Infrastructure inequality is a characteristic of urbanization

Bhartendu Pandeyab,c, Christa Brelsfordc, and Karen C. Setob

Edited by Arun Agrawal, University of Michigan, Ann Arbor, MI; received October 31, 2021; accepted February 25, 2022

Urbanization can challenge sustainable development if it produces unequal outcomes. Infrastructure is an important urbanization dimension, providing services to support diverse urban activities. However, it can lock in unequal outcomes due to its durable nature. This paper studies inequalities in infrastructure distributions to derive insights into the structure and characteristics of unequal outcomes associated with urbanization. We analyzed infrastructure inequalities in two emerging economies in the Global South: India and South Africa. We developed and applied an inequality measure to understand the structure of inequality in infrastructure provisioning (based on census data) and infrastructure availability (based on satellite nighttime lights [NTLs] data). Consistent with differences in economic inequality, results show greater inequalities in South Africa than in India and greater urban inequalities than rural inequalities. Nevertheless, inequalities in urban infrastructure provisioning and infrastructure availability increase from finer to coarser spatial scales. NTL-based inequality measurements additionally show that inequalities are more concentrated at coarse spatial scales in India than in South Africa. Finally, results show that urban inequalities in infrastructure provisioning covary with urbanization levels conceptualized as a multidimensional phenomenon, including demographic, economic, and infrastructural dimensions. Similarly, inequalities in urban infrastructure availability increase monotonically with infrastructure development levels and urban population size. Together, these findings underscore infrastructure inequalities as a feature of urbanization and suggest that understanding urban inequalities requires applying an inequality lens to urbanization.

Significance

We find that urban infrastructure inequalities are a characteristic feature of urbanization, with multiscale patterns that highlight greater inequalities at regional than at local intracity scales. Results show that the level of urban infrastructure inequality has a significant bearing on urbanization characteristics. Owing to infrastructure inequalities, urbanization as a pathway to sustainable development faces fundamental constraints. Our study raises intriguing questions about the urban infrastructure inequality implications of future urbanization and the spatial scales at which we ought to understand and address these inequalities.
urbanization. We define infrastructure inequalities as the degree of variation in physical infrastructure (availability, provisioning, or access) conditional on average infrastructure levels across geographical space. A focus on infrastructure inequalities can significantly advance our understanding of inequalities as an emergent property of urbanization and clarify 1) urbanization’s inequality implications; 2) persistence and predictability of inequality; 3) inequality-minimizing urbanization strategies; and 4) unintended multiscale consequences of reducing inequalities within urban areas.

In this study, we applied an inequality lens to infrastructure, conceptualizing infrastructure as an urbanization dimension, to understand the spatial structure and characteristics of attendant unequal outcomes and, more broadly, urban inequalities. To this end, we propose the following axioms (Fig. 1A):

1. Urbanization is a multidimensional process, with interrelated economic, environmental, institutional, socio-cultural, and technological changes, in addition to demographic change; 2. Multiple dimensions of urbanization manifest in geographical space; and 3. Urban inequalities are unequal states of the multiple dimensions of urbanization observed across spatial scales.

We examined India and South Africa, the two emerging economies in the Global South. Our goal was to understand the similarities and differences in the structure and characteristics of infrastructure inequalities vis-à-vis urbanization levels in the two countries. We examined two specific questions: 1) What are the characteristics of inequality in infrastructure provisioning and availability, and 2) How do inequalities in infrastructure provisioning and availability associate with urbanization? Using 2011 census data for over 700,000 spatial units, we examined a composite index of infrastructure provisioning: water, sanitation, housing, and electricity (SI Appendix, Table S1). We selected these four infrastructure dimensions based on their importance for sustainable development as

![Fig. 1.](https://doi.org/10.1073/pnas.2119890119)
recognized in the UN-SDGs 6, 7, and 11. Given the particularities of census data between the two countries and problems associated with comparing administrative scales between two countries, we also examined inequalities in infrastructure availability using satellite remote sensing data (i.e., Visible Infrared Imaging Radiometer Suite [VIIRS] nighttime lights [NTLs]; Materials and Methods).

