Poverty in HD: What Does High Resolution Satellite Imagery Reveal about Economic Welfare?[1]

Ryan Engstrom[2]

Jonathan Hersh[3]

David Newhouse[4]

June 2016

Abstract

This paper uses Sri Lankan data to investigate the ability of features derived from high spatial resolution satellite images to accurately predict spatial variation in poverty across small areas. Object-based features include the number and density of buildings, shadow area (building height proxy), car counts, the density of paved and unpaved roads of various widths, the type of farmland, and roof material. Spectral and texture features include a vegetation index (NDVI), Speeded-Up Robust Features (SURF), Pantex, Fourier and Gabor transformations. These features are matched to predicted poverty rates, based on estimates of household per capita household consumptions imputed into the 2011 census, for 1,251 Gram Niladhari (GN) Divisions covering a total area of roughly 3,500 sq. km. High spatial resolution satellite features explain about sixty percent of the variation in the share of the GN population in the bottom forty percent of the national distribution of predicted per capita consumption. Measures of building density are particularly strong correlates of poverty. Within satellite scenes, car counts and SURF are strongly correlated with poverty in urban areas, while the share of roads that are paved and roof type are strong correlates in rural areas. Out-of-sample predictions are less accurate but tend to preserve rank. These results suggest that features extracted from high spatial resolution imagery hold considerable promise for contributing to small area poverty estimates within areas covered by household surveys. Features can be used to extrapolate rankings out-of-sample with reasonable accuracy, but the limited size of satellite footprints constrains out-of-sample extrapolation.

Keywords: poverty estimation, satellite imagery, machine learning

JEL classification: I32, C50

---

[1] This project benefited from greatly from discussions with Sarah Antos, Ana Areias, Marianne Baxter, Sam Bazzi, Kristen Butcher, Klaus-Peter Hellwig, Kristen Himelein, Trevor Monroe, Dilip Mookherjee, Pierre Perron, Bruno Sánchez-Andrade Nuño, Kiwako Sakamoto, David Shor, Benjamin Stewart, Andrew Whitby, Nobuo Yoshida and seminar participants at the Boston University Development Reading Group and the Department of Census and Statistics of Sri Lanka. All remaining errors in this paper remain the sole responsibility of the authors. Sarah Antos, Benjamin Stewart, and Andrew Copenhaver provided assistance with texture feature classification. Object imagery classification was assisted by James Crawford, Jeff Stein, and Nitin Panjwani at Orbital Insight, and Nick Hubing, Jacqlyn Ducharme, and Chris Lowe at Land Info, who also oversaw imagery pre-processing. Hafiz Zainudeen helped validate roof classifications in Colombo. Colleen Ditmars and her team at DigitalGlobe facilitated imagery acquisition, Dung Doan and Dilhanie Deepawansa developed and shared the census-based poverty estimates, and we thank Dr. Amare Satharasinghe for authorizing the use of the Sri Lankan census data. Liang Xu provided research assistance. Zubair Bhatti, Benu Bidani, Christina Malmberg-Calvo, Adarsh Desai, Nelly Obias, Dhusynanth Raju, Martin Rama, and Ana Revenga provided additional support and encouragement. The authors gratefully acknowledge financial support from the Strategic Research Program and World Bank Innovation Labs Big Data for Innovation Challenge Grant. The views expressed here do not necessarily reflect the views of the World Bank Group or its executive board, and should not be interpreted as such.

[2] rengstro@gwu.edu Department of Geography, George Washington University, 1922 F Street, Washington DC

[3] jhersh@worldbank.org, jhersh@bu.edu, Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, and Poverty and Equity Global Practice, World Bank.

[4] dnewhouse@worldbank.org, Poverty and Equity Global Practice, World Bank, 1818 H Street, Washington DC

1. Introduction

Despite the best efforts of national statistics offices and the international development community, small area estimates of poverty and economic welfare remain rare. Partly, this reflects the large gaps in the availability of household survey data measuring economic welfare in developing countries. Between 2002 and 2011, as many as 57 countries conducted zero or only one survey capable of producing poverty statistics, and data are typically scarcest in the poorest countries. (Serajuddin et al, 2015). But even where household surveys are conducted regularly, they are typically too small to produce reliable estimates below the district level. As a result, generating estimates for smaller areas requires both a household welfare survey and contemporaneous census data, and the latter is typically available once per decade at best. Furthermore, in many conflict areas, safety concerns prohibit survey data collection altogether. Lack of timely information on living standards in small areas complicates the efforts of policymakers and aid organizations to direct scarce resources to the poor, and those of their constituents to hold them accountable for doing so.

Satellite imagery has generated considerable enthusiasm as a potential supplement to household data that can help fill these severe data gaps. In recent years, private companies such as DigitalGlobe and Airbus have rapidly expanded the coverage and availability of high spatial resolution imagery (HSRI), driving down commercial prices. Planetlabs currently operates more satellites than any organization other than the US and Russian governments, and by mid-2016 will have launched enough satellites to acquire coverage of the entire globe, with imagery resolution of 4 to 5 m per pixel, on a daily basis. Continued technological advances will increasingly allow social scientists to benefit from this type of imagery, which has been utilized intensively by the intelligence and military communities for decades.

Historically, the most popular satellite based measure for economic applications has been night-time lights (NTL), which measures the intensity of human light activity captured by satellites. Country-level estimates of the relationship between remotely sensed satellite proxies and aggregate welfare have shown strong correlations between night-time lights and GDP (Henderson et al., 2009). Subsequent research, however, has suggested that much of this link is explained by the strong correlation between night-time lights and population density. The relationship between lights and wages or other measures of income appears weak at best (Mellander et al., 2013), casting doubt on its reliability as a proxy for welfare, apart from gross distinctions at the country level. Areas with low levels of night time lights exhibit substantial variation in living standards, making it difficult to use NTL to identify the poorest parts of these dimly lit areas. Furthermore, NTL is ill-suited for identifying variation in welfare within small areas because of its low spatial resolution. Even the most advanced NTL satellite, VIIRS, has a spatial resolution at nadir of 0.75 km^2 and can increase to over 1.0 squared km.[5] NTL derived from even a state of the art satellite is therefore unlikely to be informative of welfare for areas smaller than districts.

New developments in computer vision algorithms, however, are increasingly allowing analysis to extract meaningful information from daytime imagery.[6] Modern image classification tools like

[5] Pixel size can vary depending on the angle of the satellite relative to the ground site.

Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) when applied to HSRI, have the capability to classifying objects such as cars, building area, roads, crops and roof type. These objects may be more strongly correlated with local income and wealth than NTL. (Krizhevsky, Sutskever, and Hinton, 2012). For example, Marx, Stoker and Suri (2013) use daytime satellite image data in Kenya to compute measures of roof luminosity to act as a proxy for investments in housing quality.

An alternative approach to analyzing HSRI involved calculating textural and spectral variation in the imagery instead of identifying objects (Graesser et al. 2012, Engstrom et al. 2015, Sandborn and Engstrom 2016). In this approach the spatial and spectral variations in imagery are calculated over a neighborhood (a group of pixels) to characterize the local scale spatial pattern of the objects observed in the imagery. These measures, which we refer to as "texture" or "spectral" measures, capture information about an area that may not be clear from object recognition alone. Several texture measures have been developed and used for image recognition. Two examples are the histogram of oriented gradients (HoG), which captures ordered gradients in an image, and Speeded up Robust Features, which identifies points with sharp contrasts such as right angles and edges (Bay et al, 2006). Spectral features include the normalized difference vegetation index (NDVI). NDVI, which is the most popular spectral vegetation index, uses the differences in red and near infrared spectral observations to develop a measure that is highly correlated to the amount and density of vegetation. Together, these features are able to pick up irregularity of buildings, vegetation and roads that are typical of urban informal settlements and slum environments (Graesser et al. 2012, Engstrom et al. 2015).

Finally, there is an increasing interest in the use of "unsupervised learning approaches", which apply machine learning techniques directly to classify imagery on the basis of either objective or subjectively classified training data. Glaeser, Kominers, Luca, and Naik (2015) apply texture-based machine vision classification to images that are captured from Google Street View, trained using subjective ratings of the images on the basis of the perceived safety. They build a support vector machine prediction model, and show these texture features can reliably predict block level income in New York City. However, there are practical issues in using this approach in rural areas of poor countries. In particular, street view coverage in developing countries is sporadic, and will likely remain so until the autonomous vehicles that Google deploys to capture street view imagery are adapted for use in those countries.

Jean et al (forthcoming) applies a novel unsupervised learning approach, in which a convolutional neural network algorithm is used to classify imagery as more or less poor, to predict village-level average consumption in Nigeria, Uganda, Tanzania, and Malawi. The approach combines high-resolution daytime imagery from Google static maps with NTL. They first extract a set of features from the Google imagery, using a convolutional neural network trained to predict NTL. In the second step, these features are then used to predict average consumption at the cluster level, taken from living standard measurement surveys.[7] This two-step approach allows the CNN model to extrapolate relevant features of the imagery using the full range of NTL. This can improve upon the use of NTL alone to predict poverty, to the extent that features of the imagery that explain variation in higher ranges of NTL are also correlated with household welfare at lower levels of NTL. The model predicts well, even out of sample, as it

[7] It is not clear whether the consumption aggregates have been spatially deflated, and failing to deflate the aggregates would exaggerates the predictive power of population density in explaining variation in welfare.

explains about 42 percent of the variation in village per capita consumption in a pooled sample of all four countries. However, the improvement over NTL alone is modest, as the two-step method raises the r^2 in the pooled estimate from about 0.38 with night-time lights alone to 0.42. This is consistent with the model's substantial predictive power being mostly driven by differences in population density derived from NTL.

This paper contributes to this nascent literature on the use of satellite data to complement survey-based welfare measures, by investigating the ability of object and texture features derived from HSRI (High Spatial Resolution Imagery) to estimate and predict poverty rates at local levels. The area of our study covers a 3,500 square kilometer testing and training site in Sri Lanka. This area fully contains 1,251 Grama Niladhari (GN) divisions, for which poverty estimates were generated by applying data from the 2011 Census to a model estimated from the 2012/13 Household Income and Expenditure Survey.[8] The average area size of our unit of observation (GN) is 2.17 sq. km. That is about two and a half times the size of the average US census block, about one sixtieth the size of the average US census tract, and about one sixth the size of wards in Chicago, the smallest administrative unit in that city.

For each GN division, the object features extracted from imagery include the number of cars, number and size of buildings, type of farmland (plantation vs. orchard), the type of roofs, the share of shadow pixels, road extent and road material, along with textural measures of the type mentioned above. These objects are identified using a combination of Convolutional Neural Networks (CNN) and manual "heads up" digitization. Texture and spectral features were also computed directly from the imagery. Both object and texture features are matched to household estimates of per capita consumptions imputed into the 2011 census, for the 1,251 GN Divisions.

We investigate three main questions: 1) To what extent can variation in GN Division poverty rates, defined according to the 10/20/30/40th percentiles of the national distribution of predicted per capita consumption, be explained by high-resolution features derived from satellite imagery? 2) Which features are most strongly correlated with poverty and how does that vary for urban and rural areas, and between and within imagery scenes? and 3) How do prediction models built using high resolution satellite features fare when predicting into geographic adjacent areas?

We find that: i) satellite features are highly predictive, as they explain about sixty percent of the variation in the share of the GN population in the bottom forty percent of the national distribution. ii) Measures of built-up area and building density are particularly strong correlates of welfare both in urban and rural areas. Within scenes, car counts and SURF are important correlates in urban areas, while the share of paved roads and type of roof are strong correlates in rural areas; iii) Out-of-sample predictions are less accurate but tend to preserve rank. These results suggest that features extracted from high-resolution imagery hold considerable promise for contributing to small area poverty estimates within areas covered by household surveys, while, when considering the full set of tested indicators, the limited size of satellite swath widths may constrain more accurate out-of-sample extrapolation.

[8] For more details, see Department of Census and Statistics and World Bank (2015)

This paper is the first to our knowledge to apply a supervised learning approach, for a wide variety of features derived from HSRI, to estimating poverty and welfare at local levels. It is also the first to relate features derived from satellite imagery to small area estimates of poverty, generated using census data linked to an official welfare aggregate. Compared with more opaque prediction methods, identifying specific features and textures enables a greater understanding of the performance of particular features in different contexts. For example, our approach makes it straightforward to investigate how well particular features predict poverty rates in urban and rural areas and within and across scenes of imagery. In addition, this method demonstrates the potential to extract indicators from imagery that are relevant for other purposes, such as better understanding road networks, agricultural productivity, and patterns of urbanization.

The paper proceeds as follows: Section 2 presents the methodology, summarizes how the data were created and presents brief summary statistics. Section 3 examines the predictive power of high resolution satellite features (HRSF) to validate poverty in small areas at the GN Division level. Section 4 tests whether prediction models using only HRSF are sufficiently accurate to extrapolate out of sample to areas which were not used to build the prediction models. Section 5 explores robustness using alternative model specifications. Section 6 concludes.

2. Data Description and Methodology

Figure 1 shows the areas in Sri Lanka selected to be included in our sample, shown highlighted in red. We collected imagery covering 32 Divisional Secretariat (DS) Divisions, the administrative area one level higher than GNs. A random sample of DSes was not feasible. Circa the time period required VHRI of suitable quality was not available in every area of the country. We sampled imagery conditional on the VHRI that was available, intending to draw equal areas from urban, rural, and estate sectors.[9] The total sample contains 1,251 GN divisions, with each GN division containing 10,000 persons on average according to the 2011 Census, and covering an average area of 2.12 square kilometers.

2.1 Creation of GN Poverty Statistics

Poverty statistics at the GN level were estimated using household level estimates of poverty derived from the 2011 Census of Population and Housing. The method and models are identical to those used to generate official poverty estimates at the DS Division level from the census (Department of Census and Statistics and World Bank, 2015). For each household in the census, per capita consumption was estimated based on models developed from the 2012/13 Household Income and Expenditure Survey (HIES). The models, which were estimated using the PovMap 2.0 software package, employ the canonical Elbers, Lanjouw, and Lanjouw (2003) methodology. Estimates were derived from 16 regional models of log per capita consumption, with adjusted r-squared statistics ranging from 0.40 to 0.50, that were each simulated 100 times. The simulations assume normal error terms with respect to log per capita consumption, and allows for heteroscedasticity with respect to model predictors. Per capita consumption in the HIES was

[9] Sri Lanka is unique in that it classifies sectors as urban, rural, or estate. The Estate sector is a classifications by the Department of Census that refer to "plantation areas, which are more than 20 acres in extent and have 10 or more residential laborers." For the purposes of this study, except for sample stratification, the estate sector is grouped together with the rural sector.

deflated using a district-level spatial price index derived from unit value food prices for households in between the 10th and 40th percentiles of per capita consumption (Newhouse et al, 2016). While this is an imperfect adjustment for price differences across districts because it fails to account for price differences in non-food items, failing to deflate the welfare aggregate to account for spatial price differences would overestimate the extent to which population density derived from satellite data predicts variation in poverty.

To understand how the prediction performance varies according to the prevalence of poverty, we calculate relative poverty rates at artificial levels, as opposed to relying on the official poverty line.[10] We derive GN headcount poverty statistics using the Foster-Greer-Thorbecke methodology at four different thresholds of poverty -- 10/20/30/40% of national income -- according to the formula

$$povrate_{j,z_k} = \frac{1}{N_j}\sum_{i=1}^{N_j} I(y_i < Z_k)\left(\frac{Z_k - y_i}{z_k}\right)^0 = \frac{1}{N_j}\sum_{i=1}^{N_j} I(y_i < Z_k)$$

where $PovRate_{j,z_k}$ is the poverty headcount rate in GN Division j at the z_k poverty threshold, defined at 10/20/30/40 percent of national income. The variable y_i is per capita household consumption for individual i, averaged over the 100 simulations, N_j is the total population of GN j. We calculated the national poverty threshold level, z_k, as the $k \in \{10,20,30,40\}$ percentile of the national predicted per capita consumption in the census averaged across the 100 simulations. I is an indicator function that in this case takes on one if individual i's per capita household consumption lies below the relative poverty threshold Z_k. Since we only consider the headcount rate, this is equivalent to the ratio of the number of individuals below the threshold to the total population in each GN Division.

Figure ***2*** plots kernel density estimates of these four poverty thresholds. The lower thresholds (10% and 20%) show considerable skewness, with fat right tails, indicating most GN Division have very low or zero level of poverty rates at these lower thresholds. But as the threshold increases the distribution becomes much less skewed, and by the 40% relative poverty line (RPL), the distribution of poverty rates is much closer to uniform. Table 1 presents summary statistics for model variables, including the average GN relative poverty rates for the four thresholds. At the 10 percent threshold, the average GN in the sample has 6 percentage points of poverty, while at the 20% threshold, this increases to an average of 13.8. At the 30% and 40% levels this increases to 22.3 and 31 percentage points respectively.

Poverty rates in Sri Lanka exhibit considerable heterogeneity across space (Department of Census and Statistics and World Bank. 2015). Figure ***3*** shows this poverty heterogeneity by plotting density estimates of the 40% RPL by province for seven provinces, based on the sample of 32 DS Divisions. The lowest poverty rates are seen in the North Central and Western provinces; North

[10] According to the official poverty line, which was developed in 2002 and subsequently updated for inflation, 6.7 percent of the population was poor in 2012/13.

Western and Southern provinces show moderate amounts of poverty, and Central, Eastern and Uva provinces show the highest level of poverty. As the large urban areas are located in the West, poverty rates are larger in rural areas than in the urban areas, as shown in **Table 2**.[11] For all four relative poverty lines, urban GNs in the sample have poverty rates around half of the levels of rural poverty. Poverty at the 10% threshold is an average of 3.7 percentage points in urban areas and an average of 6.8 percentage points in rural GNs. Given the substantial difference between urban and rural poverty rates, it will be important to consider how the correlates of poverty will differ for the two areas.

2.2 Comparison of GN Poverty Rates and Mean NTL Reflectance

A simple visual comparison between mean NTL and GN Division poverty rates illustrates the difficulties of using NTL as a primary source of information for estimating sub-national poverty rates. Figure 4 presents a panel of three images for the Divisional Secretariat of Seethawaka: mean raw NTL (left), poverty rates derived from the 10% national income threshold (middle), and log of mean population density (right). Comparing the left and middle panels, there is very little association between GN Divisions that have low NTL reflectance and those that are high in poverty. While NTL tracks the general contours of poverty for the DS – lower poverty areas in the Northwest and higher poverty areas in the Southeast – this gross association is of limited use for practical public policy applications such as poverty targeting.

