INTERNATIONAL MONETARY FUND

# Harnessing Satellite Data to Improve Social Assistance Targeting in the Eastern Caribbean

Sophia Chen, Ryu Matsuura, Flavien Moreau, Joana Pereira

**WP/24/84**



**2024**
APR

WORKING PAPER

 WP/24/84

**IMF Working Paper**
Western Hemisphere Department

**Harnessing Satellite Data to Improve Social Assistance Targeting in the Eastern Caribbean**

**Prepared by Sophia Chen, Ryu Matsuura, Flavien Moreau, Joana Pereira***

Authorized for distribution by Christoph Duenwald
April 2024



**ABSTRACT:** Prioritizing populations most in need of social assistance is an important policy decision. In the Eastern Caribbean, social assistance targeting is constrained by limited data and the need for rapid support in times of large economic and natural disaster shocks. We leverage recent advances in machine learning and satellite imagery processing to propose an implementable strategy in the face of these constraints. We show that local well-being can be predicted with high accuracy in the Eastern Caribbean region using satellite data and that such predictions can be used to improve targeting by reducing aggregation bias, better allocating resources across areas, and proxying for information difficult to verify.



| | |
|---|---|
| JEL Classification Numbers: | C63, H12, I38, O12 |
| Keywords: | Social assistance targeting; satellite data; machine learning; Eastern Caribbean; Small Island Developing States. |
| Author's E-Mail Address: | ychen2@imf.org, ryumatsuura@u.northwestern.edu, fmoreau@imf.org, jpereira@imf.org |



* The authors would like to thank Emine Boz, Alejandro Guerson, and participants at the 7th West Indies Economic Conference, Eastern Caribbean Central Bank Economic Seminar, IMF Western Hemisphere Department Seminar, IMF Western Hemisphere Department Caribbean 1 Division Seminar for helpful comments and suggestions.

WORKING PAPERS

# Harnessing Satellite Data to Improve Social Assistance Targeting in the Eastern Caribbean

Prepared by Sophia Chen, Ryu Matsuura, Flavien Moreau, Joana Pereira[1]

# Contents

# 1. Introduction

Social assistance programs play a crucial role in protecting vulnerable populations from hardship and mitigate poverty. Yet, in most countries, resources for social assistance remain limited. Prioritizing among various segments of the population in need when designing social assistance programs is therefore one of the most important decisions facing policymakers.

The task of prioritizing social assistance—or targeting—is challenging in the Eastern Caribbean where a key problem is the limited availability of data. Most countries lack a centralized data system on income or related household information, and the presence of a substantial informal sector makes income difficult to verify.[1] Other sources of data on livelihoods and human well-being are also lacking, partly because of the relatively high costs of conducting regular surveys in small economies and limited administrative capacities. The lack of data in turn constrains efforts to characterize for whom and where social assistance is needed the most.

This problem becomes even more acute in times of immense economic and natural disaster shocks, which often requires rapid support for those most severely impacted by the shock, including the recent COVID-19 pandemic and cost-of-living crisis. The region is also highly exposed to natural disasters, ranging from hurricanes to earthquakes and volcanic eruptions, which often cause significant damage to economic activities and livelihoods.

These challenges are common among developing economies, and especially among small island developing states (SIDS). Targeting in this context most commonly relies on estimations of poverty and vulnerability based on observable and proxy data. The estimates may be based on location characteristics (e.g., households in impoverished areas or areas hit by natural disasters), demographic characteristics (e.g., older people or households with young children), and other individual and household characteristics (e.g., education and household assets). In SIDS, national and local authorities may possess tacit knowledge that could help them target vulnerable households. For instance, they may know that households living in a specific informal settlement are vulnerable. Developing more formal and data-driven approaches can help assess the accuracy of such informal forms of targeting and inform budget allocation across the country.

In countries with extensive, periodic, and representative household surveys, estimates can be derived from econometric models with survey data. However, in countries without such surveys, estimates are often extrapolated from past or peer-country experience. Such practices give rise to inaccurate estimates and—consequently—potentially large targeting errors. Conventional

[1] Jensen (2022) estimate that the tax system includes only 1–10 percent of the economically active population in low-income countries and 30–50 percent in middle-income countries, compared to 90–95 percent in high-income countries.

targeting systems are estimated to incur more than 25 percent exclusion errors (reflecting under-coverage) and inclusion errors (reflecting leakage) when compared to ground truth data from surveys (McBride et al. 2016).[2] Errors for countries lacking survey data are unknown and could possibly be much greater.

In this paper, we address these challenges by leveraging recent advances in machine learning and satellite imagery. We hypothesize that targeting can be improved in the Eastern Caribbean through granular geospatial data from satellite imagery. High resolution satellite data have become available in recent years which can usefully be analyzed with machine learning techniques to inform policy decisions. Such data could provide invaluable local information, filling the gap of traditional data. We ask two specific questions. First, can satellite imagery be used to accurately measure local well-being in the Eastern Caribbean? Second, can these measures be used to improve targeting and if so, through what channel?

We follow the approach of a recent study that provides the first global estimates of the United Nations Human Development Index (HDI) for second-level administrative units (e.g., counties, parishes, municipalities; Sherman et al. 2023). We extend the seminal work of Sherman et al. (2023) in three ways. First, we provide comprehensive parish- and (1 square kilometer) grid-level HDI predictions for the Eastern Caribbean countries and territories filling the coverage gap in Sherman et al. (2023).[3] Second, we validate the performance of the prediction model by comparing model estimations to ground truth microdata from the region. Third, we quantify the extent to which these estimates can improve social assistance targeting in various settings of geographic and hybrid targeting designs.

The HDI summarizes three key dimensions of well-being: the health of the population, human capital, and standard of living. A more comprehensive measure of well-being than income or wealth alone, it has been widely used by policymakers and academics in areas of development assistance and the global allocation of resources. The use of HDI in the context of social assistance aligns well with an increasingly popular approach to assessing vulnerability from a multidimensional perspective rather than from income alone. To facilitate the use of HDI in national social assistance policies, the development of HDI measures at the subnational level is necessary. To estimate subnational HDI, Sherman et al. (2023) develop a downscaling technique in which the model is trained on province-level data and predictions are made on a finer geographic level. Building on this approach, we evaluate a large set of prediction models

[2] In Latin America, exclusion errors appear larger than inclusion errors in the context of conditional cash transfers (Ibarrarán et al. 2017).

[3] Sherman et al. (2023) do not include the following Eastern Caribbean countries or territories in their estimation: Anguilla, Monserrat, St. Kitts and Nevis, and St. Lucia. They provide country-level estimates for the rest of the countries or territories in our sample.

including the downscaling model and evaluate their performance on difference scales for the Eastern Caribbean region.

We find that the national-level HDI estimations are well-correlated with ground-truth HDI data at the country level in the Eastern Caribbean. We also find that the parish- and district-level HDI prediction are strongly and positively correlated with income as well as other well-being indicators in countries where such disaggregated data are available. This cross-verification exercise suggests that the parish- or district- level HDI predictions can be good proxies for local well-being in these countries.[4] More importantly, this also provides a high degree of confidence in using our subnational-level estimations in other countries where such data are not available, given the high degree of similarity in the geographic, economic, and sociodemographic characteristics in these countries. Our estimation at the one kilometer by one kilometer grid-level enables a substantially more detailed understanding of well-being at the local level.

We then evaluate whether measures of local well-being can be used to improve social assistance targeting and through what channel. To do this, we simulate policy scenarios in which policymakers aim to distribute social assistance benefits to those most in need. Our first two policy simulations are akin to geographic targeting—a common targeting method. In its purest form, geographic targeting selects areas where a social assistant program will operate and gives benefits to all in those areas. This method is highly pertinent in cases where natural disasters lead to widespread losses in certain areas. One common variant is to select areas where the program will operate and use additional eligibility criteria to select households within the area that will benefit. In another variant the program operates in all areas, but policymakers choose how to allocate program resources—including by rationing the caseload and administrative resources—across the areas served. Administrative resources are geographically targeted in the latter case, not program benefits.

In the first simulation exercise, we evaluate whether local well-being estimates can improve targeting by reducing aggregation bias in the purest form of geographic targeting. We find that using the subnational-level well-being measures to select program beneficiaries can significantly reduce targeting errors, compared to using national-level measures alone. In this case improvement in targeting is achieved because granular geospatial information provides more accurate proxies for social assistance needs than aggregated data.

In the second simulation exercise, we evaluate the degree to which local well-being estimates can improve targeting by better allocation of resources across geographic areas. We find that programs with eligibility rules informed by local well-being estimates have smaller targeting

[4] Level 1 administrative boundaries are called “parishes” in Antigua and Barbuda, Barbados, Dominica, Grenada, St Kitts and Nevis, St. Vincent and the Grenadines, and “districts” in St. Lucia. In what follows, we use parish and district interchangeably.

errors than programs with uniform eligibility rules. In this case, local well-being estimates allow policymakers to allocate caseloads proportionally to need. This result also applies to situations where geospatial data are used to allocate administrative resources to various geographic areas.

In our third policy simulation, we evaluate the degree to which local well-being estimates can improve targeting by providing proxies for hard-to-verify information or tacit knowledge. Our policy scenario is akin to hybrid targeting, in which a part of the information on the individual or household's socioeconomic condition can be verified, and the other part needs to be imputed or predicted. We consider a scenario in which existing program eligibility is based on proxies for the need for social assistance, such as demographics and formal income. In another scenario, policymakers also have information on the average value of well-being in the local area. We show that local information can be combined with existing proxies in beneficiary selection to reduce targeting errors. One practical interpretation of this result is that local well-being can be used to proxy hard-to-verify information like informal income and to complement other socioeconomic indicators and formal income to inform program design. Additionally, in the context of SIDS where local authorities may employ tacit knowledge about vulnerable areas for social assistance, local well-being can serve to verify the accuracy of and complement such knowledge.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 describes the data and methodology. Section 4 compares model performance and presents baseline estimations. Section 5 discusses simulation results and policy implications. Section 6 concludes.

# 2. Related Literature

Our paper is related to a rich literature on social assistance targeting.[5] Although the general advantages and disadvantages of each targeting method are well-established,[6] the literature is not definitive in ranking among targeting methods. It recognizes that the choice of targeting method should be based on program objective, as well as context and capacities. As new data and techniques become available and capacities evolve, the practice and potential of each method are also changing. Against this background, our paper offers a valuable case study in

[5] See Banerjee et al. (2020) and Grosh et al. (2022) for comprehensive summaries of the literature.

[6] For example, proxy mean testing (PMT) requires detailed household and individual characteristics data, which are also constrained in countries with limited administrative capacities (Skoufias et al. 2001). It may also suffer from high exclusion errors due to its limited ability to capture the variation in consumption and income (Brown et al. 2018; Iasen and Lang 2015; Kidd et al. 2017). Evidence on whether data-driven targeting approaches outperform discretionary approaches is mixed. Alatas et al. (2012) show that community targeting performs worse than PMT in Indonesia. Haseeb and Vyborny (2022) document improvement by replacing discretion with administrative data in a large-scale unconditional cash transfer program in Pakistan. Banerjee et al. (2020) argues that data-driven approaches may be susceptible to distortion due to misreporting.

assessing the potential use of new data and techniques to improve targeting in countries with relatively limited data availability and administrative capacity. We show that local well-being measures derived from satellite imagery can inform geographic targeting and hybrid targeting designs.

Our paper is also related to recent papers that use big data and machine learning techniques to improve targeting (Jean et al. 2016; McBride et al. 2021). One strand of the literature focuses on the use of mobile phone and social media data to target households directly (World Bank 2021). A recent application is Aiken et al. (2022), who estimate poverty from mobile phone subscribers' data and evaluate the COVID-19 relief aid program in Togo. A limitation of this approach is that a program for those without mobile phone access cannot be designed and evaluated. The broad application of this approach also relies on policy makers' access to mobile phone data, which are usually private sector data. The lack of centralized data on mobile phone usage also makes comparing the results difficult across countries.

Another strand of the literature focuses on the use of satellite data. The idea of mapping local well-being with satellite imagery is not new. GiveDirectly—a charity that makes unconditional cash transfers available to households in extreme poverty—has used satellite imagery to identify households that live in dwellings with thatched roofs in Eastern African countries. The rationale is that household assets are a good proxy for income, and a dwelling typically represents a household's most valuable asset. In countries where the transfer program operates, poorer households tend to live in dwellings with thatched roofs while richer household tend to live in houses with corrugated tin or tiled roofs (Abelson et al. 2014; Houshofer and Shapiro 2016).[7] As satellite imagery becomes more readily available, it makes poverty mapping possible in places without recent censuses or surveys with the potential for faster update and more granularity. Our approach builds on Sherman et al. (2023), who estimate HDI at the subnational level for a large sample of countries using satellite imagery data across the globe. This approach can fill the data gap for many countries, especially the small developing states for which other forms of disaggregated data are limited. One potential concern in using this approach is that the predictions may have system biases against certain regions and social groups (Aiken et al. 2022; Kondmann and Zhu 2021; Zhang et al. 2022); therefore, it should be adapted to local contexts with predictions cross verified using disaggregated information. By adapting the approach to the Eastern Caribbean region, we take a step toward this direction, showing that local well-being mapping can be a good proxy for income and other well-being indicators in the region.

