Independent Evaluation Group

World Bank Project Performance Ratings - Codebook

September 2015

The Independent Evaluation Group (IEG)[1] is an independent unit within the World Bank Group (World Bank, International Finance Corporation, and Multilateral Investment Guarantee Agency) that reports to the Board of Executive Directors. IEG's mission is to improve development results through excellence in evaluation. IEG is responsible for assessing programs and activities, making recommendations to help enhance the Bank Group's development effectiveness, and appraising the Bank Group's self-evaluation and development risk management systems. IEG also aims to improve the Bank Group's work by identifying and disseminating lessons learned from experience.

I. Dataset Description

The **IEG World Bank Project Performance Ratings Dataset** is a collection of the ratings from all World Bank project assessments (or assessment validations) carried out independently by IEG, since it was created back in the 1970s. Spanning over 30 years, the dataset contains ratings from more than 11,260 project assessments covering about 9,400 completed projects; it is the longest-running project performance data collection of its kind. The dataset also includes references to over 6,500 source evaluation documents, including 2,500 recently declassified reports. The World Bank and IEG share a common, objectives-based project evaluation methodology, which looks at project achievements against stated objectives. The methodological approach also assesses the relevance of objectives and design, and the efficiency of resource use, in achieving objectives. IEG assigns ratings of project performance using standardized scales. Rating criteria have evolved through the years, from a single "Outcome" indicator, to the ten core ratings used today. In mid-2005, IEG's project performance ratings criteria and scales were harmonized with those used by the World Bank, to ensure greater consistency between the independent and self-assessment of projects[2].

The World Bank prepares Implementation Completion Reports (ICRs) and assigns ratings for all Board-approved projects. IEG then carries out a review of all ICRs, along with the lending or grant agreement, project appraisal documentation (or program

[1]All unit references in this codebook are made to "IEG", however the unit has had different organizational designations in the past. The Operations Evaluation Department (OED) at the World Bank was officially created in 1973. Also reporting to the Director General-Evaluation, but set under separate entities, the IFC's Operations Evaluation Group (OEG) was established in 1995, and MIGA's Evaluation Unit (OEU) in 2002. In 2005, IEG was created as the merger of these group-wide independent evaluation units.

[2] Except for "ICR Quality", "M&E Quality" and the discontinued ratings, the definitions listed in this codebook correspond to harmonized evaluation criteria for ICR and IEG evaluations, developed in 2005 by a working group comprising staff from OPCS, IEG and Regional Units.

document in the case of development policy lending), and the most recent Country Assistance Strategy. Based on the evidence presented in the ICR, IEG reviews the achievements and lessons from a project and comes to an independent judgment on its performance ratings, which may differ from those in the ICR. IEG's performance ratings for the project are communicated to the Bank, are used for official tracking of the lending portfolio, and then disclosed to the public through this dataset.

For a subset of completed projects (~20-25 percent), IEG performs an in-depth field-based evaluation, called a Project Performance Assessment Report (PPAR). Any ratings obtained on the basis of a PPAR supersede the ratings from an ICR Review, and they are also included in this dataset. PPARs are an important vehicle for learning and are used as building blocks for country, thematic or sector, and corporate evaluation studies.

The **unit of observation** for this dataset is a single project assessment, which as described above can be an independent validation of a completion report (ICR Review), or an in-depth field-based project evaluation (PPAR). In both cases, the same criteria are used to rate performance. All resulting ratings from IEG's project assessments are included in this dataset.

All dataset observations can be uniquely identified by a composite key that includes: a) the related project identifier; b) the assessment/evaluation type; and c) the assessment/evaluation date.

The database will be regularly updated to add ratings from recently reviewed projects – ratings from ICR reviews of projects recently completed and revised ratings of older projects that have been evaluated in a PPAR. In addition, the database will be updated on a quarterly or semiannual basis to address data issues or introduce improvements, in accordance to IEG's Access to Information Policy. The update schedule will be posted on the IEG web site

II. Timeline – Evolution of IEG's World Bank Project-level Assessments & Ratings

Period	Project Evaluation Milestone	Description
1976-1981	First project assessment reports (PARs) are issued	Initially, reports were issued for a sample of the projects, but portfolio coverage increased to almost 100% by 1981. These reports were prepared jointly by operations and evaluation teams. They did not explicitly rate project performance, however imputed values have been recorded in the dataset.
1982-1994	Project Completion Reports (PCRs), with IEG's validation	PCRs were prepared by operations staff and submitted to the Board of Directors, and included a cover memo signed by the

		Director General-Evaluation with IEG's independent verification. The cover memo also contained explicit project performance ratings. In addition, IEG continued to prepare, for a sample of projects, in-depth field-based Project Assessment Reports (PARs)
1995-1997	World Bank's Implementation Completion Reports (ICRs) and IEG's independent validation reports	All self-assessment reports (ICRs) issued by World Bank Operations were independently reviewed by IEG, resulting in separately issued reports. During this period, IEG issued short briefs called "Evaluation Memoranda" (EVM).
1998-2005	ICRs and Evaluation Summaries	Given the increasing numbers of ICRs being issued, IEG introduced structured computerized forms to carry out its independent reviews. The new electronic documents were referred to as "Evaluation Summaries" (ES).
2005-2010	WB and IEG harmonize their evaluation criteria	IEG and WB Operations establish common project performance assessment criteria, in order to arrive to comparable data and ratings. IEG's updated electronic forms are now referred simply as "ICR Reviews".
2011	Specialized ICR Review for Development Policy Loans (DPLs) and Investment Operations (INV)	IEG's project assessment methodology was updated to introduce distinct handling of investment and development policy lending operations.

III. Terms of Use and Proper Citation

The dataset is available to the public free of charge. Users may copy, distribute, adapt, display or include the data in other products for commercial and noncommercial purposes at no cost subject to the following limitations:

- You must always include attribution for the data used, by including proper citations, described below.
- You must not claim or imply that IEG or any other World Bank Group entity endorses your use of the data, or use IEG's or World Bank Group logo(s) or registered mark(s) in conjunction with such use.
- IEG and the World Bank Group make no warranties with respect to the data and you agree that no World Bank Group entity shall be liable to you in connection with your use of the data.

The complete and binding "Terms of Use for Datasets" are available online[3].

Citations should include name and source, as follows: "Independent Evaluation Group - World Bank Project Performance Ratings": http://ieg.worldbankgroup.org/ratings.

IV. Data Definitions

To better explain the structure of this dataset, we have arranged dataset fields in three groups:

- **Project Description Fields**: information that describes the World Bank project or intervention being assessed.
- **Rates of Return**: transcriptions of any available economic rates of return (ERRs) from original project appraisal and/or completion documents.
- **Evaluation Description Fields**: information that describes the underlying documents or reports containing IEG's assessments.
- **Project Performance Ratings**: The specific IEG assessment criteria and ratings.

Definitions – Project Description Fields

- **Project Identification Number**: The project's unique identifier. It is a six-digit number, preceded with a "P". All lending projects and AAA activities and many grants are given a Project ID Number.
- **Project Name:** The name of the project. It is often prefaced with an abbreviation of the country name.
- **Country:** The borrowing country name**.** In the case of a project which occurs in more than one country in a region, the region name is often shown.
- **Country Code:** The corresponding ISO alpha-2 standard country code
- **Region**: The corresponding World Bank administrative geographical region.
    - AFR=Africa
    - EAP=East Asia Pacific
    - ECA=Europe and Central Asia
    - LCR=Latin America and Caribbean
    - MNA=Middle East and North Africa
    - SAR=South Asia

[3] See "Terms of Use for Datasets Listed in The World Bank Data Catalog" on the World Bank's Open Data Initiative website.