What NTL does appear to approximate is population density of the underlying GN Divisions. This corroborates the findings of Mellander et al. (2013). Comparing the right and left panel there appears to be a strong association between areas that are high in NTL and those that have a larger population density. The information content of NTL is limited in that human welfare encompasses more than just the intensity of lights outputted at night. HRSI daytime imagery provides a much richer view of settlement and land use patterns, which may be able to provide a more a useful window on economic welfare.

2.3 Creation of High Resolution Spatial Features (HRSF)

The satellite imagery consists of 55 unique "scenes" purchased from Digital Globe, covering areas specified in our sample area.[12] Each "scene" is an individual image captured by a particular sensor at a particular time. Images were acquired by three different sensors: Worldview 2, GeoEye 1, and Quickbird 2. These sensors have a spatial resolution of $0.46m^2$, $0.41m^2$, and $0.61m^2$, respectively in the panchromatic band and $1.84m^2$, $1.65m^2$, $2.4m^2$ respectively in the multi-spectral bands. Pre-processing of imagery included pan-sharpening, ortho-rectification, and image mosaicking.

Object features were classified using the assistance of two technical partners: Orbital Insight and LandInfo. Orbital Insight produced object classification for three variables: The share of the GN division that is built-up (i.e. consists of buildings), the number of cars in the GN, and the share of pixels in the GN that were identified as shadow pixels, which is a proxy for the gross floor area, or height, of buildings. The classification method is similar to Krizhevsky, Sutskever, and Hinton (2012) which utilizes convolutional neural networks (CNN) to build object predictions from raw imagery. LandInfo classified the remaining objects, which included roof typ, paved and unpaved roads of different widths, railroads, and the type of agriculture. Landinfo used a combination the

[12] Particular thanks to Digital Globe is due for their "Seeing the World" program, which offered very high spatial resolution imagery at reduced rates for non-commercial purposes.

Trimble eCognitition and Erdas Imagine software platforms to classify objects, except for roads, which were classified manually.

The CNN classification algorithm used by Orbital Insight involved four steps:

1. Ingestion/Tiling
2. Model Development
3. Classifying All Pixels Using the Trained Model
4. Aggregating Prediction Results to GN Division level

The tiling stage split the large images into many small images or tiles, in order to make the modeling computationally scalable, as each tile could be distributed to a different GPU core for greater efficiency. In the model development stage, the classification model was trained and tuned. Model building began by manually classifying or labeling a sub-sample of the imagery as a positive or negative value for a given object using a crowdsourced campaign. The classified data was split into an 80% training and a 20% testing set, where the training set was used to build the model. This allowed sample prediction metrics, presented below, to be calculated using the withheld test set. Training was run for 60,000 iterations using the Nesterov solver method, a variant of stochastic gradient descent. To get a sense of the accuracy, Figure 6 shows the receiver operator characteristics, or ROC curve, summarizing the classification accuracy of the developed area building classifier. The ROC curve presents the true positive rate on the y-axis, against the false positive rate on the x axis as the discriminant threshold is varied. A classifier that is no better than random would correspond to the 45 degree line, while a perfect classifier would correspond to the point (0,1) in the ROC space. Improvements in classifier accuracy are shown as the curve moves up and to the left. Figure 6 suggests the classification algorithm is highly accurate, corresponding to a 90% accuracy rate overall. Once the model was full trained, the team applied the trained model to the full set of imagery. The results were then summarized at the GN level.[13]

Two visual examples of the CNN training classification -- for built-up area (building footprint) and cars -- are shown in figures 5 and 7 respectively. Figure 5 shows raw satellite imagery in the top panel, and classified imagery according to the CNN algorithm in the bottom panel. Areas highlighted in green in the lower panel are true positive building classifications, where areas classified by the algorithm as buildings were confirmed as such by manual identification. Areas in red are false positives that were erroneously classified by the algorithm as buildings even though the underlying satellite imagery do not reveal any developed buildings. Figure 5 shows that while the classification has a high degree of accuracy, there still exist some areas with false positives. The overall classification accuracy of 90%, however, provides sufficient signal to meaningfully include in prediction models.

Turning to cars, figure 7 shows raw satellite imagery from Colombo, where locations of cars that were identified by the CNN algorithm are highlighted by blue circles. This visual check ensures that indeed the algorithm appears to be correctly classifying many cars – and the classifications appear to occur on the road paths or in parking lots where they would be expected to appear. It does appear that there are some false negatives – areas on the image where cars visually appear that were not classified by the algorithm as cars – particularly in areas where cars and vegetation

[13] The GN shapefile was in turn based on the GN shapefile provided by DCS, modified to correct manually identified errors.

co-occur. However, given that the classification accuracy of the algorithm is also above 90%, we conclude that the measured prevalence of cars should also provide enough signal to contribute explanatory power to poverty prediction models.

LandInfo extracted roof type, paved and unpaved roads of different widths, railroads, and the type of agriculture, using a combination of an object-based image analysis (OBIA) methodology utilizing the Trimble eCognition software platform (Benz et al., 2004) and visual interpretation. OBIA is a common classification methodology applied to HSRI because it is designed to map objects that are larger than the pixel size (Blaschke, 2010). For this OBIA analysis, the first step was to segment the image into polygons that represent the features of interest based on spectral and spatial homogeneity of those features. The segments were created based on the degree of homogeneity using a range of scale, shape, and other parameters. Once an image was segmented, the objects themselves were then classified, in this study using Erdas Imagine Software. Using this methodology, all of the roofs of individual buildings within the imagery were extracted and the type of roof was classified as asbestos, aluminum, clay, or grey roofs, in part based on validation of selected roofs in Colombo. The roof types of a few selected buildings were validated by direct observation. Roads and agriculture were extracted using a combination of eCognition and manual visual interpretation. Visual interpretation was required because the roads were covered by trees and in shadow, making automated detection difficult. In addition, the imagery was not always taken during the growing season, which along with the small, irregular shapes of the agricultural plots made automated extraction impractical.

Figures 8 and 9 provide two visual examples of the classification algorithms employed to classify roof type (figure 8) and roads, railroads, and road width (figure 9). Figure 8 shows the classification for roof type, which classifies roofs as either aluminum (blue, green or white), asbestos, clay, or grey. Roof types tend to co-occur where buildings are located, and the classification type seems to correspond to the type of roof that can be visually identified. Figure 9 shows roads and railroad classification. Roads pass through the image in a variety of widths, with the road network being replicated by the classification algorithm.

Finally, textural and spectral features were created using a block size of eight pixels and scales of 8, 16, 32, and 64 meters using a methodology similar to Graesser et al. (2012). This resulted in an output image comprised of 165 bands at a spatial resolution of 12.8 or 16m depending on native, multispectral resolution of each of the sensors. The seven textural features calculated for this study were:

1. Histogram of oriented gradients (HoG), which captures edge orientations and sorts them into a histogram (Dalal and Triggs 2005).
2. PanTex, which is a built-up presence index derived from the grey-level co-occurrence matrix (GLCM) (Pesaresi et al. 2008).
3. Line support regions (LSR), which characterize line attributes (Yu et al. 1999)
4. Local binary patterns moments (LBPM), which define contiguous regions of pixel groups and sorts them into a histogram (Wang and He, 1990).
5. Fourier transform (FT) which examines pattern frequency across an image (Smith 1997).
6. Gabor, a linear Gaussian filter used for edge detection (Gabor 1946)

7. Speeded Up Robust Features (SURF), an algorithm that extracts key points (i.e., edges and corners) from an image through pyramidal Gaussian based decomposition (Bay et al., 2006).
8. The Normalized Difference Vegetation Index (NDVI), the most widely used vegetation index that provides information about the presence and abundance of vegetation (Tucker 1979).

The spectral features calculated were simply the means of the four individual bands, Blue, Green, and Near Infrared. Once the spatial and spectral features were calculated, the mean, standard deviation, and sum were determined for each GN Division. Previous research has indicated that these features are correlated with census data that indicate poverty such as slum conditions, population density, solid waste collection, unimproved sanitation (Sandborn and Engstrom, 2016) and to map informal and slum areas within cities (Graesser et al. 2012, Engstrom et. al. 2015).

Because these texture measures may be novel to readers without backgrounds in remote sensing, further description and some intuitive explanation for some of these measures may be helpful. According to Dalal and Triggs (2005), the motivation for HoG is that we can characterize objects "by the distribution of local intensity gradients or edge directions, even without precise knowledge of the corresponding gradient or edge positions". HoG calculates a smoothed histogram for each cell, from which we can capture the mean, variance, skew and kurtosis of the histograms. Pantex can be thought of as a spatial similarity index, where a cell is compared to adjacent cells in all directions that a chess queen can move: vertically, horizontally and diagonally. The Pantex level that is assigned to a cell is the minimum contrast variance – so called grey level co-occurrence -- measured in all directions. Forests will have a low Pantex level, since cells in all directions have similar contrast. Cells containing straight roads will also have low Pantex levels, as cells where the road continues will have similar contrast and Pantex assigns the minimum neighborhood value. Cities dense with many buildings in all directions will have high Pantex values, as cells in all directions have high contrast variance. In this sense, one can think of high Pantex areas as indicating areas with a high density of human activity. SURF detects local features, and is often used for object recognition. Areas with many identifiable objects – such as areas with right angles, corners, or areas with regular grid patterns– would have larger SURF values relative to areas with chaotic or irregular spacing.

2.3 Methodology

Our methodological aim is to find which high-resolution spatial features are consistent predictors of local poverty rates in absence of any other on-the-ground information on the GN itself besides its size and sectoral classification. We begin with a baseline model of the form

$$\underbrace{y_j}_{poverty\ rate\ in\ GN\ j} = \underbrace{X_j'\beta}_{satellite\ features} + \underbrace{\Delta_j\Lambda}_{GN\ controls} + \underbrace{\varepsilon_j}_{error} \quad (1)$$

where y_j is the relative poverty rate for GN j at either the 10/20/30 or 40 percent of national income, X_j is a matrix of the object, textural, and spectral HRSF, and Δ_j are a set of two parsimonious controls for each GN. The two controls contained in Δ_j are the log of the GN area,

which can be derived from the GN division shape file, and whether the GN is officially classified as urban or rural.[14]

A further complication, alluded to when discussing NDVI, is that the features are built from a composite of 55 discrete satellite images, each of which may have been taken during a different time of the year. Our measure of vegetation, NDVI, for example, may be higher for images captured during wetter months versus the drier months, irrespective of the average vegetation in a GN during the year. Other features, such as car counts, will vary greatly depending on the day of the week an image was taken (i.e., weekend or holiday vs. mid-week). The timing of Digital Globe's image capture is exogenous with respect to poverty, and will therefore introduce substantial classical measurement error when examining correlations across images.

To account for this issue we estimate the model

$$(2) \qquad \underbrace{y_j}_{poverty\ rate\ in\ GN\ j} = \underbrace{X_j'\beta}_{satellite\ features} + \underbrace{\Delta_j\Lambda}_{GN\ controls} + \underbrace{Z_s'\theta}_{imagery/scene\ FEs} + \underbrace{\delta_j}_{error}$$

which uses a series of scene ID or imagery fixed effects (FE), Z_s' indicating the unique satellite image used to generate a given GN's spatial features. One disadvantage of this approach is that in using scene FEs precludes the option of extrapolating poverty rates outside the geographic area covered by the imagery, as estimating (2) out of sample requires $\hat{\theta}$ to be known.[15] In section 3, we estimate models with and without scene ID fixed effects.

This baseline methodology presents several methodological challenges. First, while the support of y_j is contained on the interval [0,1], standard OLS regression will report standard errors and coefficient estimates with respect to an unbounded dependent variable. In later versions of this work, we hope explore the performance of models, such as fractional response regressions, which appropriately bound the dependent variable. Second, although standard errors are clustered on DS Division, failing to account for spatial auto-correlation in the error term may overstate the precision of the estimated model. In the robustness section, we also test for the presence of spatial auto-correlation. Third, the OLS model reports a linear approximation of the relationship between of X_j' and the dependent variable.

3. Poverty Validation Using High Resolution Features

In this section, we explore the ability of features derived HSRI to validate or explain variation in the incidence of poverty at the GN level. Table 4 presents estimates of model (1), using as dependent variables GN poverty rate, as defined by the 10/20/30 and 40 percentiles of national

[14] The urban/rural classification in Sri Lanka is based on an official administrative definition and greatly underestimates the extent to which the population has urbanized (World Bank, 2016)

[15] One possible methodological alternative that would permit out of sample estimates with scene ID controls is to use a simulation approach. First, one could estimate θ via random effects, that is assume $\theta \sim N(0, \sigma^2)$. With an estimate for $\hat{\sigma}^2$, $\hat{\theta}$ can either be estimated through simulation for each out of sample observation, or modeled through some flexible specification.

predicted per capita consumption in the census. Standard errors are clustered at the Divisional Secretariat (DS) level, since the sample of GN Divisions was drawn at the DS level. The model explains an impressive amount of the variation in poverty rates. R-squared values vary from 0.39 for the 10% poverty line, 0.5 for the 20% poverty line, 0.55 for the 30% poverty line, and 0.59 at the 40% poverty line. In other words, a linear model that includes only remotely sensed information, an urban dummy, and the geographic size of the GN division, explains almost 60 percent of the variation in the share of a small area in the bottom two quintiles of the national welfare distribution.

Several features appear to be strongly predictive of GN level poverty for all four poverty thresholds. According to the level of statistical significance, the strongest object-based predictors are percent of roads paved, GN area developed, and number of roofs, which are significant across all poverty lines. Coefficient estimates vary across specifications, but for the example of the built-up area variable, a 1 percentage point increase in the area of a GN that is developed, even controlling for the number of buildings, is associated with a 0.29 percentage point decrease in poverty at the 10% RPL. In all specifications save for the 10% poverty level, shadow pixels (building height proxy) and the fraction of roofs that are aluminum are statistically significant, and have estimated coefficients that are positive. The positive coefficient on shadow pixels (building height index) indicates that after controlling for built-up area and the number of buildings, areas with taller buildings contain more poverty, perhaps because of nearby slums. The fraction of roofs aluminum coefficient is also positive and statistically significant. At first glance, this is surprising, as previous analysis in Kenya documents that roofs with higher luminosity are associated with lower levels of poverty (Suri et al., 2015). This suggests that the relationship between roof luminosity and poverty is more complex, and is likely context and region dependent.

Among the textural and spectral variables, NDVI and HoG appear to be the strongest predictors. These two variables are highly statistically significant in all of the specifications and their interpretation is consistent with the importance of building density. For NDVI, the positive coefficient suggests that poorer areas tend to have lusher vegetation, which in this case is reflecting dependence on agriculture, as well as sparser built-up area and fewer buildings. For HOG, higher levels indicate greater contrast, created by a larger number of smaller building structures. Therefore, the negative correlation between HOG and poverty is consistent with the negative association between the number of roofs and poverty rates. Finally, the coefficient on Pantex is positive and statistically significant for the 30 and 40 percent poverty thresholds. This is likely also picking up the prevalence of smaller buildings, which tend to have higher Pantex values.

Conditional on the other indicators, agricultural land classification are weakly associated with poverty rates. A ten percentage point increase in the share of the GN that is classified as agriculture is associated with a decline of less than a percentage point in the headcount poverty rate, at all four thresholds. Similarly, a ten percentage point increase in the share of GN agriculture classified as paddy or plantation is associated with a 0.5 percentage point increase in poverty. In part, this could result from both the strong correlation between agricultural land classification and NDVI and random measurement error in the agricultural land classification measure, which after controlling for NDVI would make the remaining variation in agricultural land type largely uninformative.

Table 5 adds to these estimates by including in the linear model "imagery fixed effects", which are a set of binary variables indicating from which raw satellite imagery picture the HRSF are derived. Since the satellite imagery was taken at different periods during the year, seasonally varying indicators may exhibit additional variation across images that is areorthogonal to local poverty rates. Using imagery fixed effects, R-squared values increase about 10 percentage points, to 0.54 at the 10% poverty level, 0.64 at the 20% poverty level, 0.69 for the 30% poverty level, and 0.72 at the 40% poverty level. Note these R-squared values include the amount of variation in the dependent variable contributed by the imagery fixed effects. Imagery fixed effects attenuates the coefficient magnitudes of many of the variables. HOG is no longer strongly significant across the poverty thresholds, though SURF now is significant for the 30% and 40% poverty thresholds. With imagery fixed effects two agricultural variables are now significantly associated with poverty levels, percent of GN area that is agriculture and percent of GN area that is paddy. Higher paddy areas are associated with higher poverty, which the relationship is reversed for GN area related to agriculture.

2.1 Urban and Rural Linear Models

Tables 6-9 present results estimated separately for the 348 urban GNs and the 898 rural GNs. To parse out the GNs into distinct urban and rural groups, we use the official Census definition of urban and rural. In principle, a small area's urban or rural characteristic could be estimated via remote sensing using a probability model based on variables such as NDVI, Built-up density, shadow pixels (building height index), HOG or others.

R-squared values improve when we estimate separate urban and rural models, which without the imagery fixed effects vary from 0.44 to 0.63. At the lowest poverty thresholds, models are slightly more accurate in urban versus rural areas, (0.52 urban R-squared versus 0.44 at the 10% Poverty threshold) though this relationship reverses at higher thresholds (0.6 urban R-squared versus 0.63 rural at the 40% poverty threshold). Regarding coefficients, using separate urban and rural models we estimate coefficients with more precision, which confirms that there is important heterogeneity in the correlates of poverty. We see more significant variables and even more interestingly sign reversals for some variables across urban and rural models. Fraction of roofs that are clay and aluminum show different signs across urban and rural models – clay and aluminum roofs are negatively associated with poverty in urban areas, but positively associated with poverty in rural areas. NDVI and SURF both have sign reversals across models. Higher levels of NDVI (vegetation) are negatively associated with poverty in urban areas, where the reverse is true in rural areas. This speaks to the different role that vegetation plays in urban versus rural areas. In urban areas, lusher vegetation stems from luxury items such as public parks or private garden plots or lawns. In rural areas, higher vegetation levels may indicate higher levels of lower-wage agricultural labors.

As a further check of the accuracy of the linear models, we plot the predicted GN poverty rates against the true poverty rates, also known as predicted-true plots. Figures 10-13 present these for the linear models which use separate models for urban and rural provinces. Each point represents a GN, where the location on the x-axis corresponds to the true poverty rate and location on the y-axis is the model predicted poverty rates. A model that is perfectly able to predict poverty using HRSF would represent a 45 degree line, starting at the origin and ending at the upper right of the graph. Note that for all of the thresholds, the predicted true points are roughly straddling the 45

degree line. Our methods are well able to distinguish high poverty areas from low poverty ones. While noise still is apparent around the 45 degree line, poverty is accurately predicted for the majority of the GNs.

