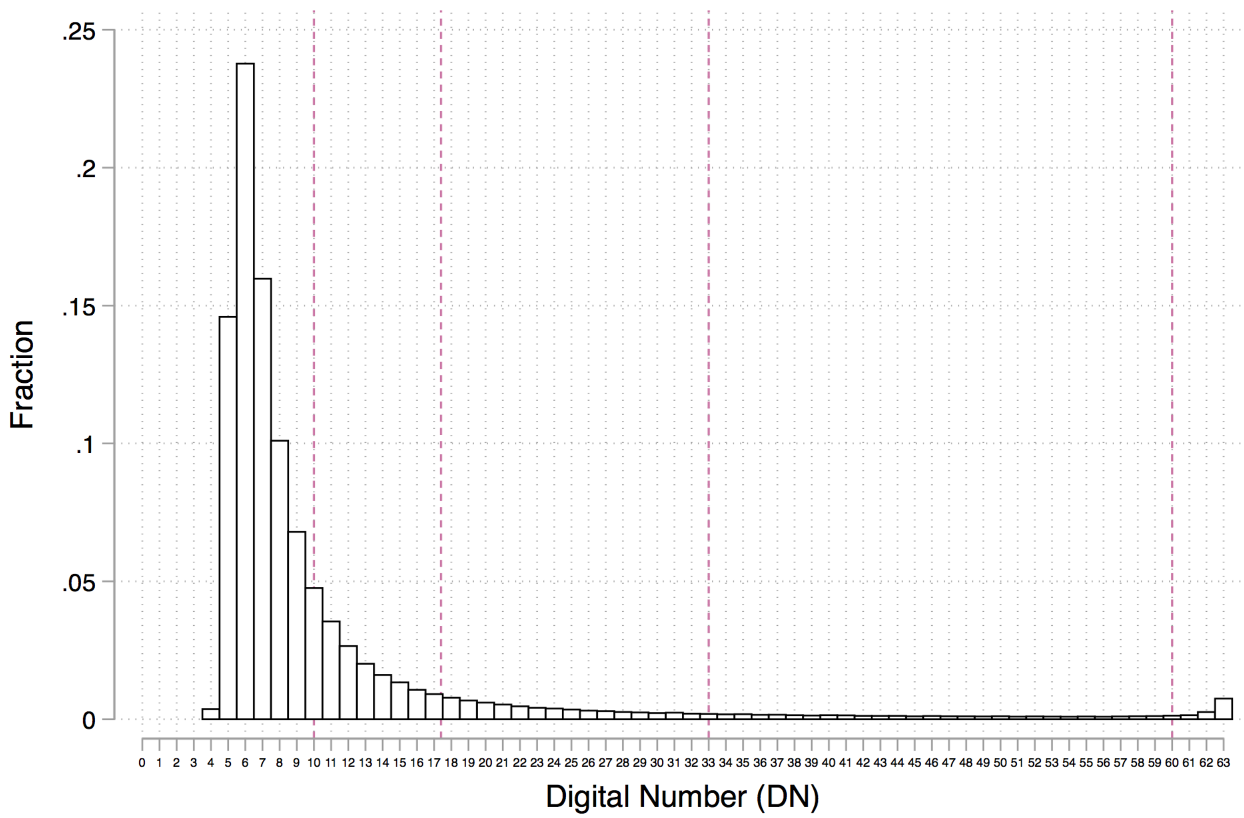
Online Appendix Tables and Figures

Table A1: Administrative Areas in India, 2011 Census

	Number	Total Population	Population Share	Mean Population	Mean Area (km^2)
Villages	640,932	833,748,852	68.9%	1,301	4.8
Towns	6,171	377,106,125	31.1%	61,109	16.6
Class 1 (>100k)	468	264,745,519	21.9%	565,696	97.6
Class 2 (50k-100k)	474	32,179,677	2.7%	67,890	20.4
Class 3 (20k-50k)	1,373	41,833,295	3.5%	30,469	14.4
Class 4 (10k-20k)	1,683	24,012,860	2.0%	14,268	9.3
Class 5 (5k-10k)	1,749	12,656,749	1.0%	7,237	5.5
Class 6 (<5k)	424	1,678,025	0.1%	3,958	4.1