    - Other=countries and areas that do not fit prior definitions, or are global (more than one region).
- **Approval Date**: The date the project or intervention was approved. The Board of Directors approves lending operations.
- **Approval Year:** This is the fiscal year that the project was approved for funding. The World Bank's fiscal year begins July 1st and ends June 30th.
- **Deactivation Date**: The date the project exits the active (supervised) portfolio. This date is determined as the earlier of either the closing date or full disbursement of all underlying loans - whichever happens first.
- **Exit Year**: The fiscal year in which the project actually closes.
- **Total Project Cost**: The estimated total cost in USD of the project, including non-Bank funding not specifically itemized in the Bank appraisal or completion reports.
- **Lending Instrument Type**: There are two basic types of lending instruments: Investment Loans, which have a long-term focus (5-10 years) and Development Policy Loans (DPLs), which have a short-term focus (1-3 years).
- **Agreement Type:** The type of legal instrument being signed by the Borrower and World Bank. Its value is derived from the product line associated with the project. Possible values are:
    - CARB - Carbon Initiative (Product line CN)
    - DRF - Debt Reduction Facility (DR)
    - GEF - Global Environment Fund (GE)
    - GEFM - Global Environment Fund Medium (GM)
    - GUAR - Guarantees (GU)
    - HIPC - Highly Indebted Poor Countries (HT)
    - IBRD - Lending IBRD amount greater than zero (PE)
    - IDA - Lending IDA amount greater than zero (PE)
    - IDF - Institutional Development Fund (IF)
    - MONT - Montreal Protocol (MT)
    - RAIN - Rainforest Initiative (RN)
    - RETF - Recipient Executed Trust Fund (RE)
    - SPF - Special Fund (SF)
- **Lending Instrument**: The type of loan provided by IBRD, or credit by IDA, to the Borrower. In addition to loans, the World Bank offers are other product lines (see “Product Lines” below).
        - ***Investment Loans***: Finance goods, works, and services in support of specific economic and social development objectives in a broad range of sectors.
            - APL - Adaptable Program Loan: Provides phased support for long-term development programs. They involve a series of loans that build on the lessons learned from the previous loan(s) in the series.

APLs are used when sustained changes in institutions, organizations, or behavior are key to successfully implementing a program.

- FIL - Financial Intermediary Loan: Provides long-term resources to local financial institutions to finance real sector investment needs. The financial institutions assume credit risk.
- ERL - Emergency Recovery Loan: Supports economic and social recovery immediately after an extraordinary event - such as war, civil disturbance, or natural disaster - that seriously disrupts a borrower's economy, or support programs to avert deterioration in economic and social services in the face of imminent emergency.
- SIM - Sector Investment and Maintenance Loan: Focuses on public expenditure programs in particular sectors. They aim to bring sector expenditures, policies, and performance in line with a country's development priorities by helping to create an appropriate balance among new capital investments, rehabilitation, reconstruction, and maintenance.
- SIL - Specific Investment Loan: Supports the creation, rehabilitation, and maintenance of economic, social, and institutional infrastructure.
- LIL - Learning and Innovation Loan: Loans of $5 million or less financing small, experimental, risky and/or time-sensitive projects in order to pilot promising initiatives and build consensus around them, or experiment with an approach in order to develop locally based models prior to a larger-scale intervention
  TAL - Technical Assistance Loan: Used to build institutional capacity in the borrower country. It may focus on organizational arrangements, staffing methods, and technical, physical, or financial resources in key agencies.

- ***Development Policy Loans****:* provide rapidly-disbursing financing to help a borrower address actual or anticipated development financing requirements of domestic or external origins. A DPL aims to help the borrower achieve sustainable poverty reduction through a program of policy and institutional actions, for example, strengthening public financial management, improving the investment climate, addressing bottlenecks to improve service delivery, and diversifying the economy.

- **Product Lines and Codes:** The World Bank offers several product lines to its clients - some primarily funded by loans and some by grants. Loans (product line code “PE”) are offered via two basic types of lending instruments - development

policy and investment, which are described above. Others major lines appearing on the dataset are:

- Global Environment Projects (GE)
- Guarantees (GU)
- IBRD/IDA (PE)
- Montreal Protocol (MT)
- Project Evaluations (EP)
- Rainforest (RN)
- Recipient Executed Activities (RE)
- Special Financing (SP)

- **Sector Board (historical)**: Until 2014, Sector Boards were bodies responsible for establishing and maintaining technical standards for a discipline, or a group of disciplines, at the World Bank. This field designates the Sector Board that had primary responsibility for the quality of the project/activity. In 2014, Sector Boards were superseded by new Global Practice Groups (GPGs).

    - List of Sector Boards:
        - Agriculture and Rural Development
        - Economic Policy
        - Education
        - Energy and Mining
        - Environment
        - Financial and Private Sector
        - Financial Management
        - Gender
        - Global Information/Communications Technology
        - Health, Nutrition and Population
        - Poverty Reduction
        - Procurement
        - Public Sector Governance
        - Social Development
        - Social Protection
        - Transport
        - Urban

Definitions – Rates of Return

During their review, IEG Evaluators gather and capture any available "rate of return" estimates at project appraisal (design phase) or at completion; they essentially transcribe rates from any cost-benefit analysis documented either in Project Appraisal Documents

(PADs) or in Implementation Completion Reports (ICRs), respectively. IEG doesn't perform re-estimates of cost-benefit analysis in project completion validation exercises. In some cases, only a fraction of project components are amenable to ERR cost-benefit analysis, so rates of return are only offered for a portion of the project's overall commitment. IEG reports these rates only if they correspond to at least 20% of project commitment.

a) ERR at Appraisal

Definition: Transcription of any economic rate of return found on project appraisal documentation (PADs). Rates of return are only reported if the underlying cost-benefit analysis performed covers or corresponds to at least 20% of committed funding.
Variable: ERRatAppraisal
Type: Alphanumeric String
Frequency Distribution: (as of August 2015)

ERR at Appraisal - Distribution Table

	Exit_FY (group)						
ERRatAppraisal (bin)	**0000**	**1960s**	**1970s**	**1980s**	**1990s**	**2000s**	**2010s**
Null	13	5	296	1,016	1,880	2,501	1,088
0		2	14	80	56	9	9
10	2	2	293	653	491	281	138
20		1	142	451	399	284	119
30			38	142	212	124	63
40			9	51	71	50	24
50			8	37	36	33	7
60			1	31	16	14	5
70				14	10	6	6
80				7	3	6	2
90			1	3	7	2	
100			1	16	3	6	
110						2	3
120							2
130						1	1
140				2	1		
150				1		1	1
160				1			
210						1	1
220						1	
240						1	
250						1	
460						1	
540					1		

-NULL indicates number of project assessments without an "ERR at Appraisal"
-Each bin label designates the lower limit of the range of numbers that is assigned to the bin. Note that the lower limit is inclusive. For example, the bin la-

b) ERR at Completion

Definition: Transcription of any economic rate of return found on project completion documents (or ICRs). Rates of return are only reported if the underlying cost-benefit analysis performed covers or corresponds to at least 20% of committed funding.
Variable: ERRatCompletion
Type: Alphanumeric String
Frequency Distribution: (as of August 5, 2015)

ERR at Completion - Distribution Table

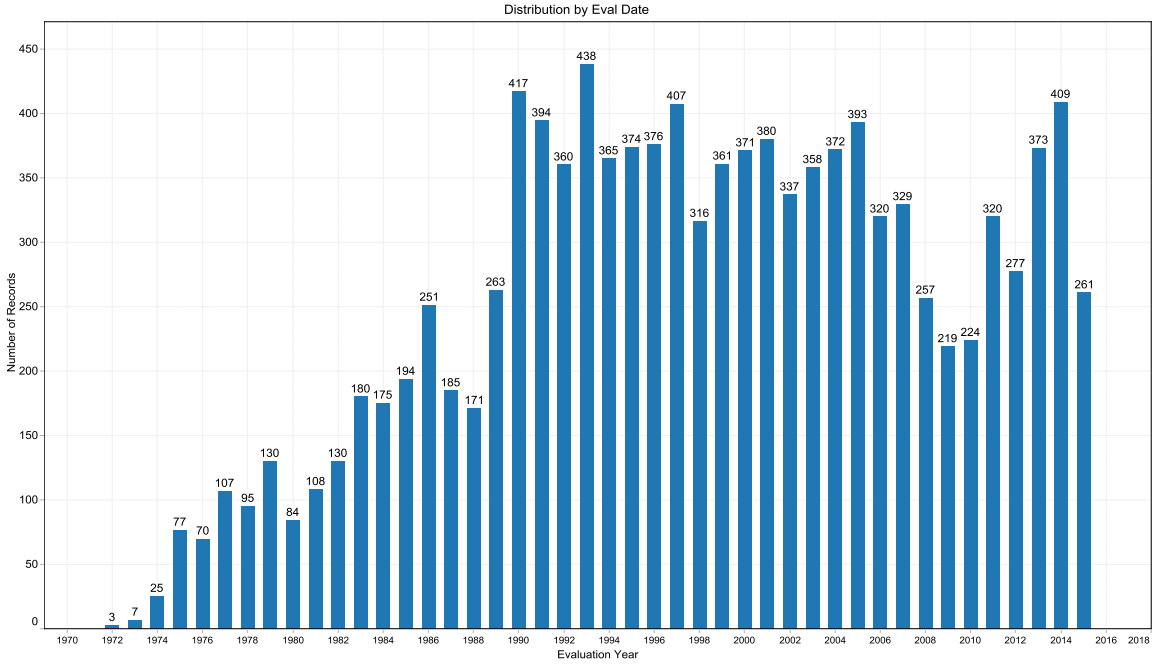
	Exit_FY (group)						
ERRatCompletion (bin)	**0000**	**1960s**	**1970s**	**1980s**	**1990s**	**2000s**	**2010s**
Null	14	2	321	1,213	2,116	2,625	1,106
-100						1	
-30						1	
-20						1	
-10			24	141	26	1	1
0	1	1	67	276	164	37	21
10		6	223	506	409	230	112
20		1	103	216	236	212	120
30			41	86	126	80	44
40			14	27	55	53	21
50			6	14	19	20	16
60				7	12	14	5
70			2	4	3	18	5
80				1	6	1	2
90			1	1	4	8	3
100				6	4	4	1
110					2	1	3
120				2		4	1
130			1	1	1	2	
140				2		3	2
150				2			
160						2	
170					2		
190						1	1
200					1		1
220						1	1
250						1	
280						1	
290						1	
300						1	
310							1
620							1
670						1	
740							1

