Supplementary Materials

Title: Infrastructure–Environment Complementarity in African Development: Spatial Thresholds and Economic Returns in Tanzania's BRI Corridors

These Supplementary Materials provide extended methodological documentation, robustness diagnostics, and additional empirical results that support and validate the findings presented in the main manuscript. All sections, tables, and figures are explicitly referenced in the main text to ensure transparency, reproducibility, and coherence for reviewers and readers.

S1. Data Construction and Harmonization

S1.1. Night-Time Light (NTL) Data Processing

Annual night-time light data were obtained from NASA's VIIRS Black Marble product (VNP46A2). Data preprocessing was implemented in Google Earth Engine (GEE) to ensure temporal consistency and noise reduction. The workflow included:
(i) removal of cloud contamination, stray light, and lunar illumination using quality assurance flags;
(ii) background noise suppression; and
(iii) inter-annual radiometric calibration to ensure comparability across the 2012–2023 period.

Processed NTL values were aggregated to the hexagonal grid level and used as a proxy for localized economic activity, consistent with established spatial development literature.

Main text linkage: Section 3.3 (Variable Selection and Multi-Source Data Harmonization)

S1.2. Environmental Quality Indicators

Vegetation health was measured using the Normalized Difference Vegetation Index (NDVI) derived from MODIS MOD13Q1 (Version 6). Sixteen-day composites were aggregated to annual means and standardized as Z-scores to facilitate interaction modeling and ecological threshold detection.

Additional environmental controls included surface water availability (MNDWI), rainfall, and a composite climate resilience index to capture broader ecological constraints influencing infrastructure performance.

Main text linkage: Sections 3.3 and 3.5

S1.3. Infrastructure Exposure and Spatial Intensity

Transport infrastructure data were compiled from Belt and Road Initiative project databases, national infrastructure records, and OpenStreetMap (OSM). Infrastructure exposure was operationalized using Gaussian Kernel Density Estimation (KDE) with a 15-km bandwidth, reflecting corridor-level connectivity effects beyond individual project footprints.

Main text linkage: Section 3.3

S1.4. Spatial Unit Construction

The study area was partitioned into 2,680 hexagonal grid cells to minimize the Modifiable Areal Unit Problem (MAUP) and ensure uniform spatial representation across corridors. All variables were harmonized to this spatial framework prior to econometric estimation.

Main text linkage: Section 3.2

S2. Spatial Econometric Diagnostics and Model Selection

S2.1. Spatial Dependence Tests

Formal diagnostics were conducted to assess spatial autocorrelation in economic activity.

Table S1. Spatial Dependence Diagnostics

Diagnostic Test	Statistic	p-value
Moran's I (NTL)	0.1478	<0.001

LM Lag	Significant	<0.01
Robust LM Lag	Significant	<0.01
LM Error	Significant	<0.01
Robust LM Error	Significant	<0.01

Note: Significant Moran's I and Lagrange Multiplier (LM) statistics confirm spatial dependence and justify the use of spatial econometric models.

Main text linkage: Section 3.4

S2.2. Model Selection and Specification

Alternative model specifications (OLS, SAR, SEM) were estimated and compared using Akaike Information Criterion (AIC), log-likelihood, and LM diagnostics.

Table S2. Model Selection Diagnostics (Renumbered from former Table S3)

Diagnostic	Statistic	p-value	Interpretation
Moran's I (Residuals)	0.1478	0.001	Residual dependence present
AIC (SDM)	7439.65	—	Best-performing model
AIC (SAR)	Higher	—	Inferior
AIC (SEM)	Higher	—	Inferior
LM Lag	Significant	<0.01	Supports SDM
LM Error	Significant	<0.01	Supports SDM

Conclusion: The Spatial Durbin Model (SDM) was selected as the preferred specification.