[7] As an initial step of the program, villages with the highest proportion of thatched roofs were chosen in the study of the impact of unconditional cash transfers in Haushofer and Shapiro (2016). The operation model used by GiveDirectly has since expanded its eligibility criteria, including physical back-checks, data back-checks, and crowd-sourced labor to confirm recipient identity and thatched-roof ownership.

# 3. Data and Methodology

To predict well-being in the Eastern Caribbean at the granular level, we combine several sources of data, in particular microdata and high-resolution satellite data. The use of satellite imagery and machine learning (SIML) to remotely predict socioeconomic variables at a detailed spatial level is growing. This approach offers a cost-effective way to obtain information that is traditionally expensive to collect through ground surveys. Although SIML estimates may not precisely replicate ground survey results, their quality has reached a level where they can be valuable in targeting aid and evaluating programs in remote areas where other sources of information are lacking. Researchers are actively studying and addressing errors in SIML predictions to enhance their accuracy.

Our prediction exercise includes the following Eastern Caribbean countries or territories: Anguilla, Antigua and Barbuda, Dominica, Grenada, Monserrat, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, and Trinidad and Tobago.

## Data description

### Socio-economic Data

National-level data on socio-economic outcomes, including HDI, per capital income, health, and education outcomes are available from the United Nations Development Program.[8] The HDI index aggregates scores in three dimensions—the population's health (proxied by life expectancy at birth), human capital (proxied by average years of schooling for adults aged 25 years or more and the expected years of schooling for children of school entering age), and standard of living (measured by the log of gross national income (GNI) per capita). The three sub-indexes are aggregated in geometric mean.

Access to detailed, up-to-date data at the subnational level in Small Island Developing States (SIDS) is a persistent challenge.[9] For the Eastern Caribbean, micro data on income are available for St. Lucia through the Population and Housing Census and for Grenada through the Labor Force Survey. Subnational-level HDI data are not available for Anguilla, Monserrat, St. Kitts and Nevis and St. Lucia. Sherman et al. (2023) provide HDI estimates for first-level administrative units for other countries or territories in the region.

[8] Source: https://hdr.undp.org/data-center/human-development-index.

[9] Developing better indicators for the distribution of Household income, consumption and savings is one of the recommendations of the G20 New Data Gap Initiative (https://www.imf.org/en/News/Seminars/Conferences/g20-data-gaps-initiative.

**Satellite data**

We use satellite data for the prediction of granular-level well-being, combining two granular satellite datasets: first, the Multi-task Observation using Satellite Imagery & Kitchen Sinks (MOSAIKS) data and second, the nighttime light data collected by the U.S. Air Force Defense Meteorological Satellite Program (DMSP) (See Sherman et al. 2023).

The MOSAIKS dataset was developed by Rolf et al (2021) to improve the efficiency and reduce the computational cost of analyses combining SIML approaches, an important constraint preventing their wider use in low and middle- income countries. Dataset of satellite images are typically large with considerable amounts of information stored as colored pixels. Learning patterns from images into outcome variables of interest is typically a computationally intensive process. MOSAIKS conveniently breaks this task into two steps. First the information contained in the daytime satellite images is encoded once into a set of variables, called “features”, which can then be used as variables entering ridge regression or alternative high-dimensional estimation techniques for a variety of applications, such as predicting population, house prices, elevation, or forest cover.

Specifically, the MOSAIKS data are organized along a global 0.01 x 0.01 equal-angle grid, which is roughly 1 kilometer by 1 kilometer near the equator. To each tile is associated a vector of 4,000 task-agnostic features that encode the information contained in the pixels inside the tile. These features are computed using random convolutions of the underlying image (Rahimi et al. 2008). Typically, each parish aggregates two or three dozen MOSAIKS tiles (Appendix Figure A2).

We also use nighttime light satellite data to complement MOSAIKS daytime data.

## Methodology

We leverage countries where data are available to train the model on the satellite data features, allowing us in turn to make predictions in countries where detailed data are not available.

We first spatially match the outcome data and MOSAIKS features and nighttime light data so that the units of the outcome variable, MOSAIKS features, and nighttime light features match each other. We then train the model by running the ridge regressions and estimate the weight for each MOSAIKS and nighttime light feature. Last, we apply the estimated weights to the MOSAIKS and nighttime light features at the geographic unit of interest to predict the outcome variable of interest at the specified geographic unit. In this paper, we specifically aggregate MOSAIKS and nighttime light features to first, the national-level and second, to the parish- and district-levels; moreover, we apply the estimated weights to (1km x 1km) grid-level MOSAIKS[10] and nighttime light features in the Eastern Caribbean to predict the outcome variable at a very granular level.

Our baseline prediction exercise builds on Sherman et al (2023) and estimates the following equation:

$$Y = \beta_0 + \beta_1 X_{MOSAIKS} + \beta_2 X_{NL} + \epsilon,$$

where $Y$ refers to the outcome variable (e.g., population density, income), $X_{MOSAIKS}$ is a matrix of daytime MOSAIKS features capturing the information from satellite imagery, and $X_{NL}$ is a matrix of nighttime light features. The coefficients $\beta_0, \beta_1, \text{and } \beta_2$ are obtained from a ridge regression, as in Rolf et al. (2021).

With these estimates at hand, we can generate out-of-sample predictions for the outcome variable of interest as long as MOSAIKS and nighttime light data are available; that is, we can obtain estimates of the outcome variables in geographical units where they are missing. For instance, we can employ all the global data or specific regional data, estimate the weight for each MOSAIKS and nighttime light feature, and apply the weights in the regions missing the outcome variable.

Moreover, we can predict the outcome variable at geographical scales different from those at which it is initially available, extending the data to coarser geographic units ("upscaling") or to finer geographic units ("downscaling"). As an example of a downscaling exercise, if the outcome variable is available only at the national level, we can compile the global data at that level, estimate the weights, and apply them to the MOSAIKS and nighttime light features at the subnational level in order to predict the outcome variable. For a detailed discussion on the methodology, see Sherman et al (2023).

We use a range of metrics to assess the accuracy of the prediction models including mean squared error (MSE), mean absolute error (MAE), and R-square. MSE measures the variance of residuals while MAE measures the average of residuals. Combined, the MSE, MAE, and R-square, provide good insights as to how the predicted values perform against the actual values of the outcome variable.

To evaluate how each model is trained, we first consider in-sample and out-of-sample performance metrics. We split the dataset into training and testing samples, train the model on the training sample, and finally verify the model prediction of the outcome variable for both the training and testing sample against its true value. Here, we consider MSE and R-square. We

[10] This is the finest geographic unit in which MOSAIKS information is available.

then assess the accuracy of the predicted values against true values of the outcome variables, whenever possible. For instance, we can train the model with GDP as the outcome variable, predict the aggregated national-level GDP for the Eastern Caribbean countries, and compare the predicted and actual values of GDP at the national level. Here, we consider mean squared MSE, MAE, and R-square for the accuracy of the predicted values.

# 4. Prediction Results

Our prediction exercise proceeds in three steps.

First, we focus on population density prediction, training and evaluating multiple models for the Eastern Caribbean. We start with population density prediction because it allows us to directly validate the models using subnational ground truth data for all countries in our sample. Moreover, the role of population density in boosting human development has been shown in an extensive theoretical and empirical literature (Lucas 1988, Glaeser 1999, Glaeser and Resseger 2010). The empirical relationship holds in the Eastern Caribbean based on microdata from St. Lucia, which shows a strong correlation between population density and income (Figure 1). We compare model performance and validate model prediction against ground truth data in the Eastern Caribbean at the national and subnational levels (e.g., parishes or districts). The population density prediction exercise allows us to select the best model for the region in the HDI prediction exercise.

Second, we train and evaluate models for HDI prediction at the national level. We validate the models against ground truth HDI data in the Eastern Caribbean.

Third, based on model performance from the first two exercises, we choose a baseline model. We generate HDI predictions at the subnational and grid levels in the Eastern Caribbean using the baseline model and validate them using available ground truth HDI data at the subnational level.[11]

## Model Estimation and Selection

We train and evaluate eleven models whose training sets vary along two dimensions: the level of aggregation and the list of countries included. For the population density prediction exercise, we use data from Gridded Population of the World (GPW) as ground truth. The models are as follows:

[11] We use subnational HDI estimates from Sherman et al. (2023) as ground truth. These are available for all the countries and territories in our sample except Anguilla, Monserrat, St. Kitts and Nevis, and St. Lucia.

1. **Global scale model**: The training sample includes a random sample of 1 million grid points across the world.[12]
2. **By-continent model**: Same as model 1 but split the grid points into 6 continents and train the model by continent.
3. **By-continent fixed model**: Same as 2 but classifies all the Eastern Caribbean countries on the North American continent instead of the South American continent. [13]
4. **Barbados-based model**: The training sample includes all the grid points in Barbados.[14]
5. **Guadeloupe-based model**: Same as Barbados-based prediction but using Guadeloupe data.
6. **Martinique-based model**: Same as Barbados-based prediction but using Martinique data.
7. **Neighbors-based model**: The training sample combines Barbados, Guadeloupe, and Martinique data.
8. **National-level model**: The training sample is the area-weighted national average for 179 countries.
9. **Subnational-level model**: The training sample is the area-weighted average at the subnational level (i.e., first-level administrative units) for 179 countries.
10. **Barbados enumeration block (EB) model**: The training sample consists of enumeration blocks from the Barbados Survey of Living Conditions.[15]
11. **St. Lucia settlement model**: The training sample consists of settlements from the St. Lucia Population and Housing Census.[16]

The main goal of this exercise is to compare model performances and use them to inform our choice of the best model for HDI estimation in the Eastern Caribbean. Note that among the 11 models, only models 8 and 9 can be used to predict HDI at the subnational level because HDI data is available at the geographic unit used in these two models but not the other models. Nevertheless, the other models provide useful benchmarks for model performance.

We estimate the coefficients for equation (2) from the 11 models.[17] For each model we present performance metrics based on data from (i) MOSAIKS (i.e., daytime); (ii) nighttime lights; and (iii) MOSAIKS and nighttime lights combined.

[12] For each grid point of MOSAIKS data, we find the nearest population density grid point.

[13] Following Sherman et al (2023), the continent classification in model 2 classifies Grenada and Trinidad and Tobago on the South America continent.

[14] To link MOSAIKS features to population density, we extract all the MOSAISKS features in a country, create Voronoi polygons for each MOSAIKS point, and obtain the population density for the polygon from GPW data.

[15] Unlike model 4, in models 10 and 11 we create Voronoi polygons for each enumeration block based on the centroid latitude and longitude of each enumeration block. For each Voronoi polygon corresponding to each enumeration block, we compute the means of MOSAIKS features that fall into the polygon.

[16] Similar to model 10, we compute the mean of MOSAIKS features that fall into each polygon that corresponds to each St. Lucia settlement.

[17] See Table 1 for the unit of observations to estimate equation (2) under each model.

Tables 1 and 2 present the in-sample and out-of-sample performance metrics for each model, respectively. In general, predictions based on MOSAIKS outperform those based on nighttime lights while predictions based on a combination of MOSAIKS and nighttime lights outperform those based on MOSAIKS or nighttime lights alone. Predictions using combined MOSAIKS and nightlights well explain both in-sample variation and out-of-sample variation. The predictions explain between 53 percent and 92 percent of the in-sample variation and between 42 percent and 78 percent of the out-of-sample variation. The only exceptions are the Barbados enumeration block model and the St. Lucia settlement model, for which the predictions explain less than 10 percent of the in-sample or out-of-sample variation. One reason for the poor performance of these two models could be errors in matching enumeration block or settlement to MOSAIKS grids. The errors could be potentially large when enumeration blocks or settlements have irregular shapes and are small (e.g., with similar size or smaller than MOSAIKS grids).

## Population predictions

*Population density prediction at the national level*

To validate the usefulness of the predictions for the Eastern Caribbean, we compare model predictions to ground truth of each country in the region.[18] To generate national-level predictions from all models except model 8, we use an upscaling exercise in which estimations of finer resolutions are scaled up to generate estimations of coarser resolutions. The scaled-up aggregation is computed by population-weighted average. Predictions from model 8 are the results from a same-scale exercise in which the model is estimated and predictions are made at the same level of aggregation.

We find strong correlations between ground truth and predictions from most models (Figure 2 and Table 3). More importantly, the national-level and subnational-level models (models 8 and 9, which we use for HDI prediction exercises) perform very well compared to other models. Using combined MOSAIKS and nightlights data, the national-level predictions explain 80 percent and 48 percent of the in-sample and out-of-sample variations, respectively; the subnational-level predictions explain 74 percent and 70 percent of the in-sample and out-of-sample variations, respectively (Tables 1 and 2). The subnational-level predictions perform particularly well out of the sample and have the second-best performance out of the 11 models (after model 2). Models 8 and 9 also perform well in the Eastern Caribbean validation exercise.

[18] The validation sample includes Antigua and Barbuda (ATG), Dominica (DMA), Grenada (GRD), St. Kitts and Nevis (KNA), St. Lucia (LCA), St. Vincent and the Grenadines (VCT), and Trinidad and Tobago (TTO).

Both models explain 27 percent of the variation, in line with the top performers among the 11 models (Figure 2 and Table 3).

