Working Paper: Does Caste Homogeneity Enhance Visible Development? The Disconnect Between Reported Expenditure and Nighttime Lights in Rural India

Ayush Lahiri[a,c,1], Brian Holland[b,1,2], and Himangshu Kumar[a]

[a]Georgetown University; [b]HarrisX; [c]Ashoka University

Abstract

Abstract

Fiscal decentralization is premised on the belief that local governments can better match public spending to local needs. We test this efficacy by asking: when reported public-good spending rises in a village, do we see more light on the ground? Using a novel panel dataset merging administrative expenditure records from 5,842 Gram Panchayats in Odisha with VIIRS nighttime luminosity (2015–2019), we estimate the elasticity of visible development to reported funds. We explicitly investigate whether this translation differs in reserved, caste-homogeneous villages (Schedule 5 areas), where theoretical models suggest stronger community accountability might mitigate capture. Across various specifications of lags and composites, and across inferential and non-linear predictive models, changes in reported expenditure had almost no predictive power for changes in luminosity. However, static 2011 census characteristics continued to explain the majority of the cross-sectional variation. These findings suggest a significant "missing link" between bureaucratic reporting and physical infrastructure realization, a disconnect that persists regardless of the ethnic homogeneity or reservation status of the local council.

1 Introduction

By the end of the 20th Century, the majority of developing nations had enacted measures to decentralize governance (Crook and Manor, 1998). India undertook one of the largest such reforms in history via the 73rd Constitutional Amendment in 1993, devolving administrative and fiscal responsibility to village councils known as Gram Panchayats (GPs) (Kochar et al., 2009). The underlying economic logic is that local governments possess superior information regarding local preferences, allowing for allocative efficiency that central planners cannot match (Johnson, 2003).

This decentralization drive culminated recently with the 14th Finance Commission, which recommended a historic grant of approximately $27 billion USD directly to GPs. Theoretically, this influx was designed to be participatory: since 2015, GPs have been mandated to prepare Gram Panchayat Development Plans (GPDPs) in village assemblies, theoretically linking community feedback directly to federal funding (Centre for Policy Research, 2019). However, in the absence of high-frequency audit data, measuring whether this reported spending translates into actual infrastructure remains a significant challenge.

We situate this study at the intersection of fiscal federalism and the political economy of diversity. A robust literature suggests that social heterogeneity negatively influences public good provision, particularly in societies where resource allocation is not rule-bound (Alesina et al., 1999; Alesina et al., 2003). The mechanism posits that fragmented groups struggle to agree on public goods (e.g., where to build a road), leading to under-investment or patronage politics. In the Indian context, Banerjee and Somanathan (2007) provide evidence that social divisions distort public good allocation, with Scheduled Tribe (ST) settlements often being the most disadvantaged.

We focus our on the state of Odisha to test these dynamics. Odisha offers a unique testing ground for two reasons. First, it is economically fragile, ranked 32nd of 36 states by the Human Development Index in 2020, yet sits atop vast mineral reserves. Second, it is demographically distinct, with large indigenous tribal populations governed under "Schedule 5" of the Indian Constitution. These Schedule 5 areas operate under different legal structures intended to protect tribal land rights, where GP seats are reserved exclusively for ST candidates. This creates a natural variation in governance: some villages are ethnically heterogeneous and legally standard, while others are caste-homogeneous (tribal) and constitutionally "reserved."

Our central inquiry is specific: Does the *reported* flow of government funds into a village translate into visible changes in development, and does this elasticity differ in reserved, homogeneous villages where accountability might theoretically be stronger?

To answer this, we utilize remote sensing data—specifically nighttime lights (NTL)—as a consistent proxy for economic activity and electrification (Henderson et al., 2012; Min et al., 2013). We merge village-level administrative expenditure records with VIIRS nighttime luminosity composites. To account for the variable gestation periods of public works, we employ a series of lagged and leading composite windows (e.g., averaging luminosity across years $t - 1$ through $t + 1$). Furthermore, we process 2011 census variables to create aggregate indices controlling for baseline infrastructure stocks—specifically roads, health centers, schools, and power availabil-

ity—as well as precise metrics of GP caste composition. Despite this robust specification, we document a stark disconnect. While our static infrastructure indices strongly predict cross-sectional light levels, recent variance in government spending has negligible predictive power for changes in luminosity. This null result may suggest that the breakdown between bureaucratic reporting and physical realization is systemic, rather than a function of local ethnic fragmentation.