Main text linkage: Section 3.4

S3. Supplementary Econometric Results

Table S3. Full Spatial Durbin Model (SDM) Estimation Results

Variable	Coefficient	Std. Error	z-Statistic	p-Value
Constant	0.5033	0.1702	2.958	0.003
NDVI	−1.3408	0.3469	−3.865	<0.001
MNDWI	0.1843	0.3168	0.582	0.561
Climate Resilience	0.00002	0.00005	0.459	0.646
Population Density	−0.1388	0.0389	−3.566	<0.001
Rainfall	0.0708	0.0016	45.675	<0.001
Infrastructure × NDVI	6.4424	0.0908	70.946	<0.001
Spatial Lag (ϱ)	0.0466	0.0148	3.157	0.002

Model diagnostics:
Observations = 2,680; Pseudo R^2 = 0.8822; Log-likelihood = −3711.83; AIC = 7439.65.

Main text linkage: Section 4.3

Table S4. Impact Decomposition from the SDM

Variable	Direct Effect	Indirect Effect	Total Effect
NDVI	**−1.3408**	**−0.0656**	**−1.4064**
Population Density	**−0.1388**	**−0.0068**	**−0.1456**
Rainfall	**0.0708**	**0.0035**	**0.0743**
Infrastructure × NDVI	**6.4424**	**0.3152**	**6.7576**

Note: Indirect effects capture spatial spillovers.

Main text linkage: Section 4.3

S4. Ecological Threshold Detection and Robustness

S4.1. Bai–Perron Structural Break Test

Table S5. Ecological Threshold Detection Results

Parameter	Estimate
Identified NDVI Breakpoint	−0.8σ
Confidence Interval	[−0.92σ, −0.67σ]
Regime Below Threshold	Insignificant returns
Regime Above Threshold	Strong positive returns
Structural Break Significance	p < 0.001

Main text linkage: Sections 3.5 and 4.4

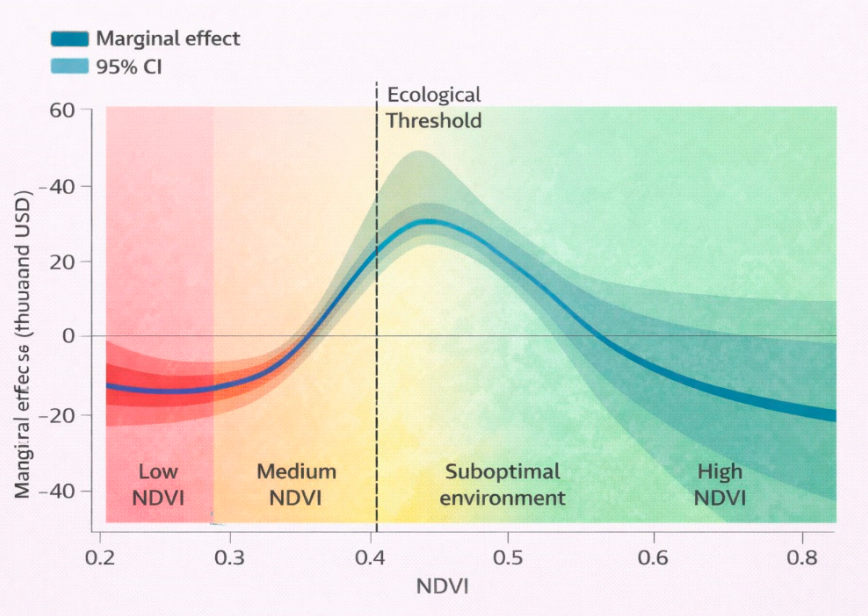


Figure S1. Marginal effects of infrastructure investment across the NDVI gradient.

