Hauke Feil

The cancer of corruption and World Bank project performance: Is there a connection?

ARTICLE

The cancer of corruption and World Bank project performance: Is there a connection?

Hauke Feil

Faculty of Economics and Social Sciences, University of Potsdam, Germany; Deutsche Bundesbank, DG Statistics, Germany

Correspondence
Hauke Feil, Deutsche Bundesbank, DG Statistics, Germany. Email: hauke.feil@bundesbank.de

Abstract

Motivation: Corruption is often cited as a central reason why development projects fail. The article tests this claim by assessing whether World Bank projects perform worse in implementation environments with a higher corruption level. The article focuses specifically on bribery between public officials and firms during the procurement of needed goods and services.

Approach and Methods: I use data from the World Bank's Enterprise Surveys to avoid the often-criticized corruption perception indices and to allow for an assessment of effects at the subnational level. The analysis builds on an assessment of the performance ratings of 1,228 World Bank projects and covers 87 different countries.

Finding: Overall, the article finds a small but statistically significant correlation between the corruption level and project performance. This result indicates that the corruption level of recipient countries should be considered during the design and implementation of projects.

Policy Implications: Nonetheless, the relatively small correlation and the low pseudo R-squareds advise not overestimating the relevance of corruption for project performance. At least for the project level, the article finds no indication that corruption is a primary obstacle to aid effectiveness.

KEYWORDS

Aid effectiveness, corruption, Enterprise Survey, project performance, World Bank

1 | INTRODUCTION

In 1996, the then president of the World Bank, James D. Wolfensohn, called corruption a cancer which needs to be tackled (Wolfensohn, 2005, p. 5). Yet, after years working at the World Bank, authors like Berkman share a pessimistic view on the prevalence of corruption in development projects in African countries. He has not worked on one project "that did not reek of corruption" (Berkman, 2008, p. 121). Other authors confirm the occurrence of corruption in World Bank projects (Olken, 2007; Winters, 2014). In addition, qualitative studies show that corruption can decrease the performance of development projects (Beekman et al., 2013; Damoah et al., 2018; Olken, 2006). This raises the question

whether corruption has a systematic effect on the performance of World Bank projects, particularly as organizations like Transparency International name corruption as one of the main obstacles to aid effectiveness (Transparency International, 2011). Yet, quantitative studies dealing with this question are surprisingly scarce. In a background paper for the World Bank Independent Evaluation Group (IEG), Kilby (1994) assessed the effect of different macroeconomic and socio-political factors on the performance of World Bank projects. Among other aspects, his models show a negative effect of corruption on project performance. Two studies by the IEG confirm this result (World Bank OED, 1999a, 1999b). In three other studies, however, the effect of corruption indicators remains mostly statistically insignificant (ADB, 2016; Buntaine & Parks, 2013; Limodio, 2011). Thus, the current state of the research is inconclusive. Furthermore, none of the existing studies differentiate between different forms of corruption, even though it is likely that the effect of corruption varies depending on the type of development project and the form of corruption. Investment projects are likely to be vulnerable to corruption in procurement procedures, while projects in the area of budget support may be more vulnerable to the misallocation of resources. Thus, to better understand the effect of corruption on project performance, it is essential to assess different forms of corruption separately (see also Kenny, 2017).

The present article seeks to address the inconclusive state of the research while focusing on one specific form of corruption, namely bribery between private firms and public officials. The level of corruption is measured by the "corr2" indicator of the Enterprise Surveys, which identifies the "percent of firms expected to give gifts to secure a government contract" in a country (World Bank, 2017b, p. 36). Based on a sample of 1228 World Bank projects, the article tests the following research question: *Do World Bank projects perform worse in implementation environments with a higher corruption level?* In so doing, the study provides important theoretical, methodological, and practical contributions. From a theoretical point of view, the article contributes to the small but growing discussion on the determinants of the performance of development projects. This discussion was started by Kaufmann and Wang (1995) and Isham et al. (1995), who assess the effect of macroeconomic policies and civil liberties on the performance of World Bank projects. Later studies include additional determinants at the level of the project, aid agency and recipient country (Bulman at al., 2017; Deininger et al., 1998; Denizer et al., 2013; Dollar & Levin, 2005; Honig, 2018; Vawda et al., 2003). Corruption, however, is one of the few potential determinants which has been largely left out of quantitative studies (see also Langley et al., 2018; World Bank IEG, 2017, pp. 171–184). The present article addresses this research gap. In addition, the results are also relevant for the methodological discussion on the measurement of corruption, particularly because the article compares the results of different corruption indicators. Furthermore, corruption indicators are heavily used by aid agencies. Multilateral Development Banks and the UK's Department for International Development use corruption indicators in their allocation formula to steer more resources to places where they are most effective (DFID, 2011; World Bank, 2001a, p. 1). In addition, the German Federal Ministry for Economic Cooperation and Development argues that good governance, measured by corruption indicators, among others, is "a precondition for effective development work" (BMZ, 2018). Thus, the article also gives insights as to whether the strong reliance of aid agencies on corruption indicators is reasonable.

The analysis reveals a small but statistically significant effect of corruption on project performance. On average, an increase of 10 percentage points in the number of firms that expect to give gifts to secure government contracts decreases the likelihood of a satisfactory project performance by 3 percentage points. Different reliability tests confirm the robustness of this result. The result is also robust when the country-level corruption values are replaced by subnational values that are calculated based on the actual implementation locations of projects. Furthermore, comparable results are obtained if the corruption indicator of the Enterprise Surveys is replaced by the Corruption Perceptions Index (CPI) given by Transparency International or the Control of Corruption index (CC) from the

Worldwide Governance Indicators. Yet, a glance at the values of the three indicators reveals that the comparable results are more a coincidence than a confirmation for a comparable measurement of corruption. The article proceeds by presenting the state of the research, the selection process of the independent variable and the analytical model. This is followed by the analysis and a short conclusion discussing the implications of the results.