-NULL indicates number of project assessments without an "ERR at Completion"
-Each bin label designates the lower limit of the range of numbers that is assigned to the bin. Note that the lower limit is inclusive. For example, the bin la-

Definitions - Evaluation Description Fields

a) IEG Evaluation Date

Definition: Describes the date of the evaluation or assessment report where project ratings are being drawn from.
Variable: IEG_EvalDate
Type: Date
Frequency distribution:



(as of August 5, 2015)

b) IEG Evaluation Type

Definition: Describes the type of evaluation document that is sourcing project ratings.
Variable: IEG_EvalType
Type: Alpha string
Values: It could be any of the following:

PCR: Project Completion Report
PCN: Project Cancellation Note
PCM: Project Cancellation Memorandum
EVM: Evaluation Memorandum
ES: Evaluation Summary, and more recently "ICR Reviews"
PAR: Project Performance Assessment Report
CSSR: Country-specific Sector Study Report

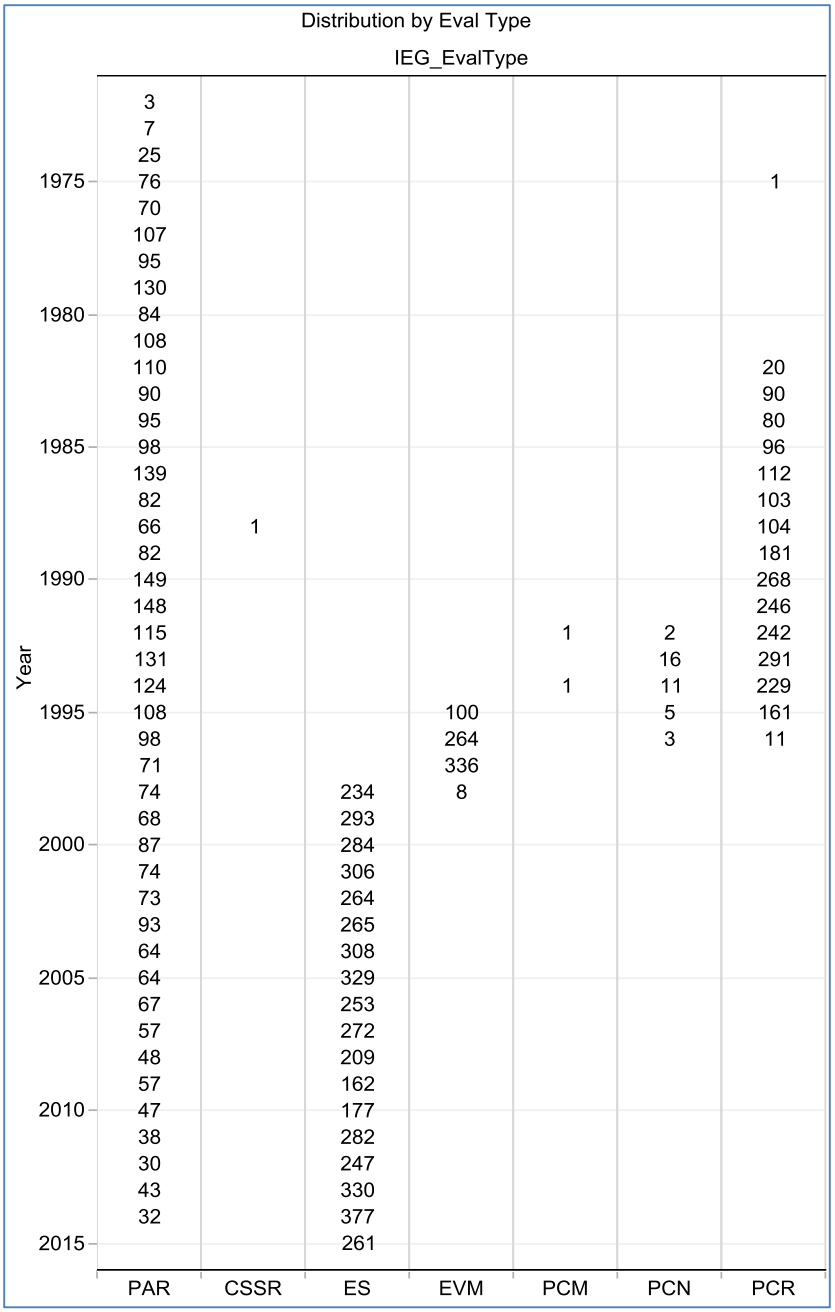
Notes:

Since its early years, IEG has been producing Project Performance Assessment Reports (PPARs), which are field-based project-level evaluations.

In the early 1980s, Operations teams began submitting Project Completion Reports (PCRs) to IEG, which in turn would review them, introducing a cover memo with its own independent assessment including project performance ratings. The updated document was then sent to the Board of Directors. This reporting mechanisms was further refined and expanded in 1994-1995, with the introduction of separate evaluation documents: a) a self-assessment completion report (ICRs) prepared by operational teams; and b) IEG's independent validation of the self-assessment - initially called Evaluation Memoranda (EVMs), later replaced by IEG Evaluation Summaries (ES) in 1998, and since 2000 they are simply called "ICR Reviews".

Project cancellation during the early 90s were documented and rated through PCNs and PCMs documents. Today all project completions (including cancellations) will trigger an independently assessed ICR Review.

Frequency distribution:



(as of August 5, 2015)

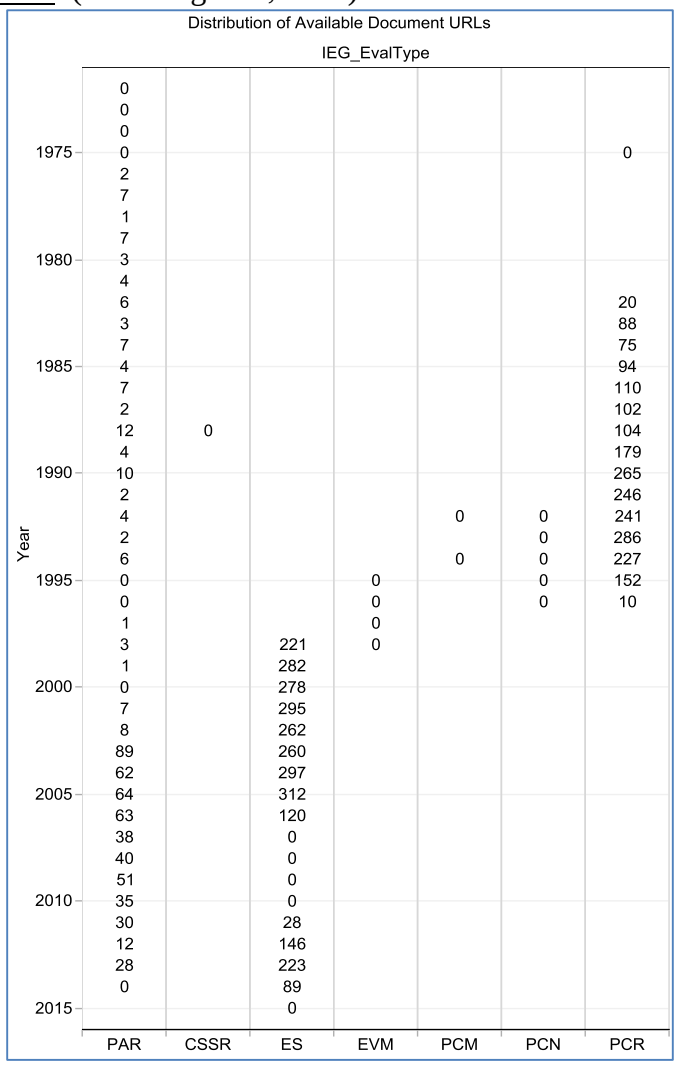
c) IEG Evaluation Document Link

The URL to the source IEG document, if available.

