Inaccessibility and Compliance: Is There a Streetlight Effect in Foreign Aid?*

Michael Denly† Benjamin Gottfried‡

March 20, 2025

For the most recent version of this paper, please visit:

https://mikedenly.com/research/aid-inaccessibility-compliance

Abstract

Why do some foreign aid projects have severe negative social and environmental externalities that often overwhelm the benefits of providing foreign aid in the first place? We argue that a key reason pertains to the inaccessibility of projects, which makes supervision more costly and less feasible. By the same token, higher-quality project leaders have agency to overcome some of these more structural constraints to supervision and development. To test the hypotheses, we use Open Street Map data to analyze whether driving times to reach projects affect compliance with World Bank social and environmental safeguard policies. Consistent with our pre-analysis plan, we find results indicating a streetlight effect phenomenon: aid externalities are a function of the extra work and structural difficulties associated with their prevention. Interaction analyses also confirm that higher-quality project leaders are sometimes—though not always—able to better mitigate negative aid externalities. At the highest possible level, our results speak to the logistical challenges behind the structural impediments to development and the extent to which individual-level agency can help overcome them.

*For funding, Michael Denly thanks Texas A&M's Lab for Economic Development Research (LEDR), the Scowcroft Institute of International Affairs at Texas A&M University, and the University of Texas at Austin's Innovations for Peace and Development lab (IPD). For advice or feedback, we thank Mirko Heinzel and Kyosuke Kikuta. We also thank the following student researchers at IPD and LEDR for assistance with the data collection: Carter Adams, Destiny Alvarez, Alyssa Ashcraft, Roxanne Bankhead, Anita Basavaraju, Sam Bennett, Mary Margaret Burniston, Mariana Caldas, Claudia Canfield, Lauren Childress, Monique Cooper, Ricardo Corona, Ji Na Gil, Colleen Goggin, Hannah Greer, Maria Fernanda Guerrero, Joelean Hall, Tara Jackson, Karan Kanatala, Samiya Javed, Judy Lane, Beomhak Lee, Julianna Leung, Vanessa Lizcano, Amanda Long, Aoife McDonnell, Ruth Mewhinney, Ana Karen Moreno Flores, Nicole Pownall, Briana Rocha, Katelyn Sammon, Mackenzie Sanderson, Kimberly Schuster, Taryn Shanes, Samantha Shoff, Mackenna Shull, Kyle Smith, Ayu Sofyan, Veronica Lulu Stracher, Adityamohan Tantravahi, Ethan Tenison, Skyler Thomas, Jade Tucker, and Jonathan Velasquez. All errors are uniquely our own.

† Assistant Professor, Texas A&M University, Bush School of Government & Public Service, mdenly@tamu.edu

‡ Master's student, Texas A&M University, Bush School of Government & Public Service, gottfried.ben@tamu.edu

How much do logistical (in)accessibility and convenience shape key foreign aid and development outcomes? By asking such a question, we aim to probe whether foreign aid is subject to a streetlight effect. The latter derives from the famous "Drunkard's Search" anecdote. In it, a drunkard nonsensically searches for his keys under the streetlight not because he lost them there, but because the streetlight makes it convenient to look for them (Hendrix, 2017, 17).

In this paper, we examine whether a streetlight effect explains the manifestation of negative foreign aid externalities. Consistent with Denly (2025), we operationalize negative aid externalities through non-compliance with the World Bank's social and environmental safeguard policies. When there is noncompliance with these risk management policies, it may not be worth undertaking the foreign aid project in the first place. For example, consider an aid project that not only builds a road to connect people to hospitals, schools, and markets but also forces mass resettlement and destroys a crucial rainforest. Such a case undermines the public good value of the road and is also relatively common (Tello, 2015). That is why all major aid donors nowadays have social and environmental safeguard policies—even if compliance with them is far from perfect (Greenstein, 2022).

To operationalize the streetlight effect, we use data on driving times to capture which aid projects are most difficult to reach and supervise. More specifically, we use Open Street Maps to measure the combined driving times from (1) the closest airport to each project's most accessible location; and (2) that project's location to the nearest location of another project, which differs on a yearly basis as new projects start and close. We incorporate the airport time to mimic a supervision mission from abroad, and we use the nearest neighbor project travel time to capture how streetlight pressures change over time. As an alternative measure, we perform spatial grid-cell analysis to map the number of nearby projects operating within the same area as well as first- and second-order neighboring areas.[1]

To complement the above analyses that focus on the structural constraints to devel-

[1]Technically, first- and second-order neighboring area refer to first- and second-order spatial lags, which we describe later in the paper.

opment, we undertake more agency-centric analyses as well. Using Denly's (2025) data on World Bank Task Team Leaders (TTL) quality, we interact the latter with the log total travel time. The interaction provides a measure of whether having higher-quality TTLs mitigate potential streetlight effect patterns.

Consistent with our pre-analysis plan,[2] we find statistical support for the proposition that larger (log) driving times yield lower levels of safeguard policy compliance. The interaction analyses also indicate that higher-quality TTLs can—though not always—mitigate negative aid externalities. We add the "not always" qualifier because results for the interaction analysis just miss conventional levels of statistical significance once we add spatial control variables.

Our results make two larger contributions. First, the paper adds a within-country, spatial dimension to the literature and highlights how five existing theories of compliance can operate all at once. Among the major theories, the streetlight effect only has little to do with legitimacy or constructivist explanations regarding norms.[3] By contrast, the links are clear to the schools on enforcement,[4] management,[5] domestic constituencies,[6] biased bureaucrats,[7] and macro-structural constraints.[8] On enforcement, a streetlight effect makes supervision and sanctions more difficult for aid providers. By definition, a streetlight effect is state capacity constraint, the key focus of the management school. With respect to the domestic constituency mechanism, a streetlight effect may explain why some domestic constituencies are subject to more compliance scrutiny than others. The macro-structural link is also clear given the streetlight effect's relation to geography. Finally, biased bureaucrats can explain

[2]See: https://osf.io/ru9dp/?view_only=e7458ea391bd4a90b7ffae2b5baba451

[3]e.g., Finnemore and Sikkink (1998)

[4]e.g., Keohane and Nye (1977), Downs, Rocke and Barsoom (1996), Heinzel and Liese (2021)

[5]Chayes and Chayes (1993)

[6]Dai (2005)

[7]e.g. Smets, Knack and Molenaers (2013), Reinsberg, Stubbs and Kentikelenis (2022), Lang, Wellner and Kentikelenis (2025)

[8]Girod and Tobin (2016) show that natural resource endowments and foreign direct investment (FDI) affect compliance levels with World Bank loans. Although natural resources are more structural than FDI (Gerring and Christenson, 2017, Chapter 5), FDI is very difficult to change in a short period of time, so it seems germane to label Girod and Tobin's (2016) contribution as structural.

lack of supervision to aid locations subject to the streetlight effects. While previous work on compliance suggests that there is value in combining theories,[9] we do not know of a concept other than the streetlight effects that unites as many theories at once.