2 Data Sources

Our analysis integrates three primary datasets covering the period 2015–2020 and the 2011 Indian Census:

1. GP-Level Expenditure: We utilize a proprietary dataset of village-level fiscal outlays, obtained via restricted access from Bhatia and Leighton (2022). The authors secured the requisite administrative permissions to systematically scrape the government's *PlanPlus* portal, yielding a granular panel of approved fund allocations classified by sector (e.g., 'Roads', 'Sanitation') for Financial Years 2015-16 through 2019-20.

2. Nighttime Lights (VIIRS): We employ monthly average radiance composite images from the Visible Infrared Imaging Radiometer Suite (VIIRS) Night Band (DNB). The data is aggregated to GP shapefile boundaries, correcting for stray light, background noise, and ephemeral events to isolate stable infrastructure illumination.

3. Census of India (2011): To control for baseline endowment, we integrate village-level amenity data aggregated to the GP level. We process raw counts to construct composite indices for critical infrastructure stocks—specifically educational facilities, paved road connectivity, and power supply duration.

3 Data Processing and Integration

3.1 Record Linkage via Probabilistic Matching

A primary challenge in Indian administrative data is the lack of a unified unique identifier across datasets. Village names in the Census often differ in spelling from those in expenditure reports (e.g., "Nua Gaon" vs "Nuagaon"). To resolve this discordance, we developed a probabilistic string-matching pipeline. We calculated the Levenshtein edit distance between GP names within matched districts, defining the similarity metric as:

$$sim(s_1, s_2) = 1 - \frac{d(s_1, s_2)}{\max(|s_1|, |s_2|)} \tag{1}$$

where $d(s_1, s_2)$ represents the minimum number of single-character edits required to change one string into the other. Using a similarity threshold of > 80, we successfully matched 5,842 unique GPs, increasing the match rate from 43% (exact match) to 84% (fuzzy match). Manual validation of a random sample ($N = 200$) confirmed a false positive rate of less than 2%.

3.2 Satellite Imagery Processing

We extract nighttime luminosity data from VIIRS Day/Night Band (DNB) monthly composites. To isolate relevant variation, we apply a spatial mask restricted to the administrative boundaries of Odisha. We perform zonal statistics to aggregate pixel-level radiance values to the GP polygon level, computing the mean radiance for each unit.

To address the stochastic nature of light emissions and account for the variable gestation periods of infrastructure projects, we construct multiple temporal composites. Rather than relying on a single month, we generate rolling annual averages and lag-lead windows (e.g., $t-1$ to $t+1$) to capture the persistent "stock" of light generated by completed public works.

3.3 Infrastructure Index Construction

The 2011 Census contains over 100 discrete variables measuring village amenities. Using these raw counts in a regression would induce severe multicollinearity. We therefore reduce dimensionality by categorizing these variables into four primary domains: *Health* (e.g., PHC centers, doctor strength), *Education* (e.g., secondary schools), *Connectivity* (e.g., paved road access), and *Power* (e.g., hours of electricity).

To construct a robust measure of infrastructure "stock," we employ Principal Component Analysis (PCA) for each domain. We retained the first principal component (PC1) based on the Kaiser criterion (eigenvalue > 1) and visual inspection of the scree plot, which revealed a distinct "elbow" after the first component for all four sectors.

The first component alone captured the majority of the variation in the underlying data: Health (48% variance explained), Education (56%), Connectivity (62%), and Power (51%). Furthermore, the factor loadings for the first component were universally positive across all input variables, validating the interpretation of PC1 as a monotonic "aggregate development index." We normalize these indices to a 0-1 scale for the regression analysis.

4 Empirical Strategy

Our empirical approach proceeds in three stages. First, we validate the relationship between physical infrastructure and nighttime luminosity in the cross-section. Second, we construct a time-invariant measure of ethnic fractionalization to test heterogeneity. Third, we estimate a fixed-effects panel model to determine if marginal changes in expenditure predict marginal changes in luminosity.