Variable: IEG_SourceDocumentURL
Type: URL
Notes:
IEG began disclosing all PPARs issued after 2002, when its first Disclosure Policy became effective. Self-assessment validations (ICR Reviews), however, remained as

internal undisclosed documents. Under IEG's new Access to Information policy, effective on July 1, 2011, all new evaluation documents prepared by IEG will be disclosed, unless they fall into specific, narrowly defined exceptions[4]. Earlier non-disclosed documents are eligible for declassification provided they don't meet the criteria for disclosure exceptions. Newly disclosed reports, as well as any declassified documents, will be linked through this field.

Frequency distribution: (as of August 5, 2015)



[4] See IEG's Access to Information Policy and implementation roadmap.

Definitions – Project Performance Ratings[5]

a) Outcome Rating

Definition**:** *the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently.* The project outcome rating is thus a composite rating based on three separate criteria – the relevance of the project's objectives and design, achievement of the objectives (efficacy), and efficiency.
Variable: IEG_Outcome
Type: Alpha string
Rating Scale:

Highly Satisfactory	There were **no** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Satisfactory	There were **minor** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Moderately Satisfactory	There were **moderate** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Moderately Unsatisfactory	There were **significant** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Unsatisfactory	There were **major** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.
Highly Unsatisfactory	There were **severe** shortcomings in the operation's achievement of its objectives, in its efficiency, or in its relevance.

In the rare instances where lack of sufficient information or other circumstances make it impossible to assign one of the above ratings, "*Not-rated*" is assigned.

Notes**:**
"Development Outcome", or simply "Outcome", was the first performance rating introduced by IEG in the 80s. Early project performance reviews (PPARs prepared in 1970-1980) didn't actually codify a rating; however a value was imputed for them when the ratings database was set up during the mid-1980s. Initially, a basic "Satisfactory" – "Unsatisfactory" scale was used.

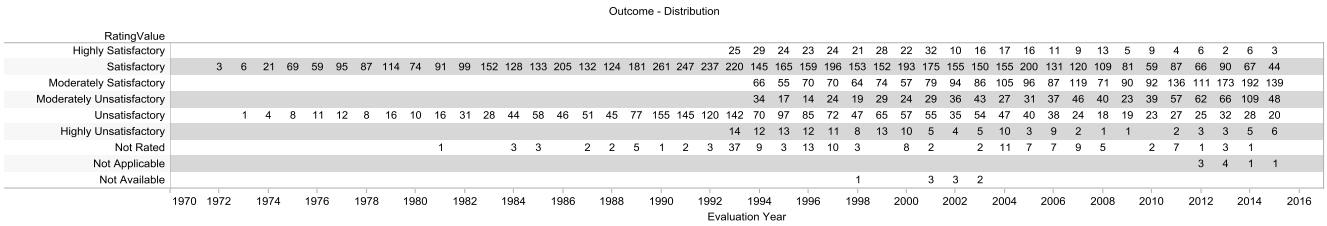
[5] This codebook includes basic ratings definitions, related data patterns, and any known discontinuities or biases. It doesn't provide extensive methodological details on how evaluators reach their conclusions and rate performance. Please refer to the document "Harmonized Evaluation Criteria for ICR and OED Evaluations", prepared jointly by IEG and the World Bank's Operations Core Services (OPCS) unit, and available on IEG's website.

In 1982 Project Completion Reports (PCRs) where introduced. These reports where prepared by operations, reviewed by IEG, and then submitted to the Board with an IEG cover note summarizing its independent assessment and performance ratings.

In 1993, the Outcome rating scale was expanded with new "Highly Satisfactory" and "High Unsatisfactory" points, in order to accommodate a broader gradation of judgments. A year later, the scale was expanded again to provide more nuanced distinctions near the center with the inclusion of "Moderately Satisfactory" and "Moderately Unsatisfactory" options. Implementation Completion Reports (ICRs), the self-assessment instrument currently in use by operations, were also introduced in 1994 and included the same 6-point Outcome rating.

Due to an error in coding form, a handful of data points from the late 90s are coded as "Marginally Satisfactory" and "Marginally Unsatisfactory". These should be interpreted as "Moderately Satisfactory" and "Moderately Unsatisfactory", respectively.

Frequency distribution:



(as of August 5, 2015)

b) Risk To Development Outcome (RDO)

Definition**:** *the risk, at the time of evaluation, that development outcomes (or expected outcomes) will not be maintained (or realized)*
Variable: IEG_RDO
Type: Alpha string
Rating Scale*:*

(i) *Negligible to Low*
(ii) Moderate
(iii) Significant
(iv) High

Whenever the lack of sufficient information, or other circumstances, makes it impossible to assign one of the above ratings, "*Non-evaluable*" should be recorded.

Notes**:**
RDO was introduced in 2007 as a replacement for the prior "Sustainability" rating.

Frequency Distribution:

RDO - Distribution

RatingValue	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
NEGLIGIBLE TO LOW	1		52	60	29	29	36	33	42	36	25	
MODERATE	1		121	109	112	89	132	110	141	159	107	
SIGNIFICANT			75	60	61	83	112	85	126	154	92	
HIGH	1		20	19	17	20	33	43	57	54	35	
NON-EVALUABLE			9	6		3	7	1	3	3	1	
Not Applicable								3	4	1	1	
LOW								1		2		

Evaluation Year

(as of August 5, 2015)

c) Bank Performance

Definition: *the extent to which services provided by the Bank ensured quality at entry of the operation and supported effective implementation through appropriate supervision (including ensuring adequate transition arrangements for regular operation of supported activities after loan/credit closing), towards the achievement of development outcomes.*

An *Overall Bank Performance* rating is determined by assessing two sub-dimensions, (i) Bank performance in ensuring quality at entry; and (ii) quality of Bank supervision.

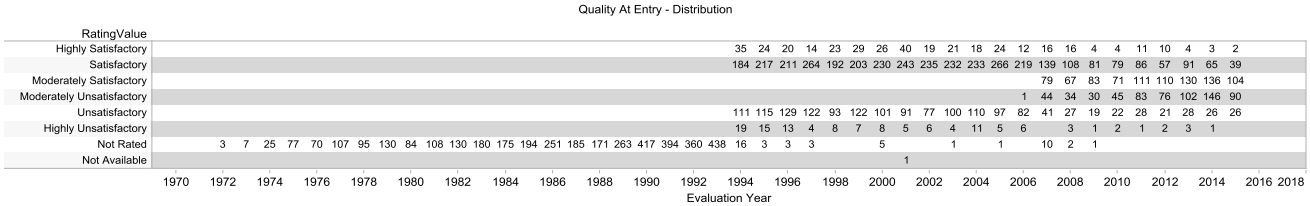
c.1) Quality at Entry

Definition**:** The extent to which the Bank identified, facilitated preparation of, and appraised the operation such that it was most likely to achieve planned development outcomes and was consistent with the Bank's fiduciary role.
Variable: IEG_BankQualityAtEntry
Type: Alpha string
Rating Scale:

Highly Satisfactory	There were **no** shortcomings in identification, preparation, or appraisal.
Satisfactory	There were **minor** shortcomings in identification, preparation, or appraisal
Moderately Satisfactory	There were **moderate** shortcomings in identification, preparation, or appraisal
Moderately Unsatisfactory	There were **significant** shortcomings in identification, preparation, or appraisal
Unsatisfactory	There were **major** shortcomings in identification, preparation, or appraisal
Highly Unsatisfactory	There were **severe** shortcomings in identification, preparation, or appraisal

Notes:
The "Quality at Entry" rating was introduced in 1994, as a four-point scale rating (Highly Satisfactory, Satisfactory, Unsatisfactory, and Highly Unsatisfactory). On July 1, 2006, the rating was expanded to its current 6-point scale.