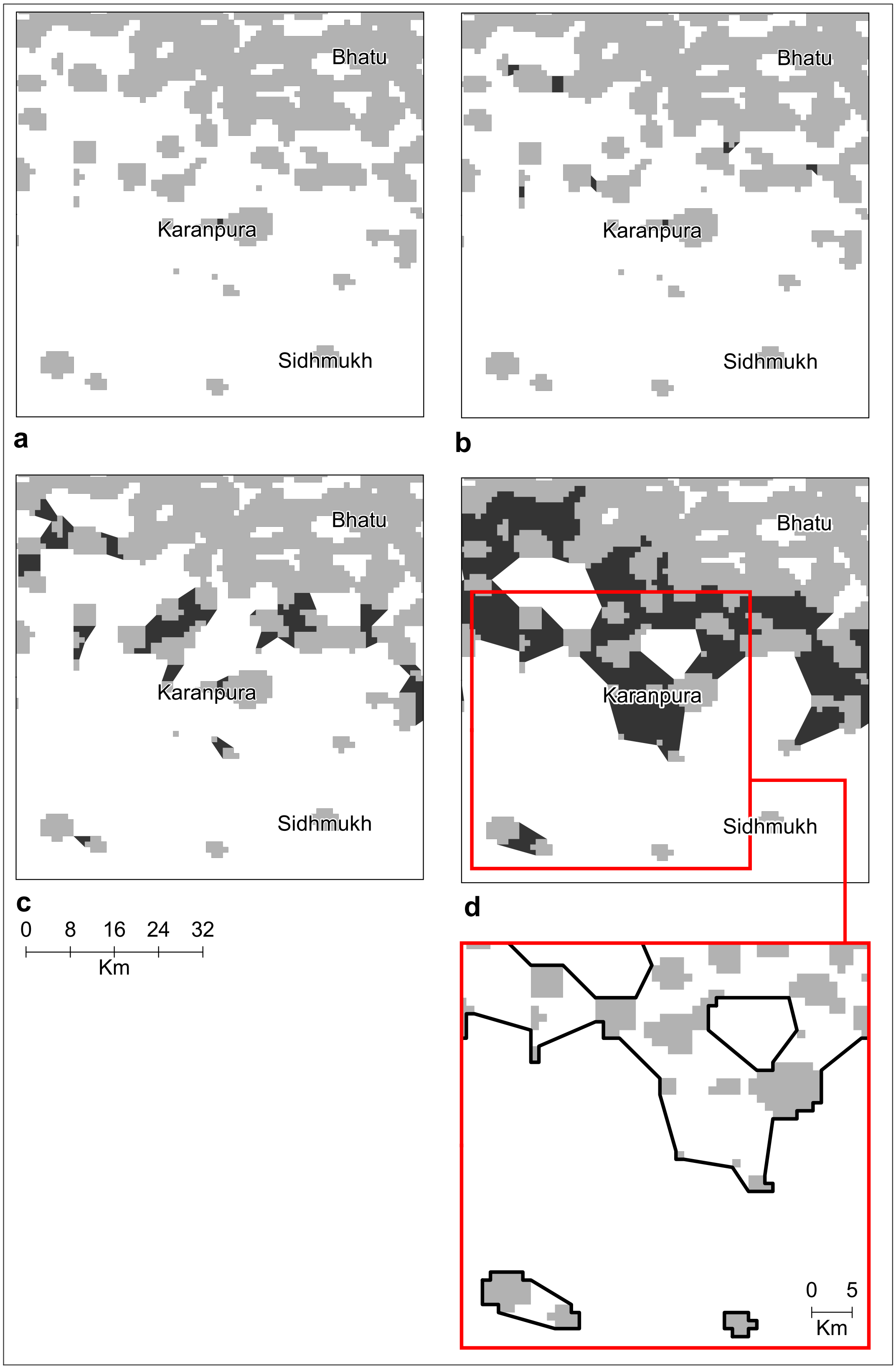
Notes: Table reports official tabulations from 2011 Census of India.

