Dynamics of Public Service Delivery: Evidence from Bangladesh

Imran Rasul, Daniel Rogger, Laura Litvine

October 2016

Abstract

We present the first descriptive look at dynamic spending patterns across a large representative sample of infrastructure projects, using unique project-year level panel data covering the universe of infrastructure projects conducted by the Government of Bangladesh between 2003 and 2013. This initial research allows us to draw two main preliminary conclusions: First, projects seem to follow a non-linear spending pattern, spending less in early stages of a project's life, and more in the latter half of a project's implementation period. This is true for both complete and incomplete projects, and shows that underspends do not appear only because projects get abandoned, but rather seem to be an issue arising early in a project's life and surviving throughout. This suggests investigating further the planning and early life of projects. Second, when comparing complete and incomplete projects, we observe that successfully completed projects overall did better at predicting this non-linear spending trend, required smaller revisions to planned spendings, and departed less from plan throughout the life of a project. Implementation dynamics and a project's completion status are therefore indeed correlated, and this relation should be explored further.

This project was conducted thanks to financial support from the International Growth Centre [VCS-VBGD-VXXXX-31303]. We are grateful to the IGC Bangladesh for their assistance and invaluable insights throughout this project, and thank the Implementation Monitoring and Evaluation Department and the Planning Commission of Bangladesh for their support in accessing historical data for this project. All errors remain our own.

Rasul: UCL and IFS [i.rasul@ucl.ac.uk]; Rogger: World Bank Research Department [drogger@worldbank.org]; Litvine: UCL [laura.litvine.13@ucl.ac.uk]

I. Introduction

The effectiveness of government bureaucracy matters: From a macro perspective, high-quality bureaucracy and effective public service delivery have been shown to be effective drivers of growth and overall patterns of economic development (Besley and Persson, 2010). Government expenditures moreover represent a large share of GDP in low-income countries (from 20-35% of GDP according to recent IMF figures), making their effectiveness crucial to fiscal balance, as well as levels of inequality. From a micro perspective, the presumption behind many program evaluations is that successful ones will be scaled up by governments, who have the capacity to do so.

Economists have long been interested in public services. Most of the existing applied work has focused on short-term aspects of delivery like procurement (e.g. Bandiera *et al.* 2009) or front-line service provision, especially that of teachers or health workers (e.g. Muralidharan and Sundararaman, 2011 ; Finan *et al.* 2015 ; Muralidharan *et al.,* 2016).

A recent literature has started examining the effectiveness of infrastructure provision by bureaucracies, but has so far focused on the completion rates of small-sized construction projects planned to be completed within a year (Rasul and Rogger, 2016; Williams, 2016). Yet, the majority of government projects span multiple years, and are subjected to political and economic uncertainty. As a result, time delays are frequently observed, as are deviations from planned costs (whether under or over budget), and effective delivery is often incomplete. A strand of literature on this issue of project completion originates from the World Bank, where the completion of each funded project is assessed (Kaufmann and Wang, 1995 ; Denizer *et al.* 2013). This literature however focuses on the performance of the World Bank bureaucracy rather than the receiving country's government, and can only study the selected sample of internationally aided projects.

We propose to start filling this gap in evidence by looking at the dynamics of public service delivery in a developing country context. We use unique project-year level panel data covering the universe of infrastructure projects conducted by the Government of Bangladesh between 2003 and 2013. The projects we study were planned, on average, to take five years to come to completion, with each one initially costed at USD34 million[1]. The study of such dynamic projects constitutes a departure from the small-scale and short term/one-off projects that much of the previous literature has focused on. Considering the universe of infrastructure projects allows us to explore potential differences between types of executing agencies, or again between projects involving foreign financing or not.

The administrative records we utilize record the planned and actual duration of a project, planned and actual spendings, and whether the project's objectives were successfully completed. We thus measure aggregate delays, cost under or overruns, and completion rates. It is easy to understand why low completion rates are problematic, and why previous literature has focused on them. Deviations from planned duration and costs can also be welfare reducing. Time and cost overruns are welfare decreasing at least in the short-run when construction is being delayed (Lewis et Bajari 2011), and fund under-utilization suggests that resources, if re-allocated, could have been used on other productive investments. Time and cost overruns of large infrastructure projects are also covered daily in the press, making it politically important for governments to address them. This study is, to the best of our knowledge, the first study to examine completion rates, and cost and time deviations from plan in a large sample of public projects.

---

[1] Financial variables, throughout, are converted using the exchange rate from Bangladeshi Taka to US Dollars in the year in which the project was planned, or expense made. For reference, in 2003 USD 1 = BDT 59.5 and in 2013 USD 1 = BDT 77.6.

Most interestingly, we also observe year-wise actual spendings, as well as other mid-project revisions to the initial plan. This allows us to try and explain aggregate project performance by the dynamics of project implementation. A key issue when examining government performance will be to choose the right benchmark to compare year-wise spendings to. Indeed, one could argue for example that underspending is the efficient thing to do for a particular project, thus making the initial budget the wrong comparison point to build a performance measure. To address this issue, we introduce and discuss several alternative benchmarks. First, we build benchmarks based on simple linear interpolations of the initially planned costs and duration, and then successively allow for deviations from planned costs, duration, or both. Second, we will use additional, more detailed, data on a subsample of projects that we know to have been successfully completed, and build a last benchmark based on these projects' spending path.

Note that we choose, as previous literature has done, to focus on infrastructure projects: they are key drivers of growth and firm productivity in low income countries.[2] Additionally, expenditures on infrastructure projects are sizable. For example, in Bangladesh, they represent on average 30% of all public expenditures - their effectiveness therefore matters for national fiscal credibility. Infrastructure projects are also politically important, as they are highly visible to the media and voters, as well as foreign donors. In particular, fund underutilization on development projects conducted by the Government of Bangladesh has lately been the focus of a lot of media and international attention (e.g. IMF, 2015 ; Financial Express, 2016 ; Byron, 2016). Finally, from a practical stand-point, the outcomes of infrastructure construction projects are more easily and objectively assessed than other types of public investments, which therefore makes them ideal candidates for study.

We discuss the existing literature in Section II, provide some background on Bangladesh and its bureaucracy in Section III, describe the data we have hand collected in Section IV, and discuss the construction of our outcome variables in Section V. Section VI presents results on dynamic explanations of aggregate performance, and Section VII concludes.

II. Literature review

II.1 Measuring bureaucratic performance

Despite the large costs and high political stakes engaged in infrastructure programmes, there exists only few large-sample studies measuring the project-level performance of public projects. The World Bank, as one of the largest infrastructure funders around the World and in particular in developing countries, is an exception. It has been evaluating the projects it funds in over 130 countries and maintaining a database of outcomes since at least 1970. This data reveals that on average, 30% of funded infrastructure projects are not satisfactorily completed (Kaufmann and Wang, 1995; Presbitero, 2016)[3], a figure that is relatively stable along the years.

---

[2] At the aggregate level, a large body of literature, starting with Aschauer (1989), has found a positive relation between infrastructure capital and TFP in the United States. Mitra et al. (2002) estimate this effect for Indian manufacturing sector. They find that infrastructure endowments explain a large part of TFP differences across Indian states. On the impact of transportation links, Banerjee, Duflo, and Qian (2012) estimate the impact of railroads in China and find that the proximity to the transportation network had a small impact on the levels of GDP per capita but no effect on growth. In India, Donaldson (2016) finds railroads increased trade and income per capita. On the effects of other infrastructure on firms, Duflo and Pande (2007) find that irrigation dams in India increase agricultural productivity and reduce rural poverty in downstream districts, while having the opposite effect in the districts where dams are built.

[3] The indicator of satisfactory completion used by the World Bank has changed along the years, from an initial 0/1 dummy to a 6-point score nowadays, but the assessment methodology has remained relatively stable. Its official

A recent literature in economics has also been studying the completion of public projects, attempting to explore, within-country, the links between practices in bureaucracies and project performance. Rasul and Rogger (2016) conduct a study in Nigeria where they correlated the completion rate of over 4700 small infrastructure projects implemented by the Federal Civil Service to management practices in executing agencies. They found that 38% of projects were never started, and 55% of projects remained incomplete overall. Williams (2016) conducted a similar study of over 14,000 small-scale infrastructure projects implemented by local governments in Ghana, and found again that 36% of projects were incomplete after three years. Taken together, these studies suggest that there exists a clear margin to improve the performance of public project implementation in such developing countries.

The economics literature on project implementation performance has focused almost entirely on project completion, and does not provide information about average time and cost overruns. Additional analysis we conducted using World Bank data similar to that used by Denizer *et al.* (2013) suggests that approximately 50% of WB-funded projects suffer some time overrun, and took on average 30% longer to complete than originally planned. The World Bank itself does not publish any statistics on time and cost overruns though, and the evidence here is very limited. Some information can be gathered from the planning literature: an ECOTEC study from 2003 on 60 large infrastructure projects conducted in 8 EU countries reports that 10% of projects were rated as having "weak" completion levels, only 23% were completed on time (66% were delayed), and only 18% were completed to budget (20% were under budget, 62% over budget). Time and cost deviations from plan thus seem to be the rule rather than the exception, even in developed countries we would expect to have more performant bureaucracies. Moreover, while cost overruns are more frequent overall, underspends are far from rare – something we will observe in our data as well[4] (ECOTEC Research and Consulting 2003).

There is also a large existing literature on Public Financial Management: Long concerned with government performance, this strand of literature uses country or sometimes agency-level macro data to compare planned and actual public expenditures and revenues. It then draws conclusions about the credibility of budget implementation at the country level. Among many examples, Addison (2013) is noteworthy. Using data collected for assessments by the Public Expenditure and Financial Accountability programme (PEFA), he finds that deviations from planned expenditures range from -45% to +44% in 45 low and middle income countries observed for 3-6 years, deviations varying widely even for the same country across different years. Additionally, underspends appear to be almost as common as overspends. Similar observations have been made when looking at deviations within-country at the agency level (Simson and Welham, 2014 ; Williams 2015). It therefore seems that variation in budgetary performance cannot be fully explained by country-level characteristics, and that within-country factors, such as practices in various agencies, need to be taken into account.

II.2 Explaining bureaucratic performance

A growing literature tries to explain the variation in public good delivery in developing countries by looking at the practices of bureaucrats and politicians. This literature explores a range of possible

---

name is Project Outcome Rating, and it is defined as "the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently." It encompasses both completion and expected benefits. For further details, see http://ieg.worldbankgroup.org/Data/ICRR_EvaluatorManualAugust2014.pdf

[4] Other examples from the planning literature include the well-cited series of papers by Brent Flyvbjerg and co-authors studying cost overruns in large infrastructure projects. In their initial 2003 paper, they introduce a sample of 258 large transport infrastructure projects conducted in 20 countries and over 70 years. They find that 90% of projects suffered cost overruns, of 28% of planned costs on average. They also find no apparent learning effect in the time covered by the study.

determinants of good public outcomes: bottom-up monitoring and collective action for public goods (Olken, 2007 ; Banerjee, Iyer, and Somanathan, 2008); top-down monitoring and audits (Olken, 2007 ; Ferraz and Finan, 2011); third-party monitoring by media and electoral incentives (Besley and Burgess, 2002 ; Finan and Mazzocco, 2016); clientelism and distributive politics (Burgess *et al.*, 2015); corruption and passive vs. active waste (Reinikka and Svensson, 2004 ; Bandiera *et al.* 2009; Luo *et al.*, 2015); and finally internal organisation and management practices within bureaucracies (Luo *et al.* 2015, Rasul and Rogger 2016).

This literature collectively highlights that even small tweaks to bureaucratic practices can have large impacts on the efficiency of public spending. In a study examining waste in public procurement of small goods in Italy's bureaucracies. Bandiera *et al.* (2009) for example build measures distinguishing active from passive waste, and show that passive waste is responsible for approximately 80% of overall funds wasted, therefore suggesting that simple improvements to management could greatly reduce wastage. Or again, in a large study of small infrastructure projects conducted in Nigeria, Rasul and Rogger (2016) show that small differences in the management practices of bureaucrats, namely providing senior staff with more autonomy, can have a potentially large beneficial impact on the completion of public projects.