Frequency distribution:



(as of August 5, 2015)

c.2) Quality of Supervision

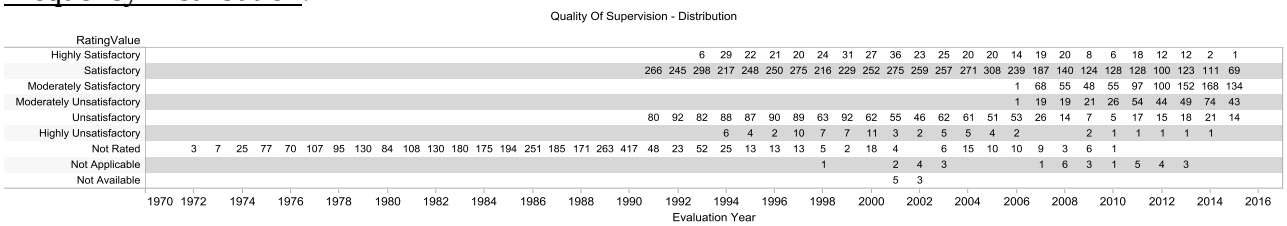
Definition:

Quality of supervision refers to the extent to which the Bank proactively identified and resolved threats to the achievement of relevant development outcomes and the Bank's fiduciary role.
Variable: IEG_BankQualityOfSupervision
Type: Alpha string
Rating Scale:

Highly Satisfactory	There were **no** shortcomings in the proactive identification of opportunities and resolution of threats.
Satisfactory	There were **minor** shortcomings in the proactive identification of opportunities and resolution of threats.
Moderately Satisfactory	There were **moderate** shortcomings in the proactive identification of opportunities and resolution of threats.
Moderately Unsatisfactory	There were **significant** shortcomings in the proactive identification of opportunities and resolution of threats.
Unsatisfactory	There were **major** shortcomings in the proactive identification of opportunities and resolution of threats.
Highly Unsatisfactory	There were **severe** shortcomings in the proactive identification of opportunities and resolution of threats.

Notes**:**
The "Quality of Supervision" rating was introduced in 1991 as a four-point scale rating (Highly Satisfactory, Satisfactory, Unsatisfactory, and Highly Unsatisfactory), and expanded on July 1, 2006 to its current 6-point scale.

Frequency Distribution:



(as of August 5, 2015)

c.3) Overall Bank Performance

Definition:

The rating of overall *Bank Performance* is based on the ratings for each of the two dimensions (i) Bank performance in ensuring quality at entry; and (ii) the quality of supervision. Quality at entry and quality of supervision should each be rated using their respective six-point rating scales, and for transparency, the individual ratings for quality at entry and quality of supervision should be presented separately. The quality at entry and quality of supervision ratings should be combined into a rating of overall *Bank Performance*.

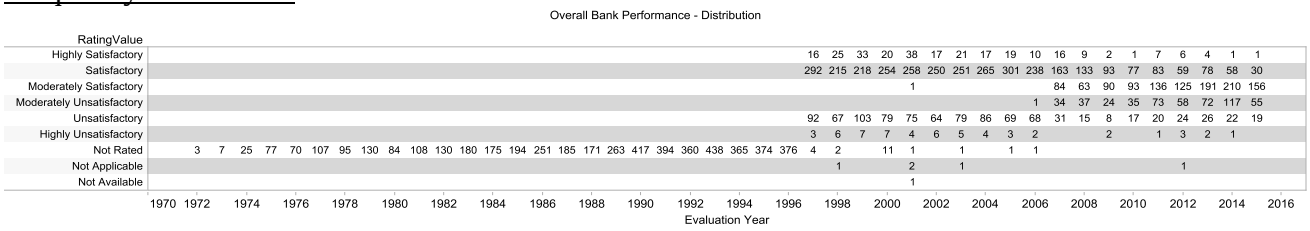
Variable: IEG_OverallBankPerf
Type: Alpha string
Rating scale:

Highly Satisfactory	Bank performance was rated Highly Satisfactory on both dimensions.
Satisfactory	Bank performance was rated Satisfactory on both dimensions, OR was rated Satisfactory on one dimension and Highly Satisfactory on the other dimension.
Moderately Satisfactory	Bank performance was rated Moderately Satisfactory on both dimensions, OR was rated Moderately Satisfactory on one dimension and Satisfactory or Highly Satisfactory on the other dimension. (Also see guidance below.)
Moderately Unsatisfactory	Bank performance was rated Moderately Unsatisfactory on both dimensions. (Also see guidance below.)
Unsatisfactory	Bank performance was rated Unsatisfactory on both dimensions, OR was rated Unsatisfactory on one dimension and Moderately Unsatisfactory on the other dimension.
Highly Unsatisfactory	Bank performance was rated Highly Unsatisfactory on both dimensions, or was rated Moderately Unsatisfactory or Unsatisfactory on one dimension and Highly Unsatisfactory on the other dimension.

Notes**:**

- The “Overall Bank Performance” rating was introduced in 1997 as a four-point scale rating (Highly Satisfactory, Satisfactory, Unsatisfactory, and Highly Unsatisfactory), and expanded on July 1, 2006 to its current 6-point scale.
- Please keep in mind that if one sub-rating of Bank performance is in the satisfactory range and the other is in the unsatisfactory range, evaluators are asked to determine overall Bank performance base on the project outcome rating. This guideline makes overall Bank performance dependent on outcome rating, and will distort attempts to regress outcome rating on Bank performance.

Frequency distribution:



(as of August 5, 2015)

d) Borrower Performance

Definition: *the extent to which the borrower (including the government and implementing agency or agencies) ensured quality of preparation and implementation, and complied with covenants and agreements, towards the achievement of development outcomes.*

An *Overall Borrower Performance* is rated by assessing two dimensions, (i) government performance (central and/or local government as relevant); and (ii) implementing agency or agencies performance. Where the government and implementing agency are indistinguishable, as is often the case for development policy operations, only an overall rating is necessary, taking into account relevant criteria from both lists below.

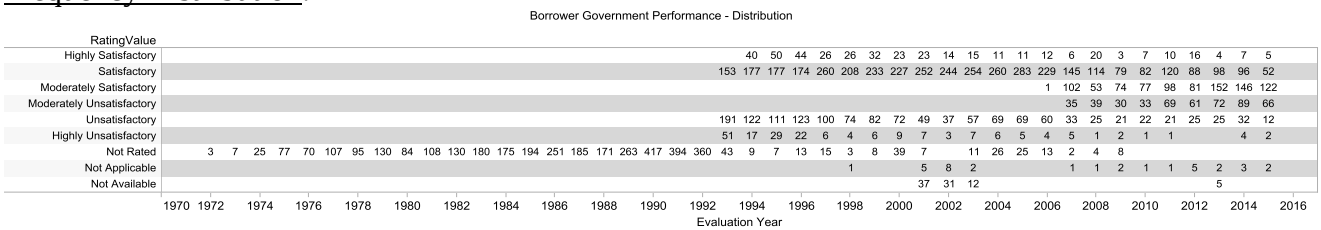
d.1) Government Performance

Variable: IEG_BorrCompliance
Type: Alpha string
Rating scale:

Highly Satisfactory	There were **no** shortcomings in government performance
Satisfactory	There were **minor** shortcomings in government performance
Moderately Satisfactory	There were **moderate** shortcomings in government performance
Moderately Unsatisfactory	There were **significant** shortcomings in government performance
Unsatisfactory	There were **major** shortcomings in government performance
Highly Unsatisfactory	There were **severe** shortcomings in government performance

Notes**:**
The current six-point scale "Government Performance" sub-rating was introduced on July 1, 2006. It replaced a 4-point scale (Highly Satisfactory, Satisfactory, Unsatisfactory, and Highly Unsatisfactory) "Government Compliance" rating, in use since 1997.

Frequency Distribution:



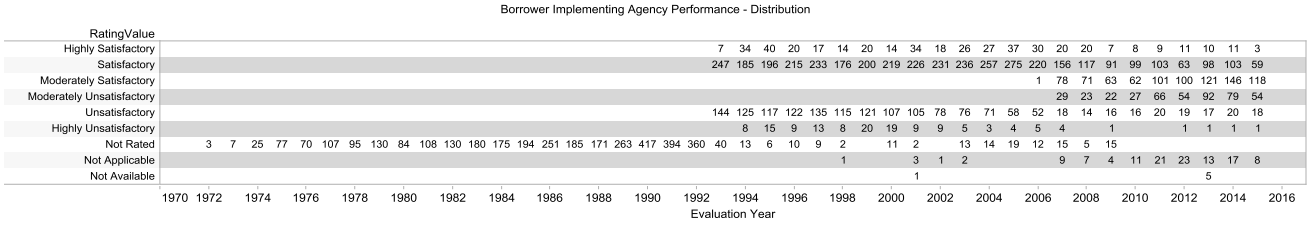
(as of August 5, 2015)

d.2) Implementing Agency or Agencies' Performance

Variable: IEG_BorrImplementation
Type: Alpha string
Rating scale:

Highly Satisfactory	There were **no** shortcomings in implementing agency or agencies' performance
Satisfactory	There were **minor** shortcomings in implementing agency or agencies' performance
Moderately Satisfactory	There were **moderate** shortcomings in implementing agency or agencies' performance
Moderately Unsatisfactory	There were **significant** shortcomings in implementing agency or agencies' performance
Unsatisfactory	There were **major** shortcomings in implementing agency or agencies' performance
Highly Unsatisfactory	There were **severe** shortcomings in implementing agency or agencies' performance

Frequency distribution:



(as of August 5, 2015)

d.3) Overall Borrower Performance

Definition:

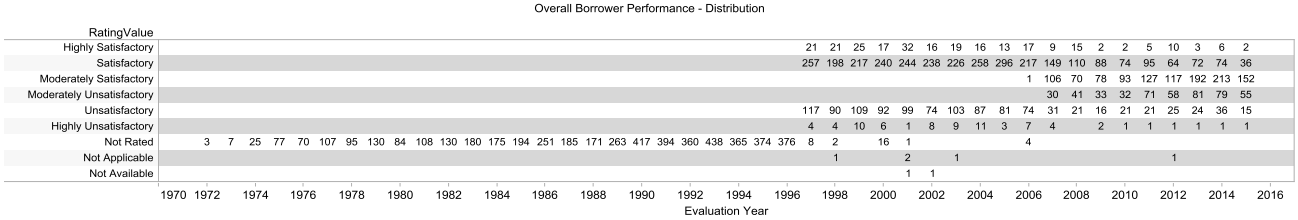
The rating of *Borrower Performance* is based on the ratings for each of the two dimensions (i) government performance; and (ii) implementing agency or agencies' performance. Government performance and implementing agency or agencies' performance should each be rated using their respective six-point rating scales, and for transparency, the individual ratings for them should be presented separately. The ratings for government performance and implementing agency or agencies' performance should be combined into a rating of *Borrower Performance*, per guidance below.
Variable: IEG_OverallBorrPerf
Type: Alpha string
Rating scale:

Highly Satisfactory	Performance was rated *Highly Satisfactory* on both dimensions.
S*atisfactory*	Performance was rated *Satisfactory* on both dimensions, OR was rated *Satisfactory* on one dimension and *Highly Satisfactory* on the other dimension.
Moderately Satisfactory	Performance was rated *Moderately Satisfactory* on both dimensions, OR was rated *Moderately Satisfactory* on one dimension and *Satisfactory* or *Highly Satisfactory* on the other dimension. (Also see guidance below.)
Moderately Unsatisfactory	Performance was rated *Moderately Unsatisfactory* on both dimensions. (Also see guidance below.)
Unsatisfactory	Performance was rated *Unsatisfactory* on both dimensions, OR was rated *Unsatisfactory* on one dimension and *Moderately Unsatisfactory* on the other dimension.
Highly Unsatisfactory	Performance was rated *Highly Unsatisfactory* on both dimensions, OR was rated *Moderately Unsatisfactory* or *Unsatisfactory* on one dimension and *Highly Unsatisfactory* on the other dimension.

Notes**:**

- The "Overall Borrower Performance" rating was introduced in 1997 as a four-point scale rating (Highly Satisfactory, Satisfactory, Unsatisfactory, and Highly Unsatisfactory), and expanded on July 1, 2006 to its current 6-point scale.
- •Please keep in mind that if one sub-rating of Borrower performance is in the satisfactory range and the other is in the unsatisfactory range, evaluators are asked to determine overall Borrower performance base on the project outcome rating. This guideline makes overall Borrower performance dependent on outcome rating, and will distort attempts to regress outcome rating on Borrower performance.

Frequency distribution:



(as of August 5, 2015)

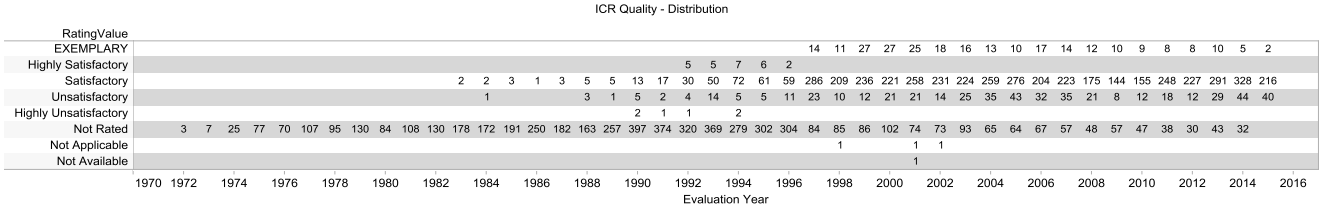
e) Quality of the Implementation Completion Report (ICR)

Definition: As the ICR review is almost entirely based on the information found in the ICR, the reliability of IEG's ratings based on the desk review depends critically on the accuracy and quality of the evidence provided in the ICR. For this reason, IEG rates the quality of the Implementation Completion and Results Report.
Variable: IEG_ICRQuality
Type: Alpha string
Rating scale:
The quality of the ICR is rated on a three-point scale:

(i) Exemplary,
(ii) Satisfactory, or
(iii)Unsatisfactory.

Notes**:**
The "ICR Quality" rating was introduced in 1997 on a 3-point scale. The database however, also contains assessments flagged as "ICR Quality" that predate 1997, and even the 1994 introduction of ICRs. These assessments are coded on an old 4-point scale (Highly Satisfactory, Satisfactory, Unsatisfactory, Highly Unsatisfactory), and are likely to represent quality assessments of Project Completion Reports (PCRs), which precede ICRs.

Frequency distribution:



(as of August 5, 2015)

f) Quality of Monitoring and Evaluation (M&E)

Definition**:** The M&E quality rating is based on an assessment of three main elements: (i) M&E design; (ii) M&E implementation; and (iii) use of M&E data. Monitoring and evaluation are distinct and the rating is informed by both the quality of monitoring and the quality of evaluation. The evaluator is asked to discuss separately each of the three elements of M&E quality and to arrive at an overall *quality of M&E rating* on a 4-point scale.
Variable: IEG_MEQuality
Type: Alpha string
Rating scale:
The quality of M&E is rated on a four-point scale:

High	M&E with a strong influence on project performance and with very few weaknesses. M&E rated "High" would be more than sufficient to satisfactorily assess the achievement of the stated objectives and would contribute to testing the logic of the theory of change to a high level. A "High" overall rating would be a best practice case.
Substantial	Would not be considered a best practice, but it would be an effective and efficient M&E performance with only some weaknesses in a few areas and still with significant impact on the project or program. M&E rated "Substantial" would be sufficient to satisfactorily assess the achievement of the stated objectives and to test the logic of the theory of change.
Modest	An M&E performance with significant weaknesses in a number of areas and with limited impact on the project or program.
Negligible	An M&E performance with many weaknesses and very little impact on the project or program.

Notes:

- The M&E Quality rating was introduced on July 1, 2006 as a four-point scale.
- Please note that when insufficient M&E information is provided by the Bank for IEG to arrive at a clear rating, IEG will downgrade relevant ratings. This rule introduces a dependency between Outcome and M&E Quality ratings in projects with poor M&E data.

Frequency Distribution:

M&E Quality - Distribution

RatingValue	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
HIGH			6	13	5	5	9	6	8	5	2	
SUBSTANTIAL			79	68	75	40	92	77	91	100	67	
MODEST		2	125	112	79	102	164	151	196	220	156	
NEGLIGIBLE		1	39	50	30	28	47	39	59	50	36	
NON-EVALUABLE			7	5	2	3	6	2	2	1		

Evaluation Year

(as of August 5, 2015)

g) Institutional Development Impact (Discontinued)

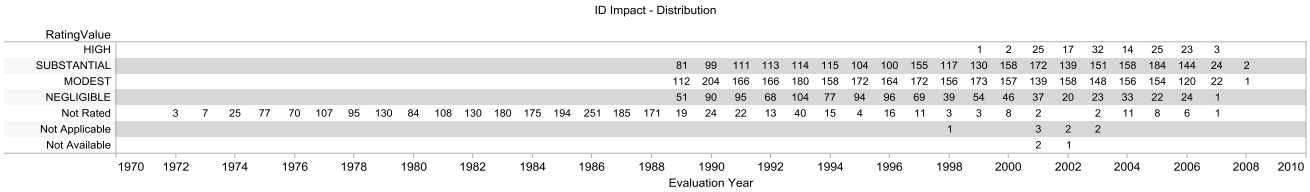
Definition: The extent to which a project improves the ability of a country or region to make more efficient, equitable and sustainable use of its human, financial, and natural resources through: (a) better definition, stability, transparency, enforceability, and predictability of institutional arrangements and/or (b) better alignment of the mission and capacity of an organization with its mandate, which derives from these institutional arrangements. IDI includes both intended and unintended effects of a project.

