Improving Institutional Performance: Foreign Aid Evaluation and Determinants of Foreign Aid Project Success Ratings

Susan Sawyer O'Keefe

Professor Michelle Connolly, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2023

Acknowledgements

I would like to thank Professor Connolly for always supporting me and providing me with feedback. Thank you to my thesis seminar classmates for being there to bounce ideas off. Lastly, thank you to my family for always supporting me.

Abstract

In this paper, I use a regression model to predict project outcome ratings for international aid projects by 12 multilateral and bilateral aid agencies taking place in 183 recipient countries. The influential factors considered are project duration, project size, evaluation type, evaluation lag, donor ratings, and country-level indicators of development. I find a significant relationship supporting differences in project outcome ratings for projects evaluated by an independent evaluation agency, a resource that some banks use to access project performance by an unbiased party. I also examine the significance of other project-level factors and compare these to trends identified in past literature on foreign aid project effectiveness.

JEL classification: H43, O22
Key words: Development project success, project effectiveness

1. Introduction

Aid effectiveness, the degree of success or failure of international aid, is a complex topic that has been highly contested by academics and those within the foreign aid community. The discussion of aid's effectiveness depends on understanding about the objectives of aid and the methods to achieve success. Micro-level analysis on outcomes of project aid paints a positive picture for foreign aid's "success", where the most common measure of success is whether these projects achieve their immediate objectives (Riddell, 2007). Most donor agencies record success ratings of upwards of 75%, but this conclusion that most projects experience successful outcomes isn't necessarily supported at the macro-level.

This paper will look micro-level project data on summaries of results from aid development projects to examine the methods used to evaluate individual aid success. Overall designs for these interventions are established by donor agencies before projects are approved. The project plan document articulates the "theory of change," which includes a budget and combination of qualitative and quantitative goals that would serve as criteria to indicate project success (Clements, 2020). Project approval indicates that the design is expected to outline a process to reach a project's targets, although project designs vary in specificity and what they leave to the discretion of project managers across projects and donor practices.

The subjectivity of the criteria introduces a regulatory environment where it's difficult to compare project ratings because donor agencies organize their own evaluation systems and usually evaluators determine how to measure and apply each criterion – meaning that no requirements to use consistent units of measurements even for projects with similar impacts (Clements, 2020). The relationship of interest that this paper aims to illuminate is between project ratings across different types of evaluators. To do so, as the evaluation process of projects is contingent upon the preferences and

criteria of individual donors, it's important to understand the extent to which donor specific factors contribute to the variability in ratings which we consider heavily in our analysis.

The regression used in this paper examines the heterogeneity in project ratings across the following 3 categories of evaluator type:

1. Internal evaluation refers to assessments conducted by in-house by staff members of the donor organization responsible for implementing the project, based on the organization's own criteria and methodologies.
2. External evaluations involve the donor contracting an external evaluator to provide an impartial assessment of the project's outcomes. These are typically conducted by experts who possess specialized knowledge in evaluating development assistance and are expected to provide an objective assessment. External evaluators are usually hired on a project-by-project basis to evaluate a specific measure while the donor relationships between donor and independent evaluation offices are usually permanent.
3. Independent evaluation offices entail the use of separate, autonomous agencies that have typically been established for the purpose of evaluating projects. These agencies have long-lasting relationships and operate independently from the donor organization, aiming to be impartial in their assessments. They promote transparency by making their reporting on evaluation findings publicly available. One example of an independent evaluation agency is the World Ban Independent Evaluation Group (IEG), which provides independent evaluation services to the World Bank Group to evaluate their work to ensure oversight and accountability.

Given this variation in practice and difference of interests for different evaluator types, the potential bias that may arise when project evaluations are contracted to external evaluators or assigned

to independent offices could compromise the credibility of the evaluation process. External evaluators may be motivated to providing favorable conclusions to increase the likelihood getting contracted in the future (Clements, 2020). Similarly, projects evaluated internally have the potential to be positively biased by vested interests of donor agency staff. We hypothesize that ratings for projects evaluated by the internal process will be higher than ratings for projects evaluated externally and projects evaluated by an independent evaluation office. We expect that the decrease in rating will be strongest for projects evaluated by the independent office.