*Population density prediction at the subnational level*

We repeat the validation exercise at the subnational level. Models 1 to 7 use an upscaling exercise to generate subnational-level predictions. Model 8 uses a downscaling exercise in which coefficients estimated from coarser (national-level) resolutions are scaled down to generate estimations of finer resolutions (subnational-level). Model 9 generates the predictions using a same-scale exercise. We find that the correlations between ground truth and the predictions are generally higher at the subnational level than at the national level (Figure 3 and Table 4). For example, the three global-scale models (models 1-3) explain over 70 percent of the variation at the subnational level, compared to 30 to 35 percent at the national level. The subnational-level model (model 9) also performs well, explaining 47 percent of the variation. This is a significant improvement compared to the national-level predictions by the same model, which explains 27 percent of the variation.

In all, the results so far show that the national-level and subnational-level models (models 8 and 9) can predict population density in the Eastern Caribbean with high accuracy. The subnational-level model performs better than the national-level model overall in terms of the accuracy of out-of-sample predictions as well as performance stability in different levels of aggregation.

### HDI predictions

*HDI predictions at the national level*

To generate HDI predictions, we train and evaluate variants of models 8 and 9. In this case, the training sample in model 8 is the area-weighted national average and the training sample for model 9 is a combination of subnational and—if subnational HDI data are not available—national averages.

In addition to ground truth HDI data, our validation exercise also uses each of the three components of HDI (health, education, and income per capita) as well as total income (measured by the log of GNI). The results with these additional measures shed light on whether the predicted HDI is a good overall measure of well-being and if so, which dimensions of well-being it captures particularly well.

Tables 5 and 6 present the in-sample and out-of-sample performance metrics. As in population density prediction, using a combination of MOSAIKS and nightlights improves model performance compared to using MOSAIKS or nightlights alone. Predictions from both models using combined MOSAIKS and nightlights data explain over 80 percent of in-sample and out-of-sample variations across countries. The performance with respect to other well-being measures are similarly good, with predictions explaining over 75 percent of the in-sample variation and over 60 percent of the out-of-sample variation. The only metric lower than this range is the predictions of model 8 in explaining health out of the sample (46 percent). Overall, model 9 outperforms model 8 marginally in all metrics.

We validate the HDI predictions for the Eastern Caribbean against the ground truth HDI of each country in the region (Figure 4 and Table 7). Model 9 outperforms model 8 in HDI predictions by a large margin. Model 9 predictions explain 45 percent of the variation in the region while model 8 explains 12 percent. Model 9 predictions also explain between 30 to 42 percent of the variations of other measures. Education is an exception, for which model 9 predictions explain only 9 percent of the variation.

*HDI predictions at the subnational level*

Our results on population density and HDI predictions show that both the national-level and subnational-level models (models 8 and 9) can predict with high accuracy. The subnational-level model performs better than the national-level model overall, especially for HDI predictions in the Eastern Caribbean. Based on these results, in what follows we use model 9 to generate baseline predictions of subnational HDI in the Eastern Caribbean. Figure 5 shows the baseline predictions.

To evaluate the accuracy of these predictions, we compare them to ground truth subnational-level data in the region. We do this exercise at the district level in St. Lucia using the 2010 Population and Housing Census and at the parish level in Grenada using the 2019Q4 Labor Force Survey data. We cannot do this exercise for other countries in the region because of the lack of ground truth data.

Figure 6 shows the results. We find that the predicted HDI, GNI, and income are strongly and positively correlated with income from the St. Lucia Census at the district level. The correlation coefficients are 0.57, 0.62, and 0.62, respectively. The same holds for Grenada, albeit with a weaker correlation coefficient of 0.31.

We also generate HDI predictions at the grid level using model 9 with a downscaling exercise.[19] These predictions are used in the policy simulation exercises in the next section.

[19] This is the same disaggregation level as MOSAIKS features are available at the finest resolution.

# 5. Policy Simulations

With predicted local HDI at hand, we then examine how these predictions can be used to improve the targeting of social assistance programs in the Eastern Caribbean. To do this, we conduct simulation exercises for several policy scenarios, designed to illustrate the various channels through which granular predictions can help with targeting. First, we evaluate whether local HDI predictions can improve targeting by reducing aggregation bias in the purest form of geographic targeting. Second, we evaluate the usefulness of these predictions in a variant of geographic targeting through better allocation of resources across geographic areas. Specifically, we examine whether policymakers informed with local well-being estimates can design location-specific eligibility rules to reduce targeting errors. Third, we evaluate the usefulness of these predictions in a hybrid targeting setting, examining whether policymakers can use local well-being to proxy hard-to-verify information (e.g., informal labor income) to reduce targeting errors.

**Reducing aggregation bias in geographic targeting**

The tenet of geographic targeting is to select the neediest areas where social assistance programs will operate or where caseloads and administrative resources will be allocated. The decision needs to be based on spatial information for which disaggregate information may not be available. Aggregation bias occurs when small geographic units (e.g., a village) are assigned to the information of a larger unit (e.g., a district or a country) but the assignment does not reflect the specific conditions of the smaller unit. This problem arises when heterogeneity appears within the larger unit. For instance, consider a district that contains both a large rural areas with relatively low local well-being and a coastal strip with developed tourism infrastructure with relatively high local well-being.

Our first policy experiment aims to illustrate how local well-being predictions (and granular geospatial information in general) can be used to reduce aggregation bias. The first-best approach to assess potential gain from granular information is to collect ground truth data on small geographical units and assess results based on less disaggregated information. Information on small geographical units, however, is not available in the Eastern Caribbean region—this is precisely the reason that our predictions would be valuable. A second-best approach is to treat disaggregated predictions as ground truth data and compare results based on various levels of disaggregated information. This is the approach we adopt. We use the grid-level predicted values as ground truth and compare simulation results based on two levels of aggregation: parish level and national level.

We consider a hypothetical scenario in which policymakers aim to distribute social assistance benefits to the neediest population. Geographic targeting is used in the absence of household-level information. Here we consider an eligibility rule based on absolute values of HDI: the program selects localities whose HDIs are below a certain absolute value. This rule is pertinent in situations in which the goal of poverty eradication is supported by a flexible budget envelope. Later we consider situations under a budget constraint.

We compare two policy scenarios, which differ by the granularity of available information.

- **Policy 1 (geographic targeting based on national information):** Policymakers who have information on the national mean and standard deviation of HDI simulate the grid-level HDI for each parish assuming a log-normal distribution with mean and standard deviation matched to the national-level statistics.

- **Policy 2 (geographic targeting based on subnational information):** Policymakers who have information on the mean and standard deviations of local HDI for each parish simulate the grid-level HDI for each parish, assuming a log-normal distribution with mean and standard deviation matched to the national-level statistics.

Two common metrics on the accuracy of targeting are exclusion and inclusion errors. With given program eligibility rules, the exclusion error measures the fraction of the population that is eligible but incorrectly classified as non-eligible while the inclusion error measures the fraction of the population that is noneligible but incorrectly classified as eligible. Formally, let exclusion and inclusion errors be defined as follows:

$$Exclusion\ error = FN/(FN + TP),$$
$$Inclusion\ error = FP/(FP + TN),$$

where $TP, TN, FP$ and $FN$ denote, respectively, the number of true positives ($TP$), true negatives ($TN$), false positives ($FP$), and false negatives ($FN$). Exclusion and inclusion errors depend on the eligibility threshold implied by targeting rules. This can be illustrated with an exclusion–inclusion curve (EIC), which maps the full space of eligibility thresholds. The EIC encompasses all types of targeting from universal programs (i.e., the entire population is eligible) to non-existent programs (i.e., no one is eligible) and everything in between. In this sense the EIC maps the fundamental trade-off between exclusion and inclusion errors.

We calibrate the simulations to each country's ground truth data and compare the performance of Policies 1 and 2 using the EIC, in which the upper-left corner corresponds to programs with no beneficiaries (i.e., maximum exclusion error and zero inclusion error). The lower-right corner corresponds to universal programs (zero exclusion error and maximum inclusion error).

In the case of St. Lucia, the EIC for Policy 2 lies strictly below the EIC for Policy 1 EIC, suggesting that simulation 2 outperforms Policy 1 for all choices of eligibility thresholds (Figure 7). In other words, for any given budget (that allows the program to cover a given number of beneficiaries), beneficiary selection based on more granular information (i.e., parish-level) instead of more aggregated information (i.e., national-level) leads to better targeting. In the context of geographic targeting policy design, the reduction in targeting errors results from a reduction in aggregation biases based on granular geospatial information. The improvement is especially valuable for localities with low HDI but are in proximity to localities with high HDI. These localities may not receive social assistance if aggregated information is used to determine program eligibility, but they will be eligible if fine resolution HDI information is used.

We find variation in model performances across countries. Appendix Figure A3 shows the results for other countries. The variation reflects the relative value of subnational level data in a country. Intuitively, countries with more within-country variation in local well-being benefit more from subnational data. A comparison of the correlation between subnational predictions and ground truth across countries illustrates this point (Appendix Table A1).

Another way to understand the improvement of policy 2 over policy 1 is to note that policy 2 better reflects variation across localities. An important aspect of the variation comes from the disproportionate concentration of those in need of social assistance in some areas. Figure 8 illustrates this point using simulated data for St. Lucia. We again use St. Lucia as an example because the simulated data can be validated with district-level income data available from the St. Lucia Census. The simulated HDI from policy 1 is distributed quite evenly across different districts, implying that based on the data, the social assistance program selects equal numbers of beneficiaries in each district despite their differences. In contrast, the simulated HDI from policy 2 reflects differences across districts. Areas with the lowest HDI, therefore, tend to fall in districts with a lower mean or a higher standard deviation or both.

### Improving resource allocation

Our second policy experiment aims to illustrate how local well-being predictions can be used to improve targeting in a variant of geographic target. Specifically, we ask whether these predictions can help allocate program resources proportionally to areas in need.

Like policies 1 and 2, we consider a hypothetical scenario in which policymakers aim to distribute social assistance benefits to the neediest population. We again assume the ground-truth is HDI at the grid level. Different from policies 1 and 2, here eligibility rules are based on the relative values of HDI. The program selects localities whose HDIs are below a certain percentile. This rule is pertinent in situations in which the total number of beneficiaries is constrained by a budget envelope.

We compare two policies. In both cases, policymakers have information on parish-level mean and standard deviations of HDI, as in policy 2.

- **Policy 3 (nationally defined eligibility rule)**: Localities in the lowest (nationally defined) $P$ percentile in each parish are eligible for the program.

- **Policy 4 (parish-specific eligibility rule)**: Localities in the lowest parish-specific percentile are eligible for the program.

Policy 3 reflects an even distribution of program or administrative resources or both. For example, it captures a situation in which the same amount of administrative resources is allocated to each parish and as a result the program achieves the same coverage ($P$ percentile) across parishes. In contrast, policy 4 reflects an allocation rule based on parish-specific needs.

We impose two additional rules in policy 4. First, policy 4 selects the same total number of beneficiaries as policy 3. Second, the selected beneficiaries satisfy the same absolute threshold. In other words, if a locality is chosen in one parish, a locality with the same HDI in a different parish would also be chosen. The first rule makes policies 3 and 4 comparable under the same budget envelope. The second rule reflects an equity consideration of the policy design. Both rules together ensure that beneficiary selection in policy 4 has a unique solution.

Figure 9 shows EIC from policies 3 and 4 for St. Lucia. The policy 4 curve lies strictly below the policy 3 curve, suggesting that policy 4 outperforms policy 3 for all choices of eligibility thresholds. A direct implication of this result is that local well-being predictions can be used to inform locality-specific eligibility rules to improve targeting. The result also applies to situations where local well-being predictions or other geospatial information is used to allocate administrative resources across the country. All else equal, allocating more caseloads and administrative resources to areas with greater needs can lead to better targeting outcome. Appendix Figure A4 shows qualitative similar results for other countries in the region.

We again illustrate the policy improvement using simulated data on St. Lucia in Figure 10. Here the simulated distributions of HDI are the same in both policies, but district-level HDI information helps to inform beneficiaries in each district in policy 4. In policy 3, all the districts have the same coverage despite their differences in HDI mean and standard deviation, implying that the HDI threshold of program eligibility (dash lines in panel A) is higher in better-off districts and lower in less well-off districts. In other words, in a better-off district, a locality with a relatively high HDI will be selected to the program whereas in a less well-off parish, a locality with the same HDI may not be selected to the program. In policy 4, the coverage of beneficiaries reflects cross-district differences in HDI. It is higher in less well-off districts and lower in better-off districts. HDI disproportionally falls in districts with a lower mean or a higher standard deviation or both. Targeting is improved by eliminating beneficiaries with high HDI in better-off districts as well as admitting beneficiaries with low HDI in less well-off districts.

**Proxying for hard-to-verify information**

Our third policy experiment aims to illustrate how local well-being predictions can be used to proxy hard-to-verify information in a hybrid targeting setting.