2 | STATE OF THE RESEARCH AND THE MEASUREMENT OF CORRUPTION

Corruption is a central topic in multiple disciplines. However, despite "extensive research devoted to the subject, there is limited research on the potential impact of corruption on project failure" (Damoah et al., 2018, p. 17). One of the few contributions to the literature stems from Tanzi and Davoodi (1998). The authors show that countries with higher corruption levels tend to have a lower quality of infrastructure. Galilea and Medda (2010) and Jiménez et al. (2017) also find a negative correlation between the level of corruption and the performance of public–private partnership projects. A few studies also focus specifically on development projects. Based on interviews with project practitioners in Ghana, Damoah et al. (2018) find that corruption negatively affects the performance of development projects. A quantitative study done by Kilby (1994) also shows a negative correlation between the corruption level in the recipient country and the performance of World Bank projects. Two studies by the World Bank Operations Evaluation Department obtain a similar result (World Bank OED, 1999a, 1999b). However, there are also studies with divergent findings. In their paper on World Bank projects, both Buntaine and Parks (2013) and Limodio (2011) include corruption indicators as control variables. In most of their models, the effect of the corruption indicators remains statistically insignificant. An assessment by the Asian Development Bank also yields no statistically significant effect between corruption and project performance (ADB, 2016). Thus, at least when it comes to development projects, the current state of the research is inconclusive. Furthermore, none of the existing studies differentiate between different forms of corruption, even though it is likely that different types of development projects are affected by different forms of corruption. Investment projects, for example, are likely to be vulnerable to corruption in procurement procedures. In contrast, projects in the area of budget support usually do rely on the procurement of goods and services. However, they may be more vulnerable to corruption in the form of misallocation of resources. To better understand the effect of corruption on project performance, it is therefore essential to assess different forms of corruption separately.

The present article addresses the conclusive state of the research by testing whether World Bank projects perform worse in implementation environments with a higher level of corruption. In so doing, it focuses on one specific form of corruption, namely bribery between private firms and public officials. Bribery in this context is defined as the "explicit exchange of money, gifts in kind, or favours for rule breaking or as payment for benefits that should legally be costless or be allocated on terms other than willingness to pay" (Rose-Ackerman & Palifka, 2016, p. 8). The corruption form of bribery is selected, as it is the one discussed most in the literature. The main theoretical argument stems from Tanzi and Davoodi (1998), who argue that corruption decreases the performance of infrastructure projects because the payers of bribes will try to recover their expenses during the contract execution (see also Collier et al., 2016; Compte et al., 2005). This can happen in three different ways. First, the payer can include the bribe in the official offer or can overprice certain services, which raises the costs of projects. Second, in co-operation with public officials, the payer can try to adjust the design of the project during the implementation to increase earning opportunities, which causes more expensive and

more complex projects. Third, the payer can reduce the implementation costs by lowering the quality of work or by using low-quality materials. This leads to inferior project implementation. Furthermore, corrupt public officials may have incentives to approve projects with deficits or to reduce maintenance of the built infrastructure because they could then take a second bribe for the repairs (Tanzi & Davoodi, 1998, pp. 42–47). Thus, corruption can lead to inadequate project implementation, which reduces the performance of projects. Damoah et al. (2018) identify an additional negative mechanism caused by corruption, namely time overrun. They argue that corrupt public officials will do everything they can to acquire bribes. If contractors are unwilling to pay these, they may try to force them. One way to do so is to refuse to approve necessary documents until bribes are paid. This can cause delays or even the abandonment of projects (Damoah et al., 2018, pp. 22–24). Davis (2004) and Kenny (2010) highlight that it is not only the procurement of contracts that is susceptible to corruption. During the execution of a contract, officials may demand bribes for their services, which increases the cost of projects. At the same time, contractors may bribe public officials to look the other way when they lower the quality of work or overprice their services.

Thus, according to the theoretical arguments in the literature, development projects perform worse in more corrupt environments due to negative effects caused by the payment of bribes by firms to public project officials during the implementation of projects. In particular, the theoretical arguments relate to the procurement of goods and services that are needed for the project. This requires a measurement of corruption that focuses on this specific form of corruption. The three standard corruption indicators used in the literature are unable to fulfil this criterion. The CPI and CC are both composite indices that cover multiple different forms of corruption and are based on a broad range of different sources. Thus, while the two indices may give an overview of the general corruption level in a country, they say little on the level of one specific form of corruption (see also Andersson & Heywood, 2009; Rose-Ackerman & Palifka, 2016, pp. 19–21). In other words, a high corruption value on the CPI does not imply that there is a high level of corruption in the procurement process of the respective country. The third corruption indicator often used in the literature, the Corruption Index of the International Country Risk Guide, focuses on the risk of corruption for foreign investment. Yet, even this more focused measurement of corruption includes forms of corruption that are not relevant for the research question (see PRS Group, 2018, pp. 4–5). Furthermore, the country coverage of the indicator is limited when it comes to developing countries.

This makes it necessary to select another corruption measurement for the analysis. The most suitable ones are the corruption indicators of the Enterprise Surveys. In these firm-level surveys conducted by the World Bank and other organizations, local business owners and managers are asked to assess the business environment in which they are active. The surveyed firms are selected according to a stratified random sampling process and each survey covers between 150 and 1800 firms in a country, depending on the size of the economy. Among other aspects, the questionnaire includes a section on corruption. For the analysis, the paper uses the "corr2" indicator, which identifies the "percent of firms expected to give gifts to secure a government contract" (World Bank, 2017b, p. 36). This indicator is selected because it corresponds perfectly with the research question and the main causal mechanism described in the literature.

The indicator also addresses two concerns that authors have raised regarding the validity and reliability of the CPI and CC.[1] The first concern is that both indicators build on assessments of experts and business leaders. These can be highly subjective, given that the corruption level may be assessed

---

[1]Corruption indicators based on surveys are also not without their critics. Kenny, for example, shows that "survey evidence is subject to considerably uncertainty." However, he also highlights that "perception measures appear to perform even worse in measuring the extent or impact of corruption in sector outcomes" (Kenny, 2009, p. 315).

based on different standards and definitions of corruption (see Andersson & Heywood, 2009; Donchev & Ujhelyi, 2014; León et al., 2013; Rose-Ackerman & Palifka, 2016, p. 20). Other authors criticize the fact that the CPI and CC rely heavily on the perception of Western or remote business leaders and experts, which could lead to biased estimates due to "consistent but erroneous" judgements on how countries work (Razafindrakoto & Roubaud, 2010, p. 1067). The selected "corr2" indicator addresses this concern by asking respondents a precisely defined question on one specific form of corruption.[2] This procedure provides respondents with common definitions and standards for their answers and helps to reduce the level of subjectivity in the responses. Furthermore, the respondents are local business owners and managers from the country in question and, since 2008, only firms that actually have attempted to secure a government contract are asked to answer the respective question. Thus, the responses to the surveys stem from relevant local actors who are able to assess how the local procurement system works in practice. The second concern is that country-level indices like the CPI or CC "cannot satisfactory [*sic*] explain the within-country variation of corruption" (Reinikka & Svensson, 2002, p. 135). Multiple studies find strong differences in the level of corruption within countries (see Goel & Nelson, 2011; Langbein & Sanabria, 2013; Nguyen & van Dijk, 2012; Olken, 2009). This aspect is essential, as World Bank projects are often only implemented in certain areas of a country. Thus, country-level indicators may represent an imprecise measurement of the corruption level in a project's implementation environment. However, contrary to the CPI and CC, the Enterprise Surveys provide data for subnational levels. This makes it possible to run a reliability test based on the data for the subnational areas in which a World Bank project is implemented.

