the IEG's validation/audits. We also get the same results using a data sample that includes more recent years (beyond the year 2011 covered in Denizer et al). However, to control for this potential effect, we follow the approach of Denizer et al., who use dummy variables for the type of evaluation to capture the possible differences in the average outcome ratings between IEG's PPAR and the task teams' ICR ratings.

Finally, Denizer, et al. note that although the subjectivity in the rating of IEG evaluations themselves may also be a source of error, several factors point to their plausibility. Specifically, IEG is independent of the rest of the Bank's management and reports directly to the World Bank's Board of Executives. Its review procedures are independent and its experienced staff draw from cross-country and cross-project experience to inform specific project assessments and apply common standards. In addition, most IEG evaluations have been public since the 1990s and IEG pays close attention to criticism and comments from external experts, civil society and academia. However, IEG is staffed by current Bank staff and future Bank staff as there is some rotation in and out of IEG, although this is considerably lower than in other parts of the Bank. There are also likely to be informal sources of communication between IEG and World Bank staff which could affect the ratings process. Nevertheless, the outcome ratings capture the overall experience and insights of World Bank and IEG staff on how well projects are performing. Finally, the data are not likely to capture the overall or aggregate impact of aid given there are likely complementarities between projects as well as the potential scope for crowding in or out public spending.

Table 1 provides summary statistics for variables used in the regression analyses. The summary statistics are divided according to the three main periods analyzed in this paper: 1975-2015 to examine long-term trends; the period 1985-2015 (to examine more closely the drivers of the trends over the period when the Bank initiated ex-post rating of project performance); and 1995 to 2015 used for most of the regression analysis. A more restricted sample is analyzed also that covers projects that reported ERRs in both the PAD and at ICR. Project performance is also defined in terms of a binary variable (that was in effect over the period 1975 to 1995) and a categorical variable (that came into effect from 1995).

Overall, the sample characteristics are similar across the various time periods (1975 to 2015; 1985 to 2015 and 1995 to 2015). In addition, regression analysis using data from 1985 to 2015 (the earliest date for which there are complete data) yield very similar results to the regressions covering the 1995 to 2015 period, which are presented in the Annex tables.

Two things are noticeable. First, even though Denizer et al. include budget support (or development policy lending) in their regressions (and we do not), our summary statistics are consistent for similar variables. Second, the differences between our full and restricted samples are statistically significant for some variables but not for others using two-sample t-statistic tests (Table 1b). For project duration, CPIA, and inflation, the means are not statistically different from each other for the years 1975-2015. The differences for CPIA increase in later years and become statistically significant. Additionally, those projects that reported ERRs in both the Project Appraisal Document (PAD) and the Implementation Completion Report (ICR) had on average larger commitment, higher ratings in project outcomes, higher GDP growth rates, and lower spending (as a share of total commitment) for project preparation and supervision. Projects that report an ERR in the PAD and ICR (see Table 1) are rated higher on average relative to those in the full sample. These differences provide some support for the hypotheses that (on average) projects that put more attention on the economic analysis (i.e. report an ERR in the PAD and ICR as a proxy for quality of economic analysis) tend to perform better.

Table 1a: Summary Statistics

Full Sample	1975-2015 Mean	Std. Dev.	# of obs	1985-2015 Mean	Std. Dev.	# of obs	1995-2015 Mean	Std. Dev.	# of obs
Project Outcome Ratings									
IEG Satisfactory (1)/Unsatisfacotry (0) Rating	0.72	0.45	7,386	0.70	0.46	5,963	0.72	0.45	4,021
IEG 6-point Rating							4.00	1.17	4,021
Evaluation Characteristics									
Years between completion and evaluation	2.31	1.79	7,085	2.12	1.57	5,718	1.71	1.22	3,799
Dummy=1 for detailed PPAR Review	0.36	0.48	7,436	0.26	0.44	6,006	0.18	0.38	4,052
Macro Variables (averaged over project life)									
Log of Real GDP per capita growth, averaged over proj.	2.58	3.68	7,188	2.58	3.76	5,895	3.19	3.79	4,038
Log CPI averaged over project life	2.36	1.25	6,878	2.30	1.34	5,704	2.13	1.17	3,916
CPIA score (1=bad, 6=good)				3.40	0.70	5,924	3.56	0.53	3,992
Basic Project Characteristics									
Logaritnm of project net commmitment (US m$)	3.15	1.26	7,436	3.28	1.26	6,006	3.39	1.26	4,052
Years from Approval to Completion	6.31	1.86	7,436	6.44	1.87	6,006	6.49	1.93	4,052
Prepartion Costs/Commitment, (%)	1.88	4.72	7,436	2.22	5.12	6,006	2.59	5.66	4,052
Supervision Costs/Commitment, (%)	2.64	4.82	7,392	3.04	5.22	6,005	3.70	5.82	4,051
Dummy=1 if EA present in both PAD and ICR	0.39	0.49	7,436	0.33	0.47	6,006	0.28	0.45	4,052
Logarithm of the absolute difference between the ERR in PAD and that in ICR	-1.39	1.39	2,880	-1.41	1.34	1,993	-1.52	1.34	1,122
Share of High-CBA projects, %	66.0			63.3			55.4		
Early Warning Indicators									
Dummy=1 if project was restructed in first half							0.06	0.23	3,853
Dummy=1 if potential problem project flag in 1rst half				0.34	0.47	4,967	0.32	0.47	4,041
Limisted Sample	**1975-2015 Mean**	**Dev.**	**obs**	**1985-2015 Mean**	**Dev.**	**obs**	**1995-2015 Mean**	**Dev.**	**obs**
Project Outcome Ratings									
IEG Satisfactory (1)/Unsatisfacotry (0) Rating	0.78	0.42	2,880	0.77	0.42	1,993	0.82	0.39	1,122
IEG 6-point Rating							4.30	1.06	1,122
Evaluation Characteristics									
Years between completion and evaluation	2.36	1.66	2,752	2.16	1.51	1,907	1.53	0.88	1,052
Dummy=1 for detailed PPAR Review	0.42	0.49	2,880	0.27	0.44	1,993	0.12	0.33	1,122
Macro Variables (averaged over project life)									
Real GDP per capita growth, %	2.76	3.34	2,735	2.83	3.36	1,935	3.67	3.27	1,116
Log CPI averaged over project life	2.38	1.16	2,621	2.28	1.29	1,896	2.12	1.22	1,098
CPIA score (1=bad, 6=good)				3.42	0.72	1,967	3.61	0.51	1,110
Basic Project Characteristics									
Logaritnm of project net commmitment (US m$)	3.40	1.19	2,880	3.69	1.15	1,993	3.94	1.11	1,122
Years from Approval to Completion	6.32	1.82	2,880	6.56	1.82	1,993	6.68	1.87	1,122
Prepartion Costs/Commitment, (%)	1.01	2.18	2,880	1.31	2.46	1,993	1.47	2.41	1,122
Supervision Costs/Commitment, (%)	1.40	2.34	2,860	1.69	2.65	1,993	2.02	2.92	1,122
Dummy=1 if EA present in both PAD and ICR	1.00	0.00	2,880	1.00	0.00	1,993	1.00	0.00	1,122
Logarithm of the absolute difference between the ERR in PAD and that in ICR	-1.39	1.39	2,880	-1.41	1.34	1,993	-1.52	1.34	1,122
Share of High-CBA projects, %	97.1			95.9			93.1		
Early Warning Indicators									
Dummy=1 if project was restructed in first half							0.06	0.24	1,037
Dummy=1 if potential problem project flag in 1rst half				0.33	0.47	1,503	0.33	0.47	1,121
Dummy=1 if problem project flag in first half				0.40	0.49	893	0.40	0.49	893

Sources: This table reports summary statistics on measured project outcomes (0/1 or 1-6 scale), as well as summary statistics on all correlates of project outcomes reported in Tables 6 to 15.

Table 1b: T-Statistics for the Difference in Means -between Full and Restricted[1] (ERR in PAD and ICR only) sample

	1975-2015	1985-2015	1995-2015
IEG 6-point Rating			6.50***
Evaluation Characteristics			
Years between Completion and Evaluation	1.00	1.10	-4.20***
Macro Variables (averaged over project life)			
Log of Real GDP per capita growth, averaged over project life	2.30**	2.50***	3.80***
Log CPI averaged over project life	0.80	-0.40	-0.04
CPIA score (1=bad, 6=good)		1.00	2.80***
Basic Project Characteristics			
Logarithm of project net commitment (US m$)	9.10***	12.60***	13.20***
Years from Approval to Completion	0.20	2.50***	2.90***
Preparation Costs/Commitment, (%)	-9.50***	-7.70***	-6.50***
Supervision Costs/Commitment, (%)	-13.20***	-11.10	-9.30***

Source: Staff calculation. *** represents significant level and 1% and ** at 5%.
Note: (1) The restricted sample contains projects that have reported ERRs in both the PAD and ICR.

5. Overview of Results

Both OLS and logit regressions are employed to examine the impact of the quality of economic analysis (correcting for both country and project level factors). Regression results are reported in the Annex and show that country level correlates alone explain around 10 percent of the variation in project performance. Project level factors explain an additional 10 percent of the variation in project performance, similar to results estimated by Denizer et al. (see Tables A.6 to A.9, regression 1).

Country level correlates of project performance. For each project, the average over the life of the project of the CPIA and GDP growth per capita are the key variables used in the regressions. We also include a dummy for when the CPIA rating schedule was modified in 1998 (the scale was changed from 5 to a 6 point scale). Regression results show that this change has no significant impact. Overall, the country level correlates all have the expected signs. A higher level of institutional development and growth is correlated with better project performance.

Project level correlates of performance. This paper focuses primarily on the impact of project economic analysis on project performance. However, the analysis also corrects for other project level effects which might have a bearing on project performance. Specifically, the regressions correct for the size of the project, its duration (the time between project appraisal and completion), the track record of the task team leader (in terms of ratings for previous projects managed by the same team leader), cost of preparation and supervision as a share of the total size (or net commitments for the project), whether the project was rated unsatisfactory in the first half of its life (actual problem project), whether there is potential that the project performance will deteriorate as indicated by the flagging of 3 or more risk factors (potential problem project in the first half of the project life), and whether the project was restructured in the first half of its life. The "early warning variables" (i.e. potential, actual problem project and restructuring) are all negatively and significantly associated with project performance. These results are also consistent with Denizer et al. Overall, the results are robust across all model specifications.

Results for the task team leader (TTL) effect and duration of the project are also robust across all regressions and consistent with Denizer et al. In the case of the TTL effects, the regressions show that the higher the rating for other projects managed by the same task team leader, the higher the project rating of the current project. In the case of project duration, performance declines the longer the time period between approval and completion. It is interesting to note that project duration is positively correlated with the likelihood of reporting an ERR in the PAD (Tables A.2 to A.4); yet, project duration is negatively associated with project performance (Tables A.6 to A.9[3]). An explanation could be that more problematic projects take longer to prepare and require more time for supervision. But, it does not explain the direction of causation. For example, do projects become problematic and as a result require more attention (including to the economic analysis) or is more (time) attention given to projects that are expected to be problematic?

