Ministerial Turnover and Performance of World Bank Education Projects

Biniam Bedasso

Abstract

Government ministers can play such a significant role in the implementation of development projects under their portfolio that a high turnover of ministers may have implications for aid effectiveness. This paper examines the link between ministerial continuity in borrower governments and the performance of World Bank education projects implemented between 2000 and 2017 in 114 countries. I use a combination of quantitative and qualitative methods to trace the link between number of ministers during project implementation and project outcome ratings. There is a statistically significant and qualitatively meaningful negative correlation between ministerial turnover and project performance. Delays caused by transition and reshuffling of senior managers by new education ministers are shown to constitute possible causal mechanisms. There is some evidence that strong supervision by World Bank staff could mitigate the negative implications of ministerial turnover on project outcome.

Ministerial Turnover and Performance of World Bank Education Projects

Biniam Bedasso
Center for Global Development

Contents

List of Figures

List of Tables

1. Introduction

Government ministers are often the political custodians as well as administrative authorizers of development projects under their portfolio. The influence of an individual minister over the strategic directions as well as day-to-day operations of a ministry can be so far-reaching that a change of the officeholder could cause significant disruptions. This is particularly true in weakly-institutionalized environments where government offices are highly personalized. The impact of such disruptions can be rather pronounced in the case of specific and time-bound activities such as aid financed projects constituting a major part of a ministry's operations. Even in the absence of disruptions, frequent turnover could undermine performance because ministers are sometimes appointed with little prior experience in the respective sector and they may need to learn on the job which requires some level of continuity. But the net effect of a change in ministers does not necessarily have to be negative since it could also mean the replacement of a poorly performing minister with a better one.

This paper examines the link between ministerial continuity in borrower governments and the performance of World Bank education projects. The education portfolio is more amenable to such analysis than most other World Bank financing because education projects can often be linked directly to distinct ministries in borrower governments. But, in addition to an in-depth examination of education projects, the paper also looks at cabinet-level continuity and World Bank project performance more broadly in order to shed light on cross-sectoral dimensions. I employ a mixed-methods approach combining both quantitative analysis of a large projects dataset and qualitative analysis of a sub-sample of projects. Quantitatively, I estimate the correlation between ministerial turnover and project performance across a large sample of projects. Qualitatively, I employ a case study approach on a small number of selected projects to validate the quantitative model as well as identify potential causal mechanisms.

The quantitative analysis is based on the dataset of World Bank project performance ratings produced by the Independent Evaluation Group (IEG) between 2000 and 2017, combined with the global dataset on members of cabinet compiled by the WhoGov project. The first leg of the quantitative analysis aims to establish the broader relationship between leadership continuity and project performance by estimating the correlation between government-wide cabinet retention during project implementation and performance ratings of World Bank projects. The second leg of the quantitative analysis zooms in on education as a special case by linking education projects to the total number of education ministers that held office over the course of project implementation. The qualitative analysis examines a sample of 23 education projects with the highest or lowest numbers of education ministers per project-year to provide further evidence complementing the quantitative findings.

The notion of project success as measured by the IEG ratings is different from project "impact" since it is narrowly focused on progress towards specific project objectives and is often assessed subjectively. Therefore, it is an imperfect measure of effectiveness. But it is still an important

intermediate measure since a project normally needs to be operationally successful in order to have the desired impact. Moreover, the evidence is largely correlational since it is difficult to identify a source of exogenous variation in ministerial turnover for the full set of projects in the quantitative analysis. However, the qualitative analysis provides useful insights on the nuances of the relationships inside the black box of correlations.

The results show that projects implemented during periods of higher cabinet stability are on average more likely to succeed. Ministerial continuity is even more strongly associated with project success in the case of education projects. By contrast, ministerial continuity has no statistically significant relationship with project performance in the health sector. The fragmentation of the education portfolio between multiple ministries is also negatively correlated with project performance. However, the case studies show that the turnover of individual minsters is more saliently linked to performance in project assessment narratives than fragmentation of authority across ministries. The qualitative evidence also reveals that delays caused by transition and associated changes in senior managers at ministries of education constitute potential causal mechanisms linking ministerial turnover to project performance. There is some evidence that strong supervision by World Bank staff could help mitigate the negative implications of ministerial turnover on project outcome.

This paper brings together two disparate literatures—one on World Bank project performance and another on ministerial turnover—with the aim of generating further insights on how political dynamics at the national level could impact aid effectiveness at a project level. Existing literature on World Bank project performance often treats country-level institutional factors and project-specific operational factors dichotomously. Moreover, there has been surprisingly little effort to understand project-specific determinants of success on the side of the borrower government compared to the number of studies conducted to identify success factors on the World Bank's side. The current study attempts to bridge that gap in the literature by exploring the role of one of the most important players linking country-level institutions and project-specific operational structures: the minister in charge of the portfolio.

The rest of the paper is organized as follows. Section two provides a brief overview of the related literature. Section three describes the data and methods used for the empirical analysis. Section four presents the results of the broader analysis on cabinet stability and project performance. Section five presents the core results of the study on the link between ministerial turnover and project performance in the education sector. Section six concludes.

2. Literature review

The recent body of literature on the determinants of the success of development aid projects was preceded by an earlier strand of literature analyzing aggregate aid effectiveness at a macroeconomic level. Despite various efforts to generate credible results in a number of landmark studies

(for example, Boone, 1996), the quest to understand the causal impact of aid on key growth and development outcomes has proven elusive due to conceptual and methodological challenges. The shift towards scrutinizing factors affecting the success of specific development projects has to some extent coincided with the emergence of "the credibility revolution" in empirical economics which emphasizes the causality of relationships at the micro level. In some sense, the study of project effectiveness in donor organizations such as the World Bank can be conceived as the meso-level link bridging macro-level aid effectiveness research and micro-level impact evaluation.

The most basic distinction that is often made between various correlates of project performance relates to country-level vs. project-specific characteristics. In the case of country-level factors, a number of studies have documented that there is positive correlation between time-varying characteristics such as economic growth rate and institutional quality, and world bank project performance (Denizer et al., 2013; Briggs 2020; Bulman et al., 2017). However, these relationships might not always hold under all circumstances. For instance, there is no statistically significant correlation between bureaucratic capacity and project performance in some cases (Blum, 2014). This could be potentially because project teams take existing state capacity into account when they design World Bank loan projects. In other words, project design and supervision quality could be key variables mediating the impact of country-level characteristics such as state capacity.

There has been a boom in recent years in research on project-specific characteristics explaining outcome ratings. Much of this work is inspired by the finding in Denizer et al. (2013) showing that 80 percent of the total variation in project outcomes occurs across projects within countries, rather than between countries. Apart from a few studies that looked into factors such as project size and duration, most of the research on project-specific factors has focused on the quality of project design/preparation and supervision as determinants of project outcome. The quality of design has been proxied by a variety of indicators ranging from length of preparation period to the amount of analytical input by Bank staff (Kilby, 2015, Deininger et al., 1998). Generally, there is positive association between project preparation and outcome. An examination of the critical success factors for Word Bank projects—as identified by surveyed staff members - also places design among the top five components (Ika et al. 2012).

There is sizeable evidence demonstrating the positive link between project supervision and subsequent outcome. Kilby (2000) shows that the quality of early supervision as measured by number of staff weeks spent on supervision is positively correlated with project performance. Another way supervision quality has been captured in the literature is through Task Team Leader (TTL) characteristics. Notably, Denizer et al. (2013) use the project track-record of the TTL as a proxy for supervision quality which is positively correlated with project outcome ratings. More recent studies have delved deeper into the role of World Bank staff characteristics, including context knowledge and posting location, for project performance (Heinzel and Liese, 2021; Honig, 2020; Limodio, 2021). In addition to supervision, the monitoring and evaluation quality of projects is also shown to be positively associated with performance (Raimondo 2016, Bedasso 2021).