Second, the results add nuance to the literature on whether "good countries" or "good projects" are more important for fostering strong aid effectiveness outcomes. Denizer, Kaufmann and Kraay (2013) and Bulman, Kolkma and Kraay (2017) provide the headline findings. They highlight that 75-90% of the variation in project-level aid effectiveness outcomes pertain to within-country features.[10] These headline findings are consistent with the present paper, as the streetlight effect is fundamentally a within-country phenomenon. By the same token, the larger take-away is different. In the debate on "good countries" versus "good projects", the latter generally refer to country-specific design and implementation choices for which aid providers have agency. However, aid providers usually engage in complicated negotiations with recipients,[11] and providers generally do not have full control over which locations within a country receive aid.[12] In this light, the present paper on the streetlight effect showcases that within-country structural issues constrain assumed agency to overcome them. This take-away is consistent with Denly's (2025) argument that scholars' tendency to view many agency failures as shirking in line with the principal-agent model is myopic. Nevertheless, the present paper on the streetlight effect is distinct from Denly's (2025) analysis of project-level safeguard compliance, where there is less tension between recipient state capacity and agency. Only by analyzing compliance spatially can we better understand the agency-structure trade-off.

[9] Tallberg (2002), Börzel et al. (2010)
[10] See also Briggs (2020) and Ashton et al. (2023).
[11] Swedlund (2017), Bennon and Fukuyama (2022)
[12] e.g., Hodler and Raschky (2014), Briggs (2017), Dreher et al. (2019), Song, Brazys and Vadlamannati (2021), and Bommer, Dreher and Perez-Alvarez (2022)

1. Research Design

1.1. Pre-Analysis Plan

Our pre-registration prior to running any analysis is available at OSF.[13] To date, we do not deviate from the pre-analysis plan.

1.2. Dependent Variable

The primary outcome variable for this paper is whether each respective World Bank project complies with the institution's social and environmental safeguard policies. They concern Environmental Assessment; Resettlement; Physical Cultural Resources; Indigenous Peoples; Natural Habitats; Pest Management; Forests; International Waterways; Dams; Disputed Areas; and Environmental Action Plans. We take the variable from Denly (2025), whose coding covers non-supplementary investment lending projects that are subject to safeguard policies for the years 2007-2015.[14] Denly (2025) chooses this time period to capture the years in which World Bank safeguard-related guidance are subject to official policies, not mere directives, prior to the 2016 change in safeguard policies (see Greenstein, 2022). Similar to Buntaine (2016), Denly (2025) codes safeguard compliance on a 1-4 scale, in which 4 represents full compliance; 3 represents moderate compliance; 2 represents moderate non-compliance; and 1 represents full non-compliance. Overall, the data correspond to the safeguard policy compliance outcomes for 1,309 projects approved from 2007-2015. Figure 1 presents relevant summary statistics.

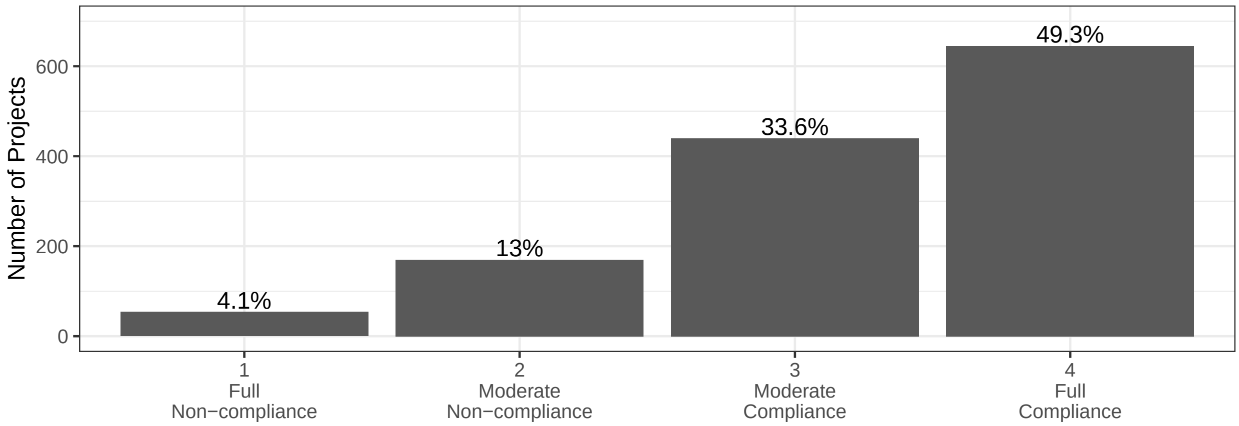
1.3. Unit of Analysis

Given our focus on whether compliance depends on spatial considerations, we explicitly take into account the locations of each projects using data from three data sources. We

[13]See: https://osf.io/ru9dp/?view_only=e7458ea391bd4a90b7ffae2b5baba451.

[14]Supplementary/additional financing projects are not subject to safeguard policy review, which is why Denly (2025) excludes them.

Figure 1: Summary Statistics of World Bank Safeguard Policy Compliance (2007-2015)