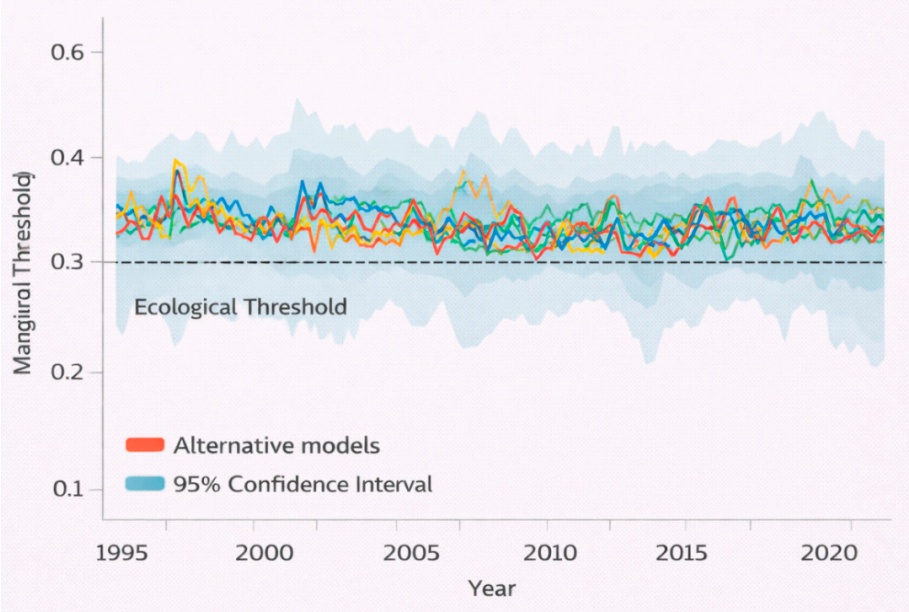


Figure S2. Sensitivity analysis demonstrating the temporal robustness of the identified NDVI threshold across alternative model specifications.

S5. Supplementary Spatial Results

S5.1. Local Indicators of Spatial Association (LISA)

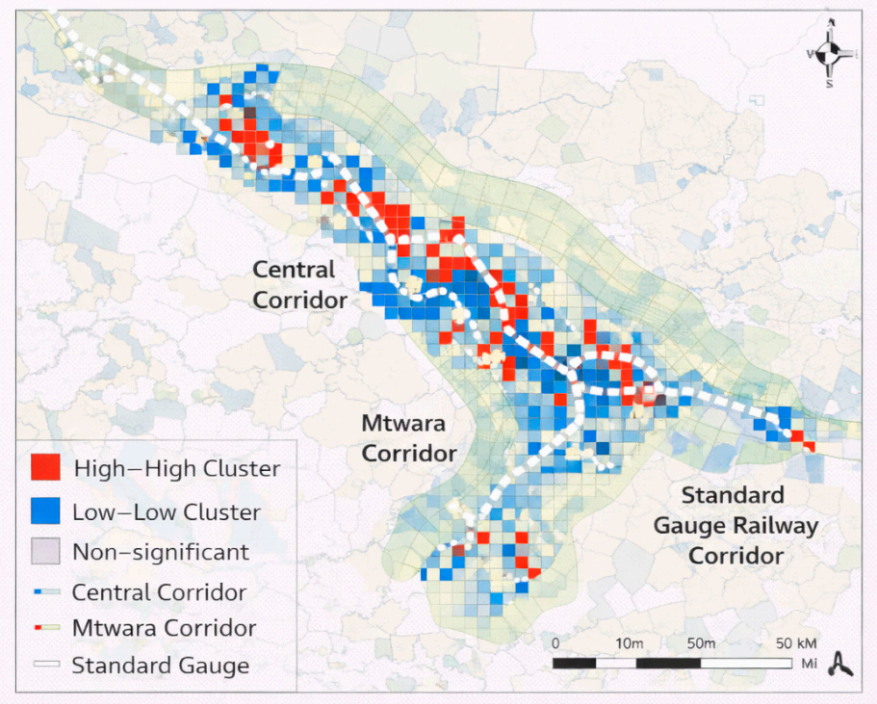


Figure S3. Spatial clustering of indirect (spillover) effects across Tanzania's BRI corridors.