Moreover, enhanced transparency mechanisms constituting Access to Information policies accompanied by independent appeals processes are shown to be causally linked to improved project performance (Honig et al., 2022).

As far as project-specific characteristics of borrower governments are concerned, the survey of critical success factors by Ika et al. (2012) identifies national coordination as one of the top five elements contributing to project performance. Moreover, the type of implementing partner seems to have implications for the success of projects. Shin et al. (2017) show that implementing-partnerships with host country governmental agencies tend to produce a less successful outcome compared to partnerships with non-state actors. Projects implemented in periods following civil conflict are less successful, probably due to latent instability in implementing organizations (Chauvet et al. 2010).

There is extensive literature on cabinet instability and ministerial turnover particularly drawing on the experiences of Westminster-type democracies. Much of the literature focuses on exploring the causes of cabinet reshuffles. For instance, Huber and Martinez-Gallardo (2008) analyse microlevel data from 19 parliamentary democracies to distinguish between factors driving the turnover of individual ministers from those causing cabinet level instability. Their findings highlight the importance of factors such as the coalition status of the government, the policymaking power of ministers and the size of the minister's party for individual turnover. In the context of less democratic countries, the size and composition of cabinets is sometimes linked to their role as an apparatus to distribute rent and manage elite alliances (see for example, Wehner and Mills, 2022; Bokobza et al. 2022). Generally, the existing evidence points to the need for accounting for underlying political factors that might drive ministerial turnover to better isolate its impact on economic outcomes such as aid effectiveness.

3. Data and methods

i. Measuring project performance

The main outcome variable of the study is the performance of World Bank projects at exit. The most commonly used performance metric for World Bank projects is the outcome rating by IEG at the completion of each Board-approved project. IEG employs objectives-based project evaluation methodology to arrive at ratings. In this context, project performance is defined as:

> *"the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently". (IEG, 2015)*

This means a project's performance is judged against three main criteria: relevance (to country priorities), efficacy and efficiency. The standard method of assessment consists of IEG's validation of the Implementation Completion Report (ICR) prepared by project staff and independent judgment on performance ratings. Around four in five projects are rated using this desk-based approach

whereas the remaining 20 percent undergo in-depth field-based evaluation which leads to a more comprehensive assessment called Project Performance Assessment Report (PPAR). In either case, the IEG assessments cannot be considered impact evaluation since they are narrowly focused on progress towards specific project objectives and do not apply standard impact evaluation methodology. However, they are a key intermediate metric towards measuring broader development impact and aid effectiveness.

The desk-based assessment is by definition less rigorous than the field-based PPAR. Hence, it is useful to take the PPAR ratings as a benchmark to assess the relative accuracy of the desk-based ratings which make up approximately 80 percent of project assessments. The other source of potential measurement inconsistency is the periodic updates introduced to the measurement methodology at a couple of points over the period covered by the data. The methodology was updated in 2005 to align the methodologies of IEG and WB operation departments and in 2011 to introduce distinct handling of investment and development policy lending operations. Figure 1 shows that the desk-based ratings track closely the PPAR ratings particularly in the period after 2005. Therefore, on average, the desk-based ratings are as accurate as the PPAR. In contrast, there is a shift in the pattern of ratings across the three periods which could be partly attributed to the change in methodology. So, it is important to control for time effects when using the ratings in regression analysis.

**FIGURE 1. The distribution of IEG outcome ratings by type of evaluation and period**
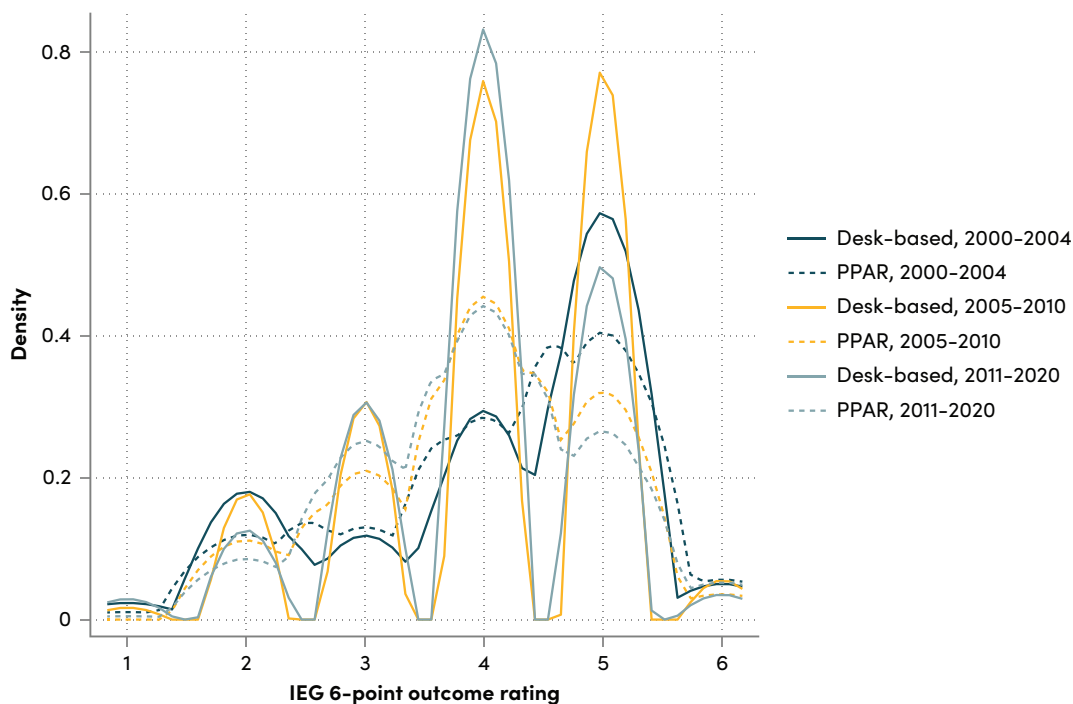
The current outcome rating framework assesses projects on a standardized 6-point scale, i.e. 1. highly satisfactory, 2. satisfactory, 3. moderately satisfactory, 4. moderately unsatisfactory, 5. unsatisfactory, 6. highly unsatisfactory. To simplify the analysis and since less than 4.2 percent of projects completed

between 2000 and 2017 received either of the two extreme ratings, I have reclassified "highly satisfactory" together with "satisfactory" and "highly unsatisfactory" together with "unsatisfactory". This means project outcome is measured on a 4-point scale for the current analysis.

FIGURE 2. Distribution of outcome ratings by sector (2000–2020)
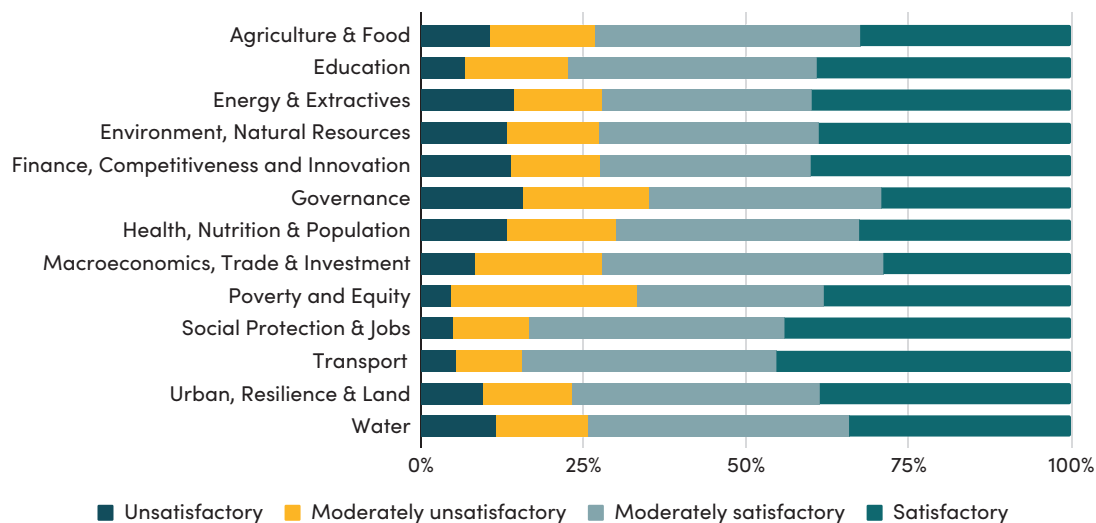
Three-quarters of projects completed between 2000 and 2020 have been rated as satisfactory or moderately satisfactory. In other words, a great majority of projects were judged to have not suffered significant shortcomings in terms of meeting their stated objectives effectively and efficiently. Therefore, on average, the most likely impact of an improvement in project performance is to turn a moderately satisfactory rating into a satisfactory rating. Figure 2 displays the variation in outcomes across sectors. Projects in sectors such as transport enjoy a generally favorable outcome whereas sectors such as governance exhibit a higher share of projects with unsatisfactory/moderately unsatisfactory outcomes. Education is one of the four sectors in which well under a quarter of projects are rated as unsatisfactory/moderately unsatisfactory.