As before, policymakers aim to select the neediest population to receive benefits, but they do not observe the data on population needs and thus need to select program beneficiaries using some proxies. These could be demographic indicators (e.g., older individuals, people with disabilities, families with school-age children, or families with a single mother), employment indicators (e.g., the unemployed), and partial information on income (e.g., self-reported income or formal income). This policy scenario encompasses some of the most commonly used targeting methods in the Easter Caribbean as in other developing economies (Grosh 2022), including demographic or categorical targeting in which eligibilities are based on sociodemographic indicators, hybrid targeting in which eligibilities are based on an income limit combined with a vulnerability score from sociodemographic characteristics, or proxy means targeting in which eligibilities are based on scores computed with a set of indicators using a formal algorithm. Regardless of the type of indicators and approaches used to rank the population’s need for social assistance, one element in common in these approaches is that the indicators are imperfect proxies for the true need.

In our simulation we investigate whether local well-being measures can be combined with existing proxies in a hybrid targeting setting. Specifically, we evaluate whether adding the local HDI value to existing proxies can reduce targeting error. As in policies 1 to 4, we assume the grid-level HDI predictions are the ground truth. To capture existing proxies in a parsimonious way, we simulate a variable (hereafter referred as the proxy variable) that is partially correlated with the ground truth.

We compare two policies:

- **Policy 5 (proxy targeting)**: Policymakers use the proxy variable to select beneficiaries. Those with a value of the proxy variable below a certain percentile are eligible for the program.
- **Policy 6 (proxy targeting combined with parish-level information)**: Policymakers rank the population from a score that combines parish level-predicted HDI values and the proxy variable.[20] Those with a score below a certain percentile are eligible for the program.

We calibrate the policy by setting the proxy variable to have a 0.25 correlation with the ground truth. The value is in the range of targeting errors estimated in the literature (Aiken et al. 2022, 2023).

Combining local well-being measures with existing proxies reduces targeting error. Figure 11 shows the EIC from policies 5 and 6 with St. Lucia data. The curve from policy 6 lies below the curve from policy 5 for almost all the percentile cutoffs, suggesting that policy 6 outperforms policy 5.

This finding extends to the other Eastern Caribbean countries (Appendix Figure A2). Policy 6 outperforms policy 5 in all countries but to different extents. Intuitively, the less accurate the proxy is relative to predicted ground truth compared to local HDI, the more valuable the information from local HDI and vice versa. The cross-country variation in Figure A4 is driven by how well local HDI predictions capture the ground truth. In countries where parish-level HDI is more correlated with the ground truth, the improvement from policy 5 to policy 6 is larger; that is, when the local well-being measures are informative about the ground truth, they will be more useful for targeting.

These results have direct implications for policy design. To the extent that local well-being measures are good proxies for population needs, they can be used to complement other observable proxies. One interpretation is that local well-being measures can proxy for unobserved and hard-to-verify information in beneficiary selection. For example, in countries with a large informal sector, informal income can be imputed based on average local income. This can then be combined with formal income to proxy for total income.

# 6. Conclusion

Prioritizing populations most in need of social assistance is one of the most important policy decisions in the Eastern Caribbean countries as in many developing countries, yet the task of

[20] We use equal weights in the simulation, but the results can be generalized to other weighting schemes.

social assistance targeting is often constrained by data limitations. Despite their small size, these countries exhibit a substantial amount of within-country heterogeneity in local well-being, including income, education, and health outcomes. Although these disparities are indicative of poverty and the need for social assistance, they are not always apparent to policy makers because of the scarcity of data.

We leverage recent advances in machine learning and satellite imagery processing to propose an implementable strategy for improving social assistance targeting in the face of information constraints. We show that local well-being can be predicted with high accuracy in the Eastern Caribbean region and that such predictions can be used to improve targeting by reducing aggregation bias, better allocating resources across different areas, and proxying for hard-to-verify information.

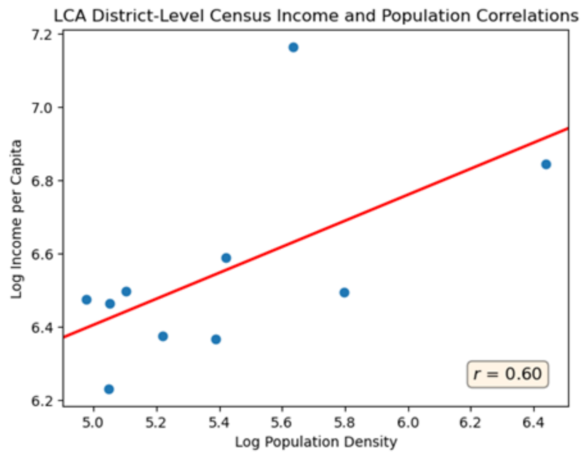
Our results have important policy implications. First, a direct implication for social assistance policy design is that local well-being measures can be used to proxy for unobserved and hard-to -verify information, especially in the Eastern Caribbean where the informal sector is large.[21] Second, allocating program resources—including by rationing the caseload and administrative resources—proportionally to areas in need can improve targeting. Understanding disparities within the country is important for social assistance policy designs. Local well-being measures (and other granular geospatial information) can provide useful information. Third, collecting granular information is important. This includes granular geospatial information as well as other dimensions of disparities. Our local HDI prediction is only a first step in this direction. As satellite imagery becomes more pervasive, the extent of information that can be brought to bear on targeting will continue to evolve. To further exploit the potential of satellite imagery in social assistance targeting, proper assessments of the satellite-informed local well-being mapping are needed and require the periodic collection of survey data with directly measured household welfare in each country.

Our paper focuses on regular social assistance programs, but a similar approach can be applied to emergent social assistance after natural disaster shocks. This is highly pertinent to the Eastern Caribbean which is highly subject to natural disasters. A fruitful extension of our exercise is to use satellite data to predict vulnerability to specific natural disasters and use the predictions to inform emergency social assistance.

[21] The informal sector is estimated to be on average 35 percent of the economy in our sample of countries (Peters 2017; IMF 2022).

# Appendix 1: Figures and Tables

**Figure 1: Correlation between income and population density in St. Lucia**




Note: This figure plots the log of income per capita and the log of population density for St. Lucia districts.
Sources: St. Lucia 2010 Population and Housing Census and authors' calculations.

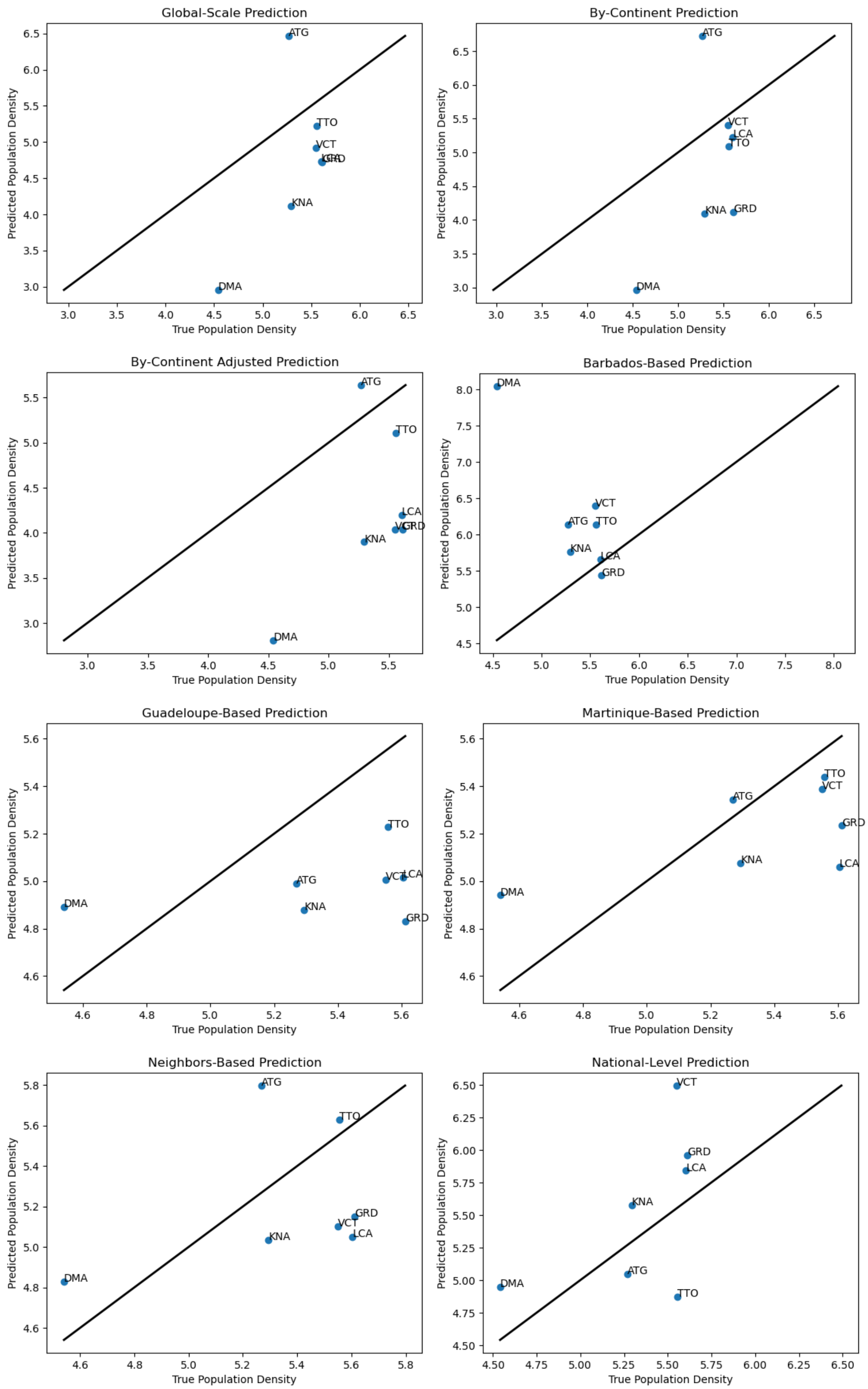
**Table 1: Population density predictions across models: in-sample performance metrics**

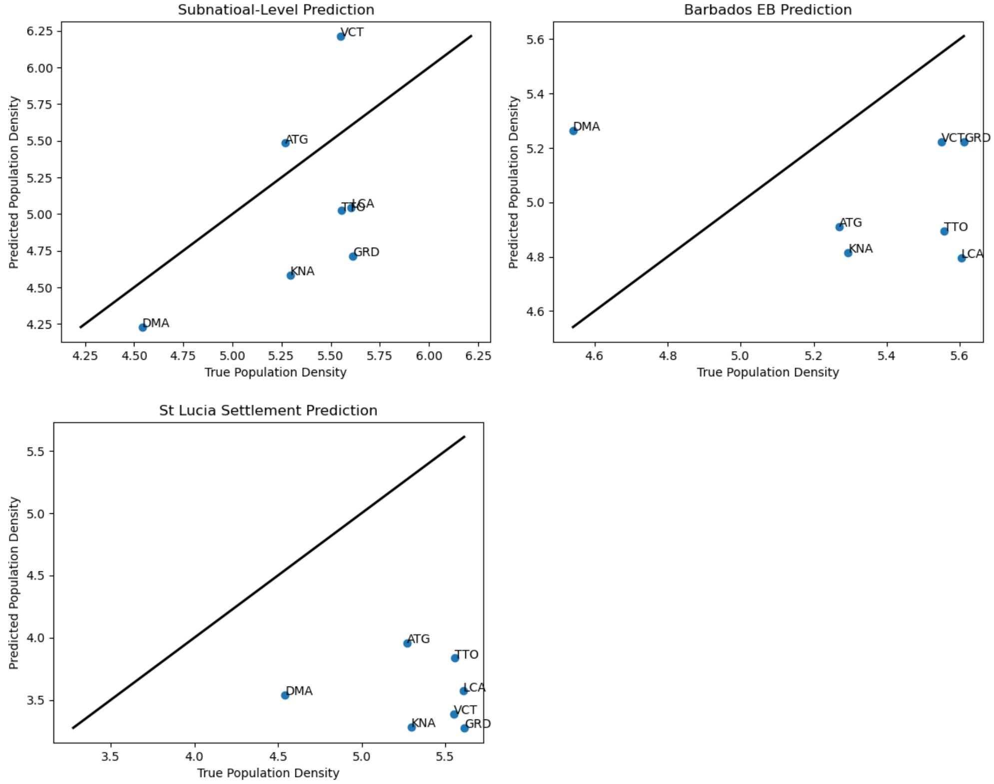
| Model | | MOSAIKS | | Nightlights | | Both | |
|---|---|---|---|---|---|---|---|
| Training Sample | Geographical Unit | MSE | R-square | MSE | R-square | MSE | R-square |
| Global-scale | Random coordinates | 1.092 | 0.624 | 2.180 | 0.249 | 0.984 | 0.661 |
| By-continent | Random coordinates | 0.324 | 0.726 | 0.523 | 0.557 | 0.248 | 0.790 |
| By-continent fixed | Random coordinates | 0.888 | 0.470 | 1.227 | 0.267 | 0.793 | 0.526 |
| Barbados-based | MOSAIKS coordinates | 0.224 | 0.880 | 1.472 | 0.210 | 0.149 | 0.920 |
| Guadeloupe-based | MOSAIKS coordinates | 0.640 | 0.588 | 1.311 | 0.157 | 0.670 | 0.570 |
| Martinique-based | MOSAIKS coordinates | 0.905 | 0.620 | 1.962 | 0.175 | 0.831 | 0.651 |
| Neighbors-based | MOSAIKS coordinates | 0.810 | 0.616 | 1.704 | 0.191 | 0.725 | 0.656 |
| National-level | National polygons | 0.312 | 0.783 | 1.287 | 0.106 | 0.284 | 0.803 |
| Subnational-level | Subnational polygons | 0.844 | 0.698 | 2.158 | 0.229 | 0.728 | 0.740 |
| Barbados Enum. B | Enumeration blocks | 7.408 | 0.088 | 6.592 | 0.185 | 7.294 | 0.102 |
| St. Lucia settlement | Settlements | 8.397 | 0.034 | 7.615 | 0.146 | 8.300 | 0.046 |