Infrastructure Inequality Measurement

Of the existing economic inequality measures (28, 29), studies have chiefly applied the Gini index or entropy-based measures to study infrastructure inequalities (30, 31). To adapt these measures to infrastructure distributions, we also must adapt the underlying assumptions and axioms. However, the properties of infrastructure distributions can differ from economic distributions in at least two ways. First, physical infrastructure growth faces resource constraints, including land, materials, energy, labor, and investments, resulting in constrained growth that often corresponds to an S-shaped logistic curve. Second, the statistical distributions of infrastructure can differ from that of economic variables (income or wealth). We expect infrastructure distributions to follow a binomial or beta distribution instead of exponential, log-normal, or power-law distributions in income or wealth distributions. Against this backdrop, infrastructure inequality measures that use statistical properties of infrastructure distributions may be more appropriate.

Consider a geographic region composed of a finite set of units (individuals, households, or locations), each with the presence or absence of infrastructure. Heterogeneity in infrastructure distribution within this region is dependent on average infrastructure, as specified by a Bernoulli’s process. In this study, we are interested in between-regions heterogeneity and how it varies with overall average infrastructure. Assuming a growth scenario where some regions have the maximum possible infrastructure and others have no infrastructure, between-region heterogeneity is maximized at every average infrastructure level, following properties of Bernoulli distribution. In contrast, in a strictly ordered growth scenario, between-region homogeneity persists when average infrastructure increases uniformly across all regions. Real-world growth, however, is likely to situate between the maximally heterogeneous and maximally ordered growth typologies, which serve as high and low reference points, respectively. In this context, heterogeneity-conditional-on-mean determines preferential growth levels (SI Appendix). As a result, we can define a measure of infrastructure inequality (I) that varies from 0 (lowest inequality) and 1 (highest inequality) using Eq. 1.

$$I = \frac{\sigma}{\sqrt{\mu (1 - \mu)}}, 0 < \mu < 1. \quad [1]$$

This inequality index can measure the level of spatial inequality in infrastructure availability or provisioning (Fig. 1B and SI Appendix, Fig. S1). It is conceptually consistent with the heterogeneity index proposed in Brelsford et al. (32). Still, the index offers a simple advantage over the heterogeneity index, as the latter requires a distribution of $\sigma$ and $\sqrt{\mu (1 - \mu)}$. We applied this measure to census and VIIRS NTL data and examined urban, rural, and overall inequality levels derived from census data and inequality across the rural-urban spectrum and selected urban areas from VIIRS NTL (Materials and Methods).

Results

Spatial Structure of Inequalities in Infrastructure Provisioning.

Results show that urbanization in India and South Africa is characterized by greater urban inequalities than rural inequalities, in addition to emergent urban-rural inequalities. This suggests a significant intertwining between urbanization and spatial inequality that goes beyond inequalities within urban areas. Infrastructure provisioning has higher inequality between urban spatial units ($I_u$) than between rural units ($I_r$) (Fig. 2A and SI Appendix, Fig. S2), contrasting expectations as provisioning tends to be more viable in urban than rural areas due to concentrated geographies. Instead, we found that $I_u$ exceeded $I_r$ by -4% and -42% in South Africa and India, respectively (Fig. 2A). Overall inequality ($I_o$) levels are comparable between South Africa and India. Still, based on the average infrastructure provisioning index, urban areas are 569.4% (51.76%) better off than rural areas in India (South Africa) (Fig. 2A). Results also showed that overall infrastructure inequality levels ($I_o$) are 4.37% greater in South Africa than in India. Putting this difference in the context of difference in income inequality, we find that infrastructure inequality is more comparable than income inequality between the two countries. A similar difference in income inequality (measured using the Gini index) is -27%. More pronounced infrastructure inequality than (estimated) income inequality in India can explain the difference in income inequality. In sum, a few winners and several losers to accumulating infrastructure gains characterize the examined state of urbanization in India and South Africa.

We found significant heterogeneity in urban inequality levels at each subnational scale but with greater average inequality at the national ($I_o$) and regional scales ($I_o$ and $I_u$) than at the local scales ($I_o$) (Fig. 2B and C). While much of the current focus on urban inequality is at the local scale, results show spatial scale dependence such that infrastructure inequalities decrease from national and regional scales to local scales. Still, we found that the average inequality levels are lower in India than in South Africa at each comparable administrative scale (Fig. 2C).