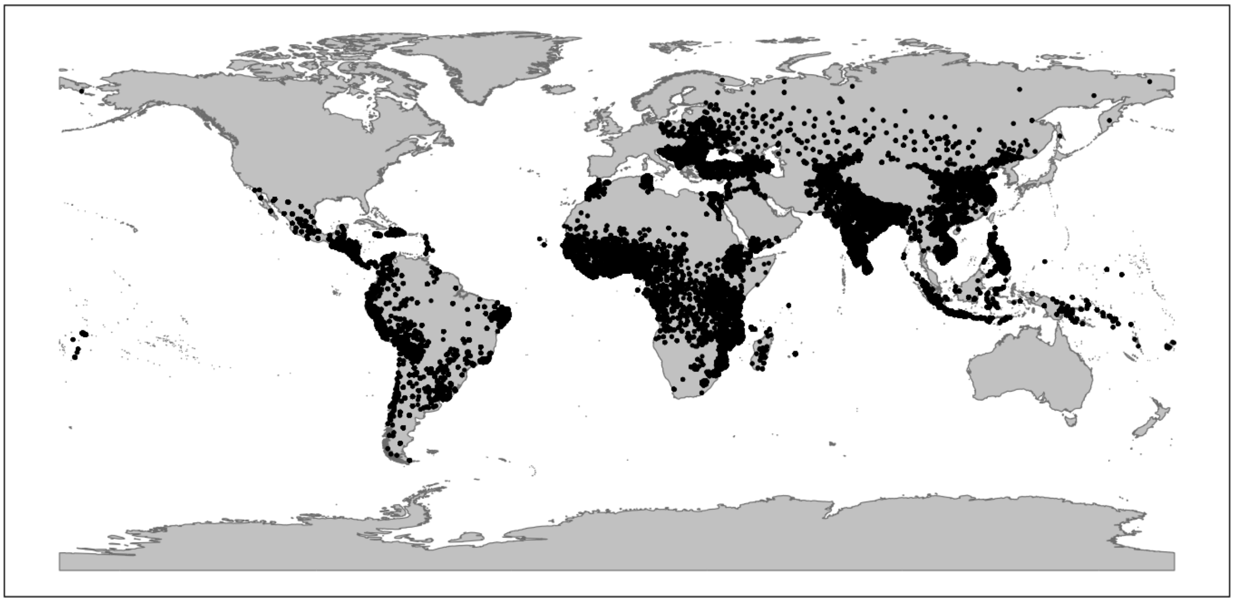
Source: Denly (2025)

draw our baseline spatial data from Bomprezzi et al.'s (2025) Geocoded Official Development Assistance Dataset (GODAD). In cases where GODAD does not include the respective latitudes and longitudes of each project, we take them from the World Bank's external website as well as a transparency request. In total, we are able to obtain location data for all but 21 of the 1,309 projects in Denly's (2025) dataset, corresponding to missingness rate of 1.6%. Figure 2 provides a map of all projects in the dataset.

Because the number of active projects within a given area differs each year as new projects start and older ones close, our primary dataset is spatially balanced. To create the dataset, we create separate rows for each year in which each project location is active, starting with the project approval year and ending with the project closing year. Then, after making location-specific travel time calculations (see Section 1.4.1 below), we use the PRIO GRID to aggregate the project locations data into roughly 55 × 55 kilometers grid cells, corresponding to 0.5 degrees latitude by 0.5 degrees longitude (Tollefsen, Strand and Buhaug, 2012). After aggregating, there are 105,689 project-cell-years in the data.

The resulting project-cell-year dataset offers numerous benefits. First, using grid cells enables us to capture projects located in similar areas with a consistent aggregator. Second, the project-cell-year dataset enable us to capture within-project variation in local condi-

Figure 2: Locations of World Bank Projects



Sources: Bomprezzi et al. (2025) and the World Bank.

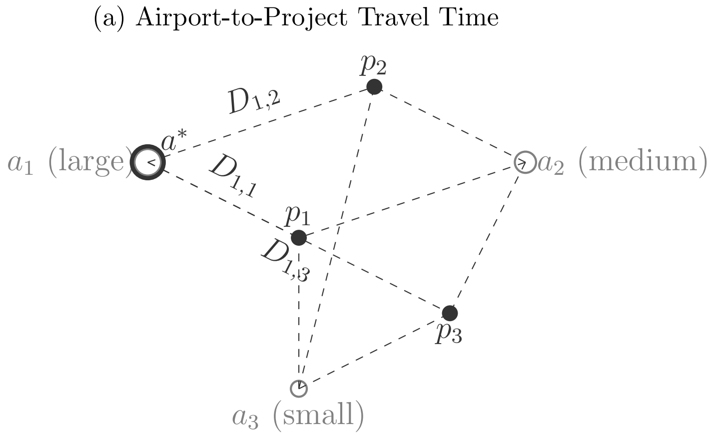
tions. Third, it allow us to merge in numerous spatially-relevant control variables that are not available at the spatial point level but are available at the raster level (see Section 1.4.4). Although the project-cell-year dataset results in repetition of the dependent variable outcomes for the same project, it has a simple solution: weighting by the regressions by the number of times that each project appears in the data. We further address weighting and other model-related issues in Section 1.5.

1.4. Independent Variables

1.4.1. Travel Time Variables

Our main independent variable to capture the streetlight effect is the natural log of travel times to access neighboring World Bank project locations. To meet the computational requirements for these calculations (see Appendix C), we partition the Open-Source Routing Machine into regional servers using maps from OpenStreetMap Foundation (2025). Our calculations, which account for the actual shapes of the roads, stops, and traffic, use three

Figure 3: Airport Selection and Travel Time Algorithms



Note: The above figure provides a graphical illustration of the combined optimal travel time calculations. As we detail in Appendix A.2, we first use an airport selection algorithm. It evaluates project locations p_1, p_2, p_3 (filled circles) against airports a_1 (large), a_2 (medium), and a_3 (small). Then, the algorithm sums the travel times, $D_{j,i}$ to obtain the total travel time for each airport, S_j. Finally, the algorithm selects a^* if $S_j > 0$ and is the lowest. Then, per Appendix A.3, we calculate the nearest neighbor project location, o^*, by minimizing the project-to-project travel time in seconds, $N_{i,*}$, from the closest other project in the same country, o, among all of its locations o_1, o_2, o_3. The total travel time equals $D_{*,i} + N_{i,*}$, for which we take the log to minimize the influence of outliers. Although we present the above figure using straight lines for readability purposes, our actual routes follow the actual shapes of the roads and take into account stops and traffic patterns. Appendix A provides the full mathematical presentation of the calculations.

algorithmic steps that we detail in Appendix A and Figure 3.

In Step 1, we use geolocated data from Megginson (2025) on 48,000 worldwide airports to measure the closest airport to each project's most accessible location. In doing so, we direct the algorithm to search within each country of the project and prioritize large airports over medium-sized ones and medium-sized airports over small-sized ones. The rationale behind the prioritization is that TTLs usually fly into to a larger airport, assuming one exists.

In Step 2, we calculate the driving time to the nearest location of another project, which differs on a yearly basis as new projects start and close. As with the airport-to-project calculations, the project-to-neighboring project calculations are country-specific. In other words, these calculations do not allow for the crossing of country boundaries, even if the route is faster. We adopt such a strategy because World Bank Task Team Leaders rarely supervise projects from different countries in *one supervision mission* without taking a different flight. Although we explain our TTL-specific approach below and in Appendix B, it is not the best way to capture the streetlight effect, so our primary strategy does make such distinctions. Harking back to the Drunkard's Search anecdote, the drunkard reverts to looking the streetlight because it is convenient, not because it helps him find his keys. In a similar way, the best way to capture the streetlight effect is whether there is/are a nearby project(s) (i.e., any streetlight), not specific types of projects (i.e., specific streetlights).