2. Literature Review

The work in this paper builds upon literature on evaluations of foreign aid projects, which suggests that at many agencies evaluations may be ambiguous and skew towards being positively biased although this relationship has not been analyzed in detail (Clements, 2020). The research in this paper introduces the dimension of evaluator type as a significant factor that could influence the ratings attributed to project evaluations.

Looking at aid effectiveness at a larger-scale, economists' analysis is centered around macroeconomic variables such as growth, investment, and poverty to measure development achievement. The results of econometric studies on aid's effectiveness at improving measures of development are varied: while some yield affirmative results some find the effectiveness of aid on macroeconomic activity to be nonexistent or even negative. Overall, foreign aid literature delivers "no consensus as to whether higher aid inflows accelerate the achievement of the desired development outcomes" (Pickbourn & Ndikumana, 2016, 397). One potential explanation for the micro-macro paradox that we observe would be widespread positive bias in project-level evaluations, many of them conducted by the donor staff that implemented the project. Throughout this paper we will explore the potential for biases in project ratings between different evaluation types and across donors.

The lack of consensus in both objectives and methods for aid as well as the complexity of determinants of success and failure make measuring effectiveness difficult. In the early decades of development assistance economic cost-benefit analysis of economic growth gains were the most widely used approximation to a standard project evaluation framework (Clements, 2020). This provided a more objective approach for accountability where being unbiased was a question of the validity of these estimation of economic gains as well as the appropriateness of economic gains as the proper measure of development.

Since then, the cost-benefit approach of project evaluation is often still incorporated into evaluation criteria, but cost-benefit approach as the overall standard evaluation framework for development assistance has fallen out of favor in the aid community and has been replaced by the Organization for Economic Cooperation and Development's (OECD's) Development Assistance Committee (DAC) six evaluation criteria: relevance, effectiveness, impact, efficiency, coherence, and sustainability (OECD, 2019). Determining allocation of ratings on the range of DAC criteria (for example from "highly unsatisfactory" to "satisfactory") is "problematic" because it requires a significant element of judgment, creating an opportunity for positive bias. The way that criteria are defined in relation to each other creates also conceptual confusion and project ratings weighs each criteria evenly without analyzing the relationship between variables. For example, the criteria consider "relevance" on par with "impact" rather than considering a causal relationship between the two indicators. Similarly, it's difficult to quantify a measure of a variable like "sustainability" separately than efficiency and impact because sustainability is concerned with measuring the likely longevity of the benefits of a project after a donor.

In this paper we include country-specific measures of GDP per capita and control of corruption to explain some of the rating variation for projects. The notion that economic growth and stability are favorable factors for project success is widely substantiated by empirical evidence showing a positive correlation between GDP per capita and the increased likelihood of project success (Denizer et al., 2013; Briggs, 2020).

On other factors, the results aren't as clear. Some studies, such as one meta-analysis of World Bank Projects, have found a negative correlation between project duration and project outcome ratings, although other studies have found no relationship between the two (Denizer et al., 2013; Wood et. al., 2020). Similarly, studies show conflicting conclusions about the relationship between project size and

outcome, with the World Bank project study finding smaller projects to be more successful while an analysis of Australian aid projects found larger projects to be more successful (Denizer et al., 2013; Wood et. al., 2020). Similarly, at least two studies on standard project traits found that projects with a longer duration received lower ratings on average, although another study of Australian project data failed to support this relationship (Denizer et al., 2013; Feeny & Vuong, 2017; Wood et al., 2020). It's understandable that associations with variables such as project size or duration could be mixed because of the endogeneity problem this type of data presents. For example, while one may expect larger projects to have better outcomes because of their funding and resources, larger projects may also be harder to complete and have stricter evaluation guidelines.

3. Data

The dataset I will be using is called the Project Performance Database (PPD) version 2.1 (Hoing, et al., 2022). The PPD is one of the world's largest databases of development projects which includes the project outcome ratings of holistic project performance produced by donor staff and independent evaluation teams that capture the extent to which projects achieve their objectives and allocate resources efficiently. This dataset contains data on 21,198 unique foreign aid projects (20,686 projects in the database contain project ratings) from 12 bilateral and multilateral aid agencies from 1956 to 2016 taking place in 183 recipient countries. Project data varies in completeness across donors, so many of these limits the sample of observations from the data that we incorporate into our model.