This existing literature has however so far mostly examined short-term projects, and studied aggregate end-of-project data. It therefore does not say much about the dynamics of project implementation. Denizer *et al.* (2013) is a notable exception, and the paper perhaps closest to ours. Using data on over 6500 World Bank-funded projects evaluated between 1983 and 2011, they study the determinants of "satisfactory" project completion, as measured by the Project Outcome Rating developed by the World Bank. They include both country-level determinants (GDP, quality of institutions index, political regime) and project-level determinants, and find that country-level factors explain at most 20% of the variation in project completion. This supports the need to focus on within-country explanations. The project-level determinants they look at include both fixed factors like initial project size, duration or whether a project is spread over several sectors, and dynamic factors such as early warning flags, mid-project revisions, or again team-leader turnover. They find that the probability that a project is assessed as successful decreases significantly with project size, does not change with project duration, and is marginally decreased if a project covers more than one sector. Most interestingly, they find that dynamics indeed matter - early revisions significantly increase the probability of a good outcome, while early flagged issues decrease it. Team leader turnover is also found to be an important dynamic factor, each additional team leader decreasing the probability of a good outcome four times more than would an early flag.

In this paper, we adopt an approach similar to that of Denizer *et al.* (2013) to study the dynamics of public service delivery. We widen the focus to a representative sample of all infrastructure projects in Bangladesh and contribute new estimates of the correlations between project characteristics and outcomes.

III. Background: Public project implementation and monitoring in Bangladesh

We focus our attention on infrastructure projects implemented by the Government of Bangladesh. Infrastructure investments represent a sizable 30% of all public expenditures each year on average, and are of high public and political interest, especially in developing countries.

Projects, as defined in this context, include all tasks undertaken under a single budget code by the Government of Bangladesh. It is the primary unit of reporting for the Government and a natural unit for study. A project could for example cover all activities involved in the construction of a bridge, including surveying, consulting, land purchase, procurement, and actual construction. Infrastructure projects, which all involve either the construction of new infrastructure or the maintenance of existing structures, are of one of 17 types defined by the Government. The most frequent are transport projects (representing 18% of all projects in our sample), physical planning (16%), and water supply and housing projects (15%). They are all conducted by one or more executing agencies, and headed by a Project Director, who is usually a senior staff member of the lead agency in charge.

The infrastructure projects in our sample are conducted by 207 executing agencies. All agencies report to one of 53 Ministries, and can be of three broad types: 50% are departments of central ministries, 46.5% are autonomous public agencies reporting to central ministries, and 3.5% are local government agencies reporting to the Ministry of Rural Development and Local Government. Figure A1 provides a visual summary of this hierarchy, of agency types, listing the main agencies under each type.

Local government agencies vary most strikingly from departments of central ministries and autonomous agencies: local government agencies are headed by locally elected officials (mayors), who nominate their own senior staff. Junior staff are recruited through the central Civil Service, but most often spend their career within the same local government. There is therefore very little mobility across agencies for bureaucrats in local governments. This is confirmed by a survey of civil servants conducted by the World Bank in 2001, which revealed that staff in local governments were on average less experienced, had had less exposure to other branches of the civil service, and had fewer higher-education degrees than staff in central or autonomous public agencies.

On the other hand, there seem to be no notable difference in the rules defining recruitment and management at central and autonomous agencies. Recruitment of junior staff is conducted through the centrally managed Bangladesh Civil Service exams, while senior staff are recruited either through internal promotion, or external recruitment drives. Promotions often involve being detached to a different agency or even Ministry, and staff therefore gather more experience and exposure. Salary scales are centrally decided, but autonomous agencies seem to be granted more flexibility when deciding on the benefits they offer their staff. Preliminary research so far has however only covered rules and regulations at these agencies, and we realize that an important next step would be to investigate potential differences in management practices.

Most importantly for us, all three types of agencies have to follow the same monitoring and evaluation procedures, which are centrally managed by the Ministry of Planning. In order for funds to be disbursed, each project must, each year, first submit a request for funds to the Planning Commission. They then have to report monthly, quarterly and annually on expenses and progress made to the Implementation Monitoring and Evaluation Department (IMED). Upon completion, all projects must also submit a Project Completion Report to the IMED. Since 2003, all forms to be submitted to the IMED have been standardized, which provides us with uniform, comparable data across all agencies and projects.

IV. Data overview

We exploit three main sources of data: (i) publications by the Planning Commission of planned expenditures at the project-year level; (ii) annual reports of actual expenditures submitted to the IMED; (iii) evaluation reports prepared by the IMED upon conclusion of a project. We provide more details on each of these sources below.

The Planning Commission, a department of the Ministry of Planning, publishes each year an Annual Development Programme (ADP) detailing approved planned expenditures on all infrastructure and technical assistance projects in a financial year (FY). The ADP is made public at the beginning of the financial year, which in Bangladesh runs from July 1st to June 30th. Spending targets might get revised along the year. If so, revised plans are then included in a Revised ADP published in November, five months into the financial year. On top of annual financial allocations, the ADP also includes more general project information such as the executing agency in charge, the project type and sector, planned duration, funding arrangements, total costs and eventual revisions to total costs. We have collected all ADP and RADP publications from FY2001-2 to FY2013-14.

The Implementation Monitoring and Evaluation Department (IMED) receives standardized monthly, quarterly and annual reports on expenses made on each project. Starting in 2003 they have been entering, at the end of each financial year, the aggregate expenditures and progress made for all projects. These then get published in an internal report. We have been given access to this data for all project-years between FY 2003-2004 and FY 2012-2013.

Upon project completion, the IMED compiles an evaluation using the completion report handed in by the implementing agency and conclusions from their own visit of the project site. In this report, IMED officials record whether the works were completed, and a completion rate from 0-100%, although such completion rates are only available for about 1/3 of incomplete projects though. We were able to access and hand-code reports for projects evaluated between FY 2003-2004 and FY 2011-2012. Our data suggests that 77% of terminated projects get evaluated, most likely because the agency did not submit a completion report to the IMED. Table A1 compares the characteristics of evaluated and non-evaluated projects. Projects for which we couldn't find an evaluation report appear to be larger, and to involve more agencies and foreign financing on average. Within our sample if evaluated projects, it is still possible to make a valid comparison of complete and incomplete projects.

In order to build our study sample, we start by merging the ADP/RADP and IMED databases. We managed to match 93% of all projects in the IMED database to the ADP/RADP database to produce a project-year level panel covering 10 financial years from 2003-2004 to 2012-2013 and including 2493 infrastructure projects. This represents a total of 8285 project-year observations, for which we have data on planned and actual expenditures, as well as general project characteristics. We then merge in the evaluation reports data. This can only include projects that were concluded and evaluated in or before 2011-2012, which implies that the sample for which we have completion data is naturally censored. It counts 1218 unique projects (4119 project-year observations, or again 49% of all observations), and is our main sample for the study of successful project completion. Because the projects vary in duration, and are planned to start and end in different years, this is an unbalanced panel.

In the first set of Columns of Table 1, we summarize the main characteristics of all projects in our panel: Projects are planned to last 5 years on average, and to cost USD33.3million. Median total costs are only USD 5.9 million though, suggesting that the distribution is heavily skewed to the right, and includes few very large projects.

We also have information about project types and the agencies conducting these projects, which we summarize respectively in Tables 2 and 3[5]. We observe 17 unique project types. In Column 1 of Table 2, we report the frequency and relative proportion of each project type: The most frequent are transport (18%) - which covers road, railways and bridges construction and maintenance; physical planning, water supply and housing (16%) and agriculture (15%)[6]. As mentioned before, the projects in our sample were implemented by 207 agencies overall. When focusing only on the sample of evaluated projects as is the case in Table 3, this number goes down to 147. A large number of these agencies additionally implement only a small number of projects, and 60% of projects are conducted by the 20 largest agencies only. We provide details about these 20 agencies in Table 3. Column 1 reports the number of projects and fraction of total conducted by each organisation. The most common agencies in the sample of evaluated projects are the Department of Roads and Highways (13%), Bangladesh Water Development Board (8%) and the

---

[5] In Tables 2 and 3, the sample used is the sample of projects evaluated between FY 2003-4 and 2011-12.

[6] Similar information is presented in Appendix Figure 1, where we show the number of projects per type, representing by a different colour each of the main agencies we focus on later on.

Local Government Engineering Department (6%). Column 2 details the type of organisation they are. Recall that agencies can be of one of three types: Local government agencies, which overall conduct 61 (5%) projects; autonomous agencies responsible for 483 projects (40%); or departments of central ministries, conducting 674 (55%) projects. Finally, Column 3 reports which ministry the agencies report to, and Column 4 the type of projects they conduct.

V. Aggregate Project Performance: Measurement and Summary Statistics

V.1 Measurement

We now summarize how we construct our key variables. An important part of understanding the dynamics of project implementation though will be to keep in mind the information available at each point in time, and when additional information gets revealed to Project Directors. To try and make this clear, we introduce some simple notation before diving into project performance measures. Let *i* index projects, and *t* time periods. Time periods are one financial year. Project *i* initially gets planned at *t=0,* with information set I_0. Planners initially decide on expected duration $E[D_i|I_0] = \widehat{D_{\iota 0}}$, and expected total costs $E[C_i|I_0] = \widehat{C_{\iota 0}}$. Projects then get evaluated at T_i, when actual duration D_i and actual costs C_i are revealed.

We use three measures of aggregate performance: completion, aggregate deviation from planned duration, and aggregate deviation from planned total costs.

Project completion is defined as Y_i, and is equal to 1 if the completion rate reported during the evaluation is greater than or equal to 95%. This will be the main outcome of interest in this report.

Deviations from planned duration and total costs are respectively equal to $\frac{D_i - \widehat{D_{\iota 0}}}{\widehat{D_{\iota 0}}}$ and $\frac{C_i - \widehat{C_{\iota 0}}}{\widehat{C_{\iota 0}}}$. Constructing them in this way means that they are expressed as a fraction of the initial plan. A value of 0 would imply that a project remained perfectly in line with initial plans, while a value of 1 implies that actual duration or costs were twice as large as the plan (or 100% over time or cost). Similarly, negative values would imply that a project was shorter than planned, or underspent compared to its initial budget.

V.2 Summary Statistics

The first set of columns of Table 1 provides summary statistics on these three outcomes for the full sample of evaluated projects. 79% of projects are reported as complete, a figure in line with the 70% completion rate reported in Denizer *et al.* (2013). 23% of evaluated projects deviate from planned duration, lasting longer than planned in 83% of these cases. The average time overrun is of 1.5 years, which corresponds to a non-negligible 25% of mean planned duration. 77% of projects suffer from significant (over 5%) deviations from planned costs. Contrary to common results from studies of large infrastructure projects in developed countries (e.g. Flyvbjerg 2003), a large majority of cost deviations here (72%) are underspends[7].

To further understand the nature of these deviations from planned duration and costs, we start by presenting in Figure 1A a scatterplot representing each combination of cost and time deviations in our sample. We focus for now on projects which departed from both cost *and* time schedules, to better

---

[7] Note that we compute actual total costs by summing up actual expenditures in all implementation years. We therefore have to further censor our sample to only include projects that are fully covered by our sampling frame, i.e. start in or after 2003-4 and end before 2011-12. This explains the reduced sample sizes for cost deviations variables.

visualize the correlation between both types of deviation. We split the dots between complete and incomplete projects.

This plot makes clear that the large majority of deviations from budget are underspends, for both complete and incomplete projects, while time deviations are more evenly distributed. Adding lines of best fit to this plot to study the correlations between the two outcomes, we see that cost and time overrun seem to be significantly positively correlated (although the sample sizes here are greatly reduced). It is not surprising to see project costs increase (decrease) when a project gets delayed (shortened). These correlations are also in line with previous findings from the planning literature looking at the implementation of large infrastructure projects (Flyvbjerg *et al.* 2004).

In Figure 1B, we conduct a similar exercise, but this time plotting separately the distribution of time deviations for projects which finished on budget (left-hand panel), and cost deviations for projects which finished on time (right-hand panel). The left-hand panel confirms that time deviations are in majority time overruns, but the sample is otherwise too small to conclude anything about the full distribution. The right hand panel however shows a few interesting trends. First, even controlling for time under/overruns, we see significant deviations from planned budget, the majority of which are cost underruns. This suggests that underspending is not only linked to a project's duration being reduced and costs therefore getting cut down, but that other reasons remain to be explored. Second, looking more closely at the distributions, we see that most observations are concentrated around 0, but that a significant fraction of the sample spends less than 75% of total costs, implying cost underruns are not driven by outliers, for complete and incomplete projects alike. We also observe, which is typical of studies on infrastructure projects, a long tail to the right of the distribution, indicating that a few outliers suffered very large cost overruns of up to 300% of planned budget. Finally, the most striking feature of Figure 1B is the apparent difference between the distributions for complete and incomplete projects: incomplete projects reporting underspending a lot more frequently than complete projects.