3.2 What share of the explained variance is accounted for by different features?

The results so far have indicate that features derived from satellite imagery explain a large portion of the variation across GN Divisions in poverty, and that the coefficients on selected predictors are statistically significant across all four poverty thresholds. [16] However, the regression results reported above don't necessarily address the question of which indicator or sets of indicators explain variation in predicted poverty rates. To address this issue, we decompose the r^2 of the regression using a Shapley decomposition (Shorrocks, 2013, Huettner and Sunder, 2012, Israeli, 2007). This procedure calculates the average of the marginal r^2 of a set of one or more explanatory variables, which is the amount by which the r^2 declines when removing that set of variables, across all potential orderings of the variable sets. In other words, for a model with k sets of explanatory variables, the procedure will estimate $2^{(k-1)}$ models and average the marginal r^2 obtained for each set of independent variables across all estimated models. and insensitive to the arbitrary order in which variables are entered into the model.

Table 10 displays the results of this r^2 decompositions for national models with no image dummies. For the purposes of the decomposition, variables are grouped into eleven categories. The urban indicator and the size of the GN area, which are not derived from satellite imagery, are treated as two separate categories. The high-resolution spatial features are divided further into seven groups: Roads, Building density, Roofs, Cars, Agricultural land classification, and texture classification. Since measures of building density, broadly defined, are consistently important predictors, these variables are further separated into four categories, one for each variable in the group. The four variables are: Built-up area and the share of shadow pixels identified by orbital insight, the (log) of the number of roofs identified by LandInfo, and the normalized difference vegetation index (NDVI), which is a measure of the presence of vegetation rather than buildings. Texture variables include the remaining textural features described above, including PanTex, the Histogram of Oriented Gradients, and the Gabor filter. Although few texture features are individually statistically significant, as a group they contribute 17 to 20 percent of the explained variation of the model. Variables on roads explain about 12 percent of the variation, which in results now shown is largely accounted for by the share of roads that are paved. The remaining sets of variable explain between 5 and 10 percent of the variation in most specifications, except for the urban dummy with low relative poverty rates, which accounts for less.

The results confirm that explanatory the four measures of building density, taken together, account for the largest portion of the explained variation. These four variables contribute to about 35 percent of the model's explanatory power. The relative strength of these four variables varies according to the dependent variables. When explaining average consumption, built-up area and the log number of roofs account for the majority of this, but when modeling poverty directly, NDVI also plays a large role. Although few texture features are individually statistically significant, as a

[16] The precision of the estimates may however be overstated by the failure to account for spatial autocrrelation among the residuals.

group they contribute 17 to 20 percent of the explained variation. Variables on roads explain about 12 percent of the variation, much of which is driven by the share of roads that are paved. The rest of the sets of variable explain between 5 and 10 percent of the variation in most specifications, except for the urban dummy when the dependent variable is the relative poverty rate at the 10 and 20 percent level. Practically, these results suggest that measures of building density, a variety of textural features, and roads are strong predictors of variation in well-being. However, additional objects such as the type of roof, the prevalence of cars, and the type of agricultural land also contribute a moderate amount of explanatory power.

4. Poverty Extrapolation Using HRF

A strong motivation for using satellite imagery is to extrapolate poverty estimates into areas where survey data on welfare does not exist. While most of the data deprivation that characterizes the developing world occurs at the country level, it is also common for surveys to omit selected regions, due to political turmoil, violence, animosity towards the central government, or prohibitive expense. For example, from 2002 through 2009/10, Sri Lanka's Household Income and Expenditure Survey failed to cover certain districts in the North and Eastern part of the country due to civil conflict, and Pakistan's Household Income and Expenditure Survey excludes the Federally Administered Tribal Areas and Jammu and Kashmir. For this purpose, it is useful to assess how well a model trained in much of the country can predict welfare in areas where household-level data are unavailable.

To assess how well a model travels to a different geographic area, we use a form of "leave-one-out cross-validation" (LOOCV), a common method used to infer statistical out of sample performance (Gentle et al., 2012). In typical LOOCV, a model is fit for every observation excluding one, then the estimated relationship is used to predict into the withheld portion. This is repeated for every observation in the dataset until every observation has an associated predicted value. This ensures that for all observations i, the fitted value of $\hat{y}_i$ used to test the model is not influenced by the relationship between x_i and y_i. Our approach differs form the typical case in that for each estimation we exclude, or "leave out", an entire Divisional Secretariat (DS), an administrative sub-unit at the level immediately below the district. To give a sense of size, our sample contains 55 unique DS divisions, with an average population per DS of 235,000. This type of LOOCV is a more stringent test of out of sample performance, but one that more accurately approximates the intended use-case of extrapolating poverty into areas where data are not present. While traditional LOOCV assesses out-of-sample performance relative to a large set of single observations, omitted at random, most cases of incomplete survey coverage omit one or more regions within a country.[17] If the data generating process for poverty is geographically heterogeneous, LOOCV at the DS level will give a more accurate assessment of our methodology in practice.

[17] A further complication is uncovered regions are not selected at random and are likely differ from the surveyed regions in unobserved ways. This will contribute to prediction error, since extrapolation requires assuming that the model estimated in the surveyed regions applies to the uncovered region as well.

The choice of covariates for the prediction model is informed by the results from the validation exercise in the previous section. First, given the large differences in the urban and rural results, we estimate separate models for urban and rural areas. Secondly, we use a reduced set of covariates, including into the model only those coefficients that were statistically significant in urban or rural specific models. This resulted in a 13 variable model used to estimate and predict in rural areas excluding the constant term. [18] For urban areas, the model contained 9 variables.[19] Coefficient estimates from the reduced model retain the qualitative significance of the fuller models shown in the poverty validation stage. Tables showing the model estimation over the full sample are presented in appendix tables A1 and A2.

To assess model accuracy we use four different metrics to compare predicted poverty rates against the "true" poverty rates at the GN Division level, which are in fact themselves estimates from the 2011 census data and the 2012/13 HIES.

1. Mean squared error (MSE),. MSE is defined as $\frac{1}{N}\sum_{j=1}^{N}\left(\hat{y_j}^{CV} - y_j\right)^2$, where $\hat{y_j}^{CV}$ indicates the cross-validated predicted GN poverty rate, y_j is the true GN poverty rate, and j indexes each GN.
2. Mean Absolute Error (MAE), which measures the extent to which the predictions deviate in absolute value from the true poverty rate. MAE is defined as $\frac{1}{N}\sum_{j=1}^{N} abs(\hat{y_j}^{CV} - y_j)$.
3. Mean Weighted Absolute Error (MWAE), which is a population-weighted version of Mean Absolute Error. Formally, MWAE is defined as $\frac{1}{N}\sum_{j=1}^{N} w_j * abs(\hat{y_j} - y_j)$ where w_j is the population weight for GN j.
4. The Spearman Rank Correlation Coefficient. Spearman's rank correlation is a non-parametric method that measures only the correlation in the monotonic rank ordering between two variables, in this case the predicted and true poverty rates. [20] This metric indicates whether the GN divisions can successfully be ranked on the basis of their predicted poverty rates. A rank correlation coefficient of 1 would indicate that models using only HRF can create a perfect rank ordering of GN poverty rates using the more stringent DS leave one out cross validation.

How well do models based on satellite indicators extrapolate poverty into areas that were not used as part of the training data? Table 10 shows the results. The MSE varies from 0.0023 for the 10% poverty threshold to 0.0197 for the 40% relative poverty threshold. Since MSE measures squared deviations, it appears that the average deviation is quite small for the lowest threshold, increasing as the poverty threshold increases. Mean weighted absolute error is 2.92 percentage points of poverty for the 10% threshold, 5.57 for the 20% threshold, 8.02 for the 30% threshold, and 10.1 for the 40% threshold. To give these error rates some context, average true relative poverty rates

[18] Object features included in the rural model include log GN Area, fraction of roads that are paved, fraction of roads that are railroad, percent of valid GN area that is built up, percent of shadow pixels covering valid area, fraction of total roofs that are clay, fraction of total roofs that are aluminum, fraction of total roofs are asbestos, log count of roofs, percent of GN area allocated to agriculture, and the texture measures of NDVI, Fourier transform, and surf.

[19] Object features used in the urban model include log GN Area, fraction of valid GN area that is built up (buildings), percent shadow pixels covering valid area (building height proxy), log number of roofs count, log number of cars. Texture features include NDVI, Histogram of oriented gradients, Gabor filter, and standard deviation of the Fourier transform.

[20] Spearman's correlation coefficient between two vectors X and Y with rank ordering x and y is defined as $= 1 - \frac{6\sum {d_i}^2}{n(n^2-1)}$, where d_i is the difference in the rank ordering between x_i and y_i and n is the row dimension of X or Y.

are 5.97, 13.86, 22.3, and 31 at the 10, 20, 30, and 40 percent thresholds, respectively. Consider the GN with the median relative poverty rate at the 40% threshold, an urban GN in the Western province. At the 10% RPL, the true poverty rate is 5.52, and high-resolution features give a cross-validated estimate of 4.44, a difference of -1.14. At the 40% threshold, the true poverty rate is 28.8, and the predicted rate is 24.2 percent. These results suggest that there is substantial prediction error, amounting on average to 33 to 50 percent of average poverty rates, when extrapolating into DS divisions withheld from the estimation.

What about Spearman rank correlation, which measures the monotonic ordering of the predictions rather than the poverty levels themselves? Spearman's ρ is estimated at 0.626 for the 10% RPL, 0.677 for the 20% RPL, 0.702 at the 30% RPL, and 0.718 for the 40% poverty rate. While predicted and true values are not perfect monotonic transformations of one another, the rank ordering between the two is moderate to strongly preserved, especially for higher poverty lines.

A further test of model performance examines the LOOCV prediction residuals to see where in the distribution the predictions differ from the true poverty rates. Figure 14 plots the kernel densities for the true and estimated relative poverty rates at various relative poverty rate thresholds. The 10% relative poverty line predicts appear to underestimate the lowest poverty rate GNs, those at zero or slightly above. At poverty rates between roughly 4 percentage points and 12 percentage points, the estimates slightly over-predict, and from 15 percentage pints poverty onward underestimate the true rate. This pattern seems to be repeated at the 20/30 and 40% thresholds, which show slight over-prediction for the lowest poverty rate GNs and under-prediction for the highest rate GNs. The distributions of the predicted values appear to be compressed relative to the true poverty rate distributions.

There are several possible reasons why the predictions fail to match the distribution of the true poverty rates. First, and perhaps most problematically, the support of the dependent variable, because it is expressed as a rate, is by construction constrained to lie between zero and one. We have not accounted the bounded support of the dependent variable in our baseline methodology. Other methods designed for these purposes such as fractional response models may perform better. Secondly, so far all the estimated models have been linear. Since the true relationship between satellite features and poverty is likely to be non-linear, methods that allow for non-linearity may improve predictive performance. In future work, or future versions of this work, we hope to the address first concern using a fractional response model, and to address the second concern using multiplicative adaptive regression splines (MARS), and support vector machines.

5. Extensions and Robustness

5.1 Prediction Results Using Lasso Estimator

One alternative estimator that may be well suited for this application is a Lasso regression. The Lasso regression is a member of the family of regularized regressions which estimates a regression model with an added constraint that forces parsimony in the coefficient estimate (Tibshirani, 1996). The motivation for the shrinkage estimator is that by adding more parameters to the model, one reduces bias (in-sample error) at the expense of variance (out-of-sample error). The Lasso estimator is defined as

$$(3)\quad \beta_{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \underbrace{\sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{K} x_{ij}\,\beta_j\right)^2}_{OLS\ Sum\ of\ squared\ residuals} + \underbrace{\lambda \sum_{j=1}^{K} |\beta_j|}_{Shrinkage\ factor} \right\}$$

where $\lambda \geq 0$ is the shrinkage parameter that penalizes the absolute values of the coefficients. As we relax the penalization factor -- that is as $\lambda \to 0$ -- $\beta_{lasso} \to \beta_{OLS}$. As $\lambda \to \infty$, variables are more strongly penalized and β_{lasso} converges to the zero vector. Lasso regressions are useful as a variable selection methodology because of the sharp L_1 metric cost function to additional parameters, which creates a de-facto variable selection (Afzal et al, 2016). The value of λ^* is chosen through 10-fold cross validation, comparing RMSE across choice of λ.

In table 12 we present estimated Lasso model coefficients for a model where the dependent variable is average GN income. We present models with two values λ^*: λ_{min} the value of λ which minimizes RMSE across folds, and λ_{1se}, which is defined as λ^* plus one standard error of λ. The latter is chosen because it is even more parsimonious than the former, resulting in fewer variables chosen. The λ_{min} model shows most variables are selected for inclusion, with the exception of % of GN area that is agriculture and % of GN area that is paddy. NDVI, total cars divided by total road length, and % of GN area that is road all show large magnitudes. The large magnitude of the cars variable is surprising in that OLS results showed the cars variable to hold little predictive power in-sample. Since the Lasso estimator focuses on out of sample error it's possible cars has more predictive power out of sample than in-sample.

In the second model using λ_{min}, only 13 variables are selected: urban, fraction roads paved, built-up area, shadow pixels, fraction roofs clay/aluminum, roof count, total cars divided per road length, % of GN agriculture plantation, % of total area paddy/plantation, NDVI and HoG. We estimate even stronger magnitudes on the cars coefficient, 0.6869, and a large magnitude of NDVI, -0.207.

Figure 15 shows the shrinkage path for the variables – which shows the order in which variables are shrunk towards zero. The left-most side of the figure corresponds to an OLS model, that is without any shrinkage penalty applied. As we move to the right, penalization increases. On the furthest right all variables are forced to be zero as the "budget" to spend on variables is essentially zero. The vertical dotted gray lines correspond to the two values of λ, λ_{min} and λ_{1se}. The order in which variables are shrunk to zero can reveal usefulness of coefficients when only few variables can be included. NDVI, Roof count, and cars per road length all remain useful variables even given very high penalization, when the budget to spend on coefficients is very tight.

Table 13 presents Lasso model performance, comparing RMSE, MAE, R-squared and rank order correlation between predicted and true for Lasso models (panel A) and Lasso models using separate urban and rural models (panel B). Note these are all in sample estimates so comparison with table 11 is inexact. Mean absolute errors are relatively low at around 25% of the poverty threshold. The rank order correlation is particularly strong – using separate urban and rural models we estimate rank order correlations between 0.71 and 0.83, meaning using HRSF alone we can achieve a high rank ordering of the poverty and income levels of the small areas considered. Rsquared values are slightly higher than the OLS models, ranging from 0.41 to 0.69.

6. Conclusion

The recent expansion in the quality and availability of HSRI have sparked interest into how they can supplement traditional sources of survey data for a wide variety of applications. Given the prohibitive cost of conducting surveys sufficiently large to provide accurate statistics for small areas, one important application is combining features derived from imagery with survey data on consumption or income to generate estimates of poverty and inequality for small areas. Traditionally, small area poverty estimates have been derived from census data. Indicators derived from satellite data could at a minimum improve the accuracy of both small area estimates and the confidence intervals associated with them (Elbers, Lanjouw, and Leite, 2008). But if indicators derived from imagery explain a sufficiently large portion of variation in welfare, they could potentially serve as viable substitutes for census data, which would greatly increase the temporal frequency of small area poverty estimates. Imagery derived from satellite imagery also can potentially be used to extrapolate estimates into portions of countries that are not currently surveyed due to security concerns.

How well do indicators derived from satellite imagery predict poverty and which indicators are most important? We investigate these questions in 32 randomly selected DS Divisions in Sri Lanka by linking small area estimates of poverty rates to a large set of satellite-based indicators. The data was analyzed at the level of the Gram Niladhari (GN) division, of which there are 13,000 in Sri Lanka and over 1,200 in our sample. The results indicate that the correlation between satellite indicators and predicted welfare is remarkably strong. National models explain 40 to 60 percent in the variation in predicted GN poverty rates, depending on whether the poverty line is set at the 10^{th} percentile or the 40^{th} percentile of the national distribution of predicted per capita consumption. Separate urban and rural models do slightly better, explaining 45 to 63 percent of the variation. Lasso models do slightly better, explaining 0.4 to 0.7 of the variation. Including dummies representing each image further increases the r^2 of the model by about 10 percentage points in all specifications, yielding r^2 statistics that range from 0.55 to 0.74.

In both rural and urban areas, variables measuring building density and built-up area are the strongest predictors of variation in poverty. This includes a measure of built-up area or impervious surfaces, the normalized difference vegetation index (which can be considered an indicator of the absence of buildings), the number of roofs, and the extent of shadows, which reflects the average height of buildings. The statistical significance of both the number of roofs and built-up area in the model suggest that each indicator captures different facets of their communities that each relate to welfare. Because the model controls for the size of the GN division in addition to built-up area, the number of roofs in part captures the average size of buildings. The higher price of land in wealthier could lead to denser building patterns, which would explain the negative correlation between the number of buildings and local estimated poverty rates, after controlling for the number of square kilometers and extent of built-up area in the community. As expected, the extent and lushness of vegetation is positively correlated with poverty in rural areas, and negatively correlated with poverty in urban areas, suggesting that vegetation and gardens are a luxury in urban areas.

The analysis also included several mathematical transformations indicating the “texture” of the image, which have been utilized for optical recognition purposes and other imagery processing

applications. As a whole, these are also strongly correlated with predicted poverty rates. They generally appear to capture contrasts and sharp edges, which characterize wealthier areas, although additional analysis is needed to better understand which features explain variation in welfare, at what scale they should be calculated and what specific characteristics of the built environment they are measuring .

Besides measures of built-up area and building density, variables indicating the nature of roads and roofs also emerge as strong predictors, as does the size of the administrative area. In particular, larger GN areas, which tend to have lower population density, are positively associated with poverty rates. GN area can easily be calculated from any shapefile used to draw a poverty map, and in many cases can be linked to household survey data, but to our knowledge has never been included as a predictor when generating small area estimates.

While these results are very encouraging for the purposes of generating small area estimates, additional analysis suggests caution when extrapolating predictions into geographically adjacent areas. The average absolute error ranges from 3 to 10 percentage points depending on the relative poverty rate, which in percentage terms is one third to half of the average predicted poverty rate. Rank correlations are preserved in many cases, as the Spearman Rank Correlation ranges from 0.63 to 0.72. A likely barrier to extrapolation is geographic heterogeneity in the estimated coefficients. But another factor is differences in when different images were taken. This could impact selected indicators such as car counts, which vary according to the day of the week the imagery was taken. Indeed, the correlation between car counts and poverty in urban areas is negative and statistically significant when controlling for image dummies, but not in the full sample. Measures of NDVI also exhibit considerable seasonal variation which could also confound extrapolation to adjacent areas. This suggests that some indicators may be more biased than others when extrapolating across space, and that the size of the image is an important consideration when considering spatial extrapolation using satellite-based indicators.