ii. Measuring ministerial continuity

Most World Bank projects are managed by one or more ministries on the borrower government side. Therefore, cabinet ministers responsible for those ministries are expected to provide leadership. Depending on the number of ministries involved in a particular project, leadership continuity can be understood as cabinet level or portfolio specific ministerial continuity. I employ data presented in the WhoGov dataset compiled by Nyrup and Bramwell (2020) to measure ministerial continuity at both cabinet and portfolio levels. The dataset contains detailed information on the identities of all cabinet members in 177 countries by year during the 1966–2016 period.

I use separate measures of ministerial continuity for the two tiers of analysis. Firstly, average cabinet retention rate over the duration of project implementation is used to measure overall leadership continuity, particularly for multisectoral projects sometimes including education-related components. Secondly, in the case of education projects, the total count of education ministers over the duration of project implementation serves as an indicator of continuity. Cabinet retention rate is the percentage of cabinet members retained from one year to the next. For each project, the average is calculated over the duration of the project. Figure 3 shows that the distribution of average cabinet retention rate for projects completed between 2000 and 2020 is skewed towards stability with a median rate of 0.69. This means, during the life cycle of the median project, seven in ten cabinet ministers are retained in office year on year.

FIGURE 3. Density distribution of average cabinet retention rate during project implementation

N = 5428 projects
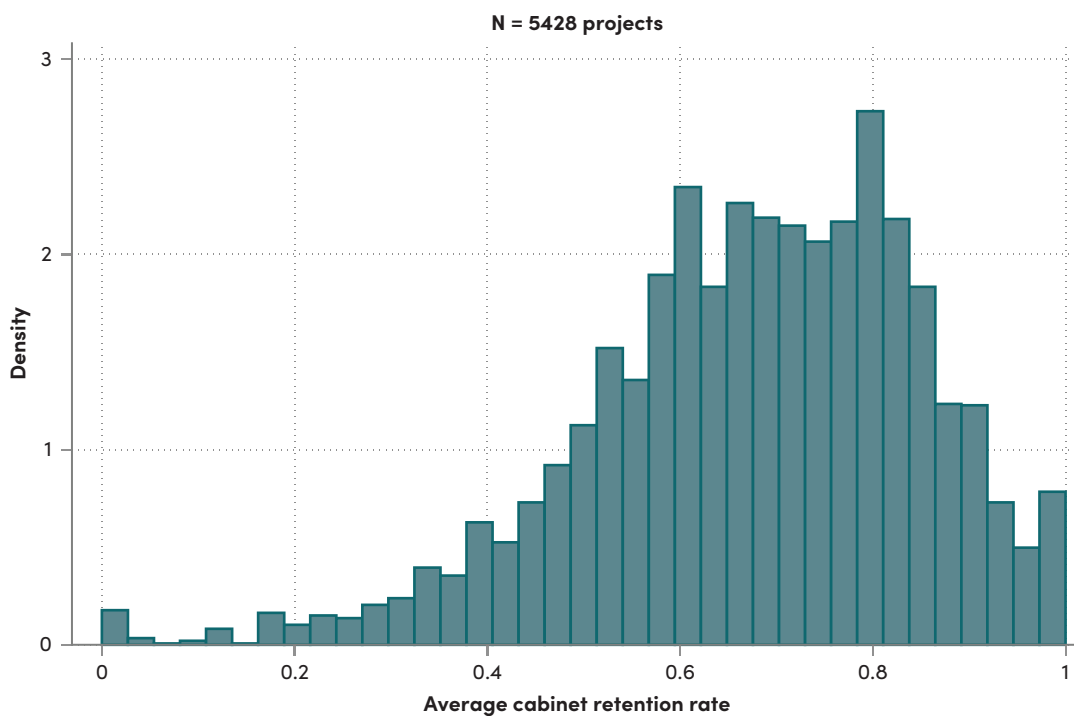
However, the pooled data can mask substantial between-country and within-country variations that might be useful to understand the source of heterogeneity in leadership continuity at a project level. Figure 4 attempts to illustrate this by depicting the level of cabinet continuity for projects implemented in three countries (China, India and Kenya) over the past two decades. Overall, projects in India are implemented in an environment characterised by higher turnover than both Kenya and China. By contrast, China features the highest level of continuity (of the three) even though the rate of retention was lower in the mid-2000s than in recent years.

**FIGURE 4. Between-country and within-country variations
in cabinet retention over project duration**
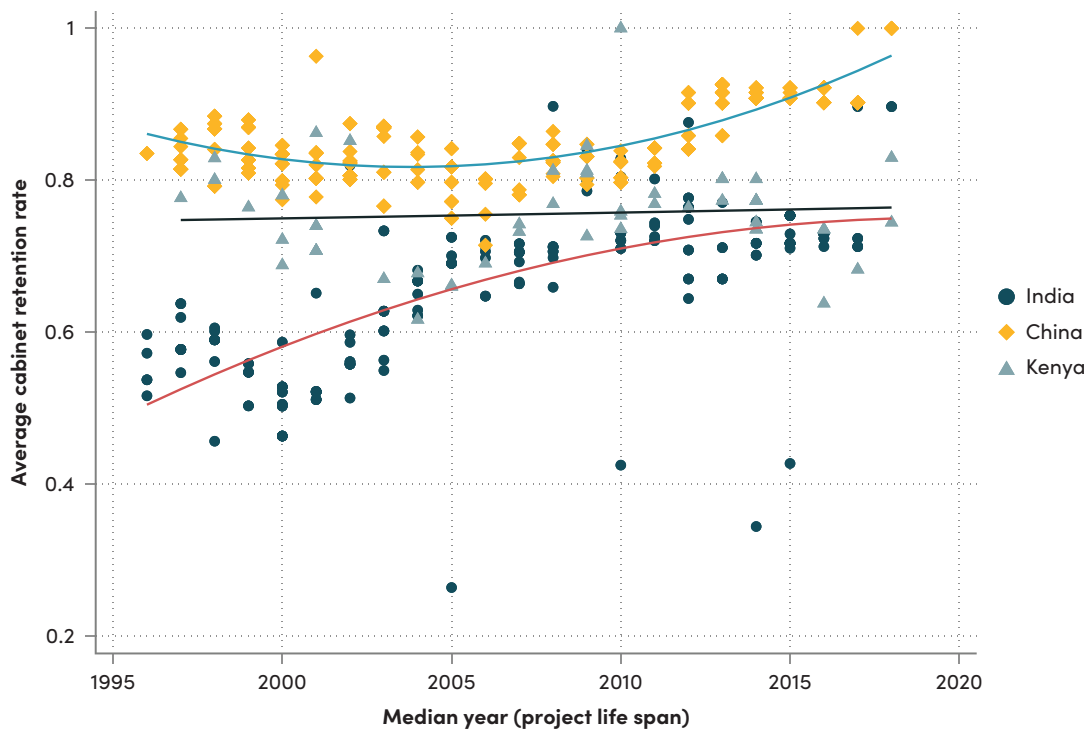
For the education-specific analysis, the total number of cabinet ministers with education related portfolios over the duration of the project serve as the main explanatory variable. There are some countries which have multiple ministries responsible for various aspects of the education sector (for example, basic education vs higher education). In some cases, ministries could be split or merged during the course of project implementation. In such instances, the change in leadership and institutional structures that come with such restructurings can have a similar effect as turnover of ministers. Therefore, the total number of education ministers, as a single variable, is composite of the number of ministers per portfolio and the number of cabinet level portfolios related to education.