3 | RESEARCH DESIGN

Like most studies on the determinants of the performance of World Bank projects, the article uses the official project "outcome rating" of the IEG as a dependent variable. This rating assesses "the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently" (World Bank IEG, 2014, p. 5). Each project is rated on a six-point scale, which goes from "Highly Satisfactory" to "Highly Unsatisfactory". Following Bulman et al. (2017), Feeny and Vuong (2017), and Assefa et al. (2014), all ratings are recoded into a binary format based on the underlying World Bank classification into satisfactory (ratings 1 to 3) and unsatisfactory projects (ratings 4 to 6).[3]

As highlighted above, the independent variable builds on the "corr2" indicator of the Enterprise Surveys. Contrary to other corruption indicators, data for the Enterprise Surveys are not available on an annual basis. Between 2006 and 2017, the Enterprise Surveys covered 20 countries each year on average.[4] In addition, while some countries have been covered multiple times, others were surveyed just once or not all. This reduced coverage makes it impossible to include the entire population of World Bank projects in the analysis. Instead, the analysis is restricted to projects that were implemented while an Enterprise Survey took place in the respective country. Thus, projects are only included in the sample if values from the Enterprise Surveys are available for at least one year of the project's implementation period.[5] Each project receives the corruption value of the Enterprise Survey that was conducted during

---

[2]See World Bank (2017b) for the formulation of the survey question.

[3]All models are recalculated based on an ordered logistic regression using the original six-point scale of the IEG. In so doing, comparable results are achieved (results not reported).

[4]Earlier Enterprise Surveys are excluded due to an incomparable methodology.

[5]The implementation period is defined as the time frame from the effectiveness date to the closing date of a project.

its implementation period. In the few cases where two Enterprise Surveys were conducted during an implementation period, the mean of the values for both surveys is assigned to a project.

The study sample is limited to projects in the area of Investment Project Financing (IPF). Development Policy Financing projects are excluded, as they focus on budget support and do not usually rely on the procurement of products and services or on the implementation of physical and social infrastructure. Thus, the theoretical arguments outlined above may not apply to this financing instrument. For better comparability among the included projects, the sample also excludes projects that became effective before the year 2000. Overall, the sample contains 1,228 projects that were implemented in 87 different countries. This number implies that the sample covers 49% of all the projects in the population.[6] A comparison with IEG project data shows that the project distribution among countries, project sectors and project types hardly differs between the sample and the population. This makes it appropriate to use statistical inference to generalize the results of the sample to the entire population of IPF projects of the World Bank.

All regression models also include different control variables,[7] which are described in detail in Table A4 in the Online Supplementary Material. Like in most studies on project performance, regional dummy variables are included to account for potential unobserved heterogeneity between regions. At the country level, the models include gross domestic product (GDP) per capita and GDP growth per capita as control variables. The first variable captures effects caused by a different level of development, while the second variable controls for potential macroeconomic shocks. In addition, it is also necessary to control for the level of civil liberties, political rights and accountability in the recipient country. Multiple factors like a free press or participatory practices can reduce the level of corruption in a country (see Lederman et al., 2005; Lyrio et al., 2018; Rose-Ackerman & Palifka, 2016, pp. 374–411). Furthermore, these factors can also affect project performance (see Blum, 2014; Dollar & Levin, 2005; Isham et al., 1997). The different models therefore include the voice and accountability indicator from the Worldwide Governance Indicators as a control variable. Furthermore, controls for four project characteristics are included. The first one is the sector of the project, as projects in some sectors may be more prone to corruption than in others. The second one is the logarithm of the financial volume of the project. The main reason for including this variable is that parts of the World Bank's financial resources are allocated according to characteristics of recipient countries. Post-conflict countries, for example, receive more resources and, thus, potentially larger projects (see World Bank, 2017a). At the same time, larger projects also may attract more corruption. In addition, all models control for the lending instrument, given that effects may differ between different types of lending instruments (see also World Bank, 2001b). The fourth project-level control variable is based on Noltze et al. (2018) and controls for the evaluation delay, which is the time between the closing date and the evaluation date of a project. The argument behind this control variable is that the later an evaluation is conducted, the more shortcomings in the implementation may become visible. The last included control variable is a dummy variable for the evaluation year of the project. The variable controls for potential changes over the period under analysis. This includes potential changes in the methods and standards of project evaluations.[8]

---

[6]The population is defined as all World Bank IPF projects that were closed and evaluated after 2005. The threshold of 2005 was selected as there are no comparable Enterprise Surveys for years prior to 2006.

[7]All country-level control variables are included by calculating project-life averages based on the effectiveness date and the closing date of projects. Thus, for each project, country-level controls indicate the mean of the values of the respective variable for the years in which the project was implemented.

[8]Given that data for the independent variable are only available for a limited number of years, the study is unable to use five-year period fixed effects as is the case in other studies with a longer time frame (Denizer et al., 2013). However, the results also remain robust if one additionally controls for the approval, effectiveness and closing year of a project.

4 | ANALYSIS

The analysis is comprised of three parts. The first part presents the main model and its result, while the second part contains six reliability tests regarding the risk of endogeneity. The last part assesses the sensitivity of the results with respect to the selected measurement of corruption. The results of the different parts are displayed in Tables 1 to 3. The respective tables hide all control variables to provide a better overview of the coefficients of the independent variables. The entire regression results are displayed in Tables A2 and A3.

Table 1 highlights the results for the main model based on different model specifications. The first four models constitute bivariate regressions between project performance and the corruption indicator, while the latter four include all control variables discussed above. All eight models highlight a statistically and economically significant correlation between project performance and the corruption level in a project's implementation environment. According to Models 1 and 2, a one-percentage point increase in the number of firms expected to give gifts to secure a government contract decreases the likelihood of a satisfactory project performance by around 0.3 percentage points. In a similar way, a 10-point decrease in the independent variable increases the likelihood of a satisfactory project performance by 3.5 percentage points. When using odds ratios instead of average marginal effects (AME), a comparable picture is obtained, regardless of whether the dependent variable is measured on a binary or six-point scale (Models 3 and 4). A one-point increase of the independent variable decreases the odds of a successful project (Model 3) or of being in a higher category of the dependent variables (Model 4) by about 1.5%. Models 5 to 8 demonstrate that the coefficient, standard error and P-value of the main independent variable only change slightly if control variables are included in the model. A look at the penultimate column of Table 1 also reveals that the corruption indicator alone explains around 1% to 2% of the variation in the dependent variable. Including the control variables in the model increases this value to around 5% to 9%, depending on the model specification. Thus, even though the effect of corruption on project performance is economically meaningful, the corruption variable alone only explains a small amount of variation in the dependent variable.