To explore further apparent differences between complete and incomplete projects, we present the same summary statistics than in the first set of columns of Table 1, but this time separately for complete and incomplete projects, respectively in the second and third sets of columns of Table 1. The p-values from tests of mean and median equality between complete and incomplete projects are then reported in the fourth set of columns. Comparing first the characteristics of complete and incomplete projects, we observe that incomplete projects were planned to last longer (p-value = 0.008). There is no significant difference in the likelihood of time delay though (85% for incomplete projects vs. 83% for complete projects), nor in either mean or median planned budget. However, incomplete projects are more likely to report significant deviations from planned costs (p-value = 0.018), 87% of them reporting discrepancies between planned and actual total costs, vs. 75% of complete projects. Additionally, more incomplete projects (86%) than complete projects (69%) reported cost deviations to be cost underruns, significantly so (p-value = 0.003).

These results have two important implications: First, fund underutilization is associated with project incompletion, which suggests that understanding and preventing it could help improve completion rates. Second, even if less frequent than for incomplete projects, underspends still represent the large majority of cost deviations (69%) for complete projects. There therefore exists a large margin to improve fund allocation without necessarily hurting completion rates. Indeed, the figures reported in Table 1 indicate that better fund allocation could have saved on average USD3.2 million (median USD 0.47 million) per complete project reporting spending less than planned.

The main hypothesis that comes to mind when thinking about the reasons behind fund underutilization by bureaucracies is that low budgetary performance could be explained by a lack of capacity. At the project level, capacity includes the skills and infrastructure needed for planning, forecasting, or again day-to-day management by planners, directors and other staff. At the national level, it is the government's capacity to raise and release funds. This hypothesis is directly linked to the idea in much of the foreign aid literature which states that developing countries might have low absorptive capacity, therefore causing larger amounts of aid (and consequently larger number of projects conducted), to actually be detrimental to the average quality of project outcomes (Presbitero 2016). In this study, we dive deeper into the origins of fund underutilization and build dynamic spending performance measures to better understand when underutilization arises.

A second hypothesis is linked to the incentives project managers and heads of executing agencies operate under. Preliminary research revealed that unspent funds remain with the agency until the end of a project, and then return to the Ministry of Finance if still unspent. There are no opportunities for agencies to move these funds to other projects, and we therefore do not believe there to be any incentive for agencies to underspend. We could still think that there are reputational or career incentives for a project director to spend less than planned, but our research reveals quite the opposite. Indeed, there is in Bangladesh a lot of media and political attention around the fraction of planned development investments actually being spent, 'ADP utilization' being used as a metric for the quality of budget implementation in a particular year. This casts some doubt on whether there are strong institutional incentives to spend less than planned.

A third, frequently discussed hypothesis is linked to the idea of active waste, suggesting that project managers give in to corrupt practices for personal gain. We could imagine, following previous literature, that managers either report more expenses than were actually made, or perhaps again that some of the unspent funds get leaked during the implementation of a project and are never returned to the Ministry of Finance. These are possibilities that we will however not be able to test, but that would deserve attention from future research.

One way to start exploring these hypotheses would be to look into how agency and sector characteristics affect project performance, as much of the institutional capacity and incentives in place are determined at the agency level. In Tables 2 and 3, we present the same summary statistics as in Table 1, respectively by project type, and for the 20 largest agencies in terms of number of projects. In both tables, we observe some significant variation in completion rates, time and cost deviations from plan, suggesting that agency characteristics and project type indeed matter in determining the success rate of infrastructure projects.

Interestingly for our current focus, Tables 2 and 3 also show that the variation in project completion across agencies and types is such that no agency succeeds or fails all the time. This variation suggests that there are project-level factors that affect success rates so that within an agency or sector, some projects get completed and others not. We turn our attention to these factors.

VI. Explaining Project Completion

VI.1 Time-invariant correlates of project completion

In order to study the correlates of project completion, we estimate the following linear probability model using the sample of evaluated projects:

$$Y_i = \alpha + \beta_1.\widehat{D_\iota} + \beta_2.\widehat{C_\iota} + \beta_3.Donor + \beta_4.FracDonor + \beta_5.Local + \beta_6.Autonomous + \beta_7.NRevisions + u_i$$

where Y_i, a dummy equal to 1 if a project is assessed as complete, is regressed first on $\widehat{D_\iota}$ and $\widehat{C_\iota}$, respectively the planned duration and total costs. We then add in variables capturing the type of financing arrangement: $Donor$ is an indicator equal to 1 if there is any participation of foreign funds, and $FracDonor$ ranges from 0 to 1 and is equal to the fraction of total planned costs that is financed by foreign donors. $Local$ and $Autonomous$ are two dummies identifying projects conducted respectively by local governments and autonomous public agencies. Coefficient estimates here are to be compared to projects conducted by departments of central ministries, which is the excluded category. Finally, we add $NRevisions$, a variable corresponding to the total number of revisions to total costs a project has undergone over its lifetime.

Projects in our sample have planned start dates ranging from 1970 to 2012. In order to avoid time effects linked to the fact that projects are conducted in different political and economic contexts, we control throughout for the year in which a project was planned to start. We also run additional versions of the model that include full sets of agency and project type fixed effects, and tests for joint significance of these sets[8]. Standard errors are corrected for heteroscedasticity throughout.

The results from this estimation are presented in Table 4. In Column 1, we include only variables corresponding to a project's initial time and cost schedule. The measured effects are both very small, and not significantly different from 0. Variables for financing arrangements, added in Column 2, do not have any significant effect on project completion either. We then add dummies for agency types in Column 3. It appears that projects conducted by autonomous agencies are more likely to be completed than projects conducted by central government agencies – this result is only marginally significant though. In Columns 4 and 5, we include project type and executing agency fixed effects. These are both consistently jointly significant, confirming the variation in completion rates we observed in Tables 2 and 3. Finally, in column 6), we add the number of times total costs were revised during a project's lifetime. In line with results in Denizer *et al.* (2013), we find that more revisions are significantly correlated with a greater likelihood of success. This suggests that learning along the life of a project will be important, and more generally implies that the dynamics of project implementation matter for project completion. We therefore focus on now describing the links between dynamic implementation and aggregate performance.

VI.2 Dynamic Performance

We build innovative measures of dynamic project performance, allowing us to assess how a project is doing at each point in its implementation. To build these measures we need to compare c_{it}, project *i*'s actual expenditures in financial year *t,* to some benchmark for planned expenditures. We do not observe in the data the year-wise spending schedule at *t=0* though*,* and have to construct these benchmarks. We detail below the two main types of benchmarks we use: (a) linear expenditure plans, and (b) projections from detailed data.

Linear expenditure plans

We assume first that the relevant benchmark corresponds to perfectly smoothed intertemporal spending, where the fraction of total costs to be spent each year is constant throughout the life of a project. We use linear projections from the initially planned costs and duration to construct a planned expenditures schedule. We construct expected expenditures for project *i* in year *t,* in the following way:

$$E[c_{it}|I_0] = \frac{E[C_i|I_0]}{E[D_i|I_0]} = \frac{\widehat{C_{\iota 0}}}{\widehat{D_{\iota 0}}}$$

These planned expenditures are then set to 0 past initially planned duration $\widehat{D_{\iota 0}}$. Because projects vary in length, we also normalize the project life cycle and use project stage instead of financial years: Project stage is expressed in % of project duration, and equal to $\frac{t}{\widehat{D_{\iota 0}}}$. We present in Panel A of Figure 2 a stylized representation of the way this benchmark expenditure plan is built.

---

[8] Note that we cannot simultaneously include agency fixed effects and agency type dummies. We therefore exclude those in Columns 4 and 5 of Table 4.

This first benchmark assumes that the only information available to the Project Director is the information set I_0, and implicitly that the right objectives to focus on are the initially planned duration and costs. It could be however that a Project Director fully anticipates that for example a project will need to run over time, or over budget, in order to complete its objectives. To allow for this, we build three additional benchmarks where (i) time deviations are fully anticipated, (ii) costs deviations are fully anticipated, and (iii) both time and costs deviations are fully anticipated.

These new benchmarks are built in a similar fashion to the first one, but respectively take into account the (i) true duration, (ii) true total costs, and (iii) true total costs and duration instead of the initial plan. Focusing on the benchmark allowing for time overruns for illustration purposes, we construct the new expenditures plan as follows:

$$E[c_{it}|I_0, D_i] = \frac{E[C_i|I_0]}{D_i} = \frac{\widehat{C_{i0}}}{D_i}$$

A similar method is then used to construct the remaining two benchmarks $E[c_{it}|I_0, C_i]$ and $E[c_{it}|I_0, C_i, D_i]$. Illustrations of this method are provided in panels B, C and D of Figure 2.

We then finally construct performance measures by comparing actual and planned expenditures in each time period. For each benchmark, we compute:

$$\frac{c_{it} - E[c_{it}|.]}{E[c_{it}|.]}$$

We use four benchmarks corresponding to four different information sets, and therefore obtain four performance measures. These measures are built in such a way that they are expressed as a fraction of planned expenditures. A value of 0 implies that a project spent according to plan, while negative (positive) values imply that actual expenditures were below (above) plan.

We plot these four performance measures by project stage in Figure 3, separately for complete and incomplete projects. The first important observation is that for all benchmarks projects tend to underspend early in their life cycle, and overspend later on. This is true for complete and incomplete projects alike, and we can thus reject the hypothesis that expenditures should be linear over time. This is true irrespective of which information set we actually use.

We also report on these figures the difference between the average performance of complete and incomplete projects at each point in time. Results from tests of significance of this difference are represented by drop lines extending to two standard errors above and below the difference estimate. These reveal that completed projects underspend less than incomplete projects early in their life cycle, and then overspend more later on. The difference between complete and incomplete projects are significantly different from 0 in later halves of projects in panels A and B. However, as expected the gap between complete and incomplete projects tightens and differences are no longer significant when looking at panels C and D, which allow for cost overruns.

Projections using detailed data

The previous section showed that even successful projects do not follow a linear spending pattern; we therefore use an additional source of data to construct alternative benchmarks and improve on the linear assumption above. Along the life of a project, executing agencies have to put together several key documents that provide valuable, detailed information about projects. The first of these documents, the Development Project Proforma (DPP), is prepared at the planning stage (t=0) to obtain approval from the Government. It provides the initial year-wise expenditure plan we were missing before – which we call $E[c_{it}|I_0, DPP]$. The DPP also includes a wealth of information that will be of interest for future research, such as for example planned spending from donors vs. government funds, project location, planned procurement and planned construction.

If a project undergoes significant revisions during the its life-cycle, the Project Director has to prepare and submit a second document, the Revised DPP (RDPP). This includes an updated year-wise expenditure plan that takes into account the information set available at the time of revision, I_t, and thus corresponds to $E[c_{it}|I_t, RDPP]$. The RDPP additionally includes a descriptive of the reasons for revisions, which again should be the focus of future qualitative research.

Finally, at the end of the project, the Project Director has to submit a Project Completion Report (PCR) to the IMED. This report includes information of immediate interest: A summary of initially planned year-wise expenditures, and final actual expenditures by year. It also includes information to be later used to further understand the reasons behind spending dynamics, such as a summary of revisions, contract dates and values for procurement, dates of visits and audits, a summary of problems faced during implementation, and finally information about the identity, tenure dates, and overall turnover of managers.

We have found, hand-coded and entered full sets of DPP-RDPP-PCR for a subset of 332 projects (2080 project-years). We therefore have full paper trails on these projects' implementation history. When looking for paper trails, we focused on agencies that implement the most projects overall so as to be able to control for potential agency effects, and ended up choosing 11 agencies, the names of which are highlighted in Table 3. We manage to match 168 projects in this subsample to the main sample of 1218 evaluated projects. In Table A2, we report the results from tests of balance between matched and unmatched projects along the main dimensions considered in analysis. Projects for which we found paper trails started significantly more recently than the ones we couldn't match, and executing agency and project sector dummies are jointly significant. None of this is surprising seeing that paper documentation tends to be destroyed by the Government of Bangladesh after five years, and that we only focused on a subset of agencies. Other than these three characteristics, matched and unmatched projects are statistically similar along all other dimensions; we can therefore consider them a representative sample of the larger sample of evaluated projects.