These findings raise a host of important questions for further work. The immediate question is whether satellite indicators can substitute for census data when constructing poverty maps in different contexts. Does the strong correlation between satellite-based indicators and predicted welfare extend to measured welfare from an expenditure survey? The predicted poverty rates examined in this analysis are based on census variables that only explain 40 to 50 percent of the variation in measured consumption. The strength of the results reported above, in which satellite-based features explain 60 to 70 percent of the variation in predicted consumption, is very encouraging. However, it is not clear that satellite-based indicators can legitimately substitute for census data to generate small area estimates if, for example, they can only explain a quarter of the variation in measured per capita consumption. In Sri Lanka's case, when the 2016 household survey becomes available, would the country be better served by re-estimating a poverty map for 2016 based on satellite indicators, or by continuing to rely on the existing estimates derived from the 2011 census? Second, it is important to better understand the extent to which these results generalize to different ecological environments, such as Africa, the Middle East, and other parts of Asia. There is no guarantee that the predictive power of building density, NDVI, and spatial features documented above, for example, will hold in all environments.

A second set of questions involves better understanding the reliability of different types of indicators and the trade-offs between the cost of generating indicators and their predictive power.

For example, would the Global Urban Footprint or the Global Human Settlement Layer, which are freely available measures of population density, perform equally well as the measures of built-up area constructed by Orbital Insight? How do the recent estimates of population and built-up area calculated by Facebook compare to other measures? Would NDVI constructed from freely available landsat imagery be a substitute from the NDVI calculated from the very high resolution imagery used in this study? Can convolutional neural networks be trained to identify roof type and/or paved roads, which in many cases for this study were digitized by hand?

A third line of research can explore which characteristics are best suited for extrapolating poverty estimates both across space and time. The ability of imagery to predict into areas currently not covered by household surveys will depend in part on which features are measured with less variation on different days, as well as the stability of coefficients across space. Related to this is the potential for selected satellite-based features to track changes in welfare over time, to help provide more current estimates of poverty. In general, the inevitable increase in the availability of imagery and feature identification algorithms, along with the encouraging results from this study, suggest that satellite imagery will become an increasingly valuable tool to help governments and stakeholders better understand the spatial nature of poverty.

References

Afzal, M., Hersh, J., and Newhouse, D. (2016). "Building a better model: Variable Selection to Predict Poverty in Pakistan and Sri Lanka". Mimeo, World Bank.

H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded Up Robust Features. Lecture Notes in Computer Science, 3951:404–417, 2006

N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition (CVPR), San Diego, CA, 2005, pp. 886-893.

Department of Census and Statistics and World Bank, 2015 "The Spatial Distribution of Poverty in Sri Lanka", available at: http://www.statistics.gov.lk/poverty/SpatialDistributionOfPoverty2012_13.pdf

Donaldson D., and Storeygard A. "Big Grids: Applications of Remote Sensing in Economics", *forthcoming, JEP*.

Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro–level estimation of poverty and inequality. Econometrica, 71(1), 355-364.

Elbers, Chris, Peter F. Lanjouw, and Phillippe G. Leite. "Brazil within Brazil: Testing the poverty map methodology in Minas Gerais." World Bank Policy Research Working Paper Series, Vol (2008).

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., & Davis, E. R. (1997). Mapping city lights with nighttime data from the DMSP Operational Linescan System. Photogrammetric Engineering and Remote Sensing, 63(6), 727-734.

Engstrom, R., Ashcroft, E., Jewell, H., & Rain, D. (2011, April). Using remotely sensed data to map variability in health and wealth indicators in Accra, Ghana. In Urban Remote Sensing Event (JURSE), 2011 Joint (pp. 145-148). IEEE.

Engstrom, R., Sandborn, A., Yu, Q. Burgdorfer, J., Stow, D., Weeks, J., and Graesser, J. (2015) Mapping Slums Using Spatial Features in Accra, Ghana. Joint Urban and Remote Sensing Event Proceedings (JURSE), Lausanne, Switzerland, 10.1109/JURSE.2015.7120494

S. W. Smith, The scientist and engineer's guide to digital signal processing. San Diego, CA: California Technical Publishing, 1997.

Gentle, J. E., Härdle, W. K., & Mori, Y. (Eds.). (2012). Handbook of computational statistics: concepts and methods. Springer Science & Business Media.

J. Graesser, A. Cheriyadat, R. R. Vatsavai, V. Chandola, J. Long, and E. Bright, "Image based characterization of formal and informal neighborhoods in an urban landscape," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 5, no.4, pp. 1164-1176, Jul, 2012.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2009). Measuring economic growth from outer space (No. w15199). National Bureau of Economic Research.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. The American Economic Review, 105(5), 491-495.

Gabor, D. (1946). Theory of Communication. Journal of the Optical Society of America-A, 2 (2),

1455-1471.

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2015). Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life (No. w21778). National Bureau of Economic Research.

Huettner, Frank, and Marco Sunder. "Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values." Electronic Journal of Statistics 6 (2012): 1239-1250.

Israeli, Osnat. "A Shapley-based decomposition of the R-square of a linear regression." The Journal of Economic Inequality 5.2 (2007): 199-212.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

Michalopoulos, S. (2012). The origins of ethnolinguistic diversity. The American economic review, 102(4), 1508.

Mooney, D. F., Larson, J. A., Roberts, R. K., & English, B. C. (2009). Economics of the variable rate technology investment decision for agricultural sprayers. In Southern agricultural economics association annual meeting, Atlanta, Georgia, January.

Mullainathan, S. (2014, August). Bugbears or legitimate threats?:(social) scientists' criticisms of machine learning?. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 4-4). ACM.

Newhouse, David; Suarez Becerra, Pablo; Doan, Dung. 2016. Sri Lanka Poverty and Welfare: Recent Progress and Remaining Challenges. World Bank, Washington, DC. © World Bank. https://www.openknowledge.worldbank.org/handle/10986/23794 License: CC BY 3.0 IGO.

Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. Review of Economics and Statistics, 94(1), 20-36.

M. Pesaresi, A. Gerhardinger, and F. Kayitakire, "A robust built-up area presence index by anisotropic rotation-invariant textural measure," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 1, no. 3, pp. 180-192, Oct, 2008.

Sandborn, A. and Engstrom, R (In Press) Determining the Relationship Between Census Data and Spatial Features Derived From High Resolution Imagery in Accra, Ghana. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS) Special Issue on Urban Remote Sensing

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N., & Dabalen, A. (2015). Data deprivation: another deprivation to end. World Bank Policy Research Working Paper, (7252).

Shorrocks, Anthony F. "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value." Journal of Economic Inequality (2013): 1-28.

Tucker CJ (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8: 127-150.

Watmough, G. R., Atkinson, P. M., Saikia, A., & Hutton, C. W. (2016). Understanding the Evidence Base for Poverty–Environment Relationships using Remotely Sensed Satellite Data: An Example from Assam, India. World Development, 78, 188-203.

Xie, M., Jean, N., Burke, M., Lobell, D., & Ermon, S. (2015). Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping. arXiv preprint arXiv:1510.00098.

Marx, B., Stoker, T. M., & Suri, T. (2013). The Political Economy of Ethnicity and Property Rights in Slums: Evidence from Kenya.

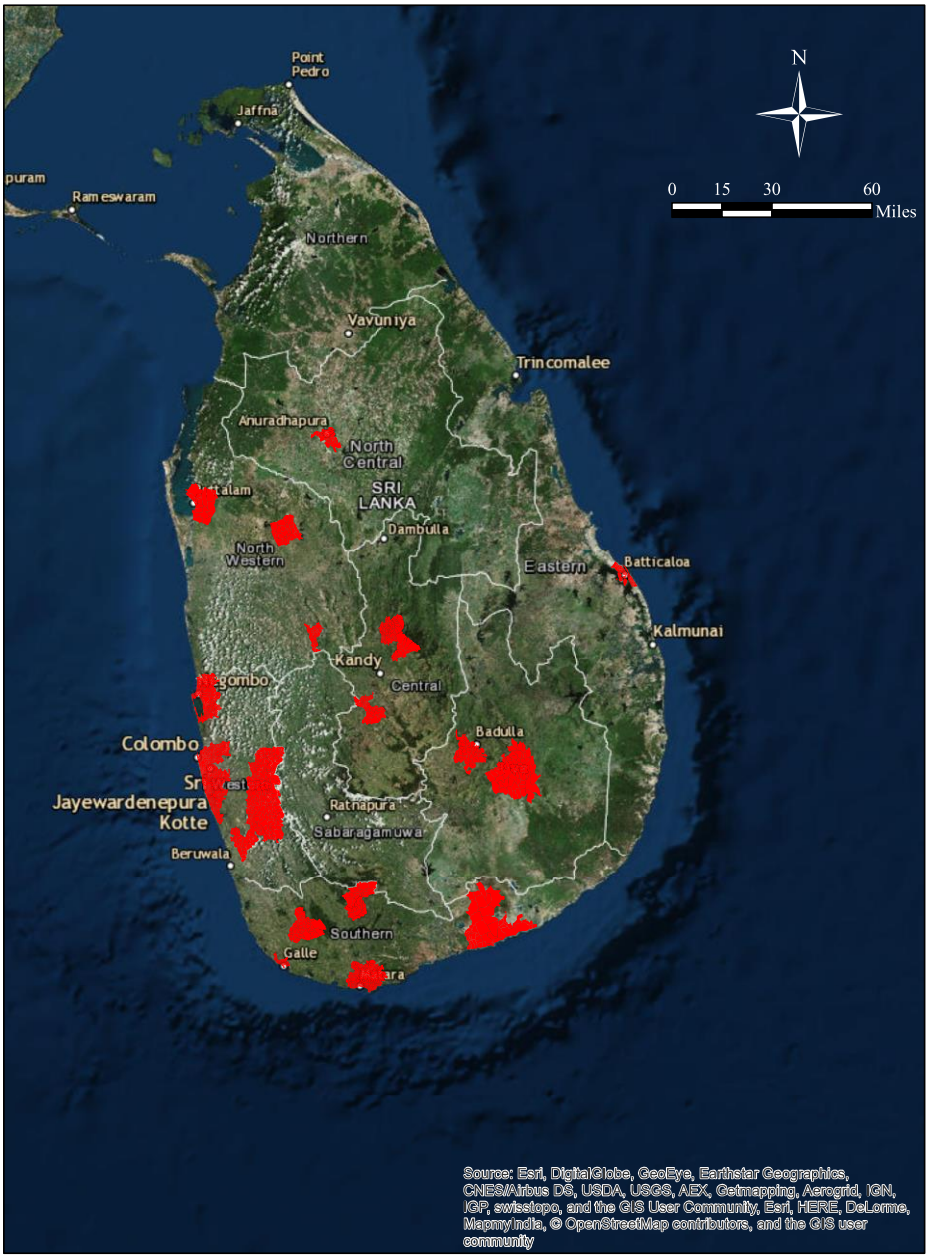
W. P. Yu, G. W. Chu, and M. J. Chung, "A robust line extraction method by unsupervised line clustering," Pattern Recognition, vol. 32, no. 4, pp. 529-546, Apr, 1999.

L. Wang, and D. He, "Texture classification using texture spectrum," Pattern Recognition, vol. 23, no. 8, pp. 905-910, 1990.

Wong, T. H., Mansor, S. B., Mispan, M. R., Ahmad, N., & Sulaiman, W. N. A. (2003, May). Feature extraction based on object oriented analysis. In*Proceedings of ATC 2003 Conference* (Vol. 2021).

Figures

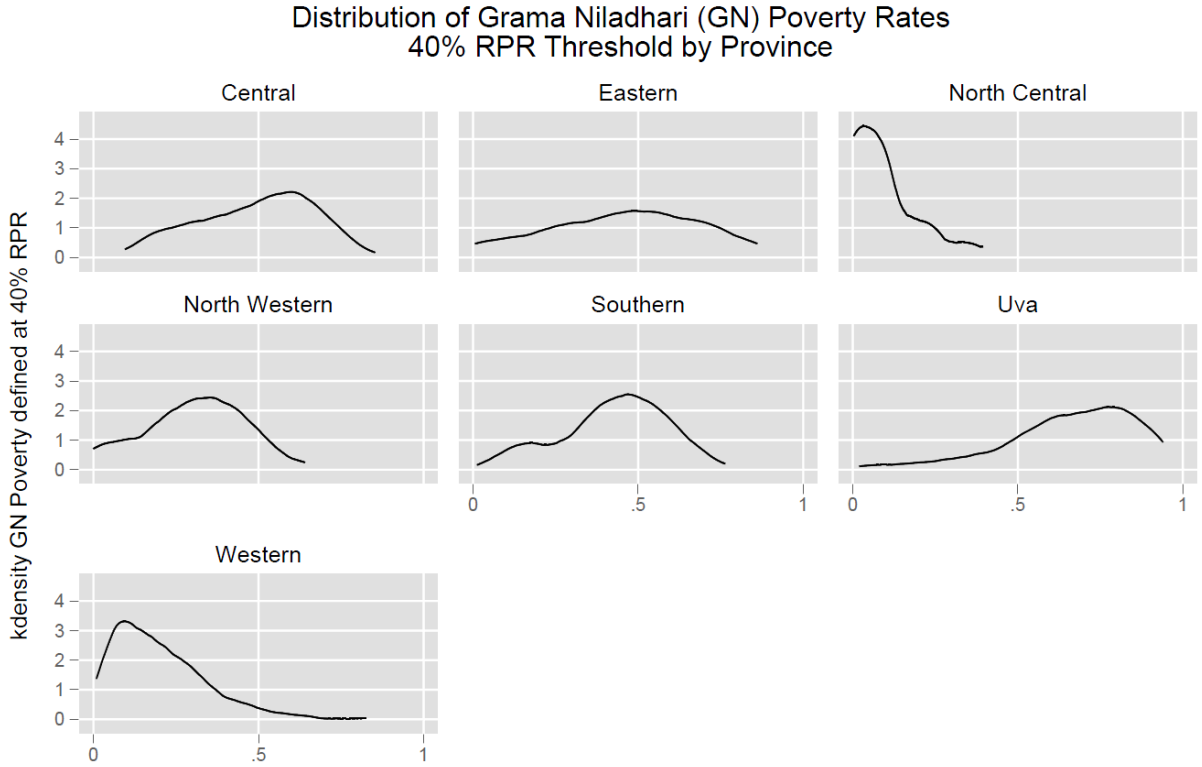
Figure 1: Coverage Area of High Resolution Satellite Imagery



Notes: Highlighted red areas shows Divisional Secretariats (DS) selected for inclusion into our sample.

Figure 2

Figure 3



Note, distribution of GN Division poverty rates is based on sample of 32 DS Divisions and not necessarily representative of the Provinces in general.

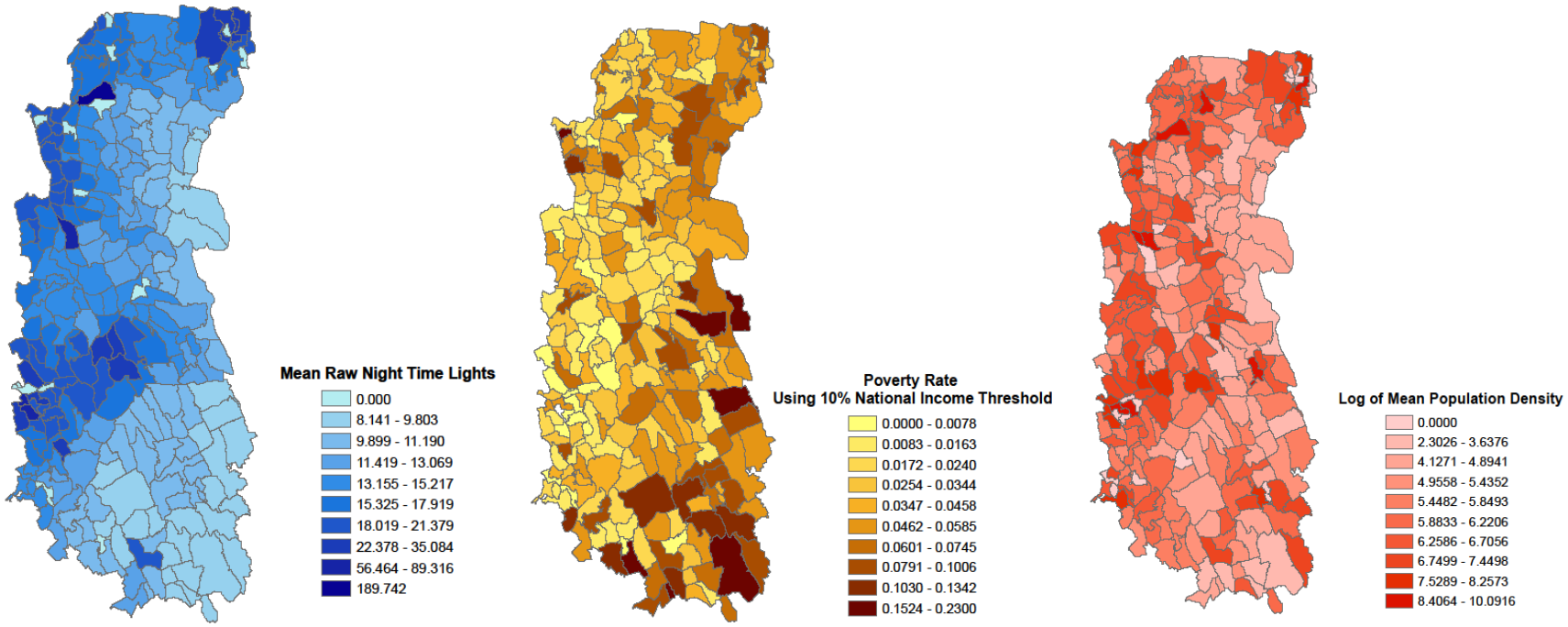


Figure 4: Comparison of Mean Night Time Lights (NTL), Poverty Rate, and Mean Population Density, Seethawaka, Sri Lanka

Figure 5: Example Developed Area (Buildings) Classification

Notes: above image shows raw (top) and classified (bottom) for developed area building classifier from raw satellite imagery. Areas in green show are true positive building classifications, that is areas that are buildings and classified using the developed classifier. Images in red are false positives: erroneously classified areas as buildings.