TABLE 1 Regression results for the main model

Model	Variables	Coefficient	SE	P-value	n	R2	Model specification
(1)	Corruption only	−0.0031	0.0008	0.000	1228	0.016	Logit, AME
(2)	Corruption only	−0.0032	0.0008	0.000	1228	0.019	Linear probability model
(3)	Corruption only	0.9841	0.0041	0.000	1228	0.016	Logit, odds ratio
(4)	Corruption only	0.9851	0.0036	0.000	1228	0.009	Ordered logit, odds ratio
(5)	Corruption + controls	−0.0035	0.0008	0.000	1228	0.081	Logit, AME
(6)	Corruption + controls	−0.0036	0.0008	0.000	1228	0.090	Linear probability model
(7)	Corruption + controls	0.9808	0.0043	0.000	1228	0.081	Logit, odds ratio
(8)	Corruption + controls	0.9825	0.0040	0.000	1228	0.050	Ordered logit, odds ratio

Notes: SE = standard error, R2 = (pseudo) R-squared, AME = average marginal effect. Standard errors are clustered at the country level. For easier interpretation of Models 4 and 8, the direction of the six-point IEG performance rating is reversed (one = highly unsatisfactory, six = highly satisfactory).

TABLE 2 Regression results for the endogeneity tests

Model	Indicator	Coefficient	SE	P-value	n	R2	Model specification
(9)	Corruption if				701	0.118	Logit, AME, interaction term
	less targeted project	−0.0033	0.0017	0.055			
	more targeted project	−0.0043	0.0011	0.000			
(10)	Corruption if				1228	0.079	Logit, AME, interaction term
	low project volume	−0.0036	0.0010	0.001			
	high project volume	−0.0031	0.0011	0.003			
(11)	Corruption if				1226	0.091	Logit, AME, interaction term
	low financial share of major sector	−0.0028	0.0010	0.007			
	high financial share of major sector	−0.0039	0.0009	0.000			
(12)	Corruption if				1228	0.082	Logit, AME, interaction term
	low project length	−0.0041	0.001	0.000			
	high project length	−0.0029	0.001	0.006			
(13)	Corruption if				1202	0.083	Logit, AME, interaction term
	funding source is IBRD	−0.0059	0.0016	0.000			
	funding source is IDA	−0.0027	0.0007	0.000			
	funding source is external	0.0035	0.0016	0.029			
(14)	Corruption	−0.0038	0.0008	0.000	1107	0.080	Logit, AME

Notes: SE = standard error, R2 = pseudo R-squared, AME = average marginal effect. Standard errors are clustered at the country level. The difference between the coefficients for the different categories of an interaction term is not statistically significant in any of the models with an interaction term. Comparable regression results are obtained if one assesses split samples (based on the median) of the respective variable instead of using an interaction term.

TABLE 3 Reliability tests with regard to the independent variable

Model	Indicator	Coefficient	SE	P-value	N	R2	Model specification
(15)	Corruption (2008–2017)	−0.0032	0.0010	0.001	918	0.090	Logit, AME
(16)	Corruption (subnational values)	−0.0029	0.0014	0.044	422	0.123	Logit, AME
(17)	Corruption (value of contract)	−0.0227	0.0041	0.000	1112	0.081	Logit, AME
(18)	Corruption (value of contract, subnational values)	−0.0215	0.0085	0.012	402	0.128	Logit, AME
(19)	Corruption	−0.0667	0.0148	0.000	1225	0.079	Logit, AME, x-standardization
(20)	Corruption (CPI)	−0.0664	0.0179	0.000	1225	0.077	Logit, AME, x-standardization
(21)	Corruption (CC)	−0.0611	0.0157	0.000	1225	0.076	Logit, AME, x-standardization

Notes: AME = average marginal effect, SE = standard error, R2 = pseudo R-squared. Standard errors are clustered at the country level.

The results are also robust if one excludes control variables that are highly correlated with other included variables. The correlation matrix in Table A1 reveals that this relates in particular to the voice and accountability indicator. Yet, similar results are obtained if the variable is excluded from the regression model. Furthermore, there are arguments that the level of government effectiveness could also affect both project performance and the level of corruption (see Buntaine & Parks, 2013; Hanson & Sigman, 2019; Ika & Donnelly, 2017; Ware et al., 2007). However, the results also remain robust if the voice and accountability variable is replaced by the government effectiveness indicator from the Worldwide Governance Indicators, or if both the voice and accountability and the government effectiveness variables are included in the regression model (results not reported). Thus, overall, the results strongly indicate that World Bank projects perform worse in implementation environments with higher levels of corruption.

When working with project-level data, a potential concern is the presence of endogeneity due to omitted variable bias (see Deininger et al., 1998; Denizer et al., 2013). In the present analysis, two aspects in particular could lead to biased and inconsistent regression estimates. Firstly, certain omitted variables may affect project performance and the corruption level in the recipient country equally. Secondly, certain omitted variables may interact with the effect of corruption on project performance. For example, a high corruption level may only affect the performance of projects with specific characteristics, but not the performance of others. Given the unavailability of reliable instruments to control for endogeneity (see also Denizer et al., 2013, p. 296), the following section contains a series of tests to reduce the risk of omitted variable bias. These tests focus on the variables seen as most relevant in the respective literature and include these variables either as a control variable or as an interaction term. The latter helps to assess whether the obtained results are stable among the different values of the potential confounding variable.

The first variable to be tested stems from Winters (2014). He argues that projects which are more targeted towards a few specific recipients allow for greater accountability, which in turn lowers the opportunity for corruption. A comparable argument can be made for the complexity of projects in

general. Less complex projects are easier to supervise and therefore may offer fewer opportunities for corruption. At the same time, there are indications that more complex projects have a lower performance on average (see Bulman et al., 2017; Denizer et al., 2013; Isham & Kaufmann, 1999). Models 9 to 11 test for this potential interaction. In so doing, Model 9 differentiates between more and less targeted projects based on the number of a project's implementation locations. Models 10 and 11 differentiate between less and more complex projects based on complexity measurements introduced by Denizer et al. (2013), namely the logarithm of the financial volume of projects and the share of financial resources that go to the largest sector of a project. The underlying argument is that projects with a higher financial volume or with a higher share of resources going to multiple sectors are more complex to implement. In the case of all three measurements of complexity, the sample is categorized into more or less complex projects based on the median of the respective variable.[9] For the actual test of endogeneity, the created binary variable interacts with the corruption measurement. The results of Models 9–11 in Table 2 indicate that corruption is correlated with project performance regardless of whether a project is more or less complex. Furthermore, Table A2 highlights that the interaction term between the two variables is insignificant. Together, these results provide strong evidence that the complexity of projects does not act as a confounding variable between corruption and project performance.[10]

Model 12 takes a look at the actual project length as one potential confounding variable. The theoretical argument for testing this argument is the following. In corrupt environments, it may take additional time to safeguard procurement procedures against corruption. This additional time could lead to delays in the project implementation and therefore to a lower project performance. In such a case, the level of corruption would have an influence on project performance, but not through the causal mechanism described above. Model 12 therefore tests whether the correlation between corruption and project performance differs between shorter and longer projects. The obtained results show no indication that the effect of corruption interacts with a project's length.