To explore building benchmarks using this detailed data, we start by plotting in Figure 4A the initial year-wise expenditure plan $E[c_{it}|I_0, DPP]$ by project stage, separately for complete and incomplete projects. The sample used here is the sample of 168 projects we could match to the sample of 1218 evaluated projects. We find some new evidence of the pattern we had previously observed, projects planning to spend below the 45° line early in their life cycle and over the 45° line later on. In Figure 4B, we then look for any evidence of learning along the life of a project, and plot the difference between revised and original planned expenditures in each period, again separately for complete and incomplete projects. The variable plotted here is:

$$\frac{E[c_{it}|I_t, RDPP] - E[c_{it}|I_0, DPP]}{E[c_{it}|I_0, DPP]}$$

so that it is expressed as fraction of originally planned expenditures. As before, we also plot the difference between complete and incomplete projects, with results of tests of significance indicated by drop lines at -/+ 2 standard errors. We see some evidence of learning along the life of a project, revised planned costs varying significantly from original costs. Additionally, the direction of revisions seems to be such that it accentuates the pre-existing pattern of spending below the 45° line early on, and more later in the life of a project. Finally, comparing complete and incomplete projects shows that complete projects tend to make revisions of lower magnitude throughout, and significantly so in the second half of a project's life. This could suggest either better planning or implementation, and overall confirms that dynamics seem to be linked to a project's likelihood of success.

We use this detailed data to build two alternative benchmarks for planned spendings, using respectively the initial and revised planned expenditures of projects that we know to have been completed[9].

We illustrate our method using initially planned spendings: We plot first in Figure 5 the average initially planned spendings $E[c_{it}|I_0, DPP]$ for completed projects in the paper trails subsample, which we compare to the average actual spendings c_{it} of complete and incomplete projects in the larger IMED database of evaluated projects. For comparability purposes, all three measures are plotted by project stage (expressed as % of originally planned duration) and expressed as % of original total costs. This implies we only represent this data up to the initial planned completion date ($\widehat{D_{\iota 0}}$), even though expenditures might actually have gone on beyond that point.

To build a new benchmark, we use the average initially planned spendings from the paper trails data (the green line in Figure 5) to project the fraction of total costs a project would plan to spend at each point in time. Multiplying this projected fraction by true original total costs for each project in the full IMED data, we can then obtain a measure of planned spendings $E[c_{it}|I_0, DPP]$ for all the projects in our sample. We use the same projection method to obtain revised expenditures plans $E[c_{it}|I_t, RDPP]$ for all projects.

Finally, we build two new performance measures, of the same form as our previous measures, comparing true expenditures each year to the projected benchmarks we constructed using detailed data ; these are respectively equal to $\frac{c_{it}-E[c_{it}|I_0,DPP]}{E[c_{it}|I_0,DPP]}$ and $\frac{c_{it}-E[c_{it}|I_t,RDPP]}{E[c_{it}|I_t,RDPP]}$. As before, they correspond to the departure from planned expenditures at each project stage, expressed as a fraction of planned spendings for that period.

These are plotted by project stage in Figure 6. Panel A shows that over the original project duration, actual spendings are consistently below initially planned spendings. This is true for both complete and incomplete projects; incomplete projects moreover appear to underspend more than complete projects, significantly so in the latter half of a project's life cycle. This suggests that underspending in incomplete projects is not simply the result of a project being abandoned at the end and spendings being suddenly interrupted, but rather the result of inefficiencies building up throughout the life of a project.

Panel B compares actual spendings to projected revised planned spendings. As a result of our projection method, the performance measures are too noisy to allow straightforward conclusions in the early life of a project[10]. In the second half of a project's life however, we observe once more that incomplete projects appear to underspend significantly more than complete projects. Additionally, this plot also shows that completed projects spend much closer to plan once this plan has been revised. This provides some evidence that the managers of successful projects learn and adjust expectations along the life of a project.

---

[9] When conducting this exercise, we focus on the full sample of projects in paper trails data in order to maximize sample size. We therefore build these benchmarks using data on 190 completed projects.

[10] A lot of projects saw delayed starts in our paper trails subsample, which drove the average revised spending towards 0 in early project phases (as can be seen in Figure 4). Because of the projection method we use, and of the way we build our performance measures, projects from the IMED database having positive spending in these early phases create large outlier values in the performance measures plotted in Figure 6 – hence the noise in early project phases in panel B.

VII. Summary and discussion

Previous research has showed that infrastructure projects are fundamental for growth and development, but suffer from low completion rates. While these projects are a focus of interest for politicians, the media, and international donors, little is in fact know about the determinants of successful completion. We believe it important for more research to open the black box of public project implementation. As a first step, we offered here an initial look at the dynamics involved in this process, focusing on a developing country context.

We built a representative sample of over 1200 infrastructure projects conducted by the Government of Bangladesh between 2003 and 2013, for which we have data on project characteristics, end-of-project outcomes, and planned and actual expenses along the life of the project. 79% of these projects were assessed to be successfully completed, which is in line with previous literature. 23% of projects departed from their planned duration, and 77% departed from their planned budget. These magnitudes are also in line with results from the planning literature studying infrastructure projects in developed countries. We found though that 72% of deviations from planned costs were actually underspends, whereas projects in developed countries seem to suffer mainly from cost overruns (Flyvbjerg 2003). Interestingly, it also appears that incomplete projects are more likely to suffer larger underspends than complete projects.

To better understand these aggregate results, we started by correlating them with project characteristics: Completion rates in our sample could not be explained by project size, duration, or funding arrangements, and only marginally by agency type. Project sectors and individual executing agencies do show significantly different completion rates, but our data at this point does not allow us to offer any formal explanation as to why, and would need to be completed by further research into actual practices at executing agencies. Most interestingly, these correlations showed that projects which got revised were more likely to be successfully completed. This indicates that implementation dynamics are potentially important determinants of project completion.

In perhaps the main innovation of this initial research, we presented the first descriptive statistics of within-project spending dynamics, and asked whether these could help explain completion rates. The main challenge when studying project dynamics is to choose a benchmark to compare actual spendings to. We therefore built and discussed several benchmarks, using first simple linear expenditure plans, and second projected planned spendings using detailed data we collected on a subsample of projects. All the performance variables we built show similar trends, and allow us to draw two main preliminary conclusions: First, projects seem to follow a non-linear spending pattern, spending less in early stages of a project's life, and more in the latter half of a project's implementation period. This is true for both complete and incomplete projects, and shows that underspends do not appear only because projects get abandoned, but rather seem to be an issue arising early in a project's life and surviving throughout. This suggests investigating further the planning and early life of projects. Second, when comparing complete and incomplete projects, we observe that successfully completed projects overall did better at predicting the non-linear spending trend, required smaller revisions to planned spendings, and departed less from plan throughout the life of a project. Implementation dynamics and a project's completion status are therefore indeed correlated, and this relation should be explored further.

The literature on public good provision discusses two broad types of reasons for ineffective delivery. First, personal, institutional or political incentives could explain the low performance of public projects in general, and cost deviations in particular. We cannot test this hypothesis formally, but our initial research and conversations with government officials suggest that there are no benefits in underspending or under-delivering, and we therefore believe future research should focus on the second broad hypothesis. This second explanation is linked to the idea that government agencies lack the capacity to successfully implement such projects. There are many aspects of project implementation that require the government's action though (planning, forecasting, raising revenues, budget planning, procurement, hiring, management, etc.), and it is not possible, with the data we currently have, to identify where the capacity gaps are. Future research looking to understand the determinants of successful project completion should focus first on identifying when and why things go off track in a project's life.

References

Addison, D. 2013. "The Quality of Budget Execution and Its Correlates." *Policy Research Working Paper.*

Aschauer, D. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23: 177-200.

Bandiera, O., A. Prat, and T. Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review* 99 (4): 1278-1308.

Banerjee, A., L. Iyer, and R. Somanathan. 2008. "Public Action for Public Goods." *Handbook of Development Economics* Ed. 4 (Ch. 49).

Banerjee, A.V., E. Duflo, and N. Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." *NBER WP 17897.*

Besley, T., and R. Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *The Quarterly Journal of Economics* 117 (4): 1415-1451.

Besley, T., and T. Persson. 2010. "State Capacity, Conflict, and Development." *Econometrica* 78: 1–34.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padro I Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review* 105 (6): 1817-1851.

Byron, R.K. 2016. "Only half of ADP money spent in 10 months." *Daily Star*, May 17.

Denizer, C., D. Kaufmann, and A. Kraay. 2013. "Good countries or good projects? Macro and micro correlates of World Bank project performance." *Journal of Development Economics* 105: 288-302.

Donaldson, D. 2016. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* forthcoming.

Duflo, E., and R. Pande. 2007. "Dams." *Quarterly Journal of Economics* 122: 601-646.

ECOTEC Research and Consulting. 2003. "Ex-post Evaluation of Objective 1 (1994-1999)."

Ferraz, C., and F. Finan. 2011. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." *American Economic Review* 101: 1274-1311.

Finan, F., and M. Mazzocco. 2016. "Electoral Incentives and the Allocation of Public Funds." *IZA Discussion Paper No. 9623.*

Finan, F., B. Olken, and R. Pande. 2015. "The Personnel Economics of the State." *NBER Working Papers.*

Financial Express. 2016. "July sees only 0.56pc ADP implementation." *Financial Express*, September 4.

IMF. 2015. "Bangladesh Progress Report - Poverty Reduction Strategy." *IMF Country Reports* (15/305).

Kaufmann, D., and Y. Wang. 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.

Lewis, G., and P. Bajari. 2011. "Procurement Contracting with Time Incentives: Theory and Evidence." *Quarterly Journal of Economics* 126 (3): 1173-1211.

Luo, R., G. Miller, S. Rozelle, S. Sylvia, and M. Vera-Hernandez. 2015. "Can Bureaucrats Really Be Paid Like CEOs? Schoold Administrator Incentives For Anemia Reduction In Rural China." *NBER Working Paper 21302.*

Mitra, A., A. Varoudakis, and M. Vegnazones-Varoudkis. 2002. "Productivity aand Technical Efficiency in Indian States' Manufacturing: The Role of Infrastructure." *Economic Development and Cultural Change* 50: 395-426.

Muralidharan, K., and V. Sundararaman. 2011. "Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39-77.

Muralidharan, K., J. Das, A. Holla, and A. Mohpal. 2016. "The Fiscal Cost of Weak Governance: Evience from Teacher Absence in India." *Policy Research Working Papers.*

Olken, B. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200-249.

Presbitero, A.F. 2016. "Too much too fast? Public investment scaling-up and absorptive capacity." *Journal of Development Economics* 120: 17-31.

Rasul, I., and D. Rogger. 2016. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal* Forthcoming.

Reinikka, R., and J. Svensson. 2004. "Local Capture: Evidence From A Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics* 119 (2).

Simson, R., and B. Welham. 2014. "Incredible Budgets: Budget credibility in theory and practice." *ODI Working Papers.*

Williams, M. 2016. "The Political Economy of Unfinished Development Projects: Corruption, Clientelism, or Collective Choice?" *Working Paper.*

Williams, M.J. 2015. "One size does not fit all: Budget Institutions and Performance in Ghana's Spending Ministries." *Working Paper.*

Table 1: Project Characteristics and Aggregate Performance Measures

Sample: IMED Database, Project data from 2003-2004 to 2012-2013

	(1) Full Sample				(2) Evaluated and Completed Projects				(3) Evaluated and Incomplete Projects				(4) P-value	
	N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median	Mean equality	Median equality
A. Project Completion														
Fraction of Projects Evaluated	2493	0.49												
Fraction of Evaluated Projects Reported as Completed	1218	0.79												
Completion Rate, Given Project Reported as Incomplete	115	0.77	0.22	0.86										
B. Project Duration														
Initial Expected Duration	2493	5.38	2.85	5.00	961	5.69	2.95	5.00	257	6.29	3.71	6.00	0.007	0.000
Actual Duration (for Evaluated Projects)	1218	6.06	3.20	5.00	961	5.91	3.02	5.00	257	6.60	3.74	6.00	0.002	0.000
Fraction of Projects with Discrepancies in Duration	1218	0.23			961	0.22			257	0.27			0.058	
Of Those, Share of Projects with Time Overrun	278	0.83			208	0.83			70	0.86			0.558	
Time Overrun (if Difference >0)	232	1.50	1.00	1.00	172	1.49	1.01	1.00	60	1.52	0.97	1.00	0.850	
Time Underrun (if Difference <0)	46	-1.28	0.75	-1.00	36	-1.33	0.83	-1.00	10	-1.10	0.32	-1.00	0.390	
C. Project Cost														
Initial Expected Total Cost (Million US$)	2493	33.68	131.93	5.92	961	20.59	55.31	3.87	257	22.96	66.82	3.61	0.561	0.639
Actual Total Cost (for Evaluated Projects; Million US$)	452	10.68	54.74	2.58	373	8.97	26.03	2.76	79	18.75	118.38	1.92	0.150	0.046
Fraction of Projects with Significant Discrepancies in Costs	452	0.77			373	0.75			79	0.87			0.018	
Of Those, Share of Projects with Cost Overrun	349	0.28			208	0.31			69	0.14			0.007	
Cost Overrun (if Difference >0)	96	4.45	23.56	0.50	86	2.27	7.31	0.49	10	23.18	70.15	1.01	0.007	
Cost Underrun (if Difference <0)	253	-3.45	9.13	-0.49	194	-3.23	7.76	-0.47	59	-4.16	12.70	-0.57	0.496	

Notes: This table presents summary statistics on aggregate performance measures. The sample used here includes all projects conducted between FY 2003-4 and 2012-13. Evaluated Projects correspond to projects for which we found an evaluation or completion report: This means that a completion report was submitted and the project terminated. It might however not report 100% completion: Projects are assumed Completed if reporting over 95% completion, and incomplete otherwise. Information about actual completion rates of projects reported as incomplete in available for only about 1/3 of incomplete projects. For all financial variables, amounts are presented in US $, converted using the average exchange rate of the corresponding year. Actual total costs are computed by adding all observable expenditures. We therefore need to restrict this sample to projects which start within our sampling frame, hence the reduced sample size. We allow a margin of +/- 5% when creating dummies for significant discrepancies in total costs. We report the p-values for tests of mean and median equality between complete and incomplete projects in the last 2 columns. Median equality tests are conducted using quantile regressions. Finally, figures are rounded up to 2 decimals when applicable in all columns but the last two columns, where they are rounded up to 3 decimals.