Figure 6: ROC Curve for Developed Area (Buildings) Classifier

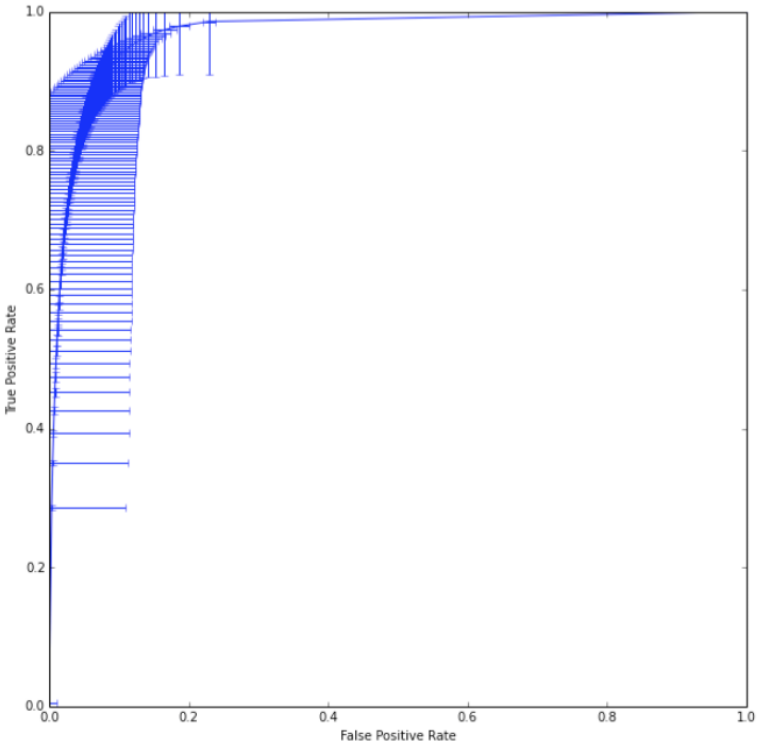


Figure 7: Example Car Classification

Figure 8: Example Roof Type Classification

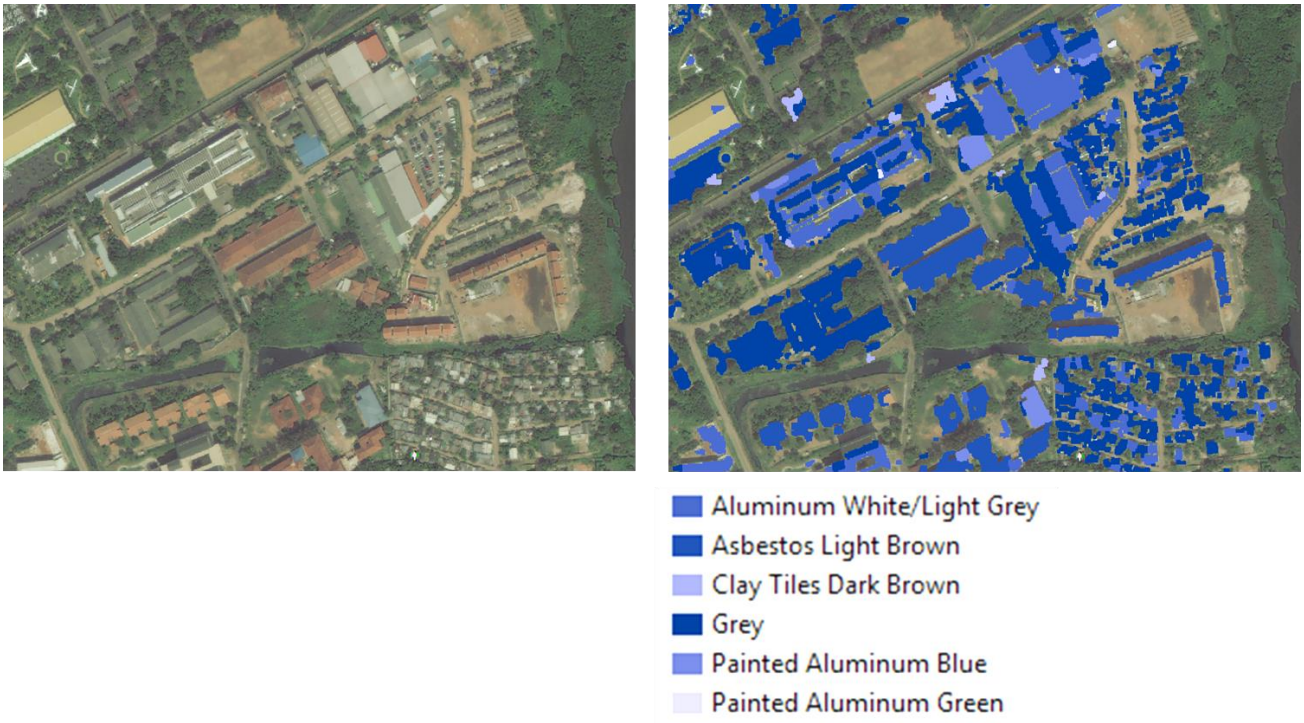


Figure 9: Example Roads and Railroads Classification

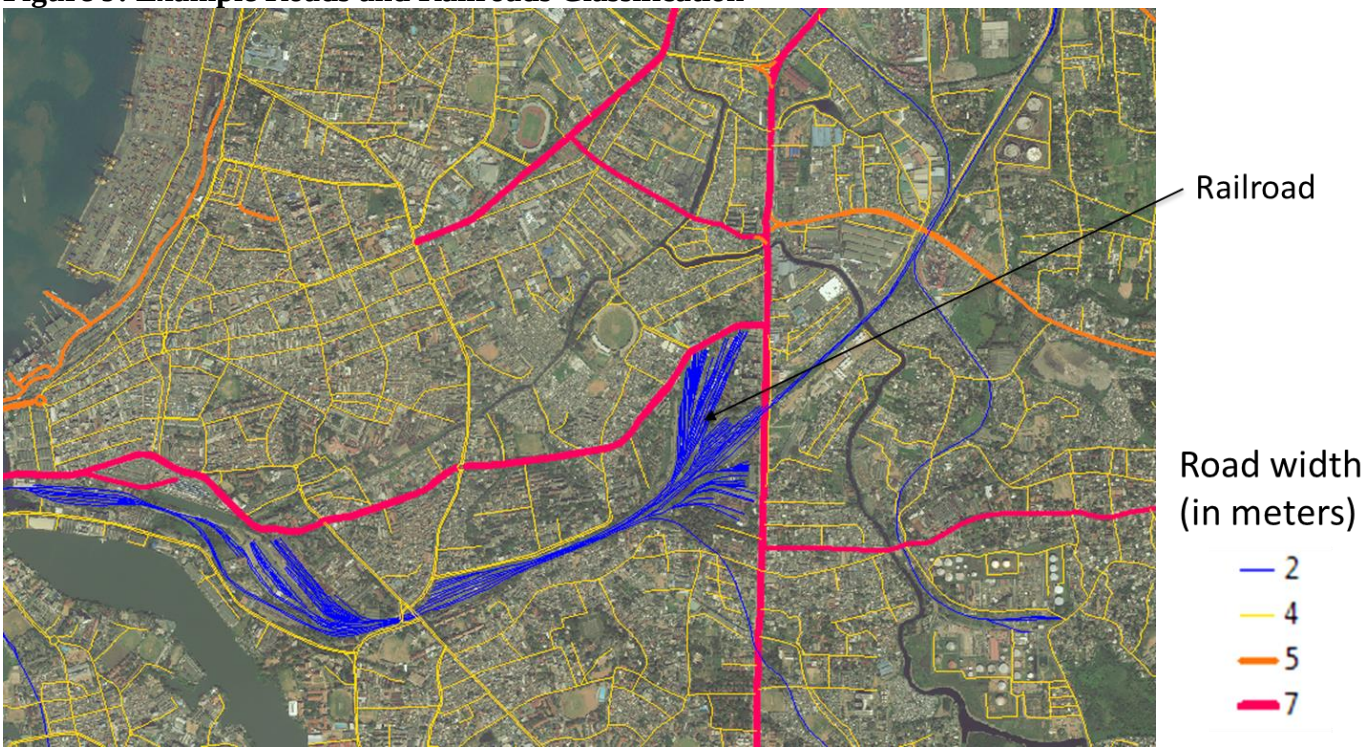


Figure 10: Predicted Versus True GN Poverty Rates Using HRSF Urban/Rural Linear Models, 10% National Income Threshold

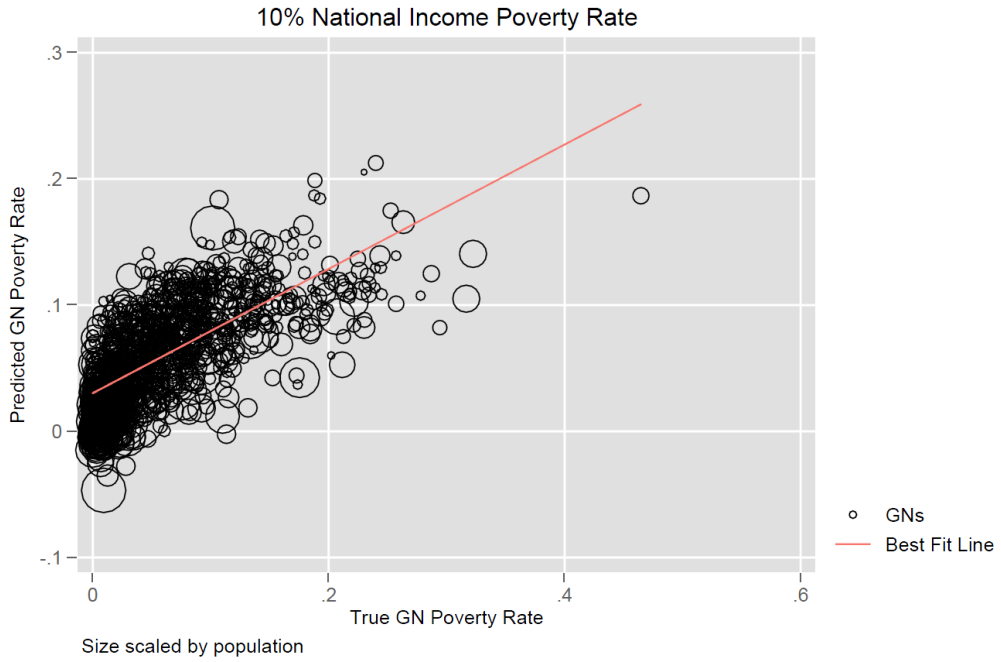


Figure 11: Predicted Versus True GN Poverty Rates Using HRSF Urban/Rural Linear Models, 10% National Income Threshold

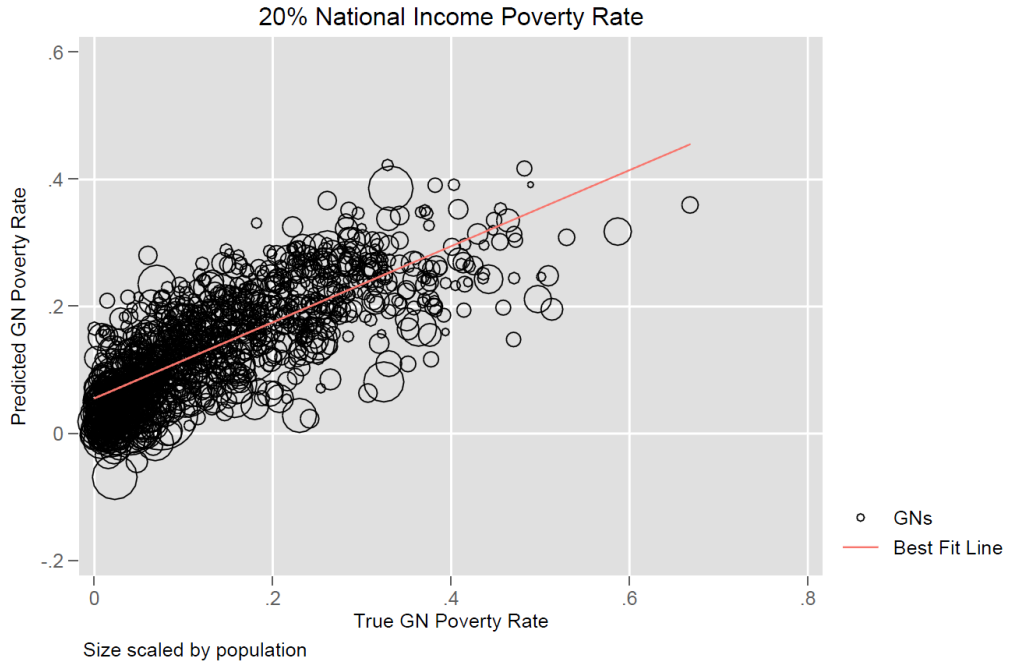


Figure 12: Predicted Versus True GN Poverty Rates Using HRSF Urban/Rural Linear Models, 10% National Income Threshold

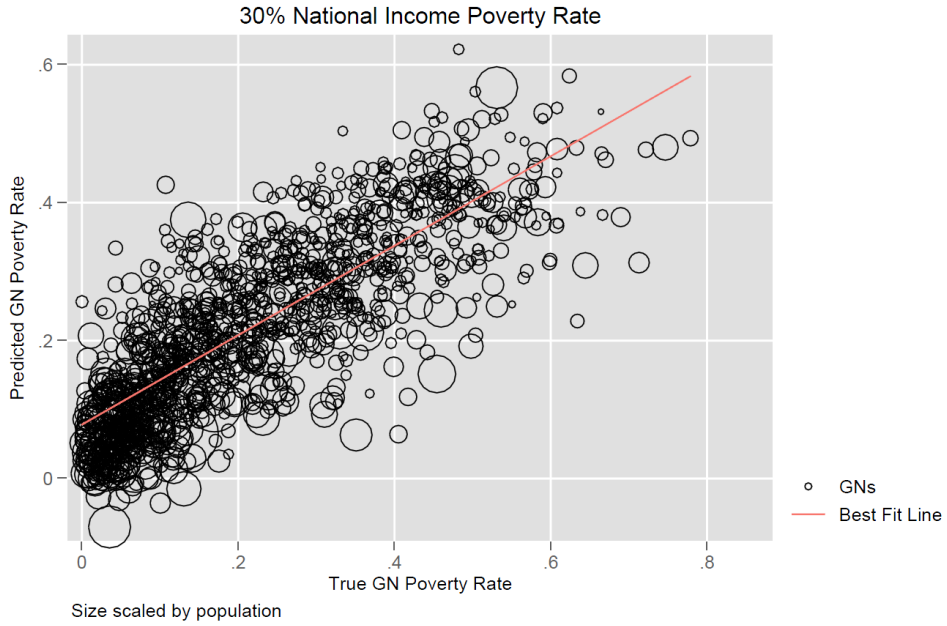


Figure 13: Predicted Versus True GN Poverty Rates Using HRSF Urban/Rural Linear Models, 10% National Income Threshold

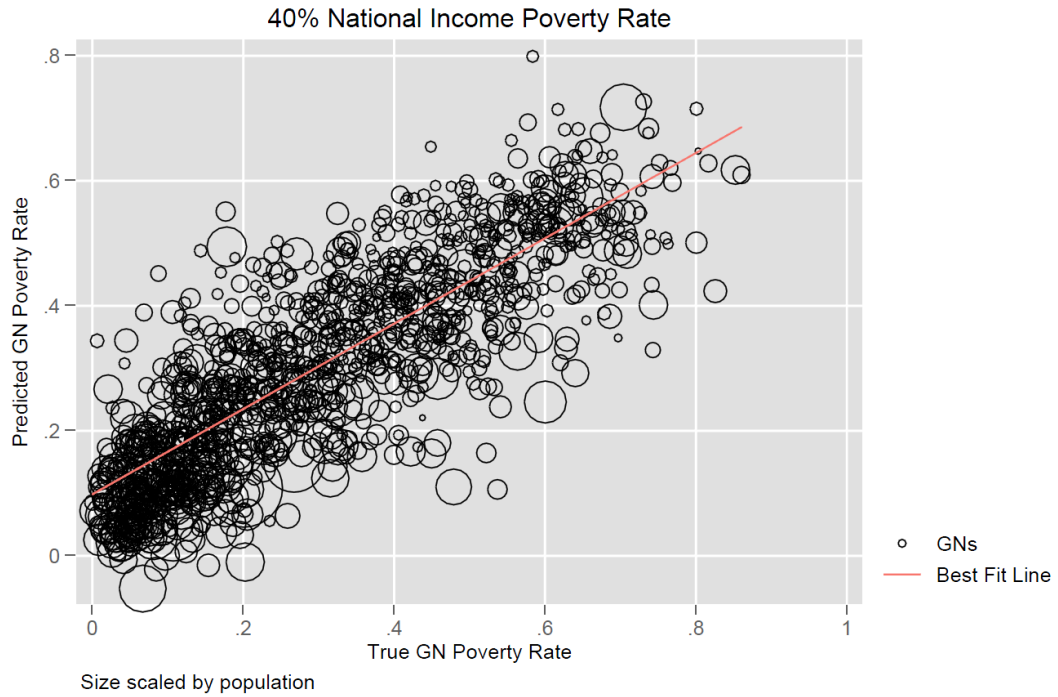


Figure 14: Kernel Density Estimates of Leave-One-Out Cross-Validated Extrapolated Estimated Poverty Rates