This rating was applied to all projects, irrespective of whether they had an explicit objective of institutional development. It was discontinued on July 1, 2006, based on the reasoning that projects with institutional development impact objectives were already being assessed on that objective as part of the outcome rating, precluding the need for a separate rating.
Variable: (disc)IEG_IDImpact
Type: Alpha string
Rating scale:
A four-point scale was used, as follows:

High	Project as a whole made, or is expected to make, a critical contribution to the country's/region's ability to effectively use human, financial, and natural resources, either through the achievement of the project's stated ID objectives or through unintended effects
Substantial	Project as a whole made, or is expected to make, a significant contribution to the country's/region's ability to effectively use human, financial, and natural resources, either through the achievement of the project's stated ID objectives or through unintended effects
Modest	Project as a whole increased, or is expected to increase, to a limited extent the country's/region's ability to effectively use human, financial, and natural resources, either through the achievement of the project's stated ID objectives or through unintended effects
Negligible	Project as a whole made, or is expected to make, little or no contribution to the country's/region's ability to effectively use human, financial, and natural resources, either through the achievement of the project's stated ID objectives or through unintended effects

Notes**:**
Was introduced in 1989, and has always used a four-point scale.

Frequency Distribution:



(as of August 5, 2015)

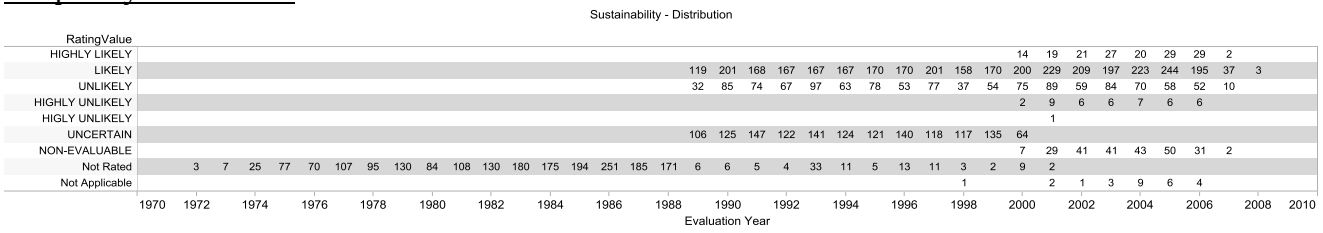
h) Sustainability (Discontinued)

Definition**:** At the time of the evaluation, the resilience to risk of future net benefits flows. This rating was superseded in July 2006 by "Risk to Development Outcomes".
Variable: (disc)IEG_Sustainability
Type: Alpha string
Rating scale:

Highly Likely	Project net benefits flow meets most of the relevant factors determining overall resilience at the "high level", with all others rated at the "substantial" level
Likely	Project net benefits flow meets all relevant factors determining overall resilience at the "substantial" level
Unlikely	Project net benefits flow meets some but not all relevant factors determining overall resilience at the "substantial" level
Highly Unlikely	Project net benefits flow meets few of the relevant factors determining overall resilience at the "substantial" level
Not Evaluable	Insufficient information available to make a judgment

Note**:**
Was introduced in 1989 as a three-point scale rating ("Likely", "Unlikely", and "Uncertain"). Changed to the 4-point scale above in 2000, and superseded by Risk to Development Outcome in 2006.

Frequency distribution:



(as of August 5, 2015)

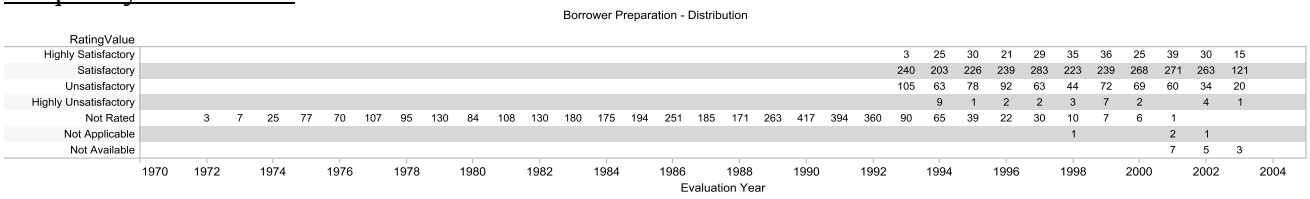
i) Borrower Preparation (Discontinued)

Definition: Rate the government/implementing agency performance on the preparation of this project. Consider specifically whether the government/implementing agency took account of economic, financial, technical, policy, and resource considerations, and ensured participation of major stakeholders in preparing the project. This rating was discontinued in 2005, with the introduction of more implementation-oriented government and execution agency ratings.
Variable: (disc)IEG_BorrPrep
Type: Alpha string
Rating scale:

Highly Satisfactory	The government/implementing agency fully took account of economic, financial, technical, policy, resource and participation considerations in preparing the project
Satisfactory	The government/implementing agency generally took account of economic, financial, technical, policy, resource and participation considerations in preparing the project
Unsatisfactory	The government/implementing agency took account only to a limited extent of economic, financial, technical, policy, resource, and participation considerations in preparing the project
Highly Unsatisfactory	The government/implementing agency largely failed to take account of economic, financial, technical, policy, resource and participation considerations in preparing the project

Notes**:**
"Borrower Preparation" was introduced in 1993 and discontinued in 2003.

Frequency distribution:



(as of August 5, 2015)

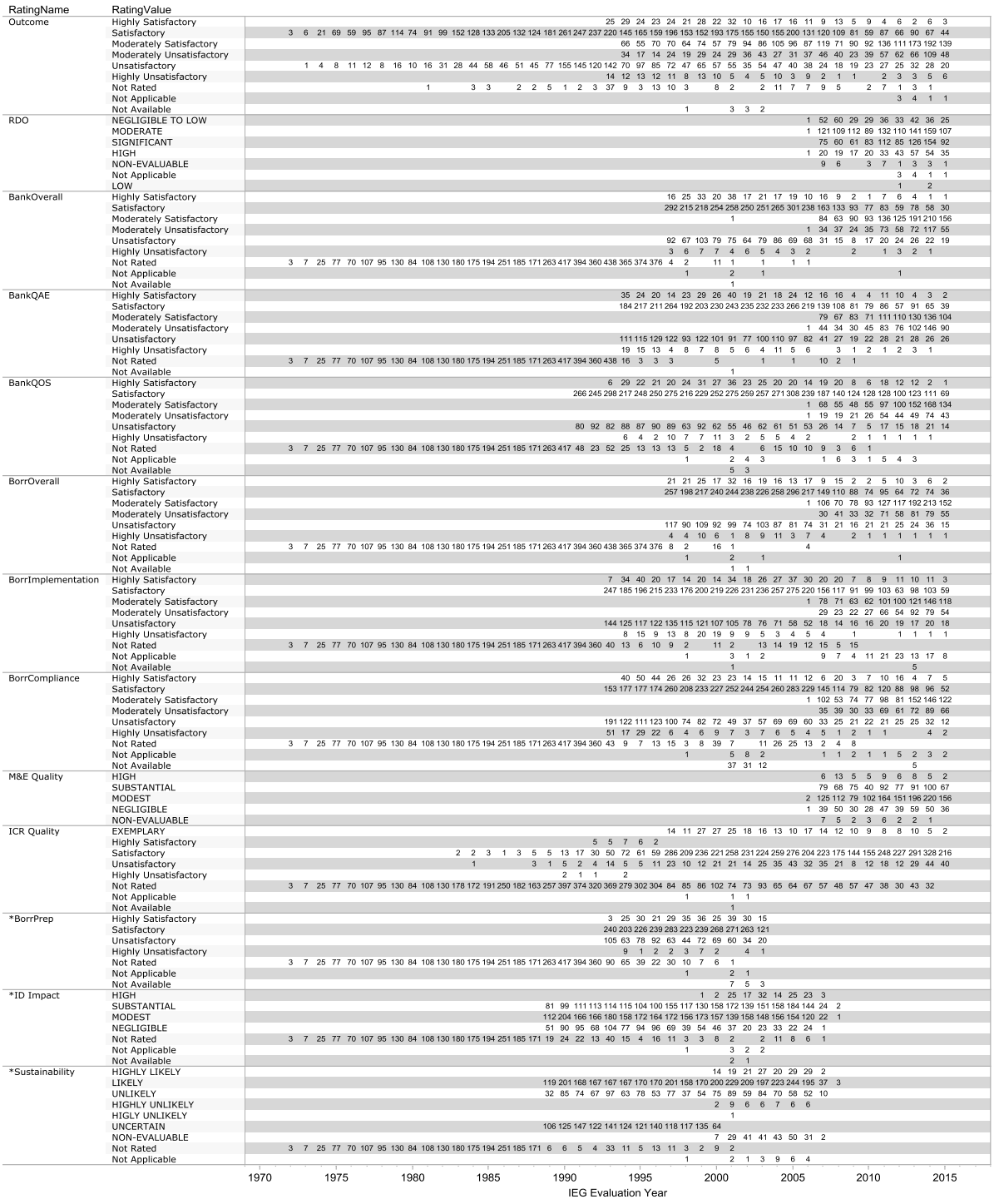
Annex A – IEG World Bank Project Performance Ratings Dataset Fields

The following table lists the fields/columns that will be presented as part of the IEG World Bank Project Performance Ratings dataset.