Table 2: Descriptive evidence on project types

Sample: IMED Database, Projects evaluated between FY 2003-04 and 2011-12

	(1) Number of Projects [Proportion]		(2) Number of Implementing Agencies	(3) Initial Expected Duration (Years)	(4) Initial Expected Total Cost (Million US$)	(5) % of Projects Completed	(6) % with Time Deviations	(7) % of which are Time Overruns	(8) Average Time Overrun (if >0)	(9) % with Cost Deviations	(10) % of which are Underspends	(11) Average Absolute Undespend (if >0)
All Project Types	1218	[1]	147	5.82	21.09	0.79	0.23	0.83	1.50	0.77	0.72	3.45
Transport	224	[0.18]	11	6.36	19.84	0.76	0.19	0.71	1.53	0.87	0.73	2.07
Physical Planning, Water Supply & Housing	196	[0.16]	34	5.49	16.64	0.84	0.27	0.83	1.39	0.73	0.78	4.93
Agriculture	187	[0.15]	27	5.28	7.76	0.80	0.17	0.81	1.28	0.76	0.72	2.73
Education & Religious Affairs	105	[0.09]	13	5.49	39.12	0.75	0.32	0.88	1.90	0.65	0.64	1.77
Water Resources	94	[0.08]	7	5.04	10.97	0.76	0.27	0.92	1.35	0.85	0.83	2.00
Rural Development & Rural Institutions	72	[0.06]	10	6.50	46.29	0.89	0.11	0.63	1.20	0.70	0.63	12.58
Power	56	[0.05]	7	8.29	80.02	0.80	0.18	1.00	1.50	1.00	0.75	17.08
Social Welfare, Women Affairs & Youth	55	[0.05]	6	5.16	3.63	0.78	0.15	0.88	1.14	0.76	0.69	0.51
Sports & Culture	55	[0.05]	14	5.40	2.80	0.82	0.44	0.92	1.55	0.53	0.56	0.24
Science, ICT	40	[0.03]	9	6.30	3.10	0.73	0.20	0.75	1.50	0.80	0.50	0.43
Communication	30	[0.02]	3	5.90	47.61	0.73	0.30	0.89	2.50	0.67	1.00	0.92
Industries	27	[0.02]	9	6.67	13.73	0.74	0.19	0.80	1.00	0.88	0.71	1.45
Oil, Gas & Natural Ressources	25	[0.02]	2	6.12	34.73	0.96	0.20	0.80	1.00	0.71	1.00	3.63
Public Administration	19	[0.02]	7	5.68	6.68	0.79	0.26	1.00	1.60	0.86	0.33	0.45
Health, Population & Family	16	[0.13]	7	4.56	12.07	0.50	0.50	0.75	1.50	0.78	0.71	3.04
Labour & Employment	9	[0.01]	1	6.11	9.17	0.67	0.00	-	-	0.67	1.00	0.17
Mass Media	8	[0.01]	4	4.75	3.92	0.88	0.38	1.00	1.00	1.00	0.50	0.17

Notes: The "project type" classification refers to the primary sector for each project. Project types highlighted in red correspond to the types included in the paper trail subsample. The sample used here is the sample of projects evaluated between FY2003-4 and 2011-12. For all financial variables, amounts are presented in Million US $, converted using the average exchange rate of the corresponding year. All time variables are reported in years. "% of Projects Completed" corresponds to an indicator = 1 if the project was reported as fully completed or with a completion rate >95% in its evaluation report. Time and cost deviations are indicators equal to 1 if the actual duration, or total costs, differ respectively from the planned duration or budget. Figures are rounded to two decimal places where relevant.

Table 3: Descriptive Evidence on Main Implementing Agencies

Sample: IMED Database, Projects evaluated between FY 2003-04 and 2011-12

	(1) Number of Projects [Proportion]		(2) Type of structure	(3) Reports to	(4) Types of projects conducted	Summary statitstics on projects conducted								
						(6) Initial Expected Duration (Years)	(7) Initial Expected Total Cost (Million US$)	(8) % of Projects Completed	(9) % with Time Deviations	(10) % of which are Time Overruns	(11) Average Time Overrun (if >0)	(12) % with Cost Deviations	(13) % of which are Underspends	(14) Average Absolute Undespend (if >0)
All Agencies	1218	[1]				5.82	21.09	0.79	0.23	0.83	1.50	0.77	0.72	3.45
Department Of Roads And Highways	162	[0.13]	Ministry department	Ministry of Roads, Transport and bridges	Transport	6.75	17.75	0.75	0.20	0.75	1.63	0.91	0.74	2.06
Bangladesh Water Development Board	98	[0.08]	Autonomous public agency	Ministry of Water Ressources	Agriculture; Water Ressources	5.14	12.22	0.76	0.28	0.93	1.40	0.86	0.79	2.92
Local Government Engineering Department	68	[0.06]	Ministry department	Ministry of Local Government, Rural Development and Co-operatives (Local Government Division)	Physical Planning; Transport; Water Ressources; Agriculture; Rural Development	6.34	48.56	0.91	0.15	0.60	1.33	0.69	0.67	6.98
Bangladesh Railway	35	[0.03]	Ministry department	Ministry of Railways	Transport	6.06	33.79	0.80	0.17	0.67	1.00	0.75	0.75	3.25
Dhaka City Corporation	35	[0.03]	Local Government	Ministry of Local Government, Rural Development and Co-operatives (Local Government Division)	Physical Planning; Education and Religious Affairs	3.77	12.60	0.86	0.29	0.90	1.00	0.53	1.00	6.17
Department Of Forests	29	[0.02]	Ministry department	Ministry of Environment and Forest	Agriculture	5.03	4.93	0.76	0.14	0.50	1.50	0.92	0.75	0.63
Public Health Engineering Directorate	27	[0.02]	Ministry department	Ministry of Local Government, Rural Development and Co-operatives (Local Government Division)	Physical Planning	6.89	27.69	0.70	0.15	0.75	1.00	1.00	0.86	11.45
Department Of Social Services	26	[0.02]	Ministry department	Ministry of Social Welfare	Health, Nutrition, Population & Family; Social Welfare	4.65	3.56	0.77	0.12	0.67	1.00	0.80	0.83	0.54
Bangladesh Agriculture Development Corporation	25	[0.02]	Autonomous public agency	Ministry of Agriculture	Agriculture; Water Ressources	4.56	7.39	0.64	0.16	0.75	1.00	0.83	0.70	3.64
Bangladesh Oil, Gas And Mineral Corporation	24	[0.02]	Autonomous public agency	Ministry of Power, Energy and Mineral Ressources	Oil, Gas, Natural Ressources	6.21	36.12	0.96	0.21	0.80	1.00	0.71	1.00	3.63
University Grants Commission	24	[0.02]	Autonomous public agency	Ministry of Education	Education and Religious Affairs	4.54	9.12	0.88	0.50	1.00	2.00	0.67	0.50	0.47
Bangladesh Power Development Board	21	[0.02]	Autonomous public agency	Power Division	Power	10.52	87.63	0.76	0.19	1.00	1.00	1.00	1.00	3.49
Department Of Agricultural Extension	21	[0.02]	Ministry department	Ministry of Agriculture	Agriculture	5.14	3.57	0.76	0.00	-	-	0.75	0.78	0.36
Bangladesh Atomic Energy Commission	20	[0.02]	Autonomous public agency	Ministry of Science and Technology	Science, ICT	6.75	3.37	0.80	0.30	0.83	1.60	0.86	0.50	0.45
Department Of Fisheries	20	[0.02]	Ministry department	Ministry of Fisheries and Animal Resources	Agriculture	5.95	6.91	0.90	0.05	1.00	1.00	0.67	0.75	1.23
Bangladesh Agricultural Research Institute	18	[0.01]	Autonomous public agency	Ministry of Agriculture	Agriculture	5.17	5.02	0.89	0.11	1.00	1.50	0.45	0.60	0.35
Department Of Women Affairs	18	[0.01]	Ministry department	Ministry of Women and children Affairs	Social Welfare	5.83	4.85	0.67	0.22	1.00	1.25	0.33	0.00	-
Rural Electrification Board	18	[0.01]	Autonomous public agency	Ministry of Power, Energy and Mineral Ressources	Power	7.39	84.84	0.72	0.06	1.00	1.00	1.00	0.75	27.47
Department Of Secondary & Higher Education	17	[0.01]	Ministry department	Ministry of Education	Education and Religious Affairs	6.71	85.45	0.47	0.29	1.00	2.60	0.67	0.50	15.08
Department Of Technical Education	17	[0.01]	Ministry department	Ministry of Education	Education and Religious Affairs	6.35	13.56	0.65	0.24	0.75	1.33	1.00	0.83	0.92

Notes: Agencies presented here are the 20 largest agencies in terms of numbers of projects. They conduct 60% of all evaluated projects. Agencies highlighted in red correspond to the types included in the paper trail subsample. The sample used here is the sample of projects evaluated between FY2003-4 and 2011-12. The "project type" classification refers to the primary sector for each project. For all financial variables, amounts are presented in Million US $, converted using the average exchange rate of the corresponding year. All time variables are reported in years. "% of Projects Completed" corresponds to an indicator = 1 if the project was reported as fully completed or with a completion rate >95% in its evaluation report. Time and cost deviations are indicators equal to 1 if the actual duration, or total costs, differ respectively from the planned duration or budget. Figures are rounded to two decimal places where relevant.