We measure project success using holistic, ex-post evaluation ratings produced by donor agency staff and evaluation experts. Project ratings are normally distributed with a mean of 4.216 and a standard deviation of 1.07.

Mean success rating varies heavily across donor. Because some of the original donor ratings were rescaled to standardize the 6-point scale in different ways depending on the data, it's important that we include donor fixed effects to avoid concern of inflation from differences in the scales across donors. See appendix table 5 for a breakdown of the original donor rating scale. Variation among project ratings could be capturing true differences in project success across donors or differences in rating scales, or a combination of both. Below we have a breakdown of mean success ratings by donor for the projects we use in our analysis.

Table 1: Mean Project Ratings by Donor

Donor Name	Mean Project Rating
African Development Bank	4.20
Asian Development Bank	3.99
UK's Department for International Development (DFID)	4.60
Global Environmental Facility (GEF)	4.33
International Fund for Agricultural Development I(FAD)	4.00
Japanese International Cooperation Agency (JICA)	4.98
World Bank (WB)	4.01
Total	4.22

Evaluator type, one of our dependent variables of interest, is somewhat correlated with donor as you can see below – some of the donors use employ one evaluation type. We were concerned that this may be cause for concern with multicollinearity but then upon checking for collinearity it was fine. For donors that use more than one evaluator type these assignments should be determined ex-ante, which agencies claim is true. This is an important assumption to avoid endogeneity. At first glance just looking at the mean outcome ratings we see that projects evaluated internally have the highest ratings, followed by projects evaluated by external evaluation agencies.

Table 2: Mean Project Ratings by Evaluation Type

Evaluator Type	Mean Project Outcome Rating	Std. Error (of the mean)	Coefficient of variation	Number of observations
External Evaluation Agency	4.64	0.04	.26	1072
Independent Evaluation Agency	4.04	0.01	0.01	6482
Internal Evaluation	4.53	0.02	0.02	2262
Total	4.22	0.01	.25	9816

The donors that use independent evaluation agencies are the Asian Development Bank, GEF, IFAD, and the World Bank. The offices that they use, respectively, are the Asian Development Bank Independent Evaluation Department (IED), the GEF independent evaluation office, the Independent Office of Evaluations of IFAD (IOE), and the which World Bank Independent Evaluation Group. These offices all aim to evaluate agency strategies, projects, and operations independently and systematically, but have varying methods of project evaluation, and success indicators which also vary heavily across projects.

Table 3: Means of Project Ratings by Donor and evaluator type (number of observations in parenthesis)