**Table 2: Population density predictions across models: out-of-sample performance metrics**

| Model | | MOSAIKS | | Nightlights | | Both | |
|---|---|---|---|---|---|---|---|
| Training Sample | Unit | MSE | R-square | MSE | R-square | MSE | R-square |
| Global-scale | Random coordinates | 1.105 | 0.621 | 2.186 | 0.250 | 0.995 | 0.659 |
| By-continent | Random coordinates | 0.343 | 0.713 | 0.536 | 0.551 | 0.262 | 0.781 |
| By-continent fixed | Random coordinates | 0.953 | 0.421 | 1.248 | 0.242 | 0.853 | 0.482 |
| Barbados-based | MOSAIKS coordinates | 0.992 | 0.548 | 1.889 | 0.138 | 0.757 | 0.655 |
| Guadeloupe-based | MOSAIKS coordinates | 1.023 | 0.490 | 1.613 | 0.196 | 0.917 | 0.543 |
| Martinique-based | MOSAIKS coordinates | 1.551 | 0.467 | 2.536 | 0.128 | 1.513 | 0.480 |
| Neighbors-based | MOSAIKS coordinates | 1.169 | 0.393 | 1.699 | 0.118 | 1.115 | 0.421 |
| National-level | National polygons | 1.092 | 0.527 | 1.946 | 0.158 | 1.202 | 0.480 |
| Subnational-level | Subnational polygons | 1.017 | 0.673 | 2.266 | 0.271 | 0.943 | 0.697 |
| Barbados EB | Enumeration blocks | 7.450 | 0.020 | 7.041 | 0.076 | 7.449 | 0.020 |
| St. Lucia settlement | Settlements | 8.685 | -0.022 | 8.021 | 0.026 | 8.705 | -0.024 |

**Figure 2: Validating national-level population density predictions in the Eastern Caribbean.**



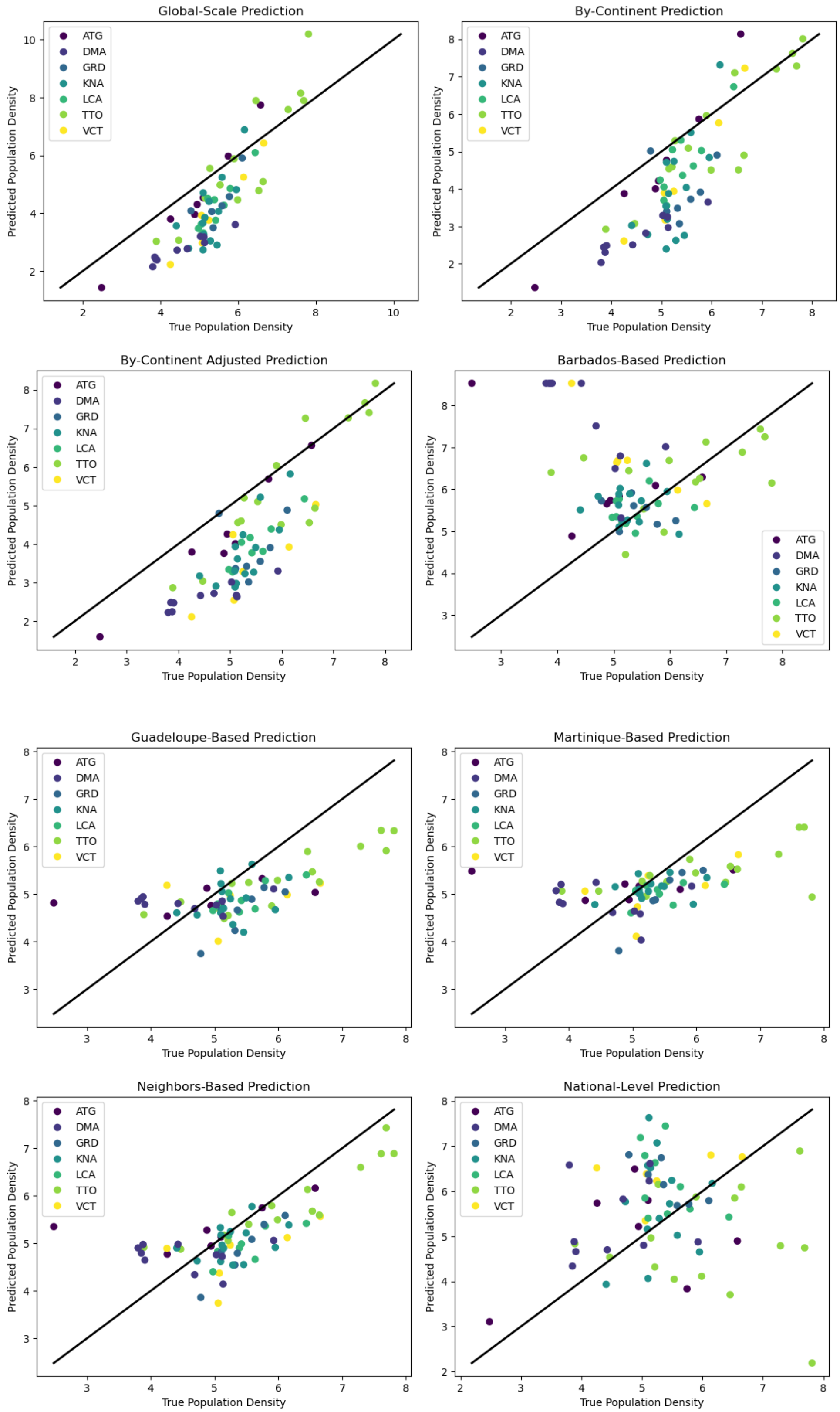


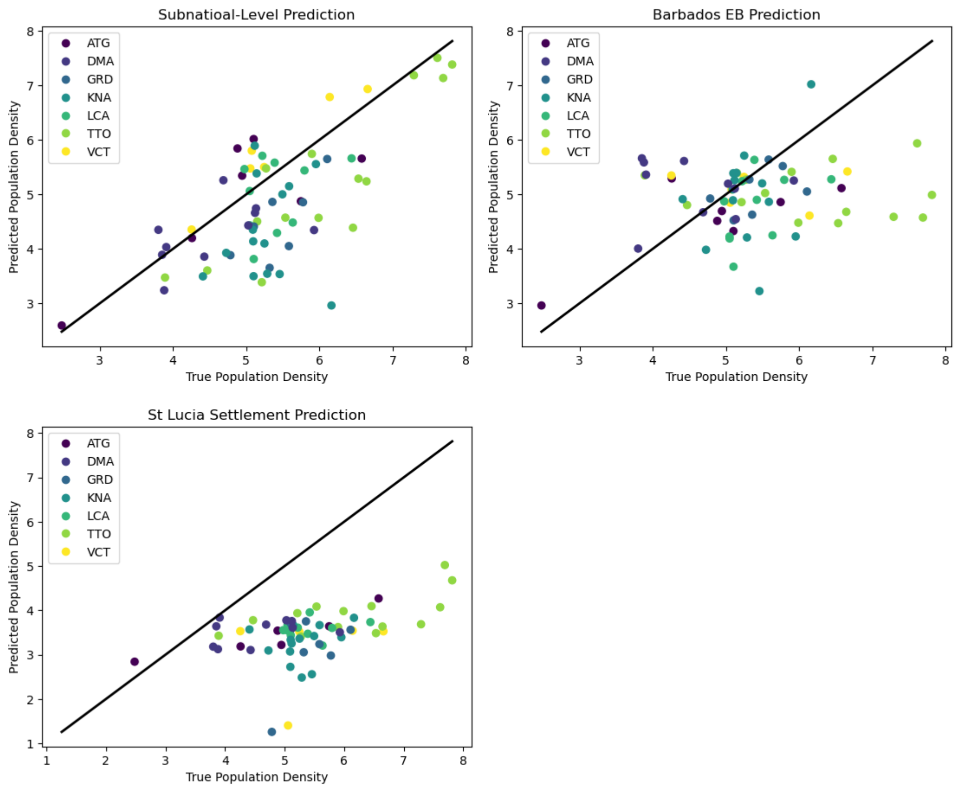

Note: This figure plots national-level population density ground truth and model predictions.
Sources: Gridded Population of the World and authors' calculations.

**Table 1: Validating national-level population density predictions in the Eastern Caribbean: summary statistics and performance metrics across models**

| Descriptives | Mean | Variance | MSE | R-square |
|---|---|---|---|---|
| Ground-truth | 5.346 | 0.146 | | |
| Global-scale | 4.705 | 1.132 | 1.092 | 0.353 |
| By-continent | 4.772 | 1.433 | 1.266 | 0.283 |
| By-continent fixed | 4.228 | 0.831 | 1.764 | 0.294 |
| Barbados-based | 6.193 | 0.752 | 1.994 | 0.793 |
| Guadeloupe-based | 4.937 | 0.020 | 0.285 | 0.078 |
| Martinique-based | 5.424 | 0.123 | 0.115 | 0.321 |
| Neighbors-based | 5.185 | 0.129 | 0.187 | 0.102 |
| National-level | 5.451 | 0.339 | 0.226 | 0.277 |
| Subnational-level | 5.053 | 0.436 | 0.360 | 0.271 |
| Barbados EB | 4.983 | 0.044 | 0.345 | 0.131 |
| St. Lucia settlement | 3.512 | 0.079 | 3.575 | 0.009 |

### Figure 3: Validating subnational-level population density predictions in the Eastern Caribbean






Note: This figure plots subnational-level population density ground truth and model predictions.
Sources: Gridded Population of the World and authors' calculations.

**Table 2: Validating subnational-level population density predictions in the Eastern Caribbean: summary statistics and performance metrics across models**

| Descriptives | Mean | Variance | MSE | R-square |
|---|---|---|---|---|
| Ground-truth | 5.357 | 0.884 | | |
| Global-scale | 4.328 | 2.723 | 1.946 | 0.761 |
| By-continent | 4.329 | 2.421 | 1.895 | 0.704 |
| By-continent fixed | 4.031 | 1.996 | 2.390 | 0.710 |
| Barbados-based | 6.156 | 1.075 | 3.101 | 0.076 |
| Guadeloupe-based | 4.945 | 0.224 | 0.711 | 0.394 |
| Martinique-based | 5.124 | 0.182 | 0.716 | 0.242 |
| Neighbors-based | 5.136 | 0.421 | 0.508 | 0.473 |
| National-level | 5.578 | 1.151 | 2.074 | 0.000 |
| Subnational-level | 4.833 | 1.171 | 0.929 | 0.467 |
| Barbados EB | 4.947 | 0.395 | 1.211 | 0.035 |
| St. Lucia settlement | 3.476 | 0.307 | 4.263 | 0.191 |

