The Geography of Development Programs*

Beyza Gulmezoglu [§] Marta Morando [†] Eddy Zou [‡]

November 11, 2026

Abstract

International organisations devote substantial resources to improving the economic lives of the poor. Yet we know little about how these resources are distributed and why. In this paper, we provide the first systematic evidence on where development programs go *within* countries. We assemble a new dataset linking roughly 5,000 geocoded World Bank projects across 130 countries over 1995-2023 to district-level nighttime lights and complete records of the personnel assigned to each project. We find that World Bank projects flow disproportionately to richer areas: only 31% of all projects land in districts significantly poorer than their neighbours within the same province; in sharp contrast, 51% of projects exhibit positive "site selection," landing in districts significantly richer than neighboring districts. The pattern is pervasive across sectors and is not explained by population-based targeting. It does not reflect a gain in performance: projects in richer areas are no more successful according to the World Bank's own independent ratings. Supply-side factors related to World Bank personnel help explain where projects are located. Through a variance decomposition, we document that World Bank project managers account for a sizeable share of the total variation in site selection, more than year, sector, and country fixed effects combined. These findings suggest that where aid goes within countries is shaped not only by the needs of the recipient country but also by the organizational structure of the agencies that deliver it.

JEL codes: O20, M12, O19, F35
Keywords: Development Programs, Aid, Site Selection, Managers, Organizations

*We thank Oriana Bandiera, Tim Besley, Robin Burgess, Michael Callen, Jacob Moscona, Imran Rasul, Anna Tompsett, Chris Udry, and David Yang for helpful comments. We also thank Anne Brockmeyer, Jed Friedman, Clement Joubert, and Bob Rijkers for valuable insights on the functioning of the World Bank, as well as seminar participants at LSE, UCL, and Warwick. We thank STICERD for financial support.

[§] Gulmezoglu: Department of Economics, London School of Economics. Email: b.n.gulmezoglu@lse.ac.uk.
[†] Morando: Department of Economics, London School of Economics. Email: m.morando1@lse.ac.uk.
[‡] Zou: Department of Economics, London School of Economics. Email: w.zou2@lse.ac.uk.

1. Introduction

Foreign aid is one of the largest transfers of resources in the world economy. International organisations such as the World Bank allocate development assistance with the objective of reducing poverty. Yet whether these resources reach those most in need is far from clear. The poorest fifth of countries receive only between 1.7 and 5.3 percent of global aid flows in a typical year (Qian, 2015). Cross-country allocations, however, provide only a partial picture of whether resources reach the poor. Income per capita in the richest region of a country is, on average, more than four times that in its poorest region, roughly the difference between the US and South Africa in 2010 (Gennaioli et al., 2013). Assessing whether development assistance reaches the poor therefore requires examining how resources are allocated *within* recipient countries.

In this paper, we provide the first systematic evidence on where development programs go *within* countries, on a global scale. Focusing on the World Bank, the largest multilateral provider of development finance, we show that projects are systematically placed in richer areas: half of all projects land in districts that are significantly richer than their neighbours in the same province, while only one-third are targeted to poorer districts. This pattern holds in every region, every year, and every sector, and it does not reflect an equity–efficiency trade-off: projects placed in richer areas are no more successful according to the Bank's own independent ratings. Instead, where a project goes depends to a large extent on who manages it. Our contribution is therefore twofold. We characterize the geography of development programs on a global scale, and we show that the organisational structure of the agencies that deliver aid, and not only the characteristics of recipient countries, shapes who benefits from it.

We begin by assembling a global, geocoded dataset linking roughly 5,000 World Bank projects across 130 countries over 1995–2023 with satellite nighttime lights and population at the level of second-order administrative units, i.e. districts (Bomprezzi et al., 2025; Chiovelli et al., 2026; Goodman et al., 2019). We complement these sources with Historical International Development Aid (HIDA), a new dataset we assembled from the full text of 141,623 World Bank project documents almost dating back to the Bank's founding in 1945. HIDA provides a longitudinal record of World Bank project teams, tracing the roles, seniority, and project assignments of more than 46,000 individuals across nearly 20,000 projects. We link geocoded World Bank projects to the personnel listed in project documents, including country directors, practice managers, and task team leaders.

We develop a measure of "site selection" that captures whether development projects are systematically located in relatively richer or poorer areas, consistent with measures in the policy experimentation literature (Wang and Yang, 2025; Allcott, 2015). For each project, we compare recipient and non-recipient districts within the same province, using nighttime lights measured before implementation as a proxy for local economic activity (Martínez, 2022).

We implement this comparison in two ways. We first test, for each project, whether average pre-period nighttime lights differ between recipient and non-recipient districts of the same country (Wang and Yang, 2025): 59% of projects are placed in significantly brighter districts, against 30% in significantly darker ones. This comparison is, however, unconditional. Recipient districts may be concentrated in provinces that are richer for reasons unrelated to project placement, such as proximity to the capital or to World Bank offices, and project allocation may be path dependent, as in the case of scale-up and follow-up projects. Our preferred specification therefore estimates, for each project, a panel regression with province-by-year fixed effects, so that project sites are compared only with other districts in the same province and year. We find that over half (51%) of World Bank projects are located in districts that are significantly richer than their neighbours at baseline. We refer to this tendency as "positive site selection." In contrast, only about a third of all projects (31%) are placed in significantly poorer areas. Positive site selection persists when we control for population, measure luminosity per capita or in totals rather than averages, and cluster standard errors at different spatial levels.

One interpretation of our results is that the World Bank or recipient governments target aid at a higher administrative level. To test this, we repeat the analysis one administrative level up. We find that 57% of projects are allocated to significantly brighter provinces within the same country, against roughly 30% allocated to poorer ones. Aid, in short, flows disproportionately to the relatively better-off parts of poor countries. We check whether positive site selection is concentrated in a particular region or period. We find that it is prevalent across time and space: it is observed in every region in which the Bank operates, and in every year in our dataset. Its degree has nonetheless declined over time. The share of positively selected projects fell from 60% in 1995 to 53% in 2022, while the share of negatively selected projects rose from 8% to 30%. The gap between the two has narrowed but never closed, and positive selection remains the dominant pattern in every year.

The degree of positive site selection varies across sectors in the direction one would expect if some projects rely on pre-existing local capacity. The share of positively selected projects is highest in infrastructure-heavy sectors such as urban development (77%) and transport (61%), and lowest in agriculture (35%) and social protection (39%), where programs are more plausibly targeted at less developed areas. Yet positive selection is substantial in every sector, including education and health (45% and 42%), where baseline infrastructure is unlikely to be a binding constraint on implementation. The sectoral gradient is consistent with an efficiency rationale: if the returns to a project increase with local capacity, placing it in a richer district may maximise its impact, albeit at the cost of equity. The prevalence of positive selection in the social sectors, however, suggests that efficiency is unlikely to be the whole story.

We find no evidence that positive site selection is driven by strategic placement: projects in richer areas perform no better than those in poorer ones, and if anything, the opposite holds. To measure project success, we use the outcome ratings of the Independent Evaluation Group (IEG), the World Bank's evaluation arm, which reports directly to the Board of Executive Directors and is independent of the teams that design and implement projects. For each closed project, the IEG rates on a six-point scale the extent to which the project achieved its stated objectives efficiently. If projects were placed in richer areas because that is where they work best, positively selected projects should receive higher ratings. Instead, the relationship between site selection and project ratings is flat and, if anything, negative: projects placed in richer districts are no more likely to meet their objectives.

Why, then, does aid systematically flow to richer areas? Building on evidence that organisations and the individuals within them shape economic outcomes across a range of settings (Bertrand and Schoar, 2003; Rasul and Rogger, 2017; Fenizia, 2022; Best et al., 2023), we turn to the lending organisation itself and, in particular, its project managers. World Bank projects are run by Task Team Leaders (TTLs), who serve as the Bank's principal point of contact for the borrower and carry substantial discretion over where a project lands (Heinzel and Liese, 2021; Moscona, 2024; World Bank, 2013; Denizer et al., 2013).[1] We find that this discretion matters substantially. Following the logic of Abowd et al. (1999), we exploit cross-country rotation of managers to decompose the variation in site selection across projects, countries, sectors, and managers. We find that TTL fixed effects

[1] Given their role as project managers, throughout the paper we refer to them as either *Task Team Leaders* or *managers* interchangeably.

alone explain 16% of it, more than year, sector, and country fixed effects combined (12%), and far more than the fixed effects of other personnel in the project hierarchy, such as practice managers, country directors, and regional vice presidents. This result is not driven by the types of projects that managers run: project size, duration, and project ratings no longer predict site selection once TTL fixed effects are included, while TTL fixed effects add twice as much explanatory power as year, sector, and country fixed effects combined. This suggests that where aid goes within countries is shaped not only by the demands of borrowers but by the organizational structure of the agencies that deliver it.

Contribution to the literature. Our findings contribute to four strands of literature. First, we contribute to research on describing and explaining patterns of allocation of development programs (Briggs, 2018). On explanations, existing work has emphasised recipient-country factors, including local social networks (Bandiera et al., 2020a) and political connections between politicians and project locations (Dreher et al., 2019), as well as domestic stability (Mueller, 2022).

Second, we contribute to the literature on the importance of the personnel of public organisations for economic and social outcomes, including productivity (Bloom et al., 2013; Fenizia, 2022; Morando, 2026), conflict (Moscona, 2024), and policy effectiveness (Barteska and Lee, 2024; Best et al., 2023; Decarolis et al., 2020; Facchetti et al., 2026). Our results demonstrate that managers not only shape policy implementation through efforts, but also through shaping where and who will benefit from a given policy.

Third, our study also relates to a large literature scrutinizing aid: previous studies have found that aid can fuel conflict (Nunn and Qian, 2014; Moscona, 2024), be captured by elites (Andersen et al., 2022), and weaken political accountability (Guiteras and Mobarak, 2015). We provide a new explanation for why the consequences of aid could vary across contexts: the geographic placement of aid is fundamentally discretionary, therefore shaping its impacts.

Finally, we connect to research on site selection in policy experimentation and diffusion (Allcott, 2015; Wang and Yang, 2025), which mostly focus on Non-Government Organisations' program evaluations, and sub-national governments' policy rollout. In particular, our findings contribute to understanding the extent to which the geographic mismatch between places and development projects (studied in McGuirk and Nunn (2024)), could be explained by the international organisations themselves.

Structure of the paper. The remainder of the paper is organised as follows. Section 2 provides institutional background on the World Bank. Section 3 describes the data and their construction. Section 4 introduces our measure of site selection and shows that World Bank projects flow to richer areas. Section 5 documents the prevalence of site selection across administrative levels, space and time, and sectors. Section 6 examines why projects are positively selected, ruling out an equity–efficiency trade-off and documenting the role of project managers. Section 7 concludes.

2. Background

This section provides contextual knowledge on the operational aspects of the World Bank as it relates to our paper. We describe how World Bank projects are allocated across and within countries, who the managers in charge of projects are, and how project outcomes are rated.

2.1. Site selection in the World Bank

The location of a World Bank project is decided in two steps: across and within countries. Across countries, allocation follows publicly available guidelines. For the Bank's concessional arm (IDA), the decision is formula-based: each country's lending envelope is a function of the Bank's own rating of the country's policies and institutions, its population, and its income per capita, with the policy rating as the main determinant.[2] For the Bank's non-concessional arm (IBRD), lending is bounded by creditworthiness and exposure limits rather than a formula.

Within countries, by contrast, the allocation of programs is, to the authors' knowledge, not governed by a common set of formal rules.[3] Formally, the borrowing government identifies and prepares the project as part of its proposal; the Bank appraises the pro-

[2]The envelope is proportional to $\text{CPR}^{\alpha} \times \text{Population} \times (\text{GNI per capita})^{-0.125}$, where CPR is the country performance rating, built from the Country Policy and Institutional Assessment (CPIA), and $\alpha \in [2, 5]$ over 1995–2014. See Independent Evaluation Group, *Approach Paper: Evaluation of the IDA Country Allocation System* (2026), Appendix B and Table E.1, https://ieg.worldbankgroup.org/sites/default/files/Data/reports/report-approach_paper-IDA_Country_Allocations.pdf; current weights at https://ida.worldbank.org/en/financing/resource-management/ida-country-performance-ratings.

[3]Notwithstanding environmental and resettlement safeguards, which constrain where projects may *not* be built. See Operational Policy 4.01 on Environmental Assessment (https://policies.worldbank.org/en/policies/all/ppfdetail/1565) and Operational Policy 4.12 on Involuntary Resettlement (https://policies.worldbank.org/en/policies/all/ppfdetail/1572).

posal, negotiates the loan agreement, and the Board of Executive Directors approves it.[4] In practice, the Bank staff in charge of the project—the Task Team Leaders described next—shape both its design and its location (Moscona, 2024). Location is fixed gradually over the project cycle: a broad region is typically named at the concept stage, districts are listed in the Project Appraisal Document, and exact sites are often left to implementation.[5] In short, de jure the recipient government chooses where projects go; de facto, the project managers play an integral role in this decision.

2.2. Task Team Leaders

Task Team Leaders (TTLs) are the World Bank's project managers, and the staff with the most influence over an individual project. The Bank's own guidance defines the TTL as "the Bank's principal point of contact for the borrower for the project" (World Bank, 2013). A TTL is appointed once the general idea of a project has been agreed between the Country Director, the recipient government, and the Bank's sectoral staff, usually by the practice manager[6] responsible for the project (Heinzel and Liese, 2021). From then on, the TTL counsels the borrower on the design of the project, hires the team, and guides the project to the end. In preparation, TTLs lead the identification, pre-appraisal and appraisal missions to the recipient country and draft the project information documents, concept notes and appraisal documents; during implementation, they lead the Bank's implementation support and produce the six-monthly implementation reviews and the mid-term review on which

[4]Operational Policy 10.00, para. 18: "The borrower prepares the Project for which it seeks Investment Project Financing." See https://ppfdocuments.azureedge.net/1598.pdf.