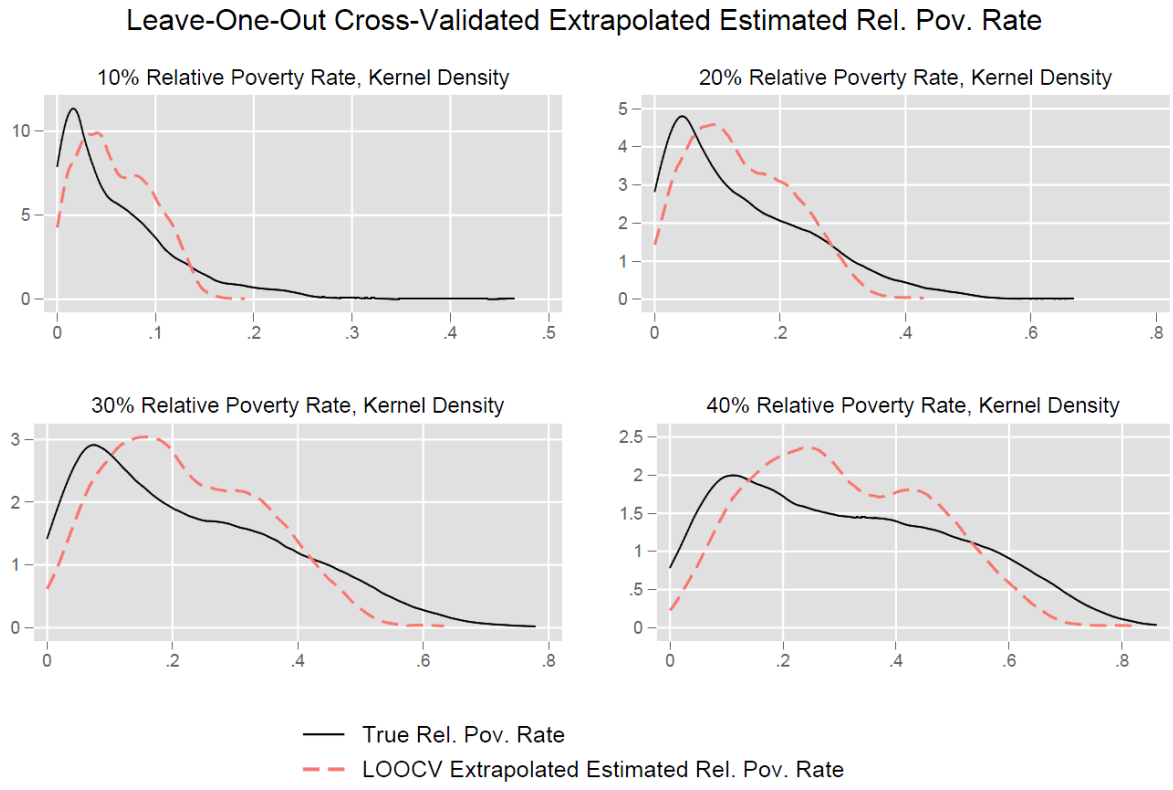
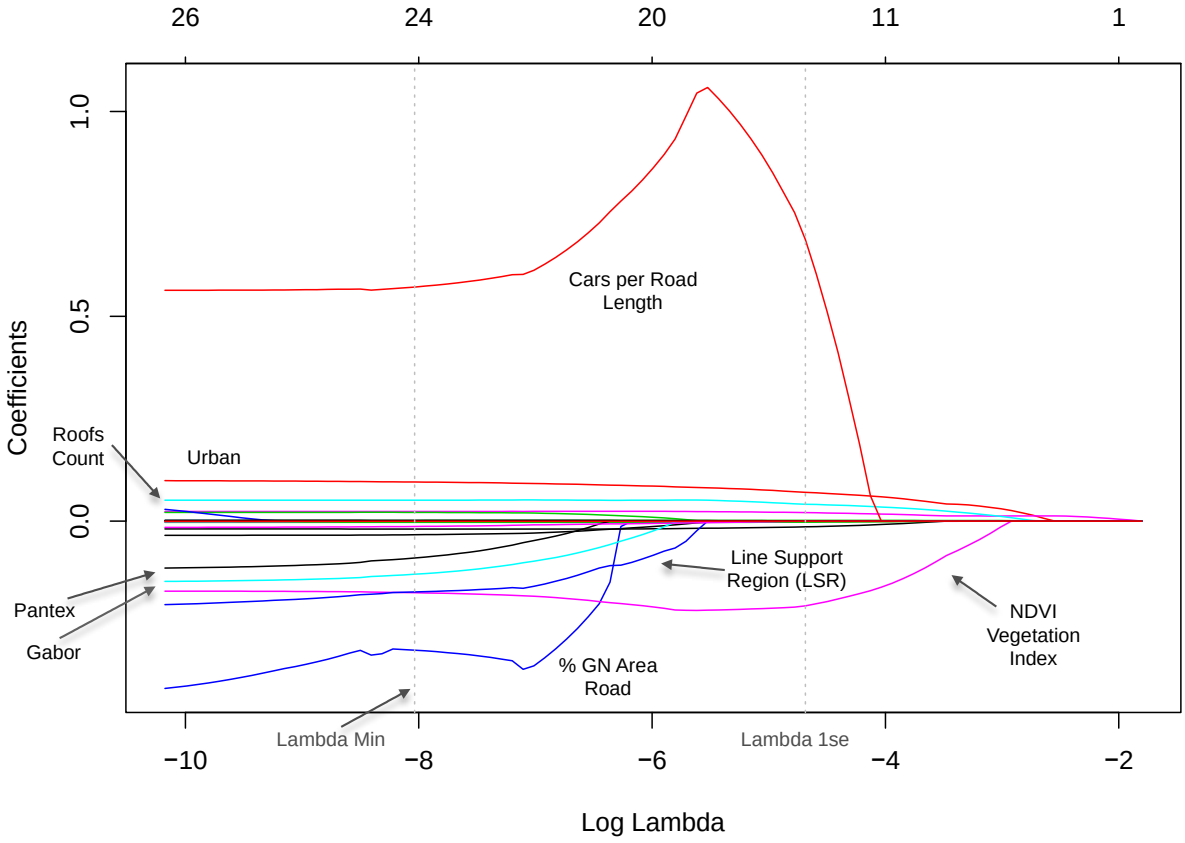


Figure 15: Lasso Model Coefficient Shrinkage Path Plot



Tables

Table 1: Summary Statistics

Variable	mean	sd	min	max
Rel. Pov. Rate at 10% Nat. Cons.	0.0596	0.058	0	0.46
Rel. Pov. Rate at 20% Nat. Cons.	0.138	0.11	0	0.67
Rel. Pov. Rate at 30% Nat. Cons.	0.223	0.16	0	0.78
Rel. Pov. Rate at 40% Nat. Cons.	0.310	0.20	0	0.86
urban	0.280	0.45	0	1
province==[1] Western	0.499	0.50	0	1
province==[2] Central	0.137	0.34	0	1
province==[3] Southern	0.246	0.43	0	1
province==[5] Eastern	0.0273	0.16	0	1
province==[6] North-Western	0.0627	0.24	0	1
province==[7] North-Central	0.0137	0.12	0	1
province==[8] UVA	0.0145	0.12	0	1
% of roads that are paved	33.90	25.9	0	100
% of GN area that is road	0.00664	0.0040	0.000011	0.035
% of roads that are railroad	1.804	6.15	0	70.6
% of GN area covered by built-up area	0.0842	0.098	0	0.57
% of valid GN area that is built up	7.061	6.19	0.052	29.2
% shadow pixels covering valid area	5.965	5.56	0.27	38.2
Fraction of total roofs that are clay	36.27	20.3	0	100
Fraction of total roofs that are aluminum	14.11	7.11	0	71.9
Fraction of total roofs are asbestos	8.195	12.0	0	71.2
log number of roofs count	6.856	0.97	0	8.87
Count of buildings per GN area	0.000669	0.00062	0	0.0047
Total cars divided by total road length	0.00527	0.0095	0.00023	0.17
Total cars divided by total GN Area	0.0000295	0.000045	0.00000014	0.00093
log number of cars	3.431	1.01	0.88	8.30
% of GN area that is agriculture	0.187	0.16	0.00000083	0.94
% of GN agriculture that is paddy	45.91	37.4	0	100
% of GN agriculture that is plantation	53.82	37.4	0	100
% of Total GN area that is paddy	8.918	11.1	0	74.7
% of Total GN area that is plantation	9.782	13.0	0	94.1
Normalized Difference Vegetation Index (NDVI)	0.448	0.20	0	0.86
Pantex (human settlements) mean contrast - scale 8	0.575	0.51	0.0025	3.95
Histogram of oriented gradients (HoG) mean - scale 64	3225.2	1937.4	31.0	10579.8
Local Binary Pattern (moments) skewness - scale 32	272.0	9.56	140.6	285.5
Line support region mean - scale 8	0.0222	0.011	-0.00000020	0.066
Gabor filter mean - scale 64	0.450	0.27	0	1.19
Fourier transform std. dev. - scale 32	22.25	4.13	0.62	29.3
Surf - scale 16	380.5	61.3	13.9	400.0
Observations	1244			

Table 2: Summary Statistics by Urban/Rural

Variable	Rural	Urban	p-value, Difference Between urban and Rural
	mean	mean	
Rel. Pov. Rate at 10% Nat. Cons.	0.0682	0.0374	1.24e-22***
Rel. Pov. Rate at 20% Nat. Cons.	0.160	0.0831	9.81e-35***
Rel. Pov. Rate at 30% Nat. Cons.	0.257	0.133	1.23e-45***
Rel. Pov. Rate at 40% Nat. Cons.	0.358	0.187	2.85e-56***
province==[1] Western	0.446	0.635	2.92e-11***
province==[2] Central	0.166	0.0603	0.0000698***
province==[3] Southern	0.289	0.135	0.00000877***
province==[5] Eastern	0	0.0977	5.29e-25***
province==[6] North-Western	0.0714	0.0402	0.112
province==[7] North-Central	0.00670	0.0316	0.000424***
province==[8] UVA	0.0201	0	3.82e-10***
% of roads that are paved	27.88	49.40	3.30e-47***
% of GN area that is road	0.00535	0.00997	1.07e-101***
% of roads that are railroad	1.152	3.484	2.00e-08***
% of GN area covered by built-up area	0.0483	0.177	3.45e-137***
% of valid GN area that is built up	4.889	12.65	1.18e-122***
% shadow pixels covering valid area	4.253	10.37	1.13e-89***
Fraction of total roofs that are clay	41.61	22.51	5.35e-54***
Fraction of total roofs that are aluminum	14.16	13.99	0.411
Fraction of total roofs are asbestos	7.446	10.12	0.00000102***
log number of roofs count	6.681	7.306	4.06e-29***
Count of buildings per GN area	0.000455	0.00122	4.62e-115***
Total cars divided by total road length	0.00481	0.00646	0.00651**
Total cars divided by total GN Area	0.0000202	0.0000533	4.99e-33***
log number of cars	3.420	3.458	0.611
% of GN area that is agriculture	0.220	0.102	2.55e-34***
% of GN agriculture that is paddy	45.54	46.88	0.230
% of GN agriculture that is plantation	54.15	52.98	0.193
% of Total GN area that is paddy	10.05	5.996	0.000000223***
% of Total GN area that is plantation	11.95	4.213	3.69e-22***
Normalized Difference Vegetation Index (NDVI)	0.468	0.399	3.22e-09***
Pantex (human settlements) mean contrast - scale 8	0.479	0.823	1.52e-33***
Histogram of oriented gradients (HOG) mean - scale 64	2839.4	4218.5	2.91e-35***
Local Binary Pattern (moments) skewness - scale 32	272.4	270.7	0.0138*
Line support region mean - scale 8	0.0202	0.0273	9.77e-22***
Gabor filter mean - scale 64	0.409	0.557	2.67e-19***
Fourier transform std. dev. - scale 32	21.51	24.14	5.20e-26***
Surf - scale 16	376.6	390.5	0.00000422***
Observations	896	348	1244

Table 3: Semi-partial Correlation Between HRSF and GN Poverty Rates, by Poverty Rate Threshold

Variables	10% Rel. Pov. Rate		20% Rel. Pov. Rate		30% Rel. Pov. Rate		40% Rel. Pov. Rate	
	Semi-partial Correlation	p-value	Semi-partial Correlation	p-value	Semi-partial Correlation	p-value	Semi-partial Correlation	p-value
Urban	-0.011	0.6249	-0.0395	0.0524	-0.0597	0.0019***	-0.0807	0***
log GN Area	0.0614	0.0062***	0.0599	0.0034***	0.0557	0.0037***	0.0498	0.007***
Object Classified Variables								
% of roads that are paved	-0.1193	0***	-0.1406	0***	-0.1402	0***	-0.1364	0***
% of GN area that is road	0.0364	0.1043	0.0405	0.047***	0.0382	0.0468***	0.0303	0.1009
% of roads that are railroad	0.0147	0.5122	-0.0061	0.7655	-0.011	0.566	-0.0062	0.7349
% of valid GN area that is built up	-0.1074	0***	-0.1005	0***	-0.1108	0***	-0.1233	0***
% shadow pixels covering valid area	0.0999	0***	0.1215	0***	0.1368	0***	0.1475	0***
Fraction of total roofs that are clay	0.0492	0.0282***	0.0455	0.0255***	0.0466	0.0153***	0.0511	0.0057***
Fraction of total roofs that are aluminum	0.0705	0.0017***	0.0701	0.0006***	0.0672	0.0005***	0.0661	0.0004***
Fraction of total roofs are asbestos	-0.0564	0.012***	-0.0826	0.0001***	-0.0809	0***	-0.0684	0.0002***
log number of roofs count	-0.1268	0***	-0.1531	0***	-0.151	0***	-0.1452	0***
Total cars divided by total road length	-0.0398	0.0759	-0.0387	0.0575	-0.0413	0.0317***	-0.0342	0.0643
Total cars divided by total GN Area	0.0166	0.4575	0.0137	0.5018	0.0145	0.4515	0.012	0.5152
log number of cars	0.012	0.591	0.0175	0.3916	0.0138	0.4716	0.0089	0.6298
% of GN area that is agriculture	-0.0149	0.5072	-0.0109	0.5914	-0.0067	0.7281	-0.0045	0.8075
% of GN agriculture that is paddy	0.0302	0.1785	0.0192	0.3472	0.0071	0.7099	0.0057	0.7577
% of GN agriculture that is plantation	0.0332	0.1391	0.0225	0.2699	0.0108	0.5723	0.0099	0.5919
% of Total GN area that is paddy	0.0017	0.9393	-0.0013	0.9473	-0.0052	0.7862	-0.0066	0.7209
% of Total GN area that is plantation	0.0101	0.6517	0.012	0.5563	0.0094	0.6244	0.0079	0.6683
Texture Classified Variables								
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	0.1527	0***	0.1636	0***	0.1616	0***	0.1546	0***
Pantex (human settlements) mean contrast - scale 8	0.0478	0.0331***	0.0431	0.0346***	0.0458	0.0172***	0.0424	0.0216***
Histogram of oriented gradients (HOG) mean - scale 64	-0.0906	0.0001***	-0.0943	0***	-0.0985	0***	-0.0955	0***
Local Binary Pattern (moments) skewness - scale 32	-0.0373	0.0962	-0.0331	0.104	-0.0259	0.1777	-0.0209	0.2569
Line support region mean - scale 8	-0.0418	0.0622	-0.0185	0.3645	-0.015	0.4351	-0.0084	0.6502
Gabor filter mean - scale 64	0.1125	0***	0.1012	0***	0.0951	0***	0.0887	0***
Fourier transform std. dev. - scale 32	0.0747	0.0009***	0.067	0.001***	0.0563	0.0034***	0.0538	0.0036***
Surf - scale 16	-0.0557	0.013***	-0.0588	0.004***	-0.0482	0.0122***	-0.0435	0.0187***

Table 4: Predicting Poverty Rates using High Res Features, Linear Models

Variable	10% Poverty Rate		20% Poverty Rate		30% Poverty Rate		40% Poverty Rate	
	b	t	b	t	b	t	b	t
urban	-0.0020	[-0.19]	-0.014	[-0.66]	-0.030	[-0.99]	-0.050	[-1.25]
log GN Area	0.010*	[2.22]	0.020*	[2.30]	0.026*	[2.18]	0.029*	[2.02]
% of roads that are paved	-0.00033***	[-3.87]	-0.00077***	[-4.45]	-0.0011***	[-4.47]	-0.0013***	[-4.30]
% of GN area that is road	1.08	[1.03]	2.35	[1.22]	3.13	[1.19]	3.06	[0.98]
% of roads that are railroad	0.00015	[0.38]	-0.00012	[-0.18]	-0.00031	[-0.33]	-0.00022	[-0.18]
% of valid GN area that is built up	-0.0029*	[-2.24]	-0.0052	[-1.97]	-0.0082*	[-2.20]	-0.011*	[-2.33]
% shadow pixels covering valid area	0.0024	[1.53]	0.0057*	[2.03]	0.0090*	[2.51]	0.012**	[2.78]
Fraction of total roofs that are clay	0.00021	[0.92]	0.00038	[0.86]	0.00054	[0.93]	0.00073	[1.04]
Fraction of total roofs that are aluminum	0.00074	[1.92]	0.0014*	[2.07]	0.0020*	[2.02]	0.0024*	[2.02]
Fraction of total roofs are asbestos	-0.00036	[-1.03]	-0.0010	[-1.65]	-0.0014	[-1.77]	-0.0015	[-1.56]
log number of roofs count	-0.012**	[-3.12]	-0.027***	[-4.14]	-0.038***	[-4.16]	-0.045***	[-3.89]
Total cars divided by total road length	-0.39	[-1.84]	-0.74	[-1.78]	-1.11*	[-2.13]	-1.13	[-1.66]
Total cars divided by total GN Area	41.2	[0.91]	66.6	[0.75]	99.2	[0.84]	101.5	[0.66]
log number of cars	0.0018	[0.47]	0.0050	[0.82]	0.0056	[0.66]	0.0044	[0.43]
% of GN area that is agriculture	-0.062	[-1.18]	-0.090	[-0.75]	-0.077	[-0.46]	-0.064	[-0.30]
% of GN agriculture that is paddy	0.00050*	[2.16]	0.00063	[1.09]	0.00033	[0.29]	0.00032	[0.22]
% of GN agriculture that is plantation	0.00055*	[2.58]	0.00073	[1.35]	0.00050	[0.46]	0.00056	[0.39]
% of Total GN area that is paddy	0.000073	[0.12]	-0.00011	[-0.081]	-0.00061	[-0.31]	-0.00096	[-0.39]
% of Total GN area that is plantation	0.00042	[0.98]	0.00099	[0.97]	0.0011	[0.76]	0.0011	[0.62]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	0.062*	[2.01]	0.13*	[2.49]	0.18**	[2.70]	0.22**	[2.80]
Pantex (human settlements) mean contrast - scale 8	0.022	[1.78]	0.038	[1.96]	0.057*	[2.36]	0.065*	[2.25]
Histogram of oriented gradients (HOG) mean - scale 64	-0.000016*	[-2.12]	-0.000033*	[-2.59]	-0.000048**	[-3.13]	-0.000057**	[-3.39]
Local Binary Pattern (moments) skewness - scale 32	-0.00032	[-0.72]	-0.00056	[-0.68]	-0.00061	[-0.56]	-0.00061	[-0.47]
Line support region mean - scale 8	-0.33	[-1.27]	-0.29	[-0.59]	-0.33	[-0.51]	-0.23	[-0.31]
Gabor filter mean - scale 64	0.070	[1.60]	0.12	[1.64]	0.16	[1.72]	0.19	[1.76]
Fourier transform std. dev. - scale 32	0.0034	[1.60]	0.0060	[1.53]	0.0071	[1.33]	0.0083	[1.24]
Surf - scale 16	-0.00013	[-1.44]	-0.00028	[-1.58]	-0.00032	[-1.30]	-0.00036	[-1.13]
Constant	-0.021	[-0.19]	0.048	[0.20]	0.15	[0.42]	0.23	[0.53]
Obs	1244		1244		1244		1244	
R Squared	0.39		0.50		0.55		0.59	