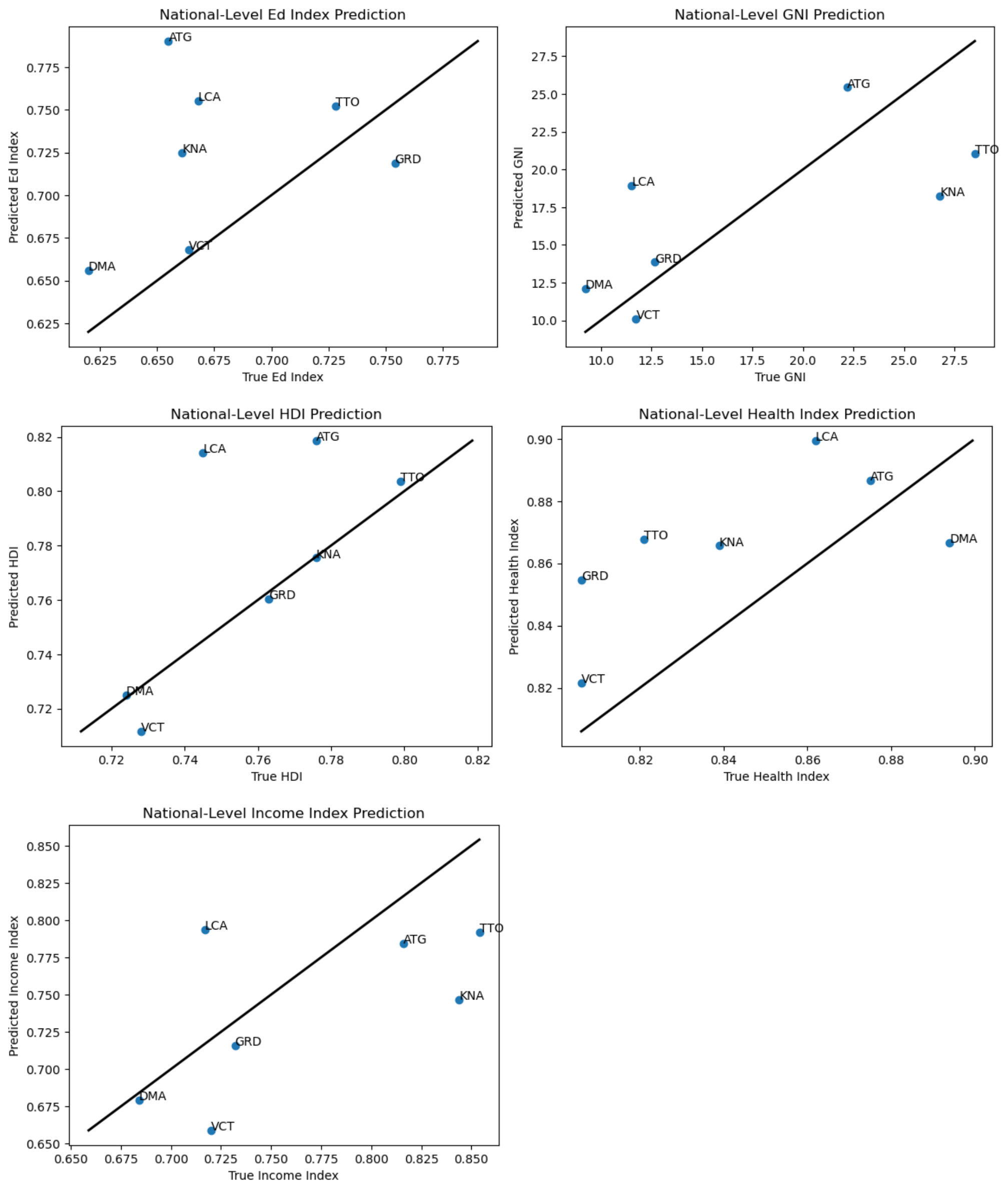
**Table 5: HDI predictions at the national level: in-sample performance metrics**

| | | MOSAIKS | | Nightlights | | Both | |
|---|---|---|---|---|---|---|---|
| Variable | Geographical Unit | MSE | R-square | MSE | R-square | MSE | R-square |
| HDI | National-level | 0.003 | 0.894 | 0.008 | 0.657 | 0.003 | 0.867 |
| | Subnational-level | 0.004 | 0.860 | 0.010 | 0.618 | 0.003 | 0.888 |
| GNI | National-level | 92.373 | 0.768 | 238.488 | 0.402 | 100.321 | 0.748 |
| | Subnational-level | 44.633 | 0.817 | 128.099 | 0.474 | 39.572 | 0.838 |
| Health | National-level | 0.003 | 0.763 | 0.006 | 0.576 | 0.003 | 0.805 |
| | Subnational-level | 0.003 | 0.807 | 0.008 | 0.483 | 0.003 | 0.816 |
| Income | National-level | 0.009 | 0.734 | 0.012 | 0.618 | 0.006 | 0.824 |
| | Subnational-level | 0.006 | 0.812 | 0.013 | 0.583 | 0.005 | 0.850 |
| Ed | National-level | 0.004 | 0.883 | 0.013 | 0.578 | 0.005 | 0.844 |
| | Subnational-level | 0.006 | 0.833 | 0.016 | 0.576 | 0.005 | 0.859 |

**Table 6: HDI predictions at the national level: out-of-sample performance metrics**

| | | MOSAIKS | | Nightlights | | Both | |
|---|---|---|---|---|---|---|---|
| Variable | Geographical Unit | MSE | R-square | MSE | R-square | MSE | R-square |
| HDI | National-level | 0.005 | 0.778 | 0.006 | 0.718 | 0.004 | 0.811 |
| | Subnational-level | 0.005 | 0.813 | 0.010 | 0.590 | 0.003 | 0.858 |
| GNI | National-level | 170.990 | 0.534 | 161.413 | 0.561 | 140.825 | 0.617 |
| | Subnational-level | 108.781 | 0.656 | 206.053 | 0.348 | 105.854 | 0.665 |
| Health | National-level | 0.006 | 0.390 | 0.004 | 0.546 | 0.005 | 0.457 |
| | Subnational-level | 0.004 | 0.754 | 0.008 | 0.454 | 0.003 | 0.767 |
| Income | National-level | 0.013 | 0.586 | 0.008 | 0.730 | 0.008 | 0.733 |
| | Subnational-level | 0.008 | 0.738 | 0.013 | 0.545 | 0.006 | 0.792 |
| Ed | National-level | 0.007 | 0.765 | 0.012 | 0.594 | 0.006 | 0.813 |
| | Subnational-level | 0.008 | 0.780 | 0.016 | 0.533 | 0.006 | 0.825 |

**Figure 4: Validating national-level HDI predictions in the Eastern Caribbean**

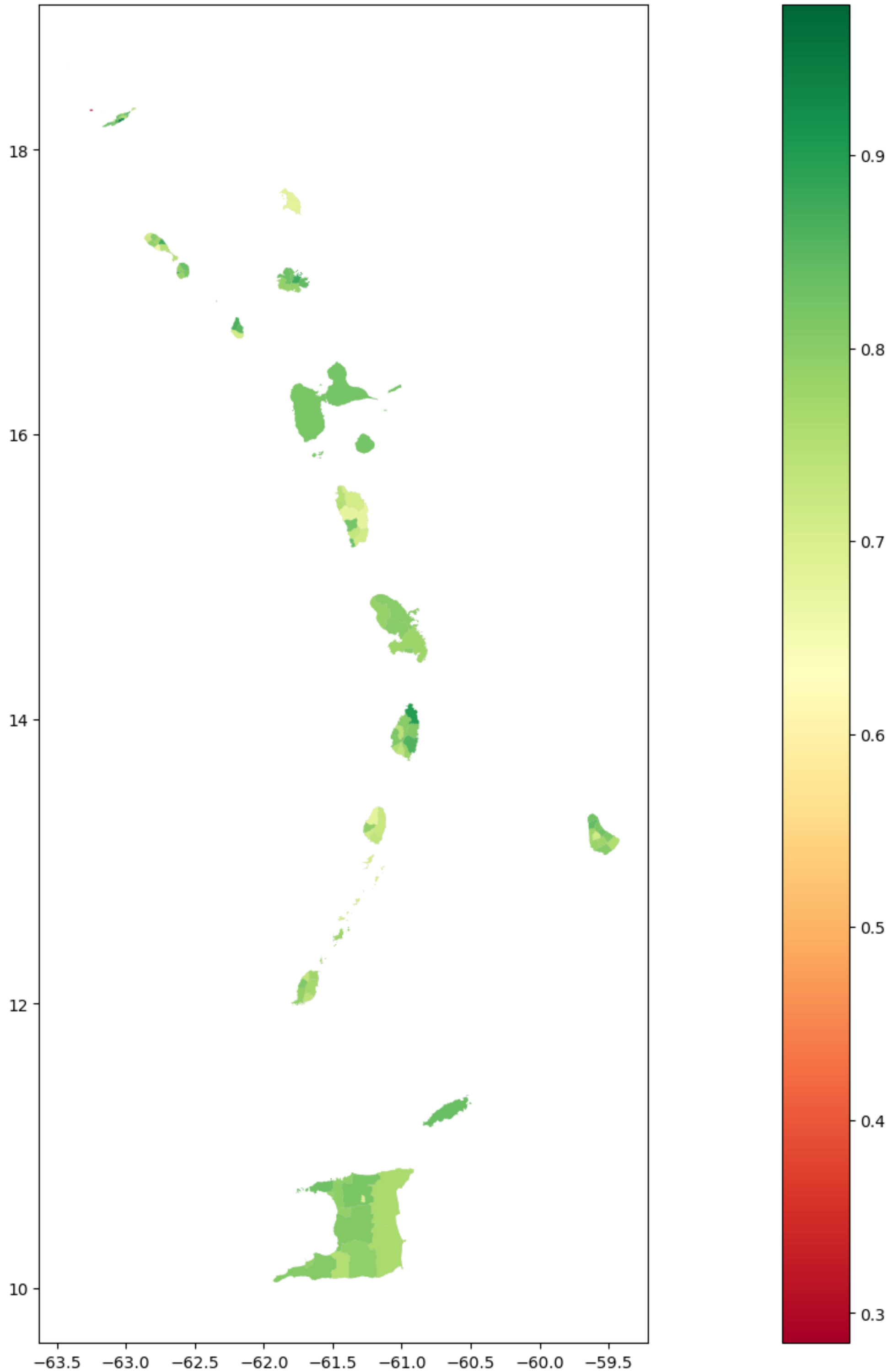



Note: This figure plots national-level HDI ground truth and model predictions.
Sources: United Nations Development Program and authors' calculations.

**Table 7: Validating national-level HDI predictions in the Eastern Caribbean**

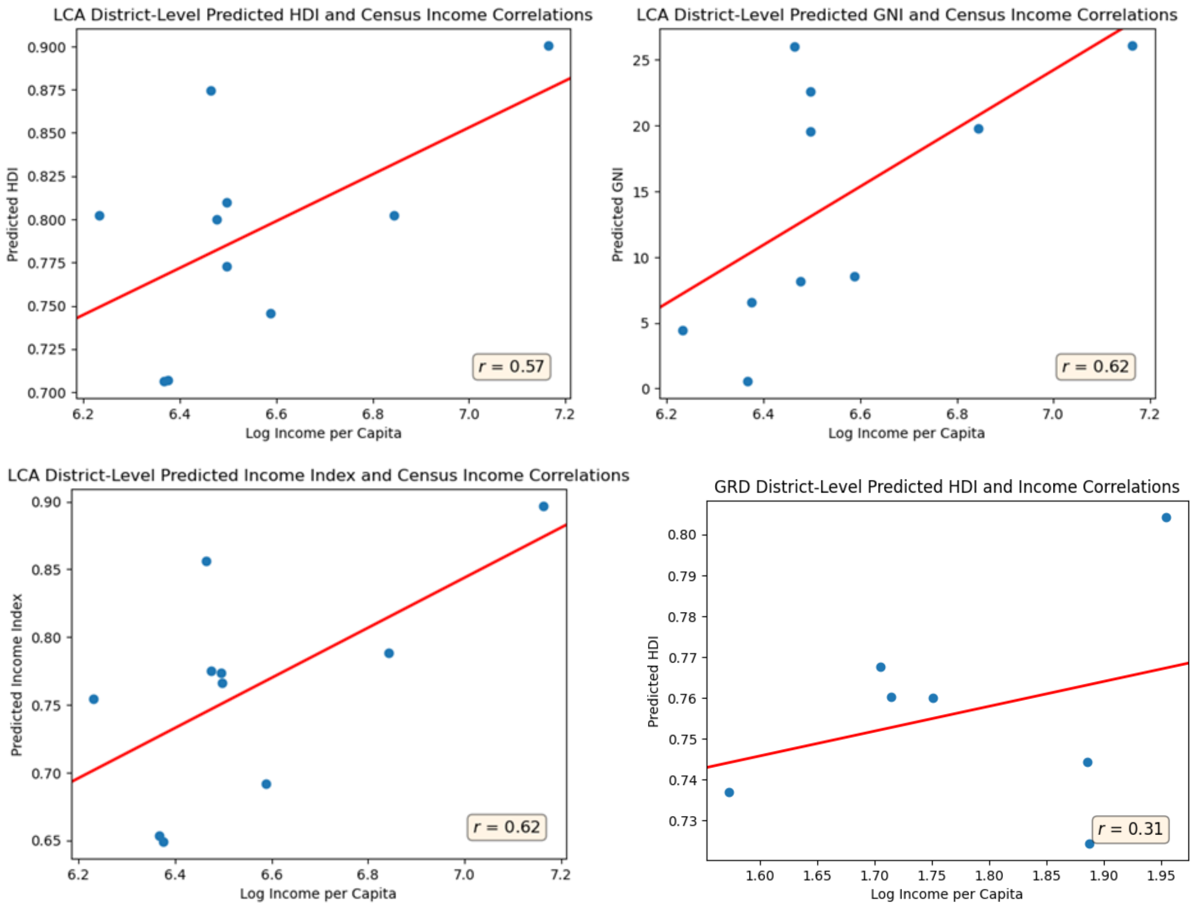
| Variable | Geographical Unit | MSE | R-square |
|---|---|---|---|
| HDI | National-level | 0.002 | 0.117 |
| | Subnational-level | 0.001 | 0.452 |
| GNI | National-level | 33.964 | 0.445 |
| | Subnational-level | 36.740 | 0.362 |
| Health | National-level | 0.003 | 0.005 |
| | Subnational-level | 0.001 | 0.416 |
| Income | National-level | 0.003 | 0.312 |
| | Subnational-level | 0.004 | 0.293 |
| Education | National-level | 0.005 | 0.153 |
| | Subnational-level | 0.005 | 0.085 |

**Figure 5: Subnational-level HDI prediction**



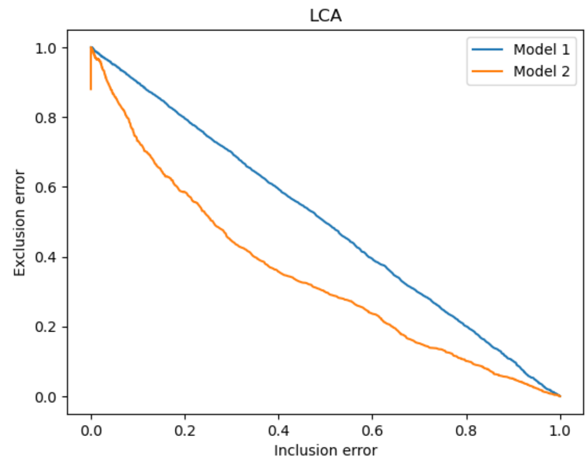

Note: This figure plots baseline predictions of subnational HDI values in the Eastern Caribbean region.
Source: Authors' calculations.