South African cities are significantly more unequal than Indian cities. Results showed greater inequality ($I_u$) within South African cities than within Indian cities (Fig. 2D). More significant spatial fragmentation and unevenness in South African cities than in Indian cities are also evident from a visual inspection of infrastructure-provisioning levels within cities (SI Appendix, Figs. S3 and S4). Results showed no significant association between urban inequality in infrastructure provisioning and population size (Fig. 2D). Instead, a core-periphery structure exists in some Indian cities with relatively higher inequality levels (i.e., infrastructure-provisioning levels are higher in the city center and lower in the peripheries) (SI Appendix, Figs. S3 and S4).

Spatial Structure of Inequalities in Infrastructure Availability.

Results obtained for infrastructure availability from VIIRS NTL show greater inequality in South Africa than in India at the national and urban scales (Fig. 3A and B and SI Appendix, Fig. S5). Furthermore, we found that the multiscale characteristics of inequality from satellite data are consistent with that from census data. Inequality levels decrease from coarser to finer spatial scales, defined using lattice grids with varying resolutions (SI Appendix, Fig. S6 and Materials and Methods). Still, inequality decreases faster from coarser to finer spatial resolutions in India than in South Africa (Fig. 3C). The rate of change in inequality levels with spatial scale suggests that India’s inequality is more concentrated at coarser scales than finer scales. In contrast, inequality in South Africa is relatively
more pronounced across spatial scales. These results emphasize differences in India and South Africa’s inequality-generating processes and suggest that the appropriate scale to address inequality in India may be different from South Africa. In addition to the spatial scale dependence, NTL inequality shows an upper bound dependence such that inequality levels decrease from lower values of NTL upper bounds to higher values (SI Appendix, Fig. S6). Nonetheless, our results above are also robust to varied upper bounds.

Despite the consistencies with census-based inequality estimates, NTL inequality measurement showed that brightly lit areas have higher inequality than dimly lit areas (Fig. 4). As expected, we found that inequality measured from NTLs across Indian and South African cities positively correlates with population size (SI Appendix, Fig. S7). The correlation is lower in South Africa than in India. This suggests that inequality in South African cities is higher overall and less associated with city size than inequality in Indian cities. In addition, we found that (log) inequality in NTLs is positively but nonlinearly associated with (log) average NTLs (SI Appendix, Figs. S8 and S9), irrespective of spatial scales and upper bound radiance. We found no evidence of an inverted-U relationship between urban inequality and average development levels.

Multidimensional Urbanization and Inequalities in Urban Infrastructure Provisioning. Inequality measurements from census and satellite data evidenced significant urban inequality. Here, we tested whether and how inequality is associated with multidimensional urbanization (Materials and Methods). Of the four principal components (PCs) constructed from key urbanization variables, we selected three based on variance explained (>90%). The interpretation of the components, however, differed between India and South Africa, which can be attributed to the difference in urbanization characteristics between the two countries.

In India, the first three PCs explain 92.68% variation in the four variables (SI Appendix, Table S2). PC 1 captures multidimensional urbanization with positively loaded average urban infrastructure provisioning, urban demographic share, and average NTL radiance. Urban average wealth is positively loaded in PC 2, whereas PC 3 scales positively with average infrastructure provisioning and negatively with urban demographic share. In South Africa, the three selected PCs explain 93.50% of the total variation (SI Appendix, Table S3). PC 1 in South Africa also captures multidimensional urbanization but is positively loaded with average NTL radiance and average urban income. PC 2 captures average urban infrastructure provisioning with negative loading, and PC 3 scales positively with urban demographic share.