Finally, the project size is positively and significantly correlated with project performance. This could be because larger projects tend to be more high profile and receive more resources and attention. As a result, they perform better. This result contrasts with Denizer et al., who show a negative correlation between project size and performance. Their result may reflect the fact that they include budget support operations in their analysis, while in this paper we examine only investment projects. It may also reflect some missing variable effects. Tables A.6 and A.7 show a positive correlation between the size of the project and project performance when the frequency of ERR reported in both the PAD and ICR is included as an explanatory variable. However, when the discrepancy or divergence between ERRs reported in the PAD and ICR is included as an explanatory variable (Tables A.8 and A.9), the sign of the coefficient associated with the project size is occasionally negative (and insignificant). This suggests some cross effects between the project size and the ERR divergence variable used to proxy the quality of project economic analysis. Regressions in Tables A.2 to A.4 confirm that the size of the project is positively and significantly correlated with the frequency of ERR reporting in the PAD and ICR both individually and jointly. In contrast, when one examines the determinants of the discrepancy between ERR reported in the PAD and ICR (Table A.5), the results show that project size (measured in terms of log commitment) is negatively and significantly correlated with the discrepancy between the ERR reported in the PAD and ICR, seemingly indicating that larger projects are associated with higher quality of EA.

Model 1 Results: Is there a continued decline in the frequency of project economic analysis (EA)?

Ideally, if an ERR has been estimated and reported in the PAD, then an ERR should be reported in the ICR (and vice versa). However, the data indicate that this is not always the case. There are a significant number of cases where an ERR is reported in the PAD but not in the ICR; or where an ERR has not been reported in the PAD but has been reported in the ICR. On average 38.7 percent of the Bank's projects have

[3] Note that multinomial regressions were also estimated for the categorical ratings (1 to 6) used as dependent variables in Tables A.6 and A.8. Results from the multinomial regression mirror the results reported in Tables A.6 and A.8. For ease of interpretation only the OLS results are reported here.

ERR analysis at Appraisal and Completion, but there are variations across regions (Table 2). While SAR and EAP have 48.5 and 47.3 percent, respectively, of projects that report an ERR in the PAD and ICR, in the Africa Region 33.3 percent of projects have reported an ERR in the PAD and ICR.

Table 2: Percent of projects with ERRs at Appraisal or Completion, by Region, 1975-2015

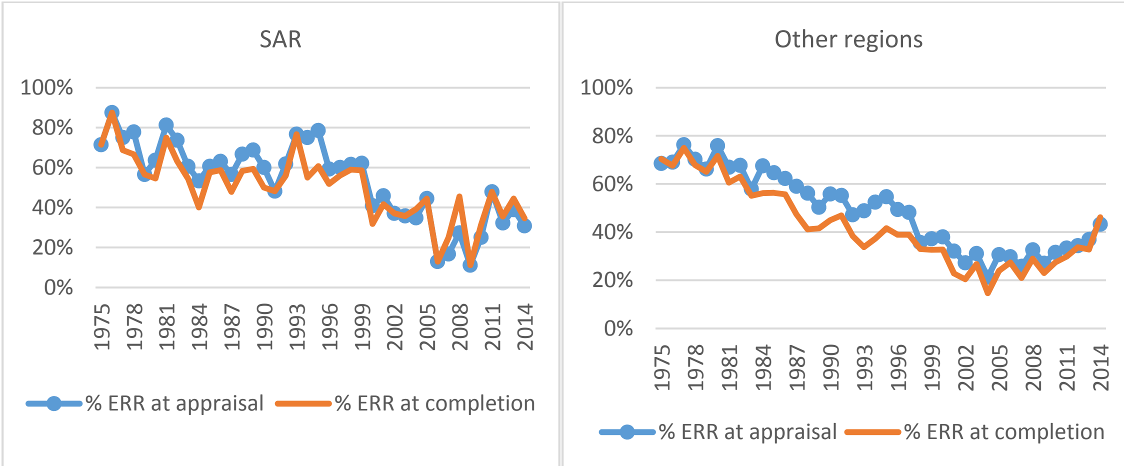
			Percent of total projects (%)				
	Projects with Appraisal ERR and Completion ERR	**Projects with Appraisal ERR only**	**Projects with Appraisal ERR (with/ without Completion ERR)**	**Projects with completion ERR only**	**Projects with either Appraisal or Completion ERR**	**Projects with no ERRs**	**Total number of project**
AFR	33.6	10.0	43.6	2.0	45.6	54.4	2,186
EAP	48.4	6.8	55.2	2.4	57.6	42.4	1,260
ECA	32.7	8.2	40.9	3.0	43.9	56.1	892
LCR	35.1	9.0	44.1	3.1	47.2	52.8	1,501
MNA	41.4	9.4	50.8	1.5	52.3	47.7	660
SAR	47.4	6.2	53.6	2.6	56.1	43.9	937
World Bank	38.7	8.5	47.2	2.4	49.7	50.3	7,436

Source: IEG data set and staff calculation

In the SAR and EAP regions, a little over half of all projects have reported ERRs in the PAD or ICR (56.1 and 57.6 percent, respectively). This is higher than for other regions where this ranges from 45.2 to 52 percent. In contrast, in SAR and EAP, about 43.9 and 42.4 percent, respectively, of the projects do not report an ERR either at the appraisal or at the completion stage. This number ranges from 48 to 55 percent for other regions (Table 2).

Model 1, described above, examines whether the frequency of ERRs has been declining in recent years. Figure 1 shows a long-term declining trend, but also suggests there has been a slight uptick over the past decade. Table 3 shows the declining trends by "high-CBA" and "low-CBA" sectors. The IEG study (World Bank 2010) found that the changes were driven mainly by changes within the "high-CBA" sectors (i.e. transport, energy, water, agriculture, and urban development) and also a shift toward "low-CBA" sectors. This paper updates that analysis and finds that this result still holds (Table 4).

Figure 1: The long-term trend of ERR presence in PAD and in ICR, respectively



Sources: IEG database and staff calculation.

Table 3: Trends in Reporting of ERR in PAD for "High-CBA" and "Low-CBA" sectors

				South Asia								Other Regions				
	1975~ 79	1980~ 84	1985~ 89	1990~ 94	1995~ 00	2000~ 04	2005~ 09	2010~ 15	1975~ 79	1980~ 84	1985~ 89	1990~ 94	1995~ 00	2000~ 04	2005~ 09	2010~ 15
Globle Practice								High	CBA GP							
Agriculture	84.6	80.0	73.1	75.0	85.0	70.4	47.1	48.0	90.7	91.8	82.2	68.2	58.8	43.1	43.1	55.1
Energy and Extractives	69.2	85.7	67.6	68.2	82.4	80.0	60.0	50.0	79.7	74.8	50.3	58.0	69.3	68.2	52.5	58.7
Social and Urban Development	100.0	100.0	0.0	50.0	70.0	0.0	14.3	0.0	40.0	80.6	85.7	65.3	56.4	29.4	18.4	17.3
Transport & ICT	92.3	88.9	86.7	100.0	82.4	72.7	61.5	62.5	94.7	95.0	85.4	81.4	85.6	76.7	66.1	69.9
Water	..	60.0	66.7	100.0	77.8	44.4	0.0	71.4	78.9	70.0	52.7	37.5	71.2	55.8	52.3	60.7
Total for High CBA GP	83.0	81.6	71.7	74.2	81.8	67.2	43.5	47.7	89.0	86.6	73.0	66.2	68.1	55.4	47.0	52.7
								Low	CBA GP							
Education	0.0	0.0	0.0	0.0	0.0	0.0	5.6	53.8	0.0	0.0	0.0	0.0	1.1	1.7	6.7	18.9
Environment & Natural Resources	..	..	..	..	80.0	0.0	..	0.0	..	0.0	..	..	31.6	13.6	22.2	30.4
Finace, Markets, Trade & Competitiveness	0.0	0.0	14.3	33.3	11.1	0.0	9.1	28.6	0.0	4.1	1.7	9.1	4.7	7.1	10.5	24.6
Governance	..	0.0	0.0	0.0	..	..	0.0	16.7	0.0	0.0	0.0	2.7	5.9	4.5	3.7	6.9
Health Nutriton & Population	..	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	5.7	17.0
Macroeconoimcs & Fiscal Management	..	..	..	0.0	0.0	0.0	0.0	..	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.7
Social Protection & Labor	..	..	..	..	0.0	..	..	0.0	..	..	0.0	0.0	12.8	2.8	5.7	7.1
Total for Low CBA GP	0.0	0.0	1.9	1.6	7.5	0.0	3.3	19.7	0.0	1.0	0.3	1.5	4.2	4.1	6.8	14.4
Total	69.8	66.0	63.3	66.0	64.2	39.4	21.6	36.1	69.2	66.7	58.3	51.6	44.4	29.8	29.2	36.0

Source: IEG database and Staff calculations.

Table 4: Contributions of High and Low-CBA GPs to the declining ERR reporting (1985-2015)

	Percentage point change*	Contribution by sector (%)
	South Asia Region	
(1) Contribution within High-CBA sectors	-23.9	85.9
(2) Contribution within Low-CBA sectors	4.8	-17.1
(3) Contribution from sector shifts (from High to Low-CBA)	-7.9	28.4
(4) Residual	-0.8	2.8
SAR-Total Reduction in Reporting	-27.8	100.0
	World Bank	
(1) Contribution within High-CBA sectors	-25.7	73.1
(2) Contribution within Low-CBA sectors	2.5	-7.1
(3) Contribution from sector shifts (from High to Low-CBA)	-11.9	33.9
(4) Residual	0.0	0.1
World Bank –Total Reduction in Reporting	-35.2	100.0

Sources: IEG, 2015, investment project only, and staff calculation

Percentage point changes are calculated as: (1) = ΔERR_High_CBA*share of High_CBA_1975; (2) = ΔERR_low_CBA*share of Low_CBA_1975; (3) = ΔERR between High and Low CBA in 2014*Δshare of High_CBA; where ERR is the percent of projects with ERR calculations, and all changes are changes between 1975~79 and 2010~14 unless noted otherwise, where ΔERR is the difference in percent of projects reporting ERR at appraisal.

Table 4 uses the same methodology presented in the IEG 2010 study to examine which sectors have contributed most to the above trends for the South Asia Region and also for the whole Bank. There has been a moderate increase in ERR reporting among low-CBA GP projects. However, this is not large enough to compensate for the reduction in reporting among high-CBA sectors and the shift in the portfolio toward low-CBA projects. For South Asia, of the total reduction of 28 percent in projects reporting ERRs at entry (or in the Project Appraisal Document or PAD), 24 percentage points arise from a reduction in reporting within high-CBA sectors. This accounts for 86 percent of the total difference. Across all regions (including SAR), the total reduction is 35 percent and 25.7 percentage points arise from a reduction in reporting within high-CBA sectors, which accounts for 73 percent of the total reduction. The increase in low-CBA projects in the portfolio slightly offsets the reduction in the frequency of ERRs reported in the PAD, more so in SAR compared to the whole Bank.

Table 5 shows that over time, the reduction in ERR reporting within high-CBA sectors appears to be contributing increasingly to the total reduction in ERR reporting. Specifically, between 1985 and 1995, 15 percent of the overall reduction in ERR reporting was due to a reduction in reporting within high-CBA sectors; 88 percent of the reduced reporting of ERRs was due to shifts from high- to low-CBA projects, reflecting the Bank's shift towards more projects in the social sectors and more programmatic approaches. By 1995-2015, 92 percent of the reduction in reporting was due to a reduction in reporting within the high-CBA sectors. Also, over time, increased reporting within the low-CBA sectors has played a greater role in offsetting the lower reporting within low- and high-CBA sectors.