Figure A1: Density of Nighttime Lights for $1km$ Pixels, All India



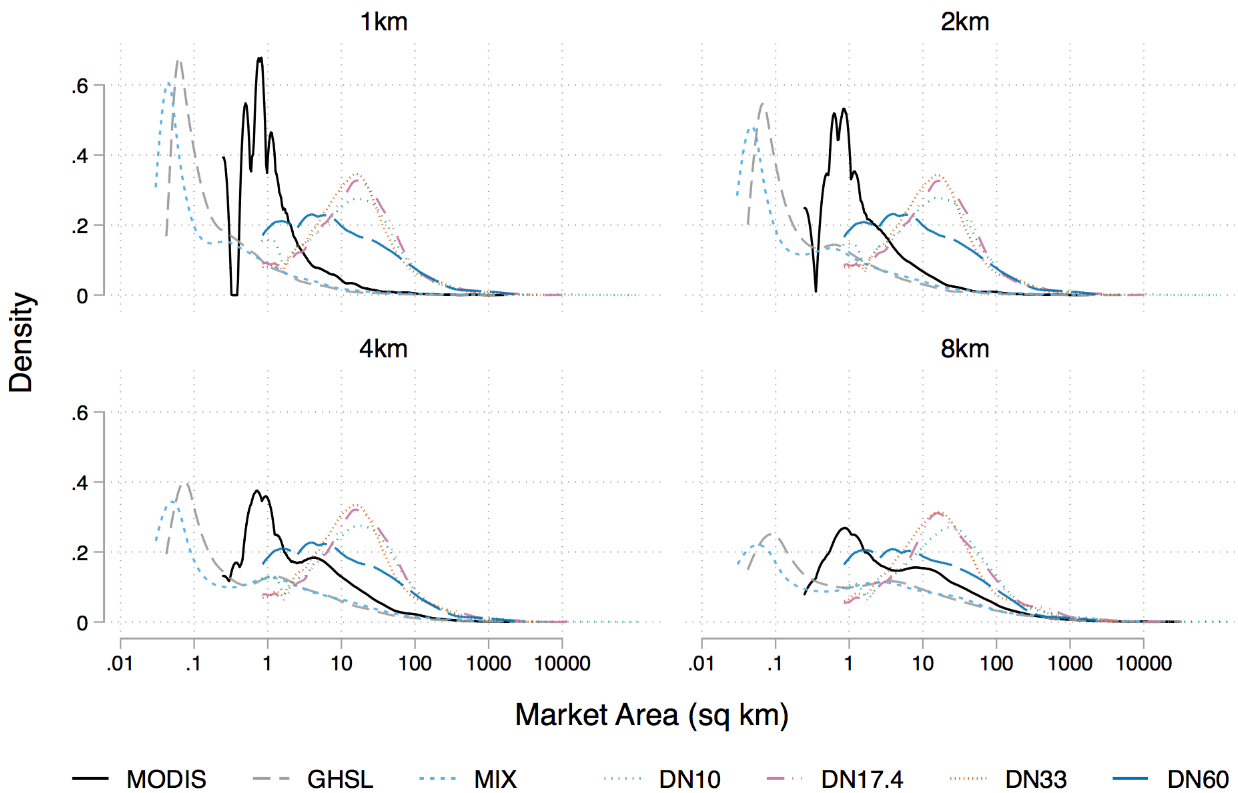
Notes: Vertical lines denote the 90th, 95th, 99th, 99.5th percentiles of DNs. Histogram formed using a 3% random sample of pixels.