Column Name	Type
ProjID	Project Description Field
ProjName	Project Description Field
Region	Project Description Field
CtryCode	Project Description Field
CtryName	Project Description Field
ApprovalDate	Project Description Field
ApprovalFY	Project Description Field
SectorBoard	Project Description Field
DeactDate	Project Description Field
AgreementType	Project Description Field
LendingProjectCost	Project Description Field
ExitFY	Project Description Field
LenInstrType	Project Description Field
LendingInstr	Project Description Field
ProdLineCode	Project Description Field
ProdLine	Project Description Field
ERRatAppraisal	Rate of Return
ERRatCompletion	Rate of Return
IEG_EvalDate	Evaluation Description Field
IEG_EvalType	Evaluation Description Field
IEG_Outcome	Project Performance Rating
IEG_RDO	Project Performance Rating
(disc)IEG_IDImpact	Project Performance Rating (discontinued)
IEG_BankQualityAtEntry	Project Performance Rating
IEG_BankQualityOfSupervision	Project Performance Rating
IEG_OverallBankPerf	Project Performance Rating
(disc)IEG_BorrPrep	Project Performance Rating (discontinued)
IEG_BorrImplementation	Project Performance Rating
IEG_BorrCompliance	Project Performance Rating
IEG_OverallBorrPerf	Project Performance Rating
IEG_ICRQuality	Project Performance Rating
(disc)IEG_Sustainability	Project Performance Rating (discontinued)
IEG_MEQuality	Project Performance Rating
IEG_SourceDocumentURL	Evaluation Description Field

Annex B – All Ratings - Distribution by Year of Evaluation

(as of August 5, 2015)



PLEASE NOTE: Some ratings in dataset erroneously have a "Not Rated" value assigned for years prior to the rating's introduction, instead of a null value. Consquently, a rating like "Bank Overall", which was introduced in 1997, has "Not Rated" values preceding that year.

Annex C – Computing Aggregates with IEG Ratings –Conventions

IEG and World Bank management regularly use IEG's World Bank project performance ratings, as follows:

- They are used by IEG's Results and Performance series, as well as by other Country and Sector/Thematic evaluation studies, to track trends in World Bank project performance
- They are part of the World Bank Group's Corporate Scorecard (see link), tier III indicators of Bank Group performance
- Rating trends are closely monitored by quality assurance and portfolio management teams in World Bank regional and thematic practice units
- Rating trends are regularly included in Quarterly Board Reports (QBRs) and similar Bank Management reporting.
- Researchers and academia use the performance rating series published by IEG on the World Bank's Open Data Initiative (see link)

Most of these uses involve aggregating ratings in meaningful groupings (e.g. by region, investment type) and across a time dimension (e.g. fiscal year). The following section provide details on current conventions for preparing aggregate ratings reports.

Trend Analysis with Ratings Aggregates

IEG follows certain conventions when using or reporting on ratings aggregates. Analysts using this dataset to explore project performance trends should use this guidance as a starting point.

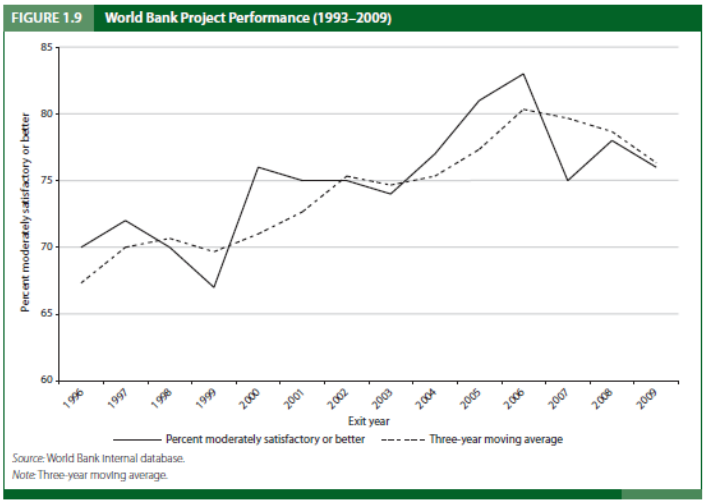


Figure 1. Ratings Aggregate Example: "World Bank project performance (1993-2009)", RAP 2010

1. For any project, the most recent assessment will supersede earlier ones. The dataset contains over 11,200 assessments of about 9,400 World Bank projects[6]. As described earlier IEG carries out a performance validation exercise for all completed projects, so the corresponding ICR Reviews (ICRRs) will contribute their scores to the project performance ratings dataset. Similarly, for the subset of projects that are evaluated through a more in-depth Project Performance Assessment Report (PPAR), we will also have updated performance ratings in the dataset. For aggregate reporting, more recent ratings (PPARs) should supersede earlier ones (based on ICR reviews), and become the project's "current" ratings.

In rare occasions, IEG may perform a post-completion validation (ICRR) after a field-based, more in-depth evaluation (PPAR). In such cases, the ICRR ratings would only supersede those of the PPAR if it is issued at least six (6) months after the PPAR's date.

2. Trend analysis based on aggregates often includes dimensional breakdowns. The dataset already includes dimensional data useful for chart breakdowns, such as Region, Country, Product Line, etc.

3. World Bank and IEG reports often use "Exit FY" as time dimension. The dataset includes three time variables: approval date, deactivation (or portfolio exit) date, and evaluation date. The time dimension can also be expressed as a fiscal year (FY), which helps associate aggregates to specific budgetary cycles.[7] Reporting by fiscal year, or a banding of several fiscal years, also increases the number of observations per aggregate calculation.

Yearly project performance reporting done by IEG and other World Bank units is based on assessing and comparing the cohort of recently completed projects with prior cycles. Trend reports are often based on Exit FYs, so comparisons involve yearly completion cohorts. Projects exiting in a given fiscal year may have been approved as recently as 2 years ago (e.g. development policy lending) to as long as 10 years ago (for an investment project with multiple extensions).

4. The six-point ratings scale for project outcome and the other ratings is often collapsed into a dichotomous variable for the purpose of tracking trends. For example, when calculating aggregates for Outcome rating across a range of fiscal years, the aggregate indicator is the share of all relevant projects (projects rated during a given fiscal year) that were rated "moderately satisfactory" or better. This ratio indicator is therefore based on two simple counts: a) the numerator indicates the total number of projects rated "moderately satisfactory" or better; and b) the denominator indicates the total number of projects rated by IEG. The table below summarizes the approach used by IEG for aggregating all other ratings, several of which have unique scales.

[6] As of August 5, 2015

[7] At the World Bank Group, fiscal years run from July 1st to June 30th. So, for example, "FY11" runs from July 1, 2010, to June 30, 2011.

Rating	Aggregate Indicator (% of projects)	Numerator	Denominator
"Outcome"	"% Outcome Satisfactory"	Number of projects rated "Highly Satisfactory", "Satisfactory", or "Moderately Satisfactory"	Number of projects rated by IEG (e.g. rating is not "Not Rated", "Not Available" or null)
"RDO"	"% RDO Moderate or better"	Number of projects rated "Moderate" or "Negligible to Low"	Number of projects rated by IEG (e.g. rating is not "Non-evaluable" or null)
"Bank Performance", or "Borrower Performance"	"% Bank Satisfactory", or "% Borrower Satisfactory"	Number of projects rated "Highly Satisfactory", "Satisfactory", or "Moderately Satisfactory	Number of projects rated by IEG (e.g. rating is not "Not Rated", "Not Applicable", "Not Available" or null)
"ICR Quality"	"% ICRs Satisfactory"	Number of projects rated "Exemplary" or "Satisfactory"	Number of projects rated by IEG
"M&E Quality"	"% M&E Quality Substantial"	Number of projects rated "High" or "Substantial"	Number of projects rated by IEG (e.g. rating is not "Non-evaluable" or null)