Dependent variable is log of GN Poverty Rate Defined at X% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 5: Predicting Poverty Rates using High Res Features, Linear Models with Image Fixed Effects

Variable	10% Poverty Rate		20% Poverty Rate		30% Poverty Rate		40% Poverty Rate	
	b	t	b	t	b	t	b	t
urban	0.0088	[1.07]	0.015	[0.96]	0.015	[0.67]	0.011	[0.37]
log GN Area	0.013**	[3.24]	0.024**	[3.33]	0.029**	[2.86]	0.029*	[2.45]
% of roads that are paved	-0.00028***	[-3.58]	-0.00065***	[-4.27]	-0.00093***	[-4.52]	-0.0012***	[-4.79]
% of GN area that is road	1.78**	[2.68]	3.11*	[2.41]	3.89*	[2.16]	4.01	[1.84]
% of roads that are railroad	0.00012	[0.39]	-0.00017	[-0.32]	-0.00043	[-0.59]	-0.00050	[-0.55]
% of valid GN area that is built up	-0.0028**	[-3.30]	-0.0052**	[-3.32]	-0.0087***	[-4.27]	-0.012***	[-4.89]
% shadow pixels covering valid area	-0.00017	[-0.095]	0.00063	[0.21]	0.0021	[0.58]	0.0039	[0.91]
Fraction of total roofs that are clay	0.000037	[0.23]	0.00014	[0.50]	0.00027	[0.70]	0.00037	[0.84]
Fraction of total roofs that are aluminum	0.00052	[1.82]	0.00097	[1.99]	0.0013*	[2.02]	0.0015*	[2.06]
Fraction of total roofs are asbestos	0.00014	[0.30]	0.00033	[0.40]	0.00055	[0.56]	0.00074	[0.69]
log number of roofs count	-0.010**	[-2.82]	-0.023***	[-3.81]	-0.028***	[-3.62]	-0.031**	[-3.31]
Total cars divided by total road length	0.039	[0.29]	0.16	[0.56]	0.11	[0.32]	0.25	[0.52]
Total cars divided by total GN Area	-23.4	[-0.73]	-51.6	[-0.82]	-44.7	[-0.55]	-45.0	[-0.41]
log number of cars	0.0029	[0.98]	0.0058	[1.12]	0.0057	[0.73]	0.0039	[0.39]
% of GN area that is agriculture	-0.0071	[-0.22]	-0.076	[-1.75]	-0.14*	[-2.07]	-0.20*	[-2.17]
% of GN agriculture that is paddy	0.00017	[1.24]	0.00035	[1.76]	0.00018	[0.41]	0.00026	[0.41]
% of GN agriculture that is plantation	0.00012	[0.95]	0.00027	[1.54]	0.00012	[0.29]	0.00020	[0.34]
% of Total GN area that is paddy	-0.00021	[-0.46]	0.00020	[0.31]	0.00061	[0.74]	0.00087	[0.83]
% of Total GN area that is plantation	0.00018	[0.70]	0.0011***	[3.64]	0.0018**	[2.84]	0.0023*	[2.59]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	-0.047	[-1.80]	-0.064	[-1.35]	-0.085	[-1.29]	-0.094	[-1.19]
Pantex (human settlements) mean contrast - scale 8	-0.0067	[-0.59]	-0.022	[-1.05]	-0.021	[-0.79]	-0.018	[-0.58]
Histogram of oriented gradients (HOG) mean - scale 64	0.0000053	[1.00]	0.0000087	[0.88]	0.0000065	[0.50]	0.0000015	[0.097]
Local Binary Pattern (moments) skewness - scale 32	-0.00016	[-0.83]	-0.00036	[-0.97]	-0.00040	[-0.81]	-0.00040	[-0.65]
Line support region mean - scale 8	-0.20	[-1.07]	-0.062	[-0.20]	0.015	[0.037]	0.18	[0.41]
Gabor filter mean - scale 64	0.015	[0.61]	0.043	[0.92]	0.063	[0.93]	0.099	[1.31]
Fourier transform std. dev. - scale 32	-0.00035	[-0.26]	-0.0014	[-0.55]	-0.0032	[-0.92]	-0.0044	[-1.07]
Surf - scale 16	0.00012	[1.84]	0.00023	[1.79]	0.00039*	[2.10]	0.00049*	[2.11]
Constant	-0.036	[-0.53]	0.0012	[0.0093]	0.098	[0.50]	0.21	[0.82]
Sat. Imagery Timing FEs	Yes		Yes		Yes		Yes	
Obs	1244		1244		1244		1244	
R Squared	0.54		0.64		0.69		0.72	

Dependent variable is log of GN Poverty Rate Defined at X% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 6: HRF Urban/Rural Poverty Models (10% Nat. Consumption)

Variable	Urban		Rural		Urban		Rural	
	b	t	b	t	b	T	b	t
log GN Area	0.039*	[2.71]	0.013*	[2.64]	0.052***	[4.21]	0.015**	[3.10]
% of roads that are paved	-0.000096	[-0.62]	-0.00044***	[-4.95]	-0.00013	[-1.48]	-0.00040***	[-4.83]
% of GN area that is road	1.63	[1.11]	0.75	[0.64]	2.53	[1.75]	1.44	[1.57]
% of roads that are railroad	0.00052	[0.98]	-0.00039	[-1.94]	0.00047	[1.25]	-0.00023	[-1.26]
% of valid GN area that is built up	-0.0016	[-1.57]	-0.0010	[-0.82]	-0.00099	[-1.46]	-0.0015	[-1.28]
% shadow pixels covering valid area	-0.00050	[-0.19]	0.0028*	[2.65]	-0.0035	[-1.23]	0.0016	[1.62]
Fraction of total roofs that are clay	-0.00081*	[-2.27]	0.00042*	[2.35]	-0.00055	[-1.13]	0.00031*	[2.09]
Fraction of total roofs that are aluminum	-0.00065	[-1.00]	0.0012**	[2.99]	0.00046	[0.70]	0.00082*	[2.39]
Fraction of total roofs are asbestos	-0.00035	[-0.47]	-0.00026	[-1.46]	0.00030	[0.32]	0.000098	[0.39]
log number of roofs count	-0.046***	[-4.12]	-0.0090**	[-2.70]	-0.039**	[-3.73]	-0.0094*	[-2.42]
Total cars divided by total road length	-0.31	[-0.69]	-0.18	[-1.05]	0.0053	[0.013]	0.12	[0.81]
Total cars divided by total GN Area	86.9	[0.85]	-21.2	[-0.24]	52.1	[0.62]	-47.2	[-0.83]
log number of cars	-0.0059	[-1.10]	0.00027	[0.055]	-0.013**	[-3.13]	0.0022	[0.58]
% of GN area that is agriculture	0.15	[0.75]	-0.065*	[-2.15]	-0.37	[-1.87]	0.0079	[0.19]
% of GN agriculture that is paddy	0.00042	[1.05]	0.00012	[0.51]	-0.00100	[-1.73]	-0.00013	[-1.16]
% of GN agriculture that is plantation	0.00029	[0.77]	0.00033	[1.41]	-0.0011	[-1.74]	-0.000044	[-0.35]
% of Total GN area that is paddy	-0.0025	[-1.21]	0.00042	[1.10]	0.0031	[1.50]	-0.00014	[-0.28]
% of Total GN area that is plantation	-0.0017	[-0.83]	0.00033	[0.96]	0.0029	[1.75]	-0.000047	[-0.13]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	-0.15*	[-2.64]	0.085**	[2.89]	-0.28*	[-2.12]	-0.0084	[-0.43]
Pantex (human settlements) mean contrast - scale 8	0.031	[1.75]	-0.024	[-1.33]	-0.013	[-0.90]	-0.0066	[-0.40]
Histogram of oriented gradients (HOG) mean - scale 64	-0.000032*	[-2.33]	-0.0000060	[-0.87]	0.0000071	[0.90]	0.0000021	[0.35]
Local Binary Pattern (moments) skewness - scale 32	0.00078	[1.27]	-0.00031	[-0.86]	0.0014	[1.31]	-0.00024	[-1.31]
Line support region mean - scale 8	-0.14	[-0.45]	-0.64*	[-2.69]	-0.15	[-0.46]	-0.43*	[-2.09]
Gabor filter mean - scale 64	0.16*	[2.20]	0.053	[1.33]	-0.0054	[-0.16]	0.024	[0.69]
Fourier transform std. dev. - scale 32	0.0080*	[2.39]	0.0042**	[3.07]	0.00018	[0.15]	0.00065	[0.48]
Surf - scale 16	0.000066	[0.46]	-0.00021**	[-3.45]	0.00057**	[3.42]	0.000038	[0.70]
Constant	-0.50*	[-2.25]	-0.066	[-0.68]	-0.70*	[-2.24]	-0.038	[-0.60]
Sat. Imagery Timing FEs	No		No		Yes		Yes	
Obs	348		896		348		896	
R Squared	0.52		0.44		0.64		0.56	

Dependent variable is log of GN Poverty Rate Defined at 10% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 7: HRF Urban/Rural Poverty Models (20% Nat. Consumption)

Variable	Urban		Rural		Urban		Rural	
	b	t	b	t	b	t	b	t
log GN Area	0.066*	[2.74]	0.026**	[3.13]	0.087***	[4.12]	0.030***	[3.78]
% of roads that are paved	-0.00029	[-1.05]	-0.00099***	[-5.27]	-0.00028	[-1.50]	-0.00086***	[-5.41]
% of GN area that is road	2.67	[1.05]	2.41	[1.09]	4.37	[1.70]	2.37	[1.23]
% of roads that are railroad	0.00047	[0.56]	-0.00095*	[-2.09]	0.00040	[0.68]	-0.00065	[-1.64]
% of valid GN area that is built up	-0.0036*	[-2.14]	-0.0014	[-0.54]	-0.0029*	[-2.22]	-0.0012	[-0.55]
% shadow pixels covering valid area	0.000078	[0.017]	0.0065**	[2.90]	-0.0046	[-0.95]	0.0032	[1.64]
Fraction of total roofs that are clay	-0.0013	[-2.06]	0.00077*	[2.15]	-0.00076	[-0.81]	0.00067*	[2.49]
Fraction of total roofs that are aluminum	-0.0011	[-1.03]	0.0023**	[3.27]	0.00070	[0.66]	0.0016**	[2.77]
Fraction of total roofs are asbestos	-0.00033	[-0.26]	-0.00096*	[-2.69]	0.00052	[0.29]	0.00033	[0.72]
log number of roofs count	-0.085***	[-4.28]	-0.024***	[-4.50]	-0.072**	[-3.41]	-0.023***	[-4.15]
Total cars divided by total road length	-0.23	[-0.26]	-0.30	[-0.81]	0.25	[0.32]	0.29	[1.00]
Total cars divided by total GN Area	78.0	[0.41]	-94.3	[-0.61]	35.7	[0.21]	-71.3	[-0.59]
log number of cars	-0.0088	[-0.94]	0.0025	[0.32]	-0.021**	[-2.96]	0.0039	[0.58]
% of GN area that is agriculture	0.54	[1.32]	-0.13	[-1.66]	-0.45	[-1.12]	-0.054	[-0.90]
% of GN agriculture that is paddy	0.00069	[0.93]	-0.000017	[-0.023]	-0.0015	[-1.41]	-0.000066	[-0.20]
% of GN agriculture that is plantation	0.00051	[0.71]	0.00041	[0.57]	-0.0017	[-1.41]	0.00011	[0.30]
% of Total GN area that is paddy	-0.0069	[-1.61]	0.00093	[1.00]	0.0036	[0.85]	0.00036	[0.47]
% of Total GN area that is plantation	-0.0049	[-1.20]	0.0011	[1.46]	0.0035	[1.06]	0.00072	[1.48]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	-0.29**	[-3.01]	0.17***	[3.70]	-0.45	[-2.02]	0.0069	[0.19]
Pantex (human settlements) mean contrast - scale 8	0.043	[1.51]	-0.040	[-1.07]	-0.027	[-1.14]	-0.0097	[-0.28]
Histogram of oriented gradients (HOG) mean - scale 64	-0.000050*	[-2.49]	-0.000019	[-1.40]	0.000013	[0.93]	-0.0000031	[-0.27]
Local Binary Pattern (moments) skewness - scale 32	0.0016	[1.45]	-0.00049	[-0.71]	0.0020	[1.20]	-0.00047	[-1.44]
Line support region mean - scale 8	0.029	[0.063]	-0.74	[-1.69]	-0.039	[-0.078]	-0.45	[-1.13]
Gabor filter mean - scale 64	0.26*	[2.54]	0.100	[1.38]	-0.0026	[-0.044]	0.070	[1.10]
Fourier transform std. dev. - scale 32	0.013*	[2.26]	0.0084**	[3.26]	-0.00016	[-0.073]	0.0010	[0.40]
Surf - scale 16	0.00019	[0.76]	-0.00047***	[-4.09]	0.00096**	[3.52]	0.000065	[0.62]
Constant	-0.84*	[-2.44]	-0.058	[-0.28]	-1.04*	[-2.20]	-0.057	[-0.55]
Sat. Imagery Timing FEs	No		No		Yes		Yes	
Obs	348		896		348		896	
R Squared	0.58		0.55		0.70		0.66	

Dependent variable is log of GN Poverty Rate Defined at 20% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 8: HRF Urban/Rural Poverty Models (30% Nat. Consumption)

Variable	Urban		Rural		Urban		Rural	
	b	t	b	t	b	t	b	t
log GN Area	0.079*	[2.54]	0.032**	[2.83]	0.10**	[3.64]	0.037**	[3.35]
% of roads that are paved	-0.00049	[-1.32]	-0.0014***	[-5.46]	-0.00043	[-1.44]	-0.0012***	[-5.57]
% of GN area that is road	2.99	[0.92]	3.51	[1.22]	5.56	[1.69]	2.50	[0.95]
% of roads that are railroad	0.00057	[0.51]	-0.0015*	[-2.46]	0.00041	[0.53]	-0.0011*	[-2.20]
% of valid GN area that is built up	-0.0061*	[-2.67]	-0.0035	[-0.95]	-0.0056**	[-2.91]	-0.0039	[-1.49]
% shadow pixels covering valid area	0.0015	[0.28]	0.0098**	[3.20]	-0.0041	[-0.67]	0.0045	[1.88]
Fraction of total roofs that are clay	-0.0015	[-1.64]	0.0010*	[2.21]	-0.00074	[-0.52]	0.00090*	[2.47]
Fraction of total roofs that are aluminum	-0.0012	[-0.78]	0.0030**	[3.39]	0.0011	[0.82]	0.0021**	[2.80]
Fraction of total roofs are asbestos	-0.00026	[-0.18]	-0.0013*	[-2.42]	0.00039	[0.17]	0.00068	[1.05]
log number of roofs count	-0.10***	[-3.98]	-0.033***	[-4.29]	-0.082*	[-2.78]	-0.028***	[-3.88]
Total cars divided by total road length	0.041	[0.034]	-0.57	[-1.21]	0.65	[0.59]	0.21	[0.58]
Total cars divided by total GN Area	47.5	[0.18]	-143.7	[-0.67]	1.25	[0.0055]	-45.0	[-0.26]
log number of cars	-0.016	[-1.26]	0.0041	[0.38]	-0.030**	[-3.26]	0.0041	[0.40]
% of GN area that is agriculture	0.82	[1.46]	-0.15	[-1.55]	-0.58	[-1.08]	-0.11	[-1.63]
% of GN agriculture that is paddy	0.0010	[1.08]	-0.00062	[-0.43]	-0.0014	[-0.98]	-0.00041	[-0.64]
% of GN agriculture that is plantation	0.00084	[0.86]	0.0000039	[0.0027]	-0.0016	[-1.01]	-0.00017	[-0.25]
% of Total GN area that is paddy	-0.010	[-1.72]	0.00094	[0.80]	0.0045	[0.81]	0.00065	[0.71]
% of Total GN area that is plantation	-0.0070	[-1.24]	0.0013	[1.57]	0.0049	[1.10]	0.0012*	[2.21]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	-0.39**	[-3.17]	0.24***	[3.91]	-0.55	[-1.91]	0.0059	[0.11]
Pantex (human settlements) mean contrast - scale 8	0.055	[1.46]	-0.048	[-0.81]	-0.026	[-0.77]	0.0079	[0.19]
Histogram of oriented gradients (HOG) mean - scale 64	-0.000063*	[-2.44]	-0.000033	[-1.66]	0.000010	[0.52]	-0.000011	[-0.70]
Local Binary Pattern (moments) skewness - scale 32	0.0020	[1.36]	-0.00056	[-0.57]	0.0021	[1.01]	-0.00058	[-1.32]
Line support region mean - scale 8	0.048	[0.077]	-0.94	[-1.63]	-0.062	[-0.087]	-0.57	[-1.14]
Gabor filter mean - scale 64	0.32*	[2.62]	0.14	[1.51]	0.026	[0.29]	0.090	[0.98]
Fourier transform std. dev. - scale 32	0.015	[2.05]	0.011**	[3.19]	-0.0011	[-0.37]	0.00085	[0.27]
Surf - scale 16	0.00037	[1.12]	-0.00063***	[-3.71]	0.0013**	[3.50]	0.00014	[1.03]
Constant	-1.03*	[-2.16]	0.051	[0.16]	-1.14	[-1.93]	0.040	[0.23]
Sat. Imagery Timing FEs	No		No		Yes		Yes	
Obs	348		896		348		896	
R Squared	0.59		0.60		0.70		0.70	

Dependent variable is log of GN Poverty Rate Defined at 30% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 9: HRF Urban/Rural Poverty Models (40% Nat. Consumption)