[5]This description is based on our reading of project documents. The Punjab Rural Sustainable Water Supply and Sanitation Project (P169071) provides a good example. The concept-stage document (September 2018) names southern Punjab as the project area but states that "the exact location of sub-projects is, as yet, unclear" (https://documents1.worldbank.org/curated/en/377341538557978297/pdf/Concept-Project-Information-Document-Integrated-Safeguards-Data-Sheet-Punjab-Rural-Sustainable-Water-Supply-and-Sanitation-Project-P169071.pdf). The Project Appraisal Document (May 2021) lists sixteen districts and one tehsil in each: "Within each district, the most disadvantaged tehsil was chosen, and all its villages included in the Project" (para. 51 and Table 1, https://documents1.worldbank.org/curated/en/223501624327374392/pdf/Pakistan-Punjab-Rural-Sustainable-Water-Supply-and-Sanitation-Project.pdf). The order in which villages are served is then decided during implementation, "under the randomized rollout of program villages in program tehsils" (Implementation Status and Results Report, February 2025, https://documents1.worldbank.org/curated/en/099020725174017983/text/P169071-7d887d5f-0af5-4f29-9f9d-7961dd80932e.txt).

[6]Practice managers are responsible for overseeing regional sector portfolios, managing technical teams, and directing client country dialogue. The role was previously known as "sector manager".

restructuring decisions rest (World Bank, 2013; Heinzel and Liese, 2021).[7] This gives TTLs substantial discretion over location, as outlined in the literature (Moscona, 2024) and confirmed in qualitative interviews we conducted with current and former World Bank staff.

The way TTLs are assigned to projects limits the concern that they are matched to richer or more dynamic areas. Assignment is determined by TTLs' current availability and sectoral expertise: unlike Country Directors, TTLs do not specialize in countries but in one or more sectors. This organizational structure generates substantial cross-country rotation of TTLs, who move to a new assignment every three to five years (Denizer et al., 2013; Kilby, 2013; Heinzel and Liese, 2021).[8] Crucially, at the time a project is assigned to a TTL, only a general description of the project is known, and details such as its location and design are still undecided (Moscona, 2024). Together, these features limit concerns that managers are systematically assigned to richer or poorer areas, or projects that are positively correlated with local economic activity.

2.3. IEG project ratings

Our measure of project success is the outcome rating assigned by the World Bank's Independent Evaluation Group (IEG), which is produced independently of the team that ran the project. The outcome rating measures "the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently," and combines three criteria: the relevance of the project's objectives and design, the extent to which the objectives were achieved (efficacy), and efficiency.[9] Alongside the outcome rating, the IEG rates the quality of monitoring and evaluation, organization, preparation, and execution, thus providing a detailed measure of project management quality.[10]

The rating process has two stages. First, the Bank unit responsible for the project conducts a self-evaluation: in consultation with the project manager, the unit appoints an evaluation team, who provide comments and feedback. This process ends with an Imple-

[7] The decisions they do not take themselves—additional financing, submission to the Board, and Board approval—rest with the Country Director, the Regional Vice President and the Board, respectively, but TTLs draft the documents on which those decisions are based (Heinzel and Liese, 2021).

[8] The rotation follows the Bank's so-called 3–5–7 rule: staff must remain in an assignment for at least three years, are encouraged to seek reassignment to another vice-presidential unit after five, and are actively rotated by regional managers after seven (Independent Evaluation Group, 2015, p. 62).

[9] Independent Evaluation Group, *World Bank Project Performance Ratings Codebook* (September 2015), p. 14, https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ieg-wb-project-performance-ratings-codebook_092015.pdf

[10] IEG Codebook, p. 17–18, on the Bank Performance and Quality at Entry ratings.

mentation Completion and Results Report (ICR), which assesses the project's performance and assigns ratings on a six-point scale, from 1 (highly unsatisfactory) to 6 (highly satisfactory). Second, the IEG validates the self-evaluation and assigns its own rating. It prepares an Implementation Completion and Results Report Review, an independent, desk-based validation of the ICR's evidence, content, and ratings.[11] Based on the ICR, the lending agreement and appraisal documents, and interviews with the task team leader, the IEG assigns its own ratings using the same evaluation criteria used by the World Bank, and these may differ from the ICR's.[12]

The IEG's independence is institutional. It operates under different incentives than other units of the World Bank: it is fully independent of management and reports directly to the World Bank Group's Executive Board.[13] To further ensure independence, the IEG conducts Performance Assessment Reports on a random subset of approximately 20–25 percent of projects (Limodio, 2021).[14]

3. Data construction

We assemble a global dataset linking World Bank development programs, nightlight luminosity, and population. All the datasets we use are geo-coded, allowing us to link each to spatial polygons of administrative units. We use the Geocoded Official Development Assistance Dataset (GODAD) for development programs (Bomprezzi et al., 2025), Chiovelli et al. (2026) for nightlight data, and the GeoBoundaries dataset for data on administrative borders and district-level population counts (Runfola et al., 2020). Our resulting dataset spans 130 countries and three decades (1995-2023).

Finally, we go beyond existing records and construct a new dataset, Historical International Development Aid (HIDA) capturing the full history of project appraisal, project design, project implementers (both contractors and World Bank personnel), across nearly 20,000 projects implemented by the World Bank from 1950 to today. Our motivation for doing so is twofold. First, GODAD extends AidData's coverage but records no information on task team leaders, an integral component of our analysis (AidData, 2017); pairing HIDA with GODAD retains the wider sample. Second, HIDA goes beyond both what Aid-

[11] See https://ieg.worldbankgroup.org/ieg-search-icrr.

[12] IEG codebook (p.2): "the IEG comes to an independent judgment on its performance ratings, which may differ from those in the ICR."

[13] IEG Codebook, p. 1.

[14] The IEG Codebook (p. 2) puts the share at 20–25 percent.

Data and GODAD deliver, and provides a rich account of the information on the identity and composition of teams implementing and designing each of the development projects.

In ongoing work, we're extending the dataset in three ways. First, we're constructing a panel dataset of World Bank staff recording titles, projects, and country placements over the years, linked with project outcome scores. This will serve to characterize which types of project managers are more or less likely to place World Bank projects in poorer (or richer) areas, and link management styles to career outcomes. Second, we're compiling citations across different project documents to characterize "project learning" within the World Bank. Third, we're extracting information on counterfactual locations the World Bank considered for project implementation to use in the causal estimation of the effect of WB project allocation.

3.1. The Geocoded Official Development Assistance Dataset (GODAD)

We leverage GODAD containing subnational geo-coded data on World Bank projects implemented between 1995 and 2023 (Bomprezzi et al., 2025).[15] The dataset offers a rich set of information for the projects, including start year, disbursement amounts, main lenders, sectors, among more.

Most crucial to our analysis, for each project, they provide information on the project geo-location, its latitude and longitude, together with a precision code indicating the spatial accuracy of that location.[16] We restrict our sample to projects with precision code 3 and below to ensure that project locations can be reliably assigned at least to the ADM2 level, which is the primary level of our empirical analysis.[17] We exclude multicountry and emergency projects from our analysis, leaving us with 4,731 geo-coded World Bank projects, covering 130 countries, spanning nearly three decades. Figure 1 demonstrates the geographic coverage our dataset.

Table 1 summarizes the characteristics of the projects in our sample. The median project reaches five districts in two provinces, and at least a quarter are confined to a

[15] Data available at https://godad.uni-goettingen.de.

[16] Precision codes range from 1 (exact coordinates, e.g., a populated place) to 9 (location known only at the country level), with intermediate codes identifying locations known to lie within a given ADM2 (code 3) or ADM1 (code 4) unit (Bomprezzi et al., 2025).

[17] Throughout the paper, we will refer to ADM2 units as "districts" and ADM1 as "provinces," where ADM1 are broader geographical units than ADM2. ADM2 are the smallest contiguous spatial units consistently recorded in the project data, and they are equivalent to a US county.

Table 1: Summary Statistics: Project Characteristics

	N	Mean	SD	Min	P25	Median	P75	Max
No. of districts (ADM2)	4,731	10.3	18.2	1	1	5	12	387
No. of provinces (ADM1)	4,731	4.9	6.5	1	1	2	6	78
Disbursements (m, 2014 USD)	3,708	110.1	171.2	0.1	21.4	52.1	124.2	3,184.5
Duration (years)	4,704	6.4	2.6	0.0	5.1	6.5	7.9	16.9
IEG outcome rating (1–6)	2,364	4.03	1.03	1.00	3.75	4.00	5.00	6.00
Start year	4,731	2010.2	7.6	1995	2005	2011	2017	2023

Notes. The table summarizes project characteristics and uses one observation per project. The sample matches what we use for the site selection analysis outlined in Section 4, and is the 4,731 single-country World Bank projects geo-coded at precision code 3 or below (GODAD), excluding disaster and emergency projects; each row uses all non-missing observations. Districts and provinces are the geoBoundaries ADM2 and ADM1 units in which the project has at least one location. Disbursements are GODAD total project disbursements in millions of constant 2014 US dollars, reported for closed projects only: GODAD records disbursements to date, so active projects are only partly disbursed. Duration is the number of years between board approval and closing. The IEG outcome rating runs from 1 (highly unsatisfactory) to 6 (highly satisfactory) and is assigned only to closed, evaluated projects; projects evaluated twice receive the average of the two ratings. Start year is the year of board approval.

single district. Projects are large in terms of cost and long-lived. Among closed projects, the median disburses $52 million (in constant 2014 dollars) and the mean $110 million, and the typical project runs for six and a half years between board approval and closing. Half of the projects in our sample, 2,364, have been evaluated by the IEG. Their average outcome rating is 4.0 on the six-point scale, corresponding to "moderately satisfactory," with a standard deviation of one point. Project start years span 1995 to 2023, with a median of 2011.

3.2. Nightlight Data

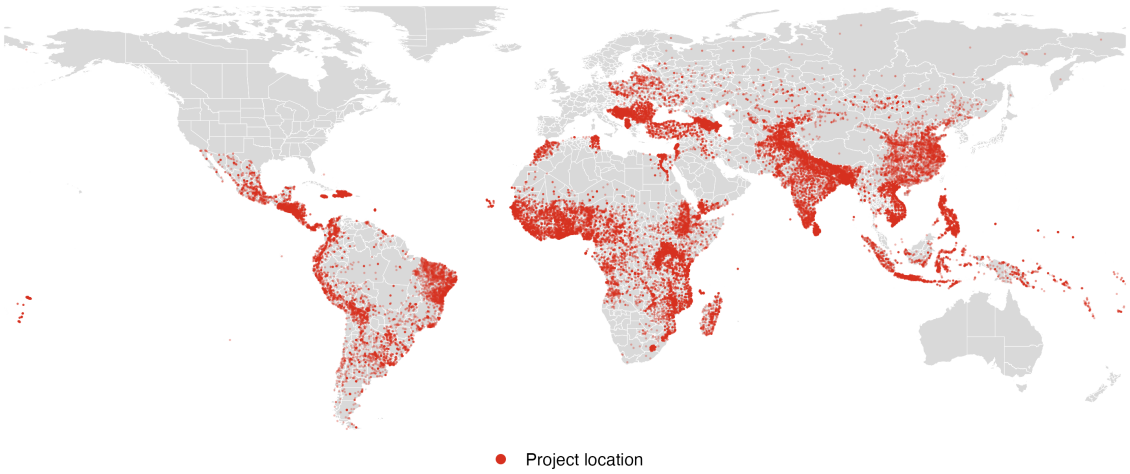
To capture local economic activity, we use harmonized nighttime light (NTL) data from Chiovelli et al. (2026).[18] NTL data are an attractive proxy for economic activity because they are recorded by satellite rather than self-reported, making them less prone to manipulation than official statistics, and because they are available at fine subnational resolution (Martínez, 2022; Henderson et al., 2012).

We prefer Chiovelli et al. (2026) over other available sources for three reasons (Li et al., 2020; Goodman et al., 2019). First, it corrects the three well-known flaws of the pre-2013 DMSP-OLS imagery: inconsistent calibration across the six satellites, top coding of the brightest pixels (digital numbers are capped at 63), and *blooming*, whereby light

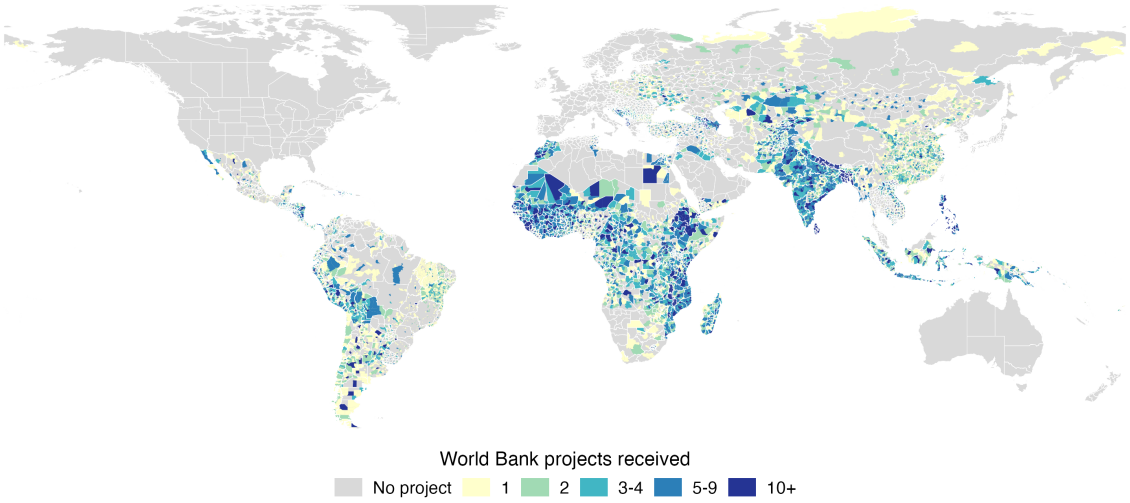
[18]Data available at https://zenodo.org/records/20558766.