Due to the presence of spatial autocorrelation in urban inequality levels across districts (municipalities) in India (South Africa) (SI Appendix, Fig. S10), we turned to spatial regression models. We found that spatial lag models are generally more appropriate than spatial error models based on log-likelihood and Akaike information criterion (AIC) estimates (SI Appendix, Table S4). These models show that urban inequality levels are higher in Indian districts with higher multidimensional urbanization and average wealth (Table 1). In contrast, districts with high average infrastructure provisioning and low urban demographic shares have low urban inequality levels. Surprisingly,
we found an insignificant correlation between urban inequality and urban demographic shares (SI Appendix, Table S4). In South Africa, average urban infrastructure provisioning is significantly associated with urban inequality, and municipalities with higher urban inequality have lower average urban infrastructure provisioning (P value < 0.1). Surprisingly, we found insignificant associations between urban inequality and wealth (income) inequality in India (South Africa).

Furthermore, we found that religious and cultural diversity in India and racial diversity in South Africa do not explain spatial variations in urban inequality levels. Our results are consistent even after controlling for physical geography with topographical variables (Table 1). We found a small but statistically significant association between (log) average elevation and urban inequality levels in India and South Africa (P value < 0.01). These findings (re)emphasize a significant association between inequality and urbanization. However, results suggest that we need a multidimensional conceptualization of urbanization to better understand the urbanization-inequality intertwining.

**Discussion**

Results have five broad implications for our understanding of infrastructure inequalities in the context of urbanization and sustainable development. First, we can better understand and address equitable urbanization by bringing infrastructure inequality to the fore. Infrastructure inequality has been given less attention than economic inequality. Consequently, inequalities in infrastructure provisioning, access,
and quality need greater attention to advance urban sustainability and can be a key lever to advancing distributive equity.

Second, inherent infrastructure inequalities emphasize multiscale distributive equity as a central imperative for urbanizing countries to progress toward sustainable development. As regions urbanize and infrastructure availability and provisioning expand exclusively within urban areas, we expect a widening urban-rural gap (33). We also expect emergent inequalities in urban infrastructure provisioning when urban infrastructure growth lags with respect to increases in urban population, as our results show. As urban infrastructural amenities proliferate beyond urban areas, rural areas are also subject to inequalities and face a dual sustainable development challenge of infrastructure deficits and inequalities. Across the rural-urban continuum, we observed greater inequalities in infrastructure availability in more urbanized regions than in less urbanized regions. In sum, infrastructure inequalities are inevitable and omnipresent, requiring understanding them at scale. To this end, new multiscale indices and temporal analyses are needed, which can build upon the measurement framework introduced in this study to examine infrastructure inequalities with spatially and temporally consistent VIIRS NTL.

Third, focusing on average infrastructure levels, individual urban areas, or one spatial scale at a time are insufficient toward equitable urbanization. Developmental progress asserted with increasing average infrastructure levels for some communities can hide associated increases in spatial inequalities. This is particularly relevant for India considering its urbanization characteristics, where urban regions with greater average provisioning levels have greater inequalities than those with lower average provisioning levels. Consequently, assessing progress with concomitant changes in inequality is important. Furthermore, emergent patterns of urban inequalities vis-à-vis average infrastructure provisioning levels can be context dependent. For instance, even though South African cities have greater inequalities than Indian cities, there is significant heterogeneity in inequality levels between cities of each country.

Fourth, inequalities in infrastructure provisioning and availability tend to be spatial scale dependent because our inequality index is independent of population size, with greater inequality at the national and regional scales than at the local scale. Some of this greater coarser-scale inequality may have resulted from the distributional impacts of past national and subnational investments and policies that intentionally or unintentionally

### Table 1. Spatial lag (models 1 and 3) and error (models 2 and 4) model estimates for urban inequality in India (district-level; models 1 and 2) and South Africa (municipality-level; models 3 and 4)