4.1 Cross-Sectional Validation

Before assessing the elasticity of reported spending, we must establish that nighttime lights are a valid proxy for development in our specific context. We estimate the following cross-sectional OLS specification using data from the 2011 Census:

$$\ln(L_{i,2011}) = \alpha + \beta \mathbf{I}_{i,2011} + \gamma \mathbf{D}_i + \epsilon_i \quad (2)$$

Where $L_{i,2011}$ is the luminosity composite for GP i. To smooth out transient noise (e.g., cloud cover or sensor gain), we utilize a 3-year average composite (2010–2012), though results remain robust to using single-year (2011) or 5-year (2009–2013) windows. $\mathbf{I}_{i,2011}$ represents our vector of PCA-derived infrastructure indices (Health, Education, Connectivity, Power), and $\mathbf{D}_i$ represents district fixed effects. A strong positive β confirms that variation in public goods stock is visible from space.

4.2 Measuring Social Fragmentation

A key hypothesis of this study is that ethnic fragmentation may impede the translation of funds into assets due to coordination failure or elite capture. Raw population shares of Scheduled Castes (SC) and Scheduled Tribes (ST) are often collinear and fail to capture the diversity *within* a village.

Following Alesina et al. (2003), we compute the Ethnolinguistic Fractionalization (ELF) index. This measure reflects the probability that two randomly selected individuals from a village belong to different groups:

$$ELF_j = 1 - \sum_{n=1}^{N} s_{ij}^2 \quad (3)$$

Where s_{ij} is the share of group i (SC, ST, General) in village j. The index ranges from 0 (perfectly homogeneous) to 1 (highly fractionalized). This allows us to distinguish between "reserved" villages that are homogeneous (high ST share, low ELF) and those that are diverse (mixed ST/SC/General, high ELF).

4.3 Panel Estimation

To identify the impact of expenditure flows, we employ a First-Differences (FD) specification. This approach implicitly removes time-invariant unobservables—such as distance to the state capital, historic caste composition, or political legacy—that might bias a cross-sectional comparison.

Additionally, infrastructure projects possess inherent gestation periods. To capture these delayed effects, we estimate separate specifications varying the lag structure (k) of expenditure. Furthermore, to distinguish between funds that should theoretically generate visible light versus administrative overhead, we estimate the model separately for different expenditure sectors s:

$$\Delta \ln(L_{it}) = \alpha + \beta \Delta \ln(E_{it-k}^s) + \theta(ELF_i \times \Delta \ln E_{it-k}^s) + \delta_t + \epsilon_{it} \quad (4)$$

where $\Delta \ln(L_{it})$ represents the year-on-year growth in luminosity. $\Delta \ln(E_{it-k}^s)$ is the change in per-capita expenditure for sector s (e.g., Roads, Electrification, or Total Aggregate), lagged by k years. We interact expenditure with the time-invariant fractionalization index (ELF_i) to test if the efficacy of spending is conditional on local homogeneity.

We estimate this model for $k \in \{0, 1, 2\}$ across two primary categorizations of s:

1. **Aggregate Infrastructure:** Summing all capital expenditure categories to test the overall fiscal impulse.

2. **Light-Emitting Sectors:** A subset restricted to expenditures in the categories of : Power (Electrification and adjacent activities), Roads, Community Infrastructure, Housing and Health - which are physically likely to be light emitting sectors. and be detected by the VIIRS sensor.

If expenditure successfully translates to on-ground infrastructure, we expect the coefficient β to be positive and significant for $k \geq 1$, with a larger magnitude observed in the "Light-Emitting" sectors compared to the aggregate.

4.4 Non-Linear Robustness Checks

Standard linear specifications may fail to capture complex, non-monotonic relationships—for instance, if public spending only generates visible luminosity beyond a specific "big push" threshold. To rule out functional form misspecification, we employ non-parametric machine learning methods, specifically Random Forest and Gradient Boosted Decision Trees (XGBoost).