Figure 1: Spatial Coverage of the GODAD Project Sample

(a) WB project locations



(b) Number of projects received per district



Notes. Figure shows the spatial coverage of the GODAD project sample. The sample is the 5,053 geo-coded projects with precision code 3 and below, implemented between 1995 and 2023 across 14,616 districts in 130 countries. The grey background is the world's ADM2 districts. ADM1 borders are drawn in white. Panel A plots one red dot per project location. Panel B shades each district by the number of distinct projects it receives. Districts in grey never receive a World Bank project in our dataset.

spills from lit pixels onto neighbouring dark ones. Second, it links the corrected DMSP data together with the sharper post-2013 VIIRS series into a single comparable panel covering 1992–2023, whereas most products rely on only one sensor and so cannot track change over three decades. Third, Chiovelli et al. (2026) show that this adjusted series predicts local development—wealth, education, and electrification across African DHS surveys and censuses in Mozambique, Indonesia, and India—and outperforms the unadjusted series of Li et al. (2020), with the largest gains over time and at fine spatial resolution: exactly the within-country, ADM2-level variation our design exploits.

The raw data are distributed as annual global rasters of DN luminosity at the DMSP source resolution (30 arcseconds, roughly 1 km at the equator). We overlay each annual raster on second-level administrative boundaries (ADM2) from the geoBoundaries CGAZ dataset and average the DN of all pixels falling within a district to obtain its *mean nightlight intensity*. This yields a district-year panel spanning 1992-2023.[19]

Throughout, our main dependent variable is the logarithm of one plus a district's mean nightlight intensity, $\log(1 + \overline{Light}_{d,t})$. The distribution of mean luminosity is highly right-skewed and concentrated at the lower bound: after the blooming and top-coding corrections, which raise the share of unlit pixels, a sizeable share of district-years record little or no light. We therefore follow Michalopoulos and Papaioannou (2013, 2014) in log-transforming average nighttime light intensity. So as not to lose the observations with no recorded nighttime light, we add a constant before taking the logarithm, which keeps these district-years in the sample rather than dropping them. Retaining these observations is sensible: a district recording no light plausibly still hosts economic activity that falls below the sensors' detection limit or is removed by the blooming correction, rather than none at all (Hodler and Raschky, 2014).

3.3. Historical International Development Aid

We construct a new dataset, Historical International Development Aid (HIDA), covering every World Bank project whose documentation is publicly available, from 1950 to today. The raw material is the full text of the Bank's own project documents, downloaded through the World Bank API: 141,623 texts for roughly 20,000 projects, spanning appraisal documents, completion reports, environmental assessments, among various other document types. We build each HIDA dataset in three steps. First, a keyword screen isolates the pas-

[19] Figure A.1 summarizes the geographic coverage of the nightlight dataset over the last 25 years.

sages of each document that can carry the target information. Second, a large language model reads the screened passages and returns a structured record; every extracted value carries the verbatim sentence that supports it, so it can be audited against the source. Third, we validate the output against existing datasets and hand-coding. Appendix C details the corpus and the extraction method.

The first component is a project organigram. Development projects are run by teams, not only by the task team leader. We extract every named Bank staff member from the documents, with their role, seniority, and the stage of the project cycle at which they appear, and link people across projects with a person identifier. The dataset names 46,606 distinct people across 18,494 projects and identifies the task team leader for 87 percent of projects with an appraisal document, together with the practice manager, country director, and regional vice president, and whether the appraisal team is retained through supervision (Appendix C.3). It is this component that we use in Section 6.2 to decompose the variation in site selection across the project hierarchy.

Table 2 summarizes the staff we observe at each layer of the project hierarchy. Each layer oversees a broader portfolio than the one below it. The median practice manager oversees three projects, the median country director four, and the median regional vice president 18, spanning seven countries. We observe a large number of TTLs on more than one project. In the analysis sample, 1,048 TTLs lead at least two projects, together accounting for 3,172 projects, or 69 percent of the projects with an identified TTL. Moreover, TTLs move across countries: four in five TTLs with repeated projects work in at least two countries, consistent with the rotation described in Section 2.2. Together, these features allow us to estimate the contribution of each layer of the hierarchy, and of TTLs in particular, to the variation in site selection in Section 6.2.

HIDA allows us to extend our analysis both in scope and in depth. First, AidData's coverage stops in 2014 (AidData, 2017), whereas GODAD continues until 2023 (Bomprezzi et al., 2025). HIDA allows us to use the more extensive option, GODAD, by supplying the absent personnel information, which we can link with project IDs. Second, HIDA records information on the full project team beyond the task team leader, including those higher up than the TTL such as the regional vice president, country director, and practice managers.

Three further components are not yet in use but underpin our ongoing work. First, we're constructing a panel dataset of World Bank staff to map the complete career transitions of approximately 46 thousand people extracted from 52 thousand project documents

Table 2: Summary Statistics: World Bank Staff

	People	Projects per person	Countries per person	Staff on 2+ projects	Share of projects led by repeat staff	Share of repeat staff in 2+ countries
Panel A: Analysis sample						
Regional vice presidents	78	54.8 (18)	10.3 (7)	61	100%	97%
Country directors	422	10.3 (4)	2.5 (2)	294	97%	74%
Practice managers	747	6.1 (3)	4.2 (2)	506	95%	95%
Task team leaders	2,473	1.9 (1)	1.5 (1)	1,048	69%	78%
Panel B: All HIDA projects						
Regional vice presidents	139	76.2 (4)	10.4 (3)	94	100%	89%
Country directors	722	15.6 (4)	3.0 (2)	482	98%	76%
Practice managers	1,236	9.9 (4)	5.6 (3)	853	97%	96%
Task team leaders	4,916	2.7 (2)	1.9 (1)	2,866	84%	78%

Notes. The unit of observation is a staff member in a given role. For each project and role, we assign the staff member named most often in that role in the appraisal documents. Panel A covers the 4,659 projects in Table 1 with staff data; Panel B covers all 13,261 World Bank projects in HIDA that name at least one of the four roles at appraisal. *Projects per person* is the number of projects on which the staff member holds the role, and *countries per person* is the number of distinct borrowing countries across these projects, counting a multi-country project as one country; both report the mean, with the median in parentheses. Repeat staff are those observed on at least two projects. The last two columns report the share of projects whose staff member in that role is a repeat staff member, and the share of repeat staff observed in at least two countries. Person identifiers merge spelling variants of the same name. Residual variants may split an individual across identifiers, so the number of people is an upper bound.

covering years 1950 to today. This dataset can serve to estimate project manager fixed effects along various project outcomes and allocation decisions, building on previous work in the literature (Bertrand and Schoar, 2003; Limodio, 2021). More specifically, we plan to characterize which types of project managers are more or less likely to place World Bank projects in poorer (or richer) areas, and link management styles to career outcomes. The second records project learning: documents routinely cite earlier projects, whether as a phase they continue, a pilot they scale up, or a lesson they draw. We collect each citation as a link from the later project to the earlier one, coded for why it is cited and whether the cited project is described as a success or a failure, giving 75,690 citations in total. This will let us trace how project designs and locations propagate through the Bank's own record of what has worked. The third records counterfactual locations. Appraisal documents often discuss candidate sites that were considered and not chosen: rejected plots, districts ranked below a targeting cutoff, areas dropped because another donor was already present. We collect each such location with the stated reason for rejection, coded as technical (terrain, land, access) or economic and social (cost, returns, local opposition), yielding 3,678 rejected sites and 27,390 ranked-out areas. These locations give us a set of places the Bank itself judged plausible but did not select, which can be used as plausible

counterfactuals against which selected sites can be compared (Greenstone et al., 2010).

4. World Bank Programs Flow to Richer Areas

In this section we document that World Bank projects are systematically placed in relatively richer areas *within* countries, as proxied by nighttime light luminosity. We refer to this pattern as "positive site selection." We first describe how we measure site selection, and then present the results of two complementary tests: a simple comparison of means and a panel regression with province-by-year fixed effects. We close the section with robustness checks.

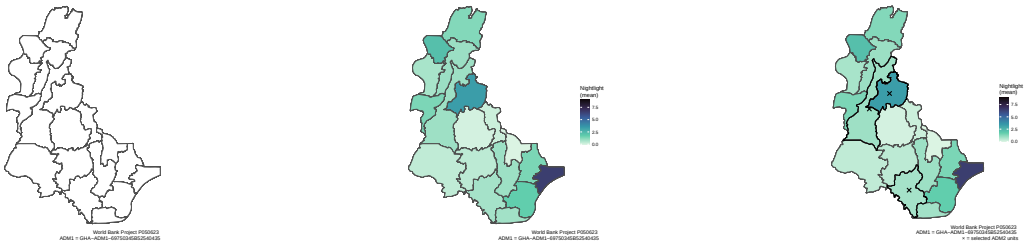
4.1. Measuring Site Selection

We conceptualize site selection as a statistically significant relationship between the project location and pre-existing local economic activity (Wang and Yang, 2025; Allcott, 2015). We link geo-coded World Bank project data matched to second-level administrative units (ADM2) (Bomprezzi et al., 2025) with grid-level data on nightlight luminosity averaged over administrative unit spatial polygons (Chiovelli et al., 2026), allowing us to empirically characterize aid allocation patterns on a global scale. In our analysis, we exclude multi-country projects and projects that constitute post-disaster/emergency responses. Our resulting dataset spans 4,731 World Bank projects across 130 countries and nearly three decades (1995–2023).

The core idea behind our site selection analysis is to compare districts that received a given WB project, to those that did not, in terms of their luminosity. Indeed, in practice, we map nightlight luminosity measures and geocoded WB programs to district-level boundaries for each project in our dataset. Then, we compare the districts that received the project to those that did not, in terms of their baseline economic conditions measured as the average luminosity in the years prior to project allocation. Figure 2 demonstrates how our measure works in practice. As an illustrative example, the figure focuses on the Volta region of Ghana that received a "Road Sector Development Project" in year 2001 (project ID = P050623).[20] Panel (a) plots the district boundaries, Panel (b) links the nightlight data to each district and illustrates the average luminosity, and finally Panel (c) indicates to which districts within the Volta region the aforementioned project was allocated to.

[20]The details of World Bank projects are available online at https://documents.worldbank.org/.

Figure 2: Illustration of the Site-Selection Measure



(a) ADM2 boundaries (b) Baseline nightlight (c) Selected ADM2 units

Notes. The panels illustrate how project sites are compared with non-project districts within the same province. The figures focus on the Volta region (ADM1) of Ghana, which received a "Road Sector Development Project" in year 2001 (project ID = P050623). The example first shows district (ADM2) boundaries, then pre-project nighttime light intensity, and finally the districts selected for project implementation.

4.2. A Simple Means Test

We begin our analysis with a simple exercise and test directly whether recipient and non-recipient districts differ in pre-implementation economic activity, following Wang and Yang (2025). For each project, we restrict the sample to the districts of the recipient country and to the years *before* the project's approval, and test the null hypothesis of no difference in average pre-period nighttime lights between recipient and non-recipient districts,

$$H_0: \quad \mathbb{E}[Y_d \mid \text{Treated}_d = 1] = \mathbb{E}[Y_d \mid \text{Treated}_d = 0]\,, \tag{1}$$

using a two-sample Welch (unequal-variance) *t*-test. Notice that since we restrict our analysis to single-country projects, i.e. we exclude projects that span multiple countries from our analysis, this comparison is *within* country. Furthermore, we use nightlight data up until the project start year, provided by GODAD, in order to capture *baseline* economic activity. The resulting project-level statistic is

$$t_i = \frac{\bar{Y}_i(1) - \bar{Y}_i(0)}{\sqrt{\hat{S}_i^2(1)/n_{i,1} + \hat{S}_i^2(0)/n_{i,0}}}, \tag{2}$$

where $\bar{Y}_i(1)$ and $\bar{Y}_i(0)$ are the sample means of pre-period average nightlight luminosity among recipient and non-recipient districts, $\hat{S}_i^2(1)$ and $\hat{S}_i^2(0)$ the corresponding sample variances, and $n_{i,1}$, $n_{i,0}$ the numbers of district-year observations in each group. A positive t_i indicates that a project is placed in districts that are brighter, and hence more economi-

cally developed, than non-recipient districts prior to implementation. As in Wang and Yang (2025), conducting a separate test for each project is conservative relative to a pooled test with multiple treated districts. While our baseline outcome is the logarithm of one plus a district's average nighttime light intensity; in Section 4.4 we show the results are robust to measuring luminosity as total or per-capita intensity, and in levels rather than logs.

Testing the pre-period nightlight luminosity between recipient and non-recipient districts reveals that most World Bank projects exhibit "positive site selection": the districts that receive World Bank projects are, on average, richer than those that do not. Across 4,699 projects[21], the average t-statistic is 7.0,[22] and 65% of projects are allocated to brighter areas, as proxied by nighttime light luminosity. Of these, roughly 90% mark statistically significant differences at the conventional 95% level, so that 59% of all projects are significantly positively selected, against 30% that are significantly negatively selected. Figure 3 plots the distribution of these project-level t-statistics, with mass concentrated well to the right of the $+1.96$ threshold.