Figure A2: Combining Polygons to Form Markets



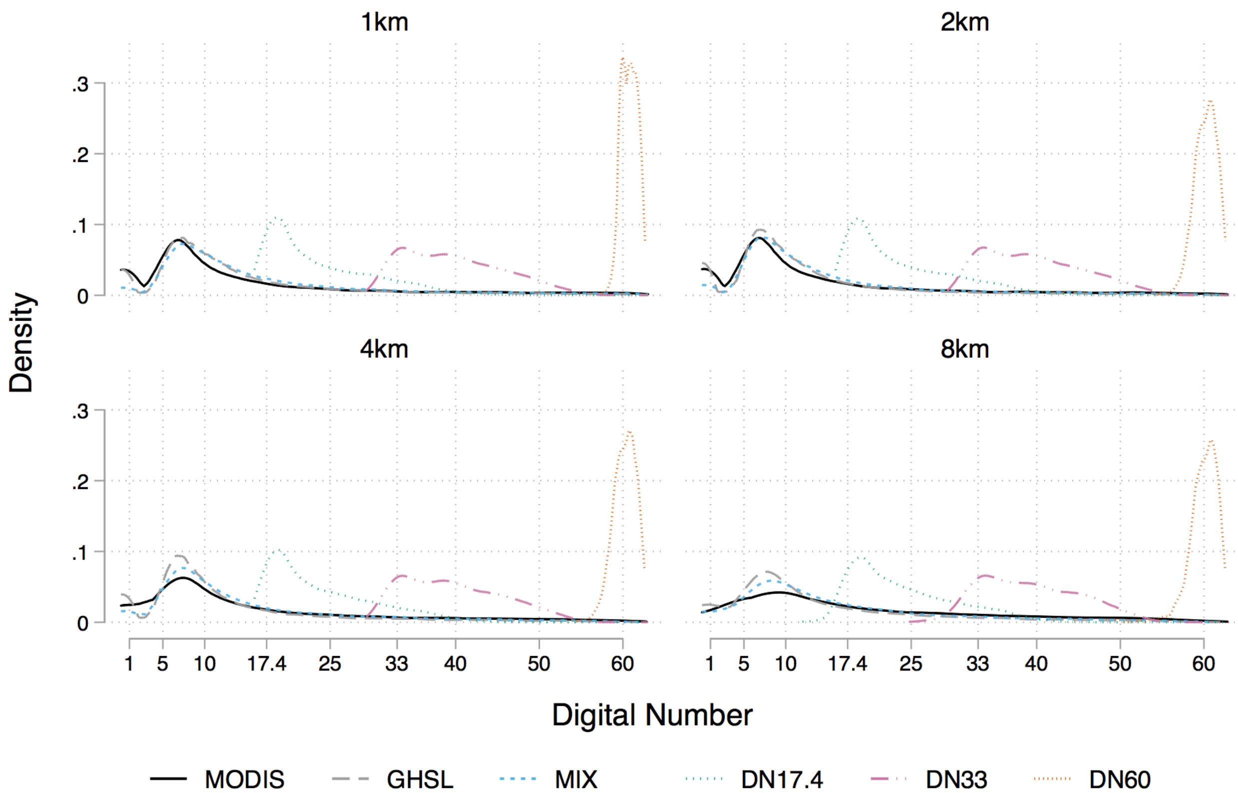
Notes: Panel (a) illustrates DN10 threshold markets. Panels (b-d) shows $2km$, $4km$ and $8km$ buffers, respectively. The last panel shows the aggregated $8km$ buffered markets.

Figure A3: Distribution of Land Area



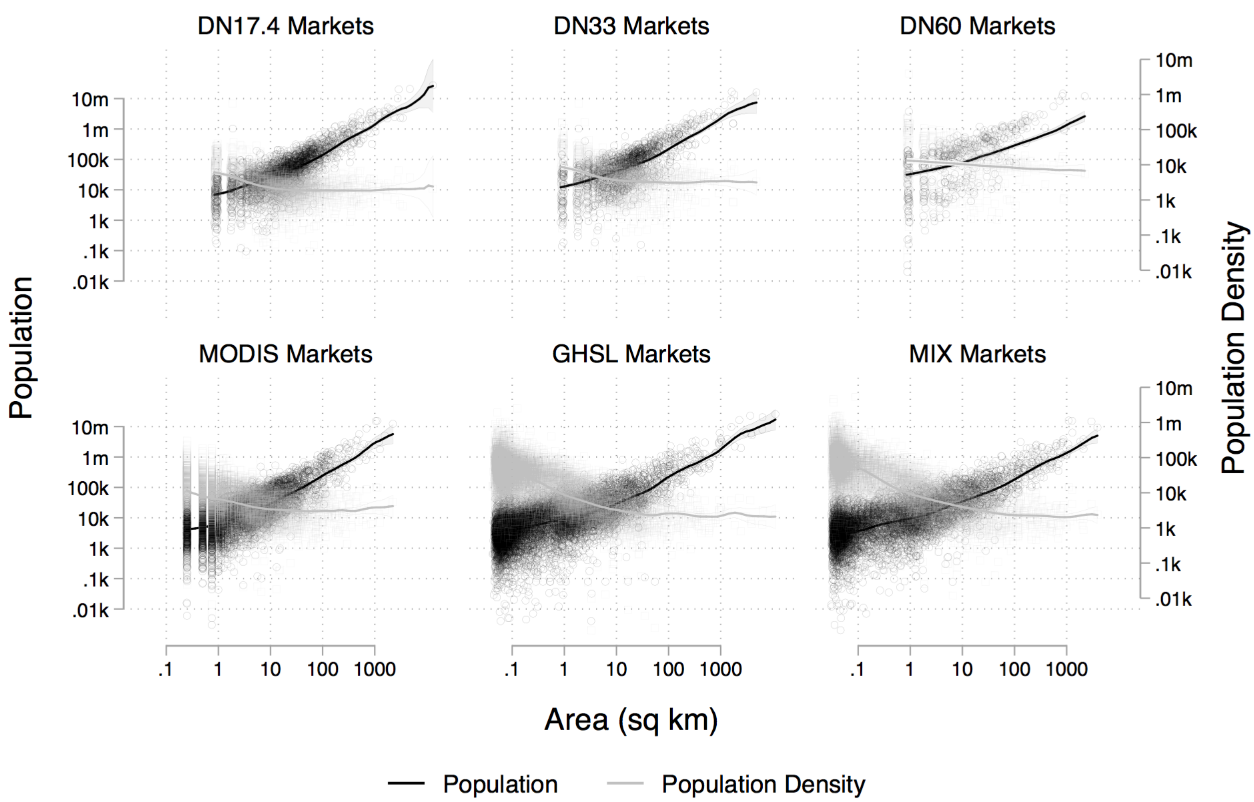
Notes: Figure reports the distribution of market land area, by market definition.

