

Good Countries or Good Projects?

Comparing Macro and Micro Correlates of World Bank
and Asian Development Bank Project Performance

David Bulman

Walter Kolkma

Aart Kraay



WORLD BANK GROUP

Development Research Group

April 2015

Abstract

This paper examines the micro and macro correlates of aid project outcomes in a sample of 3,821 World Bank projects and 1,342 Asian Development Bank projects. Project outcomes vary much more within countries than between countries: country-level characteristics explain only 10–25 percent of project outcomes. Among macro variables, country growth and the policy environment are significantly positively correlated with project outcomes. Among

micro variables, shorter project duration and the presence of additional financing are significantly correlated with better project outcomes. In addition, the track record of the project manager in delivering successful projects is highly significantly correlated with project outcomes. There are few significant differences between the two institutions in the relationship between these variables and project outcomes.

This paper is a product of the Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at akraay@worldbank.org.

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

GOOD COUNTRIES OR GOOD PROJECTS? COMPARING MACRO AND MICRO CORRELATES OF WORLD BANK AND ASIAN DEVELOPMENT BANK PROJECT PERFORMANCE

David Bulman (Harvard Kennedy School)

Walter Kolkma (Independent Evaluation, Asian Development Bank)

Aart Kraay (World Bank)

JEL Codes: O11, O12, O22

Keywords: World Bank, Asian Development Bank, Project Success

1818 H Street NW, Washington, DC. The authors are grateful to Adele Casorla for help compiling the data of the Independent Evaluation Department of the Asian Development Bank, and to Adele Casorla, Luis Serven, Ganesh Rauniyar, Vinod Thomas, Jiro Tominaga, and Hans Van Rijn for helpful comments on earlier drafts. The views expressed here are the authors' and do not reflect the official views of the World Bank, the Asian Development Bank, their Executive Directors, or the countries they represent.

1. Introduction

A large literature has studied the effectiveness of development aid, with two broad areas of interest. The first strand has focused on the country-wide effects of aggregate aid inflows, typically on aggregate outcomes such as per capita GDP growth. Out of necessity, this literature has emphasized the role of aggregate country characteristics, such as the overall policy and institutional environment, as intermediating the effects of aid on growth. The second strand in this literature has emphasized the micro side of aid effectiveness, based on the in-depth evaluation of specific aid-financed development interventions at the project level. Out of necessity, this literature has mostly emphasized the role of project characteristics in determining the effectiveness of specific interventions.

However, much less is known about the *relative* importance of country versus project characteristics in driving aid effectiveness. Yet understanding the role of country versus project characteristics is of considerable importance to large aid donors that implement many projects in a broad cross-section of countries. For example, large multi-country donors need to decide how aid is allocated both across countries as well as within countries across specific projects, and they also need to determine the role of country and project-level characteristics in their decision rules.

Systematic analysis of large numbers of projects implemented by multilateral development banks, and assessed using reasonably common standards, allows us to shed light on this important issue. In this paper, we build on earlier work by Denizer, Kaufmann, and Kraay (2013) (DKK), who investigate the macro and micro correlates of World Bank project outcomes using data from over 6,000 World Bank projects. Although DKK find that project outcomes are strongly correlated with country-level macro institutions and economic conditions, country-level factors account for only 20 percent of the total variation in project outcomes. The remaining 80 percent of the variation occurs within countries across projects. DKK investigated the relationship between a large set of project characteristics and project outcomes, and documented the role of factors such as project size, project length, the effort devoted to project preparation and supervision, and early-warning indicators that flag problematic projects during the implementation stage in explaining the within-country variation in project outcomes. DKK also documented a strong role for project manager effects in driving project outcomes. Geli, Kraay, and Nobakht (2014) (GKN) extend these results to develop an empirical model for predicting eventual project evaluation, for the set of 'active' projects under implementation in the World Bank's portfolio.

This paper extends the analysis of DKK and GKN by comparing the correlates of project success in the Asian Development Bank with those in the World Bank. This allows us to study similarities and differences across institutions in the relationship between project outcomes and country and project characteristics. To our knowledge few studies have compared the organization-specific determinants of project success. One notable exception is Honig (2014), who examines the relationship between a particular country characteristic (fragility) and development organization characteristics (autonomy of local implementers and autonomy of the institution itself from political interference) in determining aid effectiveness, finding that more “autonomous” international development organizations are more successful operating in fragile states than organizations with less autonomy.

Using data from 3,821 World Bank projects and 1,342 ADB projects, we again find that project success rates vary more within countries than across countries. In the two institutions, country-level characteristics explain only 10-25% of project success, indicating an important role for project-specific factors in understanding project outcomes. Among country-level factors, we find that, consistent with DKK, GDP growth and a good policy environment are positively correlated with project success. In contrast with DKK, we find that across projects in Asia for both institutions, civil liberties and political freedom at the country level are negatively correlated with project outcomes. With regard to the micro correlates of project success, we find that projects that take longer to implement are less likely to be successful. We also find that the difference between actual and initially-planned funding is positively correlated with project outcomes, likely reflecting the fact that projects that are not doing well are closed early. Additionally, we find that the track record of the project manager (known as the “task team leader” in the World Bank, and the “project officer” in the Asian Development Bank), defined as the success rate of the project manager on other projects, is a very strong correlate of eventual project outcomes. Leading indicators of project success, in the form of negative project ratings by staff during the first half of a project, are also correlated with eventual project outcomes. In terms of differences across institutions, in most cases we cannot reject the null hypothesis that the magnitude of the relationship between these macro and micro correlates and project outcomes is the same across the two institutions.

Beyond its immediate antecedents in DKK and GKN, this paper contributes to a growing literature that has studied aid effectiveness by analyzing outcomes of public spending projects financed by multilateral development banks. An early contribution was the 1991 World Development Report (World Bank 1991), which noted higher economic rates of return on World Bank financed projects in

countries with good policy performance, based on background research that was eventually published as Isham and Kaufmann (1999). Related work by Isham, Kaufmann, and Pritchett (1997) documented the significance of political rights and civil liberties for rates of return in a global sample of World Bank financed projects. More recently, papers such as Dollar and Levin (2005) and Guillamont and Laajaj (2006) consider other country-level factors such as policy quality and macroeconomic volatility in accounting for project-level success. As we have already noted, however, the vast majority of the variation in project outcomes occurs within countries across projects, indicating at best a limited role for country-level factors in accounting for project level success rates. The second part of our paper, which explores this, builds on several recent studies that have also examined project-level correlates of project outcomes, including Dollar and Svensson (2000), Kilby (2000), Chauvet, Collier, and Fuster (2006), and Kilby (2011, 2012). Relative to these studies, we emphasize (i) the distinction between cross-country versus within-country variation in project outcomes,¹ (ii) the role of project manager quality in driving project outcomes, and (iii) the role of early warning signals of project outcomes coming from internal project monitoring processes.

The rest of this paper proceeds as follows. Section 2 describes the source of project outcome ratings for the WB and the ADB and provides some basic data description. Section 3 investigates the relationship between project outcomes and a variety of ‘macro’ country-level variables and ‘micro’ project-level variables, emphasizing comparisons between the WB and the ADB. Section 4 uses a smaller sample of projects for both institutions where we have information on the identity of the project manager of the project, to investigate the role of project management in project outcomes and the role of leading indicators in the form of negative outcome ratings during the first half of a project’s life. Section 5 discusses policy implications and concludes.