Table 5: Comparison of Trends in ERR Reporting with the IEG Report

	Percentage point change	Contribution by sector (%)
	1985-1995	
(1) Contribution within High-CBA sectors	-2.5	15.1
(2) Contribution within Low-CBA sectors	0.6	-3.5
(3) Contribution from sector shifts (from High to Low-CBA)	-14.3	88.3
(4) Residual	0.0	0.0
All projects	-16.2	100.0
	1995-2015	
(1) Contribution within High-CBA sectors	-9.3	92.3
(2) Contribution within Low-CBA sectors	3.3	-32.6
(3) Contribution from sector shifts (from High to Low-CBA)	-4.4	43.9
(4) Residual	0.4	-3.7
All projects	-10.1	100.0

Sources: IEG, 2015, for both investment and budget support projects, and staff calculation

Regression analysis (Table A.2 to A.4) shows that the above decline in the frequency of ERR reporting is strongly significant. The full effect of the time trend (linear and quadratic specification of the approval year), suggests that the overall trend is declining; however, an examination of the marginal effect indicates that there is a turning point around the years 2003/2004. South Asia is also benchmarked against other regions in the regression using region dummy variables. The regression results show that the frequency of ERR reported in PADs is not statistically different between South Asia and East Asia and also between South Asia and Middle East and North Africa. In contrast, South Asia exhibits significantly greater frequency in ERR reporting compared to Africa and slightly more compared to Latin America. The positive and strongly significant results for the sector dummies simply highlight the difference between high and low CBA projects.

The size of the project, duration, and project cost are positively associated with ERR reporting at entry (Tables A.2 and A.4). For the ICR regression (Table A.3), extension of the closing date is positively correlated with ERR reporting. These are all indicators that could be (to various degrees) proxies for the level of attention, or visibility of a project and therefore lend support to the hypothesis that the quality of economic analysis tends to be better for projects with higher visibility. Also the team effects are significant in almost all the regressions.[4]

[4] The exceptions are regressions in Table A.2 and Table A.5. Table A.2 uses ERR reporting only in the PAD as a proxy for quality. In contrast, team effects are significant across regressions that use reporting in both the PAD and ICR as the proxy for quality. Another exception is Table A.5 which attempts to explain the divergence between ERR reported in the PAD and ICR.

Model 2 Results: Is the discrepancy or divergence between ERR estimated at appraisal and completion worsening?

Table 6 summarizes the average differences in estimated ERRs between appraisal and completion. For the Bank as a whole, the average mean difference is -12.1 percent and the average median difference is -15.4 percent. Across regions of the Bank, the average ERRs by region fall within 2 to 5 percentage points of one another at appraisal, and up to 9 percentage points at project completion, with ECA exhibiting the highest estimated ERRs, followed by SAR. Table 6 also shows by how much the ERR at appraisal exceeds the ERR at completion for each region. For SAR, there is a -8.7 percent difference between the two estimates (for all other regions this difference ranges from -18 to +3.6 percent). Similarly, the median difference of estimated ERRs between appraisal and completion for SAR is -16 percent (and for all other regions it ranges from -22 to -7 percent).

Table 6: Average ERRs and their differences by Region, 1975-2015

	Average ERR at appraisal	Average ERR at completion	Mean difference between the two ERRs, %	Median difference between the two ERRs, %	Standard Deviation of the differences	total # of projects with any ERRs
AFR	25.3	19.4	-18.0	-21.8	75.7	997
EAP	24.9	21.3	-10.1	-10.7	52.7	726
ECA	28.5	28.3	3.6	-7.2	78.6	392
LCR	25.4	20.9	-15.8	-22.0	56.8	709
MNA	22.8	19.5	-15.8	-12.5	62.2	345
SAR	25.7	21.6	-8.7	-16.0	63.8	526
World Bank	25.4	21.4	-12.1	-15.4	65.5	3,695

Source: IEG data set and staff calculation

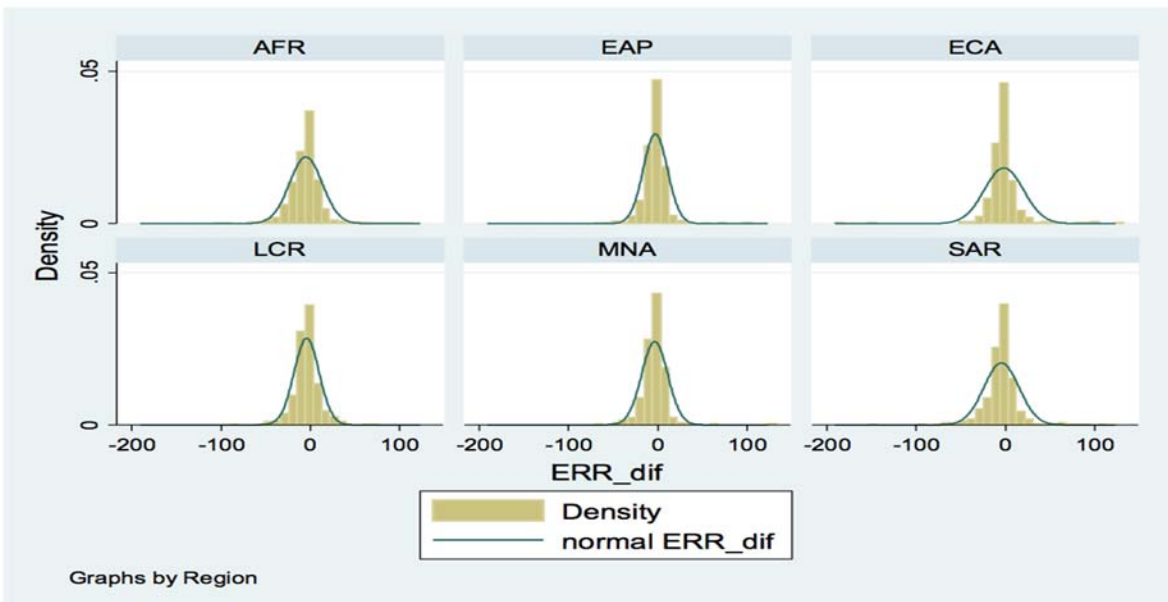
The difference between the two ERRs is calculated for every project with two ERRs as 100(ERR at completion-ERR at appraisal)/ERR at appraisal. The mean difference was then taken on the percent of differences.

The differences in estimates are also summarized graphically in Figure 2. Ideally, the distribution of the difference between ERR at appraisal and completion would be centered on zero with a narrow spread. This, however, varies across regions. The distribution is slightly squatter and wider for SAR and ECA compared to other regions while EAP and MNA seem to have the narrowest distribution.

Table 7 presents the data by GP. The mean difference and median differences are lowest for the Education, Environment, and Transport/ICT. Social and Urban Development also have relatively low average deviations between appraisal and completion. The high-CBA sectors of water and agriculture exhibit higher deviations between ERRs estimated at appraisal and completion compared to those in the soft or low-CBA sectors. The results for the low-CBA sectors may be a reflection of the small sample size of this group in the full data set (based on Table 1, only 3 to 7 percent of low-CBA projects report ERRs in the PAD and ICR). It could also reflect a pattern noted in IEG 2010 that the estimation or reporting of an

ERR at completion for low performing projects is much less likely to occur than for high performing projects. The IEG 2010 report notes that the probability of calculating an ERR for a Highly Unsatisfactory project, at completion, is around 20 percent for high-CBA projects and practically 0 for low-CBA projects. Conversely, the probability of calculating an ERR for a Highly Satisfactory project at completion is around 90 percent for a high-CBA project and 100 percent for a low-CBA project. It could be that teams presume that an unsatisfactory project will have an extremely low ERR and therefore see no need to estimate one. Going forward, a systematic calculation of ERRs (for both poor and good performing projects in the PAD and ICR) along with project related factors that may be affecting performance could help in gaining a better understanding of the appropriate range of ERRs.

Figure 2 Distributions of the differences between ERR at appraisal and ERR at completion, 1975-2015



Source: IEG data set and staff calculation.
Note: the difference is defined as ERR at completion – ERR at appraisal

Regression analysis (corrected for country and project related factors) is also used to further examine the above trends and is reported in Table A.5. The variable Approval Year is small and insignificant, indicating little trend in ERR divergence. However, with only 3 to 5 percent of variation explained, it is difficult to statistically explain what is driving the difference. One key factor could be a lack of uniform standards for assumptions and in particular the tendency not to reflect in the EA the identified risks and project design issues that would have a bearing on the ERR.

Table A.5 also shows that SAR exhibits a significantly higher dispersion between ERR estimated at appraisal and completion compared to the EAP region. SAR is not statistically different from the other regions (after correction for country and project characteristics). The regression also shows that the difference between high- and low-CBA projects with respect to the reported ERRs in the PAD and ICR is not significant. Task team effects and inflation are not strongly correlated with the absolute difference in ERRs reported in the PAD and ICR. The logarithm of total project commitment (or project size) is strongly and negatively correlated with the absolute difference in reported ERRs in the PAD and ICR. The larger the project, the smaller this difference. This again suggests that more care may have been given to the quality of economic analysis in larger projects, independent of the task team effect, which is insignificant. The coefficient on the inflation variable becomes significant when the project size is also entered into the regression, suggesting some cross effects between these two variables.

Table 7: Average ERRs and their differences by GP, 1975-2015

	Average ERR at appraisal	N	Average ERR at completion	N	Mean difference between the two ERRs, %	N	Median difference between the two ERRs, %	Standard Deviation of the differences
Agriculture	24.7	1,284	17.5	1,103	-25.8	1,050	-26.5	60.3
Education	19.8	35	22.2	31	11.7	19	0.2	71.2
Energy and Extractives	22.5	661	19.3	607	-9.4	582	-14.5	74.7
Environment & Natural Resources	20.2	29	24.2	26	12.2	21	0.0	68.4
Finace, markets, Trade & competitiveness	33.3	46	29.5	27	49.9	23	25.0	110.0
Governance	84.8	14	113.1	6	-22.2	4	-25.8	57.1
Health Nutriton & Population	43.3	24	42.5	16	-10.4	9	0.0	40.8
Macroeconoimcs & Fiscal Management	156.0	1	0.0	-	0.0	-	-	-
Social Protection & Labor	29.9	13	25.3	9	18.2	4	-9.4	72.1
Social, Urban, Rural and Resilience	23.1	195	23.0	164	-4.6	136	-10.8	66.1
Transport & ICT	30.0	951	28.1	852	-0.5	830	-5.8	59.2
Water	15.6	257	13.2	216	-13.9	199	-16.7	66.7
Poverty and Equity	-	-	-	-	-	-	-	-
World Bank	25.4	3,510	21.4	3,057	-12.1	2,877	-15.44	65.5

Source: IEG database and staff calculation.

The difference between the two ERRs is calculated for every project reporting an ERR at appraisal and completion as: 100(ERR at completion – ERR at appraisal)/ERR at appraisal. The mean difference was then taken on the percent of differences.

Model 3 Results: Does EA matter to the success of a project, after controlling for country, project and team effects?

The results for model 3, which examines how the frequency and discrepancy in ERRs are correlated with project performance, are summarized in Tables A.6 to A.9. OLS regressions are estimated for the case where the dependent variable is the project rating on the 6 point scale. Logit regressions are used for the binary dependent variable where project performance is represented as either successful or unsuccessful.

Tables A.6 and A.9 show that reporting the ERR in both the PAD and ICR is strongly and significantly associated with project performance. Tables A.8 and A.9 show that the discrepancy between ERR reported in the PAD and ICR is negatively and significantly associated with project performance. Specifically, the wider the discrepancy (after controlling for the optimism bias at project preparation), between the ERR reported in the PAD and ICR, the lower the project performance. These regressions provide strong evidence for the impact of the quality of project economic analysis on project performance, with results that are robust across the various regression specifications and estimation methods.