4.3. Panel Fixed Effects Regressions

We next turn to a panel fixed effects regression to characterize the direction of subnational aid flows more precisely. Indeed, the simple means test above compares recipient and non-recipient districts unconditionally: it pools all pre-project years and does not net out common time shocks or systematic differences across provinces. For instance, recipient districts may be concentrated in provinces that are richer for reasons unrelated to project placement, such as proximity to the national capital or WB offices. Moreover, project allocation may exhibit spatial path dependency, most directly in the form of scale-up or follow-up projects, but also as a result of establishing local ties and gaining organizational familiarity with the setting.

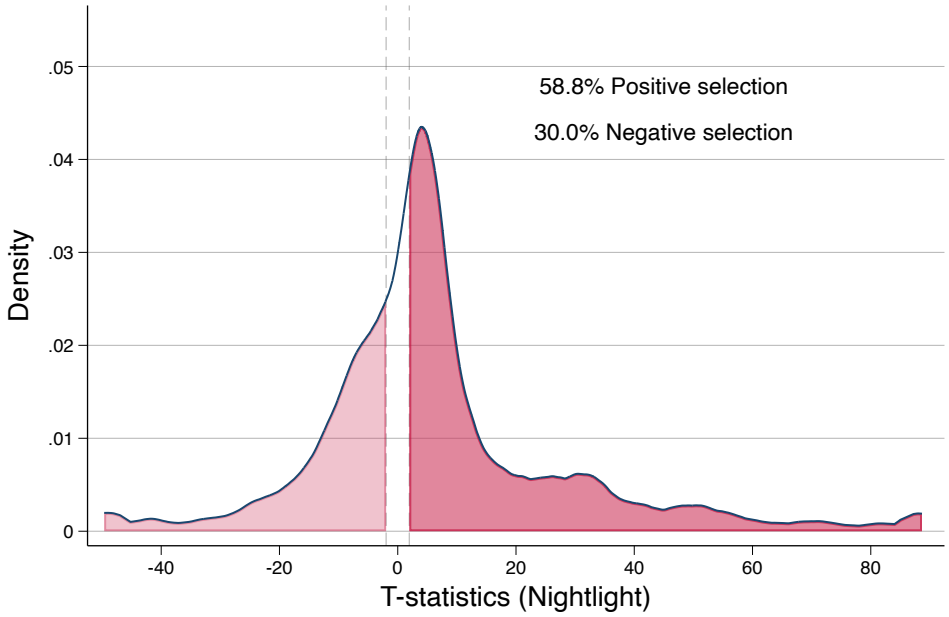
To address these concerns and isolate a cleaner relationship between pre-existing economic activity and World Bank project allocation, we proceed with a regression framework. Specifically, for each project i, we restrict the dataset to its recipient country[23] and to the years before its start (to capture baseline economic development), and we estimate:

[21] The number differs from the baseline of 4,731 projects as the two-sample test is undefined when a group contains fewer than two district-year observations.

[22] We winsorize the project-level t-statistics at the 1st and 99th percentiles throughout.

[23] We restrict our analysis to projects operating in single countries.

Figure 3: Distribution of T-Statistics from Simple Means Test



Notes.The figure shows the distribution of the project-level t-statistics from Equation 2, testing the null hypothesis that project and non-project locations face the same pre-implementation economic conditions, measured through average nighttime light intensity. The dashed vertical lines correspond to -1.96 and 1.96, the critical values for a two-sided test at $\alpha = 5\%$. The values are winsorized at the 1st and 99th percentiles of the distribution. The shaded regions mark the density beyond these thresholds, and the reported percentages give the shares of significantly negatively ($t < -1.96$) and positively ($t > 1.96$) selected projects among those with an identified estimate.

$$Y_{d,p,t} = \alpha + \beta \cdot \text{Treated}_{d,p,t} + \delta \cdot \text{Previously Treated}_{d,p,t} + \gamma_{p,t} + \varepsilon_{d,p,t}, \tag{3}$$

where d indexes districts, p indexes provinces, and t years. $\text{Treated}_{d,p,t}$ is an indicator for the district(s) that receive the project. $Y_{d,p,t}$ is the log of one plus the district's average nightlight intensity in the years *before* project implementation. Importantly, we add province-by-year fixed effects, $\gamma_{p,t}$, thereby exploiting variation *only in the same province and the same year*. The within-province comparison alleviates concerns about other underlying geographic differences between selected and non-selected project locations, such as distance to World Bank offices or to major cities. Moreover, they absorb any province-specific shock in a given year; such as a regional boom, a drought, or a province-wide policy change. Finally, to account for potential path-dependency within districts we include a dummy variable, $\text{Previously Treated}_{d,p,t}$ which indicates whether districts had already

received a World Bank project in an earlier year.

The estimated coefficient $\hat{\beta}$ captures differential economic conditions between project and non-project sites. From these regressions, we obtain the project-level *t*-statistic on $\hat{\beta}$, computed using heteroskedasticity-robust (White) standard errors, testing the null hypothesis that prior economic conditions are indistinguishable between the project and non-project sites. Values of the *t*-student statistics above 1.96 indicate that projects are located in significantly more economically advantaged areas, measured prior to project allocation, than non-project locations. Throughout the paper, we define this pattern in the data as site selection. In total we extract 4,438 *t*-statistics—one for each project for which $\hat{\beta}$ is identified within the province-by-year comparison. Note that, once again, this procedure is conservative as a pooled test with multiple treated areas would reject the null hypothesis more often (Wang and Yang, 2025).

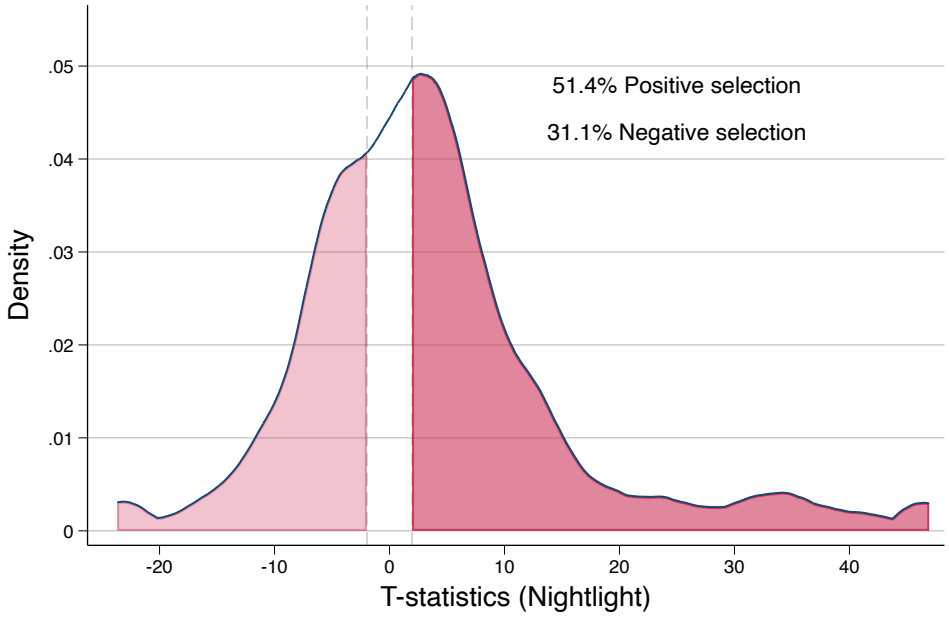
The regression results confirm that the majority of World Bank projects exhibit positive site-selection: 51% of World Bank projects are allocated to economically advantaged areas, while only 31% are targeted towards economically disadvantaged places. Project sites thus tend to have significantly higher pre-implementation nightlight intensity than non-project sites within the same province. Figure 4 plots the distribution of the project-specific *t*-statistics comparing the economic conditions between areas with a World Bank project and those without, as described in Eq. 3. The dashed gray lines correspond to $\pm$ 1.96, marking where one would reject the null hypothesis of no site selection.

4.4. Robustness Checks

We conduct a series of robustness checks to ensure that the prevalence of positive selection amongst WB projects is not spurious. First, we check whether restricting to populous areas, controlling for district-level population in 2001, or measuring luminosity per capita eliminates positive selection. Second, we repeat our analysis changing the dependent variable from logs to levels, and from average nightlight to sums. Finally, we swap our robust standard errors for standard errors clustered at the district and province level. The pattern of positive site selection remains dominant across all these checks, with varying magnitudes.

Uninhabited vs unlit. We use the logarithm of one plus average nightlight luminosity per district as our main dependent variable. One worry in doing so is that we might be conflating less-developed and *uninhabited* districts. Put differently, one would not expect

Figure 4: Distribution of T-Statistics($\hat{\beta}$) from Regression Analysis



Notes. The figure shows the distribution of the project-level t-statistics on $\hat{\beta}$ from Equation 3, testing the null hypothesis that project and non-project locations face the same pre-implementation economic conditions, measured through average nighttime light intensity, estimated with province-by-year fixed effects and heteroskedasticity-robust standard errors. The dashed vertical lines correspond to -1.96 and 1.96, the critical values for a two-sided test at $\alpha = 5\%$. The values are winsorized at the 1st and 99th percentiles of the distribution. The shaded regions mark the density beyond these thresholds, and the reported percentages give the shares of significantly negatively ($t < -1.96$) and positively ($t > 1.96$) selected projects among those with an identified estimate.

the World Bank to allocate programs to unpopulated areas, yet such areas are also dark. To ensure this is not driving our results, we re-estimate the main specification on progressively smaller samples, first restricting to districts with known 2001 population (the earliest year available in our demographic data) and then dropping those with fewer than $1{,}000$, $5{,}000$, and $10{,}000$ inhabitants. Because second-level administrative units are typically populous, these filters bind rarely and leave the result essentially unchanged: the positive share holds at roughly 51% across every threshold. As a further check, shown in Figure A.2, we account for population directly in two ways: one specification controls for log population in 2001, and another replaces the outcome with per-capita luminosity. Positive selection remains the dominant pattern, if somewhat attenuated (around 44% of projects positively selected under the per-capita outcome), indicating that the tilt toward brighter areas is not merely a tilt toward more populous ones.

Nightlight mean vs sum. To show our results are not sensitive to the choice of dependent variable, we repeat the analysis using total rather than average intensity, and in levels rather than logs. As Figure A.3 shows, positive selection is robust throughout and, if anything, stronger for total luminosity, rising to 54% (and 62% in logs), compared with 48% for average intensity in levels and 51% for our main log-average outcome.

Clustering SEs. Our preferred specification uses heteroskedasticity-robust standard errors. Clustering at the province or district level mechanically raises the share of null results—the significantly positive share falls from 51% under robust standard errors to 36% when clustering by province and 28% when clustering by district—but within a single country there are often only a handful of provinces, so the cluster-robust variance estimator is unreliable when clusters are few. We therefore report robust standard errors as our baseline and treat the clustered results as a conservative bound.

5. The Prevalence of Site Selection

The previous section establishes that half the World Bank projects exhibit positive site selection, landing in districts that are richer than their neighbours in the same province. In this section we ask how general this pattern is. We show that positive site selection is observed in every continent and in every year of our sample, and that our findings hold one administrative level up, across provinces within countries. We close by showing that positive site selection is present in every sector, and that its variation across sectors is informative about why aid flows to richer areas.

5.1. Site Selection across Space and Time

We begin by documenting that positive site selection is not confined to a particular region or period. Figure 5a averages the project-level t-statistics over the projects each province receives. Three facts emerge. First, positive site selection is a global phenomenon, observed in every region the Bank operates in. Second, there's geographic variation in the degree of site selection: positive averages dominate across much of Sub-Saharan Africa, the Middle East, China, and Brazil, while negative averages are concentrated in a smaller set of places, such as Argentina, Egypt, Zimbabwe, and much of Russia. Finally, substantial heterogeneity remains within countries. Where projects land therefore varies substantially *within* the same borrower, a pattern that the borrower's characteristics alone cannot explain and to

which we return in Section 6.2.

We next look at the evolution of our site selection measure over time. Figure 5b plots the mean t-statistic by project start year, and shows that site selection has persisted throughout the past three decades. The mean lies above 1.96 in every year from 1995 to 2022 except 2009 (1.77), and shows no clear trend; if anything, a somewhat cyclical pattern emerges. Figure 5c shows that the share of significantly positively selected projects fell from 59% among projects starting in 1995–2004 to 49% among those starting afterwards, while the share of significantly negatively selected projects roughly doubled, from 18% to 36%. Positive selection nonetheless remains the dominant pattern in every year.