2. World Bank and Asian Development Bank Project Outcome Ratings

We begin with a sample of 5,038 WB projects exiting the World Bank’s portfolio since 1995. This is the same set of projects used in GKN. For the WB as a whole, over 10,000 projects have been completed since the late 1940s. However, we focus on this more recent set of projects because of the more complete information on project characteristics available for them, most notably information on

¹ For related work on this issue in a very different context of aid projects in different communities in Pakistan, see Khwaja (2009).

the identity of the project manager. For the ADB we work with a sample of 1,696 projects exiting the ADB’s portfolio since 1973, as obtained from its evaluation databases. Eliminating a few outliers, and restricting the regression sample to projects with full data across the main variables of interest limits the sample to 1,342 ADB projects and 3,821 World Bank projects, for a total of 5,163 projects. The subsequent discussion and analysis refers to this restricted sample.

Table 1: Distribution of Projects across Sectors

	<i>World Bank</i>		<i>World Bank - Asia</i>		<i>Asian Development Bank</i>	
	<u># of projects</u>	<u>Project value</u>	<u># of projects</u>	<u>Project value</u>	<u># of projects</u>	<u>Project value</u>
<i>Agriculture</i>	11.9%	9.7%	14.8%	11.7%	25.9%	15.6%
<i>Transport</i>	12.0%	13.8%	14.6%	16.8%	16.8%	20.7%
<i>Public admin.</i>	22.8%	19.1%	15.4%	10.5%	3.0%	6.0%
<i>Energy</i>	9.0%	13.9%	12.6%	19.8%	14.4%	19.3%
<i>Education</i>	10.4%	7.8%	10.1%	7.5%	8.2%	4.8%
<i>Finance</i>	5.7%	10.5%	5.8%	10.3%	9.2%	13.9%
<i>Water</i>	8.0%	6.8%	8.7%	7.3%	10.7%	8.1%
<i>Industry</i>	7.1%	9.3%	7.3%	8.8%	4.2%	4.0%
<i>Health</i>	12.0%	8.2%	9.5%	6.6%	3.3%	2.6%
<i>Other</i>	1.2%	1.0%	1.2%	0.7%	4.4%	5.0%
Total number / value (billion 2005 USD)	3821	420.7	1146	167.6	1342	130.2

Notes: This table presents the distribution of projects (both number of projects and the total value of projects by initial ADB/WB commitment) in the World Bank and ADB, as well as the distribution of World Bank projects in ADB countries. Project values are determined by initial commitments in constant 2005 USD.

WB and ADB lending and grant-making activities are organized by projects. Most often these projects finance particular public sector activities, such as infrastructure projects, health and education initiatives, and a myriad of potential other development-oriented government actions supported by aid donors. In some cases projects simply take the form of budget support, adding some conditions that governments need to meet in order for disbursement to occur. To give a sense of the diversity of these projects, **Error! Reference source not found.** reports the distribution of projects across sectors.²

² World Bank projects can be assigned to multiple sectors, with an indication of the fraction of the project’s value in each sector. We follow DKK in assigning each project to its largest sector. In the ADB data we have information on only one sector per project, which is reflected in the table above.

Compared with the WB, the ADB tends to have a larger share of projects in agriculture and in finance (even within Asia), but a smaller share of projects concentrated on public administration and health.

Projects are identified and designed through a collaborative process involving development bank staff and their counterparts in the country where the project will be implemented. A key ingredient of the World Bank project design process is the identification of the project's "development objective" (or "development outcome"), which summarizes what the project is intended to achieve. Upon completion, projects are assessed according to their success in achieving this development objective. In addition, over the course of project implementation, project managers regularly report on the status of the project using the Implementation Status and Results Report (ISR). These ISRs include the project manager's assessment of whether the project is making good progress relative to its development objective, using the same rating scale that is used to ultimately assess the project (as discussed below). We use these ISR-DO ratings as an interim assessment of overall project quality. Analogously, project managers in the ADB regularly fill out interim Project Performance Reports and provide "Impact and Outcome" (IO) ratings, to predict the possible achievement of impact and outcome by completion date based on current assumptions and risks (this was the system from 2000 to 2010). Below we discuss in more detail these interim ratings and their usefulness in predicting eventual project outcomes.

Upon completion, WB projects are self-assessed by the project manager, in the form of an Implementation Completion Report (ICR). In the WB, during the post 1995 period that we consider, all Implementation Completion Reports were desk-reviewed by the WB's Independent Evaluation Group (IEG). The summary evaluation at this stage consists of a six-category rating of the project's outcome relative to its development objective (*Highly Unsatisfactory/Unsatisfactory/Moderately Unsatisfactory/Moderately Satisfactory/Satisfactory/Highly Satisfactory*). In addition, roughly 25 percent of projects are subject to more detailed reviews based on field missions by the Independent Evaluation Group, which produces a detailed evaluation study of the project known as a "Project Performance Assessment Report." These reports also assess the project's outcome relative to its development objective, using the same six-category scale. We use these detailed evaluation ratings for all projects where they are available, for a total of 1,022 projects, and the ICR review-based evaluations for the remaining 2,799 projects in our sample. We in addition convert the six-point scale to a binary *Successful/Unsuccessful* classification by grouping the three gradations of each together.

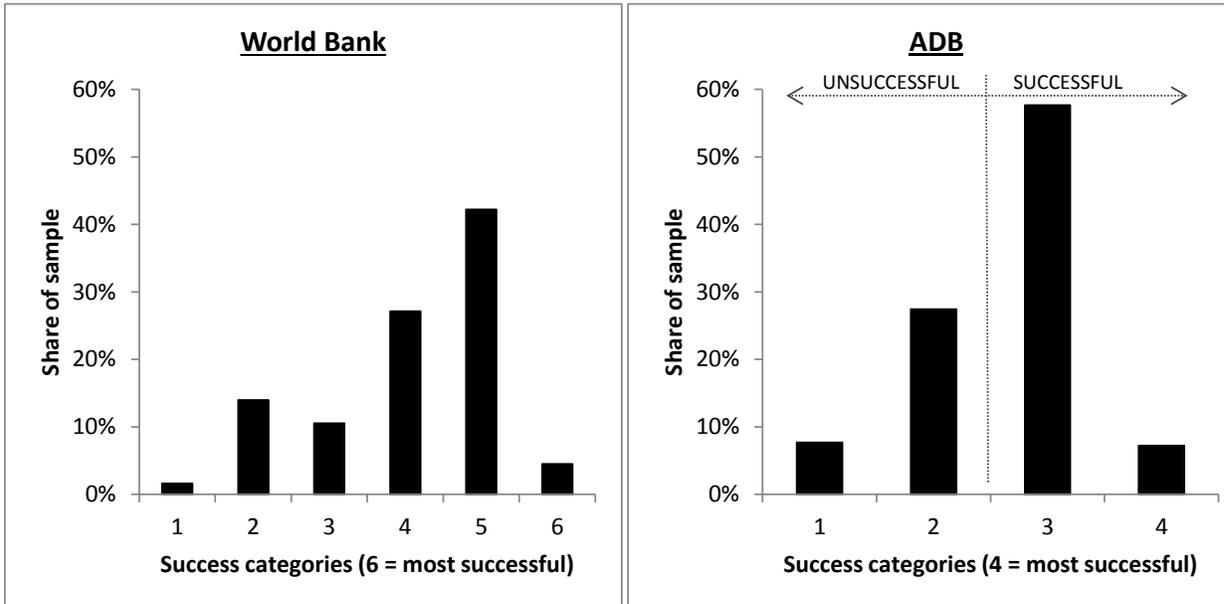
The ADB follows a broadly similar process. Self-evaluations by project managers are known as Project Completion Reports (PCRs). They are done for all completed projects which incurred expenditures. Since 2007, PCRs have been desk reviewed by the ADB's Independent Evaluation Department using a Project Completion Validation Report (PVR). In addition, nearly 50 percent of projects in our sample receive detailed evaluations in the form of Project Performance Evaluation Reports (PPER). Most of these were done prior to 2007 when the validation system started. As in the World Bank data, we use the most detailed evaluation available, i.e. the PPER if available, otherwise the PVR, and if neither is available we rely on the staff self-assessment in the form of the PCR. Finally, we note that prior to 1995, we have evaluations only for projects that were subject to a full PPER. Staff self-assessments through PCRs prior to 1995 were not formally rated, and are therefore not included in our sample for analysis.³

ADB projects evaluated before 2000 were rated on a three-point scale (*Unsuccessful/Partly Successful/Generally Successful*). After 2000, the ADB split the *Generally Successful* category in two, resulting in a four-point scale (*Unsuccessful/Less than Successful/Successful/Highly Successful*). To generate a binary success rating for use in our combined analysis of WB and ADB projects, we define the first two categories as *unsuccessful* in both the pre-2000 and post-2000 periods, and all *Generally Successful/Successful/Highly Successful* projects as *successful*.