Among the project variables, TTL performance, preparation cost as a share of total commitment and whether a project is an actual problem project (i.e. has been rated unsatisfactory at any time between the start and ending of the project when all commitments have been fully disbursed) are all consistently significant across all the regressions. Overall, project performance is higher, the higher the ratings a task team has received in previous projects. Higher project preparation costs as a share of total commitment are associated with lower project performance. And, a project that had problems in the first half of life is likely to perform poorly at the completion stage. The other project related variables are less robust across the specifications, although in general they have the expected signs. For example, larger projects with lower duration perform better in general, but in some regressions the coefficients are negative and insignificant. The restructuring variable is negative and insignificant across most of the regressions. The parameters for the country level variables have the expected signs and most are statistically significant. GDP per capita growth and overall CPIA ratings are positively associated with project outcomes.

To test the robustness of the association between EA and project outcomes, we also ran a set of regressions spanning from 1985 to 2015 (Annex Tables A.10 and A.11). These regressions, however, did not include early warning variables because these came in effect mostly in the late 1980s or early 1990s. (Specifically, indicators of potential problems came in effect in 1989, and those for actual problem projects came in effect in 1994, and for restructuring in 1995.) Nevertheless, the results show remarkably similar signs to those reported for the 1995 to 2015 sample for both the frequency of ERRs reported in project documents and the ERR discrepancy.

6. Summary and Conclusion

Overall, the analysis indicates that the frequency of ERRs reported in project documents (as a proxy for the quality of economic analysis in the portfolio) is declining over time but with a marked uptick in the last decade. Overall, the reduction in the frequency of ERRs reported (or EA completed) stems largely from a sector shift between 1985 and 1995, but from 1995 onward, the key factor causing the reduction in ERRs reported arises from lower reporting within the high-CBA sectors (which may have worsened over time). There is no discernible trend in the discrepancy between ERR at appraisal and completion (after accounting for country and other project level effects). However, larger projects (in terms of committed amounts) are associated with higher ERR reporting; and task team effects are significantly correlated with ERR reporting and project performance but not ERR dispersion.

This paper also presents some of the factors that may be explaining the ERR discrepancy. Key among these are price inflation and the size of the project (which suggests that larger projects may tend to get more attention in terms of a preparation/supervision budget and oversight). Overall, less than 5 percent of the variation in ERR discrepancy is explained by the variables in the regressions. However, earlier assessments (and also anecdotal evidence gleaned in the course of preparing this paper) point to inconsistent data capture, lack of consistent and sector specific guidelines for EA, as well as limited resources as factors that may explain the dispersion in ERR.

The strong links of the EA to project performance suggest that EA should be viewed as more than just an approach to estimating rates of return on which to justify investment decisions. Devarajan et al. (1995) have questioned the focus on estimating a precise ERR. Instead, they argued for a shift in emphasis of EA away from the precise rate-of-return calculations toward broader sectoral analyses and public expenditure reviews. They noted three critical areas for proper project appraisal, namely: the rationale for public intervention, fiscal impact of the project and contribution to poverty reduction. These aspects have been incorporated into recent World Bank guidance on doing EA and CBA for investment project financing (IPFs).

The analysis presented in this note suggests that in addition to the elements noted by Devarajan et al (1997), project EA should be seen as a means to think through design and implementation issues with a focus on assessing ways to improve performance and manage risks to achieve targeted goals. This is a particularly practical approach given that the estimated level of ERR is rarely a factor for deciding whether to proceed with a project or not (World Bank 2010). Also, the Bank is putting more attention on how to use project risk ratings to improve the performance of the project portfolio.

The EA could be used to strengthen the discussion in the risk section of the PAD and the rationale for the project risk ratings (or the SORT). In order to achieve this, it will be important to pay closer attention to the assumptions used in calculating ERRs. There may also be a need to more systematically track ERRs reported in the PADs and ICRs to ensure these are systematically and consistently captured, which, along with other readily available country and project level variables, could greatly facilitate gaining a deeper understanding of the factors that determine the ERR.

A number of recommendations could help strengthen the link between the EA and the SORT and thereby more strongly inform project design.

I. **Automate the process for systematically tracking ERRs reported in PADs and ICRs**. In the course of this review, it was noted that some of the ERRs estimated and reported in the PADs and ICRs were not always systematically captured in the IEG database. For example, some PADs/ICRs reported multiple ERRs (sometimes by project component) that were often reflected as an average in the IEG database or not captured at all. Systematically capturing estimated ERRs would provide more refinement to the data and provide a more robust estimate of the levels of rates of returns for different sectors/projects and sub-components of a project. This system could be captured through the Bank's management information system (or Ops Portal). This would enable the systematic capture of ERRs (even when this is done only for specific components or involves multiple estimates or for project restructuring). This could then be downloaded automatically to the IEG ICR Review templates and validated by IEG at the time of their independent review.

II. **Place more emphasis on the EA as a way to systematically think through project design and how risks will impact design**. This would also include devising an approach to risk analysis that focuses on informing the design of the project or the approach to implementation. The aim would be to use the EA to devise approaches to increase the project's expected return and reduce the risk to the PDO. Risk analysis could also be used to prioritize and make explicit in the results framework what variables to monitor more closely during implementation. This analysis would then be used to discuss risk mitigation in the Risk section of the PAD and inform the project (SORT) ratings. There could also be more systematic discussions during the PCN and Decision Review meetings by requiring that reviewers comment specifically on whether the EA adequately underpins the design or approach outlined in the project; or whether it provides sufficient evidence to demonstrate that the chosen design or implementation approach is the best to maximize the project's expected return.

III. **Strengthen the database on EA in order to facilitate better analysis and identify flags, etc. for managing the portfolio**. Collect project-by-project information on cost overruns and implementation delays, etc. in the IEG database. This will enrich the IEG data set and make it possible to examine how these variables affect project outcomes and risks for different regions and GPs. It could possibly help to devise early warning flags for project supervision, implementation and mid-course correction.

IV. **Focus on strengthening assumptions and data collection at appraisal in order to reduce the difference between ERR at appraisal and at completion**. Some aspects of the divergence are due to risk, but there is also a part due to the quality of assumptions and data. To accommodate the element of risk, Herrera (2005) has recommended that the sensitivity or risk analysis at appraisal should be used to define an upper and lower bound of the ERR to be used as a benchmark at completion. In this context, the systematic collecting of data on ERRs and possible correlates (such as cost overruns, restructuring, project delays, etc.) could greatly facilitate gaining a deeper understanding of the factors that affect the ERR.

V. **Develop GP specific guidelines to doing EA (with an emphasis on how it informs project design and risk mitigation) to ensure standards and consistency across projects**.

7. References

Belli, P. 1996. Is Economic Analysis of Projects Still Useful? Policy Research Working Paper No.1689. December. World Bank.

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

Chauvet L, P. Collier, and M. Duponchel. 2010. What Explains Aid Project Success in Post-Conflict Situations? Policy Research Working Paper 5418, World Bank, Washington DC.

Dasgupta, Partha, Stephen Marglin, and Amartya K. Sen. 1972. Guidelines for Project Evaluation. New York: United Nations Industrial Development Organization

Deininger, Klaus, Lyn Squire, and Swati Basu (1998). "Does Economic Analysis Improve the Quality of Foreign Assistance?". World Bank Economic Review. 12(3):385-418.

Denizer, C., D. Kaufmann, A. Kraay. 2013. Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance, Journal of Development Economics, 105 (November 2013).

Devarajan, S.L, L. Squire, S. Suthiwart-Narueput. 1997. Beyond Rate of Return: Reorienting Project Appraisal. The World Bank Research Observer, vol. 12, no. 1 (February), pp. 35-46.

Devarajan, S., Squire, L. and Suthiwart-Narueput, S. 1995. Reviving Project Appraisal at the World Bank. World Bank Policy Research Paper, Number 1496.

Dollar, David and Victoria Levin (2005). "Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries". World Bank Policy Research Working Paper No. 3524.

Dreher, Axel, Stephan Klasen, James Raymond Vreeland, and Eric Werker (2010). "The Costs of Favouritism: Is Politically-Driven Aid Less Effective?. CESIFO Working Paper No. 2993.

Guillaumont, P., Laajaj, R. 2006. When Instability Increases the Effectiveness of Aid Projects. World Bank Policy Research Working Paper 4034, October.

Harberger, Arnold C. 1973. Project Evaluation: Collected Papers. New York: Macmillan

Herrera, S. 2005. The Economic Rate of Return of World Bank Projects. Mimeo. September 7. World Bank.

Jenkins, G.P. 1997. Project Analysis and the World Bank. The American Economic Review, Vol. 87, No. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting fo the American Economic Association. May, pp.38-42.

Isham, Jonathan, Daniel Kaufmann and Lant Pritchett. 1997. "Civil Liberties, Democracy, and the Performance of Government Projects". World Bank Economic Review. 11(2): 219-242.

Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobatón (1999). Aggregating Governance Indicators. World Bank Policy Research Working Paper No. 2195.

Kilby, C. 2014. Assessing the Impact of World Bank Preparation on Project Outcomes. Department of Economics Villanova University. Accessed from: http://sites.bu.edu/neudc/files/2014/10/paper_79.pdf.

Kilby, C. 2000. Supervision and performance: the case of World. Journal of Development Economics. 62: 233-259.

Little, I.M.D. and Mirrlees, J.A. 1991. Project Appraisal and Planning Twenty Years On. In Stanley Fischer, Dennis de Tray and Shekhar Shah, eds., Proceedings of the World Bank Annual Conference on Development Economics, 1990s. Washington, D.C. World Bank: pp.351-82.

Little, I.M.D. and Mirrlees, J.A. 1974. Project Appraisal and Planning. London: Heinemann.

Little, I. M. D., and J. A. Mirrlees. 1969. Manual of Industrial Project Analysis. Paris: OECD Development Centre.

Squire, Lyn, and Herman G. van der Tak. 1975. Economic Analysis of Projects. Baltimore, Md.: Johns Hopkins University Press for the World Bank.

Pohl, G., D. Mihaljek. 1992. Project Evaluation and Uncertainty in Practice – A Statistical Analysis of Rate of Return Divergences of 1,015 World Bank Projects. The World Bank Economic Review. Vol. 6, No. 2:255-277

Vawda, Ayesha Yaqub, Moock, J., Gittinger, J.P., and Patrinos, H. 2001. Economic Analysis of World Bank Education Projects and Project Outcomes, Policy Research Working Paper 2564, World Bank

World Bank. 2010. Cost-Benefit Analysis in World Bank Projects. Independent Evaluation Group, World Bank, Washington, DC.

World Bank. 1995. A Review of the Quality of Economic Analysis in Staff Appraisal Reports for Projects Approved in 1993. Operations Policy Department Operations Evaluation Department. May 5.