Table 4: Correlates of project completion

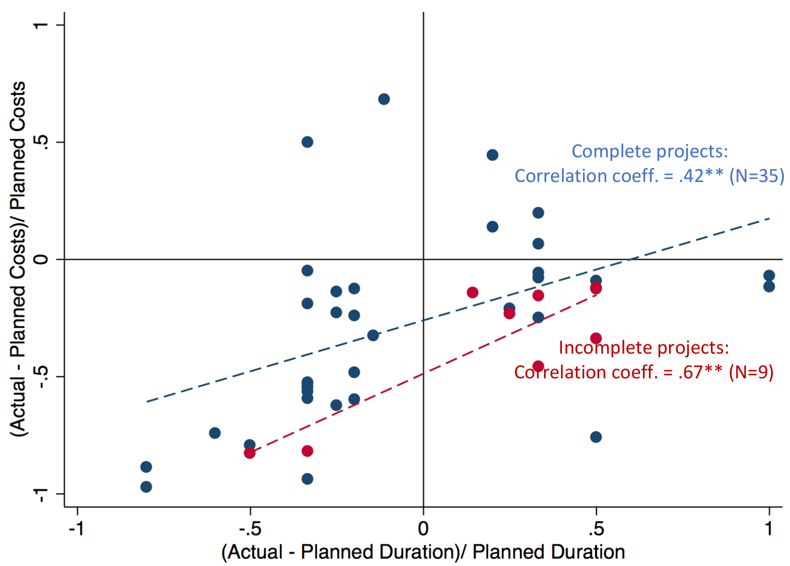
Sample: IMED Database, Projects evaluated between 2003-2004 to 2011-2012
Dependent Variable: Dummy =1 if Project is evaluated as completed
Linear Probability Estimates
Standard errors are corrected for heteroskedasticity

	(1) Schedule	(2) Financing	(3) Agency type	(4) Sector and agency FE	(5) Revisions
Planned Duration	0.003	0.002	0.004	0.004	-0.006
	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
Planned Total Costs (Million US$)	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Dummy = 1 if used donor money		-0.026	-0.035	-0.107	-0.123
		(0.093)	(0.093)	(0.105)	(0.104)
Donor Money/Total Costs		0.089	0.105	0.142	0.155
		(0.125)	(0.126)	(0.139)	(0.137)
Dummy = 1 if multiple agencies involved			0.121	0.098	0.083
			(0.099)	(0.107)	(0.108)
Dummy = 1 if Local Government Agency			0.032		
			(0.053)		
Dummy = 1 if Autonomous Public Agency			0.042 *		
			(0.025)		
Number of Revisions to Total Costs					0.056 ***
					(0.015)
Planned Start Year fixed-effects [p-value]	Yes [0.000]	Yes [0.000]	Yes [0.002]	Yes [0.067]	Yes [0.005]
Project Type fixed-effects [p-value]	No	No	No	Yes [0.000]	Yes [0.000]
Executing Agency fixed-effects [p-value]	No	No	No	Yes [0.000]	Yes [0.000]
Observations	1218	1218	1218	1218	1218

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors are in parentheses. All columns report coefficient estimates from a Linear Probability Model estimation. The dependent variable is an indicator equal to 1 if the project was successfully completed. The sample covers all projects evaluated between FY 2003-4 and 2011-12. For the vectors of Year, Project Type and Agency fixed effects, we report only the p-value from a test of joint significance. The variable for Number of Revisions counts the total number of revisions to total costs along a project's life. Figures are rounded to three decimal places.

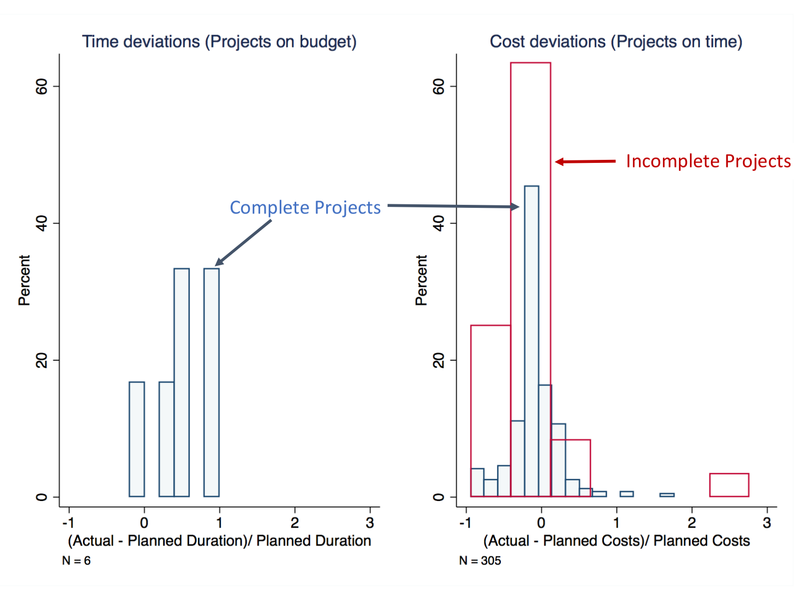
Figure 1: Distribution of aggregate performance outcomes

A. Compared distributions of time and cost deviations, by completion status



Notes: This scatterplot represents the correlation between time and cost deviations, separately for completed and incomplete projects. Dashed lines represent the best linear fit, again separately for completed and incomplete projects. The sample considered here is the sample of evaluated projects for which we have data on planned and actual total costs, and who reported non-zero time **and** cost deviations, hence the reduced sample sizes. All outcomes are expressed as a fraction of planned duration or costs respectively; 0 therefore corresponds to no deviation from plan, while a value of 1 would mean a project lasted/spent 100% more than planned.

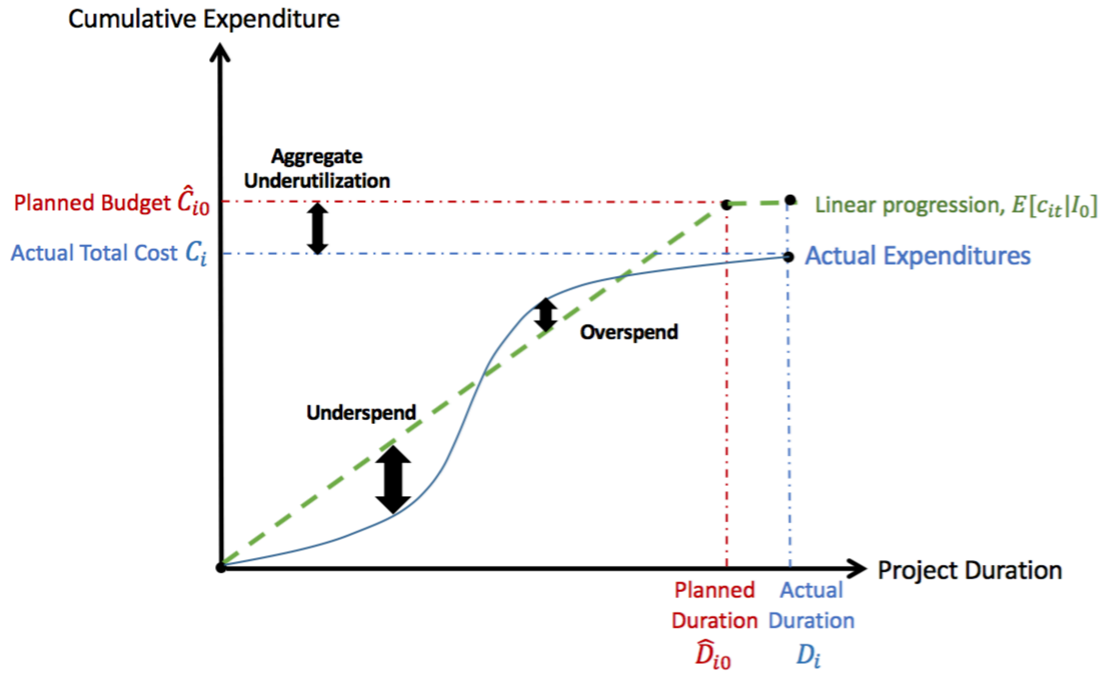
B. Distributions of time and cost deviations for projects completed on time/on budget



Notes: This figure plots the distribution of time and cost deviations, separately for completed and incomplete projects, focusing respectively on time deviations for projects finished on budget (left-hand panel) and cost deviations for projects finished on time (right-hand panel). Note that no incomplete projects were finished on budget, hence their absence in the left-hand plot. All outcomes are expressed as a fraction of planned duration or costs respectively; 0 therefore corresponds to no deviation from plan, while a value of 1 would mean a project lasted/spent 100% more than planned. The sample considered here is the sample of evaluated projects for which we have data on planned and actual total costs, excluding projects reporting non-zero time and cost deviations. The sample sizes are reported under each panel.

Figure 2: Building dynamic performance measures - Stylized description

A. Based on original costs and duration



B. Based on original costs and actual duration

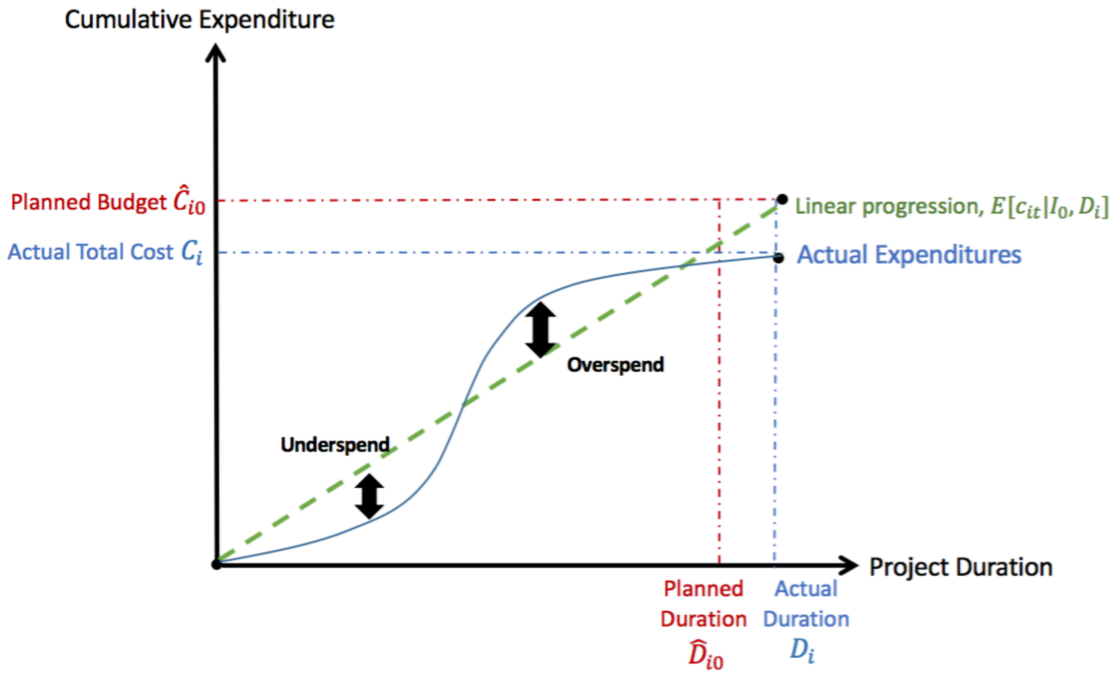
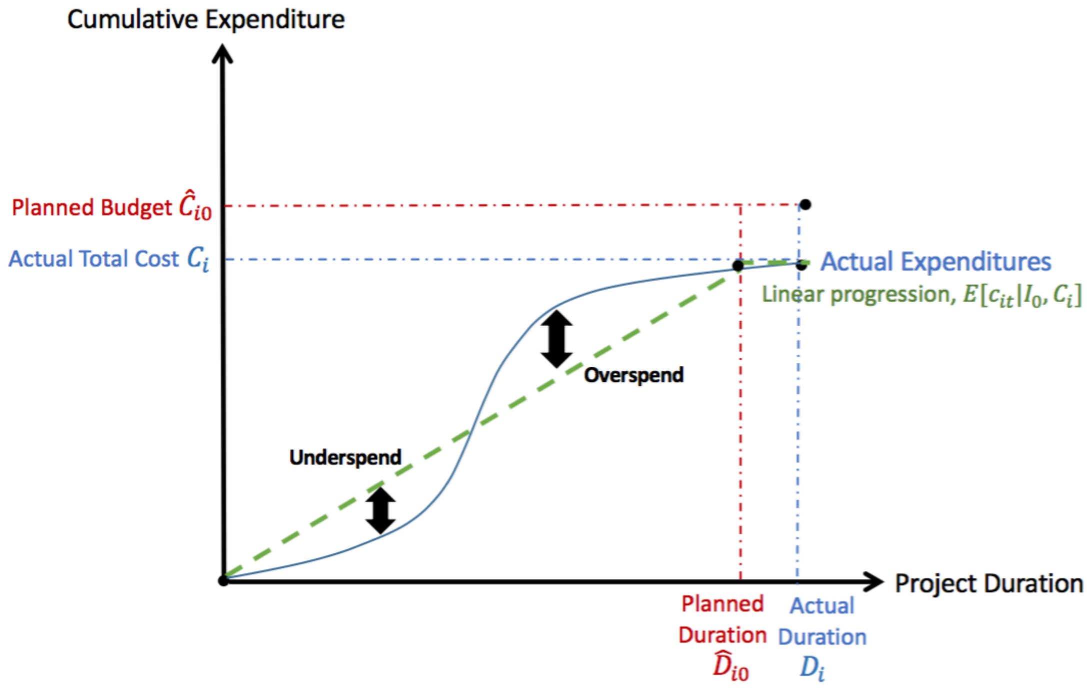


Figure 2: Building dynamic performance measures - Stylized description (ctd)