Figure 1 shows the distribution of projects across the various outcome categories, pooling all projects for the WB (left panel), and for the ADB (right panel).

³ The lack of PCR ratings prior to 1995 explains the high share of projects (48%) in our sample that receive detailed Project Performance Evaluation Reports (PPERs). From 1995 onwards, 294 out of 995 ADB projects have PPERs (30%).

Figure 1: Distribution of Project Outcome Ratings



Notes: For ADB projects evaluated before 2000, a three point scale was used (“Unsuccessful,” “Partly successful,” and “Generally successful”). The first two categories match categories post-2000, but the “Generally successful” category splits into “Successful” and “Highly successful.” To generate this graph, we distributed the pre-2000 “Generally successful” ratings into the post-2000 successful categories (“Successful” and “Highly Successful”) assuming the same distribution between the two (88% “Successful” and 12% “Highly successful”).

One important feature of this data to keep in mind is that, while there are obvious similarities in the terminology used to define the rating scale for the two institutions (using gradations of “success”), the overall success rate of projects across the WB and ADB may not be fully comparable, as evaluators at the two institutions may have somewhat different standards and norms for determining what constitutes a “successful” (ADB) or “satisfactory” (WB) rating. For instance, ADB includes a sub-rating for sustainability in the determination of the overall success rating; the WB has since the early 2000s produced a separate rating for sustainability, which they call ‘*risk to outcomes*.’ There are also differences in the distribution of projects across countries, which further complicates the comparison of overall aggregates of project success. In the empirical specifications that follow, we include a separate intercept for the WB and the ADB, which will pick up any differences in average success rates across the two institutions. Also, to pick up any differences in project success rates across both sectors and institutions, we include a full set of sector dummies for both institutions.

There are a variety of other reasonable concerns about the validity of these project outcome ratings. One general concern is that all projects are assessed relative to their development objective, rather than relative to any absolute standard. This introduces the possibility that at least some of the variation in project outcome ratings is due to differences in the ambition or attainability of the stated development objective, rather than due to any differences in actual outcomes. To some extent this problem is inevitable, given the wide variety of sectors in which the ADB and the WB operate: it would be difficult to imagine a common absolute standard that could be applied to the literally thousands of very different projects that these institutions have financed. This also means that differences in project success rates across sectors should not be taken too seriously, as they may reflect both differences in actual outcomes as well as differences in standards of setting project development objectives across sectors.

Another potential concern is that we are pooling results from two different types of evaluations for the WB, and three types for the ADB. For example, one might be concerned that project success ratings based primarily on the views of staff implementing the project (in the form of PCRs in the ADB, and their desk reviews in both institutions (the ICR and the PVR)), and that staff “close” to projects may naturally be more reticent about admitting less than satisfactory performance. To capture this possibility, in the regressions that follow we include two dummy variables for evaluation type. The first takes value equal to one if the project received a full evaluation in the form of a PPAR/PPER, and zero otherwise. The second, relevant only for ADB projects, takes value one for projects that receive PVRs, and zero otherwise. The ADB projects that receive a zero correspond to projects that receive only PCRs and are not desk reviewed (33% of projects in the ADB; in the WB, all ICRs are reviewed by the IEG in the post-1995 sample that we work with).

A final caveat we should note is that these project evaluations are by no means well-identified (in the econometric sense) impact evaluations. Rather, they are reasonably careful administrative assessments which generally rely on special project field visits held within two years of project completion and careful write-ups of the findings, with economic and financial analysis done where feasible. For several major categories of projects, such as in finance or public administration, it would be difficult to conduct a rigorous impact evaluation; many projects are furthermore dispersed over many project sites and rigorous impact evaluations would then be difficult and costly to organize. Although not rigorously impact evaluated using counterfactuals, a very significant sample in both groups

was nevertheless independently evaluated, which helps to minimize the conflicts of interest and optimism biases inherent in self-evaluations by project managers.

With these caveats in mind, Figure 2 shows trends over time in average performance across evaluation types in the World Bank and ADB.