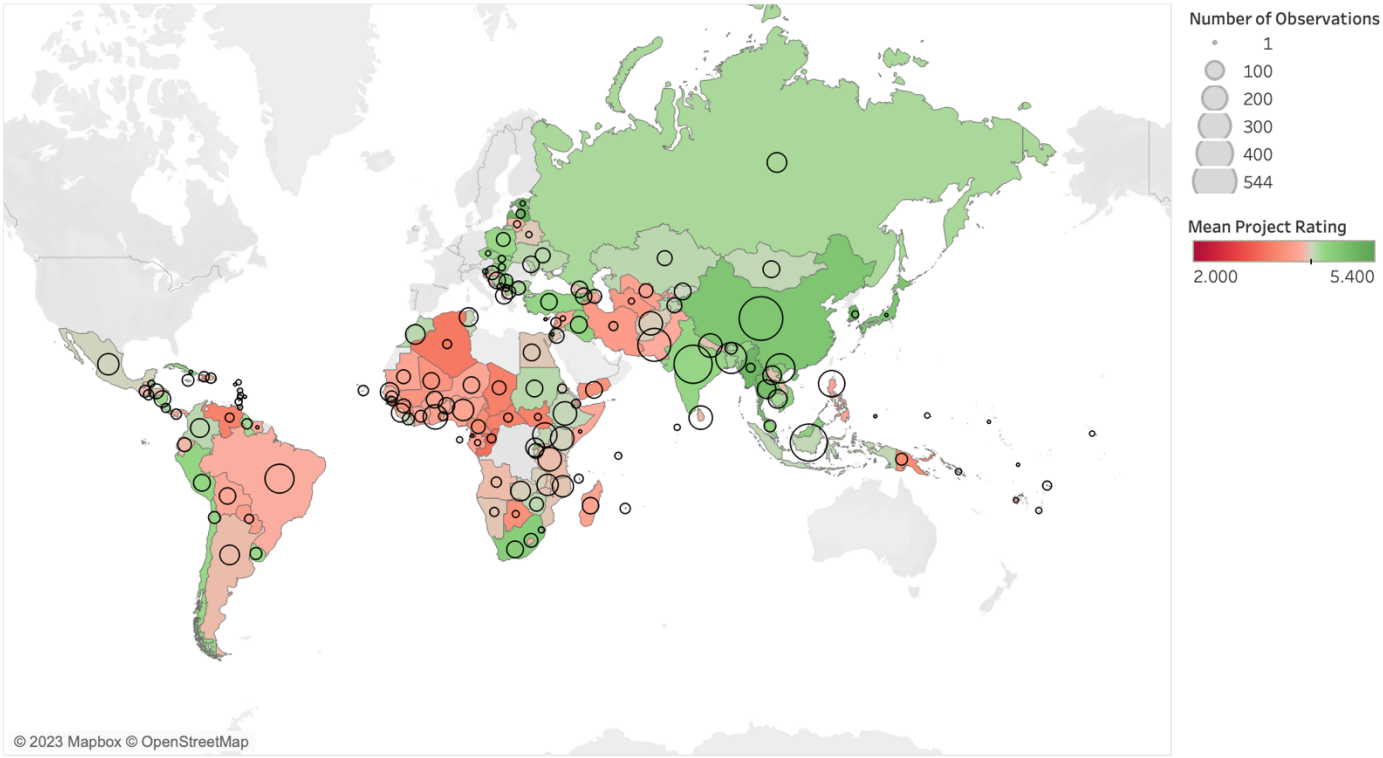
	Evaluator Type			
Donor	External Evaluation	Independent Evaluation	Internal Evaluation	Total
African Development Bank			4.20 (515)	4.20 (515)
Asian Development Bank	4.16 (430)	3.82 (446)		3.99 (876)
DFID			4.6 (1700)	4.60 (1700)
GEF		4.33 (841)		4.33 (841)
IFAD		4.00 (280)		4.00 (280)
JICA	4.96 (642)		5.17 (47)	4.98 (689)
WB		4.01 (4915)		4.01 (4915)
Total	4.64 (1072)	4.04 (6482)	4.52 (2262)	4.22 (9816)

We use the natural log of project size to estimate project size using cost of the project in USD. The original project size variable was given in multiple currencies, so we converted them all to USD using official exchange rate data from the International Monetary Fund (IMF) and standardized the values to account for previous differences in scale. Next, we used the CPI deflator to standardize all project costs to a base year for comparison and then divided the cost by a million to scale. After correcting the project size variable distribution was positively skewed, meaning the mean overestimates the most common values. We used log to normalizes the data along a bell curve that is needed to comply with the assumptions of this model.

We expect project outcomes to vary heavily across countries and include a dummy for country to capture the variation in outcome that can be attributed to country effects. In Figure 1 the color of the country indicates the mean project rating for that country for the projects in our data, and the size of the

circle covering the country indicates how many project observations were recorded in each country. We observe consistency in trends of country rating for projects in the same regions.

Figure 1 – Spatial Map of Mean Project Ratings and Number of Observation by Country



4. Multiple Linear Regression model

Model 1 for our multiple regression can be given by the following equation:

$$Project\ Rating\ (y) = \beta_0 + \beta_1(Donor) + \beta_2 Ln(Project\ Size\ USD) + \beta_3(Evaluator\ Type) + \beta_4(Evaluation\ Lag) + \beta_5(Duration) + \beta_6(Sector) + \beta_7(GDP\ per\ capita_{ic}) + \beta_8(Corruption_{ic}) + \beta_9(EvalYear) + \epsilon$$

Where the independent variable, *Project Rating,* is the overall project outcome rating for the aid project, ranging on a scale of 1-6 with 1 being highly unsatisfactory and 6 being highly satisfactory. The coefficient β_0 represents the constant value for the regression. The independent variable *EvaluatorType* is a dummy that represents whether the ratings were conducted by internally, by an external evaluator, or by an independent evaluation office. The variable *Donor* is a dummy variable for donor. *Country* is a dummy variable for country. $Ln(Project\ Size\ USD)$ is a continuous size estimated using the cost of the project standardized in USD adjusted for inflation. $Evaluation\ Lag$ is determined by the difference between the completion date of the project and the date of evaluation and given in years. *Duration* is calculated ex-post and given in years. *Sector* is a dummy variable that represents the sector. *GDP per capita* and *Corruption* are both calculated using historical year averages from World Bank data. *Corruption* is estimated based on the project country and year using the World Bank indicator for control of corruption, which is higher for countries with better control of corruption and ranges from values -2.5 to 2.5. *EvalYear* is a dummy for the year that the project evaluation took place.