8. Annexes

Table A.1: Distribution of Projects Across Sectors, 1975-2015

	Number	% of Total	Number	% of Total	Number	% of Total
Agriculture	1,851	24.93	1,418	23.65	768	19
Transport & ICT	1,146	15.43	811	13.53	513	12.69
Energy and Extractives	1,024	13.79	797	13.29	400	9.9
Social, Urban, Rural and Resilience	454	6.11	414	6.9	308	7.62
Water	436	5.87	362	6.04	254	6.28
Total high CBA project	4,911	66.1	3,802	63.4	2,243	55.5
Education	755	10.17	614	10.24	460	11.38
Environment & Natural Resources	126	1.7	124	2.07	124	3.07
Finance, markets, Trade & competitivene	575	7.74	432	7.2	318	7.87
Governance	333	4.48	316	5.27	243	6.01
Health Nutrition & Population	448	6.03	432	7.2	391	9.67
Macroeconomics & Fiscal Management	59	0.79	57	0.95	48	1.19
Social Protection & Labor	217	2.92	217	3.62	213	5.27
Poverty and Equity	2	0.03	2	0.03	2	0.05
Total low CBA project	2515	33.9	2194	36.6	1799	44.5
Total	7,426	100	5,996	100	4,042	100

Table A.2: Logit analysis, 1995-2015, Frequency of ERR in PADs as dependent variable (Model 1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres
CPIA rating	0.353***	0.238**	0.381***	0.370***	0.350***	0.303***	0.281***
	(3.85)	(2.53)	(4.12)	(3.51)	(3.77)	(2.76)	(2.95)
CPIA scale dummy	-0.190	-0.0910	-0.193	-0.0752	-0.0824	0.0861	-0.0365
	(-1.13)	(-0.54)	(-1.14)	(-0.39)	(-0.48)	(0.44)	(-0.21)
Approval fiscal year sequence	-0.502***	-0.460***	-0.495***	-0.519***	-0.563***	-0.529***	-0.507***
	(-7.19)	(-6.56)	(-7.03)	(-6.23)	(-7.90)	(-6.15)	(-7.03)
Quadratic form of approval year sequence	0.00868***	0.00801***	0.00857***	0.00896***	0.00990***	0.00935***	0.00895***
	(6.75)	(6.23)	(6.61)	(5.79)	(7.54)	(5.86)	(6.76)
East Asia and Pacific	0.0593	0.0877	0.0351	0.0726	0.0223	0.0335	0.0370
	(0.40)	(0.58)	(0.23)	(0.42)	(0.15)	(0.19)	(0.24)
Africa	-0.462***	-0.244*	-0.524***	-0.429***	-0.400***	-0.235	-0.293**
	(-3.28)	(-1.69)	(-3.67)	(-2.58)	(-2.81)	(-1.36)	(-2.00)
East Europe and Central Asia	-0.181	0.0703	-0.192	-0.106	-0.167	0.126	0.0212
	(-1.19)	(0.45)	(-1.26)	(-0.60)	(-1.09)	(0.69)	(0.13)
Latin America & Caribbean	-0.314**	-0.168	-0.315**	-0.285*	-0.176	-0.0512	-0.100
	(-2.12)	(-1.12)	(-2.13)	(-1.67)	(-1.18)	(-0.29)	(-0.66)
Middle East & North Africa	0.0516	0.182	-0.0147	0.0181	0.0833	0.113	0.124
	(0.28)	(0.98)	(-0.08)	(0.08)	(0.45)	(0.52)	(0.66)
Agriculture	2.229***	2.251***	2.232***	2.230***	2.185***	2.227***	2.218***
	(20.65)	(20.64)	(20.63)	(18.10)	(20.07)	(17.71)	(20.19)
Energy & mining	2.752***	2.654***	2.764***	2.743***	2.794***	2.725***	2.713***
	(20.40)	(19.43)	(20.44)	(17.53)	(20.41)	(16.93)	(19.61)
Transport	3.364***	3.200***	3.363***	3.319***	3.433***	3.254***	3.284***
	(24.60)	(23.13)	(24.53)	(21.44)	(24.72)	(20.32)	(23.27)
Water	2.507***	2.450***	2.490***	2.413***	2.465***	2.331***	2.415***
	(16.35)	(15.87)	(16.22)	(13.91)	(15.95)	(13.22)	(15.55)
Urban Development	1.376***	1.317***	1.390***	1.438***	1.386***	1.425***	1.347***
	(8.60)	(8.16)	(8.66)	(7.86)	(8.57)	(7.62)	(8.28)
Log net commit. (m$)		0.269***				0.238***	0.217***
		(7.33)				(5.32)	(5.63)
Project length			0.0772***			0.0773***	0.0655***
			(3.53)			(3.01)	(2.95)
TTL performance				0.0870		0.0749	
				(1.61)		(1.35)	
Log (preparation cost, us$)					0.376***	0.293***	0.268***
					(6.85)	(4.48)	(4.67)
Observations	4020	4020	4020	3065	4006	3055	4006
Pseudo R-squared	0.274	0.284	0.276	0.274	0.283	0.294	0.291

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.3: Logit analysis, 1995-2015, Frequency of ERRs in ICRs as dependent variable (Model 1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres
CPIA rating	0.319***	0.209**	0.362***	0.341***	0.329***	0.297***	0.271***
	(3.45)	(2.21)	(3.87)	(3.21)	(3.54)	(2.70)	(2.77)
CPIA scale dummy	-0.512***	-0.523***	-0.580***	-0.428*	-0.488**	-0.496**	-0.562***
	(-2.59)	(-2.62)	(-2.91)	(-1.94)	(-2.46)	(-2.20)	(-2.79)
Exit year sequence	-0.950***	-0.954***	-0.955***	-0.930***	-0.854***	-0.848***	-0.922***
	(-7.74)	(-7.70)	(-7.76)	(-6.52)	(-6.74)	(-5.66)	(-7.02)
Quadratic form of exit year sequence	0.0152***	0.0152***	0.0152***	0.0148***	0.0135***	0.0133***	0.0145***
	(7.93)	(7.91)	(7.92)	(6.61)	(6.83)	(5.63)	(7.12)
East Asia and Pacific	-0.202	-0.174	-0.230	-0.121	-0.190	-0.118	-0.170
	(-1.36)	(-1.16)	(-1.55)	(-0.70)	(-1.28)	(-0.67)	(-1.11)
Africa	-0.854***	-0.603***	-0.919***	-0.770***	-0.869***	-0.598***	-0.676***
	(-6.03)	(-4.13)	(-6.41)	(-4.59)	(-6.10)	(-3.43)	(-4.57)
East Europe and Central Asia	-0.524***	-0.217	-0.532***	-0.473***	-0.521***	-0.181	-0.230
	(-3.45)	(-1.38)	(-3.50)	(-2.68)	(-3.43)	(-0.99)	(-1.46)
Latin America & Caribbean	-0.575***	-0.412***	-0.584***	-0.521***	-0.586***	-0.387**	-0.417***
	(-3.91)	(-2.75)	(-3.96)	(-3.04)	(-3.97)	(-2.21)	(-2.76)
Middle East & North Africa	-0.364**	-0.208	-0.436**	-0.263	-0.356*	-0.157	-0.256
	(-1.97)	(-1.11)	(-2.34)	(-1.22)	(-1.92)	(-0.71)	(-1.35)
Agriculture	2.359***	2.387***	2.359***	2.447***	2.366***	2.478***	2.377***
	(20.40)	(20.36)	(20.35)	(18.27)	(20.40)	(18.17)	(20.13)
Energy & mining	2.743***	2.625***	2.749***	2.807***	2.727***	2.688***	2.624***
	(19.83)	(18.70)	(19.84)	(17.36)	(19.69)	(16.34)	(18.62)
Transport	3.289***	3.105***	3.282***	3.326***	3.279***	3.143***	3.106***
	(24.14)	(22.48)	(24.05)	(21.23)	(24.01)	(19.69)	(22.24)
Water	2.482***	2.412***	2.461***	2.500***	2.467***	2.400***	2.376***
	(15.84)	(15.24)	(15.67)	(13.90)	(15.71)	(13.17)	(14.97)
Urban Development	1.485***	1.424***	1.500***	1.582***	1.482***	1.532***	1.430***
	(8.73)	(8.27)	(8.79)	(8.04)	(8.70)	(7.65)	(8.28)
IEG evaluation method (PPAR=1)		0.317***				0.307***	0.300***
		(8.49)				(7.03)	(7.42)
Log net commit. (m$)			0.0811***			0.0848***	0.0650***
			(3.62)			(3.19)	(2.78)
Project length				0.132**		0.102*	
				(2.38)		(1.80)	
TTL performance					0.00925***	0.00837**	0.00568*
					(3.15)	(2.35)	(1.86)
Closing extension (months)					0.00925***	0.00837**	0.00568*
					(3.15)	(2.35)	(1.86)
Observations	4020	4020	4020	3065	4009	3057	4008
Pseudo R-squared	0.259	0.274	0.262	0.266	0.261	0.284	0.278
t statistics in parentheses	* p<0.10, ** p<0.05, *** p<0.01			Estimation method: Logit			

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.4: Logit analysis, 1995-2015, Frequency of ERRs in both the PAD and ICR as dependent variable (Model 1)[5]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres	Pres/Nopres
CPIA rating	0.512***	0.387***	0.542***	0.573***	0.513***	0.504***	0.431***
	(5.24)	(3.86)	(5.50)	(5.12)	(5.21)	(4.33)	(4.24)
CPIA scale dummy	-0.478***	-0.377**	-0.487***	-0.352*	-0.413**	-0.250	-0.365**
	(-2.70)	(-2.11)	(-2.74)	(-1.74)	(-2.31)	(-1.22)	(-2.02)
Approval fiscal year sequence	-0.522***	-0.482***	-0.516***	-0.529***	-0.581***	-0.530***	-0.518***
	(-7.11)	(-6.54)	(-6.96)	(-6.05)	(-7.74)	(-5.86)	(-6.81)
Quadratic form of approval year sequence	0.00893***	0.00832***	0.00882***	0.00899***	0.0101***	0.00919***	0.00903***
	(6.58)	(6.13)	(6.45)	(5.53)	(7.28)	(5.46)	(6.44)
East Asia and Pacific	-0.165	-0.149	-0.189	-0.131	-0.199	-0.170	-0.191
	(-1.08)	(-0.96)	(-1.23)	(-0.74)	(-1.30)	(-0.94)	(-1.23)
Africa	-0.823***	-0.573***	-0.884***	-0.753***	-0.776***	-0.555***	-0.629***
	(-5.57)	(-3.77)	(-5.92)	(-4.33)	(-5.22)	(-3.07)	(-4.09)
East Europe and Central Asia	-0.508***	-0.220	-0.518***	-0.444**	-0.489***	-0.178	-0.251
	(-3.22)	(-1.35)	(-3.28)	(-2.43)	(-3.08)	(-0.93)	(-1.53)
Latin America & Caribbean	-0.672***	-0.520***	-0.677***	-0.645***	-0.557***	-0.441**	-0.473***
	(-4.37)	(-3.32)	(-4.39)	(-3.63)	(-3.58)	(-2.42)	(-3.01)
Middle East & North Africa	-0.280	-0.132	-0.345*	-0.227	-0.255	-0.118	-0.194
	(-1.46)	(-0.68)	(-1.78)	(-1.02)	(-1.32)	(-0.52)	(-0.99)
Agriculture	2.442***	2.471***	2.442***	2.506***	2.393***	2.505***	2.438***
	(19.13)	(19.15)	(19.11)	(17.08)	(18.67)	(16.83)	(18.82)
Energy & mining	3.015***	2.905***	3.026***	2.997***	3.034***	2.930***	2.933***
	(20.36)	(19.32)	(20.40)	(17.43)	(20.32)	(16.69)	(19.40)
Transport	3.484***	3.304***	3.481***	3.478***	3.527***	3.353***	3.343***
	(24.11)	(22.59)	(24.05)	(21.00)	(24.18)	(19.76)	(22.59)
Water	2.575***	2.515***	2.558***	2.515***	2.540***	2.433***	2.486***
	(15.48)	(15.01)	(15.36)	(13.13)	(15.20)	(12.56)	(14.78)
Urban Development	1.394***	1.330***	1.407***	1.507***	1.386***	1.459***	1.341***
	(7.24)	(6.86)	(7.29)	(6.86)	(7.16)	(6.56)	(6.89)
Log net commit. (m$)		0.315***				0.283***	0.281***
		(7.92)				(5.87)	(6.71)
Project length			0.0746***			0.0814***	0.0678***
			(3.21)			(2.98)	(2.87)
TTL performance				0.139**		0.121**	
				(2.40)		(2.03)	
Log (preparation cost, us$)					0.312***	0.187***	0.178***
					(5.33)	(2.69)	(2.89)
Observations	4020	4020	4020	3065	4006	3055	4006
Pseudo R-squared	0.282	0.296	0.284	0.285	0.288	0.304	0.299

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

[5] These regressions were also estimated for high CBA projects only. All the coefficients along with their signs were found to be similar to what is reported here.