Figure 5 presents the distribution of the total number of education ministers per a year of project. It also displays the distribution of the total number of health ministers which is presented here to provide a comparator from a related human development sector as a reference. The median number of ministers presiding over the education portfolio per a year of an education project's implementation is 0.57. In contrast, the median number of health ministers per a year of implementation of a health project is 0.33. The relatively high number of education ministers can be attributed to a combination of high turnover of education ministers and the fragmentation of the education portfolio across multiple ministries.

FIGURE 5. Density distribution of number of ministers for education and health
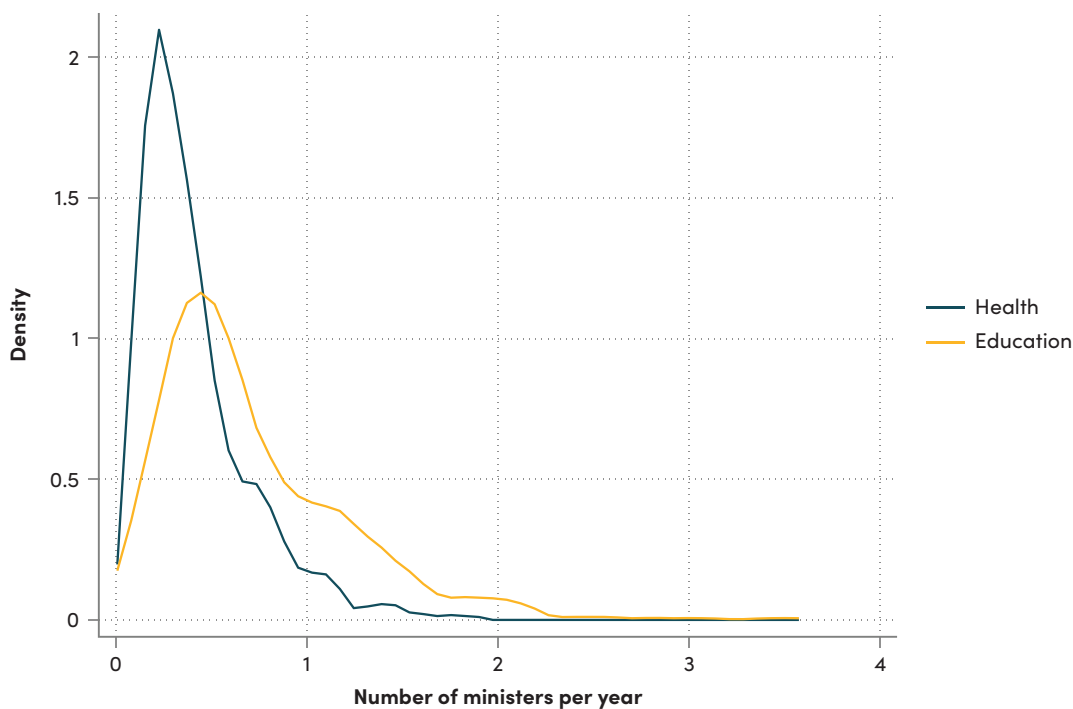
iii. Methods

The study employs a mixed methods approach consisting of large sample regression analysis followed by small-N case studies. I first estimate the correlation between ministerial continuity and project outcome rating using an ordinal regression model of the form:

$$\Pr(Outcome_j = i) = \Pr(k_{i-1} < \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \beta_k x_{kj} + u_j < k_i) \tag{1}$$

where the probability of observing project outcome i corresponds to the probability that the estimated linear function composed of parameters β_1, β_2, ..., β_k, a vector of independent variables X_j and a normally distributed error term u_j falls within the range of cutpoints k_1, k_2, ..., k_{I-1}. I is the number of possible outcomes which is four in the current case.

The average marginal effects, with respect to ministerial continuity, derived from the coefficients in (1) can be interpreted as the partial correlation between ministerial continuity and project performance. Ministerial continuity cannot be considered exogenous to project performance because some unobserved underlying dynamic could be behind both variables. Therefore, the best outcome that can be expected from the regression analysis in the current setup is a precise-enough measure of association between ministerial continuity and project outcome.

The qualitative case studies complement the quantitative estimates by substantiating the link between continuity and project outcome through systematic examination of evaluation narratives. The universe of cases for the qualitative analysis is 49 education projects that were completed between 2000 and 2017 and were subject to in-depth field-based PPAR. I restrict the case studies to those projects with PPAR because the narrative in the reports is more independently constructed than in the case of projects with only desk-based assessments. Therefore, they are more reliable to conduct comparative case studies. The cases are selected according to the "extreme cases on the independent variable" criteria by picking a quarter of projects on either end of the ministerial continuity spectrum. In other words, 13 education projects with the highest number of ministers per project-year and 10 others with the lowest number of ministers per project-year are selected for the qualitative analysis.

4. Cabinet turnover and World Bank project performance: cross-sectoral overview

This section presents the results from the baseline analysis to establish the overall relationship between cabinet continuity and World Bank project performance across all sectors including multisectoral projects. By tracing the link between cabinet retention rate and project performance, it highlights the relationship between overall government continuity and the performance of aid projects. This is useful to capture the implications of pervasive instability which is manifested at a cabinet level and is not confined to a specific portfolio for the performance of all sorts of projects regardless of sector.

Table 1 presents a set of results based on the ordered probit estimation of the correlates of performance ratings of all projects completed between 2000 and 2017. In addition to average cabinet retention, all specifications control for key project-specific variables such as design and supervision quality as well as country-specific variable such as government effectiveness index and per capita income, most of which are shown to be correlated with project performance in the literature. Country fixed effects are included to absorb the effects of time-invariant country level omitted variables whereas year fixed effects are included to account for changes in rating methodology over time.

TABLE 1. Correlates of project performance: all sectors

Average Marginal Effect on Probability of "Satisfactory" Outcome (Ordered Probit)	(1)	(2)	(3)	(4)
Average cabinet retention	.155*** (.041)	.096** (.044)		.094** (.045)
Average cabinet retention: first half of project			.025 (.032)	
Average cabinet retention: second half of project			.108*** (.028)	
Number of heads of state per year		−.143*** (.046)	−.110*** (.049)	−.123*** (.048)
Regime type: Polity democracy score				.007*** (002)
Political stability index				.029* (.016)
Project design quality: satisfactory	.412*** (.016)	.414*** (.016)	.341*** (.014)	.412*** (.016)
Project supervision quality: satisfactory	.351*** (.014)	.349*** (.013)	.341*** (.014)	.352*** (.014)
Government effectiveness index	.105*** (.034)	.093*** (.034)	.121*** (.037)	.064* (.037)
Per capita GDP	−.088*** (.039)	−.081** (.039)	−.082** (.042)	−.066* (.040)
Other controls:	– Project volume – Project length – Year fixed effects – Country fixed effects			
Psuedo R2	0.31	0.31	0.31	0.31
N	4866	4829	4543	4685

Notes: Standard errors are in parenthesis. ***p<.01, **p<.05, *p<.1.