Model 2 is also a multiple regression and functions similarly to Model 1 but includes a dummy variable for *Country* to pick up effects by country and doesn't include GDP per capita or corruption control indicators. Model 2 is given by the following equation:

$$Project\ Rating\ (y)$$

$$= \beta_0 + \beta_1(Donor) + \beta_2 Ln(Project\ Size\ USD) + \beta_3(Evaluator\ Type) + \beta_4(Evaluation\ Lag) + \beta_5(Duration) + \beta_6(Sector\ ) + \beta_7(Country\ ) + \beta_8(EvalYear) + \in$$

5. Findings and analysis

Table 4 displays the multiple regression on project outcome rating.

Table 4 – Regression On Project Rating

==	(Model 1) Project Outcome Rating	(Model 2) Project Outcome Rating
Project size (log, $USD)	.075***	.057***
	(.008)	(.009)
Project duration (years)	-.029***	-.026***
	(.004)	(.004)
Evaluation Lag (years)	-.047***	-.043***
	(.008)	(.008)
GDP per capita	.084**	
	(.038)	
Corruption Index	.18***	
	(.023)	
Donor:		
African Development Bank	-.455**	-.318*
	(.18)	(.181)
Asian Development Bank	-.1*	-.198***
	(.053)	(.055)
DFID	.064	.057
	(.173)	(.175)
GEF	.483***	.42***
	(.058)	(.059)
IFAD	.147**	.161**
	(.068)	(.068)
JICA	.8***	.694***
	(.09)	(.09)
Evaluator Type:		
External Evaluation	-.36**	-.352**
	(.157)	(.157)
Independent Evaluation Office	-.599***	-.588***
	(.171)	(.172)
Country fixed effects	No	Yes
Sector fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
_cons	3.926***	3.798***
	(.742)	(.729)
Observations	9616	9616
R-squared	.13	.179

Standard errors are in parentheses
*** p<.01, ** p<.05, * p<.1*

First, we examine the relationship between evaluator type and project outcome rating. In the base regression model, for projects evaluated by independent evaluation offices, the project outcome rating is

expected to decrease by 0.559, assuming all other variables are held constant. This decrease supports the initial hypothesis that independent evaluation agencies may have harder evaluation criteria when compared to evaluations conducted by internal staff or by external agencies, and this difference is statistically significant ($p<0.01$). This relationship remains statistically significant ($p<0.01$) when we add country fixed effects to the model and remove the GDP and corruption indicators, although the change associated with independent evaluations is very slightly smaller -0.588.

Projects evaluated by an external office, which are different from independent evaluation agencies as we clarified above because they're contracted by the aid agencies rather than being assigned as part of an independent process, are associated with a smaller and less-significant decrease in ratings. In our first regression we find that holding all other factors constant, for projects evaluated externally, their outcome ratings are predicted to be lower than projects evaluated internally by .36 points. ($p<0.05$). When we add in country fixed effects that positive significant relationship remains present.

Many of the coefficients captured by the donor dummy are both large and statistically significant for all banks other than the DFID. This was expected given that although ratings were standardized to a 6-point scale to compare across donors, evaluation criteria as well as project performance are expected to vary greatly across these aid agencies. We use the World Bank as the constant, and in our first model we find that the African Development Bank and the Asian Development Bank are associated with a statistically significant decrease in ratings ($p<$). IFAD, GEF, and JICA are associated with a statistically significant increase in project ratings when compared to the world bank constant. This is telling that the variation on project rating varies this heavily between donors.

When we introduce country fixed effects there is an overall decrease in the variation in outcome explained by the donor variable, but the overall trends remain consistent. Compared to the other independent variables, donor effects is the one that changes the most after adding in country fixed

effects. We expect that this decrease in variation is due to patterns in donor-country pairings that would skew donor ratings depending on how difficult it is to achieve project goals in the countries the donor most frequently partners with. The biggest coefficient change in the model with country effects is seen in the African Development Bank from -.455 to -.318, this is likely because project ratings are lower for African countries, where this bank operates.