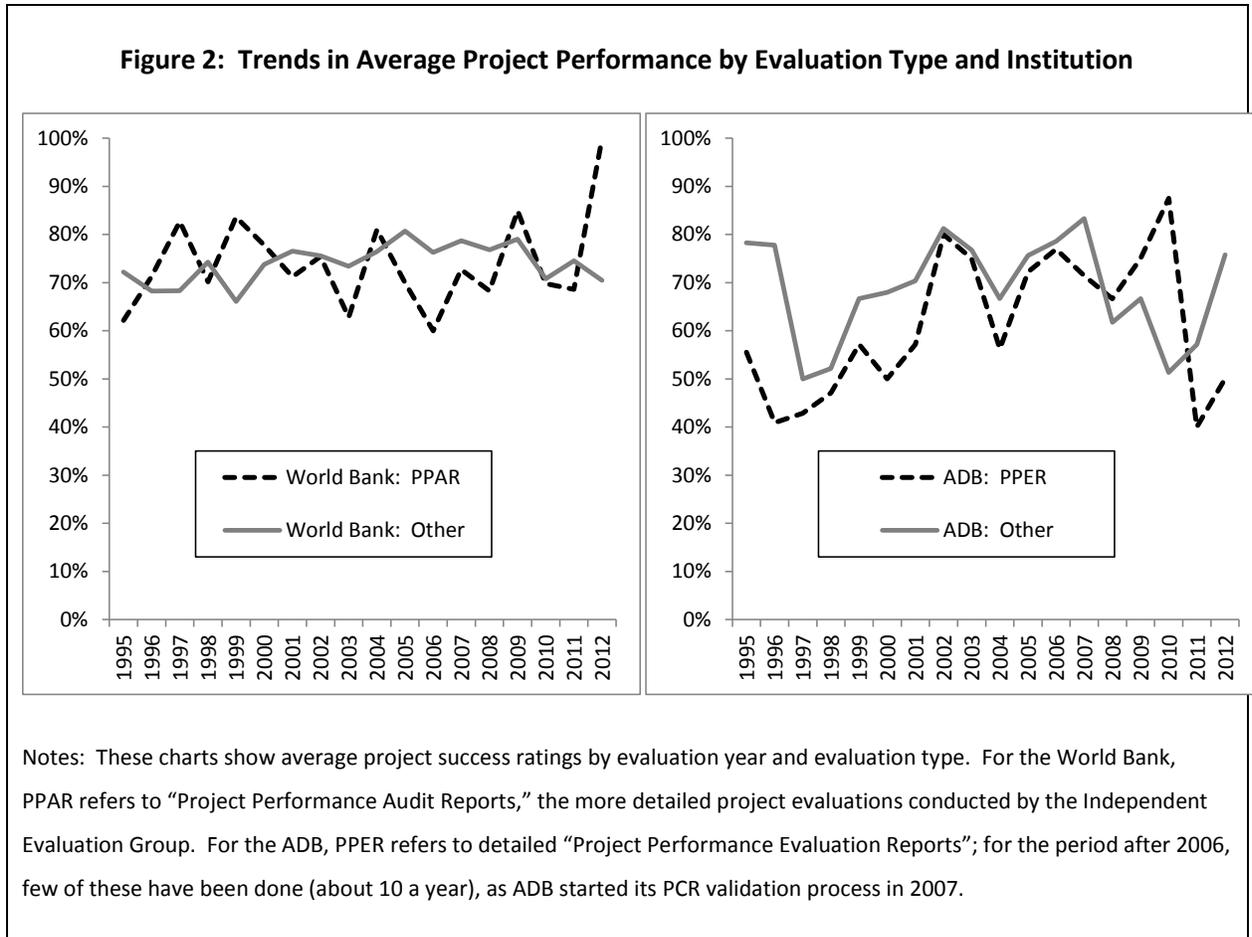


Figure 3 gives a first sense of the extent of variation in project outcomes across Asian countries in which both the WB and the ADB are active. The concentration of observations in the upper left hand quadrant reflects the higher mean success rate across World Bank projects, demonstrating that this higher mean success is not simply an artifact of different project distribution across countries for the two institutions. However, different project distribution across Asian countries does account for some of the difference in overall success rates shown in Figure 2. For instance, in the sample 17% of World

Bank Asia projects are in China, which has a high project success rate for both institutions (90%), while only 7% of ADB projects are in China. If we assume ADB country success rates but apply World Bank Asia country distribution of projects, the ADB overall success rate would rise from 64.9% to 71.0%; similarly, applying ADB country distribution to World Bank country success rates lowers the World Bank Asia success rate from 78.3% to 76.0%.⁴

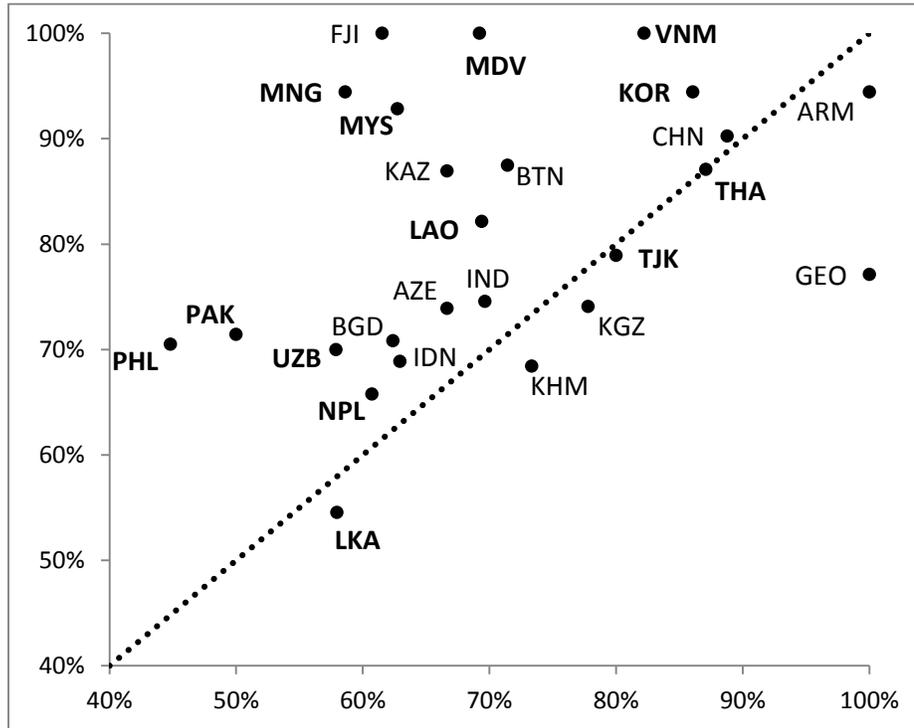
Although Figure 3 shows that there is non-trivial variation across countries in average project performance for both institutions, a key feature of WB project ratings documented in DKK is that most of the variation in project outcomes occurs across projects within countries. Put differently, while there clearly are cross-country differences in average project success rates, as demonstrated in Figure 3, there is also a great deal of variation within countries, with successful and unsuccessful projects coexisting in the same country.

One way to document this within-country variation directly is to consider a regression of project outcomes on country dummy variables. The R-squared from such a regression corresponds to the share of the variation in project outcomes in a given year that can be accounted for by differences in average performance across countries. Specifically, for each year from 1995 to 2007, we take the set of projects active in that year, and regress the ultimate project outcome on a set of country dummies. We do this for WB and ADB projects separately, and also for WB projects implemented in the same set of countries as the ADB. The resulting R-squareds are quite low, varying between 10 and 25 percent, and averaging 17.6% (WB), 14.3% (WB Asia), and 14.6% (ADB) in the three samples.⁵ This motivates an analysis of the project-level correlates of success, which we turn to next.

⁴ Average project success rates vary across sectors, and the sectoral composition of projects differs slightly between the ADB and the WB. However, these differences are sufficiently small that they do not account for much of the difference in overall project success rates between the two institutions. Assuming ADB sector success rates and applying the WB Asia sector distribution results in a success rate of 65.1% compared to the ADB baseline of 64.9%. Similarly, assuming the WB sector success rates and applying ADB sector distribution results in a success rate of 77.4% compared to a WB Asia baseline of 78.3%. The different treatment of sustainability issues in evaluations in World Bank and ADB in the 2000s may also explain part of the difference. Sustainability sub-ratings are generally somewhat lower than overall ratings in ADB.

⁵ We cut off the analysis in 2007 due to the more limited number of projects to analyze from 2008 onwards, which artificially drives the R-squareds higher. The data exhibit a general upward trend in the R-squareds for Asian countries after 2000, both among ADB and WB projects. However, this should not necessarily be interpreted to mean that country effects are increasingly driving project performance. Rather, the number of projects in the regressions declines every year after 1999, and with fewer observations the R-squared values move upwards. Even in the last few years, over 70 percent of the variation in ADB project outcomes is due to variation across projects within countries.

Figure 3: Distribution of Average Success Rates by Country



Notes: This chart compares average project success rates by country in the World Bank (y axis) and ADB (x axis). All countries with significant differences in success rates between the two institutions (indicated by z scores greater than two) are identified by bolded and underlined country labels. Observations with labels in gray do not have significantly different mean success rates across the two institutions.

The ISO country codes correspond to: Armenia (ARM), Azerbaijan (AZE), Bangladesh (BGD), Bhutan (BTN), Cambodia (KHM), China (CHN), Fiji Islands (FJI), Georgia (GEO), India (IND), Indonesia (IDN), Kazakhstan (KAZ), Kyrgyz Republic (KGZ), Lao People's Democratic Republic (LAO), Malaysia (MYS), Maldives (MDV), Mongolia (MNG), Nepal (NPL), Pakistan (PAK), Philippines (PHL), Republic of Korea (KOR), Sri Lanka (LKA), Tajikistan (TJK), Thailand (THA), Uzbekistan (UZB), and Vietnam (VNM).

3. Correlates of Project Outcomes

We next document the empirical relationship between project outcomes and a variety of country-level and project-level characteristics. We do so using a series of probit regressions of the binary outcome rating on these variables, pooling all WB and ADB observations, and allowing the estimated coefficients on all variables to differ across institutions. We report results for a pooled

sample of all projects from both institutions, and for a sample eliminating WB projects outside Asia.⁶ In the probit results, we report estimated marginal effects for all continuous explanatory variables, and estimated coefficients for the discrete explanatory variables.