An additional potential confounding variable is the source of the financial funding of a project. The projects in the sample are funded by three different sources. The first one is the International Bank for Reconstruction and Development (IBRD), which provides countries with loans based on an interest rate that is usually lower than on the private market. The second source is the International Development Association (IDA), the World Bank's concessional lending arm. The IDA offers a mix of grants and (nearly) interest-free credits to eligible countries. External programmes like the Global Environment Facility are the third source of funds. Such programmes always offer grants to eligible countries. Based on the different conditions for the funding, one can argue that recipient countries may care more about the performance of projects for which they have to repay the funds with domestic taxpayer money than about the performance of projects that are funded by external foreign sources. This also implies that recipient countries may be more inclined to fight corruption for IBRD projects than for projects paid for by external grants. Model 13 tests for this aspect based on an interaction term between the corruption level and the source of a project's funding. Projects with blended financing (mix of IBRD and IDA funds) are excluded from the regression. The results of the model show no empirical support for the argument just described.

---

[9]The World Bank does not provide implementation locations for all of its projects. The sample of Model 9 is therefore limited to 700 projects. Binary classifications are used in the interaction terms to enhance the illustration of the results. Furthermore, multiplicative interaction models are often accompanied by serve pitfalls, which could result in unreliable estimates (see Brambor et al., 2006; Pattillo et al., 2007).

[10]The results of the main model are also robust if the three variables are included in the form of a linear control variable instead of a binary interaction term (results not shown).

One final potential source of endogeneity is the content of a project itself. Some projects have the explicit goal of reducing corruption and improving the public procurement system in the recipient country. Thus, a well performing anti-corruption project could lead to a lower corruption level in the recipient. In a similar way, anti-corruption projects could have a lower average performance in countries with high corruption due to overoptimistic goals when it comes to fighting corruption. In such cases, simultaneity or reverse causality between the dependent and independent variables would lead to biased estimates of the effect of corruption on project performance.[11] To encounter this risk, Model 14 excludes all projects that include anti-corruption measures as an objective. The respective projects are identified based on their theme codes. Usually, each World Bank project receives up to five theme codes to describe its goals and objectives. The model excludes all projects that received "Other accountability/anti-corruption" or "Public expenditure, financial management and procurement" as one of their five theme codes. Overall, this excludes 115 projects. The results obtained for Model 14 remain comparable to those of the main model.

The third part of the analysis assesses the reliability of the independent variable. The first test is related to a change in the methodology of the Enterprise Surveys. Since 2008, the question on which the independent variable relies is put only to firms that "have secured or attempted to secure a government contract in the last 12 months" (World Bank, 2017b, p. 28). In 2006 and 2007, in contrast, all selected firms were asked to answer the question. This could lead to incorrect estimates for these two years, as firms that never attempted to secure a government contract may have inaccurate impressions as to whether it is necessary to make informal payments or give gifts to secure a government contract. Model 15 therefore reruns the main model but excludes cases that rely solely on values from Enterprise Surveys conducted in 2006 and 2007. The result of Model 15 indicates that excluding these cases hardly changes the estimated effect of corruption. The second reliability test addresses the criticism that country-level corruption indicators neglect potential within-country variation in corruption. This is especially relevant, as World Bank projects are often only implemented in certain areas of a country. However, contrary to the CPI and CC, the Enterprise Surveys provide data for subnational levels. This allows a test of whether the results are robust if the country-level corruption values are replaced by values for subnational levels. The subnational corruption values are assigned to projects based on three steps. In the first step, the geographic co-ordinates of the implementation locations of a project are used to assess in which subnational divisions of the Enterprise Survey a project was implemented. In the second step, each project receives the corruption values for all the subnational divisions in which it was implemented. In the last step, the mean of these received values is calculated and used as a measurement for the corruption level in a project's implementation environment. An example might be the Enterprise Survey for country X uses six subnational divisions. Project Y is implemented in two of these subnational divisions. The corruption value assigned to project Y is therefore the mean of the corruption values for the two respective subnational divisions. However, due to insufficient responses to the Enterprise Surveys in some of the subnational divisions (see Table A4 for further details) and a lack of geospatial data for several projects, the sample is reduced to 422 projects. The result of the respective model (Model 16) is still comparable to that of the main model. An additional robustness test exchanges the corruption indicator used by another procurement related corruption indicator of the Enterprise Surveys, namely an indicator of the "value of gift expected to secure government contract (% of contract)" (World Bank, 2017b, p. 28). Thus, contrary to the indicator used in the main model, this indicator measures the magnitude instead of the occurrence of corruption during the procurement of goods or services.

---

[11]I would like to thank an anonymous reviewer for this valuable remark.

Models 17 and 18 illustrate that when using this alternative corruption indicator, one also finds a statistically and economically significant correlation between corruption and project performance for both the country and subnational levels.