|                      | Model 1 | Model 2 | Model 3 | Model 4 |
|----------------------|---------|---------|---------|---------|
| (Intercept)          | 0.120***| 0.265***| 0.019   | 0.015   |
|                      | (0.029) | (0.036) | (0.175) | (0.187) |
| (PC1) Urbanization\_multidimensional | 0.019***| 0.023***| 0.015   | 0.105***|
|                      | (0.003) | (0.004) | (0.015) | (0.015) |
| (PC2) Average wealth\_DHS | 0.027***| 0.028***| 0.019   | 0.115***|
|                      | (0.006) | (0.007) | (0.019) | (0.020) |
| (PC3) High infrastructure access\_Low urbanization | -0.033***| -0.043***| -0.101  | -0.095  |
|                      | (0.007) | (0.008) | (0.019) | (0.020) |
| Wealth/Income inequality (Coefficient of Variation)) | 0.146   | 0.134   | 0.018   | 0.026   |
|                      | (0.077) | (0.094) | (0.032) | (0.033) |
| Religious and cultural diversity | -0.003  | 0.003   | 0.110   | 0.033   |
|                      | (0.008) | (0.011) | (0.019) | (0.020) |
| Racial diversity     | 0.063   | 0.068   | 0.018   | 0.026   |
|                      | (0.061) | (0.065) | (0.064) | (0.065) |
| Log elevation\_μ   | 0.009** | 0.013** | 0.062** | 0.072***|
|                      | (0.003) | (0.004) | (0.020) | (0.022) |
| Log slope,\_μ     | 0.008   | 0.010   | 0.138   | 0.170   |
|                      | (0.002) | (0.002) | (0.005) | (0.006) |
| ρ                    | 0.464***| 0.517***| 0.138   | 0.170   |
|                      | (0.042) | (0.045) | (0.009) | (0.010) |
| λ                    | 0.517***| 0.138   | 0.170   | 0.170   |
|                      | (0.045) | (0.009) | (0.009) | (0.009) |
| Number of Observations | 623     | 623     | 199     | 199     |
| Akaike information criterion (AIC) | -1349.997 | -1333.993 | -85.753 | -86.134 |
| Bayesian information criterion (BIC) | -1310.086 | -1294.082 | -52.820 | -53.201 |
| Deviance             | 3.872   | 3.922   | 6.820   | 6.792   |
| Log likelihood       | 683.998 | 675.996 | 52.877  | 53.201  |
| Pseudo-R^2           | 0.48    | 0.47    | 0.32    | 0.32    |

Principal components (PCs) were generated from urbanization rate, average wealth/income, average infrastructure level, and (log) average NTL radiance variables. ***P < 0.001; **P < 0.01.

†Average elevation and average slope were significantly correlated in India (Pearson’s correlation coefficient = 0.69; P value < 0.01). Consequently, we used only (log) average elevation to control for topographic effects. This correlation was significant but weak in South Africa (Pearson’s correlation coefficient = 0.30; P value < 0.01).
favored neighborhoods in some urban areas over others (34). Similarly, socio-cultural preferences can also generate the observed multiscale patterns (30). Importantly, the multiscale variations place urban inequality as a regional problem and challenge existing efforts to reduce urban inequalities by focusing on a few cities.

Finally, owing to the association between infrastructure and economic activity, urbanization (as a process of infrastructure growth) can inherently make some urban areas more economically productive than others (35). The inherent inequality can further pose fundamental constraints to urbanization as a pathway for sustainable development—we cannot have urbanization without inequality and vice versa. Nevertheless, the limited scope of the current study warrants further investigation into 1) whether inequality in infrastructure availability increases with urbanization and development levels across countries and regions and over time, 2) under what urbanizations conditions do high levels of inequality in infrastructure provisioning emerge, and 3) how do urban infrastructure inequalities become a conduit to human well-being and planetary impacts. Overall, an inequality lens applied to urbanization is necessary to answer these questions.

Materials and Methods

Data.

Census data. We used South African 2011 census data (available from https://www.statssa.gov.za) with the finest administrative scale, subplaces, as the base geographical unit with 21,298 units. Our analysis relied on aggregated data at six administrative scales: national, province, district, municipality, main-place, and subplace. Additionally, we focused on inequality in eight metropolitan municipalities. As such, the census data do not classify subplaces into urban and rural. We applied an alternative urban-rural classification based on population (≥1,000) and population density (≥1,000/km²) thresholds (36). In the absence of such a multiscale geospatial census product in India, we created a spatial database of household access to infrastructure amenities (SI Appendix, Table S1). We combined data available online from the Indian census website at the urban ward, slum, and village levels (https://censusindia.gov.in/) and a spatial dataset at three levels: state, district, and subdistrict levels. In total, we examined 681,779 urban wards, towns, slums, and villages. Note that urban wards, towns, and slums constitute urban areas, and villages constitute rural areas. Urban ward, slum, and village boundaries across India are either not readily available or lack sufficient accuracy. We restricted spatializing of the data and subsequent analysis at the national, state, district, and subdistrict levels. We also spatialized the census data for 23 select cities in India (with population size ranging from ~57,000 to ~12 million)—for which digital ward-level boundaries are available—to compare Indian cities with the eight major metropolitan municipalities in South Africa. Given the differences in the census data variables between India and South Africa, our selection of India’s infrastructure variables differed from that of South Africa’s. SI Appendix, Table S1 provides a list of all infrastructure variables used in this study. We considered infrastructure alternatives that were likely to maximize subjective well-being beyond the presence or absence of provisioning while selecting the individual variables (32). Ideally, we expected provisioning typologies in urban areas that maximize efficiency and subjective well-being, subject to budget constraints. Our analysis of infrastructure inequality indirectly evaluated this expectation.