As discussed above, differences in the rating scales used by the WB and ADB to evaluate projects imply that there may also be differences in reported overall average success rates across the two institutions. In addition, we have discussed how project success rates might differ across sectors due to differences in the types of development objectives set for projects in different sectors. To pick up these differences, we include a full set of sector dummy variables, and their interaction with a dummy for ADB projects, in all the regressions that follow (not reported for reasons of space). All specifications also include year fixed effects and their interaction with the ADB dummy, to pick up potential changes over time in evaluation standards in the two institutions (also not reported for reasons of space). Finally, all specifications include a dummy variable for projects where the outcome rating is based on the more detailed PPAR/PPER evaluation, as well as its interaction with an ADB dummy, and also a dummy variable for projects where the outcome rating is based on PVR evaluations (relevant only for ADB projects) to pick up any effects of evaluation type.

In addition to documenting the characteristics of successful projects, we also are interested in the extent that these differ between the WB and the ADB. In order to facilitate this, we add an interaction of each right-hand-side variable in the probit regression with a dummy variable indicating ADB projects, so that the interaction term picks up differences across institutions in the estimated relationship between the variable of interest and project outcomes.

In the first two columns of Table 2 we look at the relationship between three major country characteristics and project outcomes. Since most projects require several years to implement, we measure each of these country characteristics as the annual average of the variable in the country and over the years in which the project was implemented. The first is a measure of country-level policy performance, the Country Policy and Institutional Assessment (CPIA) ratings of the World Bank.⁷ The

⁶ As a further robustness check, we also estimated our specification for a sample that excluded ADB projects evaluated prior to 1995, in order to match the timing of the WB project sample. However, the results are very similar in this smaller sample, and so are not reported to conserve space.

⁷ The ADB produces a similar set of assessments, but only for poorer client countries eligible for concessional lending from the Asian Development Fund. In contrast, the WB produces its CPIA assessments for all clients, but discloses them publicly only for concessional borrowers. For concessional borrowers in Asia, the WB and ADB CPIA ratings are quite highly correlated.

CPIA rates countries on 16 criteria in four clusters: economic management, structural policies, policies for social inclusion and equity, and public sector management. CPIA ratings are very strongly positively and significantly correlated with project outcomes across all specifications. This intuitive relationship supports the findings of Dollar and Levin (2005), who demonstrate that World Bank project success rates in the 1990s increase with institutional quality in recipient countries. Similarly, Isham and Kaufmann (1999) show that the economic rate of return on public and private investments increases within countries as economic policy making improves. The same relationship between better country-level policy performance and World Bank project outcomes is documented in DKK and GKN.

We also look at country-level real GDP growth rates over the period in which the project was implemented. Echoing the finding in DKK, projects implemented in fast-growing countries are very significantly more likely to be rated as successful. An additional percentage point of average growth over a project's life is associated with a probability of project success that is 1.4%-1.6% greater at the margin.

Finally among the country-level variables, we consider the index of civil liberties and political rights produced by Freedom House. We use this particular measure to be consistent with earlier work that has used the same indicator, and also for the pragmatic reason that it is the only such measure available for the full time span and set of countries included in our data set. Freedom House scores both of these indicators on a 1-7 scale. We sum them together and reorient to arrive at a scale from 0 to 12, with higher values corresponding to higher civil liberties and political rights. In the global sample and for WB projects, there is no significant correlation between this measure and project outcomes, in the Asian sample the relationship is significantly negative, with better project outcomes in countries rated by Freedom House as having fewer civil liberties and political rights. This differs from earlier findings. DKK find an insignificant (but positive) relationship between Freedom House ratings and project success, the same as our global sample. Isham, Kaufmann, and Pritchett (1997) find a positive correlation between civil liberties and the performance of government investment projects, using a global sample.

It is noteworthy that in nearly all cases, there is no evidence of a differential relationship between the macro variable of interest and project outcomes in the ADB as compared with the WB. The estimated coefficients on the interactions of growth and the CPIA score with the ADB dummy are statistically insignificant and small. The one exception is in the full sample of projects, where the ADB interaction with the Freedom House variable is marginally significant. However, this simply is picking up an Asia effect rather than an ADB effect: as can be seen in the second and fourth columns, among Asian

countries, the negative relationship between rights and project outcomes holds, and with no evidence of a differential ADB effect.

In the next two columns, we consider the relationship between a set of project characteristics and project outcomes. The first two are measures of project size, as proxied by the logarithm of the total commitment, measured in constant 2005 \$US, and by the initial planned length of the project, measured in months from project approval to planned completion date. Both of these can be thought of as proxies for project complexity. We find that planned project size is positively (although not significantly) correlated with project success, while planned length is negatively and significantly correlated with success. In neither case is the ADB interaction significant. The positive correlation with planned size is somewhat surprising, as it is the opposite of the findings in DKK. While this correlation should be interpreted with some caution as it is not statistically significant at conventional levels, a possible intuition is that projects with greater initial commitments are given greater attention; this effect may outweigh any complexity effect (i.e., projects with larger initial commitments are likely to be more complex and thus less likely to be successful). In the ADB case, China, Vietnam, and India tend to receive large loans, and these countries, particularly China and Vietnam, also have higher project implementation capacity and greater control over local factors such as land and local government. The negative relationship between project success and planned length is more intuitive: more complex projects that are expected to take longer to implement are more likely to receive unsuccessful ratings.⁸

The next two indicators capture delays in the process of project implementation. The first, effectiveness delay, measures the time (in months) between project approval and project “effectiveness”, i.e. the time from signing the loan to the time that all conditions of the loan agreement are declared fulfilled so that disbursements can be made. The second, implementation delay, measures the difference (in months) between the actual completion date of the project and the planned completion date. We find some evidence that longer “effectiveness” delays counterintuitively signal better project outcomes. A potential channel for this effect is that in some countries, experienced executing agencies wait to fulfill all conditions until detailed project designs have been completed and all procurement packages are prepared; this then reduces later delays after effectiveness for which special “commitment” charges may be levied by the lender. The observed “delay” to declaration of

⁸ It is important to emphasize that we are looking at *planned* project length, not *actual* project length. Complex projects may get restructured or cut off and then rated unsuccessful. As components get canceled, the actual project length may be quite short.

effectiveness therefore indicates special care and attention (which then enhances the speed of subsequent implementation). More intuitively, we find evidence that projects that take longer than expected after loan effectiveness to complete are more likely to be rated poorly.⁹

We also include a variable measuring the difference between the (log) final disbursements and (log) initial commitment on the project. These differences can arise for a variety of reasons. In some cases, total disbursements are less than the initial commitment if the project was performing poorly. Conversely, projects that are performing well may receive additional financing beyond their initial commitment. Together these suggest a positive relationship between this variable and project outcomes. On the other hand, sometimes additional financing is required to complete a project due to unforeseen cost overruns. In this case the relationship between this variable and project outcomes is less clear. In spite of cost overruns, some projects are still rated successful due to their positive outcomes. Conversely, projects that at their midterm are deemed unlikely to achieve positive outcomes have loan cancellations applied towards the second half of the implementation period. We find that additional financing is strongly and significantly correlated with project success, lending support to the first interpretation of additional financing.¹⁰

Finally, we note that in all specifications, the PPAR/PPER evaluation type dummy variable's interaction with the ADB dummy is negative and significant, and the ADB PVR evaluation type dummy variable is also significantly negative. In other words, in the ADB case, evaluation ratings based on more detailed PPERs and PVRs on average result in project ratings that are significantly lower than those based on Project Completion Reports. There is no significant difference in project ratings coming from the two evaluation types (PPARs and ICR reviews) in the WB.

A challenge in interpreting these findings is that both project ratings and observed project-level variables to some extent respond to unobserved project characteristics that ultimately determine the success or failure of the project. For instance, as indicated above, projects that seem to be failing are

⁹ While this is somewhat intuitive, there is a countervailing trend as well: projects that are restructured or cut short tend to get poor ratings. For instance, the ADB Pakistan portfolio was comprehensively restructured in 2007-2010, with all projects that were discontinued receiving poor ratings. Allowing some of these projects to complete may have enabled more successful outcomes from the same set. Funds freed were invested in new projects.