Figure A4: Distribution of Minimum Nightlight DN Values



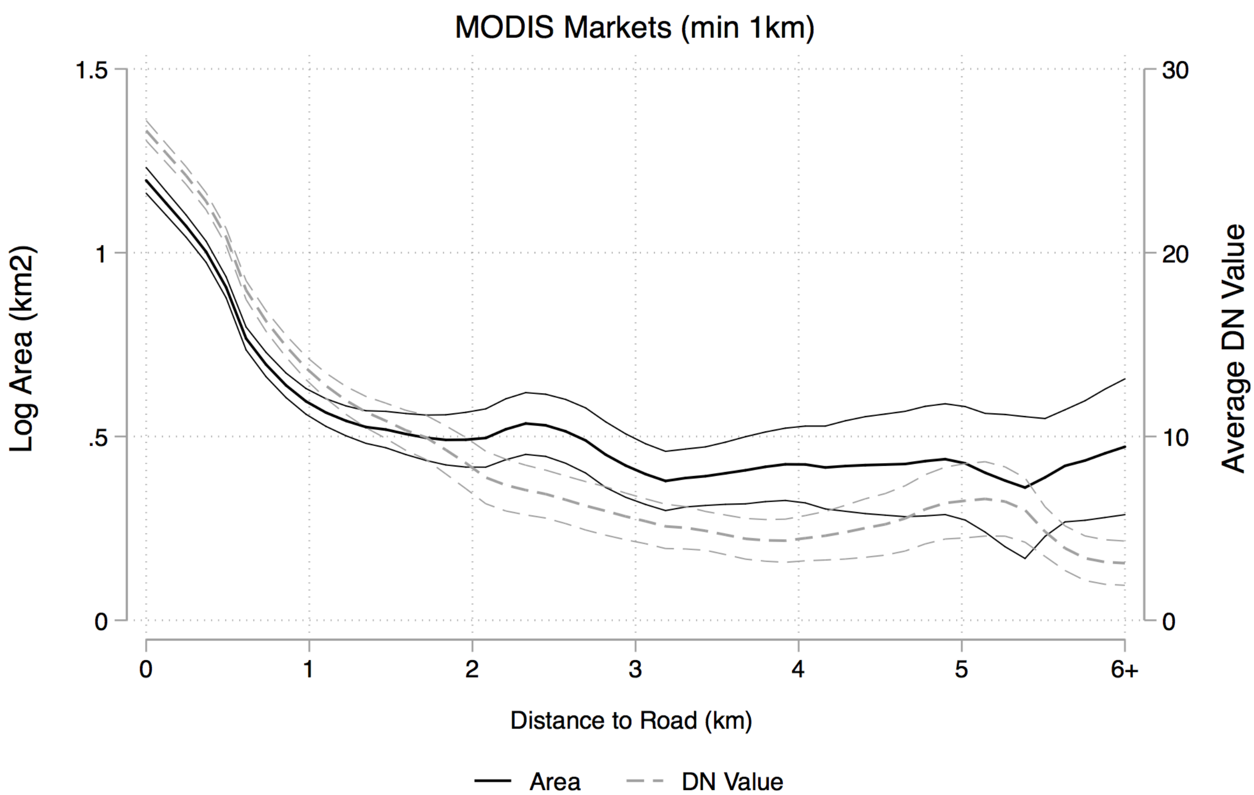
Notes: Figure reports the distribution of minimum DN values, by market definition.