This indicates that for the African country effects capture more of the variation in ratings for these countries than the differences variation by the country-indicators of GDP and corruption in the first model. This prediction is supported by the R-squared values of the models above, where we observe that the model with country fixed effects has a higher R-squared, meaning that the amount of variance in outcome ratings can be better explained when looking at country effects than the proxy country indicators of GDP and corruption.

Looking at evaluation lag, we found that projects with longer lags have significantly lower outcome ratings. Specifically, holding all other variables constant for every one-year increase in evaluation lag we expect a -0.004 decrease in project rating. There are two likely possible explanations for this phenomenon: first that that visible project impacts may decrease during the lag between project completion and evaluation. It's also possible that projects that are being evaluated more promptly could be better supervised and have received more resources. We had originally anticipated that projects being evaluated by certain evaluation agencies implemented specific timelines. To avoid potential confounding variable bias between evaluator type and evaluation lag, we checked for correlation between evaluation lag and evaluator type as well as evaluation lag and donor agency and found that the correlation coefficients were low, indicating that the variations in project lag weren't directly related to either variable.

Similarly, for project duration we found that a one-year increase in project duration was associated with a .209 decrease in project outcome rating, while holding all other variables constant. One limitation of looking at project duration is that this variable captures the real duration of the project, so we don't know the original intended length and thus how to separate the effect of a project longer than intended.

The relationship between country GDPs per capita and corruption indicator score at each project's respective time exhibits the expected relationship, that projects in countries with higher GDP per capita and better control of corruption are associated with higher outcome ratings. An increase in GDP per capita of $10,000 is associated with an increase of 0.084 in the predicted outcome rating ($p<0.05$). Similarly, an increase of one point in the World Bank control of effectiveness indicator (values range from -2.5 to 2.5) is associated with a .18 increase in outcome rating. This is consistent with our expectations from literature on country government role in project performance, where corruption is found to impede project performance. It's important to note that these two indicators are positively related (countries with higher control of corruption estimates tend to have higher GDP) with a correlation coefficient of 0.449. This means that including both may cloud how much variation is truly attributable to each indicator but using both allows us to gauge both a level of wealth and corruption that exists within the country which impact the foreign aid process in different ways.

For a 10% increase in project size, the difference in expected mean outcome rating is 0.0031 and found to be statistically significant ($p < 0.01$). This means that bigger projects, size being measured by the cost of the project, tend to receive slightly higher ratings. This makes sense in the context that one could expect more funding to be associated with higher quality and outcomes, although this isn't very telling because the cost of the project alone isn't very telling when not contextualized with the goals and expectations of the project. The project size variable is also impacted by donor resources.

The coefficients associated with evaluation year, sector, and country were mostly insignificant and didn't overall uncover any trends not consistent from what we expected. The estimated coefficients pertaining to country-level effects were found to exhibit a greater magnitude compared to the indicators for GDP per capita and corruption indicating the relative prominence of country-level effects in this context. However, the relationship was not as strong as the relationship between the indicators and in most cases the country-effect was not statistically significant.

Looking at trends in sector effects, the only sector that was associated with a loosely statistically significant ($p < .1$) difference in outcome was the emergency response sector, an area that is historically associated with higher project outcomes in foreign aid. Overall, this indicates that across projects sector does not seem to significantly effect project outcome ratings. Over the years we observe a downward trend in project evaluations, suggesting either that later projects were less successful or that later projects were evaluated more strictly.

For the base model the constant, the Y intercept of my model, is 3.926 which represents the predicted value of project outcome rating when all other variables are held constant.

6. Conclusion

In conclusion, the findings of this thesis support the initial hypothesis that evaluator type, donor agency, evaluation lag, project duration, and country-level indicators such as GDP per capita and corruption have significant impacts on project outcome ratings. The results indicate that projects evaluated by independent evaluation offices tend to have lower outcome ratings compared to projects evaluated by internal staff or external agencies contracted by aid agencies. Donor agencies also play a significant role, with some agencies associated with lower outcome ratings and others associated with higher ratings compared to the World Bank constant. Moreover, longer evaluation lag and longer project duration are both associated with lower outcome ratings, suggesting that timely evaluation and shorter project durations may contribute to better project performance.