In Step 3, we sum the driving times from Step 1 and Step 2 and take the natural log to obtain the final log airport-project-neighbor total travel times. Table 1 provides relevant summary statistics. We supplement the overall statistics with country-level ones for Bangladesh, Ethiopia, and Kazakhstan to provide face validity to the intricate travel time calculations. Ostensibly, Bangladesh is a much smaller country than Kazakhstan, so it is intuitive that average driving times are shorter in Bangladesh. Similarly, the size of Ethiopia is somewhere between the sizes of Bangladesh and Kazakhstan, making it logical that Ethiopia's summary statistics fall in the middle. With respect to large maximum values for the overall panel, they pertain to projects in the Brazilian Amazon, where the road network is extremely sparse. In any event, we are not concerned by these large values. Even if TTL supervision missions to such places normally involve taking a small plane or helicopter to an airport that is outside of Megginson's (2025) list of circa 48,000 airports, it remains that such a location is very difficult or costly to access and supervise. Furthermore, because our final travel time variables incorporate natural logarithms, we minimize the influence of outliers in very remote locations.

Consistent with our pre-analysis plan, we also calculate TTL-specific travel times that distinguish by the name of the project Task Team Leader (TTL) in each country-year, using data from Denly (2025). The data come from the name of the TTLs at each World Bank project's Implementation Status Reports (ISRs). In the rare cases where TTLs switch within a given calendar year, we take the first chronological TTL in the list. Given that Denly's (2025) data only extend through 2021, we fill down the names of the TTL for projects that are still running in 2022 and 2023. Such instances are also rare: since Denly's (2025) safeguard

Table 1: Summary Statistics for Driving Times

Variable	N	Mean	SD	Min	Max
Panel A: Overall					
Airport to Location	180,116	405.97	581.45	0.00	18,754.20
Location to Nearest Neighbor	179,955	96.03	359.68	0.00	16,961.20
Total Travel Time	176,791	503.80	864.23	0.00	35,715.40
Panel B: Bangladesh					
Airport to Location	8,778	170.38	80.73	12.10	620.80
Location to Nearest Neighbor	8,778	26.05	29.20	0.00	603.30
Total Travel Time	8,778	196.43	94.89	25.10	1,224.10
Panel C: Ethiopia					
Airport to Location	3,096	411.91	227.53	9.40	899.20
Location to Nearest Neighbor	3,096	99.20	92.58	0.00	784.60
Total Travel Time	3,096	511.11	281.97	9.60	1,575.70
Panel D: Kazakhstan					
Airport to Location	837	812.60	594.05	1.20	2,489.50
Location to Nearest Neighbor	837	202.99	205.02	0.00	1,476.50
Total Travel Time	837	1,015.59	696.19	29.60	3,966.00

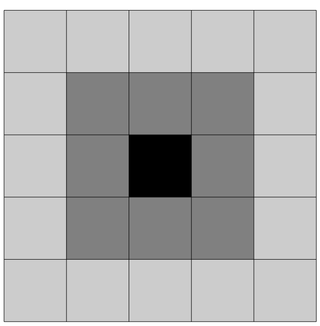
Note: The unit of analysis for these summary statistics is All driving times are in minutes. See Appendix A for the full mathematical details behind our algorithmically-derived calculations.

compliance data refer to projects approved from 2007 to 2015, there are very few projects still running in 2022 or 2023, as World Bank projects take around 5 years to complete. Despite our efforts to avoid missingness, it is impossible to avoid it for these TTL-specific variables. The reason is simple: TTLs do not always supervise more than one project in a respective country at the same time. In total, we are only able to capture 17% of the 180,000 country-location-years with the TTL-specific variables. Given this low share, we use both standardized mean differences (SMDs) and p-values analyze the balance of all pre-treatment covariates in Table B1. We observe that none of the covariates are balanced, as p-values are all < 0.001 and SMDs all exceed $\pm$ 0.5. Against this backdrop, it is clear that the TTL-specific data are not a random sample of all projects and likely pose selection problems that make inference tenuous. Consequently, while we still conduct the relevant analyses to be consistent with our pre-analysis plan, we only report these analyses in Appendix B.

1.4.2. Spatial Grid Count Variables

As an alternative primary independent variable, we calculate the number of active projects within the same year that are (a) immediate neighbors within the same focal point grid cell; (b) first-order neighbors/spatial lags; and (c) second-order neighbors/spatial lags. Consistent with Figure 4, our first- and second-order categories aggregate projects from the previous category. In turn, the aggregated first-order spatial lag category also includes the number of projects in the focal point. Similarly, the aggregated second-order spatial lag includes the number of projects in both the focal point grid cell and first-order neighbors. As with the travel time variables, we also supplement the grid count variables with their TTL-specific counterparts.

Figure 4: Aggregated Spatial Lag Diagram



■ Focal Point

■ Aggregated First-Order

(Focal point + immediate neighbors)

□ Aggregated Second-Order

(Focal point + neighbors + neighbors' neighbors)

1.4.3. Project-Level Variables

Our measure of TTL quality for the interactions comes from Denly (2025), who replicates and extends the earlier coding from Denizer, Kaufmann and Kraay (2013) and Bulman, Kolkma and Kraay (2017). Given that larger projects may require more attention from TTLs, we also control for the commitment amount. Additionally, we control for the safeguard category, given that projects with higher risks may prompt TTLs to dedicate more time to their supervision. We do not control for supervision costs due to the fact that doing so would invoke post-treatment bias—or what Angrist and Pischke (2008) call a "bad control".

1.4.4. Spatial Control Variables

A key set of spatial control variables comes from the HYDE project (Klein Goldewijk, Beusen and Janssen, 2010; Goldewijk et al., 2017). In particular, we use the following different grid-cell-year aggregates: (log) population counts per square kilometer; (log) population density per square kilometer; (log) urban population counts; (log) rural population counts; and (log) built-up area. Unlike other social science studies using HYDE data aggregated by the PRIO GRID (e.g., Briggs, 2018, 2021), we do not use five-year intervals and linear interpolation to fill in missing values. Instead, we rasterize the original HYDE data to obtain yearly values by grid cell. Given that the HYDE variables are all highly correlated, we only include one variable at a time in the regressions. For the meantime, our estimations only include the log population counts per square kilometer to be consistent with Briggs (2018, 2021).