5.2. Site Selection across Provinces

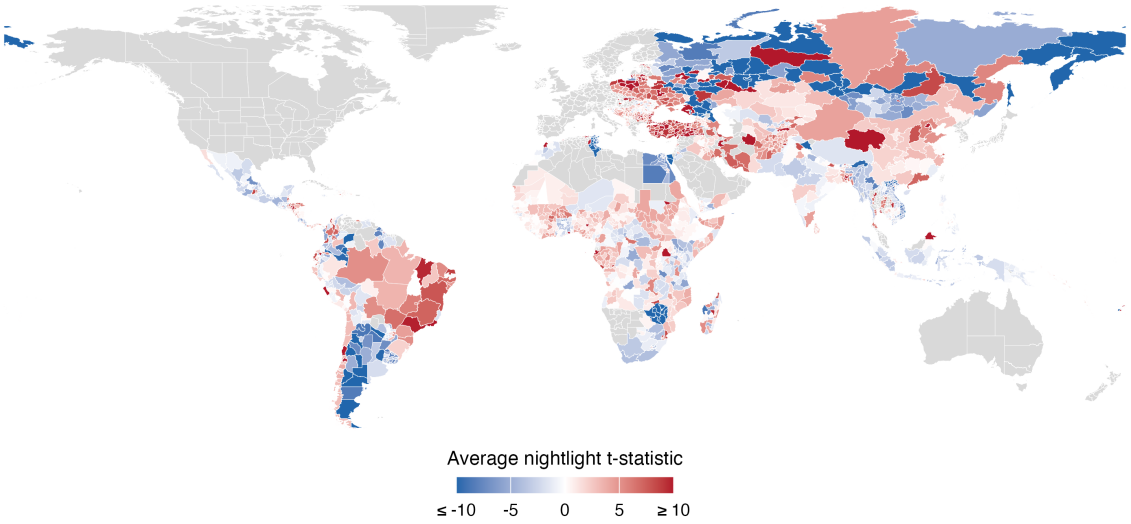
It might be that the World Bank, or recipient countries, are targeting development programmes at a higher administrative unit than districts. To test this, we ask: Does site selection operate only across districts within a province, or also across provinces within a country? We repeat the analysis one administrative level up: we aggregate districts to provinces and compare provinces that received a project to those that did not, within the same country and year. Positive selection is, if anything, sharper at this level: 57% of projects are allocated to significantly brighter provinces, against roughly 30% allocated to darker provinces. The World Bank therefore not only overwhelmingly allocates its projects to richer districts within provinces, but also to richer provinces within countries.

5.3. Site Selection across Sectors

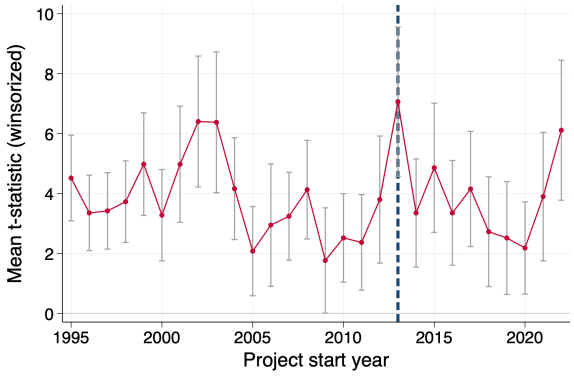
Finally, we ask whether positive site selection is specific to particular types of projects. Figure 6 plots the distribution of the project-level t-statistics separately for each sector. Two patterns emerge. First, the degree of positive selection varies across sectors in the direction one would expect if some projects rely on pre-existing local capacity. The share of positively selected projects is highest in sectors that plausibly require a minimum level of infrastructure: 77.4% in urban development, 61.3% in transport, and 55.7% in energy. It is lowest in agriculture (35.3%) and social protection (38.5%), where one would expect programs to target less developed areas; these are the only two sectors in which negatively selected projects outnumber positively selected ones. Second, positive selection is nonetheless sizeable in every sector, including those in which baseline infrastructure is unlikely to be a binding constraint on implementation: 44.9% of education projects and 41.6% of health

Figure 5: Site Selection across Time and Space

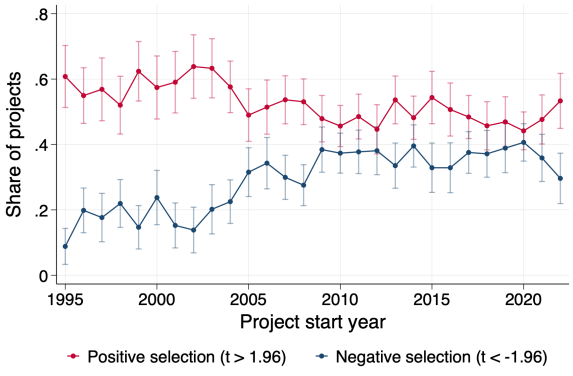
(a) Average site selection by province



(b) Average site selection by start year



(c) Share of selected projects by start year



Notes. All panels use the project-level t-statistics on $\hat{\beta}$ from Equation 3, winsorized at the 1st and 99th percentiles of the project distribution, as in Figure 4. Panel (a) shades each province by the mean t-statistic of the projects it receives; a project implemented in several provinces enters each of them. The sample is the 4,438 projects with an identified estimate, spread over 1,989 provinces in 126 countries. The colour scale is clipped at ± 10, roughly the 5th and 95th percentiles of the province means, and provinces beyond these cutoffs take the end colours. Provinces in grey receive no project with an identified estimate. Panels (b) and (c) cover the 4,432 projects starting between 1995 and 2022; the six projects starting in 2023 are omitted. Panel (b) plots the mean t-statistic by start year with 95% confidence intervals. The vertical dashed line marks 2013, the year the nightlight series switches from DMSP to VIIRS imagery (Chiovelli et al., 2026). Panel (c) plots, by start year, the share of projects significantly positively ($t > 1.96$) and negatively ($t < -1.96$) selected, with 95% confidence intervals.

projects are placed in districts significantly richer than their neighbours.

The cross-sector gradient is consistent with an efficiency rationale: if some projects are more productive where local capacity is higher, placing them in richer districts may maximise their returns. The prevalence of positive selection in the social sectors, however, suggests that efficiency considerations are unlikely to be the whole story. Section 6 tests the efficiency rationale directly, using the World Bank's own ratings of project success, and then turns to the managers who shape where projects go.

6. Why Does Aid Flow to Richer Areas? The Role of Managers

The previous sections show that the World Bank, the largest multilateral provider of development finance, systematically places its projects in richer areas, both across districts within provinces and across provinces within countries, and that it does so in every region, year, and sector. We next ask, why does aid flow to richer areas?

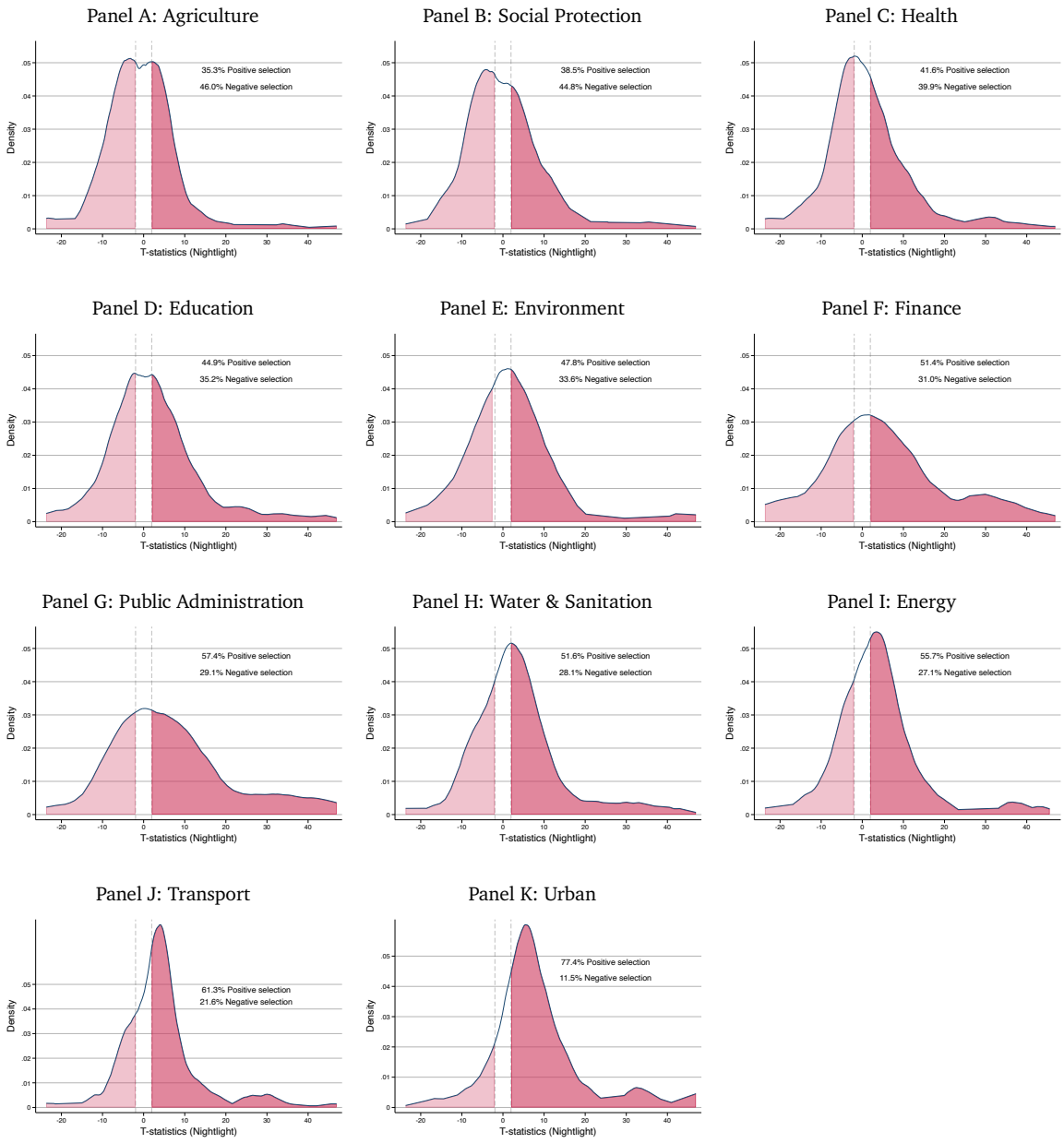
Positive selection need not imply misallocation. If the returns to a project increase with local capacity, as the sectoral gradient in Section 5.3 suggests, placing projects in richer areas may be efficient, albeit at the cost of equity. Alternatively, positive selection may reflect who decides where projects go. As described in Section 2.2, the location of a project is largely determined after a task team leader has been assigned to it, which leaves managers substantial discretion over site selection.

We examine both explanations in turn. We first show that projects placed in richer areas are no more successful according to the ratings of the Independent Evaluation Group, which rules out an equity–efficiency trade-off as the main explanation. We then show that the identity of the task team leader explains a larger share of the variation in site selection than year, sector, and country fixed effects combined, and that this pattern is not driven by the characteristics of the projects that managers are assigned.

6.1. Is There an Equity–Efficiency Trade-off?

We first ask whether the World Bank places its projects strategically, where they are most likely to generate the highest returns. Projects may be more likely to succeed in economically more advantaged areas. For instance, richer areas likely have better road connections,

Figure 6: Distribution of T-Statistics($\hat{\beta}$) by Sector



Notes.The panels show the distribution of project-level t-statistics on $\hat{\beta}$ from Equation 3, split by the project's main sector (classified from the World Bank project documentation). Panels are ordered from the sector with the largest share of negatively selected projects (Agriculture, 46.0 percent) to the smallest (Urban, 11.5 percent). The dashed vertical lines mark the critical values -1.96 and 1.96. Within each panel, the percentages report the shares of significantly positively ($t > 1.96$) and negatively ($t < -1.96$) selected projects among that sector's identified estimates, and the density omits t-statistics below the 1st and above the 99th percentile of the overall distribution.

which allow the WB team, as well as the country officials, to reach implementation areas more frequently and promptly. If positively selected projects are more effective, that is, generate higher returns, this would point to an equity–efficiency trade-off. If not, under the assumption that lowest-consumption households have the highest marginal social valuation, the welfare-maximising strategy would be for the World Bank to allocate programs to less developed areas (Hanna et al., 2026).

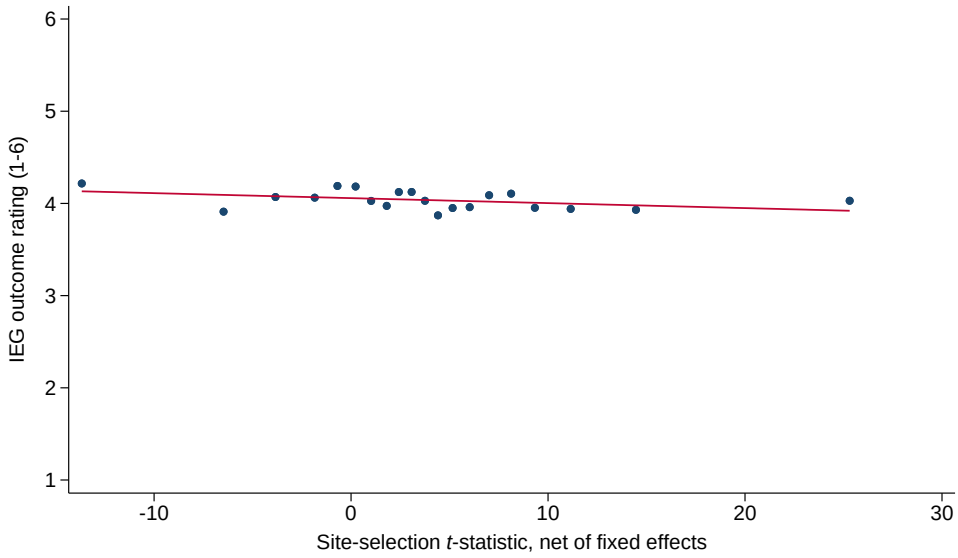
We can measure project success directly with the project score developed by the IEG. As outlined in Section 2.3, the Independent Evaluation Group is the World Bank's independent evaluation arm: for every project that closes, it reviews the Implementation Completion and Results Report prepared by the operational team and assigns an outcome rating, capturing the extent to and efficiency in which the project achieved its stated development objectives. The rating runs on a six-point scale, from 1 (highly unsatisfactory) to 6 (highly satisfactory); in our sample the average project scores 4.03, between moderately satisfactory and satisfactory, with a standard deviation of 1.03, as displayed in Table 1.