The main finding is that there is a positive and statistically significant association between cabinet continuity during project implementation and the probability of attaining a "satisfactory" outcome (i.e. the highest performing 37 percent education projects). The baseline specification in column 1 shows that a one standard deviation change in average cabinet retention is associated with 3.4 percentage point change in the probability of satisfactory outcome. However, the average marginal effect of cabinet retention declines when the number of heads of state during a project is controlled for in column 2. This means some of the observed correlation in column 1 is due to complete change in administration beyond cabinet reshuffles. Column 4 accounts for political factors—namely regime type and political stability—that are shown to be linked to cabinet reshuffle in the literature but are also generally believed to be correlated with government performance. However, their inclusion does not seem to attenuate the cabinet retention coefficient.

Figure 6 shows that the overall marginal effect of cabinet retention is limited to flipping a project from "moderately satisfactory" to "satisfactory". Although this may make the role of cabinet stability seem less important, in reality, it is relevant for a great majority of projects since 75 percent of

education projects fall in the two categories of outcome. In column 3, average cabinet retention rate for the first and second half of the project lifespan are entered separately as explanatory variables. This is intended to check if cabinet stability exhibits some form of non-linear effect as supervision quality which is shown to have a larger effect at the early stages of project implementation (Kilby 2000). However, the results show that it is rather in the later part of project implementation that cabinet stability seems to correlate significantly with project outcome.

FIGURE 6. Predicted probability of various project outcomes by rates of cabinet retention
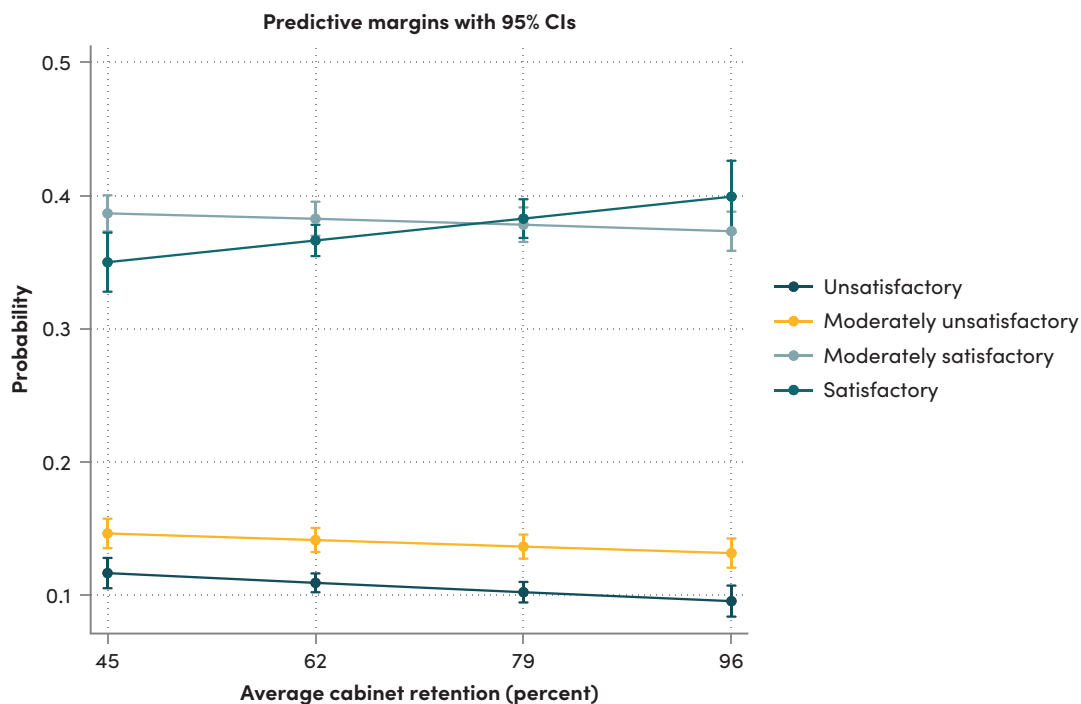
Even if the comparison of PPAR-based and desk-based ratings in the previous section has shown that there is overall measurement consistency in IEG ratings, there could still be some unobserved and nonrandom bias. Therefore, it is useful to do a robustness check using alternative performance scores compiled by independent external parties. One such measurement that covers World Bank projects approved between 1981 and 2012 was constructed by Malik and Stone (2017). They compute their own average performance score for each project based on a series of objective indicators extracted from project reports. Even though the sample size drops significantly due to mismatch in the period covered by the two datasets, the overlap is enough to run a robustness check on the main estimation in Table 1.

TABLE 2. Robustness check with alternative outcome measurement (OLS estimation)

Outcome Variable: Project Performance (Numeric Scale, 0–4)	(1)	(2)	(3)
Average cabinet retention rate	.374** (.171)	.350* (.181)	.325* (.182)
Additional controls (in addition to the baseline specification in Table 1, column 1)	Report-type fixed effects	Head of state turnover, Report-type fixed effects	Head of state turnover, Political stability index, Report-type fixed effects
Adjusted R-Squared	0.27	0.27	0.27
N	2465	2446	2432

Notes: Standard errors are in parenthesis. ***p<.01, **p<.05, *p<.1.

The results in Table 2 demonstrate that the positive correlation between cabinet continuity and World Bank project performance holds even when alternative outcome scores are employed. The magnitude of the correlation remains small with a standard deviation change in average cabinet retention rate being associated with 1.5 percent change in standardized performance score.

5. Ministerial continuity and performance of education projects

This section zooms in on the relationship between the performance of education projects and the number of education ministers that have been in office during project implementation. The section is divided into two parts with the first one focusing on the results of the regression analysis and the second on the complementary case studies.

i. Regression results

Table 3 presents the correlates of outcome ratings of education projects completed between 2000 and 2017. I have also run the baseline estimation for health projects to check if similar patterns are present in a comparator human development sector. The headline result shows that a high number of education ministers during project implementation is associated with lower project performance. One standard deviation increase in number of ministers responsible for education portfolio during project implementation is associated with 11 percentage points decline in probability of "satisfactory" project outcome.

TABLE 3. Correlates of project performance: education and health

Average Marginal Effect on Probability of "Satisfactory" Outcome (Ordered Probit)	Education	Health
Total number of ministers during project life span	−.034*** (.008)	−.015 (.012)
Number of heads of state during project	.028* (.017)	−.018 (.019)
Project design quality: satisfactory	.412*** (.061)	.388*** (.051)
Project supervision quality: satisfactory	.352*** (.059)	0.341*** (.052)
Government effectiveness index	−.041 (.034)	−.223* (.115)
Per capita GDP	.352*** (.108)	.279* (.157)
Other controls:	– Project volume – Project length – Project length X Number of ministers interaction – Year fixed effects – Country fixed effects	
Pseudo R2	0.57	0.46
N	407	392

Notes: Standard errors are in parenthesis. ***p<.01, **p<.05, *p<.1.

Making inference based only on the total number of ministers could be misleading because longer projects are likely to run through the tenures of more ministers which might not necessarily be linked to higher turnover. Therefore, I include a term interacting total number of ministers with project length to account for this. The negative effect of ministerial turnover is robust to the inclusion of number of heads of state during project implementation (which has a positive and marginally significant coefficient). This indicates that the effect of ministerial continuity is independent from regime stability.

Unlike in education, there is no statistically significant correlation between number of ministers and project performance in the health sector as shown in column 2 of Table 2. This could be attributed to one or both of the following potential explanations. Firstly, the content and structure of health projects make them more immune to high turnover at the most senior levels of the borrowing government than education projects. Secondly, the variance in the capabilities of ministers appointed to health portfolios is usually smaller than that of education ministers which helps to minimize the disruption caused by turnover. A testable hypothesis in this regard is that education ministers are often political appointees whereas health ministers are supposed to have some level of technical expertise in the field.

One of the channels through which ministerial turnover could affect project outcome is disruptions in coordination and oversight. Therefore, it is useful to examine whether bank supervision could help mitigate the negative effect of turnover by compensating for some of the gap on the government side.