To account for ethnicity, we use the measure of excluded ethnic groups from Vogt et al. (2015). We considered using Denly et al.'s (2022) natural resource measure, but it had too much missingness to include. Finally, we use UCDP's spatial conflict data from Pettersson, Högbladh and Öberg (2019) to account for the effect that conflict may have on ability to travel to an area.

1.4.5. Determining the Final Set of Independent Variables with a DAG

In a future iteration of the paper, we will construct a Directed Acyclic Graph (DAG) to characterize the relationships between the various independent variables. After analyzing potential mediation and collider relationships, we will specify a final set of covariates to include in the model. For now, we include separate specifications with the bivariate regressions, models including all project-level controls, and models including all spatial controls.

1.5. Statistical Models

Consistent with Berman et al. (2017) and Denly et al. (2022), we use a spatial heteroskedastic and autocorrelation consistent (HAC) linear regression model:

$$\text{compliance}_{(k,i,t)} = \beta_{\text{log travel time or nearby projects}_{(k,i,t)}} + \beta_{\text{controls}_{(k,i,t)}} + FE + \epsilon_{(k,i,t)} \quad (1)$$

where subscript k refers to the respective grid cell, i corresponds to the project, t denotes the time/year, ϵ is an error term, and FE corresponds to fixed effects. With respect to the latter, we consider fixed effects at the country and year levels. Within-grid cells effects are irrelevant for the question and estimand of interest, so we do not include grid cell fixed effects. Fundamentally, the study concerns whether distances away from projects yield lower compliance levels. Because grid cell fixed effects would pull interpretation back within the same grid cell, the resulting estimand would become difficult to interpret and irrelevant for the research question.

Given our focus on spatial driving time to nearby projects, the most appropriate standard error specification entails the use of Conley (1999) standard errors. To account for spatial dependence, we specify distance cutoffs for potential spatial correlation of 150 kilometers, thereby capturing grid cells within the second to third order spatial lag of each other (see Figure 4). As a robustness measure, we also consider spatial cutoffs of 75 kilometers not only because it is half the 150 threshold, but also because the mean project-to-neighbor distance is 82 kilometers. Using such a structure with Conley standard errors has three key benefits. First, Conley standard errors capture the spatial nature of our estimations, which would be prone autocorrelated residuals without an appropriate spatial specification. Second, given that locations from the same project tend to be located within the same area, Conley standard errors allow for addressing project-specific serial correlation—i.e, similar to clustering the standard errors at the project level. Third, using Conley standard errors avoids ambiguous decisions concerning whether to cluster at the level(s) of the grid cell, project, country, year, or the 24 combinations thereof.[15]

We also address the fact that Denly (2025) only measures the dependent variable, compliance with safeguard policies, once for each project. Our solution entails weighting the regressions by the number of times the project appears in the project-cell-year data. Although the solution is simple, it accomplishes the end goal of ensuring an appropriate effective sample size. In turn, the weighting solution ensures accurate calculation of standard errors and t-statistics, as failure to weight would result in inflation of the latter and deflation of the former.

Finally, our interaction models use a spatial similar HAC model as Equation (1):

$$\begin{aligned}\text{compliance}_{(k,i,t)} =& \beta_{\text{log travel time or nearby projects}_{(k,i,t)}} + \beta_{\text{TTL Quality}_{(k,i,t)}} + \\ & \beta_{\text{log travel time or nearby projects} \times \text{TTL Quality}_{(k,i,t)}} + \beta_{\text{controls}_{(k,i,t)}} + FE + \epsilon_{(k,i,t)}\end{aligned} \tag{2}$$