¹⁰ In the ADB, few projects get additional funding, as the 'supplementary financing' policy was difficult to comply with in the past and acted as a deterrent. While this policy changed recently (2008), very few projects end up with officially approved additional financing, and fewer are evaluated or validated, implying that for ADB projects, the likely interpretation is that underperforming projects are cut short, with disbursements less than initial commitments.

likely to be cut short; may be more likely to receive additional funding in order to achieve success (or alternatively receive less money if such funds are seen as throwing good money after bad); and may receive less attention and are unable to finish on time. For these reasons, the partial correlations between project outcomes and project characteristics should be interpreted with some caution. For a more detailed discussion of the extent to which a causal interpretation can be assigned to these relationships, see Section 6 of DKK.

Table 2: Correlates of Project Outcomes

Dependent variable is binary success (0,1) in all specifications

<i>Independent variables:</i>	(1)	(2)	(3)	(4)
	Full	Asia	Full	Asia
CPIA rating	0.150*** (8.03)	0.141*** (3.15)	0.145*** (7.61)	0.156*** (3.26)
---ADB interaction	-0.0436 (-1.26)	-0.0354 (-0.66)	-0.0282 (-0.77)	-0.0404 (-0.71)
Real GDP per capita growth	1.645*** (6.20)	1.416** (2.45)	1.660*** (6.09)	0.957 (1.55)
---ADB interaction	0.0298 (0.04)	0.243 (0.28)	-0.838 (-1.10)	-0.142 (-0.15)
Freedom House rating	0.00163 (0.57)	-0.0119** (-2.12)	0.00162 (0.56)	-0.0115** (-2.01)
---ADB interaction	-0.0123* (-1.91)	0.00130 (0.16)	-0.00887 (-1.34)	0.00429 (0.52)
Dummy for PAR/PPER evaluations (d)	0.000538 (0.03)	0.0105 (0.32)	-0.0233 (-1.33)	-0.00366 (-0.11)
---ADB interaction (d)	-0.0974** (-2.21)	-0.106** (-2.01)	-0.110** (-2.41)	-0.128** (-2.36)
Dummy for ADB PVR evaluations (d)	-0.206*** (-2.99)	-0.202*** (-2.99)	-0.203*** (-2.85)	-0.200*** (-2.85)
Log(total commitment)			0.00996 (1.55)	0.00441 (0.33)
---ADB interaction			-0.0107 (-0.69)	-0.00514 (-0.27)
Planned project length			-0.00153*** (-4.06)	-0.00139* (-1.85)
---ADB interaction			0.00142 (1.64)	0.00128 (1.19)
Effectiveness delay			0.00109 (0.59)	0.00849* (1.86)
---ADB interaction			0.00273 (0.73)	-0.00471 (-0.84)
Implementation delay			-0.00129*** (-2.62)	-0.000769 (-0.78)
---ADB interaction			-0.00156** (-2.01)	-0.00205* (-1.79)
Log(additional funding)			0.143*** (10.90)	0.252*** (7.01)
---ADB interaction			0.137*** (3.59)	0.0256 (0.51)
Observations	5155	2480	5155	2480

Notes: Table reports marginal effects from probit regression. "(d)" indicates discrete change of dummy variable from 0 to 1. T statistics are reported in parentheses. *** (**) (*) denote significance at the 1 (5) (10) percent level. All regressions include year fixed effects and year interactions with the ADB dummy.

4. Project Managers, Leading Indicators, and Project Outcomes

Thus far, we have discussed the project and country-level correlates of project success, noting mostly the similarities across World Bank and ADB projects. Following DKK, in this section we focus on the roles of project managers as well as leading indicators of project outcomes. For each project, we calculate a project manager track record variable that reflects the weighted average success rate of all projects that the project manager has worked on, excluding the current project, with weights proportional to the amount of time the project manager worked on each project. To see how this works, consider a hypothetical project manager who worked on projects A, B, and C, and we want to calculate the “track record” of the project manager for project C. This will be a weighted average of the success rating of projects A and B, with weights proportional to the fraction of each project’s life that the project manager was responsible for those two projects. For the ADB sample, we only have data on project managers from 2000 onwards, limiting the sample and also weighting the project manager track record variable towards more recent years. In other words, if a project runs from 1995-2005, the track record variable for the ADB only reflects the project managers from 2000-2005 and their average project success on other projects from 2000-2010. For both institutions, adding project manager track record limits our analysis to projects that have managers who work on other projects, reducing our sample to 2,664 World Bank projects and 579 ADB projects.

The following regressions also include an indicator for the frequency of project manager turnover. To do this, we create a variable capturing the number of managers per project year over the life of a project. In the sample, managers of ADB projects turn over more frequently than for World Bank projects: on average, ADB projects have 0.74 managers per project year, while World Bank projects have 0.44 managers. Equivalently, ADB project managers serve an average of 1.6 years per project, while WB project managers serve an average of 2.9 years per project.¹¹ Including the managers per project year variable and its interaction with the ADB dummy variable enables an analysis of the effects of project manager turnover in both World Bank and ADB projects.

¹¹ Note that the figures in this sentence are not the inverses of the figures in the previous sentence. This is because the average across project of years per project manager is not the same as the average across projects of project managers per year.

In this section, we also consider an interim indicator of project success. In principle, staff and management can respond to early unsatisfactory ratings to turn around projects that are in trouble. By including interim indicators of project success, we can control for some of the unobserved project characteristics that determine project success, making it easier to interpret the partial correlations between project outcomes and other project characteristics discussed in the previous section. In the World Bank, these interim ratings correspond to ISR-DO (Development Outcome) ratings; in the ADB they correspond to IO (Impact and Outcome) ratings referring to the expected achievement of impact and outcome by the project completion date. Similar to the project manager data, we only have ADB IO rating data from 2000 onwards, and can thus only identify a negative rating in the first half of a project's life if the project's midpoint is after 2000. This limits us to 380 more recent ADB projects. For both institutions, the variable takes the value one when a negative rating is reported at any time during the first half of the project. For the ADB, this "negative" rating corresponds to an "Unsatisfactory" or "Partly Satisfactory" rating as opposed to a "Satisfactory" or "Highly Satisfactory" rating. This occurs in 9.1% of ADB projects, versus 13.2% of WB projects. A striking feature of these interim assessments is that they are quite optimistic –upon completion a significantly larger proportion of these projects are rated as unsatisfactory (31.1% and 26.0% in the ADB and World Bank, respectively). This likely reflects not only project quality deteriorating in the second half of implementation, but also excessive optimism about ultimate project outcomes on the part of project managers rating their own projects.¹²

Including the project manager track record, managers per project year, and warning rating variables to our initial sample leaves us with 2,759 observations, 2,385 World Bank projects and 374 ADB projects.¹³ The probit results reported in Table 3 below follow the approach in Table 2 above, using this more limited sample. To ensure that the limited sample includes projects similar to the full sample, the first two columns of Table 3 present the results for the full sample above for project and country-level correlates, but use only the limited sample of projects. As can be seen comparing the first two columns of Table 3 to columns 3 and 4 in Table 2, the limited sample does not materially change the main findings discussed above.

¹² Optimism bias is a well-known feature of project management (Flyvbjerg 2006; Siemiatycki 2010; Kolkma 1999).

¹³ All World Bank projects in our sample have first half ratings. Including the warning ratings reduces the ADB sample significantly while including the project manager track record variable reduces the sample size for both institutions. The regression results reported below focus only on a sample with observations for both variables. Running separate regressions for each increases the number of observations but does not significantly alter the results.