Figure 7 shows that the impact of ministerial turnover depends on the quality of bank supervision. In projects with high quality supervision, the probability of attaining a satisfactory outcome declines only marginally with the total number of education ministers. The corresponding decline in the probability of attaining a satisfactory outcome is much more dramatic in projects with low quality supervision. Conversely, the probability of ending up with an unsatisfactory outcome increases significantly with number of ministers under low quality supervision.

As noted above, one of the factors contributing to the total number of education ministers is the number of independent ministries responsible for various parts of the education sector. Over 46 percent of the education projects in the current sample were implemented in a context where there were more than one ministry in charge of education. The fragmentation of authority across multiple ministries is unlikely to have much impact on strictly subsector-specific projects—for example higher education—when one of the ministries is responsible for that particular portfolio. However, a large number of projects implemented in the presence of multiple ministries have sector-wide elements that could fall under the mandates of more than one ministry. Therefore, it is important to understand the impact of such fragmentation on project performance separately from the turnover of individual ministers.

FIGURE 7. Heterogeneity of effect with respect to quality of supervision
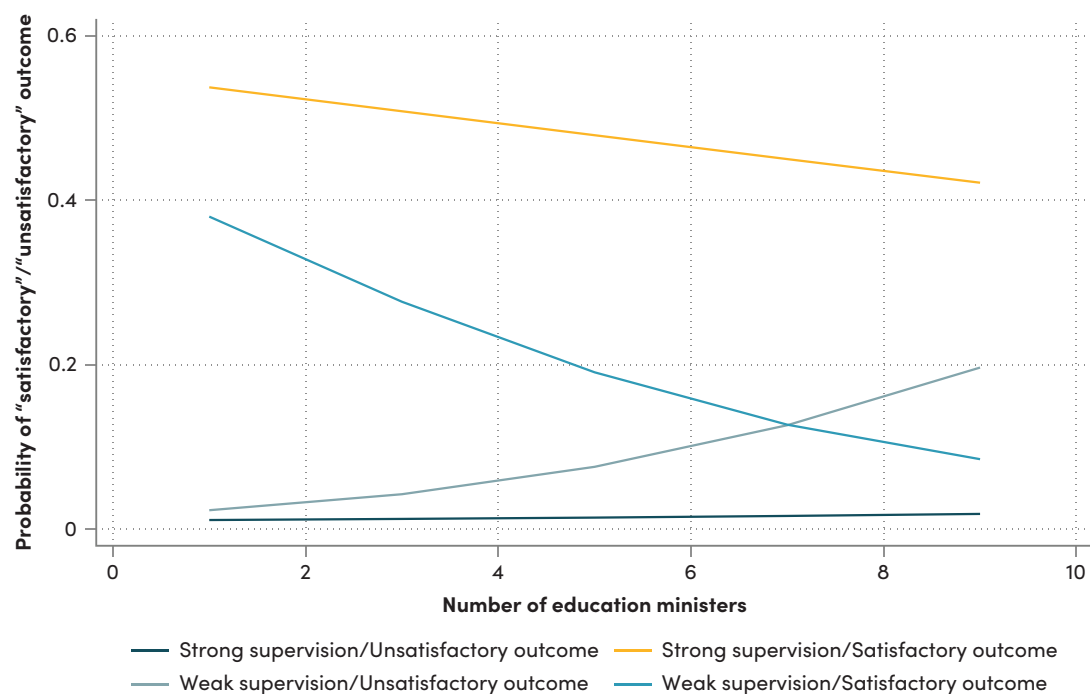
In contexts that are prone to frequent restructurings, it is not only the fragmentation of authority across various ministries that could influence performance but also whether the portfolio split during the implementation of a project. During the course of implementation of 21 percent of education projects in the current sample, the number of ministries responsible for education

increased implying some form of split. Such splits could potentially have negative implications for project performance due to disruptions in the institutional structure created as a result of a split. But, on the flipside, a split could also be beneficial to a subsector-specific project if it helps streamline activities while institutional restructuring is undertaken smoothly.

Table 4 shows that there is a strong negative correlation between the maximum number of ministries that have been responsible for education and project outcome. A one standard deviation increase in the number of ministries is associated with a 12 percentage point decline in the likelihood of attaining satisfactory outcome. However, there is a possibility that this could be driven by the underlying clientelist tendency in a government which often leads the proliferation of cabinet positions (Wehner and Mills, 2022). This is addressed by controlling for cabinet size in column 2 which shows that the effect of fragmentation of the education portfolio is independent of overall cabinet size and associated political dynamics. On the other hand, there is no statistically significant correlation between the occurrence of split during projection implementation and performance ratings. This could be interpreted as indicating that the benefits of streamlining due to a split offset the negative implications of potential disruption and coordination failure caused by it.

### TABLE 4. Education portfolio fragmentation and project performance

Average Marginal Effect on Probability of "Satisfactory" Outcome (Ordered Probit)	(1)	(2)	(3)
Maximum number of ministries responsible for education during project period	−.101*** (.007)	−.092*** (.025)	
Education portfolio split mid–project			−.015 (.041)
Median cabinet size during project period		−.004 (.003)	
Project design quality: satisfactory	.437*** (.062)	.435*** (.061)	.433*** (.061)
Project supervision quality: satisfactory	.333*** (.051)	.336*** (.052)	.331*** (.050)
Government effectiveness index	.020 (.081)	−.031 (.081)	.022 (.082)
Per capita GDP	.288*** (.106)	.277*** (.106)	.317*** (.109)
		– Project volume – Project length – Number of heads of state – Project length X Number of ministers interaction – Year fixed effects – Country fixed effects	
Pseudo R2	0.56	0.56	0.54
N	409	406	409

Notes: Standard errors are in parenthesis. ***p<.01, **p<.05, *p<.1.

ii. Case studies

The case studies are based on 23 projects selected on the criteria of "extreme cases on the independent variable" as shown in Figure 8. The qualitative data on these projects is compiled from the narrative sections of the Implementation Completion Report (ICR) prepared by project staff and the Project Performance Assessment Report (PPAR) prepared by IEG. The analysis aims to validate the quantitative model presented in the last section and identify potential causal mechanisms through systematic analysis of the narratives.

FIGURE 8. Distribution of case study projects in the pool of education projects with PPAR
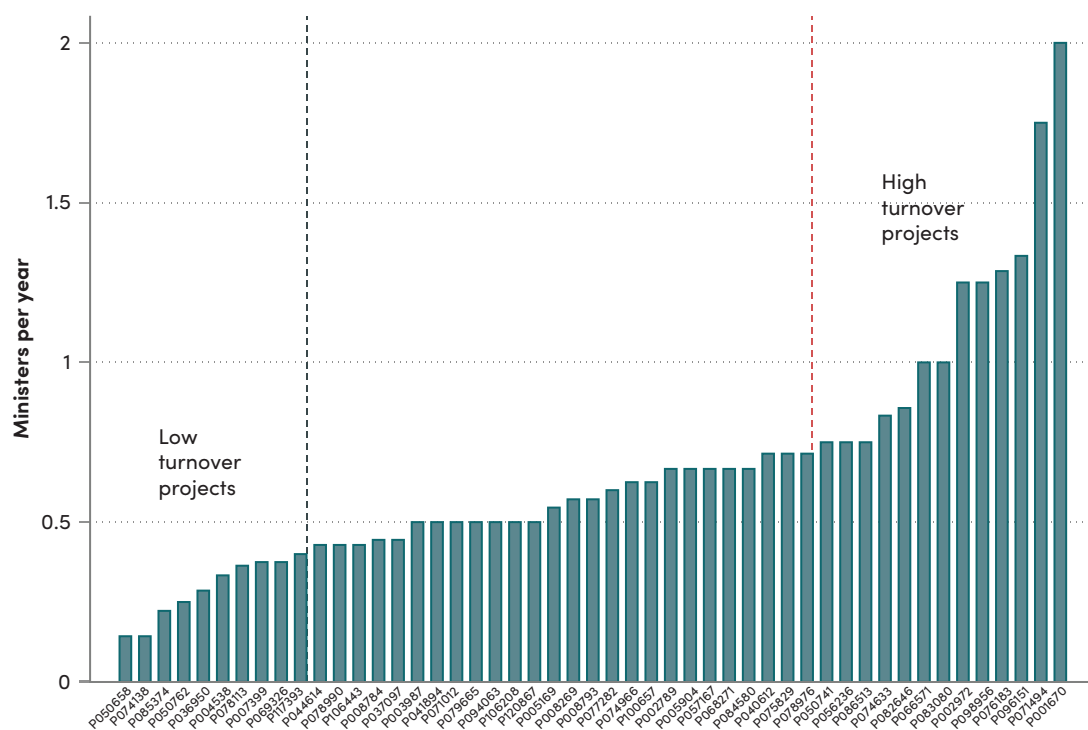
Non-spuriousness of relationship