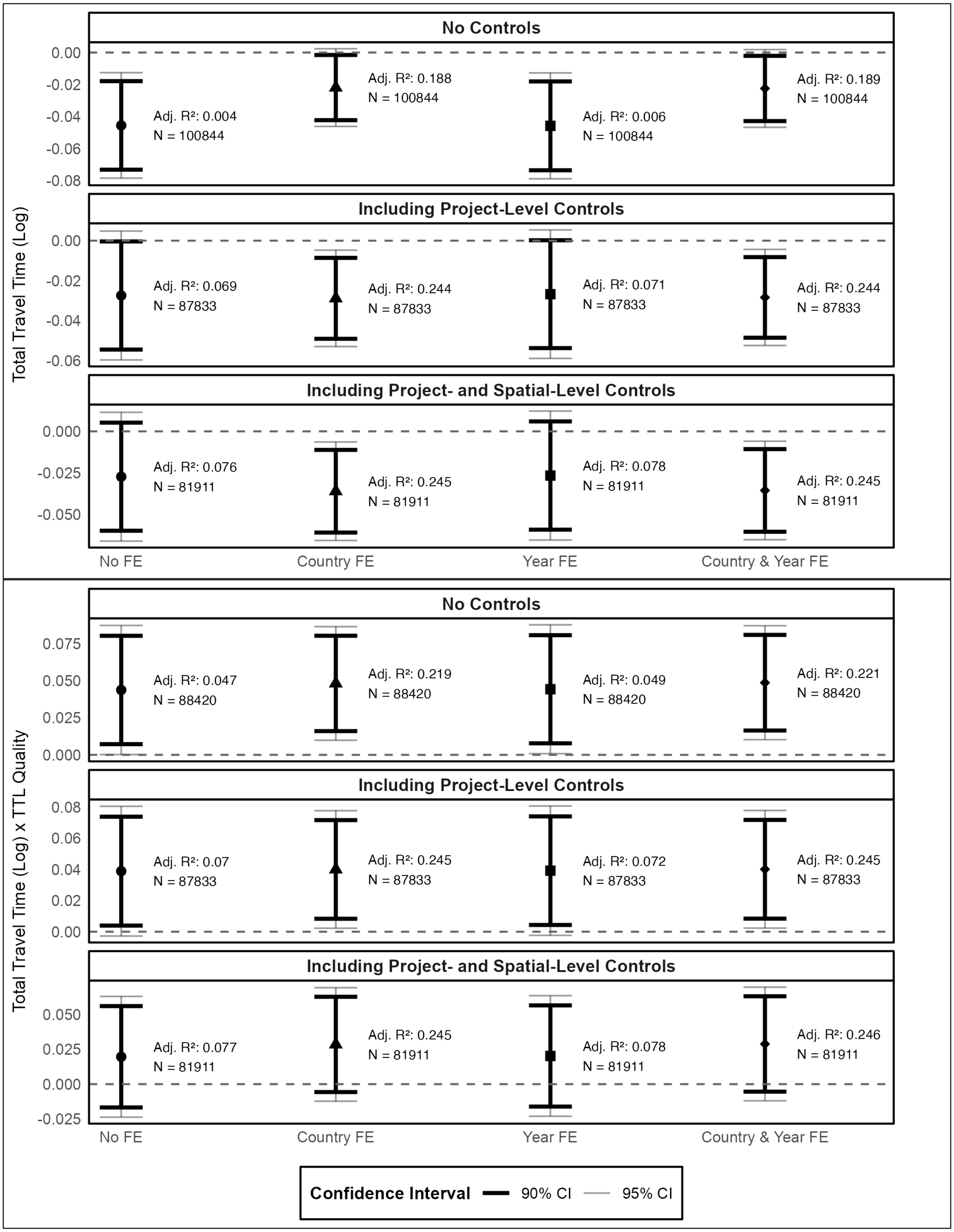
2. Results

2.1. Travel Time Results

Figure 5 presents the main results to test the streetlight effect hypothesis concerning whether larger (log) travel times decrease compliance with the World Bank's social and environmental safeguards policies. Overall, although some specifications indicate lower levels of statistical significance than others, the results are all directionally consistent and suggestive of the hypothesis. Indeed, most specifications suggest that a one percent increase in travel time yields a reduction in compliance levels by circa 0.02-0.04. Given that compliance is measured on a 1-4 scale, that corresponds to 0.5-1% decrease in compliance. By extension, a 10% increase in travel time results in a 5-10% decrease in compliance levels, which is a substantively meaningful effect size.

[15] $4 \times 3 \times 2 \times 1 = 24$

Figure 5: Main Travel Time Results



Among all of the models, the most essential ones include country fixed effects. Given that the driving times calculations do not allow for routes to leave the respective project country, it is very difficult to interpret the models without country fixed effects. We can further ascertain their importance by comparing the Adjusted R^2 values for each model. Of particular note is that the Adjusted R^2 is essentially identical in each model containing both country fixed effects as well as those with both country and year fixed effects. This suggests that the year fixed effects, which also contain the most volatile results, are not driving outcomes, are likely unnecessary for estimation, and may be producing noise. Against this backdrop, we are confident that the effect is real.

The bottom panel of Figure 5 also reports the results of the interaction models. The ones with no controls and only project-level controls suggest that higher-quality TTLs have agency to overcome the structural constraints of larger log travel times effects. However, the effects dissipate once we add spatial-level controls to complement those at the project level. Not only do the effect sizes decrease from 0.04 to 0.02 in these models with all covariates, but the effects also become no longer statistically significant at the 90% percent confidence interval. Overall, the results suggest that higher-quality TTLs only have some inconsistent agency to overcome the streetlight effect, which is a product of structure and agency.

2.2. Spatial Grid Results

Coming soon.

3. Conclusion

Coming soon.

References

Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, New Jersey: Princeton University Press.

Ashton, Louise, Jed Friedman, Diana Goldemberg, Mustafa Zakir Hussain, Thomas Kenyon, Akib Khan and Mo Zhou. 2023. "A Puzzle with Missing Pieces: Explaining the Effectiveness of World Bank Development Projects." *World Bank Research Observer* 38:115–146.

Bennon, Michael and Francis Fukuyama. 2022. "The Obsolescing Bargain Crosses the Belt and Road Initiative: Renegotiations on BRI Projects." *Oxford Review of Economic Policy* 38(2):278–301.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner and Mathias Thoenig. 2017. "This Mine is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review* 107(6):1564–1610.

Bommer, Christian, Axel Dreher and Marcello Perez-Alvarez. 2022. "Home Bias in Humanitarian Aid: The Role of Regional Favoritism in the Allocation of International Disaster Relief." *Journal of Public Economics* 208(104604):1–16.

Bomprezzi, Pietro, Axel Dreher, Andreas Fuchs, Teresa Hailer, Andreas Kammerlander, Lennart C. Kaplan, Silvia Marchesi, Tania Masi, Charlotte Robert and Kerstin Unfreid. 2025. "Wedded to Prosperity? Informal Influence and Regional Favoritism.".

Börzel, Tanja A., Tobias Hofmann, Diana Panke and Carina Sprungk. 2010. "Obstinate and Inefficient: Why Member States Do Not Comply with European Law." *Comparative Political Studies* 43(11):1363–1390.

Briggs, Ryan C. 2017. "Does Foreign Aid Target the Poorest?" *International Organization* 71(01):187–206.

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

Briggs, Ryan C. 2020. "Results from Single-Donor Analyses of Project Aid Success Seem to Generalize Pretty Well across Donors." *Review of International Organizations* 15(4):947–963.

Briggs, Ryan C. 2021. "Why Does Aid Not Target the Poorest?" *International Studies Quarterly* 65(3):739–752.

Bulman, David, Walter Kolkma and Aart Kraay. 2017. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank and Asian Development Bank Project Performance." *Review of International Organizations* 12:335–363.

Buntaine, Mark T. 2016. *Giving Aid Effectively: The Politics of Environmental Performance and Selectivity at Multilateral Development Banks.* Oxford: Oxford University Press.

Chayes, Abram and Antonia Handler Chayes. 1993. "On Compliance." *International Organization* 47(2):175–205.

Conley, Timothy G. 1999. "GMM Estimation with Cross-sectional Dependence." *Journal of Econometrics* 92(1):1–45.

Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59(02):363–398.

Denizer, Cevdet, Daniel Kaufmann and Aart Kraay. 2013. "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance." *Journal of Development Economics* 105(2013):288–302.

Denly, Michael. 2025. "Preventing the Negative Externalities of Development: Aid Compliance, State Capacity, and Incomplete Contracts.".
URL: *https://mikedenly.com/research/aid-compliance-externalities*

Denly, Michael, Michael G. Findley, Joelean Hall, Andrew Stravers and James Igoe Walsh. 2022. "Do Natural Resources Really Cause Civil Conflict? Evidence from the New Global Resources Dataset." *Journal of Conflict Resolution* 66(3):387–412.

Downs, George W., David M. Rocke and Peter N Barsoom. 1996. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50(3):379–406.

Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C. Parks, Paul A. Raschky and Michael J. Tierney. 2019. "African Leaders and the Geography of China's Foreign Assistance." *Journal of Development Economics* 140(August 2018):44–71.

Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52(4):887–917.