The last models assess whether comparable results are obtained if the "corr2" indicator of the Enterprise Surveys is replaced by the CPI or CC. Given that each indicator measures the corruption level on a different scale, the comparison requires an x-standardization of the three corruption indicators. Furthermore, the comparison is limited to years and countries for which data for all three indicators are available. The results of Models 19–21 show that the x-standardized coefficients of the three indicators are close to one another. On average, an increase of one standard deviation in the corruption indicators decreases the likelihood of a satisfactory project performance by 5.8 to 7.5 percentage points. Comparable results are achieved if one looks at fully standardized coefficients (see Long & Freese, 2014; Mood, 2010; results not shown). Thus, the results indicate that all three corruption indicators are correlated in a similar strength with the dependent variable. Based on this result, one may assume that the three corruption indicators measure a comparable construct. However, a glance at Figure 1 reveals that this is not the case. For some countries, the indicators of the Enterprise Surveys and CPI show comparable corruption levels. This includes countries like Chile or Uruguay at the lower end of the scale and the Democratic Republic of the Congo and Kenya at the higher end of the scale. For other countries, in contrast, the values of the CPI are considerably higher than those of the Enterprise Surveys. These countries include Cambodia, Zimbabwe, Nicaragua, Ethiopia, Honduras, and Bosnia and Herzegovina. An assessment of the correlation between the three indicators shows a similar picture. While the CPI and CC are highly correlated at a level of 0.92, the correlations of the two variables with the values of the Enterprise Surveys only reach 0.45 and 0.52 (see Table A1). The most likely explanation for the divergent values for some of the countries is that both the CPI and CC also include other forms of
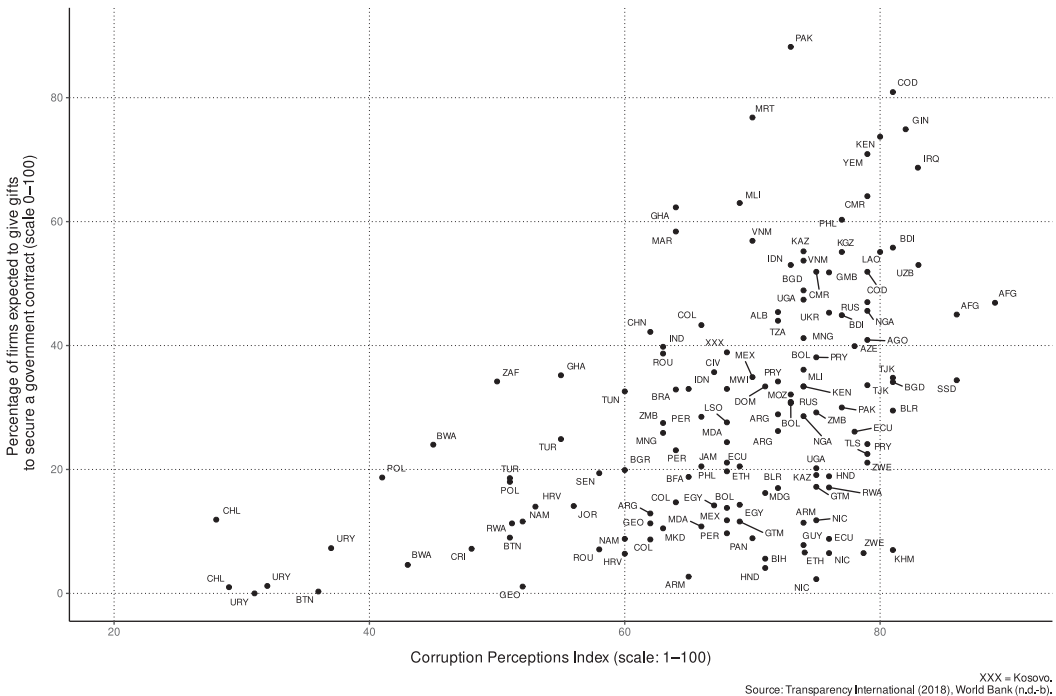

FIGURE 1 Comparison of the corruption values of the Enterprise Surveys and the corruption Perceptions Index
Source: Transparency International (2018), World Bank (n.d.-b).

corruption that are not relevant for the research question. Thus, Figure 1 also highlights the importance of selecting an adequate indicator for the assessment of a research question. This is especially true, as it is merely a coincidence that the analysis obtains comparable results for the three corruption indicators.

5 | CONCLUSION

The present article asks whether the performance of World Bank projects is affected by the level of corruption in their implementation environments. In so doing, the article focuses on one specific form of corruption, namely bribery between firms and public officials. The analysis finds a small but statistically significant effect of corruption on the performance of World Bank investment projects. On average, an increase of 10 percentage points in the number of firms that expect to give gifts to secure government contracts decreases the likelihood of a satisfactory project performance by three percentage points. This result is robust among different reliability tests. This result contributes to the ongoing discussion on the determinants of the performance of development projects (see Assefa et al., 2014; Bulman et al., 2017; Denizer et al., 2013; Feeny & Vuong, 2017), in particular as it sheds light on a country-level determinant which has been largely left out of existing studies. In so doing, Figure 1 highlights the importance of using indicators that correspond to the research question, especially for complex topics like corruption. The systematic effect of corruption on project performance also confirms the need to consider the risk of corruption during the design and implementation of projects. However, the rather small coefficients and the low r-squares also emphasize that one should not overestimate the effect corruption has on project performance. The corruption level in the implementation environment matters for project performance but is hardly one of the most important determinants when it comes to explaining or predicting the performance of World Bank projects. This raises the question of whether it is reasonable to include corruption indicators in aid allocation formulas. In particular, as aid allocation according to corruption indicators risks creating a corruption trap, since this shifts resources away from corrupt countries in which reforms are often impossible to achieve without aid (see Andersson & Heywood, 2009, pp. 747–748). Thus, other forms of aid allocation, like results-based aid, might be a more effective way to deal with the negative effect of corruption on the performance of development projects (see Kenny, 2017). Finally, it is also important to keep in mind that this article only focuses on investment projects and bribery between firms and public officials. Further studies are needed for other types of aid and other forms of corruption, such as budget support and the misallocation of resources. So far, large-n studies on such topics are noticeable by their absence.

REFERENCES
Asian Development Bank. (2016). *Analyzing the determinants of project success: A probit regression approach*. https://www.adb.org/sites/default/files/linked-documents/4-Analyzing-the-Determinants-of-Project-Success-A-Probit-Regression-Approach.pdf

Andersson, S., & Heywood, P. M. (2009). The politics of perception: Use and abuse of transparency international's approach to measuring corruption. *Political Studies*, *57*(4), 746–767. https://doi.org/10.1111/j.1467-9248.2008.00758.x.

Assefa, Y., Rivera, O. P., & Vencatachellum, D. (2014). Macro and micro determinants of project performance. *African Evaluation Journal*, *2*(1), 1–9. https://doi.org/10.4102/aej.v2i1.86.

Beekman, G., Bulte, E. H., & Nillesen, E. E. M. (2013). Corruption and economic activity: Micro level evidence from rural Liberia. *European Journal of Political Economy*, *30*(2013), 70–79. https://doi.org/10.1016/j.ejpoleco.2013.01.005.

Berkman, S. (2008). *The World Bank and the gods of lending*. Kumarian Press.

Blum, J. R. (2014). What factors predict how public sector projects perform? *A review of the World Bank's public sector management portfolio*. https://openknowledge.worldbank.org/handle/10986/17299

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. (2018). *Good governance – an enabling environment for development* [Background statement]. https://www.bmz.de/en/issues/guteregierung/hintergrund/index.html

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, *14*(1), 63–82. https://doi.org/10.1093/pan/mpi014.

Bulman, D., Kolkma, W., & Kraay, A. (2017). Good countries or good projects? Comparing macro and micro correlates of World Bank and Asian Development Bank project performance. *The Review of International Organizations*, *12*(3), 335–363. https://doi.org/10.1007/s11558-016-9256-x.

Buntaine, M. T., & Parks, B. C. (2013). When do environmentally focused assistance projects achieve their objectives? Evidence from World Bank post-project evaluations. *Global Environmental Politics*, *13*(2), 65–88. https://doi.org/10.1162/GLEP_a_00167.