The first step in the qualitative analysis of the project narratives is to verify whether ministerial turnover and fragmentation actually featured as a salient factor in the ICR and PPAR. This would provide some evidence as to whether the correlations estimated in the previous sections are nonspurious. Theoretically, ministerial turnover is expected to be cited in relation to lack of satisfactory project performance while ministerial continuity is mentioned in relation to good performance. However, project assessment reports are likely to focus more on factors that might have undermined performance than enumerate factors that have contributed to the success of a project. Therefore, the observability of ministerial continuity as a factor in project narratives is expected to be asymmetrical in favour of less satisfactory projects.

FIGURE 9. Salience of ministerial turnover in project assessment narratives
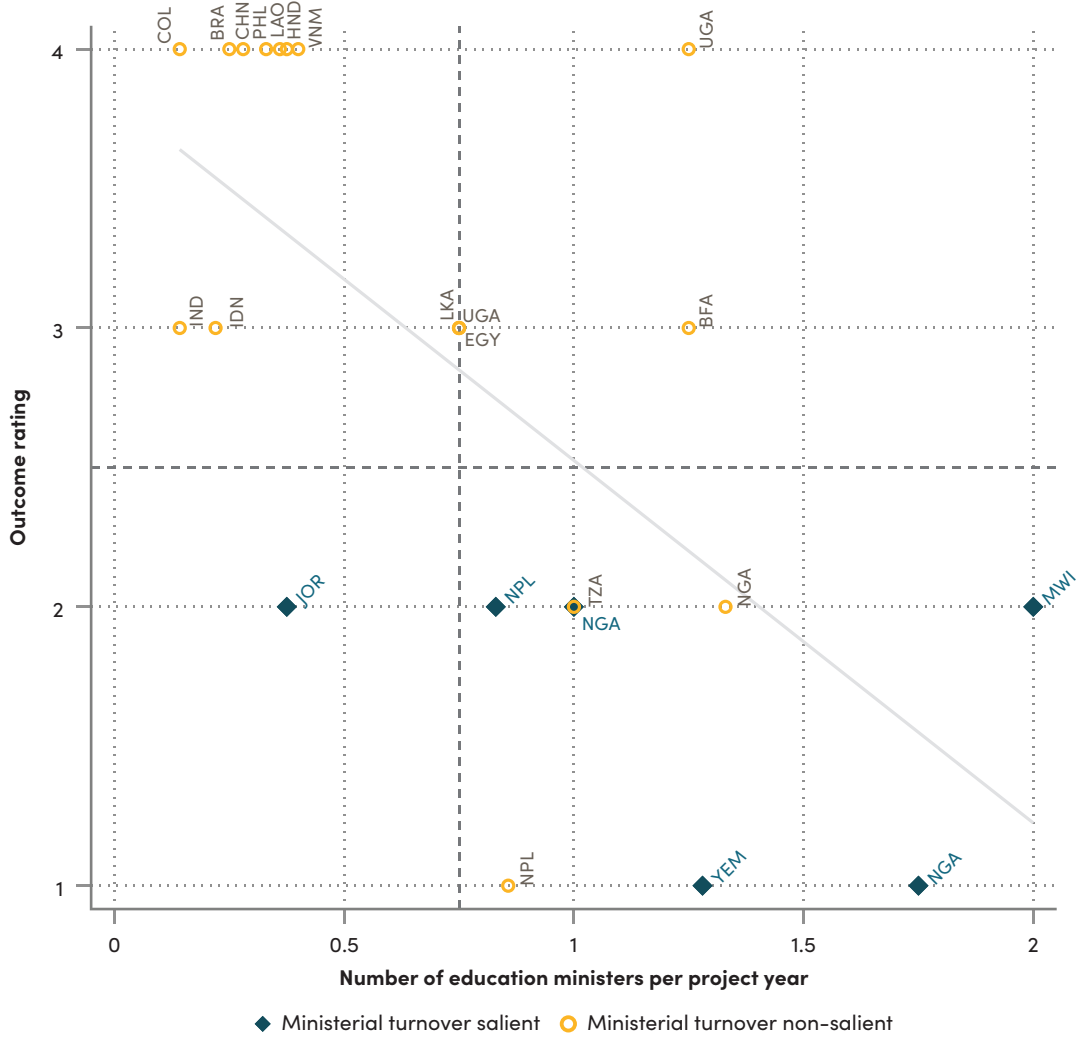
The frequent turnover of education ministers and the associated lack of continuity in senior leadership was cited in six of the 23 projects reviewed for the case studies. More importantly, it is the turnover of minsters—not fragmentation of authority across multiple ministries—that is mentioned as a factor undermining performance. Figure 9 shows that all of the cases in which ministerial turnover was mentioned are rated as less than satisfactory. Moreover, five of the six cases fall in the lower-right quadrant where low project performance coincides with above median number of education minsters. Therefore, this not only demonstrates the salience of ministerial turnover as a relevant factor in low performing projects but also validates the measurement used in the quantitative analysis.

Alternative explanations

The project assessment narratives could provide a partial overview of alterative or additional explanations for project outcome to the extent that they were deemed relevant enough by project staff or IEG evaluators to be included in the reports. This could be useful to understand the extent of omitted variable bias in the regression models in the previous section. The additional explanations could also reveal factors that moderate the impact of ministerial turnover, resulting in heterogeneity of effect. As such, the assortment of alternative explanations for poor/good performance across projects with high ministerial turnover are summarised below:

1. High turnover and low performance projects: the project assessment reports reveal that factors that are considered to be behind unsatisfactory performance, in addition to ministerial turnover, include design problems such overambitious design, political interference and capacity constraints. The quantitative model in the previous section accounts for quality of design as rated by IEG which is shown to have the biggest explanatory power of all included predictor variables. However, there could be some bias resulting from omission of the other factors.

2. High turnover and high performance projects: the alternative explanations in this category of projects are linked to factors that may have countervailed the effects of high turnover to produce satisfactory outcome. These include careful design, lessons distilled from previous similar projects implemented by the Bank, political ownership and robust implementation capacity.

Possible confounders

One of the reasons it is difficult to draw causal inference based on the type of correlational evidence generated through the quantitative analysis in the previous section is the existence of potential confounders. The project assessment reports can flag factors that may have influenced both ministerial turnover and project outcome even though they cannot be expected to provide an exhaustive treatment of potential confounders.

In this regard, among the factors that are identified in the project narratives as having impacted project outcomes, political transition and regime instability can plausibly be linked to ministerial turnover. Political transition and instability have been more common among projects that were implemented in high turnover environments (5 out of 13 projects) than among projects that were implemented under more continuity (2 out of 10 projects). The quantitative analysis partially accounts for the effect of political transition and stability by controlling for number of heads of state and median political stability index during project implementation. The fact that the coefficient of ministerial continuity declines when number of heads of state is controlled for is indicative of the confounding effect of regime transition.