Table 3: Correlates of Project Outcomes Including Project Manager Records and “Warning” Ratings*Dependent variable is binary success (0,1) in all specifications*

	(1)	(2)	(3)	(4)
<i>Independent variables:</i>	Full	Asia	Full	Asia
CPIA rating	0.143*** (5.66)	0.160*** (2.86)	0.0824*** (3.05)	0.0823 (1.43)
---ADB interaction	-0.0145 (-0.15)	-0.0432 (-0.42)	0.0340 (0.35)	0.0202 (0.20)
Real GDP per capita growth	1.784*** (5.34)	1.035* (1.65)	1.037*** (2.97)	0.392 (0.58)
---ADB interaction	1.059 (0.73)	1.550 (1.09)	1.236 (0.85)	1.609 (1.14)
Freedom House rating	-0.000671 (-0.19)	-0.0112* (-1.83)	-0.00546 (-1.40)	-0.0116* (-1.78)
---ADB interaction	-0.0115 (-0.99)	0.000150 (0.01)	-0.00826 (-0.71)	-0.000481 (-0.04)
Dummy for PAR/PPER evaluations (d)	-0.0150 (-0.67)	0.0188 (0.50)	-0.0600** (-2.41)	-0.0271 (-0.63)
---ADB interaction (d)	-0.377*** (-2.62)	-0.408*** (-2.70)	-0.343** (-2.23)	-0.359** (-2.21)
Dummy for ADB PVR evaluations (d)	-0.322*** (-3.08)	-0.292*** (-2.99)	-0.320*** (-2.98)	-0.282*** (-2.88)
Log(total commitment)	0.0125 (1.53)	-0.00308 (-0.21)	0.00933 (1.07)	0.0146 (0.93)
---ADB interaction	-0.000318 (-0.01)	0.0142 (0.52)	-0.00536 (-0.20)	-0.0111 (-0.41)
Planned project length	-0.00148*** (-3.05)	-0.000771 (-0.88)	-0.00173*** (-3.11)	-0.00201** (-2.06)
---ADB interaction	0.000958 (0.61)	0.000301 (0.19)	0.000825 (0.52)	0.00121 (0.74)
Effectiveness delay	-0.00102 (-0.46)	0.00929* (1.87)	-0.000299 (-0.12)	0.00789 (1.45)
---ADB interaction	-0.00235 (-0.38)	-0.0123* (-1.71)	-0.00307 (-0.50)	-0.0109 (-1.47)
Implementation delay	-0.000945 (-1.58)	0.000422 (0.37)	-0.00205*** (-3.09)	-0.00289** (-2.29)
---ADB interaction	0.00150 (1.00)	0.0000800 (0.05)	0.00237 (1.55)	0.00317* (1.81)
Log(additional funding)	0.137*** (8.70)	0.192*** (5.22)	0.0312* (1.87)	0.0762** (2.24)
---ADB interaction	0.243*** (3.32)	0.153** (2.04)	0.327*** (4.59)	0.239*** (3.37)
Project manager track record			0.216*** (6.21)	0.261*** (4.12)
---ADB interaction			0.0867 (0.85)	0.00543 (0.05)
Project manager turnover			-0.133** (-2.48)	-0.206** (-2.13)
---ADB interaction			0.193** (2.12)	0.259** (2.22)
Negative rating in first half (d)			-0.693*** (-27.07)	-0.692*** (-13.00)
---ADB interaction (d)			0.229*** (22.69)	0.197*** (12.08)
Observations	2756	1128	2756	1128

Notes: Table reports marginal effects from probit regression. "(d)" indicates discrete change of dummy variable from 0 to 1. T statistics are reported in parentheses. *** (**) (*) denote significance at the 1 (5) (10) percent level. All regressions include sector and year fixed effects and their interactions with the ADB dummy.

Columns 3 and 4 of Table 3 include the project manager track record variable and the first half rating variable along with the interaction of both variables with an ADB dummy. Column 3 corresponds to the full range of observations, while Column 4 reflects only World Bank and ADB projects in countries in which the ADB operates, similar to above. In both the full and Asia sample, the coefficient on project manager track record is large and significant, with no significant difference between the World Bank and ADB. Note that these results imply that an increase in project manager track record by one standard deviation (0.28) increases the chances of project success by 6.0%; in comparison, a one standard deviation (0.44) increase in the mean CPIA score increases the chances of project success by only 3.6%. This result is similar to that in DKK, who find that project manager track record and recipient country policy quality have comparably large impacts on project outcomes. Looking only at the Asia-only sample, project manager track record appears to have a significantly larger impact than country institutional quality: a one standard deviation increase in project manager track record (0.27) increases the chances of project success by 7.0%, versus a 2.8% increase for a one standard deviation (0.34) increase in the CPIA score.

The project manager turnover variable and its ADB interaction are significant in both the full and Asia-only specifications. The combined coefficient is negative, while the ADB interaction is positive. The coefficients have similar magnitudes and opposite signs, implying no effect of manager tenure length on ADB project success. Although we can only speculate on the different observed effects in the WB and ADB, one potential explanation is that project manager turnover in WB projects is more likely to be driven by poor project performance, while in the ADB turnover is driven by other institutional factors. In this case, project outcomes would not be correlated with project manager turnover in the ADB, and the negative correlation in the WB could be due to new managers not being able to fully address the problems in projects they were assigned to during the implementation period.

The first half rating that serves as a leading indicator of project outcomes is highly correlated with project failure. All else equal, in the combined sample, a negative rating in the first half of a project is associated with a 70% higher probability of project failure, and only 19% of projects with a negative first half rating go on to become successful. This relationship is however significantly different between the WB and the ADB: in the ADB a negative rating in the first half reduces the probability of a satisfactory rating by only 46%. These differences reflect a balance of two opposing forces. On the one hand, as noted above a significantly larger proportion of WB projects receive negative ratings in their

first half: 13.2% of projects in the WB vs. 9.1% of projects in the ADB. On the other hand, conditional on a poor rating in the first half, ADB projects are much more likely to eventually end up with a successful final rating (65% percent of projects rated as unsuccessful in the first half in the ADB, versus 14% percent in the WB).

One possible explanation of these findings is that early problem identification in the ADB sample may be more likely to lead to added effort to turn around a project; in other words, the ADB is better at responding to bad interim ratings. However, an alternative explanation is that ADB project managers may be less willing to report problems early on. For example, ADB project managers may only be willing to admit to bad interim ratings if projects have only “small” or more “solvable” practical problems (such as procurement delays, etc.), and underreport more intractable problems related to ultimate outcomes. In other words, the positive observation that of those projects flagged as problems in the first half, the ADB has a better "turnaround rate," may not be entirely good news, to the extent that the projects being flagged as problems by project managers are ones where the problems are relatively easy to fix.

It is also true that fewer ADB projects are flagged as possible problems than in the WB. This could possibly indicate lower candor on the part of ADB project managers, because overall project success rates are not so different between the two institutions. Note that this also means that relatively more projects in the ADB are rated satisfactory in the first half but get an unsatisfactory outcome rating at the end. In total, 31% of first half satisfactory ADB projects ultimately have a bad outcome, versus 17% of WB projects rated as satisfactory in the first half of project life.¹⁴

Including the first half rating as a leading indicator of “bad projects” also helps to control for some of the omitted variable bias problems that clouded the interpretation of other project-level correlates discussed above. This is because this indicator captures information observable to the project manager at the time about likely deficiencies in the project that are hard to measure ex post. Nevertheless, for the most part the earlier findings continue to hold: larger projects are still more likely to succeed, though the effect is smaller and no longer significant, which may just be due to the smaller sample; project length is still negative and significant; effectiveness delay is still positive and significant;

¹⁴ Note also that in the ADB sample, we are missing many projects that actually had first half warnings (for instance, if a project runs from 1995-2005, we would only have a 1/5 chance of identifying a flag in the first half, as the ADB flag-based monitoring system only started in 2000). This might help explain why a smaller proportion of ADB projects receive negative first half ratings, but it does not explain why ADB projects are more likely to turn around.

implementation delay is still negative and significant; first half unsatisfactory ratings raise the likelihood of an unsuccessful rating, and additional funding is still positive and significant. The ADB interaction effects remain the same whether or not the leading indicator is included.