Collier, P., Kirchberger, M., & Söderbom, M. (2016). The cost of road infrastructure in low- and middle-income countries. *The World Bank Economic Review*, *30*(3), 522–548. https://doi.org/10.1093/wber/lhv037.

Compte, O., Lambert-Mogiliansky, A., & Verdier, T. (2005). Corruption and competition in procurement auctions. *The RAND Journal of Economics*, *36*(1), 1–15. https://www.jstor.org/stable/1593751

Damoah, I. S., Akwei, C. A., Amoako, I. O., & Botchie, D. (2018). Corruption as a source of government project failure in developing countries: Evidence from Ghana. *Project Management Journal*, *49*(3), 17–33. https://doi.org/10.1177/8756972818770587.

Davis, J. (2004). Corruption in Public Service Delivery: Experience from South Asia's Water and Sanitation Sector. *World Development*, *32*(1), 53–71. https://doi.org/10.1016/j.worlddev.2003.07.003.

Deininger, K., Squire, L., & Basu, S. (1998). Does economic analysis improve the quality of foreign assistance? *The World Bank Economic Review*, *12*(3), 385–418. https://doi.org/10.1093/wber/12.3.385.

Denizer, C., Kaufmann, D., & Kraay, A. (2013). Good countries or good projects? Macro and micro correlates of World Bank project performance. *Journal of Development Economics*, *105*(2013), 288–302. https://doi.org/10.1016/j.jdeveco.2013.06.003.

DFID. (2011). *Bilateral aid review: Technical report*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214110/FINAL_BAR_20TECHNICAL_20REPORT.pdf

Dollar, D., & Levin, V. (2005). *Sowing and reaping: Institutional quality and project outcomes in developing countries*. https://openknowledge.worldbank.org/handle/10986/8885

Donchev, D., & Ujhelyi, G. (2014). What do corruption indices measure? *Economics & Politics*, *26*(2), 309–331. https://doi.org/10.1111/ecpo.12037.

Feeny, S., & Vuong, V. (2017). Explaining aid project and program success: Findings from Asian Development Bank interventions. *World Development*, *90*(2017), 329–343. https://doi.org/10.1016/j.worlddev.2016.10.009.

Galilea, P., & Medda, F. (2010). Does the political and economic context influence the success of a transport project? An analysis of transport public-private partnerships. *Research in Transportation Economics*, *30*(1), 102–109. https://doi.org/10.1016/j.retrec.2010.10.011.

Goel, R. K., & Nelson, M. A. (2011). Measures of corruption and determinants of US corruption. *Economics of Governance*, *12*(2), 155–176. https://doi.org/10.1007/s10101-010-0091-x.

Hanson, J. K., & Sigman, R. (2019). *State capacity and World Bank project success*. http://www-personal.umich.edu/~jkhanson/resources/hanson_sigman16.pdf

Honig, D. (2018). *Navigation by judgment: Why and when top-down management of foreign aid doesn't work*. Oxford University Press.

Ika, L. A., & Donnelly, J. (2017). Success conditions for international development capacity building projects. *International Journal of Project Management*, *35*(1), 44–63. https://doi.org/10.1016/j.ijproman.2016.10.005.

Isham, J., & Kaufmann, D. (1999). The forgotten rationale for policy reform: The productivity of investment projects. *The Quarterly Journal of Economics*, *114*(1), 149–184. https://doi.org/10.1162/003355399555972.

Isham, J., Kaufmann, D., & Pritchett, L. H. (1997). Civil liberties, democracy, and the performance of government projects. *The World Bank Economic Review*, *11*(2), 219–242. https://doi.org/10.1093/wber/11.2.219.

Isham, J., Narayan, D., & Pritchett, L. (1995). Does participation improve performance? Establishing causality with subjective data. *The World Bank Economic Review*, *9*(2), 175–200. https://doi.org/10.1093/wber/9.2.175.

Jiménez, A., Russo, M., Kraak, J. M., & Jiang, G. F. (2017). Corruption and private participation projects in Central and Eastern Europe. *Management International Review*, *57*(5), 775–792. https://doi.org/10.1007/s11575-017-0312-4.

Kaufmann, D., & Kraay, A. (2018). Worldwide Governance Indicators. http://info.worldbank.org/governance/wgi/#home

Kaufmann, D., & Wang, Y. (1995). Macroeconomic policies and project performance in the social sectors: A model of human capital production and evidence from LDCs. *World Development*, *23*(5), 751–765. https://doi.org/10.1016/0305-750X(95)00003-U.

Kenny, C. (2009). Measuring corruption in infrastructure: Evidence from transition and developing countries. *The Journal of Development Studies*, *45*(3), 314–332. https://doi.org/10.1080/00220380802265066.

Kenny, C. (2010). *Publishing construction contracts and outcome details* (World Bank Policy Research Working Paper No. 5247). http://documents.worldbank.org/curated/en/456291468147859471/pdf/WPS5247.pdf

Kenny, C. (2017). *Results not receipts: Counting the right things in aid and corruption*. Center for Global Development.

Kilby, C. (1994). *Risk management: An econometric investigation of project-level factors* (A background paper for The Annual Review of Evaluation Results 1994 of the World Bank, Operations Evaluation Department). World Bank.

Langbein, L., & Sanabria, P. (2013). The shape of corruption: Colombia as a case study. *The Journal of Development Studies*, *49*(11), 1500–1513. https://doi.org/10.1080/00220388.2013.800858.

Langley, W., Peng, X., & Behrens, J. (2018, July 24). Uncovering factors of project success: A literature review. *IEG*. http://ieg.worldbankgroup.org/blog/uncovering-factors-project-success-literature-review

Lederman, D., Loayza, N. V., & Soares, R. R. (2005). Accountability and corruption: Political institutions matter. *Economics & Politics*, *17*(1), 1–35. https://doi.org/10.1111/j.1468-0343.2005.00145.x.

León, C. J., Araña, J. E., & de León, J. (2013). Correcting for scale perception bias in measuring corruption: An application to Chile and Spain. *Social Indicators Research*, *114*(3), 977–995. https://doi.org/10.1007/s11205-012-0185-7.

Limodio, N. (2011). *The success of infrastructure projects in low-income countries and the role of selectivity* (World Bank Policy Research Working Paper No. WPS 5694). http://documents.worldbank.org/curated/pt/857331468176676870/pdf/WPS5694.pdf

Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using Stata* (Third ed.). Stata Press.

Lyrio, M. V. L., Lunkes, R. J., & Taliani, E. T. C. (2018). Thirty years of studies on transparency, accountability, and corruption in the public sector: The state of the art and opportunities for future research. *Public Integrity*, *20*(5), 512–533. https://doi.org/10.1080/10999922.2017.1416537.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, *26*(1), 67–82. https://doi.org/10.1093/esr/jcp006.