Furthermore, the findings highlight the importance of country-level factors in influencing project outcome ratings. Country fixed effects were found to have a higher explanatory power in the models compared to proxy country indicators such as GDP per capita and corruption. Specifically, higher GDP per capita and better control of corruption were associated with higher outcome ratings, indicating the role of country government and local conditions in project performance. Additionally, project size was found to be positively correlated with higher outcome ratings, suggesting that larger projects tend to receive higher ratings, possibly due to increased funding associated with higher quality or more impactful projects.

Overall, this research contributes most uniquely to development assistance literature by exposing the variation in project outcome ratings that can be attributable to evaluator type to help understand that the differences in standards between both donors and evaluation agency type has the potential to skew the interpretation of the success of the project. to the understanding of the complex factors that influence project outcome ratings in the context of international development projects. These findings have

implications for policymakers, practitioners, and evaluators involved in international development projects, providing insights into the factors that can impact project success and offering potential areas for improvement in project evaluation and implementation practices. Based on these findings further investigation could be conducted to understand in detail how the evaluation processes differ between evaluator types and institute future controls to standardize the evaluations across evaluator. These findings indicate that creating a standardized evaluation method or agency that could span across donors would help promote fairness and transparency in project ratings.

Sources

Briggs, R. C. (2019). Results from single-donor analyses of project aid success seem to generalize pretty well across donors. The Review of International Organizations, 15(4), 947–963. https://doi.org/10.1007/s11558-019-09365-x

Clements, Paul, 2020. "Improving learning and accountability in foreign aid," World Development, Elsevier, vol. 125(C). https://ideas.repec.org/a/eee/wdevel/v125y2020ics0305750x19303183.html

Denizer, Cevdet, Daniel Kaufmann, and Aart Kraay. 2013. “Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance.” Journal of Development Economics 105: 288–302.

Doucouliagos, Chris and Paldam, Martin, (2005), Aid Effectiveness on Accumulation. A Meta Study, Economics Working Papers, Department of Economics and Business Economics, Aarhus University, https://EconPapers.repec.org/RePEc:aah:aarhec:2005-12

Feeny, S., & Vuong, V. (2017). Explaining Aid Project and Program Success: Findings from Asian Development Bank Interventions. World Development, 90, 329–343. https://doi.org/10.1016/j.worlddev.2016.10.009

Honig, D., Lall, R. and Parks, B.C. (2022), When Does Transparency Improve Institutional Performance? Evidence from 20,000 Projects in 183 Countries. American Journal of Political Science. https://doi.org/10.1111/ajps.12698

"Kenny, Charles. 2006. What Is Effective Aid? How Would Donors Allocate It?. Policy Research Working Paper; No. 4005. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/9296 License: CC BY 1.0 IGO."

Official Development Assistance (ODA) - OECD. (2019) (n.d.). https://www.oecd.org/dac/financing-sustainable-development/development- finance-standards/official-development-assistance.htm

Pickbourn, L., & Ndikumana, L. (2016). The impact of the sectoral allocation of foreign aid on gender inequality. *Journal of International Development*, *28*(3), 396-411.

Riddell, R. C. (2008). *Does foreign aid really work?*. OUP Oxford.

Wood, T, Otor, S, Dornan, M. (2020), Australian aid projects: What works, where projects work and how Australia compares. *Asia Pac Policy Stud*. 2020; 7: 171– 186. https://doi.org/10.1002/app5.300

Wood, T., Otor, S., & Dornan, M. (2022). Why are aid projects less effective in the Pacific?. *Development policy review : the journal of the Overseas Development Institute*, *40*(3), e12573. https://doi.org/10.1111/dpr.12573

Appendix

Table 5. Original Donor Rating Scale

Donor	Original Rating Scale
African Development Bank	(1-4)
Asian Development Bank	(1-4)
UK's Department for International Development (DFID)	(1-5)
Global Environmental Frontier (GEF)	(1-6)
Global Fund for AIDS, Tuberculosis, and Malaria (GFATM)	(1-4)
International Fund for Agricultural Development (IFAD)	(1-6)
Japanese International Cooperation Agency (JICA)	(1-4)
World Bank	(1-6)