**Figure 6: Validating subnational HDI predictions for St. Lucia and Grenada**




Note: This figure plots subnational-level per capita income and related model predictions.
Sources: St. Lucia Population and Housing Census, Grenada Labor Force Survey, and authors' calculations.

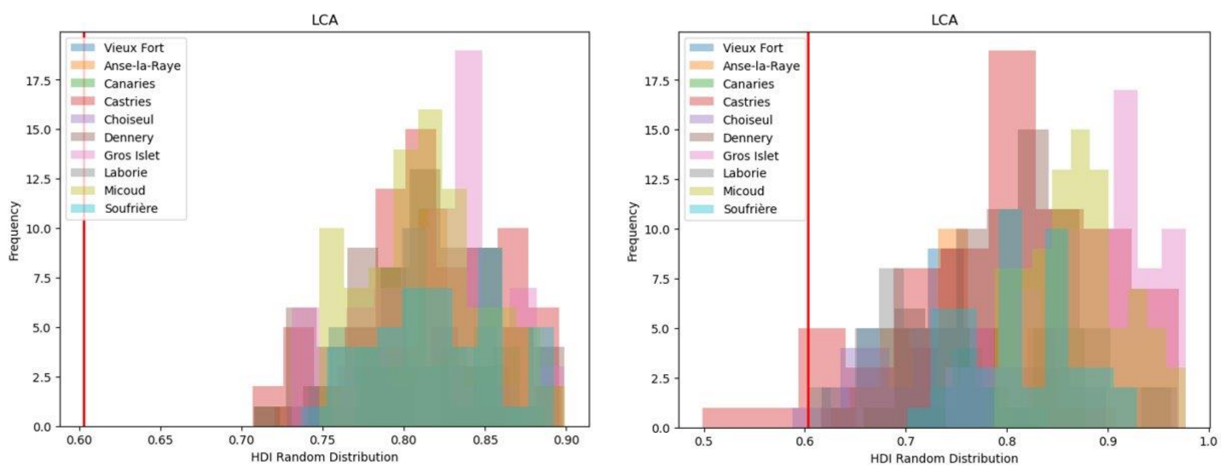
**Figure 7: Reducing Aggregation bias. Exclusion-Inclusion Curves for Policies 1 and 2**




Note: This figure plots the EIC for policies 1 and 2.
Source: Authors' calculations.

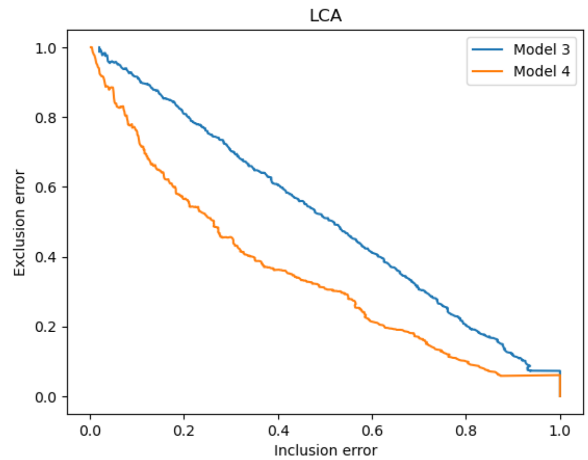
**Figure 8: Illustrative outcomes of Policies 1 and 2**

Panel A: Policy 1 Panel B: Policy 2



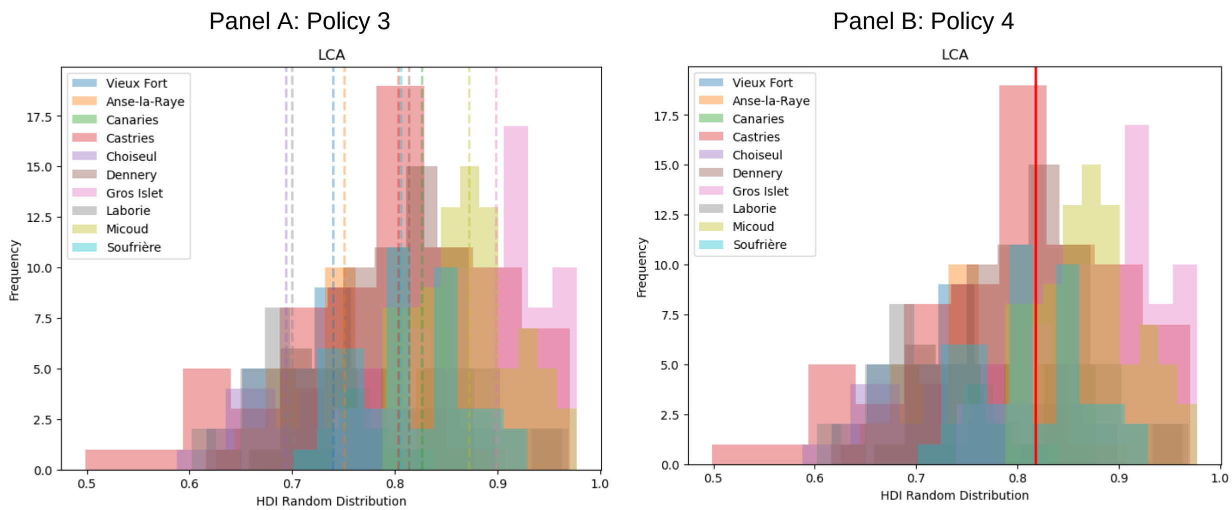

Note: This figure plots the distribution of simulated St. Lucia HDI for policies 1 and 2.
Source: Authors' calculations.

**Figure 1: Improving Resource allocation. EIC for policies 3 and 4**



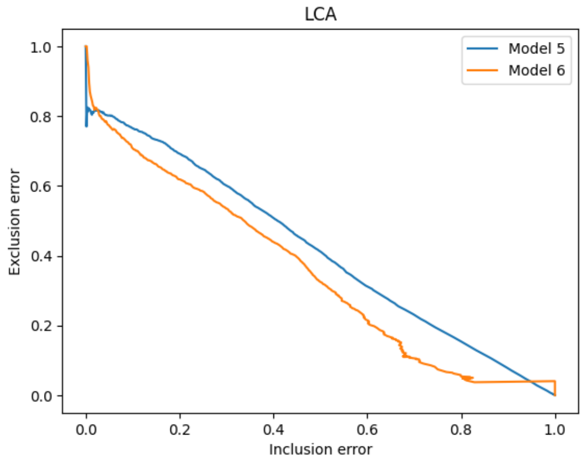

Note: This figure plots the EIC for policies 3 and 4.
Source: Authors' calculations.

**Figure 10. Illustrative outcomes of Policies 3 and 4**




Note: This figure plots the distribution of simulated St. Lucia HDI for policies 3 and 4.
Source: Authors' calculations.

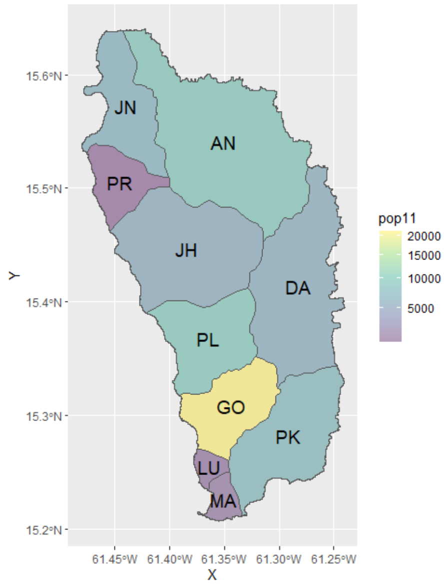
**Figure 11: EIC for Policies 5 and 6**




Note: This figure plots the EIC for policies 5 and 6.
Source: Authors' calculations.

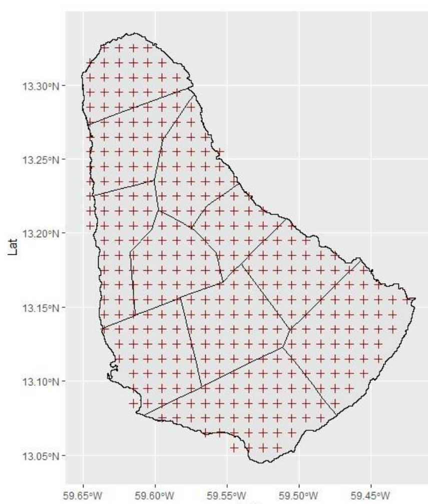
# Appendix 2: Additional Results

### Figure A1. Example of Geographical Heterogeneity



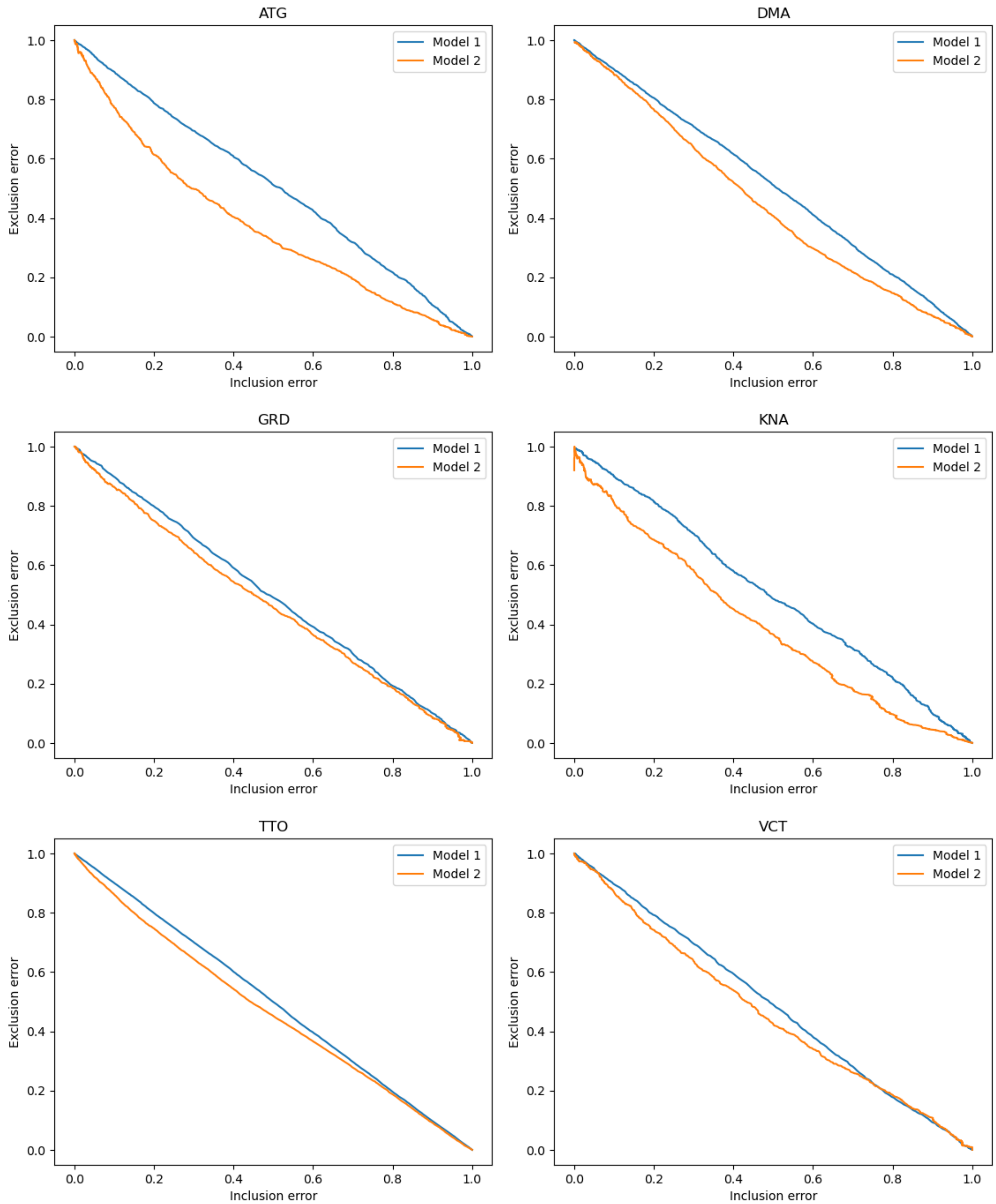

Note: The figure shows cross-district variation in population density in Dominica.
Sources: Gridded Population of the World and authors' calculations.

### Figure A2: Illustrating the resolution of MOSAIKS' grid




Note: This figure illustrates the resolution of MOSAIKS grid for Barbados. The red crosses denote the corners of each MOSAIKS grid.
Sources: Sherman et al. (2023) and authors' calculations.