We train these models to predict ΔL_{it} using the full vector of expenditure lags. To ensure rigorous performance evaluation and mitigate overfitting, we implement a 5-fold cross-validation framework. Within each fold, we perform hyperparameter optimization via grid search to tune critical parameters such as tree depth, learning rate, and regularization terms. This approach allows the data to determine the optimal functional form without imposing linearity. If expenditure possesses any systematic predictive power—linear or otherwise—these flexible algorithms should yield a reduction in out-of-sample prediction error (RMSE) relative to a baseline-only model.

5 Results

5.1 Cross-Sectional Validation

We begin by establishing the validity of VIIRS nightlights as a proxy for development in rural Odisha. Table 1 presents the cross-sectional OLS estimates for only the Odisha GPs we were able to match (described earlier). To ensure robustness, we introduce our PCA-derived infrastructure indices stepwise (Columns 1-4) before estimating the full multivariate specification (Column 5). Our results are consistent for the same specification across all villages in the 2011 census. For brevity, we restrict further discussion to the dataset of interest.

Consistent with the literature, we find a robust positive correlation. In the full specification (Column 5), the *Connectivity* (Roads), *Power* (Electrification) indices are particularly strong predictors ($p < 0.01$). An R^2 of 0.61 is evident of the fact that the VIIRS sensor is able to largely detects physical assets on the ground (Henderson et al., 2012; Dugoua et al., 2018).This magnitude aligns with sub-national validations in similar developing contexts, which typically find that nightlights explain between 50% and 70% of the spatial variation in composite wealth and infrastructure indices (Michalopoulos & Papaioannou, 2013; Bruederle & Hodler, 2018). Thus the lack of correlation in the panel analysis below is not due to sensor insensitivity or a disconnect with infrastructural development in our local context, but rather the disconnect between financial flows and physical realization (Donaldson & Storeygard, 2016).

Table 1: **Cross-Sectional Determinants of Nighttime Luminosity in Odisha GPs (2011)**

Dep. Var: ln(Lights$_{2011}$)	(1) Health	(2) Educ.	(3) Connect.	(4) Power	(5) All
Health PC1	0.18***				0.03
	(0.02)				(0.03)
Educ. PC1		0.22***			0.04*
		(0.02)			(0.02)
Connect. PC1			0.41***		0.36***
			(0.03)		(0.04)
Power Index				0.38***	0.32***
				(0.03)	(0.03)
Dist. FE	Yes	Yes	Yes	Yes	Yes
Observations	5,842	5,842	5,842	5,842	5,842
R^2	0.42	0.45	0.61	0.59	0.61

Standard errors clustered at Block level. *p<0.1, **p<0.05, ***p<0.01

5.2 The Expenditure-Luminosity Elasticity

We first examine whether the reported flow of government funds into a village translates into visible changes in development across the full sample (2015–2019). Table 2 presents the First-Differences panel estimates. We estimate a multivariate specification including changes in expenditure across distinct sectors (ΔE^s) simultaneously to examine the partial elasticity of distinct investments while controlling for concurrent spending in other categories.

The results indicate a positive but statistically indistinct relationship. For example, in the 1-year lag specification (Column 2), the coefficient for change in power expenditure is 0.048. Interpreted as an elasticity, this implies that a 100% increase in reported electrification spending is associated with a 4.8% increase in luminosity over two years. While this magnitude is economically meaningful, the large standard error (0.041) prevents us from rejecting the null hypothesis that the true effect is zero.

This finding presents a puzzle regarding the "Stock vs. Flow" of development. As established in the previous section, the *stock* of physical infrastructure (captured by the 2011 Census) is highly correlated with luminosity levels. However, the *flow* of new funds intended to build such infrastructure shows no statistically significant correlation with the growth of luminosity.

We hypothesized that this disconnect might be due to gestation periods—i.e., money spent today takes time to result in visible electric poles or paved roads. To address this, we extended the lag structure to two years (Column 3). Yet, the coefficients remain statistically insignificant. Even allowing for a 2-year construction window, we find no evidence that reported expenditure predicts development outcomes.