Table A.5: OLS analysis, 1995-2015, Log of ERR Difference as dependent variable, High-CBA investment projects only (Model 2)

	(1)	(2)	(3)	(4)	(5)	(6)
	ERR dif	ERR dif	ERR dif	ERR dif	ERR dif	ERR dif
Over estimation of ERR in PAD (1 if PAD ERR>ICR ERR	0.146*	0.172**	0.147*	0.154*	0.180*	0.176**
	(1.80)	(2.09)	(1.82)	(1.69)	(1.95)	(2.15)
CPIA	-0.105	-0.0487	-0.0753	-0.189*	-0.106	-0.00143
	(-1.11)	(-0.46)	(-0.79)	(-1.76)	(-0.88)	(-0.01)
CPIA scale dummy	0.128	0.0702	0.0848	0.161	0.0485	0.0143
	(1.03)	(0.55)	(0.68)	(1.15)	(0.34)	(0.11)
Approval fiscal year sequence	0.00399	0.00537	0.000877	0.0103	0.00949	0.00217
	(0.47)	(0.62)	(0.10)	(1.02)	(0.93)	(0.25)
East Asia and Pacific	-0.408***	-0.397***	-0.414***	-0.333**	-0.332**	-0.403***
	(-3.09)	(-3.00)	(-3.15)	(-2.17)	(-2.17)	(-3.05)
Africa	0.0670	0.0693	-0.0182	0.0732	-0.0170	-0.0271
	(0.48)	(0.49)	(-0.13)	(0.45)	(-0.10)	(-0.19)
East Europe and Central Asia	0.0755	0.0170	-0.0178	0.172	0.00592	-0.0992
	(0.52)	(0.11)	(-0.12)	(1.03)	(0.03)	(-0.63)
Latin America & Caribbean	0.0410	-0.00983	0.00385	0.189	0.0878	-0.0621
	(0.29)	(-0.07)	(0.03)	(1.16)	(0.52)	(-0.42)
Middle East & North Africa	-0.119	-0.139	-0.194	-0.109	-0.239	-0.222
	(-0.68)	(-0.78)	(-1.10)	(-0.54)	(-1.14)	(-1.24)
Agriculture	0.00326	0.00239	-0.0520	-0.0153	-0.0823	-0.0611
	(0.03)	(0.02)	(-0.50)	(-0.13)	(-0.69)	(-0.58)
Energy & mining	-0.0285	-0.0180	-0.0256	0.0192	0.0376	-0.0155
	(-0.25)	(-0.16)	(-0.23)	(0.15)	(0.29)	(-0.14)
Water	0.0666	0.0791	0.0269	0.0814	0.0641	0.0312
	(0.47)	(0.56)	(0.19)	(0.49)	(0.38)	(0.22)
Urban Development	-0.0224	-0.0185	-0.0719	-0.0619	-0.111	-0.0761
	(-0.11)	(-0.09)	(-0.36)	(-0.28)	(-0.49)	(-0.38)
Average Inflation over project life, log(%)		0.0630			0.0660	0.0760*
		(1.52)			(1.42)	(1.83)
Log net commit. (m$)			-0.105**		-0.121**	-0.119***
			(-2.53)		(-2.53)	(-2.83)
TTL performance				-0.00251	0.0243	
				(-0.04)	(0.42)	
Observations	1042	1020	1042	843	827	1020
R squared	0.033	0.036	0.039	0.039	0.050	0.043