Figure A5: Population versus Land Area, $4km$ Buffer



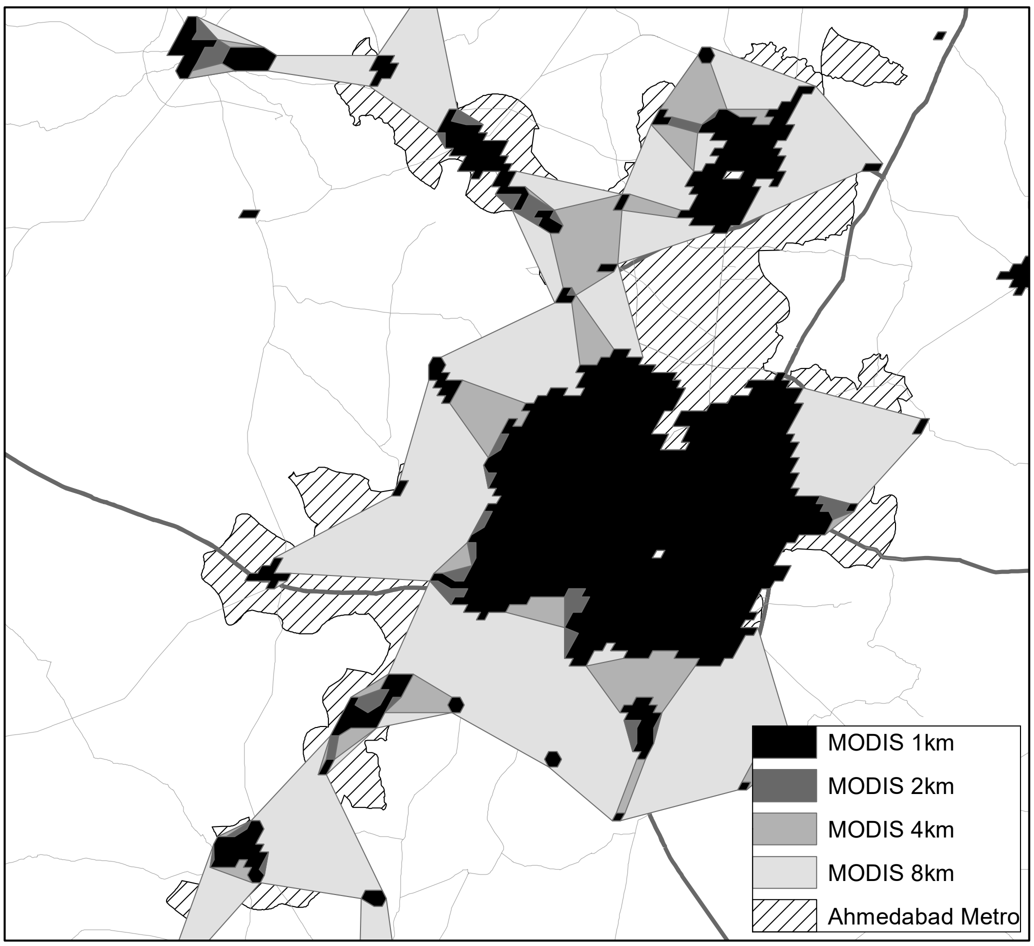
Notes: Figures reports relationship between market size, population and population density. Markets are buffered at $4km$. Population from 2011 Census.