5. Conclusions and Policy Implications

This paper has compared the country- and project-level correlates of success for a sample of 3,821 World Bank projects and 1,342 Asian Development Bank projects. The results on macro and micro correlates of project success that we find generally support those in previous literature, particularly DKK. In terms of our institutional comparison, we cannot reject the null hypothesis that the magnitude of the relationship between most of these macro and micro correlates and project outcomes is the same across the WB and ADB. Such a finding tends to support the robustness of the country-level findings, and implies that many of the project-level correlates are generalizable to aid projects overall rather than reflecting particular institutional rules and norms of the WB and the ADB. This further highlights the robustness of the project-level findings, and helps support a broader conclusion: institutional aid allocation decisions should be made based on project-level characteristics in addition to the current system that is based on country-level policy and institutional characteristics.

Our findings with regard to the macro correlates of project success are similar to those of DKK. We find that country-level characteristics explain only 10-25% of project success, indicating an important role for project-specific factors in understanding project outcomes. Among country-level factors, we find that GDP growth and a good policy environment are positively correlated with project success. The significance of CPIA scores helps to justify the IDA's Performance Based Allocation system and the ADB's Asian Development Fund system which use these scores to determine cross-country aid allocation. However, our finding for projects in Asia, that civil liberties and political diversity at the country level are not positively correlated with project outcomes, differs from both DKK and Isham, Kaufmann, and Pritchett (1997). The implication, that some one-party polities in Asia do not have less policy implementation capacity, is not surprising given the countries in question – in particular, China and Vietnam – but the broader policy implications of such a finding are limited. We would not suggest that civil liberties and political freedoms are orthogonal to project performance, but simply that certain centralized countries with strong bureaucracies have managed to implement projects successfully despite lower scores in these areas.

Our main findings with regard to project-level correlates are worth reiterating. We find that the difference between actual and initially-planned funding is positively correlated with project outcomes, likely reflecting the fact that projects that are not doing well are closed early. Intuitively, we find evidence that projects taking longer to implement are also more likely to be rated poorly. To the extent that project length is a proxy for project complexity, this suggests that project complexity may be associated with lower success – projects should not try to do too much. It also suggests that extending projects to attempt to achieve goals in spite of hitches during implementation may not always be successful, although there is also the possibility that outcomes could be even worse if bad projects are closed without trying to solve remaining problems. We find some evidence that the delay between project approval and the meeting of conditions for loan effectiveness on a project signals better project outcomes. As discussed above, a potential channel for this effect is that some countries and executing agencies do not meet conditions until detailed project designs have been completed and all procurement packages are prepared; therefore, the observed “delay” may in fact indicate concern to reduce commitment charges later, as well as increased care and attention in project planning. This reinforces the lesson that it helps if detailed project designs and procurement packages are prepared prior to project implementation.

Leading indicators of project success, in the form of negative project ratings by staff during the first half of a project, are also correlated with eventual project outcomes. The implication is that these leading indicators do not significantly contribute to turning around projects. While intuitive, this finding implies that either greater effort should be made to turn around projects, or, if significant efforts are already implemented in response to warning ratings, then these projects should be cut short rather than trying to turn them around if these efforts are unlikely to succeed.

Finally, and perhaps most importantly, our finding that the track record of the project manager is a strong correlate of eventual project outcomes implies that aid organizations could improve aid effectiveness by devoting more effort to project manager selection, training, screening, and supervision. This could be complemented by giving more project management responsibilities to project managers with good track records on past projects. While this finding is in some ways an intuitive one – project managers matter – it is one that is often overlooked in discussions of the importance of institutional rules and implementation agency quality.

Generally, the importance of project-level correlates of success and the low degree of variation explained by country-level correlates suggest a greater role for utilizing project-level correlates in

determining within-country aid allocation. Both DKK and Gelb (2010) note that finding ways to allocate aid on project-level indicators is important as a way to provide countries with greater incentives to ensure project success, as well as to scale up successful projects in countries that have worse policy and institutional environments. Our findings add weight to this conclusion by highlighting that the project-level correlates of project success are largely the same in both WB and ADB institutional settings, implying that these findings are more broadly relevant. Although this paper highlights several micro correlates of project success, considerably more attention should be given to identifying additional within-country project-level correlates of success.

References

- Deininger, Klaus, Lyn Squire, and Swati Basu (1998). "Does Economic Analysis Improve the Quality of Foreign Assistance?" *World Bank Economic Review*. 12(3):385-418.
- Denizer, Cevdet, Daniel Kaufmann and Aart Kraay (2013). "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Outcomes." *Journal of Development Economics*.
- Dollar, David and Jakob Svensson (2000). "What Explains the Success and Failure of Structural Adjustment Programs?" *The Economic Journal*. 110():894-917.
- Dollar, David and Victoria Levin (2005). "Sowing and reaping: institutional quality and project outcomes in developing countries," Policy Research Working Paper Series 3524, The World Bank.
- Flyvbjerg, Bent (2006). From Nobel Prize to Project Management: Getting Risks Right. *Project Management Journal*, vol. 37, no. 3, pp. 5-15.
- Gelb, Alan (2010). "How Can Donors Create Incentives for Results and Flexibility for Fragile States? A Proposal for IDA". Center for Global Development, Working Paper No. 227.
- Geli, Patricia, Aart Kraay and Hoveida Nobahkt (2014). "Predicting Active World Bank Project Outcomes". World Bank Policy Research Working Paper No. 7001.
- Guillaumont, Patrick and Rachid Laajaj (2006). "When Instability Increases the Effectiveness of Aid Projects". World Bank Policy Research Working Paper No. 4034.
- Honig, Dan (2014). "Letting the Driver Steer: Organizational Autonomy and Country Context in Delivering Better Aid."

Isham, Jonathan and Daniel Kaufmann (1999). "The Forgotten Rationale For Policy Reform: The Productivity of Investment Projects," *The Quarterly Journal of Economics*, MIT Press, vol. 114(1), pp. 149-184.

Isham, Jonathan, Daniel Kaufmann, and Lant Pritchett (1997). "Civil Liberties, Democracy, and the Performance of Government Projects," *World Bank Economic Review*, World Bank Group, vol. 11(2), pp. 219-42, May.

Khwaja, Asim Ijaz (2009). "Can Good Projects Succeed in Bad Communities?" *Journal of Public Economics*. 93: 899-916.

Kilby, Christopher (2000). "Supervision and Performance: The Case of World Bank Projects". *Journal of Development Economics*. 62: 233-259.

Kilby, Christopher (2011). "The Political Economy of Project Preparation: An Empirical Analysis of World Bank Projects". Villanova School of Business Economics Working Paper No. 14.

Kilby, Christopher (2012). "Assessing the Contribution of Donor Agencies to Aid Effectiveness: The Impact of World Bank Preparation on Project Outcomes". Villanova School of Business Economics Working Paper No. 20.

Kolkma, Walter (1999). Work in Progress: The Hidden Dimensions of Monitoring and Planning in Pakistan. West Lafayette: Purdue University Press.

Siemiatycki, Matti (2010). Managing Optimism Biases in the Delivery of Large-Infrastructure Projects: A Corporate Performance Benchmarking Approach. *European Journal of Transport and Infrastructure Research*, (1), March 2010, pp. 30-41.

World Bank (1991). World Development Report 1991: The Challenge of Development. New York: Oxford University Press.