Satellite data. Differences in how infrastructure provisioning is measured between the two countries can confound the analysis and interpretations drawn. Similarly, comparing administrative scales, such as Indian states with South African provinces, can also bias the analysis, given differences in population size and area coverage. As a robustness check and complementing our analysis of infrastructure distributions from the census data, we also examined inequality measurement using the VIIRS 2012 to 2019 time-series NTL dataset. Despite our use of the 8-y time-series data, this study only reports findings from a cross-sectional analysis of India’s and South Africa’s infrastructure distributions for the year 2012 (SI Appendix). Available nighttime remote sensing products remove the effects to natural sources to yield radiance from anthropogenic sources (37). In principle, anthropogenic NTL sources can be classified as stable (e.g., buildings and streetlights) and ephemeral (e.g., vehicular traffic, gas flaring, and fires). Our VIIRS NTL time series processing removed the effects of ephemeral events and isolated stable annual average NTL emissions by 1) applying an outlier-removal algorithm, 2) fitting a seasonality-trend-residual model to each pixel time series, and 3) calculating an annual average NTL radiance (SI Appendix). Consequently, the stable annual average NTL radiance is expected to retrieve outside energy infrastructure from street lights and exterior building lights, which we interpreted as infrastructure availability.

Methods. We analyzed the data in R (https://www.r-project.org/) using ggplot2, sf, rgdal, Hmisc, spdep, spatialreg, raster, tmap, and dplyr packages and in python (https://www.python.org/) programming languages using numpy, scpy, pandas, geopandas, osgeo, scikit-image, matplolib, and rasterio packages.

Measuring and analyzing basic infrastructure inequality. We created an infrastructure provisioning index (IP) using a geometric mean of water, sanitation (and drainage in India), housing, and electricity provisioning rates (I) for all the fine-scale spatial units (j), functionally consistent with the human development index using Eq. 2.

$$IP_j = \prod_{j=1}^n \text{Provisioning Rate.}$$

By construction, IP ranges from 0 (lowest development) to 1 (highest development). Next, we applied the inequality measurement framework using the average infrastructure provisioning index (IPs) levels and heterogeneity (IPh) therein (SI Appendix, Figs. S11 and S12). Here, we calculated population-weighted moments of the distribution to calculate inequality, using the Hmisc package (38).

We described inequality (Iα) characteristics across spatial scales (s) and sectors (c): urban (u), rural (r), and overall (o). We analyzed national-scale (s = 1) variations in Ius, Isr, and Io. In doing so, we compared differences in Isr levels in the two countries with differences in income inequality levels. Next, we examined urban inequality at subnational scales [i.e., state/province (s = 2), district (s = 3), and subdivsions/municipality (s = 4)]. Additionally, we examined inequality levels (Iαs) for 23 select cities in India and eight major metropolitan municipalities in South Africa (s = 5).

Measuring and analyzing infrastructure inequality from satellite data. Inequality measurement using VIIRS NTL is more involved than using census data. NTL is a reliable proxy for human and economic activity (39-41) and has been used to measure human development (42). Still, there is a lack of consensus on its interpretation. Intuitively, ground sources of artificial NTL include street lights, building lights, and vehicle lights (43). We propose that these sources form a part of the total infrastructure stock, and as a corollary, NTL can be considered a proxy for urban infrastructure.