**Figure A3: EIC for policies 1 and 2 in the other Caribbean Countries**

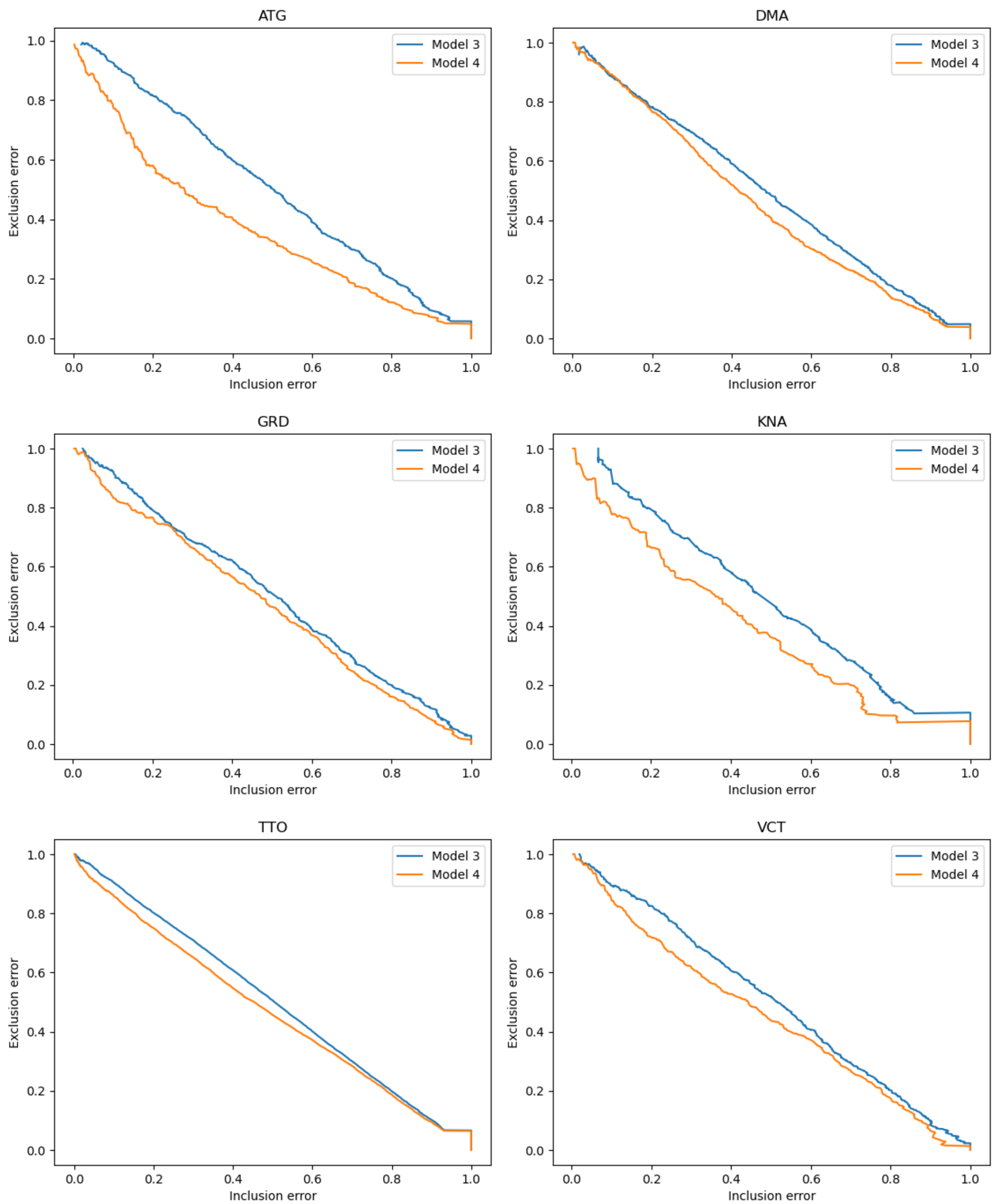



Note: This figure plots the EIC for policies 1 and 2. ATG: Antigua and Barbuda, DMA: Dominica, GRD: Grenada, KNA: St. Kitts and Nevis, TTO: Trinidad and Tobago, VCT: St. Vincent and the Grenadines.
Source: Authors' calculations.

**Table A1. Information value of subnational data**

| | ATG | DMA | GRD | KNA | LCA | TTO | VCT |
|---|---|---|---|---|---|---|---|
| Correlation | 0.45 | 0.24 | 0.16 | 0.38 | 0.48 | 0.26 | 0.19 |

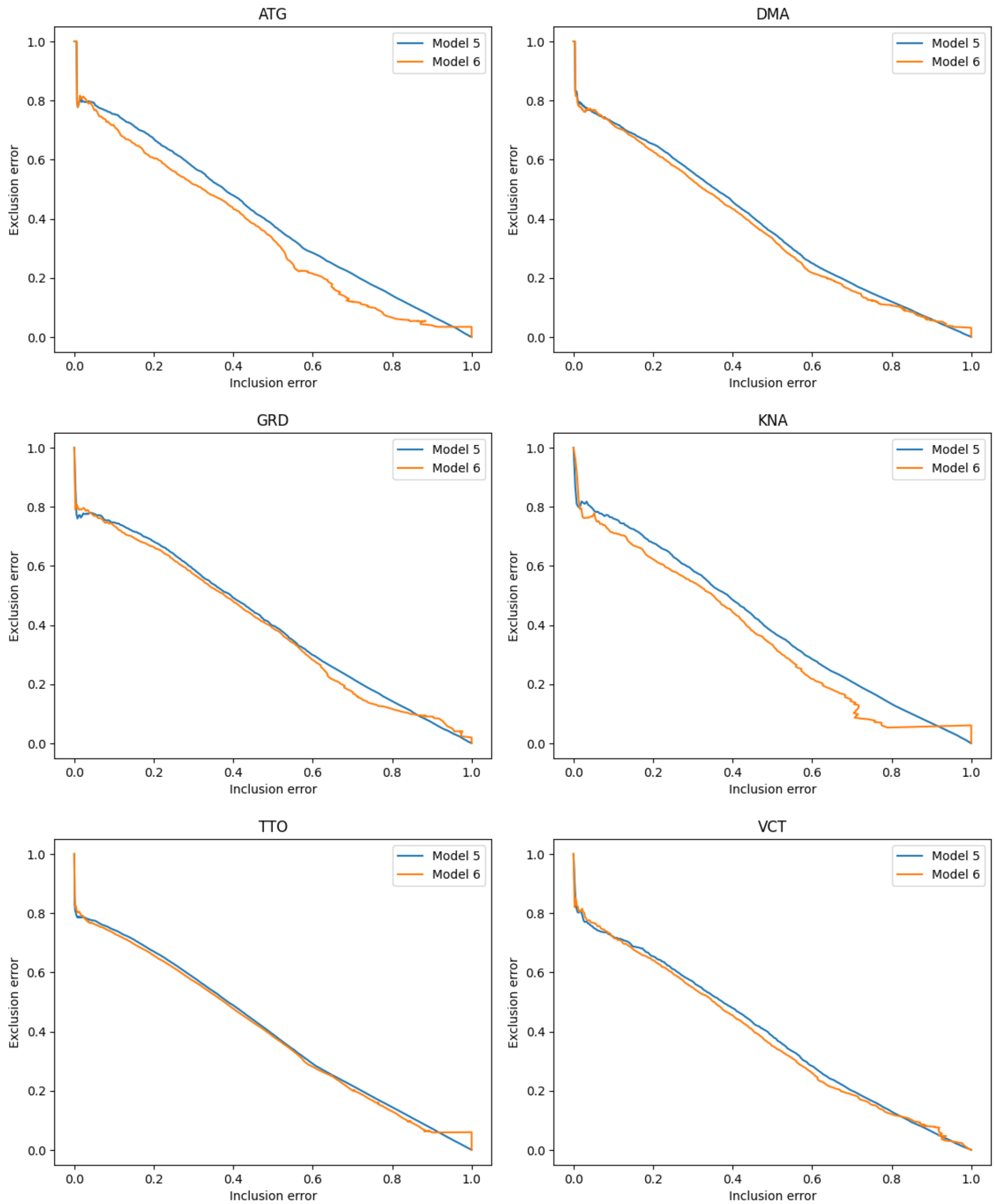
**Figure A4: EIC for policies 3 and 4 for other Caribbean countries**




Note: This figure plots the EIC for policies 3 and 4. ATG: Antigua and Barbuda, DMA: Dominica, GRD: Grenada, KNA: St. Kitts and Nevis, TTO: Trinidad and Tobago, VCT: St. Vincent and the Grenadines.

Source: Authors' calculations.

**Figure A3: EIC for policy experiment 5 and 5 for other Caribbean countries**




Note: This figure plots the EIC for policies 5 and 6. ATG: Antigua and Barbuda, DMA: Dominica, GRD: Grenada, KNA: St. Kitts and Nevis, TTO: Trinidad and Tobago, VCT: St. Vincent and the Grenadines.
Source: Authors' calculations.

# References


Aiken, Emily, Suzanne Bellue, Dean Karlan, Christopher R. Udry, and Joshua Blumenstock. 2022. "Machine Learning and Mobile Phone Data Can Improve the Targeting of Humanitarian Assistance," NBER Working Paper 29070.

Aiken, Emily, Guadalupe Bedoya, Joshua Blumenstock, and Aidan Coville. 2023. "Program Targeting with Machine Learning and Mobile Phone Data: Evidence from an Anti-Poverty Intervention in Afghanistan." *Journal of Development Economics*, 161.

Abelson, Brian, Kush R. Varshney, and Joy Sun. 2014. "Targeting Direct CashTransfers to the Extremely Poor." DataKind DataCorps, New York.

Burke, M., A. Driscoll, D. B. Lobell, S. Ermon. 2021. "Using Satellite Imagery to Understand and Promote Sustainable Development." *Science* 371(6535)..

Chi, Guanghua. Han Fang, Sourav Chatterjee, and Joshua E. Blumenstock. 2022. Microestimates of Wealth for All Low-and Middle-Income Countries. *Proceedings of the National Academy of Sciences* 119(3) e2113658119.

Glaeser, Edward L. 1999. "Learning in Cities," *Journal of Urban Economics*, 46(2): 254-277.

Glaeser, Edward L. and Matthew G. Resseger. 2010. "The Complementarity Between Cities and Skills," *Journal of Regional Science*, 50(1): 221-244.

Grosh, Margaret, Phillippe Leite, Matthew Wai-Poi, Emil Tesliuc. 2022. Revisiting Targeting in Social Assistance: A New Look at Old Dilemmas. Washington, DC: World Bank.

Houshofer, Johannes and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to The Poor," *The Quarterly Journal of Economics*, 131(4): 1973-2042.

Ibarrarán, P., Medellín, N., Regalia, F., Stampini, M., Parodi, S., Tejerina, L., ... & Vásquez, M. (2017). Así funcionan las transferencias condicionadas. Inter-American Development Bank.

International Monetary Fund (IMF). 2022. Staff Report for the 2022 Article IV Consultation with Member Countries on Common Policies of the Eastern Caribbean Currency Union.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1): 3-42.

Jensen, Anders. 2022. "Employment Structure and the Rise of the Modern Tax System." *American Economic Review* 112 (1): 213–34.

McBride, Linden and Austin Nichols. 2016. "Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning. *The World Bank Economic Review*, 32(3): 531–550.

McBride, L., C. B. Barrett, C. Browne, L. Hu, Y. Liu, D. Matteson, Y. Sun, & J. Wen. 2021. Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning. *Applied Economic Perspectives and Policy.*

Mohanasundaram, S., K. Kasiviswanathan, C. Purnanjali, I. P. Santikayasa, S. Singh, Downscaling Global Gridded Crop Yield Data Products and Crop Water Productivity Mapping Using Remote Sensing Derived Variables in the South Asia. *International Journal of Plant Production*, 1–16 (2022).

Pablo Ibarrarán, Nadin Medellín, Ferdinando Regalia, Marco Stampini, Sandro Parodi, Luis Tejerina, Pedro Cueva, and Madiery Vásquez eds. 2017. How Conditional Cash Transfers Work. Washington DC: Inter-America Development Bank.

Amos Peters. 2017. "Estimating the Size of the Informal Economy in Caribbean States." Inter-America Development Bank Technical Note IDB-TN-1248.

Rahimi, A. & Recht, B. 2008. "Weighted Sums of Random Kitchen Sinks: Replacing Minimization with Randomization in Learning." *Advances in Neural Information Processing Systems* 21.

Rolf, Esther, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Vaishaal Shankar, Miyabi Ishihara, Benjamin Recht, and Solomon Hsiang. 2021. , "A Generalizable and Accessible Approach to Machine Learning with Global Satellite Imagery." *Nature Communications* 12: 1–11.

Sherman, Luke, Jonathan Proctor, Hannah Druckenmiller, Heriberto Tapia, and Solomon M. Hsiang. 2023. "Global High-Resolution Estimates of the United Nations Human Development Index Using Satellite Imagery and Machine-learning," NBER Working Paper 31044.

World Bank. 2021. World Development Report 2021: Data for Better Lives. Washington, DC: World Bank.