Main text linkage: Section 4.2

S5.2. Residual Diagnostics

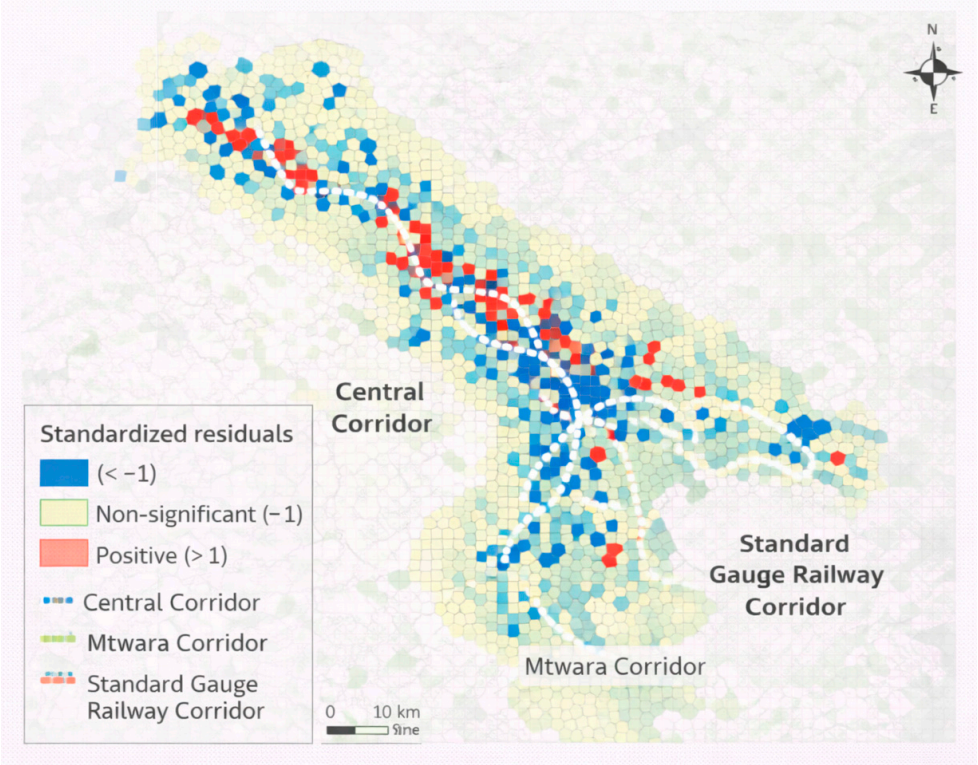


Figure S4. Spatial distribution of model residuals confirming the absence of systematic spatial bias after accounting for spatial dependence.

Main text linkage: Section 4.2

S6. Corridor Investment Classification

Table S6. Corridor Investment Zone Classification

Zone	Area Share	Representative Districts	Key Characteristics
High-Priority	63%	Dar es Salaam, Morogoro, Dodoma	High NDVI, strong spillovers
Medium-Priority	22%	Tabora, Singida	Moderate NDVI
Restoration-First	15%	Mtwara, Lindi	NDVI < −0.8σ

Main text linkage: Section 4.5

Table S7. Alternative Model Specifications

Model	Key Finding
OLS	Underestimates spatial spillover effects
SAR	Omits interaction effects
SEM	Captures error dependence only
SDM	Captures full infrastructure–environment complementarity

Conclusion: Only the Spatial Durbin Model (SDM) consistently identifies infrastructure–environment complementarity.

S7. Descriptive Statistics and Variable Definitions

Table S8. Variable Definitions and Sources

Variable	Description	Source	Spatial Resolution
NTL	Proxy for localized economic activity	VIIRS Black Marble	500 m
NDVI	Vegetation health indicator	MODIS MOD13Q1	1 km
MNDWI	Surface water availability	Sentinel-2	10 m
Rainfall	Annual precipitation	WorldClim v2.1	1 km
Population Density	Human population concentration	WorldPop	100 m
Infrastructure Density	Transport infrastructure exposure (KDE-based)	OSM / BRI sources	Corridor-scale

Note: All variables were harmonized to a common hexagonal grid framework to ensure consistency in spatial econometric estimation and ecological threshold analysis.

S8. Computational Implementation and Reproducibility

All data processing and econometric analyses were conducted using open-source software, including Python (PySAL, spreg, GeoPandas, Rasterio) and R (splm). The complete computational workflow and replication materials are archived in a public GitHub repository maintained by the authors. A persistent DOI will be assigned via Zenodo upon publication to ensure long-term accessibility and reproducibility.

S8.1. Replication Guide

A step-by-step replication guide (README.md) accompanies the code repository, detailing software installation, preprocessing order, model estimation steps, and procedures for reproducing all tables and figures.