Nguyen, T. T., & van Dijk, M. A. (2012). Corruption, growth, and governance: Private vs. state-owned firms in Vietnam. *Journal of Banking & Finance*, *36*(11), 2935–2948. https://doi.org/10.1016/j.jbankfin.2012.03.027.

Noltze, M., Euler, M., & Verspohl, I. (2018). *Sustainability in German development cooperation*. https://www.deval.org/files/content/Dateien/Evaluierung/Berichte/2018/DEval_Evaluierungssynthese_EN_web.pdf

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, *90*(4–5), 853–870. https://doi.org/10.1016/j.jpubeco.2005.05.004.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249. https://doi.org/10.1086/517935.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, *93*(7–8), 950–964. https://doi.org/10.1016/j.jpubeco.2009.03.001.

Pattillo, C., Polak, J., & Roy, J. (2007). *Measuring the effect of foreign aid on growth and poverty reduction or the pitfalls of interaction variables* (IMF Working Paper, No. WP/07/145). https://www.imf.org/external/pubs/ft/wp/2007/wp07145.pdf

PRS Group. (2018). *ICRG methodology*. https://www.prsgroup.com/wp-content/uploads/2018/01/icrgmethodology.pdf

Razafindrakoto, M., & Roubaud, F. (2010). Are international databases on corruption reliable? A comparison of expert opinion surveys and household surveys in sub-Saharan Africa. *World Development*, *38*(8), 1057–1069. https://doi. org/10.1016/j.worlddev.2010.02.004.

Reinikka, R., & Svensson, J. (2002). Measuring and understanding corruption at the micro level. In D. D. Porta, & S. Rose-Ackerman (Eds.), *Corrupt exchanges: Empirical themes in the politics and political economy of corruption* (pp. 135–146). Nomos Verlagsgesellschaft.

Rose-Ackerman, S., & Palifka, B. J. (2016). *Corruption and government: Causes, consequences, and reform*. Cambridge University Press.

Tanzi, V., & Davoodi, H. (1998). Corruption, public investment, and growth. In H. Shibata & T. Ihori (Eds.), *The welfare state, public investment, and growth* (pp. 41–60). Springer Japan.

Transparency International. (2011). *Policy position 06/2011: Making aid effective: An anti-corruption agenda*. https:// www.transparency.org/whatwedo/publication/policy_position_6_2011_making_aid_effective_an_anti_corru ption_agenda

Transparency International. (2019). *Corruption Perceptions Index. Corruption around the world in 2019*. https://www. transparency.org/en/cpi

Vawda, A. Y., Moock, P., Gittinger, J. P., & Patrinos, H. A. (2003). Economic analysis of World Bank education projects and project outcomes. *International Journal of Educational Development*, *23*(6), 645–660. https://doi.org/10.1016/ S0738-0593(03)00100-7.

Ware, G. T., Moss, S., Campos, J. E., & Noone, G. P. (2007). Corruption in public procurement: A perennial challenge. In J. E. Campos & S. Pradhan (Eds.), *The many faces of corruption: Tracking vulnerabilities at the sector level* (pp. 295–334). International Bank for Reconstruction and Development and World Bank.

Winters, M. S. (2014). Targeting, accountability and capture in development projects. *International Studies Quarterly*, *58*(2), 393–404. https://doi.org/10.1111/isqu.12075.

Wolfensohn, J. D. (2005). *Voice for the world's poor: Selected speeches and writings of World Bank president James D. Wolfensohn, 1995–2005*. http://documents.worldbank.org/curated/en/717111468137730697/pdf/343620PAPE R0Vo101OFFICIAL0USE0ONLY1.pdf

World Bank. (2001a). *Linking IDA support to country performance: Recent experience and emerging issues*. https://ida. worldbank.org/sites/default/files/pdfs/country_performance.pdf

World Bank. (2001b). *World Bank lending instruments: Resources for development impact*. http://siteresources.world bank.org/INTBULGARIA/Resources/Lending_Instr_Eng.pdf

World Bank. (2017a). *Additions to IDA resources: Eighteenth replenishment. Towards 2030: Investing in growth, resil-ience and opportunity*. http://documents.worldbank.org/curated/en/348661486654455091/pdf/112728-correct-file-PUBLIC-Rpt-from-EDs-Additions-to-IDA-Resources-2-9-17-For-Disclosure.pdf

World Bank. (2017b). *Enterprise surveys: Indicator descriptions*. https://www.enterprisesurveys.org/content/dam/enter prisesurveys/documents/Indicator-Descriptions.pdf

World Bank. (n.d.-a). *World Bank Projects and Operations* [Data set]. https://datacatalog.worldbank.org/dataset/world -bank-projects-operations

World Bank. (n.d.-b). *Enterprise Surveys Data* [Data set]. http://www.enterprisesurveys.org/data

World Bank. (n.d.-c). *IBRD Statement of Loans - Historical Data* [Data set]. https://finances.worldbank.org/Loans -and-Credits/IBRD-Statement-Of-Loans-Historical-Data/zucq-nrc3

World Bank. (n.d.-d). *IDA Statement Of Credits and Grants - Historical Data* [Data set]. https://finances.worldbank.org/ Loans-and-Credits/IDA-Statement-Of-Credits-and-Grants-Historical-Dat/tdwh-3krx

World Bank. (n.d.-e). *World Development Indicators* [Data set]. https://datacatalog.worldbank.org/dataset/world-devel opment-indicators

World Bank. (n.d.-f). *Explore the Project API* [Data set]. http://search.worldbank.org/api/v2/projects

World Bank Independent Evaluation Group. (2014). *Guidelines for reviewing World Bank implementation completion and results reports: A manual for evaluators*. https://ieg.worldbankgroup.org/sites/default/files/Data/ICRR_Evalu atorManualAugust2014.pdf

World Bank Independent Evaluation Group. (2015). *World Bank project performance ratings - Codebook*. https://ieg. worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratings-codebook_092015.pdf

World Bank Independent Evaluation Group. (2017). *Results and performance of the World Bank Group 2017. An inde-pendent evaluation*. https://openknowledge.worldbank.org/handle/10986/29765

World Bank Independent Evaluation Group. (2018). *IEG Data* [Data set]. http://ieg.worldbankgroup.org/data

World Bank Operations Evaluation Department. (1999a). *1998 Annual review of development effectiveness*. http://documents.worldbank.org/curated/en/118351468780618912/pdf/multi-page.pdf

World Bank Operations Evaluation Department. (1999b). *Development effectiveness in health, nutrition, and population. Lessons from World Bank Experience*. http://documents.worldbank.org/curated/en/694581468770364130/pdf/multi-page.pdf

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.