If the positive selection we document reflected strategic targeting–the World Bank placing programs where they are most likely to succeed–we would expect the projects with the highest *t*-statistics to be the ones the IEG rates most highly, that is, a positive relationship between the two. Note that the rating measures achievement against a project's own objectives rather than its social return, so a modestly ambitious project in a well-connected district can score well. Note that this works in favour of finding the positive relationship, since objectives are more easily met where implementation is easier.

Figure 7 shows that this is not the case: the relationship between IEG scores and site selection is flat, and if anything it slopes the wrong way for a targeting story. Absorbing start-year-by-country and sector fixed effects, a one standard deviation increase in the *t*-statistic is associated with a 0.057 standard deviation *lower* rating (s.e. 0.031), with a 95% confidence interval of $[-0.119, 0.004]$.[24] We can therefore rule out that projects placed in economically advantaged areas are meaningfully more successful. Put differently, site selection does not correlate with project success, which implies that positive selection is not a by-product of where World Bank programs work best, and hence cannot be rationalised

[24]The unconditional relationship shows the same pattern: a slope of -0.0034 (s.e. 0.0020) across the 2,237 rated projects, or -0.036 standard deviations of the rating per standard deviation of the *t*-statistic (s.e. 0.022). Nor does it change when we look at one side of the distribution at a time. Among projects with $t > 0$, the slope is 0.0010 (s.e. 0.0056) on 870 projects, and among projects with $t < 0$ it is -0.0200 (s.e. 0.0128) on 411 projects, both with the same fixed effects. Neither is statistically distinguishable from zero, so the pooled null is not masking offsetting slopes on either side of $t = 0$.

Figure 7: Project Success and Site Selection



Notes.The figure is a binned scatter plot of the project's IEG outcome rating—1 highly unsatisfactory to 6 highly satisfactory, mean 4.04 and standard deviation 1.03—against the project-level t-statistic on $\hat{\beta}$ from Equation 3, in twenty equal-sized bins, with a line fitted through the underlying project-level data. The sample is the 2,237 projects with both an identified t-statistic and an IEG outcome rating; t-statistics are winsorized at the 1st and 99th percentiles of the full distribution, as in Figure 4. The plot absorbs start-year-by-country and sector fixed effects, so singleton cells drop out and it rests on 1,618 projects; because both variables are residualized on these effects, the horizontal axis is the t-statistic net of the fixed effects. The fitted slope is -0.0054 with a heteroskedasticity-robust standard error of 0.0030. The IEG rates only closed and evaluated projects, so rated projects start between 1995 and 2017.

as an equity–efficiency trade-off.[25]

6.2. Task Team Leaders Explain Site Selection

A key organizational determinant of effectiveness and decision-making is the quality of managers (e.g., Minni (2024); Bertrand and Schoar (2003); Bandiera et al. (2007, 2020b); Fenizia (2022)), as well as the degree of discretion they have. This naturally extends to development programs; indeed, previous studies have found that better project management can reduce violent conflict in aid receiving regions from the World Bank (Moscona, 2024), and career driven politicians can influence policy experiments' success in China (Wang and

[25]One caveat is that the IEG rates only closed and evaluated projects, which leaves 2,237 of our t-statistics and excludes the most recent cohorts. Reassuringly, whether a project is rated is itself unrelated to its t-statistic ($p = 0.40$ unconditionally, and $p = 0.89$ within start-year-by-country and sector cells), so the projects we can score are not selected on the dimension we study.

Yang, 2025).

We ask whether World Bank TTLs influence where aid is directed within countries, consistent with the qualitative evidence on their discretion over program design and location, as outlined in Section 2.2. We are particularly interested in whether TTLs play a role in explaining site selection patterns. To this end, we conduct a variance decomposition exercise, adopting the canonical framework from Abowd et al. (1999), and sequentially regress our measure of site selection on year, sector, country fixed effects, and subsequently on personnel fixed effects for regional vice presidents, country directors, practice managers, and, finally, task team leaders (TTLs). The variance decomposition is enabled by the high degree of mobility of managers across countries through rotations, enabling us to separately estimate the contributions of different fixed effects. Figure 8 reports the adjusted R^2 from project-level regressions of the site-selection t-statistic as fixed effects are added sequentially.

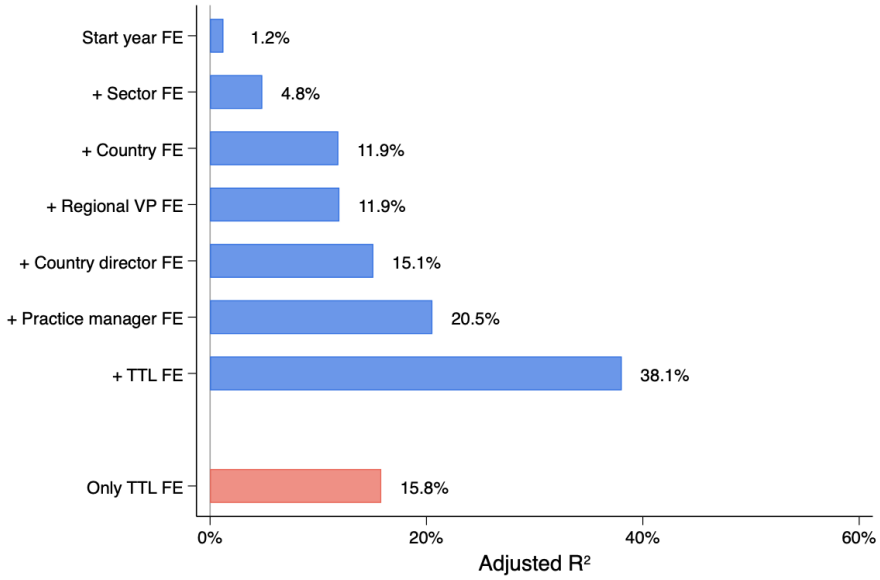
We find that TTLs are the largest contributor in explaining the variation in site selection. Indeed, TTL fixed effects, by themselves, explain 15.8% of the variation in site selection, more than year, sector, and country fixed effects combined (11.9%). Adding personnel fixed effects down the hierarchy raises this gradually: regional vice-president, country director, and practice manager fixed effects bring the adjusted R^2 to 20.5%. Adding TTL fixed effects then raises it to 38.1%, an increase of 17.6 percentage points, the largest single jump in the sequence. Personnel in other roles, such as regional vice-presidents, country directors, and practice managers, do not explain as much variation.

Our finding is in line with the institutional background: TTLs are the managers in charge of defining the main characteristics of the development programs, including their location. Overall, this evidence suggests that the allocation of development programs over space is not only influenced by the characteristics of the borrowing country, but also by the internal structure of international organizations, and the discretion of who is in charge of the projects' implementation.

6.3. Project Characteristics Matter, but Managers Matter More

Manager fixed effects could pick up the *type* of project each manager is given rather than the manager. For example, if some TTLs systematically run larger, longer, or better-performing projects, and these characteristics predict where projects land, the variance decomposition in Figure 8 would attribute to managers what belongs to their projects. To

Figure 8: Explaining variation in T-Statistics($\hat{\beta}$)



Notes. This figure shows the adjusted-R^2 from a project-level regression of site-selection on project characteristics. Each row –except the last one– adds an additional set of covariates to the model, starting from year fixed effects and sequentially adding sector and country fixed effects, followed by regional vice president, country director, and practice manager fixed effects, and lastly, manager fixed effects. The last row only controls for manager fixed effects.

test this, we run a horse race regression between observable project characteristics and the fixed effects of each layer of the Bank's hierarchy:

$$t_i = X_i'\phi + \mu_{s(i)} + \tau_{y(i)} + \lambda_{c(i)} + \theta_{m(i)} + \varepsilon_i, \tag{4}$$

where t_i is the site-selection t-statistic of project i from Equation 3[26] and X_i contains project characteristics: log disbursements, duration in years, and the IEG outcome rating. $\mu_{s(i)}$, $\tau_{y(i)}$ and $\lambda_{c(i)}$ are sector, start-year and country fixed effects, added one at a time. $\theta_{m(i)}$ is the fixed effect of the person holding a given role at appraisal—regional vice president, country director, or TTL—each added separately on top of the sector, year and country fixed effects, so that every column compares one layer of the hierarchy against the same baseline. Standard errors are clustered by country. Since the IEG rates only closed projects, our sample is restricted to $N = 2,207$ projects with a rating.

[26]We winsorize at the 1st and 99th percentiles as throughout.

Without manager fixed effects, longer and better-rated projects go to poorer districts. Panel A of Table 3 and Panel (a) of Figure 9 show that each additional year of duration lowers the *t*-statistic by 0.66 to 0.92, and each additional point on the six-point IEG scale by 0.56 to 0.76; both are significant at the 5% level in the first four columns. The rating result is in line with Figure 7: if anything, the projects the IEG rates most highly are the least positively selected. Disbursements are insignificant without country fixed effects and positive once they are added (0.72, s.e. 0.30), so that within a country, larger projects go to richer districts. Adding regional vice president and country director fixed effects (columns 5 and 6) leaves the estimates on duration and the IEG rating largely unchanged and attenuates the estimate on disbursements, which remains significant only at the 10% level.

With TTL fixed effects, none of these relationships survives. In column 7, the coefficients on disbursements, duration and the IEG rating are all statistically indistinguishable from zero. The link between ratings and site selection therefore runs *across* managers rather than across the projects of the same manager: TTLs whose projects are rated highly are also those who place their projects in poorer districts. Because TTL fixed effects are identified only from managers with at least two projects, column 7 runs on 1,337 projects led by 482 TTLs, against 2,201 in column 4. Panel B of Table 3 and Panel (b) of Figure 9 re-estimate every column on this smaller sample. While the project characteristic coefficients lose significance earlier than in the last specification[27], we see that with the addition of TTL fixed effects they are all centered around zero.

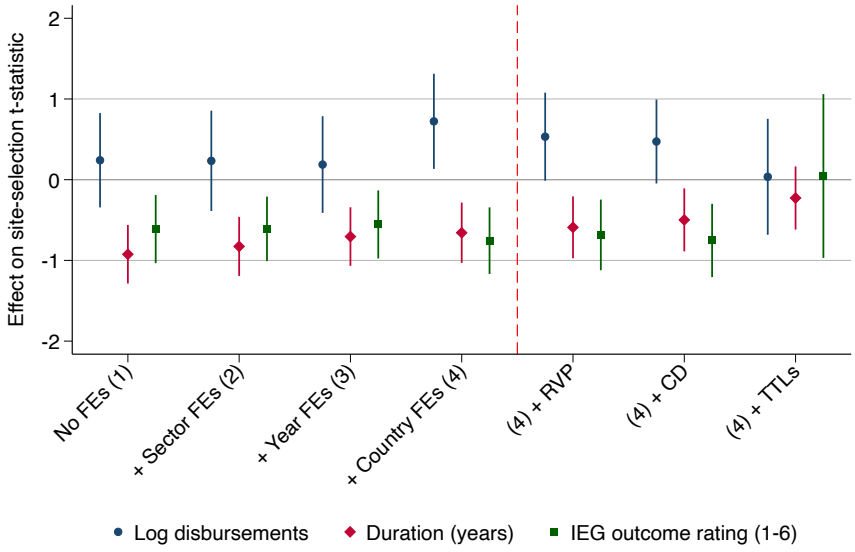
This exercise also confirms that TTLs also explain far more of the variation in site selection than any other layer of the hierarchy, or project-specific characteristic. On the common sample of Panel B, sector, start-year and country fixed effects together raise the adjusted R^2 by 0.10 (from 0.018 to 0.116); regional vice president fixed effects add a further 0.01, country director fixed effects 0.04, and TTL fixed effects 0.18.[28] Managers thus add twice as much as sector, year and country combined, and about five and thirteen times as much as country directors and regional vice presidents.

[27] The IEG rating remains significant in columns 1–6 and disappears in column 7; duration loses significance at the 5% level once country fixed effects are added (column 4), and disbursements are insignificant throughout.

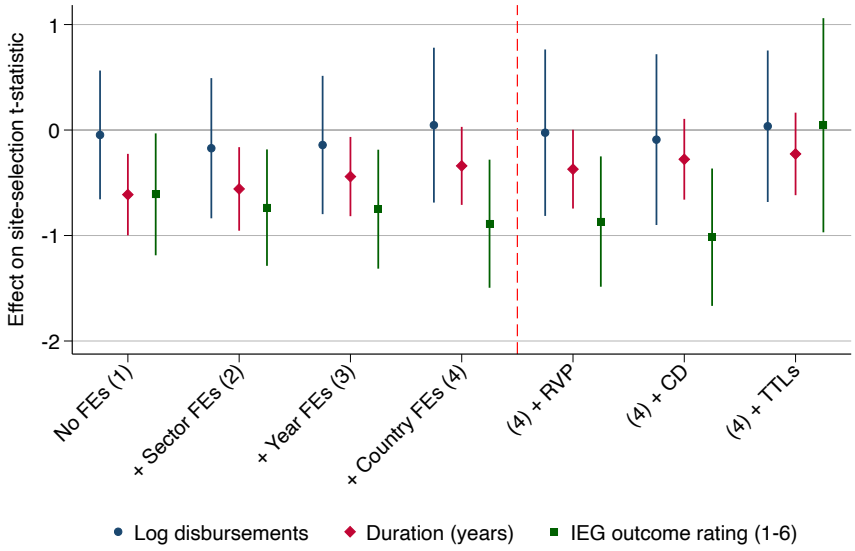
[28] Columns 5 and 6 of Panel B run on slightly fewer projects (1,284 and 1,257), as some projects have no recorded regional vice president or country director, or are singletons of those fixed effects.