Variable	Urban		Rural		Urban		Rural	
	b	t	b	t	b	t	b	t
log GN Area	0.089*	[2.58]	0.037**	[2.79]	0.11**	[3.21]	0.040**	[3.11]
% of roads that are paved	-0.00077	[-1.67]	-0.0016***	[-5.63]	-0.00068	[-1.64]	-0.0014***	[-6.14]
% of GN area that is road	3.00	[0.80]	3.61	[1.18]	6.34	[1.73]	1.84	[0.64]
% of roads that are railroad	0.00063	[0.50]	-0.0015*	[-2.02]	0.00029	[0.33]	-0.0013	[-1.97]
% of valid GN area that is built up	-0.0090**	[-3.21]	-0.0053	[-1.08]	-0.0087**	[-3.44]	-0.0064*	[-2.08]
% shadow pixels covering valid area	0.0037	[0.58]	0.012**	[3.34]	-0.0029	[-0.40]	0.0054	[2.01]
Fraction of total roofs that are clay	-0.0017	[-1.44]	0.0014*	[2.43]	-0.00087	[-0.46]	0.0011*	[2.66]
Fraction of total roofs that are aluminum	-0.0011	[-0.60]	0.0036**	[3.52]	0.0014	[0.86]	0.0024**	[2.96]
Fraction of total roofs are asbestos	-0.00025	[-0.15]	-0.0013	[-1.74]	0.00013	[0.048]	0.0011	[1.46]
log number of roofs count	-0.12***	[-3.85]	-0.039***	[-3.95]	-0.090*	[-2.47]	-0.030***	[-3.55]
Total cars divided by total road length	0.38	[0.24]	-0.48	[-0.75]	1.10	[0.79]	0.41	[0.87]
Total cars divided by total GN Area	15.5	[0.046]	-127.6	[-0.52]	-33.1	[-0.12]	23.2	[0.12]
log number of cars	-0.025	[-1.64]	0.0024	[0.19]	-0.041**	[-3.38]	0.0010	[0.085]
% of GN area that is agriculture	1.22	[1.67]	-0.16	[-1.36]	-0.47	[-0.72]	-0.15	[-2.02]
% of GN agriculture that is paddy	0.0010	[0.92]	-0.00073	[-0.38]	-0.0016	[-0.86]	-0.00040	[-0.44]
% of GN agriculture that is plantation	0.00077	[0.68]	0.000100	[0.051]	-0.0018	[-0.89]	-0.000076	[-0.082]
% of Total GN area that is paddy	-0.015	[-1.93]	0.0011	[0.75]	0.0032	[0.47]	0.00086	[0.83]
% of Total GN area that is plantation	-0.010	[-1.43]	0.0014	[1.40]	0.0041	[0.74]	0.0015*	[2.36]
Normalized Difference Vegetation Index (NDVI) local mean - Scale 64	-0.46**	[-3.21]	0.27***	[4.01]	-0.64	[-1.87]	0.011	[0.17]
Pantex (human settlements) mean contrast - scale 8	0.056	[1.20]	-0.071	[-0.88]	-0.029	[-0.67]	0.027	[0.52]
Histogram of oriented gradients (HOG) mean - scale 64	-0.000066*	[-2.19]	-0.000042	[-1.67]	0.000013	[0.49]	-0.000023	[-1.29]
Local Binary Pattern (moments) skewness - scale 32	0.0025	[1.32]	-0.00063	[-0.54]	0.0029	[1.10]	-0.00069	[-1.29]
Line support region mean - scale 8	0.28	[0.41]	-1.04	[-1.53]	0.16	[0.19]	-0.70	[-1.22]
Gabor filter mean - scale 64	0.34*	[2.67]	0.19	[1.71]	0.032	[0.28]	0.15	[1.47]
Fourier transform std. dev. - scale 32	0.015	[1.73]	0.015**	[3.25]	-0.0046	[-1.16]	0.0018	[0.54]
Surf - scale 16	0.00057	[1.50]	-0.00078**	[-3.53]	0.0016**	[3.53]	0.00013	[0.86]
Constant	-1.15	[-1.96]	0.12	[0.29]	-1.35	[-1.88]	0.16	[0.71]
Sat. Imagery Timing FEs	No		No		Yes		Yes	
Obs	348		896		348		896	
R Squared	0.60		0.63		0.70		0.74	

Dependent variable is log of GN Poverty Rate Defined at 40% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Table 10: Shapely Decomposition of Share of Variance Explained of National Model R^2 by HRSF Subgroup

	Average predicted per capita consumption in GN	10% poverty rate	20% poverty rate	30% poverty rate	40% poverty rate
Urban	8.6	2.5	3.6	4.6	5.8
Log of GN Area	6.7	7.9	7.7	7.5	7.3
Road variables	11.7	12.1	12.2	11.8	11.5
Building density variables	36.4	37.8	36.9	36.4	36.1
Of which: Built-up area	18.6	12.8	11.7	12.4	13.3
Log Number of roofs	8.8	9.4	10.8	10.4	10.0
Shadow	5.2	3.8	4.0	4.3	4.6
NDVI	3.8	11.8	10.4	9.3	8.3
Roofs	9.3	7.6	7.5	7.5	7.8
Cars	5.2	4.5	4.4	4.3	4.2
Agricultural land variables	4.9	6.3	7.3	7.5	7.3
Texture variables	17.3	21.4	20.4	20.5	20.0
Total r^2	0.64	0.39	0.50	0.55	0.59

Notes: Road variables include percent of roads that are paved, percent of GN area that is road, and percent of GN that is railroad. Roof variables include fraction of total roofs that are clay, aluminum, and asbestos. Cars includes total road length, cars per sq km, and log number of cars. Agricultural land variables include percent of GN area that is agriculture,, paddy, and plantation, and percent of agriculture that is paddy . Texture variables include Pantex scale 8, Histogram of oriented gradients scale 64, Local Binary Pattern Skewness scale 32, Line Support Region scale 8, Gabor filter scale 64, Fourier transform st. dev scale 32, and Speeded Up Robust Features scale 16.

Table 11: Poverty Extrapolation Prediction Performance, Using Leave-One-Out Cross-Validation (LOOCV)

Metric	**10% Poverty Rate**	**20% Poverty Rate**	**30% Poverty Rate**	**40% Poverty Rate**
Mean Squared Error (MSE)	0.0023	0.0075	0.0136	0.0197
Mean Absolute Error (MAE)	0.0339	0.0638	0.0887	0.1082
Mean Weighted Absolute Error (MWAE)	0.0292	0.0557	0.0802	0.1014
Spearman Rank Correlation Between Predicted and True Poverty Rates	0.6258	0.6765	0.7023	0.7176
Avg. True Rel. Pov Rate	0.0597	0.1386	0.2228	0.3105

Table presents out of sample estimates for of extrapolated poverty rate prediction into withheld Divisional Secretariat (DS) administrative districts. We estimate 55 models, each time withholding one of 55 DS units to reserve as an out of sample test. Using the relationship between poverty and satellite variables estimated using the 54 variables as training data, we predict into the withheld DS – so called "leave-one-out" cross-validation (LOOCV). Separate urban and rural models are used for inference and prediction.

Table 12: Lasso Model Coefficients, Lambda Plus 1 Standard Error and Minimum Lambda Model

	Lambda 1se Model Coefficient	Min Lambda Model Coefficient
Intercept	8.895056	9.350
Log of GN Area	0	-.033419
Urban	.070097	.095442
% of roads that are paved	.001726	.001563
% of GN area that is road	0	-.315284
% of roads that are railroad	0	-.001206
% of valid GN area that is built up	.020641	.023529
% shadow pixels covering valid area	-.014037	-.019426
Fraction of total roofs that are clay	-.001117	-.000956
Fraction of total roofs that are aluminum	-.00270	-.002976
Fraction of total roofs are asbestos	0	.001131
log number of roofs count	.041118	.051034
Total cars divided by total road length	.686951	.571406
log number of cars	0	.021239
% of GN area that is agriculture	0	0
% of GN agriculture that is paddy	0	0
% of GN agriculture that is plantation	-.000570	-.000290
% of Total GN area that is paddy	.000443	.001906
% of Total GN area that is plantation	-.000211	-.000326
Normalized Difference Vegetation Index (NDVI)	-.2070	-.175064
Pantex (human settlements) mean contrast	0	-.090118
Histogram of oriented gradients (HOG) mean	.000014	.000065
Local Binary Pattern (moments) skewness	0	.000151
Line support region mean	0	-.173226
Gabor filter mean	0	-.129829
Fourier transform std. dev.	0	-.013185
Surf	0	.000161
Sat. Imagery Timing FEs	No	No
N	1,244	1,244

Dependent variable is average GN income.

Table 13: Lasso Model Performance Comparison

Panel A: Lasso Model Performance

Dependent Variable	Root Mean Squared Error (RMSE)	Mean Abs Error (MAE)	Rank Order Correlation, Predicted Versus True	Rsq
10% Poverty Rate	0.0482	0.0354	0.6424	0.3119
20% Poverty Rate	0.0840	0.0639	0.7173	0.4584
30% Poverty Rate	0.1107	0.0859	0.7501	0.5270
40% Poverty Rate	0.1314	0.1043	0.7660	0.5609
Average GN Consumption	0.1692	0.1297	0.7913	0.6133

Panel B: Lasso Model Performance, Separate Urban and Rural Models

Dependent Variable	Root Mean Squared Error (RMSE)	Mean Abs Error (MAE)	Rank Order Correlation Predicted Versus True	Rsq
10% Poverty Rate	0.0444	0.0314	0.7174	0.4162
20% Poverty Rate	0.0763	0.0570	0.7797	0.5529
30% Poverty Rate	0.1015	0.0778	0.7989	0.6022
40% Poverty Rate	0.1172	0.0915	0.8169	0.6506
Average GN Consumption	0.1508	0.1158	0.8324	0.6929

Satellite imagery fixed effects not included.

Appendix Tables

Rural models used to Extrapolate Poverty using Cross-Validation

Variable	10% Rel. Poverty Rate		20% Rel. Poverty Rate		30% Rel. Poverty Rate		40% Rel. Poverty Rate	
	b	t	b	t	b	t	b	t
log GN Area	0.014***	[5.42]	0.033***	[6.87]	0.043***	[6.73]	0.049***	[6.55]
% of roads that are paved	-0.00042***	[-5.55]	-0.00094***	[-6.87]	-0.0013***	[-6.99]	-0.0015***	[-6.80]
% of roads that are railroad	-0.00040	[-1.38]	-0.00099	[-1.87]	-0.0015*	[-2.17]	-0.0016	[-1.96]
% of valid GN area that is built up	-0.0020**	[-2.64]	-0.0041**	[-2.97]	-0.0079***	[-4.25]	-0.011***	[-5.21]
% shadow pixels covering valid area	0.0046***	[6.60]	0.011***	[8.79]	0.017***	[9.99]	0.021***	[10.8]
Fraction of total roofs that are clay	0.00037***	[3.74]	0.00068***	[3.77]	0.00092***	[3.82]	0.0013***	[4.53]
Fraction of total roofs aluminum	0.0011***	[3.96]	0.0023***	[4.46]	0.0030***	[4.43]	0.0036***	[4.56]
Fraction of total roofs are asbestos	-0.00043**	[-2.68]	-0.0014***	[-4.75]	-0.0020***	[-5.05]	-0.0020***	[-4.42]
log number of roofs count	-0.0098***	[-4.63]	-0.025***	[-6.36]	-0.033***	[-6.36]	-0.040***	[-6.57]
% of GN area that is agriculture	-0.017	[-1.56]	-0.0060	[-0.31]	-0.0031	[-0.12]	0.00016	[0.0053]
Normalized Difference Vegetation Index	0.11***	[14.2]	0.22***	[16.0]	0.32***	[16.8]	0.38***	[17.0]
Fourier transform std. dev. - scale 32	0.00096	[1.17]	0.0013	[0.89]	0.00030	[0.15]	0.00063	[0.27]
Surf - scale 16	-0.00013*	[-2.39]	-0.00025*	[-2.57]	-0.00028*	[-2.10]	-0.00033*	[-2.11]
Constant	-0.12**	[-3.08]	-0.25***	[-3.47]	-0.28**	[-2.84]	-0.26*	[-2.27]
Obs	897		897		897		897	
R Squared	0.40		0.49		0.52		0.54	

Dependent variable is log of GN Poverty Rate Defined at X% of national consumption
t-stats in brackets, clustered at Divisional Secretariat (DS) level
* p<0.05, ** p<0.01, *** p<0.001

Urban models used to Extrapolate Poverty using Cross-Validation

Variable	Urban		Rural		Urban		Rural	
	b	t	b	t	b	t	b	t
log GN Area	0.029***	[4.34]	0.055***	[4.95]	0.070***	[4.70]	0.082***	[4.49]
% of valid GN area that is built up	-0.0021**	[-2.74]	-0.0046***	[-3.62]	-0.0076***	[-4.46]	-0.011***	[-5.37]
% shadow pixels covering valid area	-0.00018	[-0.31]	0.00048	[0.51]	0.0020	[1.60]	0.0040*	[2.56]
log number of roofs count	-0.038***	[-6.39]	-0.070***	[-7.01]	-0.088***	[-6.54]	-0.098***	[-5.98]
log number of cars	-0.0049	[-1.41]	-0.012*	[-2.06]	-0.022**	[-2.83]	-0.033***	[-3.50]
Normalized Difference Vegetation Index (NDVI)	-0.14***	[-6.41]	-0.27***	[-7.13]	-0.35***	[-7.01]	-0.41***	[-6.74]
Histogram of oriented gradients (HOG) mean	-0.0000086*	[-2.48]	-0.000018**	[-3.12]	-0.000024**	[-3.04]	-0.000025**	[-2.69]
Gabor filter mean - scale 64	0.091***	[3.65]	0.17***	[4.09]	0.22***	[3.90]	0.25***	[3.62]
Fourier transform std. dev. - scale 32	0.0074***	[7.07]	0.014***	[7.97]	0.018***	[7.83]	0.021***	[7.37]
Constant	-0.18*	[-2.28]	-0.34*	[-2.52]	-0.39*	[-2.18]	-0.42	[-1.89]
Obs	348		348		348		348	
R Squared	0.44		0.52		0.53		0.54	

Dependent variable is log of GN Poverty Rate Defined at X% of national consumption

t-stats in brackets, clustered at Divisional Secretariat (DS) level

* p<0.05, ** p<0.01, *** p<0.001

Appendix

Sampled divisional secretariats include Ambagamuwa, Ambalantota, Ambanpola, Bandaragama, Biyagama, Bulathsinhala, Colombo, Dehiwala, Devinuwara, Dodangoda, Doluwa, Dompe, Galle Four Gravets, Hali Ela, Hambantota, Homagama, Horana, Ingiriya, Kaduwela, Kalutara, Kamburupitiya, Katana, Kattankudy, Kelaniya, Kesbewa, Kirinda Puhulwella, Kolonnawa, Kotapola, Kothmale, Kurunegala, Madurawala, Maharagama, Malimbada, Manmunai North, Matara Four Gravets, Moratuwa, Nagoda, Negambo, Nuwara Eliya, Nuwaragam Palatha East, Padukka, Panadura, Panvila, Puttalam, Rathmalana, Rattota, Seethawaka, Sri Jayawardanapura Kotte, Thihagoda, Thimbirigasyaya, Tissamaharama, dapalatha, Udunuwara, Ukuwela, and Uva Paranagama

Figure A1: Graphical Depiction of Relative Size of Sri Lankan Administrative Division

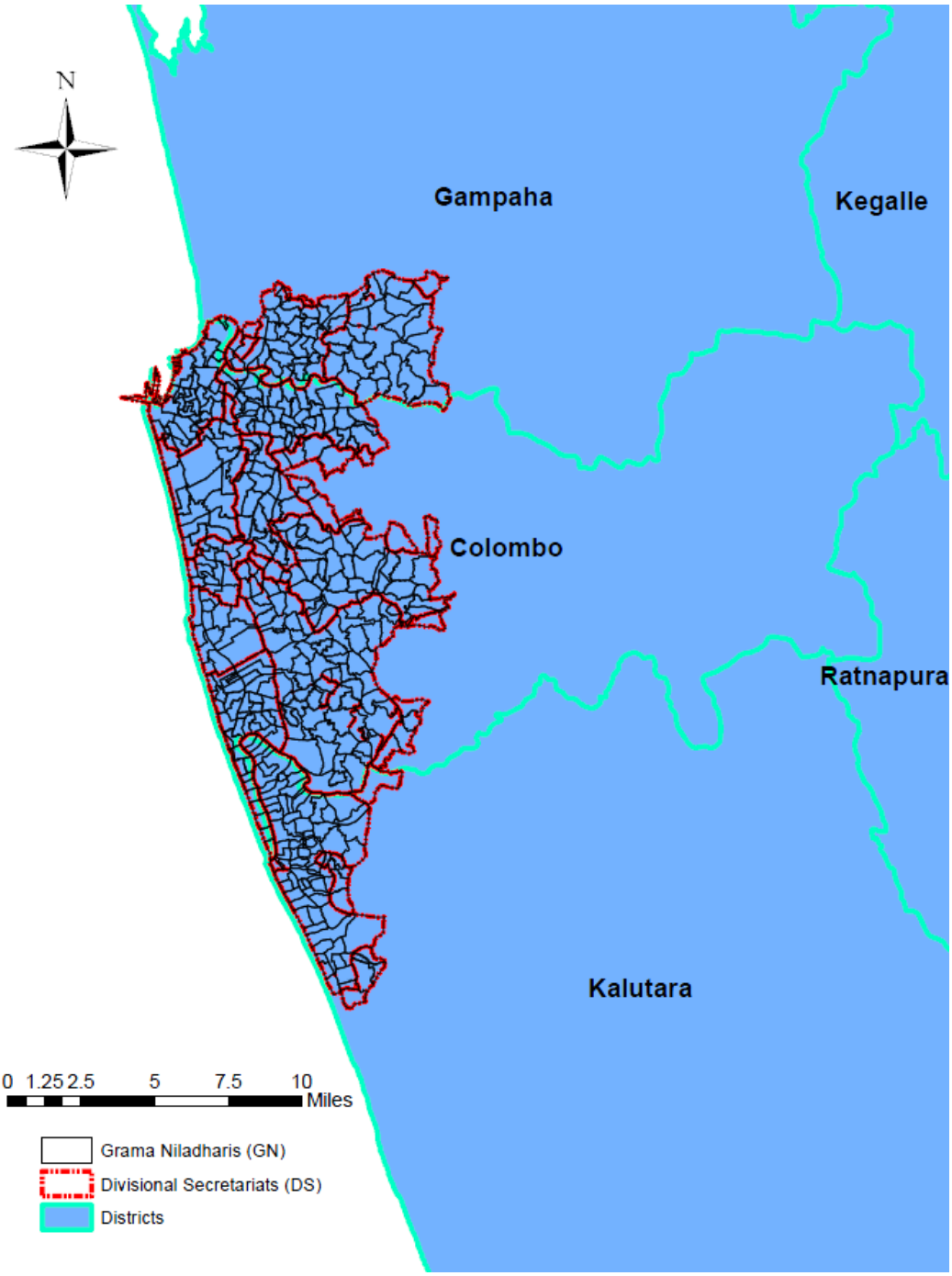


Figure A2: Count of Administrative Divisions of Sri Lanka by Type