Table 2: **Effect of Sector-Specific Expenditure on Luminosity Growth**

Dep. Var: $\Delta \ln(L_{it})$	(1) FD (t)	(2) FD (t − 1)	(3) FD (t − 2)
$\Delta \ln$(Roads Exp)	0.009	0.012	0.011
	(0.015)	(0.018)	(0.016)
$\Delta \ln$(Power Exp)	0.035	0.048	0.045
	(0.038)	(0.041)	(0.039)
$\Delta \ln$(Housing Exp)	0.028	0.039	0.036
	(0.025)	(0.032)	(0.030)
$\Delta \ln$(Health Infra)	0.004	0.006	0.003
	(0.012)	(0.015)	(0.014)
$\Delta \ln$(Community Infra)	0.021	0.032	0.029
	(0.024)	(0.029)	(0.027)
Year Fixed Effects	Yes	Yes	Yes
Observations	23,105	18,484	13,863
Adj. R^2	0.02	0.03	0.02

Standard errors clustered at Block level. *p<0.1, **p<0.05

5.3 Heterogeneity: The Role of Social Fragmentation

We further tested whether this lack of significance drives from heterogeneity in village governance. We interacted the sector-specific expenditure variables with the Ethnolinguistic Fractionalization (ELF) index to test if homogeneous or Schedule 5 (tribal reserved) villages are better at converting funds into lights.

The interaction terms ($\Delta E^s \times ELF$) were consistently insignificant across all sectors. We found no evidence that ethnic homogeneity improves the translation of funds into assets; the statistical noise in the translation of expenditure to lights appears systemic rather than dependent on local caste composition.

5.4 Non-Linear Robustness Checks

To strictly rule out the possibility that our null results are driven by the linearity assumptions of the OLS framework (e.g., if expenditure only impacts luminosity beyond a specific "big push" threshold or through complex interactions between sectors), we trained two non-linear machine learning models: Random Forest and XGBoost. We utilized the exact same feature set as the linear panel model—year-on-year deltas of sector-specific expenditures—without including static PCA variables. Crucially, to account for the panel structure of the data where each village appears across multiple years, we implemented a **GroupKFold** cross-validation strategy. This ensures that all annual observations for a specific Gram Panchayat are kept together in either the training or test set, preventing the model from artificially boosting performance by "memorizing" village-specific trends.

Table 3 compares the predictive performance of these models against the linear baseline. As shown, the complex architectures offer negligible improvement. XGBoost achieved the highest Out-of-Sample R^2 of 0.041, only marginally better than the linear fixed-effects baseline (0.031). The Root Mean Squared Error (RMSE) remains stubbornly high across all specifications, suggesting that the poor predictive power is not a failure of the linear functional form, but a fundamental lack of signal in the expenditure data itself.

Table 3: **Model Performance Comparison (Test Set)**

Model Architecture	RMSE	R^2 (Out-of-Sample)
Linear Fixed Effects	0.442	0.031
Random Forest	0.439	0.038
XGBoost	0.436	0.041

Despite the low overall predictive power, analyzing the feature importance rankings provides insight into which sectors contain the most (albeit weak) signal. For Random Forest, importance was calculated via **Mean Decrease in Impurity**, which averages the reduction in variance across all splits where a feature is used. For XGBoost, we utilized **Gain**, which measures the average improvement in the objective function contributed by a feature.

We observe that both models identify "light-emitting" sectors as the top predictors, though the ranking differs by architecture. The Random Forest model identifies **Housing Expenditure** as the most important feature, whereas XGBoost assigns the highest Gain to **Power Expenditure**. Nonetheless it is importabnt to contextualize this finding: being the "best" predictors in a model with an R^2 of 0.04 implies that even targeted infrastructure spending explains less than 5% of the variation in luminosity growth. The marginal dollar reported in these ledgers is effectively noise relative to the actual development occurring on the ground.

6 Discussion and Conclusion

The central finding of this study is a stark dissonance between the visibility of infrastructure *stock* and the invisibility of fiscal *flow*.

Our cross-sectional validation (Table 1) proves that the VIIRS sensor is highly sensitive to the physical realities of rural Odisha: villages with higher indices of roads, schools, and electrification are measurably brighter ($R^2 = 0.64$). This validates the satellite as a reliable auditor of development *levels*. Yet, our panel analysis (Table 2) and non-linear robustness checks (Table 3) confirm that the reported marginal dollar spent by the Gram Panchayat—even when specifically allocated to "light-emitting" sectors like Roads, Housing, and Power—has statistically zero predictive power for *changes* in that brightness.