Figure 9: Project Characteristics and Site Selection, Adding Fixed Effects

(a) Sample not fixed



(b) Sample fixed to TTLs with at least two projects



Notes. The figure plots the coefficients from Equation 4, with 95% confidence intervals from standard errors clustered by country, across the seven columns of Table 3. The dependent variable is the project-level site-selection t-statistic on $\hat{\beta}$ from Equation 3. Columns (1)–(4) add sector, start-year and country fixed effects in turn; the last three columns each add one personnel fixed effect to column (4): regional vice president, country director, or TTL, all at appraisal. The red dashed line separates the columns without and with personnel fixed effects. Panel (a) lets the sample change across columns; Panel (b) fixes it to the 1,337 projects in the estimation sample of the TTL column.

Table 3: Project Characteristics and Site Selection, Adding Fixed Effects

	Without Org. Characteristics				+ Regional VP	+ Country director	+ TTLs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: sample not fixed							
Log disbursements	0.241	0.233	0.188	0.723**	0.533*	0.473*	0.036
	(0.295)	(0.314)	(0.302)	(0.298)	(0.275)	(0.262)	(0.362)
Duration (years)	-0.924***	-0.827***	-0.705***	-0.657***	-0.590***	-0.498**	-0.227
	(0.183)	(0.185)	(0.183)	(0.189)	(0.193)	(0.197)	(0.197)
IEG outcome rating (1-6)	-0.612***	-0.609***	-0.555**	-0.756***	-0.684***	-0.753***	0.045
	(0.213)	(0.202)	(0.213)	(0.208)	(0.221)	(0.229)	(0.512)
Sector FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Start-year FE	No	No	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	Yes	Yes	Yes	Yes
No. of sectors		13	13	13	13	13	13
No. of start years			23	23	23	23	22
No. of countries				115	115	115	105
No. of regional VPs					49		
No. of country directors						173	
No. of TTLs							482
Observations	2,207	2,206	2,206	2,201	2,114	2,099	1,337
R^2	0.039	0.088	0.114	0.206	0.233	0.331	0.624
Adjusted R^2	0.038	0.082	0.098	0.148	0.154	0.209	0.298
Panel B: sample fixed to column 7 (TTLs with at least two projects)							
Log disbursements	-0.046	-0.172	-0.142	0.046	-0.025	-0.091	0.036
	(0.308)	(0.335)	(0.331)	(0.370)	(0.398)	(0.408)	(0.362)
Duration (years)	-0.612***	-0.558***	-0.441**	-0.340*	-0.371*	-0.277	-0.227
	(0.195)	(0.200)	(0.189)	(0.186)	(0.188)	(0.193)	(0.197)
IEG outcome rating (1-6)	-0.610**	-0.736***	-0.750***	-0.888***	-0.868***	-1.016***	0.045
	(0.291)	(0.278)	(0.284)	(0.306)	(0.311)	(0.328)	(0.512)
Sector FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Start-year FE	No	No	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	Yes	Yes	Yes	Yes
No. of sectors		13	13	13	13	13	13
No. of start years			22	22	22	22	22
No. of countries				105	105	103	105
No. of regional VPs					38		
No. of country directors						135	
No. of TTLs							482
Observations	1,337	1,337	1,337	1,337	1,284	1,257	1,337
R^2	0.020	0.083	0.114	0.208	0.250	0.337	0.624
Adjusted R^2	0.018	0.072	0.089	0.116	0.130	0.154	0.298

Notes. The dependent variable is the project-level site-selection *t*-statistic on $\hat{\beta}$ from Equation 3, winsorized at the 1st and 99th percentiles of the full distribution, as in Figure 4. The sample is projects with an IEG outcome rating. Columns (2)–(4) add sector, start-year and country fixed effects in turn; columns (5)–(7) each add one personnel fixed effect to column (4): regional vice president, country director, or TTL, all at appraisal. Panel A lets the sample change across columns. Panel B re-estimates every column on the estimation sample of column (7), so column (7) is identical in both panels; columns (5) and (6) of Panel B lose projects with no recorded regional vice president or country director, and singletons of those fixed effects. The "No. of" rows report the number of unique values of each absorbed fixed effect in the estimation sample. Standard errors clustered by country in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7. Conclusion

Despite the scale of development assistance, we know relatively little about where it goes within recipient countries and what determines its allocation. This paper examines whether the geographic allocation of World Bank projects is consistent with the Bank's stated objective of targeting the bottom 40% of the population in every country.[29]

We find that it is not: World Bank projects flow disproportionately to richer areas. Over half (51%) of projects are located in districts with significantly higher baseline nighttime light intensity than non-recipient districts in the same province, while only 31% are located in significantly poorer ones. The pattern is robust to alternative specifications, holds one administrative level up, and is present in every region, year, and sector. It does not reflect an equity–efficiency trade-off: projects placed in richer areas perform no better according to the World Bank's independent ratings. Instead, the identity of the project manager matters. Task team leader fixed effects explain 16% of the variation in site selection, more than year, sector, and country fixed effects combined, and this result is not driven by the characteristics of the projects that managers run.

These findings suggest that the geographic allocation of aid is shaped not only by the characteristics of borrowing countries but also by the personnel of the international organisations that deliver it. Where aid goes within countries is therefore, in part, an organisational outcome, and one that the agencies themselves can influence. In ongoing work, we exploit the rotation of task team leaders across countries to estimate the causal effect of managers on project placement, characterise which managers direct projects to poorer areas, and study whether these choices affect their careers.

[29] Independent Evaluation Group, *Growth for the Bottom 40 Percent: The World Bank Group's Support for Shared Prosperity* (Washington, DC: World Bank, 2017), https://documents1.worldbank.org/curated/en/176811511360825388/pdf/120635-WP-PUBLIC.pdf.

References

John M Abowd, Francis Kramarz, and David N Margolis. High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333, 1999. ◁ pp. 3 and 28

AidData. WorldBank_GeocodedResearchRelease_Level1_v1.4.2 Geocoded Dataset, 2017. URL http://aiddata.org/research-datasets. Accessed on May 15, 2025. ◁ pp. 8 and 13

Hunt Allcott. Site Selection Bias in Program Evaluation. *The Quarterly Journal of Economics*, 130 (3):1117–1165, 03 2015. ◁ pp. 2, 4, and 15

Jørgen Juel Andersen, Niels Johannesen, and Bob Rijkers. Elite Capture of Foreign Aid: Evidence from Offshore Bank Accounts. *Journal of Political Economy*, 130(2):388–425, 2022. ◁ p. 4

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment. *The Quarterly Journal of Economics*, 122(2): 729–773, 2007. ◁ p. 27

Oriana Bandiera, Robin Burgess, Erika Deserranno, and Imran Rasul. Development Policy through the Lens of Social Structure. 2020a. ◁ p. 4

Oriana Bandiera, Andrea Prat, Stephen Hansen, and Raffaella Sadun. CEO Behavior and Firm Performance. *Journal of Political Economy*, 128(4):1325–1369, 2020b. ◁ p. 27

Philipp Barteska and Jay Euijung Lee. *Bureaucrats and the Korean Export Miracle*. PhD thesis, London School of Economics and Political Science, 2024. ◁ p. 4

Marianne Bertrand and Antoinette Schoar. Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, 118(4):1169–1208, 2003. ◁ pp. 3, 14, and 27

Michael Carlos Best, Jonas Hjort, and David Szakonyi. Individuals and Organizations as Sources of State Effectiveness. *American Economic Review*, 113(8):2121–2167, August 2023. doi: 10.1257/aer.20191598. URL https://www.aeaweb.org/articles?id=10.1257/aer.20191598. ◁ pp. 3 and 4

Nicholas Bloom, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51, 11 2013. ◁ p. 4

Pietro Bomprezzi, Axel Dreher, Andreas Fuchs, Teresa Hailer, Andreas Kammerlander, Lennart Kaplan, Silvia Marchesi, Tania Masi, Kerstin Perlik, and Charlotte Robert. Wedded to Prosperity? Spousal Favoritism in Foreign Aid and Regional Development. CEPR Discussion Paper 18878, CEPR, 2025. URL https://godad.uni-goettingen.de/uploads/Spouses.pdf. Version 3. ◁ pp. 1, 8, 9, 13, and 15

Ryan C Briggs. Poor Targeting: A Gridded Spatial Analysis of the Degree to Which Aid Reaches the Poor in Africa. *World Development*, 103:133–148, 2018. ◁ p. 4

Giorgio Chiovelli, Stelios Michalopoulos, Elias Papaioannou, and Tanner Regan. Illuminating the Global South. *The Economic Journal*, page ueaf134, 2026. ◁ pp. 1, 8, 10, 12, 15, 23, and A.2

Francesco Decarolis, Leonardo M Giuffrida, Elisabetta Iossa, Vincenzo Mollisi, and Giancarlo Spagnolo. Bureaucratic Competence and Procurement Outcomes. *The Journal of Law, Economics, and*

Organization, 36(3):537–597, 05 2020. ◁ p. 4

Cevdet Denizer, Daniel Kaufmann, and Aart Kraay. Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance. *Journal of Development Economics*, 105: 288–302, 2013. doi: 10.1016/j.jdeveco.2013.06.003. ◁ pp. 3 and 7

Axel Dreher, Andreas Fuchs, Roland Hodler, Bradley C Parks, Paul A Raschky, and Michael J Tierney. African Leaders and the Geography of China's Foreign Assistance. *Journal of Development Economics*, 140:44–71, 2019. ◁ p. 4

Elisa Facchetti, Arianna Ornaghi, and Eddy Zou. Middle Managers in the Public Sector: Evidence from Policing. Working paper, 2026. URL https://www.dropbox.com/scl/fi/0ym4ns5xe56n946dl2fz8/police_manager_draft.pdf?rlkey=t8wefuogbldqe8spa86bl6le3&dl=0. ◁ p. 4

Alessandra Fenizia. Managers and Productivity in the Public Sector. *Econometrica*, 90(3):1063–1084, 2022. doi: https://doi.org/10.3982/ECTA19244. ◁ pp. 3, 4, and 27

Nicola Gennaioli, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. Human Capital and Regional Development. *The Quarterly Journal of Economics*, 128(1):105–164, 2013. doi: 10.1093/qje/qjs050. ◁ p. 1

Seth Goodman, Ariel BenYishay, Zhonghui Lv, and Daniel Runfola. GeoQuery: Integrating HPC Systems and Public Web-Based Geospatial Data Tools. *Computers & Geosciences*, 122:103–112, 2019. doi: 10.1016/j.cageo.2018.10.009. ◁ pp. 1 and 10

Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118 (3):536–598, 2010. ◁ p. 15

Raymond P Guiteras and Ahmed Mushfiq Mobarak. Does Development Aid Undermine Political Accountability? Leader and Constituent Responses to a Large-Scale Intervention. Technical report, National Bureau of Economic Research, 2015. ◁ p. 4

Rema Hanna, Emily Breza, Benjamin A. Olken, Abhijit Banerjee, Arun G. Chandrasekhar, and Gideon Moore. Target What? The Foundations of Poverty Targeting. 2026. ◁ p. 26

Mirko Heinzel and Andrea Liese. Managing Performance and Winning Trust: How World Bank Staff Shape Recipient Performance. *The Review of International Organizations*, 16:625–653, 2021. doi: 10.1007/s11558-021-09414-4. ◁ pp. 3, 6, and 7

J Vernon Henderson, Adam Storeygard, and David N Weil. Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2):994–1028, 2012. ◁ p. 10

Roland Hodler and Paul A Raschky. Regional Favoritism. *The Quarterly Journal of Economics*, 129 (2):995–1033, 2014. ◁ p. 12

Independent Evaluation Group. Learning and Results in World Bank Operations: Toward a New Learning Strategy, Evaluation 2. Technical report, World Bank, Washington, DC, 2015. URL https://ieg.worldbankgroup.org/sites/default/files/Data/reports/chapters/LR2_full_report_revised.pdf. ◁ p. 7

Christopher Kilby. An Empirical Assessment of Informal Influence in the World Bank. *Economic Development and Cultural Change*, 61(2):431–464, 2013. doi: 10.1086/668278. ◁ p. 7

Xuecao Li, Yuyu Zhou, Min Zhao, and Xia Zhao. A Harmonized Global Nighttime Light Dataset 1992–2018. *Scientific Data*, 7(1):168, 2020. ◁ pp. 10 and 12

Nicola Limodio. Bureaucrat Allocation in the Public Sector: Evidence from the World Bank. *The Economic Journal*, 131(639):3012–3040, 01 2021. ◁ pp. 8 and 14

Luis R. Martínez. How Much Should We Trust the Dictator's GDP Growth Estimates? *Journal of Political Economy*, 130(10):2731–2769, 2022. doi: 10.1086/720458. ◁ pp. 2 and 10

Eoin F McGuirk and Nathan Nunn. Development Mismatch? Evidence from Agricultural Projects in Pastoral Africa. Technical report, National Bureau of Economic Research, 2024. ◁ p. 4

Stelios Michalopoulos and Elias Papaioannou. Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, 81(1):113–152, 2013. ◁ p. 12

Stelios Michalopoulos and Elias Papaioannou. National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, 129(1):151–213, 2014. ◁ p. 12

Virginia Minni. Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs. Working paper, SSRN, November 2024. Available at SSRN: https://ssrn.com/abstract=5289177. ◁ p. 27

Marta Morando. Make or Buy the State? Consulting in the Public Sector. Job market paper, London School of Economics and Political Science, November 2026. URL https://martamorando.com/papers/jmp_morando.pdf. ◁ p. 4