C. Based on actual costs and original duration



D. Based on actual costs and duration

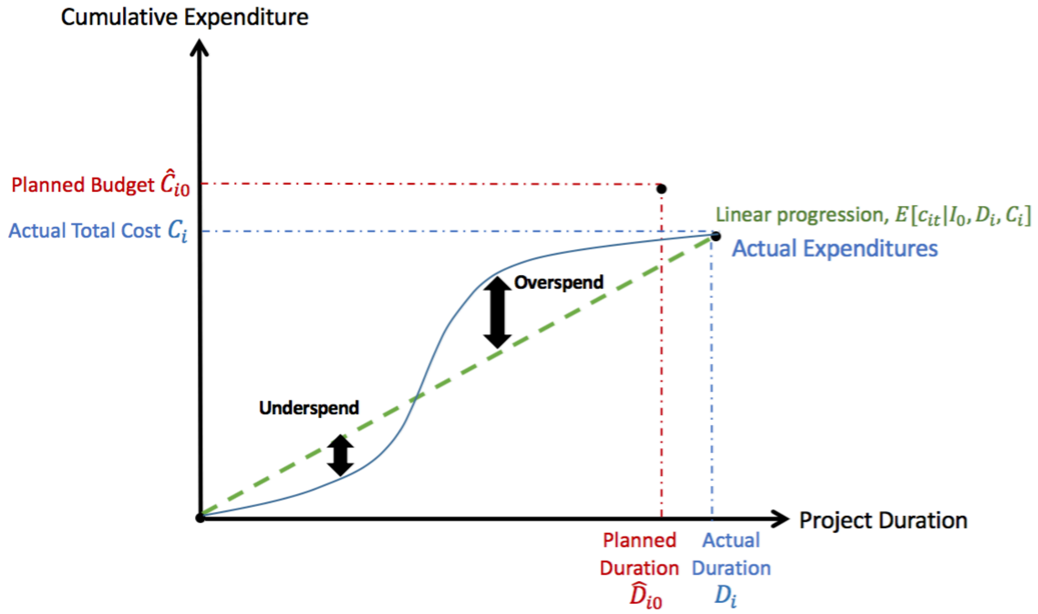
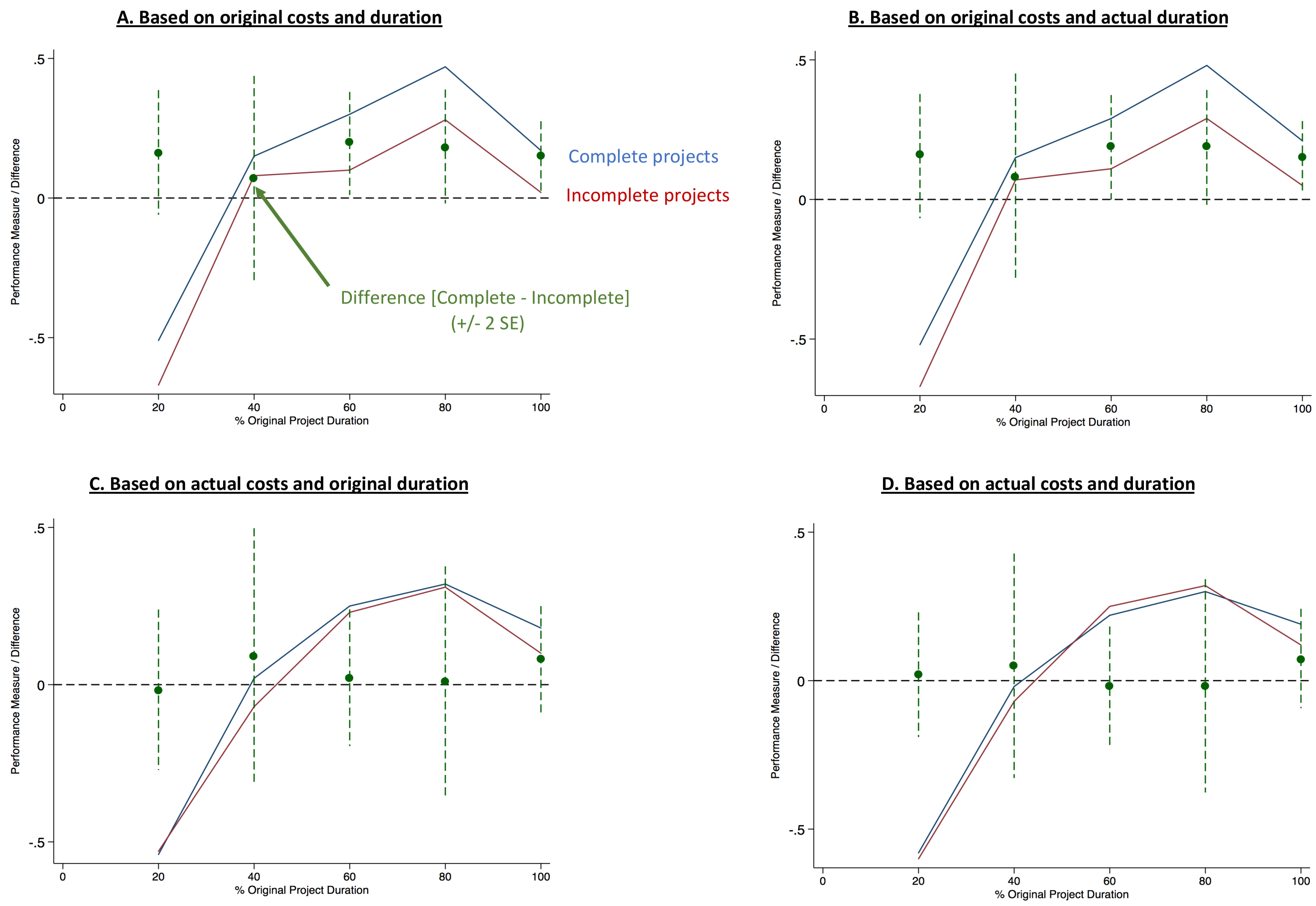


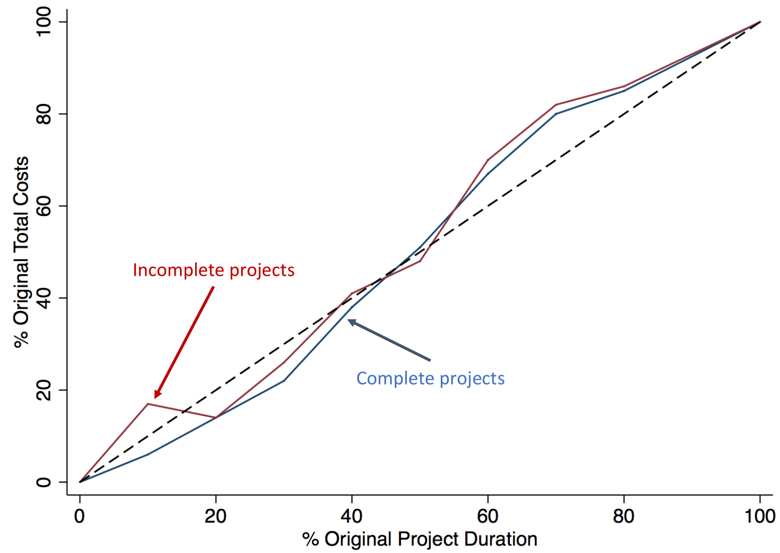
Figure 3: Dynamic performance measures, by project stage (% of originally planned duration)



Note: This figure reports plots of the dynamic performance measures, by project stage (expressed as % of originally planned duration). The measures are built by comparing actual expenditures in a project stage with linear interpolations of total costs based on (a) originally planned total costs and duration, (b) originally planned total costs and actual duration, (c) actual costs and originally planned duration, and (d) actual total costs and duration. Practically speaking, we compute *(Actual - Planned Expenditures) / Planned Expenditure*. Zero means a project spent according to budget, negative values represent underspends, and positive values overspends. An overspend for example of .5 would mean that a project in a year spend 50% more than planned. In all panels, performance measures are only plotted up to the end of the originally planned duration, for comparability purposes. We also report the difference between the average for complete and incomplete projects for each quintile of planned duration, and show drop lines representing +/- 2 standard errors to summarize the results from a test of mean equality.

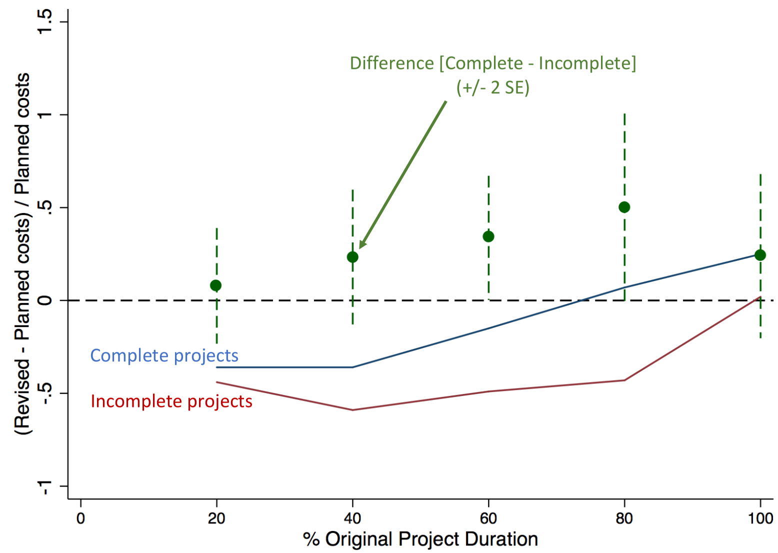
Figure 4: Initial and Revised Planned Costs, by project stage (Paper trail data)

A. True initially planned costs, by project stage and completion status



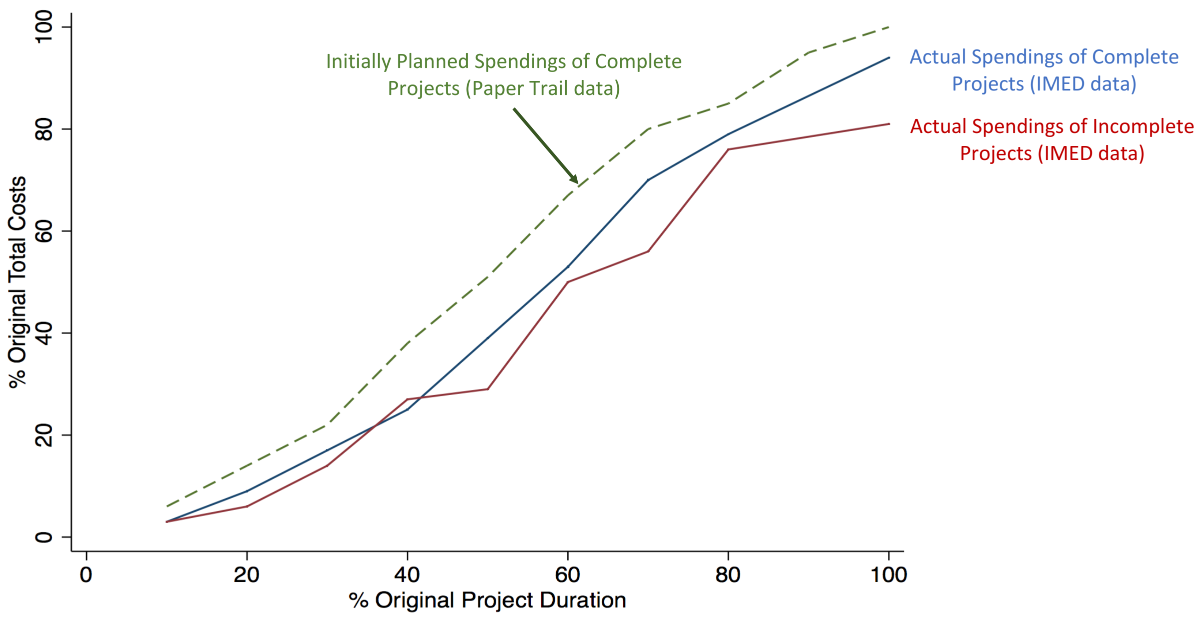
Notes: This plot shows true initial cumulative planned spendings, by project stage, and separately for completed and incomplete projects. We use the detailed data available on 168 evaluated projects, and average over project stages, as defined by Year of Implementation/Originally planned duration

B. (Revised - Originally planned) costs, by project stage and completion status



Notes: This graph plots the difference between Revised and Initially Planned Spendings, as a fraction of Initially Planned Spendings. We use the detailed data available on 168 evaluated projects, and average over project stages, as defined by Year of Implementation/Originally planned duration. We also plot the difference of differences between completed and incomplete projects (Blue - Red line), and show drop lines representing +/- 2 standard errors to summarize the results from tests of mean equality at each duration quintile.