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.6: OLS analysis, 1995-2015, Project Outcome (scale 1-6 as independent variable) and ERR Reported in both PAD and ICR, as explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating
GDP PC growth	0.0211***	0.0202***	0.0204***	0.0203***	0.0214***	0.0206***	0.0215***	0.0181***	0.0166***	0.0198***	0.0180***	0.0175***
	(3.39)	(3.28)	(3.34)	(3.30)	(3.21)	(3.37)	(3.52)	(2.96)	(2.69)	(3.19)	(2.77)	(2.87)
CPIA rating	0.367***	0.340***	0.302***	0.326***	0.316***	0.321***	0.304***	0.214***	0.268***	0.310***	0.179***	0.186***
	(8.86)	(8.28)	(7.38)	(7.88)	(6.87)	(7.89)	(7.46)	(4.99)	(6.24)	(7.32)	(3.58)	(4.16)
CPIA scale dummy	-0.0972	-0.110	-0.118	-0.0902	-0.0888	-0.127	-0.111	-0.122	-0.0598	-0.0686	-0.0285	-0.0695
	(-1.22)	(-1.39)	(-1.51)	(-1.14)	(-0.99)	(-1.62)	(-1.42)	(-1.56)	(-0.45)	(-0.81)	(-0.19)	(-0.52)
Exit year sequence	-0.0291***	-0.0289***	-0.0301***	-0.0270***	-0.0300***	-0.0281***	-0.0220***	-0.0263***	-0.0268***	-0.0293***	-0.0180***	-0.0229***
	(-5.52)	(-5.53)	(-5.81)	(-5.12)	(-4.88)	(-5.44)	(-4.22)	(-5.08)	(-5.18)	(-5.60)	(-2.86)	(-4.33)
East Asia and Pacific	0.0957	0.112	0.132*	0.122*	0.118	0.136**	0.134**	0.118*	0.0262	0.100	0.0978	0.0807
	(1.39)	(1.64)	(1.94)	(1.79)	(1.50)	(2.02)	(1.98)	(1.76)	(0.37)	(1.45)	(1.22)	(1.16)
Africa	-0.169***	-0.117*	-0.0139	-0.0973	0.00737	-0.0876	-0.0611	-0.0818	-0.165**	-0.136**	0.0794	-0.0618
	(-2.67)	(-1.86)	(-0.22)	(-1.54)	(0.10)	(-1.40)	(-0.98)	(-1.31)	(-2.55)	(-2.12)	(1.03)	(-0.94)
East Europe and Central Asia	0.0806	0.120*	0.239***	0.126*	0.111	0.181***	0.202***	0.133**	0.0281	0.111	0.163**	0.164**
	(1.18)	(1.77)	(3.47)	(1.86)	(1.42)	(2.69)	(2.99)	(1.98)	(0.41)	(1.62)	(2.03)	(2.36)
Latin America & Caribbean	0.0758	0.122*	0.190***	0.125*	0.148*	0.135**	0.162**	0.173***	0.0310	0.109	0.150*	0.112
	(1.13)	(1.83)	(2.86)	(1.87)	(1.94)	(2.05)	(2.45)	(2.61)	(0.45)	(1.60)	(1.88)	(1.61)
Middle East & North Africa	-0.132	-0.106	-0.0350	-0.0834	0.000949	-0.0767	-0.0622	-0.0749	-0.180**	-0.143*	0.0352	-0.0807
	(-1.56)	(-1.26)	(-0.42)	(-0.99)	(0.01)	(-0.92)	(-0.75)	(-0.90)	(-2.07)	(-1.67)	(0.35)	(-0.93)
IEG evaluation method (PPAR=1)	-0.0274	0.0461	0.0204	0.0424	0.0658	0.0248	0.0310	0.0445	-0.0158	0.0426	0.0161	-0.0416
	(-0.56)	(0.94)	(0.42)	(0.86)	(1.22)	(0.51)	(0.64)	(0.91)	(-0.31)	(0.84)	(0.29)	(-0.81)
ERR present in both PAD & ICR		0.416***	0.361***	0.422***	0.401***	0.387***	0.375***	0.413***	0.348***	0.382***	0.288***	0.296***
		(8.74)	(7.58)	(8.85)	(7.56)	(8.20)	(7.93)	(8.77)	(6.95)	(7.80)	(5.20)	(5.92)
Log net commit. (m$)			0.131***								0.0337	0.0519**
			(8.41)								(1.37)	(2.49)
Project length				-0.0271***							-0.0107	-0.0112
				(-2.80)							(-0.99)	(-1.16)
TTL performance					0.184***						0.174***	
					(7.95)						(7.18)	
Preparation cost over commit, %						-0.0343***					-0.0180***	-0.0170***
						(-9.65)					(-3.41)	(-3.58)
Supervision cost over commit, %							-0.0322***				-0.0193***	-0.0121**
							(-9.94)				(-3.00)	(-2.48)
Potential prob. dummy (1=if raised in the 1rst half of the proj.)								-0.387***			-0.148***	-0.127***
								(-9.32)			(-2.79)	(-2.71)
Actual prob. dummy (1=if raised in the 1rst half of the proj.)									-0.525***		-0.415***	-0.452***
									(-14.07)		(-9.00)	(-11.06)
Restructuring dummy (1=if restricted in the 1rst half of the proj.)										-0.194**	-0.00188	-0.0108
										(-2.54)	(-0.02)	(-0.14)
Observations	3975	3975	3975	3975	3054	3975	3974	3965	3455	3781	2665	3443
R squared	0.104	0.121	0.136	0.122	0.134	0.141	0.142	0.140	0.166	0.116	0.206	0.193
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.7: Logit analysis, 1995-2015, Project Outcome (Dummy for Successful Outcome as independent variable) and ERR Reported in both PAD and ICR of explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc
GDP PC growth	0.0418***	0.0403***	0.0404***	0.0408***	0.0402**	0.0416***	0.0421***	0.0372***	0.0340**	0.0389***	0.0356**	0.0368**
	(2.98)	(2.88)	(2.89)	(2.89)	(2.51)	(2.95)	(2.99)	(2.63)	(2.22)	(2.71)	(2.09)	(2.38)
CPIA rating	0.486***	0.450***	0.387***	0.422***	0.402***	0.421***	0.390***	0.196**	0.346***	0.401***	0.171	0.178*
	(5.59)	(5.16)	(4.42)	(4.81)	(3.98)	(4.79)	(4.43)	(2.13)	(3.51)	(4.40)	(1.40)	(1.69)
CPIA scale dummy	-0.105	-0.125	-0.138	-0.0852	-0.0867	-0.166	-0.135	-0.144	-0.0706	-0.114	-0.0152	-0.117
	(-0.66)	(-0.77)	(-0.85)	(-0.53)	(-0.45)	(-1.02)	(-0.83)	(-0.88)	(-0.24)	(-0.65)	(-0.04)	(-0.38)
Exit year sequence	-0.0232**	-0.0227**	-0.0245**	-0.0193*	-0.0271**	-0.0218**	-0.00916	-0.0177	-0.0164	-0.0223**	0.00373	-0.00609
	(-2.18)	(-2.12)	(-2.27)	(-1.79)	(-2.09)	(-2.02)	(-0.83)	(-1.64)	(-1.42)	(-2.06)	(0.25)	(-0.50)
East Asia and Pacific	0.145	0.180	0.237	0.200	0.148	0.249*	0.238	0.203	0.0316	0.152	0.129	0.164
	(0.99)	(1.21)	(1.58)	(1.35)	(0.85)	(1.66)	(1.59)	(1.35)	(0.19)	(0.99)	(0.64)	(0.96)
Africa	-0.353***	-0.264**	-0.0881	-0.222*	-0.105	-0.205	-0.160	-0.203	-0.414***	-0.318**	-0.00266	-0.220
	(-2.76)	(-2.04)	(-0.67)	(-1.71)	(-0.67)	(-1.58)	(-1.22)	(-1.55)	(-2.82)	(-2.38)	(-0.01)	(-1.44)
East Europe and Central Asia	0.137	0.203	0.411***	0.217	0.216	0.336**	0.373**	0.232	0.0269	0.180	0.279	0.288*
	(0.95)	(1.40)	(2.74)	(1.49)	(1.25)	(2.27)	(2.51)	(1.57)	(0.17)	(1.20)	(1.41)	(1.73)
Latin America & Caribbean	0.209	0.289**	0.416***	0.297**	0.286*	0.325**	0.375***	0.392***	0.115	0.251*	0.329*	0.293*
	(1.48)	(2.02)	(2.87)	(2.08)	(1.71)	(2.26)	(2.60)	(2.71)	(0.70)	(1.69)	(1.68)	(1.75)
Middle East & North Africa	-0.235	-0.195	-0.0779	-0.148	-0.0666	-0.138	-0.128	-0.147	-0.369*	-0.274	-0.0322	-0.194
	(-1.39)	(-1.15)	(-0.45)	(-0.87)	(-0.33)	(-0.81)	(-0.74)	(-0.85)	(-1.91)	(-1.55)	(-0.14)	(-0.98)
IEG evaluation method (PPAR=1)	-0.0816	0.0237	-0.0172	0.0156	0.0580	-0.0175	-0.00280	0.0149	-0.109	-0.00408	-0.0613	-0.158
	(-0.80)	(0.23)	(-0.16)	(0.15)	(0.49)	(-0.17)	(-0.03)	(0.14)	(-0.93)	(-0.04)	(-0.45)	(-1.32)
ERR present in both PAD & ICR		0.763***	0.670***	0.775***	0.741***	0.719***	0.691***	0.770***	0.677***	0.683***	0.558***	0.584***
		(7.11)	(6.17)	(7.21)	(6.00)	(6.65)	(6.38)	(7.09)	(5.45)	(6.12)	(3.87)	(4.62)
Log net commit. (m$)			0.239***								0.0233	0.0857
			(7.18)								(0.37)	(1.63)
Project length				-0.0573***							-0.0341	-0.0257
				(-2.82)							(-1.30)	(-1.14)
TTL performance					0.358***						0.384***	
					(7.32)						(6.69)	
Preparation cost over commit, %						-0.0638***					-0.0309**	-0.0289**
						(-6.77)					(-2.22)	(-2.52)
Supervision cost over commit, %							-0.0596***				-0.0512***	-0.0283**
							(-7.84)				(-3.38)	(-2.40)
Potential prob. dummy (1=if raised in the 1rst half of the proj.)								-0.742***			-0.326***	-0.280***
								(-8.74)			(-2.68)	(-2.71)
Actual prob. dummy (1=if raised in the 1rst half of the proj.)									-1.038***		-0.849***	-0.919***
									(-12.36)		(-7.75)	(-9.77)
Restructuring dummy (1=if restricted in the 1rst half of the proj.)										-0.342**	0.0687	0.0378
										(-2.23)	(0.35)	(0.22)
Observations	3962	3962	3962	3962	3050	3962	3961	3952	3443	3769	2661	3431
R squared	0.061	0.072	0.083	0.074	0.084	0.085	0.088	0.088	0.107	0.069	0.139	0.126
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.8: OLS analysis, 1995-2015, Project Outcome (scale 1-6 as independent variable) and ERR discrepancy between PAD and ICR as explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating	1-6 Rating
GDP PC growth	0.0211***	0.0269**	0.0244*	0.0282**	0.0357***	0.0246*	0.0232*	0.0226*	0.0269**	0.0319**	0.0328**	0.0257**
	(3.39)	(2.12)	(1.91)	(2.22)	(2.68)	(1.96)	(1.84)	(1.79)	(2.09)	(2.48)	(2.40)	(1.99)
CPIA rating	0.367***	0.199**	0.196**	0.183**	0.177**	0.191**	0.182**	0.0975	0.180**	0.132	0.170*	0.118
	(8.86)	(2.54)	(2.49)	(2.29)	(2.09)	(2.46)	(2.33)	(1.19)	(2.05)	(1.58)	(1.68)	(1.28)
CPIA scale dummy	-0.0972	-0.152	-0.158	-0.138	-0.0494	-0.198	-0.163	-0.159	-0.142	-0.176	-0.171	-0.150
	(-1.22)	(-1.05)	(-1.10)	(-0.96)	(-0.32)	(-1.39)	(-1.14)	(-1.11)	(-0.51)	(-1.18)	(-0.56)	(-0.52)
Exit year sequence	-0.0291***	-0.0223*	-0.0237*	-0.0195	-0.0245	-0.0229*	-0.0157	-0.0166	-0.0120	-0.0159	-0.00416	-0.00659
	(-5.52)	(-1.66)	(-1.77)	(-1.43)	(-1.61)	(-1.72)	(-1.17)	(-1.24)	(-0.86)	(-1.16)	(-0.24)	(-0.45)
East Asia and Pacific	0.0957	0.105	0.114	0.115	0.0179	0.111	0.121	0.0998	-0.00529	0.0800	-0.0656	0.0140
	(1.39)	(1.02)	(1.10)	(1.11)	(0.16)	(1.09)	(1.18)	(0.97)	(-0.05)	(0.75)	(-0.53)	(0.13)
Africa	-0.169***	-0.0395	0.00151	-0.0185	0.00287	-0.00980	0.00133	-0.00850	-0.0433	-0.0761	0.0973	-0.00670
	(-2.67)	(-0.37)	(0.01)	(-0.17)	(0.02)	(-0.09)	(0.01)	(-0.08)	(-0.38)	(-0.69)	(0.71)	(-0.06)
East Europe and Central Asia	0.0806	0.132	0.181	0.141	0.0660	0.193*	0.205*	0.141	0.0833	0.147	0.0733	0.157
	(1.18)	(1.20)	(1.59)	(1.28)	(0.54)	(1.76)	(1.85)	(1.30)	(0.74)	(1.31)	(0.55)	(1.34)
Latin America & Caribbean	0.0758	0.183*	0.200*	0.187*	0.225*	0.184*	0.201*	0.234**	0.116	0.200*	0.149	0.128
	(1.13)	(1.66)	(1.80)	(1.70)	(1.83)	(1.68)	(1.83)	(2.12)	(0.96)	(1.76)	(1.09)	(1.06)
Middle East & North Africa	-0.132	0.156	0.193	0.179	0.150	0.183	0.184	0.151	0.0364	0.167	0.108	0.0959
	(-1.56)	(1.15)	(1.41)	(1.31)	(1.00)	(1.37)	(1.37)	(1.12)	(0.25)	(1.19)	(0.66)	(0.65)
IEG evaluation method (PPAR=1)	-0.0274	0.0797	0.0693	0.0827	0.0458	0.0585	0.0757	0.0713	0.162	0.189*	0.115	0.152
	(-0.56)	(0.80)	(0.70)	(0.83)	(0.43)	(0.60)	(0.77)	(0.72)	(1.34)	(1.77)	(0.86)	(1.27)
Over estimation of ERR in PAD (1 if PAD ERR>ICR ERR)		-0.430***	-0.430***	-0.424***	-0.444***	-0.415***	-0.418***	-0.430***	-0.418***	-0.426***	-0.410***	-0.400***
		(-6.98)	(-6.99)	(-6.87)	(-6.64)	(-6.80)	(-6.83)	(-7.03)	(-6.46)	(-6.70)	(-5.74)	(-6.18)
ERR difference between PAD & ICR, log (%)		-0.0723***	-0.0690***	-0.0726***	-0.104***	-0.0660***	-0.0686***	-0.0674***	-0.0432*	-0.0613***	-0.0773***	-0.0413*
		(-3.15)	(-3.00)	(-3.17)	(-4.19)	(-2.90)	(-3.01)	(-2.96)	(-1.86)	(-2.64)	(-3.02)	(-1.78)
Log net commit. (m$)			0.0537*								-0.0210	-0.0375
			(1.70)								(-0.39)	(-0.78)
Project length				-0.0218							0.000212	-0.0139
				(-1.25)							(0.01)	(-0.79)
TTL performance					0.145***						0.176***	
					(3.51)						(3.71)	
Preparation cost over commit, %						-0.0595***					-0.0663***	-0.0623***
						(-4.54)					(-3.10)	(-3.15)
Supervision cost over commit, %							-0.0417***				0.0121	-0.00868
							(-3.66)				(0.55)	(-0.45)
Potential prob. dummy (1=if raised in the 1rst half of the proj.)								-0.316***			-0.0834	-0.100
								(-4.11)			(-0.86)	(-1.16)
Actual prob. dummy (1=if raised in the 1rst half of the proj.)									-0.450***		-0.411***	-0.390***
									(-6.64)		(-5.12)	(-5.39)
Restructuring dummy (1=if restricted in the 1rst half of the proj.)										-0.110	-0.0555	-0.0192
										(-0.83)	(-0.38)	(-0.14)
Observations	3975	1112	1112	1112	895	1112	1112	1111	885	1028	714	882
R squared	0.104	0.172	0.174	0.173	0.205	0.188	0.182	0.185	0.247	0.177	0.298	0.269
Adjusted R-squared	0.091	0.130	0.131	0.130	0.156	0.146	0.140	0.143	0.198	0.130	0.235	0.214
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.9: Logit analysis, 1995-2015, Project Outcome (Dummy for Successful Outcome as independent variable) and ERR discrepancy between PAD and ICR as one of explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc
GDP PC growth	0.0418***	0.0784**	0.0778**	0.0825**	0.0958**	0.0748*	0.0721*	0.0672*	0.0892**	0.0959**	0.0961*	0.0930**
	(2.98)	(2.04)	(2.02)	(2.14)	(2.22)	(1.94)	(1.87)	(1.73)	(2.01)	(2.41)	(1.88)	(2.03)
CPIA rating	0.486***	0.181	0.181	0.144	0.0962	0.175	0.162	-0.0855	0.170	-0.0187	0.202	0.0246
	(5.59)	(0.82)	(0.82)	(0.65)	(0.38)	(0.79)	(0.73)	(-0.37)	(0.60)	(-0.08)	(0.58)	(0.08)
CPIA scale dummy	-0.105	-0.492	-0.493	-0.443	-0.395	-0.602	-0.514	-0.457	-0.788	-0.636	-0.831	-0.826
	(-0.66)	(-1.22)	(-1.23)	(-1.10)	(-0.86)	(-1.47)	(-1.27)	(-1.13)	(-0.92)	(-1.53)	(-0.84)	(-0.88)
Exit year sequence	-0.0232**	0.0370	0.0367	0.0458	0.0377	0.0367	0.0481	0.0593	0.0862*	0.0583	0.115**	0.106**
	(-2.18)	(0.98)	(0.97)	(1.18)	(0.83)	(0.96)	(1.25)	(1.51)	(1.92)	(1.48)	(1.97)	(2.16)
East Asia and Pacific	0.145	0.0493	0.0527	0.0701	-0.0531	0.0569	0.0865	0.0359	-0.277	-0.0189	-0.304	-0.303
	(0.99)	(0.16)	(0.18)	(0.23)	(-0.15)	(0.19)	(0.29)	(0.12)	(-0.77)	(-0.06)	(-0.73)	(-0.82)
Africa	-0.353***	-0.373	-0.364	-0.310	-0.347	-0.334	-0.319	-0.303	-0.483	-0.536*	-0.281	-0.489
	(-2.76)	(-1.28)	(-1.22)	(-1.05)	(-1.03)	(-1.14)	(-1.09)	(-1.03)	(-1.34)	(-1.75)	(-0.62)	(-1.27)
East Europe and Central Asia	0.137	0.0417	0.0506	0.0734	0.0835	0.117	0.135	0.0656	-0.0456	0.0666	-0.0918	-0.0161
	(0.95)	(0.13)	(0.16)	(0.24)	(0.23)	(0.37)	(0.43)	(0.21)	(-0.12)	(0.20)	(-0.21)	(-0.04)
Latin America & Caribbean	0.209	0.143	0.148	0.146	0.457	0.137	0.165	0.252	-0.0270	0.200	0.140	-0.116
	(1.48)	(0.46)	(0.47)	(0.47)	(1.24)	(0.44)	(0.53)	(0.80)	(-0.07)	(0.60)	(0.30)	(-0.29)
Middle East & North Africa	-0.235	0.202	0.210	0.267	0.173	0.241	0.239	0.203	-0.0189	0.238	0.0451	0.0187
	(-1.39)	(0.53)	(0.55)	(0.70)	(0.40)	(0.63)	(0.63)	(0.53)	(-0.04)	(0.59)	(0.08)	(0.04)
IEG evaluation method (PPAR=1)	-0.0816	0.193	0.192	0.200	0.144	0.171	0.202	0.166	0.350	0.427	0.119	0.304
	(-0.80)	(0.67)	(0.67)	(0.70)	(0.44)	(0.59)	(0.70)	(0.57)	(0.83)	(1.33)	(0.25)	(0.71)
Over estimation of ERR in PAD (1 if PAD ERR>ICR ERR)		-1.192***	-1.192***	-1.181***	-1.326***	-1.174***	-1.178***	-1.208***	-1.282***	-1.193***	-1.391***	-1.278***
		(-6.25)	(-6.25)	(-6.18)	(-5.86)	(-6.12)	(-6.16)	(-6.29)	(-5.68)	(-6.00)	(-5.11)	(-5.52)
ERR difference between PAD & ICR, log (%)		-0.284***	-0.283***	-0.285***	-0.417***	-0.267***	-0.275***	-0.268***	-0.192**	-0.245***	-0.311***	-0.198**
		(-3.40)	(-3.38)	(-3.41)	(-4.00)	(-3.21)	(-3.31)	(-3.22)	(-2.08)	(-2.92)	(-2.59)	(-2.10)
Log net commit. (m$)			0.0116								-0.282	-0.241
			(0.13)								(-1.48)	(-1.50)
Project length				-0.0592							-0.0348	-0.0518
				(-1.17)							(-0.51)	(-0.88)
TTL performance					0.281**						0.309**	
					(2.43)						(1.97)	
Preparation cost over commit, %						-0.0960***					-0.187***	-0.170***
						(-2.95)					(-2.88)	(-2.94)
Supervision cost over commit, %							-0.0577**				0.0328	-0.000607
							(-2.08)				(0.47)	(-0.01)
Potential prob. dummy (1=if raised in the 1rst half of the proj.)								-0.772***			-0.171	-0.293
								(-3.66)			(-0.56)	(-1.11)
Actual prob. dummy (1=if raised in the 1rst half of the proj.)									-1.247***		-1.223***	-1.101***
									(-5.67)		(-4.46)	(-4.63)
Restructuring dummy (1=if restricted in the 1rst half of the proj.)										-0.124	-0.0969	0.0190
										(-0.35)	(-0.22)	(0.05)
Observations	3962	1105	1105	1105	893	1105	1105	1104	878	1021	712	875
R squared	0.061	0.144	0.144	0.145	0.167	0.152	0.148	0.157	0.210	0.149	0.252	0.228
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.10: Logit analysis, 1985-2015, Project Outcome (Dummy for Successful Outcome as independent variable) and ERR Reported in both PAD and ICR as explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc
GDP PC growth	0.0531***	0.0515***	0.0498***	0.0528***	0.0402***	0.0526***	0.0547***	0.0465***	0.0549***
	(4.49)	(4.35)	(4.21)	(4.44)	(2.65)	(4.42)	(4.58)	(3.01)	(4.57)
CPIA rating	0.435***	0.413***	0.363***	0.397***	0.405***	0.389***	0.382***	0.345***	0.358***
	(7.63)	(7.21)	(6.29)	(6.88)	(4.24)	(6.75)	(6.62)	(3.53)	(6.13)
CPIA scale dummy	0.133	0.111	0.146	0.114	0.0555	0.0357	-0.0308	-0.0418	-0.00238
	(1.31)	(1.08)	(1.41)	(1.11)	(0.35)	(0.34)	(-0.29)	(-0.25)	(-0.02)
East Asia and Pacific	0.208*	0.234**	0.272**	0.250**	0.242	0.289**	0.269**	0.314*	0.305**
	(1.76)	(1.97)	(2.27)	(2.10)	(1.42)	(2.40)	(2.24)	(1.81)	(2.52)
Africa	-0.314***	-0.233**	-0.0893	-0.198*	-0.0972	-0.178*	-0.142	0.0996	-0.0900
	(-3.03)	(-2.23)	(-0.84)	(-1.88)	(-0.64)	(-1.68)	(-1.34)	(0.63)	(-0.83)
East Europe and Central Asia	0.0709	0.136	0.277**	0.152	0.228	0.255**	0.271**	0.434**	0.314**
	(0.56)	(1.07)	(2.14)	(1.19)	(1.34)	(1.97)	(2.09)	(2.47)	(2.41)
Latin America & Caribbean	0.0267	0.101	0.164	0.114	0.329**	0.122	0.149	0.442***	0.165
	(0.24)	(0.89)	(1.43)	(1.00)	(2.02)	(1.07)	(1.30)	(2.65)	(1.44)
Middle East & North Africa	-0.139	-0.100	-0.0181	-0.0542	-0.109	-0.0550	-0.0470	0.0428	0.0125
	(-1.03)	(-0.74)	(-0.13)	(-0.40)	(-0.57)	(-0.40)	(-0.34)	(0.22)	(0.09)
IEG evaluation method (PPAR=1)	-0.0878	-0.0470	-0.0827	-0.0582	0.161	-0.0853	-0.0892	0.0787	-0.107
	(-1.18)	(-0.63)	(-1.10)	(-0.78)	(1.42)	(-1.13)	(-1.18)	(0.68)	(-1.41)
ERR present in both PAD & ICR		0.638***	0.566***	0.650***	0.794***	0.604***	0.584***	0.724***	0.588***
		(7.97)	(6.99)	(8.10)	(6.62)	(7.48)	(7.22)	(5.93)	(7.21)
Log net commit. (m$)			0.202***					0.0287	0.0360
			(7.28)					(0.53)	(0.97)
Project length				-0.0443***				-0.0571**	-0.0436**
				(-2.60)				(-2.43)	(-2.53)
TTL performance					0.926***			0.918***	
					(7.72)			(7.53)	
Preparation cost over commit, %						-0.0700***		-0.0305**	-0.0345***
						(-7.78)		(-2.28)	(-3.15)
Supervision cost over commit, %							-0.0639***	-0.0542***	-0.0378***
							(-8.83)	(-3.92)	(-3.63)
Observations	5622	5622	5622	5622	3183	5622	5621	3183	5621
Pseudo R2	0.069	0.078	0.086	0.079	0.091	0.090	0.092	0.115	0.095
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01