Causal mechanisms

The case studies can help leverage a systematic analysis of potential causal mechanisms to further substantiate the correlational evidence from the previous section. Based on existing policy literature, I have formulated three hypotheses in relation to the causal mechanisms transmitting the impact of high ministerial turnover to project performance. Subsequently, I have pieced together relevant information on the implications of ministerial turnover from the project assessment narratives to test the hypotheses.

> H1: *Substantial revisions and reprioritizations due to new leadership slow down progress.*

The review of assessment reports shows that mid-course revisions were conducted in many of the projects with high turnover. However, there is no obvious link between ministerial turnover and revisions. In most of the cases, revisions are done to fix a problem identified through supervision. Instead of establishing mid-course revision as a channel for ministerial turnover, this finding reaffirms the role of supervision in mitigating the potential negative implications of turnover.

> H2: *Ministerial turnover leads to high turnover in senior managers at the ministry of education which in turn undermines performance.*

New ministers can sometimes reshuffle existing leadership or appoint new senior managers at the ministry. In some contexts, there could be a patronage element in the appointment of new managers by new ministers. Management turnover was mentioned in four out of six projects where ministerial turnover is salient (as opposed to two out of seven projects where ministerial turnover is not salient). There is also indication that management turnover has impacted project performance negatively.

> H3: *Overall delay in execution due to transition between ministers reduces performance.*

Ministerial turnover and associated transition can cause delays in key decisions that are needed for implementation on the government side. The case studies show that material delays occurred in four out of six projects where ministerial turnover is salient (as opposed to two out of seven projects where ministerial turnover is not salient). Notably, delays are associated with slow disbursement at take-off stage and failures in inter-ministerial coordination that is often under the purview of ministers. Such delays could have negative repercussions on project performance despite strong supervision and well-functioning project office because there are decisions that can only be taken by ministers.

The role of supervision

The quantitative analysis in the previous section shows that the quality of supervision can have a moderating effect on the correlation between ministerial turnover and project outcome. The project assessment reports offer some narrative on the quality of supervision by the Bank which can be used as a basis to substantiate the quantitative evidence.

The review is based on 13 projects with high ministerial turnover. Four (out of eight) projects which ended up with unsatisfactory outcome had reasonably robust supervision systems whereas one of the five projects with satisfactory outcome was judged to have weak supervision. This implies that good supervision is neither a necessary nor sufficient condition for a satisfactory outcome in the face of high ministerial turnover. But four of five projects that succeed despite high ministerial turnover had above average supervision. This suggests that strong supervision increases the probability of success even though it is not a sufficient condition.

One of the features of the supervision system in projects that succeeded regardless of ministerial turnover relates to where task team leaders are located and how they are replaced. The presence of an in-country TTL and smooth transition between TTLs is associated with robust supervision leading to successful outcome. Project-specific capacity building at country office level is also mentioned as a factor that has contributed to satisfactory supervision despite frequent change in TTL.

6. Conclusion

As the largest international funder of education programs in low and middle income countries, the performance of World Bank education projects has significant implication for global education aid effectiveness. In many cases, the success of major reform programs hinges on the outcome of World Bank projects which often constitute the core component of the broader reform package.

How do World Bank education projects perform in an environment where there is frequent change of ministers responsible for strategic decisions as well as administrative authorization? The turnover of education ministers can be disruptive particularly when appointment is driven by political considerations than technocratic qualifications. Moreover, frequent change in ministers deprives them of the opportunity to learn on the job which is often necessary for cabinet level appointees with limited sectoral experience. Both the quantitative and qualitative evidence presented in this paper show that there is significant negative relationship between ministerial turnover and performance of World Bank education projects.

The findings indicate that education (unlike health, for example) is one of the sectors in which instability within an incumbent administration can have considerable negative implications for the performance of aid projects. This implies that it could be useful to put in place sufficient mitigation strategies at a design stage to reduce the potential impact of ministerial turnover in education projects. Safeguards to minimize delays in disbursement and inter-ministerial coordination that might be caused by transition between successive ministers are particularly important.

The results on the role of supervision in mitigating the impact of ministerial turnover add to the existing literature showing the importance of supervision for project success. Therefore, it would be beneficial to strengthen supervision capabilities particularly in environments that are likely to

suffer from high ministerial turnover. Since one of the channels through which ministerial turnover influences project performance is through associated changes in project management on the ministry's side, the Bank could improve outcomes by proactively building the capacities of project implementation units. In this regard, assigning in-country TTLs for education projects could play a role in bolstering supervision capacities in general.

References

Bedasso, B. (2021). Is Good Monitoring and Evaluation the Secret to Success for World Bank Programs? Blog, Center for Global Development (https://www.cgdev.org/blog/good-monitoring-and-evaluation-secret-success-world-bank-programs).

Blum, J. R. (2014). What factors predict how public sector projects perform? A review of the World Bank's public sector management portfolio, Policy Research Working Paper, No. 6798, World Bank.

Bokobza, L., Krishnarajan, S., Nyrup, J., Sakstrup, C. & Aaskoven, L. (2022). The morning after: Cabinet instability and the purging of ministers after failed coup attempts in autocracies. *Journal of Politics*, **84**(3), 1437–1452.

Boone, P. (1996). Politics and the effectiveness of foreign aid. *European Economic Review*, **40**(2), 289–329.

Briggs, R. C. (2020). Results from single-donor analyses of project aid success seem to generalize pretty well across donors. *Review of International Organizations,* **15**, 947–963.

Bulman, D., Kolkma, W. & Kraay, A. (2017). Good countries or good projects? Comparing macro and micro correlates of World Bank and Asian Development Bank project performance. *Review of International Organizations,* **12**, 335–363.

Chauvet, L., Collier, P. & Duponchel, M. (2010). What Explains Aid Project Success in Post-Conflict Situations? Policy Research Working Paper, No. 5418, World Bank.

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

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

Heinzel, M. & Liese, A. (2021). Managing performance and winning trust: How World Bank staff shape recipient performance. *Review of International Organizations*, **16**, 625–653.

Honig, D. (2020). Information, power, and location: World Bank staff decentralization and aid project success. *Governance*, **33**(4), 749–769.

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

Huber, J. & Martinez-Gallardo, C. (2008). Replacing Cabinet Ministers: Patterns of Ministerial Stability in Parliamentary Democracies. *American Political Science Review*, **102**(2), 169–180.

Ika, L. A., Diallo, A. & Thuillier, D. (2012). Critical success factors for World Bank projects: An empirical investigation. *International Journal of Project Management*, **30**(1), 105–116.

IEG (2015). World Bank Project Performance Ratings Codebook. Independent Evaluation Group, September 2015.

Kilby, C. (2000). Supervision and performance: The case of World Bank projects. *Journal of Development Economics*, **62**(1), 233–259.

Kilby, C. (2015). Assessing the impact of World Bank preparation on project outcomes. *Journal of Development Economics*, **115**, 11–123.

Limodio, N. (2021). Bureaucrat allocation in the public sector: Evidence from the World Bank. *The Economic Journal*, **131**(639), 3012–3040.

Malik, R. & Stone, R. W. (2017). Replication Data for: Corporate Influence in World Bank Lending. https://doi.org/10.7910/DVN/61JSZQ, Harvard Dataverse, V1, UNF:6:gDOoN+ HM0yU7KwXxJIkD8g== [fileUNF].

Nyrup, J. & Bramwell, S. (2020). Who governs? A new global dataset on members of cabinets. *American Political Science Review*, **114**(4), 1366–1374.

Raimondo, E. (2016). What Difference Does Good Monitoring & Evaluation Make to World Bank Project Performance? Policy Research Working Paper, No. 7726, World Bank.

Shin, W., Kim, Y. & Sohn, H. (2017). Do different implementing partnerships lead to different project outcomes? Evidence from the World Bank project-level evaluation data. *World Development*, **95**, 268–284.

Wehner, J. & Mills, L. (2022). Cabinet size and governance in Sub-Saharan Africa. *Governance*, **35**, 123–141.