Gerring, John and Dino P. Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press.

Girod, Desha M. and Jennifer L. Tobin. 2016. "Take the Money and Run: The Determinants of Compliance with Aid Agreements." *International Organization* 70(1):209–239.

Goldewijk, Kees Klein, Arthur Beusen, Jonathan Doelman and Elke Stehfest. 2017. "Anthropogenic Land Use Estimates for the Holocene - HYDE 3.2." *Earth System Science Data* 9(2):927–953.

Greenstein, Gus. 2022. "The Influence of Alternative Development Finance on the World Bank's Safeguards Regime." *Global Environmental Politics* 22(3):171–193.

Heinzel, Mirko and Andrea Liese. 2021. "Managing Performance and Winning Trust: How World Bank Staff Shapes Recipient Performance." *Review of International Organizations* 16:625–653.

Hendrix, Cullen S. 2017. "The Streetlight Effect in Climate Change Research on Africa." *Global Environmental Change* 43:137–147.

Hodler, Roland and Paul A. Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics* (September):995–1033.

Keohane, Robert and Joseph Nye. 1977. *Power and Interdependence: World Politics in Transition*. Boston: Little, Brown, and Company.

Klein Goldewijk, Kees, Arthur Beusen and Peter Janssen. 2010. "Long-term Dynamic Modeling of Global Population and Built-up Area in a Spatially Explicit Way: HYDE 3.1." *Holocene* 20(4):565–573.

Lang, Valentin F., Lukas Wellner and Alexander E. Kentikelenis. 2025. "Biased Bureaucrats and the Policies of International Organizations." *American Journal of Political Science* .

Megginson, David. 2025. "Our Airports.".
URL: *https://ourairports.com*

OpenStreetMap Foundation. 2025. "geofabrik.".
URL: *https://www.geofabrik.de/*

Pettersson, Therése, Stina Högbladh and Magnus Öberg. 2019. "Organized Violence, 1989–2018 and Peace Agreements." *Journal of Peace Research* 56(4):589–603.

Reinsberg, Bernhard, Thomas H. Stubbs and Alexander E. Kentikelenis. 2022. "Unimplementable by Design? Understanding (Non-)Compliance with International Monetary Fund Policy Conditionality." *Governance* 35:689–715.

Smets, Lodewijk, Stephen Knack and Nadia Molenaers. 2013. "Political Ideology, Quality at Entry and the Success of Economic Reform Programs." *Review of International Organizations* 8(4):447–476.

Song, Tianyang, Samuel Brazys and Krishna Chaitanya Vadlamannati. 2021. "Which Wheel Gets the Grease? Constituent Agency and Sub-national World Bank Aid Allocation." *Journal of Development Studies* 57(3):519–533.

Swedlund, Haley J. 2017. *The Development Dance: How Donors and Recipients Negotiate the Delivery of Foreign Aid*. Ithaca, New York: Cornell University Press.

Tallberg, Jonas. 2002. "Paths to Compliance: Enforcement, Management, and the European Union." *International Organization* 56(3):609–643.

Tello, Rodolfo. 2015. *Social Safeguards: Avoiding the Unintended Impacts of Development*. Arlington, VA: Amakella Publishing.

Tollefsen, Andreas, Håvard Strand and Halvard Buhaug. 2012. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* 49(2):363–374.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59(7):1327–1342.

Appendices

A. Mathematical Framework

A.1. Input Data

Let:

- $P = \{p_1, p_2, \ldots, p_n\}$ be the set of project locations for a given project, where each p_i refers to the latitude and longitude coordinates of each project's location.
- $A = \{a_1, a_2, \ldots, a_m\}$ be the set of airports in the same country, where each a_j uses data from Megginson (2025) to denote each airport's latitude and longitude coordinates as well as its size in terms of whether it is `large`, `medium`, or `small`.
- $O = \{o_1, o_2, \ldots, o_k\}$ be the set of other project locations in the same country and year, excluding the current project, where o_i refers to the latitude and longitude coordinates of each other project's location.

A.2. Nearest Airport Selection

For each project, the algorithm selects one airport $a^* \in A$ that minimizes the total travel time to all project locations P while ensuring at least one non-zero travel time. The algorithm evaluates airports in the following priority order: `large`, then `medium`, and finally `small`.

We define the travel time matrix D from OSRM as:

$$D_{j,i} = \text{travel time (seconds) from } a_j \text{ to } p_i$$

We compute the total travel time for airport a_j by

$$S_j = \sum_{i=1}^{n} D_{j,i}$$

Then, we select the airport a^* as:

$$a^* = \arg\min_{a_j \in A_{\text{size}}} S_j \quad \text{such that} \quad S_j > 0$$

where A_{size} represents the subset of airports for the current size being evaluated. The minimization first considers `large` airports, then `medium`, and finally `small` until a valid a^* is found. If no a^* with $S_j > 0$ exists, the related fields are set to missing. For each location p_i, if a^* is selected, we compute:

$$\text{travel time in minutes}_i = \frac{D_{*,i}}{60}$$

where $*$ denotes the index of a^*, respectively. If $D_{*,i} = 0$ or if no a^* is found, we assign it as missing.

A.3. Neighbor Calculations

For each location p_i, we identify the nearest nieghbor project $o^* \in O$ within the same country. We define the neighbor travel time N from OSRM as:

$$N_{i,l} = \text{travel time (seconds) from } p_i \text{ to } o_l$$

We next determine the nearest general neighbor, o^*, by

$$o^* = \arg\min_{o_l \in O} N_{i,l}$$

Then, we calculate the neighbor time as follows:

$$\text{neighbor time} = \frac{N_{i,*}}{60}$$

Next, we compute the total travel time in minutes as follow:

$$\text{total time minutes}_i = \text{travel time in minutes}_i + \text{neighbor time}$$