Figure A6: Proximity to Roads, Coarser MODIS Markets ($1km^2$ minimum area)



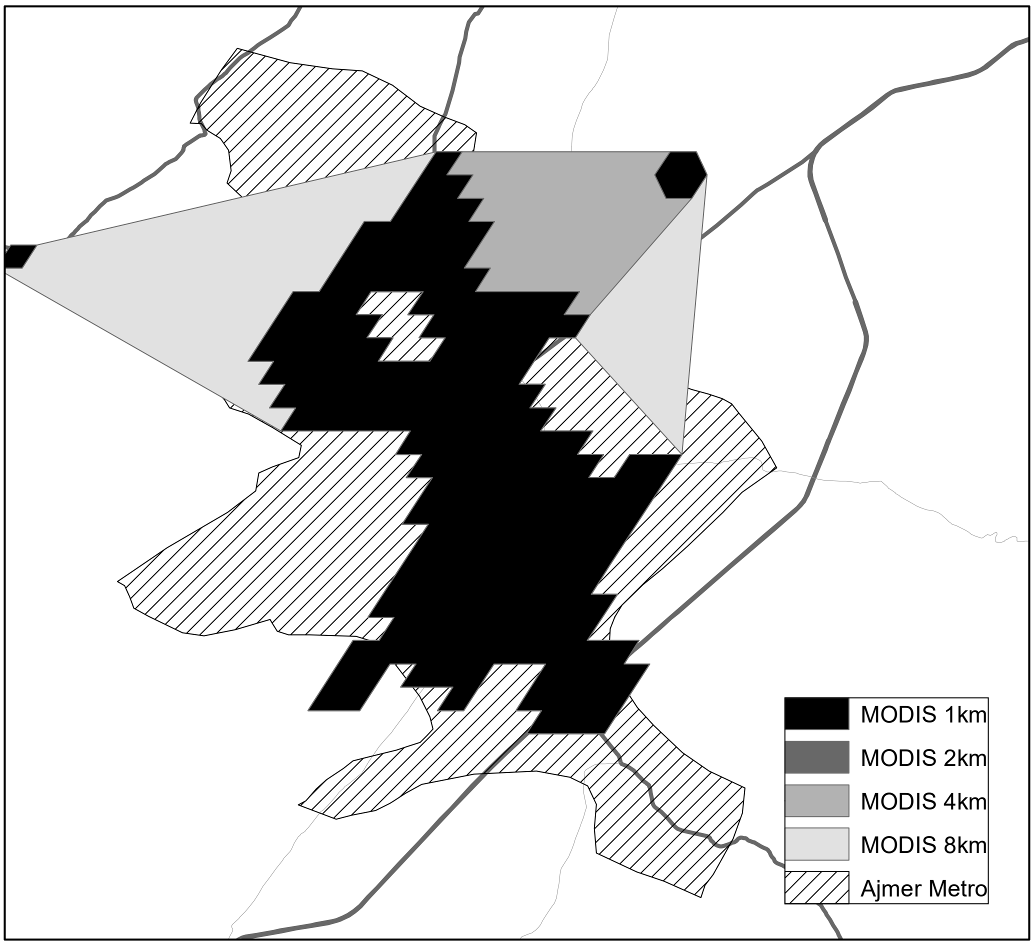
Notes: Distance to road is the shortest distance from market centroid to a primary, secondary or tertiary road. Road data obtained from OpenStreetMaps. Figure uses MODIS markets formed using a minimum threshold of $1km$ and buffered at $1km$. Figure shows 5% and 95% confidence intervals.

Figure A7: MODIS Landcover-Based Markets within Ahmedabad Metro Area



Notes: Map shows MODIS markets in the Ahmedabad metropolitan area. The black outline is the official administrative boundary of Ahmedabad from 2011 Census. Within the administrative boundary, there are 15 1km, 11 2km, 7 4km and 2 8km markets.

Figure A8: MODIS Landcover-Based Markets within Ajmer Metro Area



Notes: Map shows MODIS markets in the Ajmer metropolitan area. The black outline is the official administrative boundary of Ajmer from 2011 Census. Within the administrative boundary, there are 1 1km, 1 2km, 1 4km and 1 8km markets.