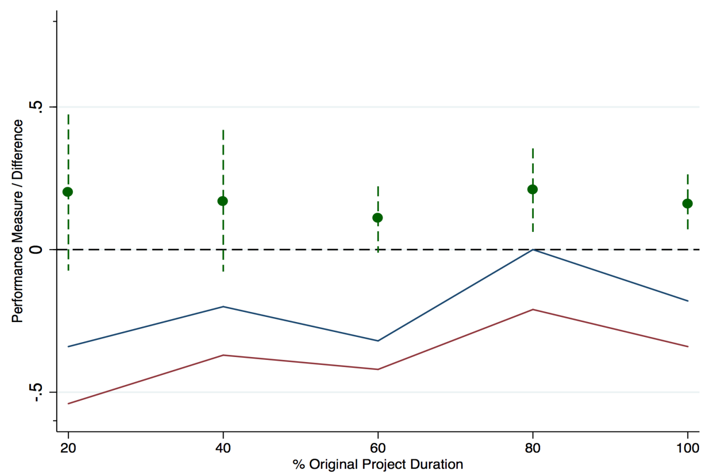
Figure 5: Compare paper trail benchmark and actual spendings



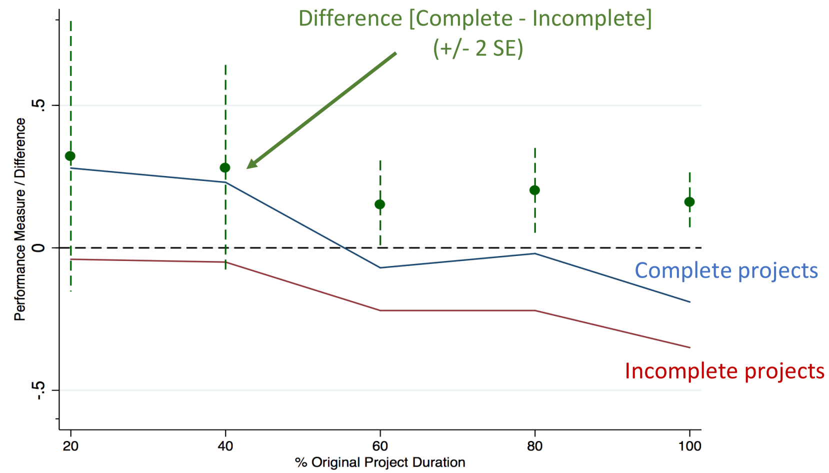
Notes: This graph compares the average initially planned spendings of completed projects in the sample of 332 projects for which we have detailed data, to average actual spendings of evaluated projects in the IMED database, separately for completed and incomplete projects.

Figure 6: Alternative dynamic performance measures, by project stage (% of originally planned duration)

A. Based on initially planned spendings



B. Based on revised planned spendings



Note: This figure reports plots of alternative dynamic performance measures, by project stage (expressed as % of originally planned duration). The measures are built by comparing actual expenditures in a project stage with benchmark initial (panel A) or revised (panel B) planned expenditures. These benchmarks are built by projecting actual initial and revised planned spendings of 190 completed projects in the subsample of projects for which we have paper trails. Practically speaking, we compute *(Actual - Planned Expenditures) / Planned Expenditure*. Zero means a project spent according to budget, negative values represent underspends, and positive values overspends. An overspend for example of .5 would mean that a project in a year spend 50% more than planned. In all panels, performance measures are only plotted up to the end of the originally planned duration, for comparability purposes.

Table A1: Balance - Evaluated vs. Unmatched Projects

Sample: IMED Database, Project data from 2003-2004 to 2011-2012 (excluding projects ongoing by the end of 2011-2012)

	Evaluated Projects			Unmatched Projects			
	N	Mean	SD	N	Mean	SD	P-value
Planned Start Year	1218	2002	4.33	348	2003	4.65	0.001
Planned Duration	1218	5.82	3.13	348	5.57	2.96	0.187
Planned Total Costs (Million US$)	1218	21.09	57.91	348	55.81	262.47	0.000
Dummy = 1 if used donor money	1218	0.17		348	0.23		0.008
Donor Money/Total Costs	201	0.64	0.24	79	0.69	0.22	0.001
Dummy = 1 if multiple agencies involved	1218	0.01		348	0.07		0.000
Dummy = 1 if Total Costs were revised	1218	0.65		348	0.54		0.000
Number of Revisions to Total Costs (if >0)	793	1.59	0.82	188	1.57	0.94	0.003
Project sector dummies (F-stat joint significance)							0.000
Executing agency dummies (F-stat joint significance)							0.000

Notes: This table presents balance tests comparing projects we found an evaluation report for ("Evaluated Projects") and those we could not find an evaluation report for ("Unmatched"). The sample excludes projects we know to be ongoing by the end of 2011-2012. We include a series of variables describing the history of revisions to total costs a project went through: an indicator equal to 1 if total costs were ever revised, dummies for early and late revision respectively equal to 1 if a revision occurred in the first, or last, half of the project's implementation; and a variable counting the total Number of Revisions along a project's life. We include vectors of sector and agency dummies, for which we report only the p-value for a test of joint significance.

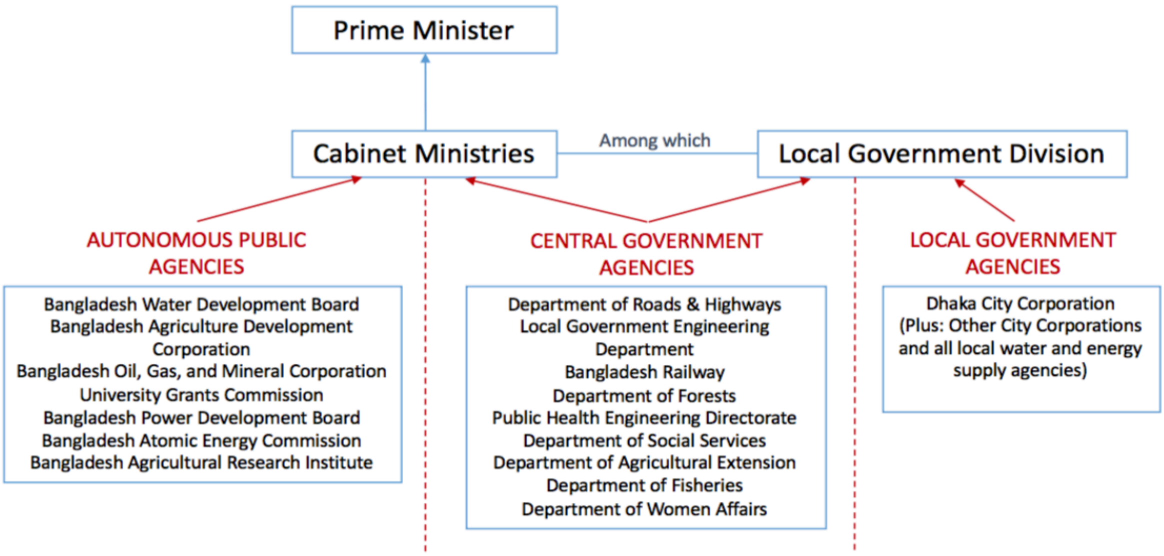
Table A2: Balance - Matched vs. and Unmatched with Paper Trail

Sample: IMED Database, Projects Evaluated between 2003-2004 and 2011-2012

	IMED Database only			IMED Database and Paper Trail			
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**P-value**
Planned Start Year	1050	2001	4.27	168	2005	3.52	0.000
Planned Duration	1050	5.85	3.14	168	5.65	3.08	0.445
Planned Total Costs (Million US$)	1050	21.92	60.97	168	15.89	32.58	0.210
Dummy = 1 if used donor money	1050	0.17		168	0.14		0.291
Donor Money/Total Costs	178	0.64	0.24	23	0.64	0.24	0.335
Dummy = 1 if multiple agencies involved	1050	0.01		168	0.01		0.650
Dummy = 1 if Total Costs were revised	1050	0.65		168	0.68		0.421
Number of Revisions to Total Costs (if >0)	679	1.58	0.81	114	1.64	0.87	0.307
Project sector dummies (F-stat joint significance)							0.000
Executing agency dummies (F-stat joint significance)							0.000

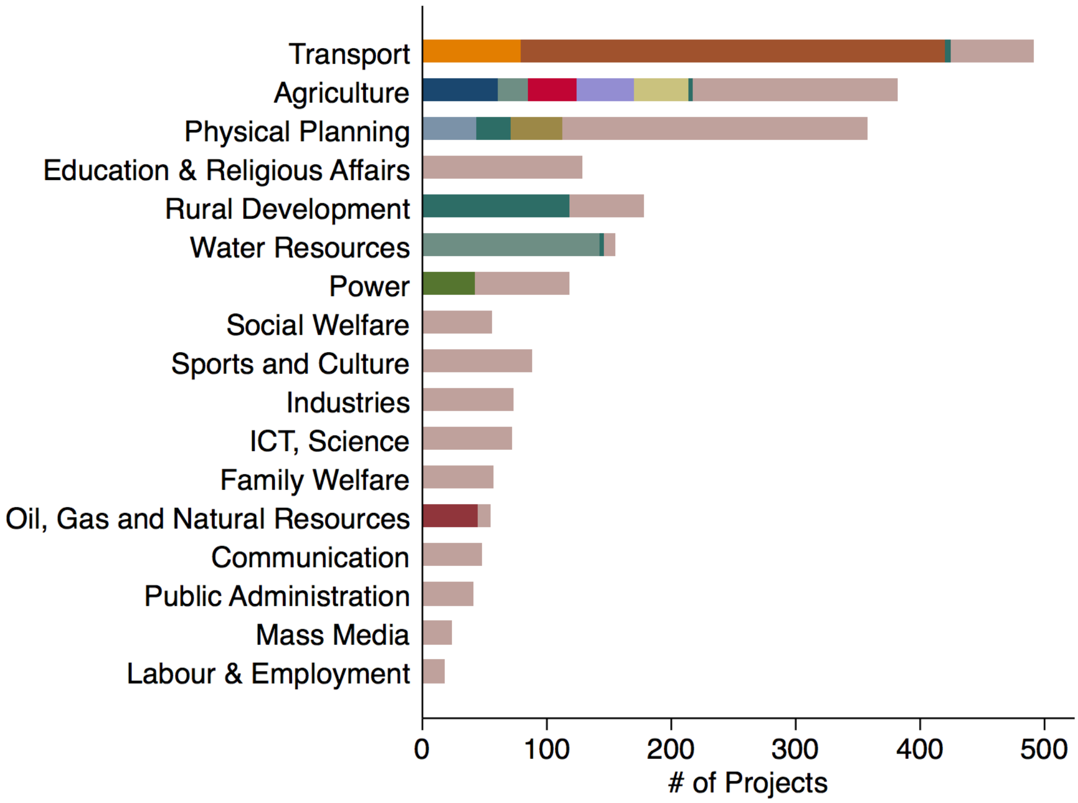
Notes: This table presents balance tests comparing projects we found full paper trails for ("IMED Database and Paper trail") and those we could not find a paper trail for ("IMED Database only"). The sample is restricted to projects evaluated between 2003-2004 and 2011-2012. We include a series of variables describing the history of revisions to total costs a project went through: an indicator equal to 1 if total costs were ever revised, dummies for early and late revision respectively equal to 1 if a revision occurred in the first, or last, half of the project's implementation; and a variable counting the total Number of Revisions along a project's life. We include vectors of sector and agency dummies, for which we report only the p-value for a test of joint significance. Note that we focused our search for paper trails on the 10 largest agencies only; selection according to sector and agency is therefore expected

Figure A1: Type of Agencies and Reporting Structure



Notes: This figure presents a visual summary of the type of agencies conducting infrastructure projects in Bangladesh. The arrows point towards the institutions agencies report to. We only list here the agencies responsible for most of the projects in the IMED database we use for analysis (these agencies are listed and detailed in Table 3).

Figure A2: Number of unique projects, by sector and agency



Notes: We plot the number of unique projects by sectors (height of bar). Each colour then represents projects implemented by one of the main agencies included in the paper trail sample. Greyish pink bars represent all other agencies. Total number of unique projects= 2482