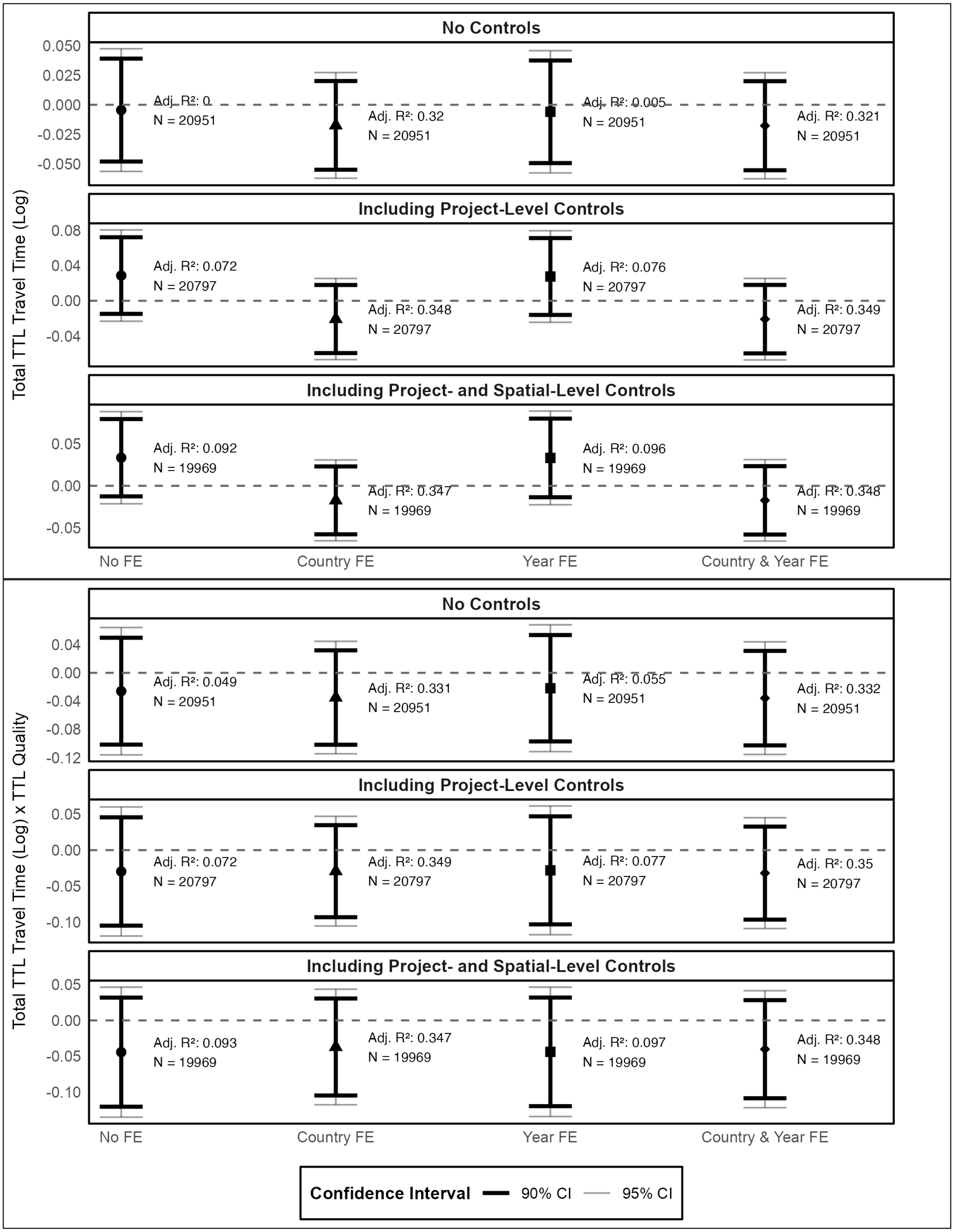
B. TTL-Specific Analyses

Table B1: Balance Table Describing the Full Sample vs. TTL-Specific Sample

Variable	Full Mean	Full SD	TTL Mean	TTL SD	p-value	SMD
TTL	4.082	0.592	4.201	0.538	<0.001	-0.210
Safeguard Category	2.056	0.723	1.924	0.846	<0.001	0.168
Commitment (Log)	18.323	1.146	18.659	1.088	<0.001	-0.300
Ethnic Groups	2.110	1.335	2.181	1.377	<0.001	-0.053
Conflict Count	0.149	0.532	0.103	0.418	<0.001	0.096
Population (Log)	8.871	1.930	9.367	1.769	<0.001	-0.268

The results from Figure B1 regarding the sample of projects where TTLs supervise more than one project in a particular country are much less robust than those of the main analysis. On the one hand, given that coefficients for these estimations are neither directionally consistent nor statistically significant, it is clear that there is no clear relationship in this sample. On the other hand, as we specified earlier, only 17% of the project-cell-year observations include TTLs supervising multiple projects within the same country. These observations are also not a random sample of the full sample (see Table B1), suggesting series selection-related problems that are likely worthy of a whole new paper. Accordingly, we prefer not to lend strong interpretation to the results from Figure B1, which we only report to be consistent with our pre-analysis plan.

Figure B1: Results for the TTL-Specific Sample



C. Local Server Setup and Data Processing Steps

C.1. Introduction

The Open Source Routing Machine (OSRM) computes optimal routes across large-scale road networks using OpenStreetMap data. Advanced algorithms, such as contraction hierarchies, accelerate query times by precomputing shortcut edges in the network. OSRM software and data are sourced from Geofabrik, which organizes OpenStreetMap data into Protocolbuffer Binary Format (PBF) files by region.

C.2. Rationale for a Local Server

The extraction and contraction phases in OSRM demand intensive computations that online test servers cannot handle. A local server processes data and serves routing queries efficiently, enabling the engine to operate on high-resolution maps without the limitations imposed by remote servers.

C.3. Data Acquisition and Preparation

We download Regional PBF files directly from Geofabrik (OpenStreetMap Foundation, 2025). This up-to-date data forms the foundation of the routing engine and ensures accuracy in representing the latest road network information.

C.4. Server Setup Process

The setup utilizes Docker in PowerShell and follows two key phases:

1. **Extraction:** We extract raw geographic data from the PBF files, converting it into a structured graph representation of the road network—an essential step for routing calculations.

2. **Contraction:** We then contract the data by precomputing shortcuts, reducing the number of nodes and edges during routing queries. This optimization significantly

speeds up the query process by minimizing the search space.

Both phases leverage parallel processing to fully utilize multi-core CPU architectures.

C.5. Hardware Considerations

Our data processing runs on a system equipped with:

- A 12th generation Intel processor (using 8 out of 14 available cores),
- 40 GB of DDR4 RAM, and
- 200 GB of swap space provided by an external solid state drive (SSD).

C.6. Routing Profiles and Computation

The `car.lua` profile computes realistic travel metrics by integrating data from Geofabrik. It calculates:

- **Travel Time Estimates:** Provided in minutes, these estimates consider speed limits and road conditions.
- **Travel Distances:** Provided in kilometers, these measurements reflect the actual distance traveled.

The algorithm factors in critical elements, including traffic signals, stop signs, and other road features that influence travel times. This approach produces realistic and practical route estimates.