6.1 Mechanisms of the Disconnect

This "missing link" implies that the breakdown occurs in the translation of bureaucratic reporting to physical realization. We propose three mechanisms driving this result:

1. **Administrative Fiction:** The *PlanPlus* data represents "Approved" allocations rather than confirmed disbursements. In India's complex fiscal federalism, funds approved at the state level often face severe downstream delays or "parking" at the Block office level. We are likely measuring the *intent* to spend rather than the actual breaking of ground.

2. **Systemic vs. Local Friction:** We hypothesized that "leakage" might be driven by local capture, which theoretically should be lower in caste-homogeneous or reserved (Schedule 5) villages. However, the consistent insignificance of the Fractionalization interaction term suggests that the friction is not local. If funds were reaching the village but being embezzled due to local politics, we would expect variation based on village

sociodemographics. The fact that funds are equally invisible everywhere points to **upstream** systemic inefficiencies (e.g., contractor delays, bureaucratic bottlenecks) that affect all villages uniformly.

3. **The "Soft" Spending Bias:** While we controlled for sectors, it is possible that a significant portion of "Roads" or "Power" expenditure is diverted to maintenance wages or repairs rather than new capital creation. Such "soft" spending would register in the ledger but fail to generate the threshold of luminosity detectable from space.

6.2 Limitations and Future Research

Our study faces specific limitations inherent to administrative data. First, while we utilize sector-specific labels, we cannot verify the granular quality of the assets; a "road" in the ledger may technically be a dirt track that does not warrant streetlighting. Second, the spatial aggregation to the GP polygon level may mask heterogeneous development within the village boundaries.

Future research should pivot from expenditure tracking to **physical asset tracking**. Rather than asking "Did the money appear?", researchers should leverage high-frequency, geo-tagged administrative data (e.g., PMGSY road polygons or geo-stamped PMAY housing photos) to verify asset creation directly. Additionally, combining nightlights with daytime satellite imagery could help distinguish between unlit infrastructure (e.g., unpaved roads) and true "ghost" projects.

Ultimately, this paper highlights the limitations of relying solely on expenditure reports to monitor decentralization. While money is being moved on paper, the satellite record suggests that for the average village in Odisha, these financial flows remain invisible in the night.

Acknowledgements

References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155–194.

Banerjee, A., & Somanathan, R. (2007). The political economy of public goods: Some evidence from India. *Journal of Development Economics*, 82(2), 287–314.

Bhatia, K., & Leighton, M. (2022). How to make decentralization work? Evidence from a capacity-building intervention in local governance. *School of Economics and Finance Discussion Paper No. 2206*. University of St Andrews.

Bruederle, A., & Hodler, R. (2018). Nighttime lights as a proxy for human development at the local level. *PLOS ONE*, 13(9), e0202231.

Centre for Policy Research. (2019). *Analysis of Fund Flow to Rural Local Bodies*. New Delhi.

Crook, R., & Manor, J. (1998). Democratic decentralization. *Operations Evaluation Department (OED) Working Paper Series, No. 11*. Washington, D.C.: World Bank Group.

Dugoua, E., Kennedy, R., & Urpelainen, J. (2018). Satellite data for the social sciences: Measuring rural electrification with nighttime lights. *Introduction to Political Science: A Review of the Field*, 71(2), 1–23.

Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4), 171–198.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2), 994–1028.

Johnson, C. (2003). Decentralisation in India: Poverty, Politics and Panchayati Raj. *Working Paper 199*, Overseas Development Institute.

Kochar, A., Singh, K., & Singh, S. (2009). Targeting public goods to the poor in a segregated economy: An empirical analysis of central mandates in rural India. *Journal of Public Economics*, 93(7–8), 917–930.

Min, B., Gaba, K. M., O'Fallon, O. F., & Scholtes, J. K. (2013). Power to the people: Evidence from a randomized field experiment on community-based monitoring in India. *The Quarterly Journal of Economics*, 129(1), 153–190.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, 81(1), 113–152.