Jacob Moscona. The Management of Aid and Conflict in Africa. Working Paper, 2024. ◁ pp. 3, 4, 6, 7, and 27

Joris Mueller. China's Foreign Aid: Political Determinants and Economic Effects. Technical report, Working paper, 2022. ◁ p. 4

Nathan Nunn and Nancy Qian. US Food Aid and Civil Conflict. *American Economic Review*, 104(6):1630–1666, 2014. ◁ p. 4

Nancy Qian. Making Progress on Foreign Aid. *Annual Review of Economics*, 7(1):277–308, 2015. ◁ p. 1

Imran Rasul and Daniel Rogger. Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*, 128(608):413–446, 04 2017. ISSN 0013-0133. doi: 10.1111/ecoj.12418. URL https://doi.org/10.1111/ecoj.12418. ◁ p. 3

Daniel Runfola, Austin Anderson, Heather Baier, Matt Crittenden, Elizabeth Dowker, Sydney Fuhrig, Seth Goodman, Grace Grimsley, Rachel Layko, Graham Melville, Maddy Mulder, Rachel Oberman, Joshua Panganiban, Andrew Peck, Leigh Seitz, Sylvia Shea, Hannah Slevin, Rebecca Youngerman, and Lauren Hobbs. geoBoundaries: A Global Database of Political Administrative Boundaries. *PLOS ONE*, 15(4):e0231866, 2020. doi: 10.1371/journal.pone.0231866. URL https://doi.org/10.1371/journal.pone.0231866. ◁ p. 8

Shaoda Wang and David Y. Yang. Policy Experimentation in China: The Political Economy of Policy Learning. *Journal of Political Economy*, 133(7):2180–2228, 2025. doi: 10.1086/734873. ◁ pp. 2, 4, 15, 16, 17, 19, and 27

World Bank. Task Team Leader Toolkit: Investment Project Financing – The Project Cycle. Technical

report, The World Bank, 2013. ◁ pp. 3, 6, and 7

The Geography of Development Programs Online Appendix

Beyza Gulmezoglu, Marta Morando, Eddy Zou

A. Additional Descriptive Information **A.2**

B. Site-Selection Validation **A.3**

C. A New Dataset: Historical International Development Aid (HIDA) **A.4**

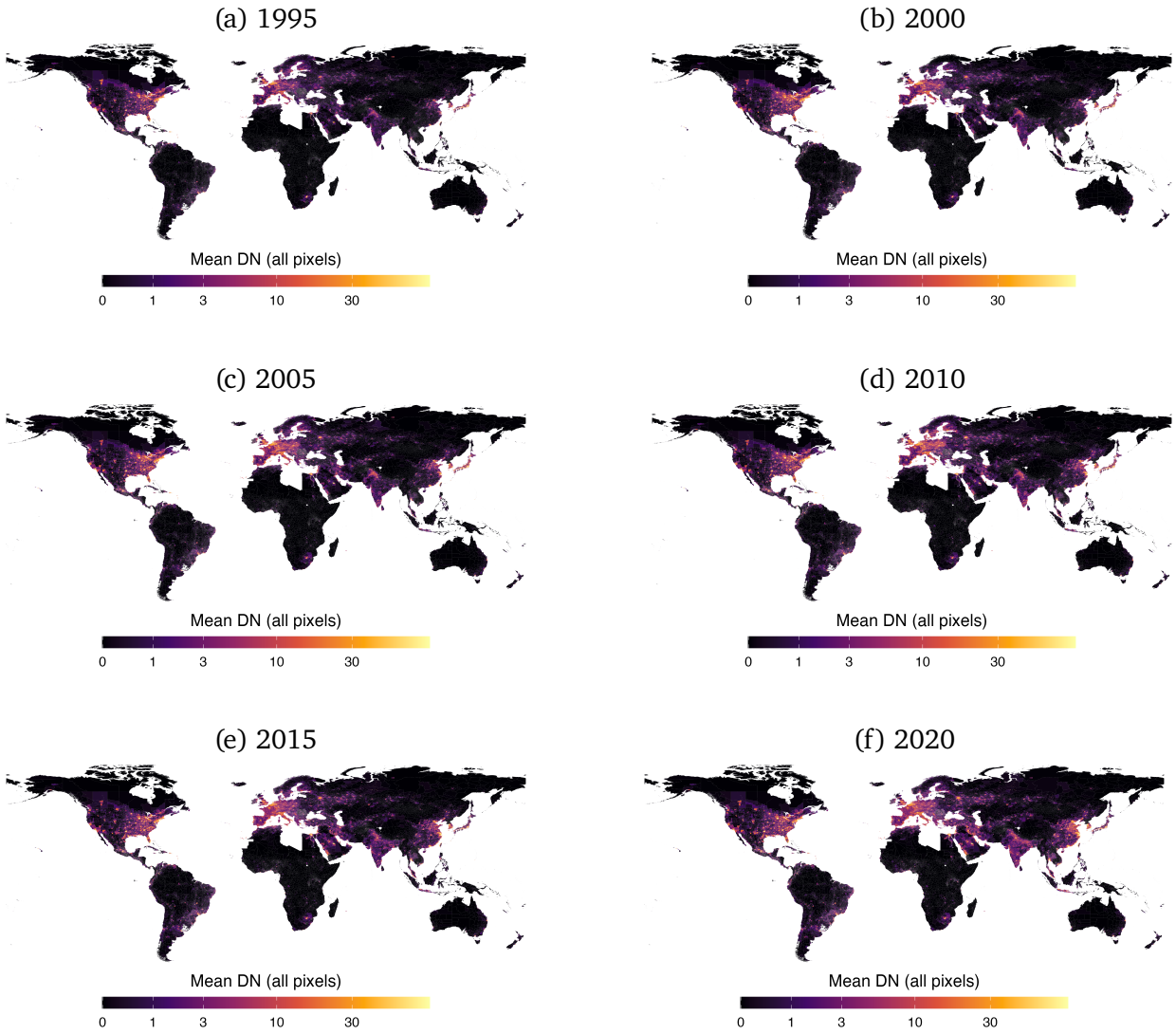
C.1. The document corpus A.4

C.2. From text to data A.4

C.3. Project teams and managers A.5

A. Additional Descriptive Information

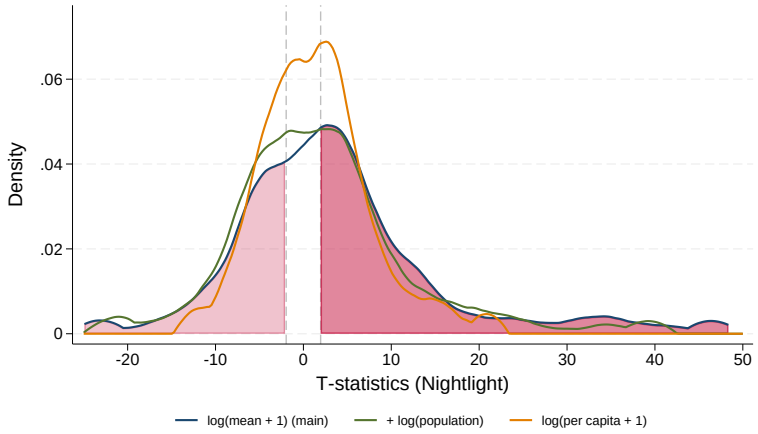
Figure A.1: Mean nighttime light intensity by ADM2 district, 1995–2020



Notes. Mean nighttime light intensity by ADM2 district, 1995–2020. Each panel maps $\log(\text{mean DN} + 1)$ across ADM2 districts, using the bloom- and topcode-corrected DMSP series (Chiovelli et al., 2026). The mean is taken over *all* pixels in a district, with unlit pixels counted as zero, so every district has a value and no threshold is applied. All six panels share a common colour scale, capped near the 99th percentile of the pooled distribution; the brightest districts are therefore drawn at the top colour.

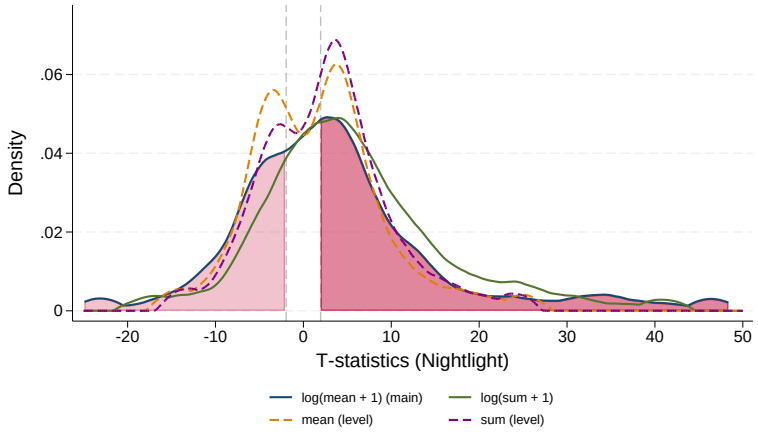
B. Site-Selection Validation

Figure A.2: Robustness: Alternative Specifications



Notes. Distribution of the project-level *t*-statistics on $\hat{\beta}$ from Equation 3, varying the specification while holding the province-by-year fixed effects and heteroskedasticity-robust standard errors fixed. The main specification (shaded) uses $\log(1 + \text{average nightlight})$ as the outcome; the alternatives add a $\log(1 + \text{population}_{2001})$ control and replace the outcome with $\log(1 + \text{per-capita nightlight})$. The two population-based specifications are estimated only on districts with known 2001 population. The dashed vertical lines mark ± 1.96; values are winsorized at the 1st and 99th percentiles. The shares of significantly positively / negatively selected projects are 51% / 31% (main), 46% / 36% (population control), and 44% / 31% (per capita).

Figure A.3: Robustness: Alternative Outcome Definitions



Notes. Distribution of the project-level *t*-statistics on $\hat{\beta}$ from Equation 3, varying the dependent variable while holding the province-by-year fixed effects and heteroskedasticity-robust standard errors fixed. The main specification (shaded) uses $\log(1 + \text{average nightlight})$; the alternatives use $\log(1 + \text{total nightlight})$ and the same two measures in levels. The dashed vertical lines mark ± 1.96; values are winsorized at the 1st and 99th percentiles. The shares of significantly positively / negatively selected projects are 51% / 31% (log average, main), 62% / 21% (log total), 48% / 34% (average, level), and 54% / 28% (total, level).

Table A.1: The document corpus

Document type	Texts	Years
Appraisal document (PAD, SAR)	13,559	1950–2026
Completion report (ICR)	8,539	1981–2026
Project information document (PID)	15,302	1994–2026
Combined PID and safeguards sheet	2,323	2015–2026
Safeguards data sheet (ISDS)	9,393	2000–2026
Environmental assessment	17,935	1977–2026
Resettlement plan	7,371	1995–2026
Project paper	12,810	2002–2026
Memorandum of the president	6,058	1950–2014
Implementation status report (ISR)	46,746	2006–2026
Program document	1,587	2002–2026
Total	141,623	1950–2026

Notes. Plain-text project documents downloaded from the World Bank Documents and Reports archive. Counts are texts recovered on disk; a project has several documents of several types.

C. A New Dataset: Historical International Development Aid (HIDA)

C.1. The document corpus

The raw material is the full text of the World Bank's own project documents, downloaded from the Bank's Documents and Reports archive. The corpus holds 141,623 documents for nearly 20,000 projects, spanning 1950 to 2026, in eleven types (Table A.1). The appraisal document is the richest single source: its data sheet names the full team and other information. The completion report names the team at closing.

C.2. From text to data

The data are built in three steps. First, a keyword screen locates the passages of each document that can carry the target information, and keeps a fixed window of text around each hit. Only these passages, a small share of the total text, are read further by a large language model. This helps minimising costs. Second, a large language model (Google's Gemini 2.5) reads the screened passages for each project and returns a structured record in a fixed schema. Every non-empty field is required to carry a sentence that supports it, so any value can be audited against the original source. Third, we normalize the categorical variables, add the source files, and write one dataset at the project level.

Table A.2: Team dataset: coverage and fill rates

	Filled
Panel A. Person-project records (529,191 rows)	
Name	100%
Role (grouped)	98%
Seniority (1–8)	99%
Project stage	98%
Staff or consultant	74%
Bank unit	28%
Printed title	16%
Panel B. Project team summary (16,642 projects)	
Task team leader at start	87%
Practice manager at start	79%
Country director at start	73%
Regional vice president at start	66%
Team retained to supervision	47%
Staff-weeks and cost	41%

Notes. The dataset names 46,606 distinct people across 18,494 projects. "Filled" is the share of person-project records (Panel A) or projects (Panel B) with a non-missing value. Title and unit are thin because most documents give a role but not the exact printed title or unit code.

C.3. Project teams and managers

Development projects are carried out by large teams, which do not solely rely on task team leaders, but other workers too, including country directors, team members, practitioner managers, as well as external. AidData and available World Bank data only record the Task Team Leader names and only for a limited subset of projects. We extract every named Bank staff member from the documents, with their role, seniority, and the stage at which they appear, and link each person across projects with a person identifier. The dataset names at least one staff member for 18,494 projects and identifies the task team leader for 87 percent of projects with an appraisal document. It contains 46,606 distinct people. We validate the task team leader against AidData and the Bank's own staff records; agreement is high where all three exist.

Snapshot: P086689 (Chile, urban transport). At appraisal the data sheet names Jorge Rebelo as task team leader, Jose-Luis Irigoyen as practice manager, Axel van Trotsenburg as country director, and Pamela Cox as regional vice president, with a team of 28. Three different task team leaders lead the project over its life, which the data flags as high leadership turnover.