Here, we examined whether inequality (NTLinequality)–heterogeneity conditional on mean–in infrastructure distributions measured from VIIRS NTL is higher in South Africa than in India at the national and urban scales. In doing so, we measured inequality as a function of spatial scale and NTL thresholds. One key advantage of satellite data over census data is their spatial consistency. Unlike administrative scales, inequality can be estimated and compared across contexts with precise spatial scales. Our experiment analyzed NTL using administrative divisions and lattice grids ranging in resolutions from 0.05° to 1°. With lattice grids, we used LandScan population distribution data for the year 2012 (44). We analyzed NTL imposing a maximum threshold in nanowatt(s) per steradian per square centimeter (nWsr⁻¹ cm⁻²) before aggregating the dataset at a given scale. This constraint assumes that the developmental gains decouple beyond a specific threshold, drawing upon the diminishing marginal utility principle. The pixel-level threshold also limits bias from high NTL emitting sites such as industries. Nevertheless, we evaluated our results at multiple such thresholds. Without an upper bound, we found that NTL exhibited an empirical regularity compatible with Taylor’s power law (45) (SI Appendix, Fig. S13). Using the threshold radiance, we applied our measurement framework to calculate NTL-based inequality at the subdistrict (municipality) scales and took a population-weighted average of inequality levels to estimate inequality at the national scale in India (South Africa).
In addition to analyzing the complete rural-urban spectrum in South Africa and India, we also analyzed urban areas with a population size greater than 300,000 in 2018 in the two countries (46). Given a lack of official urban boundaries for these urban areas, we derived urban boundaries using a random walk segmentation algorithm applied to VIIRS NTL (47) in python programming language. The algorithm starts with a point location and clips VIIRS NTL using a circular neighborhood with a radius of 300 pixels (~150 km). We then applied Otsu’s thresholding to the clipped NTL and sampled 10% of the urban pixels and nonurban pixels (48). We used these select pixels as marker inputs to the segmentation algorithm and generated a binary urban and nonurban output. We repeated this procedure and generated 500 binary outputs, based on which we identified urban boundaries using pixels that were labeled urban with over 95% probability. In cases where boundaries of two or more urban areas overlapped, we considered these urban areas forming a part of a larger urban agglomeration. Using these urban boundaries, we calculated inequality for 194 cities in India and South Africa as a population-weighted average of inequality levels.

**Regression analysis.** We investigated how multidimensional urbanization explains urban inequality measured from the census data at a given spatial scale. Specifically, we compared regression models with urban demographic share and models with multiple urbanization dimensions. Here, we focused on urban inequality at the district-level (India) and municipality-level (South Africa) based on data availability and comparability. We began with a principal component analysis (PCA) of four aspects of urbanization: 1) urban demographic share, 2) average urban infrastructure provisioning, 3) average urban wealth/income, and 4) average NTL radiance. These aspects capture urbanization’s demographic, economic, and environmental dimensions. PCA is essential as the select aspects are likely to be correlated. Next, we selected (based on cumulative variation explained >90%) and interpreted the PCs as orthogonal, urbanization dimensions. We estimated ordinary least squares (OLS) regression fit using PCs as explanatory variables. We assessed whether spatial lag and spatial error models are more appropriate than OLS models and selected and interpreted accordingly, using spatialreg package in R (SI-Appendix, Table S4) (49, 50).

In our regression, we also controlled for topography using an elevation data set (https://srtm.csi.cgiar.org/srtmdata). We used data on India’s asset-based household wealth and associated inequality from the 2015-2016 demographic and health survey (DHS).

**Data Availability.** Processed census and VIIRS NTL data, along with the R and python code used for data processing and analysis, has been deposited and are available at the public server Figshare (https://figshare.com/articles/dataset/Supplementary_Data_and_Code_Infrastructure_Inequalities_in_India_and_South_Africa/19125809/1) (51).

**ACKNOWLEDGMENTS.** This study was supported by the Yale School of the Environment, the Yale Institute for Biospheric Studies, and the Hixon Center for Urban Ecology. A portion of the work was sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC for the US Department of Energy under contract DE-AC05-00OR22725.