Table A.11: Logit analysis, 1985-2015, Project Outcome (Dummy for Successful Outcome as independent variable) and ERR discrepancy between PAD and ICR as explanatory variables (Model 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc	Suc/Unsuc
GDP PC growth	0.0531***	0.0828***	0.0821***	0.0834***	0.0831**	0.0818***	0.0789***	0.0919**	0.0834***
	(4.49)	(2.93)	(2.90)	(2.94)	(2.01)	(2.89)	(2.79)	(2.16)	(2.92)
CPIA rating	0.435***	0.169	0.164	0.162	0.156	0.152	0.149	0.153	0.166
	(7.63)	(1.48)	(1.42)	(1.39)	(0.66)	(1.33)	(1.29)	(0.63)	(1.41)
CPIA scale dummy	0.133	-0.461	-0.460	-0.466	-0.423	-0.533*	-0.547*	-0.607	-0.609*
	(1.31)	(-1.46)	(-1.46)	(-1.48)	(-1.03)	(-1.68)	(-1.72)	(-1.45)	(-1.89)
East Asia and Pacific	0.208*	0.160	0.166	0.166	0.149	0.166	0.178	0.104	0.146
	(1.76)	(0.76)	(0.78)	(0.78)	(0.45)	(0.78)	(0.84)	(0.31)	(0.68)
Africa	-0.314***	-0.319	-0.304	-0.306	-0.271	-0.313	-0.300	-0.372	-0.392*
	(-3.03)	(-1.54)	(-1.43)	(-1.45)	(-0.85)	(-1.51)	(-1.44)	(-1.07)	(-1.80)
East Europe and Central Asia	0.0709	-0.0377	-0.0259	-0.0273	0.157	-0.0129	0.0173	0.0751	-0.0428
	(0.56)	(-0.14)	(-0.10)	(-0.10)	(0.45)	(-0.05)	(0.06)	(0.21)	(-0.16)
Latin America & Caribbean	0.0267	-0.156	-0.151	-0.149	0.497	-0.172	-0.157	0.407	-0.199
	(0.24)	(-0.71)	(-0.69)	(-0.68)	(1.43)	(-0.78)	(-0.72)	(1.15)	(-0.90)
Middle East & North Africa	-0.139	0.0287	0.0414	0.0436	0.221	0.0299	0.0426	0.167	-0.0347
	(-1.03)	(0.11)	(0.15)	(0.16)	(0.53)	(0.11)	(0.16)	(0.38)	(-0.13)
IEG evaluation method (PPAR=1)	-0.0878	-0.282*	-0.285*	-0.285*	0.115	-0.311**	-0.298*	0.0878	-0.303*
	(-1.18)	(-1.83)	(-1.85)	(-1.85)	(0.37)	(-2.01)	(-1.93)	(0.28)	(-1.95)
Over estimation of ERR in PAD (1 if ERR in PAD>ERR in ICR)		-1.359***	-1.356***	-1.357***	-1.298***	-1.343***	-1.341***	-1.265***	-1.351***
		(-8.64)	(-8.60)	(-8.62)	(-5.97)	(-8.51)	(-8.51)	(-5.75)	(-8.53)
ERR difference, log (of absolute difference [ratio])		-0.522***	-0.521***	-0.523***	-0.404***	-0.518***	-0.515***	-0.390***	-0.524***
		(-7.49)	(-7.45)	(-7.49)	(-4.02)	(-7.42)	(-7.41)	(-3.85)	(-7.47)
Log net commit. (m$)			0.0217					-0.260*	-0.154*
			(0.33)					(-1.87)	(-1.77)
Project length				-0.0124				-0.0503	-0.0106
				(-0.34)				(-0.88)	(-0.29)
TTL performance					0.646**			0.667**	
					(2.20)			(2.24)	
Preparation cost over commit, %						-0.0607***		-0.134**	-0.0537
						(-2.63)		(-2.57)	(-1.41)
Supervision cost over commit, %							-0.0560**	-0.00242	-0.0496
							(-2.49)	(-0.05)	(-1.24)
Observations	5622	1820	1820	1820	945	1820	1820	945	1820
Pseudo R2	0.069	0.186	0.186	0.186	0.162	0.190	0.189	0.173	0.192
Sector dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x evaluation period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sector x approval period dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: t statistics in parentheses, * p<0.10, ** p<0.05, *** p<0.01