A Aggregating Pixels to Markets

To combine clusters of highly lit pixels, we use the Aggregate Polygons function in ArcGis. This function combines polygons within a specified buffer to form larger polygons. Appendix Figure A2 illustrates the tool with lit pixels, focusing the border between Rajasthan and Haryana, two states in India. The gray areas illustrate polygons that are contiguous sets of pixels with a DN that exceeds 10. Notice that there are many unconnected polygons. Merging two polygons forms a larger polygon that contains the land area of the original two polygons plus a land bridge that connects them, whose dimension is determined by the algorithm. The larger is the distance buffer, the larger will be the land bridges that connect polygons. Figure A2a illustrates the results of implementing a $1km$ buffer; Figures A2b through A2d implement $2km$, $4km$, and $8km$ buffers, respectively. For a sub-area within the sample geographic region, Figure A2e illustrates the resulting markets when we impose the $8km$ buffer. Notice that moving from the smallest to the largest buffer collapses the number of markets in this area from more than 20 to just 3.

B Construction of the MIX Layer

This online appendix provides an overview of the builtup classification methodology developed by Goldblatt et al. (2018) for India, Mexico, and the U.S. The methodology uses DMSP-OLS nightlight data as quasi-ground truth to train a classifier for builtup land cover using Landsat 8 imagery. The basic idea is that since lights indicate the presence of human activity, we can train a classifier that uses the spectral signature of daytime images to predict the presence of humans, as indicated by lights. The challenge of using nightlights as a source of ground truth is the blooming of lights. Goldblatt et al. (2018) correct for this blooming as follows. Using their approach and imagery for 2013, we calculate the per-band median values from a standard top-of-atmosphere calibration of raw Landsat 8 scenes. These per-pixel band values are then used to construct commonly used indices to detect vegetation (the normalized difference vegetation index, NDVI), water (the normalized difference water index, NDWI), physical structures (the normalized difference built index, NDBI), and other relevant features. We use these indexes to mask out pixels that appear with high DN from the DMSP-OLS data; the assumption is that these pixels, because they are composed mostly or entirely of water or vegetation, do not contain builtup activity and appear unlit only because of blooming. We then proceed with the classification.

The steps of the methodology are as follows:

1. Designate a pixel as *builtup* if its DN exceeds a threshold. This threshold is set at the 95^{th} percentile of pixels in the training set, which is 17.4 across all India but ranges is allowed to vary across hex-cells (discussed below).
2. Re-classify a builtup pixel as *not builtup* if the Landsat index bands (NDVI, NDWI, NDBI) indicate presence of water, dense vegetation or not builtup activity (as noted above, this corrects for the blooming).

3. Use supervised machine learning to train a classifier (a random forest with 20 trees) with the adjusted builtup/not builtup binary pixels from steps 1 and 2, and the Landsat 8 median-band values and index values as inputs.
4. Use the classifier to construct the posterior probability that a pixel is builtup, and then create binary values of builtup/not builtup status based on this probability (discussed below).
5. Evaluate the accuracy of the classifier by comparing the predicted builtup status of a pixel to a ground-truth dataset that has 85,000 human-labeled pixels that were classified as builtup or not builtup.

In (3), we allow for variation in how the reflectance of India's heterogeneous land cover is associated with urbanization by partitioning the country into an equal-area hexagonal grid with hex-cells that have center-to-center distances of 1-decimal degree, and then treat each hex-cell as an independent unit of analysis. (We also train classifiers for hex-cells that have distances of 4- or 8-decimal degrees, but find that the 1-decimal degree hex-cell is most accurate.) After training the classifier separately within each hex-cell, we mosaic the resulting local classifications to map predicted builtup land cover for the entire country. In (4), we designate a pixel as builtup if its posterior probability exceeds a given threshold that is determined by the Otsu algorithm (Otsu 1979), which is a nonparametric and unsupervised method for automatic threshold selection originally developed for picture segmentation. The method uses a discriminant criterion to identify an optimal threshold that maximizes the between-class variance. We choose the threshold to maximize the variance between builtup and not-builtup classes. In (5), which compares our predicted values of builtup status with human-labeled examples, we achieve an overall accuracy rate is 84%. The accuracy rate is defined as the sum of true positives and true negatives divided by the total sample. Note that this accuracy rate exceeds the MODIS classification accuracy by 2.5% in India; see Table 6 of Goldblatt et al. (2018).